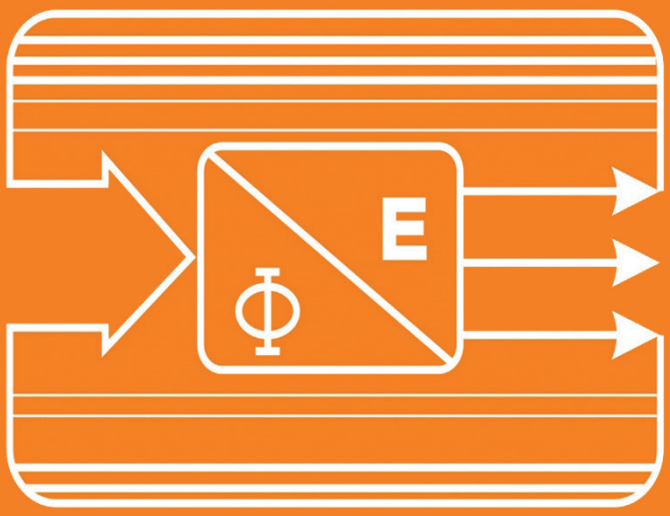
VDI/VDE-Gesellschaft
Mess- und Automatisierungstechnik

20. GMA/ITG-Fachtagung

Sensoren und Messsysteme 2019



NürnbergConvention Center
25. und 26. Juni 2019

VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA)
Informationstechnische Gesellschaft im VDE (ITG)

Tagungsband

20. GMA/ITG-Fachtagung
Sensoren und Messsysteme 2019

Bei diesem Band handelt es sich um den Kongressband der
20. GMA/ITG-Fachtagung Sensoren und Messsysteme 2019.

Dieser Band beinhaltet die Manuskripte zu den jeweiligen Vorträgen und Posterbeiträgen.

Für Form und Inhalt der Beiträge zeichnen sich die Autoren verantwortlich.

Die AMA Service GmbH übernimmt keine Gewähr für die Richtigkeit, Genauigkeit und Vollständigkeit der Angaben sowie die Beachtung privater Rechte Dritter.

Herausgeber / Publisher:

AMA Service GmbH
Von-Münchhausen-Str.49 • 31515 Wunstorf / Germany
Tel. +49 5033 9639-0
info@ama-service.com
www.ama-service.com • www.sensor-test.com
www.ama-science.org

Association for Sensors + Measurement
Linking Innovators

ISBN 978-3-9819376-0-2

© AMA Service GmbH, 2019

VDI/VDE-Gesellschaft
Mess- und Automatisierungstechnik



Willkommen zur 20. GMA/ITG-Fachtagung Sensoren und Messsysteme 2019

Ein herzliches Willkommen zur 20. GMA/ITG-Fachtagung Sensoren und Messsysteme 2019 in Nürnberg. Zum ersten Mal wird diese Veranstaltung nun in einem ungeraden Jahr stattfinden. Gleichgeblieben ist der Standort Nürnberg, der Ihnen als Teilnehmer nicht nur die Wahrnehmung hervorragender Fachvorträge aus Forschung, Wissenschaft und Industrie ermöglicht, sondern zusätzlich die Möglichkeit zum Messebesuch auf der SENSOR+TEST bietet. Die Kombination exzellenter Vorträge gepaart mit den Informationsmöglichkeiten auf der Fachmesse rund um das Thema Sensorik ist es, die diese Veranstaltung so einzigartig in Deutschland macht.

Trotz des Umbruchs im Tagungsrhythmus ist es uns im Programmkomitee unter der Leitung von Prof. Dr. Andreas Schütze (Uni Saarland) sowie Prof. Dr.-Ing. Rainer Tutsch (TU Braunschweig) wieder gelungen, ein spannendes Programm zusammenzustellen. Die Anzahl und die Qualität der Beiträge knüpfen nahtlos an die Vorjahre an. In vier parallelen Sessions werden aktuelle Lösungen aus der Sensorik und Messtechnik aufgezeigt. Begleitet von einer umfangreichen Posterausstellung haben Sie die Möglichkeit, mit allen Referenten der Tagung sowie mit den anderen Tagungsteilnehmern Ihre Fragen zu diskutieren. Auf dieser Tagung ist die Community der Messtechnik und Sensorik vertreten.

Die Themenschwerpunkte reichen von der mechanischen, optischen und photoakustischen Sensorik über Temperaturmessung, Sensorik für Lifescience und Citizen Science sowie energieautarke Messsysteme bis hin zu Modellbildung, Messunsicherheit und Zustandsüberwachung, um nur einige der Themen zu nennen. Es werden Fragestellungen der Sensorik genauso behandelt wie Lösungen zu komplexen Messaufgaben aufgezeigt werden. Gerade diese Vielfalt der Themen bietet die Chance, einen Überblick über das heute Machbare zu bekommen und daraus eigene Lösungen für aktuelle und kommende Aufgabenstellungen ableiten zu können. In Plenarvorträgen erfahren Sie darüber hinaus, was rund um das Thema Digitalisierung in der Messtechnik und Sensorik auf Sie zukommt.

Sie sind herzlich eingeladen, vom umfassenden Erfahrungsschatz der Referenten zu profitieren und wir freuen uns, Sie in Nürnberg zu begrüßen.

S. Dessel

Sascha Dessel, M.Sc.
VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik

Beim letzten Mal lasen Sie an dieser Stelle noch „immer in den Schaltjahren zeichnet die GMA – die VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik – inhaltlich für die Durchführung der Fachtagung „Sensoren und Messsysteme“ verantwortlich.“ In diesem Jahr wechselt dieser Rhythmus nun also, da im kommenden Jahr erstmals in Nürnberg die neue Tagung „SMSI – Sensors and Measurement Science International“ stattfinden wird. Ansonsten bleibt der zweijährige Rhythmus der Tagung „Sensoren und Messsysteme“ erhalten, bei der sich GMA und die Informationstechnische Gesellschaft im VDE (ITG) mit der Programmgestaltung abwechseln. Auch gewechselt hat der Zeitpunkt, der ab jetzt immer in der letzten Juni-Woche liegen wird.

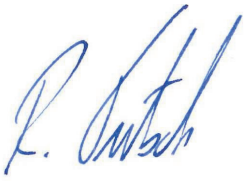
Nicht geändert hat sich jedoch die Bedeutung! Gegründet 1982 in Bad Nauheim, hat sich die „Sensoren und Messsysteme“ seit Jahrzehnten zur bedeutendsten deutschsprachigen wissenschaftlichen Veranstaltung im Bereich der Sensorik und Messtechnik entwickelt. Sie begleitet damit die seit vielen Jahren dynamische Entwicklung der Branche mit durchschnittlich über 8 % Umsatzwachstum, das auch durch einen kontinuierlichen Beschäftigungsaufwuchs begleitet wird. Und ein Ende ist nicht in Sicht, da in Zeiten der Digitalisierung die Sensoren als Quellen der Information in allen Industrie- und Lebensbereichen weiter an Relevanz gewinnen.

Wir freuen uns über die große Zahl an Beitragseinreichungen trotz der unmittelbaren zeitlichen und räumlichen Konkurrenz durch die Tagung Transducers/Eurosensors in Berlin. Die besondere Nähe zur Industrie während der Leitmesse SENSOR+TEST wird nicht nur dadurch, sondern auch durch die hohe Zahl an Beiträgen (30 %) aus und mit der Industrie unterstrichen. Wir danken herzlich allen Autoren für ihr Interesse an der Präsentation auf dieser Tagung, zudem dem Programmkomitee für die hervorragende Arbeit bei der Bewertung der Vorträge und Poster und der Gestaltung des Programms. Die besonderen Schwerpunkte der Tagung im Jahr 2019 werden Sensoren und Sensorsysteme für den industriellen Einsatz sein, z.B. Werkstoffprüfung und -charakterisierung, Zustandsüberwachung sowie Sensorik für Industrie 4.0. Gleichermaßen soll die Tagung jedoch auch über neue Entwicklungen auf den „klassischen“ Gebieten der Sensorik und Messsysteme berichten. Ergänzt wird das Programm durch aktuelle Beiträge, z.B. zu Sensoren für die Wasserstoffwirtschaft bzw. zu Citizen Science mit Sensorik. Wir sind überzeugt, damit wieder einen spannenden Überblick über das gesamte Fachgebiet präsentieren zu können.

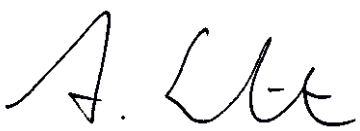
Die Tagung findet am 25. und 26. Juni 2019 parallel zur Messe SENSOR+TEST im NürnbergConvention Center statt. Nürnberg präsentiert sich so einmal mehr als DIE „Sensor-Hauptstadt“ Deutschlands und Europas.

Zuguterletzt gilt unser Dank der AMA Service GmbH für die zuverlässige und konstruktive Zusammenarbeit, die die Organisation der Tagung trotz verkürztem Zeitplan mit deutlich späterer Einreichungsfrist in bewährter Manier gemeistert hat.





Prof. Dr.-Ing. Rainer Tutsch
TU Braunschweig



Prof. Dr. Andreas Schütze
Universität des Saarlandes



Wissenschaftliche Tagungsleiter

Prof. Dr. R. Tutsch	TU Braunschweig
Prof. Dr. A. Schütze	Universität des Saarlandes, Saarbrücken

Programmausschuss

Prof. Dr.-Ing. J. Auge	Hochschule Magdeburg-Stendal
Dr. H. Bosse	PTB Braunschweig
Dipl.-Ing. H. Bödeker	AMA Service GmbH, Wunstorf
Prof. Dr. J. Czarske	TU Dresden
Prof. Dr. W. Daum	BAM, Berlin
S. Dessel	VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik, Düsseldorf
Dr. S. Eichstädt	PTB Braunschweig
Dr. P. Farber	Robert Bosch GmbH, Stuttgart
Dr.-Ing. M. Fischer	TU Braunschweig
Dr. R. Fix	Bosch Sensortec GmbH
Prof. Dr. H. Fritze	TU Clausthal
Prof. Dr.-Ing. T. Fröhlich	TU Ilmenau
Dr. A. Gasch	ABB Forschungszentrum, Ladenburg
Prof. Dr. R. Gerhard	Universität Potsdam
Prof. Dr. M. Gerken	Christian-Albrechts-Universität Kiel
Prof. Dr.-Ing. G. Gerlach	TU Dresden
Prof. Dr.-Ing. M. Heizmann	Karlsruher Institut für Technologie KIT
Prof. Dr.-Ing. B. Henning	Universität Paderborn
Dr.-Ing. D. Imkamp	Carl Zeiss IMT GmbH, Oberkochen
Dr. U. Kaiser	Endress+Hauser Management AG, Reinach (Schweiz)
Prof. Dr.-Ing. O. Kanoun	TU Chemnitz
Prof. Dr. M. Kreutzbruck	Universität Stuttgart
Dr.-Ing. R. Kumme	PTB Braunschweig
Prof. Dr.-Ing. P. Lehmann	Universität Kassel
Prof. Dr.-Ing. E. Manske	TU Ilmenau
Dr. R. Mattheis	IPHT, Jena
Prof. Dr.-Ing. R. Moos	Universität Bayreuth
Dr.-Ing. U. Neuschaefer-Rube	PTB Braunschweig
Dr.-Ing. A. Nocke	TU Dresden
Prof. Dr.-Ing. F. Puente León	Karlsruher Instituts für Technologie – KIT
Prof. Dr. L. Reindl	Universität Freiburg
Prof. Dr.-Ing C. Rembe	TU Clausthal
Dr. S. Rupitsch	Friedrich-Alexander-Universität Erlangen
Dr. T. Sauerwald	Universität des Saarlandes, Saarbrücken
Dr. A. Schäfer	Hottinger Baldwin Messtechnik GmbH, Darmstadt
Univ.-Prof. Dr.-Ing G. Scholl	Helmut-Schmidt-Universität Hamburg
Prof. Dr. G. Sessler	TU Darmstadt
Prof. Dr.-Ing. K.-D. Sommer	TU Ilmenau
Dr. E. Starke	SICK Engineering GmbH, Ottendorf-Okrilla
Prof. Dr.-Ing. H.-R. Tränkler	UniBW München
Prof. Dr.-Ing. R. Werthschützky	TU Darmstadt
Prof. Dr. J. Wöllenstein	Fraunhofer IPM, Freiburg
Prof. Dr.-Ing. S. Zimmermann	Universität Hannover
Dr. J. Zosel	KSI Meinsberg

Sensoren und Messsysteme

Festvortrag **22**
Für alle Zeiten und Culturen! Das revidierte internationale System der Einheiten
J. H. Ullrich, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig (Deutschland)

1. Plenarvortrag
„Immer der Nase nach? – Analytische Geruchssensorik im Spannungsfeld von Digitalisierung und Künstlicher Intelligenz“
A. Büttner, Fraunhofer IVV und Universität Erlangen (Deutschland)
(Beitrag lag nicht vor)

2. Plenarvortrag **30**
Mess- und Sensortechnik in der digitalen Transformation
U. Kaiser, Endress-Hauser AG, Reinach (Schweiz), K.-D. Sommer, Technische Universität Ilmenau (Deutschland)

3. Plenarvortrag **33**
Schnellste Vorgänge im Mikrokosmos: Von der Grundlagenforschung bis hin zur Krebsbekämpfung
F. Krausz, Max-Planck-Institut für Quantenoptik, Garching (Deutschland)

1.1 Optische Messtechnik 1

Sitzungsleiter: M. Fischer, Technische Universität Braunschweig (Deutschland)

1.1.1 **Ein Framework zur Simulation optischer Sensoren** **35**
S. Sauer, T. Dunker, Fraunhofer IFF, Magdeburg (Deutschland), M. Heizmann, Karlsruher Institut für Technologie (Deutschland)

1.1.2 **HoloPort – 3D-Sensor for machine tools** **43**
T. Seyler, J. Engler, Fraunhofer-Institut für Physikalische Messtechnik, Freiburg (Deutschland)

1.1.3 **3-Komponenten-Laser-Doppler-Anemometer mit Selbst-Kalibrierung** **50**
H. Müller, C. Gutsmuths, M. Eggert, H. Többen, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland)

1.1.4 **Optisches Lokalisierungssystem für autarke Sensorknoten in industriellen IoT Anwendungen** **54**
L. Hörmann, M. Pichler-Scheder, Linz Center of Mechatronics GmbH (Österreich), P. Priller, AVL List GmbH, Graz (Österreich), H.-P. Bernhard, A. Springer, Johannes Kepler Universität Linz (Österreich)

1.2 Mechanische Sensoren 1

Sitzungsleiter: T. Fröhlich, Technische Universität Ilmenau (Deutschland)

1.2.1 **Radialsymmetrischer Scherkraftaufnehmer als Transfernormal** **62**
T. Kleckers, Hottinger Baldwin Messtechnik, Darmstadt (Deutschland)

1.2.2 **Design study for a multi-component transducer for wind turbine test benches** **65**
J. Gnauert, G. Jacobs, S. Kock, D. Bosse, B. Janik, RWTH-Aachen (Deutschland)

1.2.3 **Design eines Drehmomentensensors für den MN·m-Bereich nach dem Kraft-Hebelarm-Prinzip**
J. Böseler, G. Jacobs, S. Kock, J. Gnauert, D. Bosse, RWTH Aachen (Deutschland)
(Beitrag lag nicht vor)

1.2.4 Für internationale Vergleichsmessungen optimierter Drehmoment-Transferaufnehmer TN 70
A. Schäfer, Hottinger Baldwin Messtechnik GmbH, Darmstadt (Deutschland)

1.3 Sensorik für die Life Sciences

Sitzungsleiter: S. Zimmermann, Leibniz Universität Hannover (Deutschland)

1.3.1 Erfassung biomechanischer Belastungen und Asymmetrien bei Nachwuchs-Eliteläufern mithilfe mehrachsiger Inertialsensoren 74
O. Ueberschär, J. Roediger, D. Fleckenstein, F. Warschun, N. Walter, Institut für Angewandte Trainingswissenschaft, Leipzig (Deutschland), M. Hoppe, Universität Wuppertal (Deutschland)

1.3.2 A novel non-invasive, non-conductive method for measuring respiration 80
J. Ringkamp, P. Radler, P. Lebhardt, J. Langejürgen, Fraunhofer IPA, Mannheim (Deutschland)

1.3.3 Hochminiaturisiertes nicht-invasives Messsystem zur Erfassung von Vitalparametern bei Kleinstlebewesen mit drahtloser RFID-/NFC-Ausleseschnittstelle 86
P. Moser, E. Mackensen, Hochschule Offenburg (Deutschland), F. Rank, VEGA Grieshaber KG, Schiltach (Deutschland)

1.3.4 Study of safe sensor configurations in human robot collaboration 93
W.-J. Baek, S. Huang, C. Ledermann, T. Kröger, Karlsruher Institut für Technologie (Deutschland), R. Kirsten, ABB AG Forschungszentrum, Ladenburg (Deutschland)

1.4 Prozessautomatisierung und Industrie 4.0

Sitzungsleiter: U. Neuschaefer-Rube, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland)

1.4.1 Optische Sauberkeitskontrolle in der Industriellen Bauteilreinigung – Lösungsansätze für problematische Messaufgaben 100
M. Windisch, Fraunhofer IVV Dresden (Deutschland)

1.4.2 Ein Ansatz zur medienunabhängigen thermischen Flusssensorik
M. Jägle, M. Bartel, M. Benkendorf, B. Bierer, H.-F. Pernau, K. Vanaraj, J. Wöllenstein, Fraunhofer IPM, Freiburg (Deutschland)
(Beitrag lag nicht vor)

1.4.3 Concepts for accurate electrical conductivity measurement of liquids in industrial Process Analytics 105
M. Vogt, KROHNE Messtechnik GmbH, Duisburg (Deutschland), S. Hidalgo, T. Musch, Ruhr Universität Bochum (Deutschland), M. Mallach, T. Lange, J. Förster, KROHNE Innovation GmbH, Duisburg (Deutschland)

1.4.4 Intelligente Kabeltrommel – Vom Sensor in die Cloud
T. Schildknecht, Schildknecht AG, Murr (Deutschland)
(Beitrag lag nicht vor)

2.1 Optische Messtechnik 2

Sitzungsleiter: R. Tutsch, Technische Universität Braunschweig (Deutschland)

2.1.1 Kontinuierliche Feinstaubpartikelklassierung – von Emissionsüberwachung zu Prozessanwendungen 113
A. Schladitz, SICK Engineering GmbH, Ottendorf-Okrilla (Deutschland)

2.1.2 Microwave based electrical read-out of optical sensors 119
U. Nordmeyer, N. Neumann, X. Wang, D. Plettemeier, TU Dresden (Deutschland), T. Thiel, AOS GmbH, Dresden (Deutschland), K. Kojucharow, KMDC, Dresden (Deutschland)

2.1.3 Aktorisches Miniaturspektrometer für die Gassensorik 125
M. Schädel, J. Baldauf, N. Thronicke, D. Mitrenga, D. Karolewski, H.-G. Ortlepp, T. Ortlepp, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt (Deutschland), A. Grewe, S. Sinzinger, Technische Universität Ilmenau (Deutschland)

2.1.4 Quantum cascade laser-based tunable laser absorption spectroscopy for the detection of stable isotopes of CO_2 130
P. Nitzsche, C. Dinc, K. Schmitt, J. Wöllenstein, Universität Freiburg (Deutschland)

2.2 Mechanische Sensoren 2
Sitzungsleiter: A. Schäfer, Hottinger Baldwin Messtechnik GmbH, Darmstadt (Deutschland)

2.2.1 Einfluss von DC- und AC-Versorgungsspannungen auf Dünnschicht- Drucksensoren aus ZrO_2 136
M. Mathis, FITT gGmbH, Saarbrücken (Deutschland), D. Vollberg, G. Schultes, HTW des Saarlandes, Saarbrücken (Deutschland)

2.2.2 Uncertainty evaluation of a traceable dynamic force transfer standard 140
S. Hassan, C. Schlegel, R. Kumme, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland), R. Tutsch, TU Braunschweig (Deutschland)

2.2.3 Absolut messender Winkelsensor auf Basis optischer Interferometrie 147
E. Oertel, V. Ullmann, E. Manske, Technische Universität Ilmenau (Deutschland)

2.2.4 Hybridintegration von Mikrodehnungssensoren 153
T. Frank, A. Grün, M. Kermann, A. Cyriax, A. Steinke, T. Ortlepp, Forschungsinstitut für Mikrosensorik GmbH, Erfurt (Deutschland), G. Reschke, Technische Universität Dresden (Deutschland)

2.3 Energieautarke Messsysteme
Sitzungsleiter: S. Rupitsch, Friedrich-Alexander-Universität Erlangen (Deutschland)

2.3.1 Characteristics of vibrations in domestic environments as sources for kinetic energy harvesters 159
M. Mösch, G. Fischerauer, Universität Bayreuth (Deutschland)

2.3.2 Zutrittskontrolle mit Energy Harvesting aus der Türklinkenbewegung 166
J.-M. Gruber, P. Toma, ZHAW, Winterthur (Schweiz)

2.3.3 SMERS – Energieautarkes drahtloses Sensornetzwerk zur Detektion von Katastrophen und zur Lagebilderfassung in U-Bahnsystemen 170
J. Vincke, C. Horch, F. Schäfer, Fraunhofer-Institut für Kurzzeitdynamik, Freiburg (Deutschland)

2.3.4 Realisierung eines bistabilen Feuchtesensorschalters mit Schalthysterese 177
N. Gulnizkij, G. Gerlach, TU Dresden (Deutschland)

2.4 Hochtemperatursensorik

Sitzungsleiter: L. Reindl, Albert-Ludwigs-Universität, Freiburg (Deutschland)

2.4.1 Investigation for direct sensing the state of three-way-catalysts in the exhaust gas aftertreatment
A. Wollbrink, H. Fritze, Technische Universität Clausthal, Goslar (Deutschland), C. Steiner, R. Moos, Universität Bayreuth (Deutschland)
(Beitrag lag nicht vor)

2.4.2 Berührungslose Temperaturmessung an Verbrennungsgasen bei hohen Temperaturen und hohen Drücken 183
M. Zipf, J. Manara, T. Stark, M. Arduini, H.-P. Ebert, Bayerisches Zentrum für angewandte Energieforschung, Würzburg (Deutschland), J. Hartmann, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt (Deutschland)

2.4.3 Modulare Sensor-Testplattform für Hochtemperaturanwendungen bis 300 °C 191
B. Bieske, T. Reinhold, M. Reinhard, IMMS GmbH, Ilmenau (Deutschland)

2.4.4 Impedanz-basierter NO_X-Sensor für Abgasanwendungen 197
G. Hagen, A. Müller, J. Lattus, J. Kita, R. Moos, Universität Bayreuth (Deutschland), F. Noack, CPK Automotive GmbH & Co. KG, Münster (Deutschland), D. Bleicker, CARIT Automotive GmbH&Co. KG, Münster (Deutschland)

3.1 Infrarotsensoren

Sitzungsleiter: G. Gerlach, Technische Universität Dresden (Deutschland)

3.1.1 Aktuelle Entwicklungen von MEMS IR-Sensorkomponenten 200
M. Schädel, J. Baldauf, L. Long, J. Freitag, N. Thronicke, K. Neckermann, S. Nieland, A. Grün, I. Käpplinger, T. Ortlepp, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt (Deutschland)

3.1.2 High sensitivity 80x64 Thermopile infrared array sensors 205
J. Schieferdecker, M. Schnorr, B. Forg, M. Simon, W. Leneke, F. Herrmann, C. Schmidt, Heimann Sensor GmbH, Dresden (Deutschland)

3.1.3 Infrared-based sensor system for contactless detection of wetness and ice on surfaces 210
L. Tharmakularajah, J. Döring, L. Friedrich, K.-L. Krieger, Universität Bremen (Deutschland)

3.1.4 RoHS-konforme, thermoelektrisch gekühlte Detektoren für das langwellige Infrarot auf Basis antimonidischer Übergitter 216
R. Müller, J. Niemasz, V. Daumer, R. Rehm, Fraunhofer-Institut für angewandte Festkörperphysik (IAF), Freiburg (Deutschland)

3.2 MEMS-Sensoren

Sitzungsleiter: R. Fix, Bosch Sensortec GmbH, Reutlingen (Deutschland)

3.2.1 Droplet-on-cantilever approach for determining the mass of magnetic particles 222
W. O. Nyang'au, A. Setiono, P. Puranto, M. Bertke, H. S. Wasisto, T. Viereck, E. Peiner, Technische Universität Braunschweig (Deutschland), H. Bosse, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig (Deutschland)

3.2.2 Ein neuer Ansatz für eine präzise automatische Temperaturkalibrierung von Microhotplates 230
C. Schultealbert, R. Diener, T. Baur, A. Schütze, T. Sauerwald, Universität des Saarlandes, Saarbrücken (Deutschland)

3.2.3 Wafer-Level-3D-Integrationsverfahren für hochsensitive optische Sensoren 236
L. Kalwa, S. Gläsener, J. Ruskowski, M. Figge, H. Kappert, H. Vogt, Fraunhofer-Institut für Mikroelektronische Schaltungen und Systeme IMS, Duisburg (Deutschland)

3.2.4 Programmable reference parameters for resonance locking in electro-thermal piezoresistive cantilever sensor 243
A. Setiono, M. Fahrbach, J. Xu, W. O. Nyang'au, M. Bertke, H. S. Wasisto, E. Peiner, Technische Universität Braunschweig (Deutschland)

3.3 Sensoren für eine moderne Wasserstoffwirtschaft
Sitzungsleiter: U. Kaiser, Endress+Hauser Management AG, Reinach (Schweiz)

3.3.1 Prozess- und Sicherheitssensorik für das Wasserstoff-Qualitätsmanagement 250
J. Zosel, Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, Waldheim (Deutschland), M. Eiserbeck, A. Hebestreit, Hochschule für Technik, Wirtschaft und Kultur Leipzig (HTWK) (Deutschland), T. Frosch, Leibniz-Institut für Photonische Technologien e.V. (IPHT), Jena (Deutschland), O. Kiesewetter, UST Umweltsensortechnik GmbH (UST), Geschwenda (Deutschland), M. Wienecke, Materion GmbH, Wismar (Deutschland), F. Altmann, Fraunhofer-Institut für Mikrostruktur von Werkstoffen und Systemen IMWS, Halle (Saale) (Deutschland), M. Mertig, Technische Universität Dresden (Deutschland)

3.3.2 Sensitivität und Selektivität optischer und neuartiger MEMS-Wasserstoffsensoren 256
M. Wienecke, Hochschule Wismar (Deutschland), L. Godentath, B. Zacharias, Materion GmbH, Wismar (Deutschland)

3.3.3 Diversitär-redundanter H_2-Sensor für H_2-Konzentrationen bis 100 % und einen Umgebungsdruck bis 100 bar für Elektrolyseuranwendungen
O. Kiesewetter, N. Kiesewetter, A. Kraußer, M. May, UST Umweltsensortechnik GmbH, Geschwenda (Deutschland), T. Frank, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt (Deutschland)
(Beitrag lag nicht vor)

3.3.4 Development and characterization of a highly selective hydrogen sensor system 264
P. Sood, J. Zosel, M. Mertig, W. Oelßner, A. Klockow, Kurz-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, Waldheim (Deutschland), O. Herrmann, M. Woratz, ACI Analytical Control Instruments GmbH, Berlin (Deutschland)

3.4 Magnetische Sensorik
Sitzungsleiter: M. Kreutzbruck, Universität Stuttgart (Deutschland)

3.4.1 Advanced TMR sensing solutions for reliable, precise and robust motor control applications
M. Rezé, TDK-Micronas GmbH, Freiburg (Deutschland), M. C. Meyer, TDK Europe GmbH, Leonberg (Deutschland)
(Beitrag lag nicht vor)

3.4.2 Combination of laser topography measurements with 3D magnetic field mapping for permanent magnets, magnet assemblies and PM rotors 271
K. Vervaeke, Magcam NV, Leuven (Belgien)

3.4.3 **Development of a rotating-coil scanner for superconducting accelerator magnets 275**
P. Rogacki, L. Fiscarelli, S. Russenschuck, CERN, Meyrin (Schweiz),
K. Hameyer, Universität Aachen (Deutschland)

3.4.4 **Tunnelmagnetoresistive Sensoren für Weg-, Winkel- und Positionsmessung in der elektrischen Antriebstechnik 283**
R. Slatter, Sensitec GmbH, Lahnau (Deutschland)

4.1 Werkstoffprüfung und -charakterisierung 1
Sitzungsleiter: S. Rupitsch, Friedrich-Alexander-Universität Erlangen (Deutschland)

4.1.1 **Messung mechanischer Beanspruchungen in laufenden Schleifprozessen 291**
A. Tausendfreund, G. Alexe, D. Stöbener, A. Fischer, Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft (Deutschland)

4.1.2 **An optical sensor system for characterization of ceramics based on laser speckle photometry 297**
L. Chen, U. Cikalova, B. Bendjus, S. Muench, M. Roellig, Fraunhofer-Institut für Keramische Technologien und Systeme IKTS, Dresden (Deutschland)

4.1.3 **Impactschäden schnell und einfach finden 304**
J. Rittmann, M. Rahammer, M. Kreutzbruck, Universität Stuttgart (Deutschland),
N. Holtmann, edevis GmbH, Stuttgart (Deutschland)

4.1.4 **Ultrasonic characterisation of steel fibre orientation in ultra-high performance concrete 310**
B. Ankay, K. Metje, T. Mönig, C. Zhang, T. Leutbecher, Universität Siegen (Deutschland)

4.2 Photoakustische und optische Gassensorik 1
Sitzungsleiter: S. Zimmermann, Leibniz Universität Hannover (Deutschland)

4.2.1 **Konzeption und Evaluierung eines Laboraufbaus zur photoakustischen Detektion von CH_4 mit modulierter Breitband MEMS-IR-Strahlungsquelle 318**
F. Liebold, B. Vatter, A. Magi, Micro-Hybrid Electronic GmbH, Hermsdorf (Deutschland)

4.2.2 **Miniaturisierte photoakustische Detektoren für den Nachweis fluorhaltiger Kältemittel 322**
C. Weber, M. El-Safoury, A. Eberhardt, K. Schmitt, J. Wöllenstein, Fraunhofer Institut für physikalische Messtechnik IPM, Freiburg (Deutschland)

4.2.3 **Filterrotationsspektrometer für den Nachweis von Ethen im ppb-Bereich 328**
A. Eberhardt, M.-L. Bauersfeld, Fraunhofer-Institut für Physikalische Messtechnik IPM, Freiburg (Deutschland), K. Schmitt, J. Wöllenstein, Technische Fakultät der Universität Freiburg (Deutschland)

4.2.4 **Lichtabsorption durch freie Ladungsträger in Metalloxiden: Anwendung der Brechungsindex-Modulation in der Gassensorik 334**
T. Wagner, X. Wu, X. Zhang, S. Amrehn, Universität Paderborn (Deutschland)

4.3 Machine Learning und Signalverarbeitung
Sitzungsleiter: M. Heizmann, Karlsruher Institut für Technologie (Deutschland)

4.3.1 **Erkennung von Landmarken für Roboternavigation mit Luftultraschall und neuronalen Netzen – Feature Engineering und Vergleich 339**
P. K. Kroh, S. J. Rupitsch, Friedrich-Alexander-Universität Erlangen-Nürnberg (Deutschland), R. Simon, Vrije Universität Amsterdam (Niederlande)

4.3.2 Vergleich subjektiver und automatisierter Merkmalsextraktion sowie Einsatz maschineller Lernalgorithmen zur mikromagnetischen Materialcharakterisierung 347
S. Youssef, C. Zimmer, K. Szielasko, Z. K. Suri, Fraunhofer Institut für Zerstörungsfreie Prüfverfahren IZFP, Saarbrücken (Deutschland), A. Schütze, Universität des Saarlandes, Saarbrücken (Deutschland)

4.3.3 An empirical study of vibration based bearing fault diagnosis methods 355
K. Pichler, C. Hesch, C. Kastl, Linz Center of Mechatronics GmbH, Linz (Österreich), T. Ooijevaar, Flanders Make, Leuven (Belgien)

4.3.4 Enhanced performance of thermographic asset monitoring by an interpolation-based correction algorithm 362
A. Boyaci, S. Wildermuth, H. Kaul, ABB AG, Ladenburg (Deutschland), T. Kozel, ABB s.r.o., Brno (Tschechische Republik)

4.4 Modulare Sensorsysteme für die Zustandsüberwachung
Sitzungsleiter: E. Starke, SICK Engineering GmbH, Ottendorf-Okrilla (Deutschland)

4.4.1 Zustandsüberwachung in der Automatisierungstechnik mittels maschinellem Lernen 366
S. Klein, T. Schneider, ZeMA - Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Saarbrücken (Deutschland), A. Schütze, Universität des Saarlandes, Saarbrücken (Deutschland)

4.4.2 Robuste xMR-Sensoren für die Automatisierungstechnik 373
J. Traute, Sensitec GmbH, Lahnau (Deutschland), F. Casper, J. Paul, Sensitec GmbH, Mainz (Deutschland)

4.4.3 Modularer Messsystembaukasten für FPGA-basierte Signalverarbeitung 379
D. Laumann, M. Selke, C. Nienhaus, J. Doerr, CANWAY Technology GmbH, Ostbevern (Deutschland)

4.4.4 Drahtlose Sensorik für die Automatisierungstechnik
A. Hennig, H.-C. Müller, Fraunhofer-Institut für Mikroelektronische Schaltungen und Systeme IMS, Duisburg (Deutschland)
(Beitrag lag nicht vor)

5.1 Werkstoffprüfung und -charakterisierung 2
Sitzungsleiter: A. Schütze, Universität des Saarlandes, Saarbrücken (Deutschland)

5.1.1 Mikromagnetische Detektion von Härteinhomogenitäten an Grobblechen 385
S. Youssef, M. Schuppmann, W. Bähr, S. Pushkarev, H. Theado, P. Stopp, J. Wagner, Fraunhofer Institut für Zerstörungsfreie Prüfverfahren IZFP, Saarbrücken (Deutschland), D. Molenda, O. Stawicki, ROSEN Technology and Research Center GmbH, Lingen (Deutschland)

5.1.2 A DC bias approach to the characterisation of non-linear material parameters of piezoelectric ceramics 391
D. Dreiling, N. Feldmann, B. Henning, Universität Paderborn (Deutschland)

5.1.3 Torque calibration in the MN m range under rotation in a nacelle test bench 398
P. Weidinger, G. Foyer, R. Kumme, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland), S. Kock, J. Gnauert, RWTH Aachen Universität (Deutschland)

5.1.4 Bestimmung des Brechungsindex des Schmiermittels Beruforge 152DL unter Verwendung eines dafür entwickelten Dünnschichtnormals 406
S. Metzner, T. Reuter, E. Wirthmann, T. Hausotte, Friedrich-Alexander-Universität Erlangen-Nürnberg (Deutschland)

5.2 Photoakustische und optische Gassensorik 2
Sitzungsleiter: T. Sauerwald, Universität des Saarlandes, Saarbrücken (Deutschland)

5.2.1 Ramanspektroskopie mit photonischen Hohlkernfasern am Beispiel atmosphärischer Gase
C. Niklas, G. Ctistis, F. Müller, H. Wackerbarth, Laser-laboratorium Göttingen e. V. (Deutschland)
(Beitrag lag nicht vor)

5.2.2 Entwicklung eines Hochdruckprüfstands für NDIR-Messungen zur Verunreinigungsbestimmung in Wasserstoff für Drücke bis 900 bar 412
M. Schott, Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Saarbrücken (Deutschland), A. Schütze, Universität des Saarlandes, Saarbrücken (Deutschland)

5.2.3 Zinkoxid als photostabiler Luminophor zur optischen Sauerstoffdetektion 420
M. Poeplau, S. Ester, Wöhler Technik GmbH, Bad Wünnenberg (Deutschland), B. Henning, T. Wagner, Universität Paderborn (Deutschland)

5.2.4 Photoakustisches Gasmesssystem zur Bestimmung des Schwefeldioxidgehaltes in Schiffsabgasen 425
M. El-Safoury, M. Dufner, C. Weber, H.-F. Pernau, Fraunhofer-Institut für Physikalische Messtechnik IPM, Freiburg (Deutschland), K. Schmitt, J. Wöllenstein, Universität Freiburg (Deutschland), B. Willing, RÜEGER SA, Crissier (Schweiz)

5.3 Akustische Messverfahren
Sitzungsleiter: J. Auge, Hochschule Magdeburg-Stendal (Deutschland)

5.3.1 Key design elements of a radio-acoustic-sounding system for industrial applications 430
D. Exel, B. Zagar, Johannes Kepler Universität Linz (Österreich), S. Schuster, S. Scheiblhofer, D. Zankl, V. Ganglberger, J. Reisinger, voestalpine Stahl GmbH, Linz (Österreich)

5.3.2 Berührungslose akustische Prozessüberwachung von Laserschweißen
G. Kaniak, T. Heine, R. Sommerhuber, Xarion Laser Acoustics GmbH, Wien (Österreich), C. Schröter, Optoprim Germany GmbH, Unterschleißheim (Deutschland), S. Zarini, Optoprim srl, Monza (Italien)
(Beitrag lag nicht vor)

5.3.3 Parametric spectrum analysis of backscattered ultrasound signals for the characterization of particles in suspensions 436
M. Vogt, M. Deilmann, KROHNE Messtechnik GmbH, Duisburg (Deutschland)

5.3.4 Characterization of the imaging properties of an ultrasound-doppler velocimeter using a multimode wavequide
R. Nauber, A. Klass, J. Czarske, Technische Universität Dresden (Deutschland)
(Beitrag lag nicht vor)

5.4 Zustandsüberwachung

Sitzungsleiter: E. Starke, SICK Engineering GmbH, Ottendorf-Okrilla (Deutschland)

5.4.1 Synchronisationsprobleme innerhalb eines Sensorsystems und deren Auswirkungen auf Ergebnisse des maschinellen Lernens 444
T. Dorst, S. Eichstädt, Physikalisch-Technische Bundesanstalt, Berlin (Deutschland), T. Schneider, A. Schütze, ZeMA – Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Saarbrücken (Deutschland), S. Klein, Universität des Saarlandes, Saarbrücken (Deutschland)

5.4.2 Feature-based analysis of reproducible bearing damages based on a neural network 451
A. Beering, J. Döring, K.-L. Krieger, ITEM Universität Bremen (Deutschland)

5.4.3 Ultraschallsensoren für die Instandhaltung 457
P. Holstein, N. Bader, C. Probst, SONOTEC Ultraschallsensorik Halle GmbH (Deutschland), S. Seitz, TU Dresden (Deutschland)

5.4.4 Homodyne vector network analysis as a tool for the real-time measurement of electrical material parameter distributions in the field 462
R. Peter, G. Fischerauer, Universität Bayreuth (Deutschland)

6.1 Citizen Science mit Sensoren – Sensorik für Citizen Science

Sitzungsleiter: H.-E. Endres, Fraunhofer-Einrichtung – EMFT, München (Deutschland)

6.1.1 Citizen Science: Was, Warum und Wie?
E. Griessler, Institut für Höhere Studien, Wien (Österreich)
(Beitrag lag nicht vor)

6.1.2 Luftdaten.info – von der lokalen Idee zum globalen Projekt
R. Zschiegner, Luftdaten.info, Stuttgart (Deutschland)
(Beitrag lag nicht vor)

6.1.3 Chancen und Herausforderungen von Kooperationen zwischen Forschungseinrichtungen und bürgerwissenschaftlichen Initiativen am Beispiel des Forschungsprojekts „CitizenSensor“ 469
U. Kaiser, Fraunhofer-Zentrum für Internationales Management und Wissensökonomie, Leipzig (Deutschland), K. T. Möbius, Fraunhofer-Einrichtung für Mikrosysteme und Festkörper-Technologien, München (Deutschland), T. Klages, Fraunhofer-Informationszentrum Raum und Bau, Stuttgart (Deutschland)

6.1.4 Rapid-Prototyping mikrofluidischer Komponenten ermöglicht dezentrale und personalisierte Gesundheitsanwendungen
R. Trojok, UnternehmerTUM GmbH, Garching (Deutschland)
(Beitrag lag nicht vor)

6.1.5 Citizen science für Schüler*innen: Durchführung von Umweltstudien mit Smartphone und mobiler Messtechnik 476
S. Höfner, A. Schütze, Universität des Saarlandes, Saarbrücken (Deutschland), B. Brück, Schülerforschungszentrum Saarlouis (Deutschland), M. Hirth, J. Kuhn, TU Kaiserslautern (Deutschland)

6.2 Chemische Sensoren

Sitzungsleiter: J. Wöllenstein, Fraunhofer-Institut IPM, Freiburg (Deutschland)

6.2.1 Unterscheidung und Quantifizierung von Geruchsstoffen im ppb-Bereich mit temperaturzyklisch betriebenen MOS-Sensoren 482
J. Joppich, A. Schütze, T. Sauerwald, Universität des Saarlandes, Saarbrücken (Deutschland)

6.2.2 Novel photoelectrochemical biosensor to detect catechol
J. Boudaden, A. Klumpp, Fraunhofer EMFT, München (Deutschland)
(Beitrag lag nicht vor)

6.2.3 Impedanzspektroskopie eines auf metallorganischen Netzwerken basierenden Sorptionssensors 488
H. Lensch, A. Schütze, T. Sauerwald, Universität des Saarlandes, Saarbrücken (Deutschland), N. Baskent, M. Völker, Fraunhofer-Institut für Integrierte Schaltungen, Erlangen (Deutschland), I. Wilhelm, J. Hürttlen, Fraunhofer-Institut für Chemische Technologien, Pfinztal (Deutschland)

6.2.4 Einsatz der Cyclovoltammetrie zur Steigerung der Selektivität von Festelektrolytsensoren . . 492
A. Ruchets, J. Zosel, U. Guth, M. Mertig, Kurt-Schwabe-Institut für Mess- und Sensortechnik e. V. Meinsberg, Waldheim (Deutschland), N. Donker, D. Schönauer-Kamin, R. Moos, Universität Bayreuth (Deutschland)

6.2.5 Niedertemperatur-Pellistoren aus mesoporösem Au-Pd@Co_3O_4 498
H. Gao, X. Lyu, K. Schmitt, J. Wöllenstein, Universität Freiburg (Deutschland), K. Tarantik, Fraunhofer Institut für physikalische Messtechnik IPM, Freiburg (Deutschland)

6.3 Akustische und resonante Sensoren

Sitzungsleiter: B. Henning, Universität Paderborn (Deutschland)

6.3.1 Precise viscosity and density sensing in industrial and automotive applications 504
T. Voglhuber-Brunnmaier, A. O. Niedermayer, F. Feichtinger, B. Jakoby, Johannes Kepler Universität, Linz (Österreich)

6.3.2 Accuracy study of a torsionally oscillating pipe viscometer 511
S. Clara, F. Feichtinger, T. Voglhuber-Brunnmaier, A. O. Niedermayer, B. Jakoby, Johannes Kepler Universität Linz (Österreich)

6.3.3 Bestimmung des Mischungsverhältnisses eines strömenden Gasgemisches 517
J. Stauffenberg, S. Durstewitz, M. Hofmann, T. Ivanov, M. Holz, J.-P. Zöllner, M. Ziegler, I. W. Rangelow, TU Ilmenau (Deutschland), W. Ehrhardt, Electronic & Sensors, Ilmenau (Deutschland), W.-U. Riegel, ETG – Entwicklungs- und Technologie Gesellschaft mbH, Ilmenau (Deutschland)

6.3.4 Integraler Defekterkennungssensor für Faserkunststoffverbunde mit koppelmittelfreiem luftgekoppelten Ultraschall 523
Y. Bernhardt, M. Kreutzbruck, Universität Stuttgart (Deutschland)

6.3.5 High-temperature electromechanical losses in Bulk Z-cut AlN single crystals for resonant sensors
I. Kogut, H. Fritze, Technische Universität Clausthal, Goslar (Deutschland), I. Gamov, C. Hartmann, K. Irmscher, M. Bickermann, T. Straubinger, Leibniz-Institut Für Kristallzüchtung (IKZ), Berlin (Deutschland)
(Beitrag lag nicht vor)

6.4 Messunsicherheit und Modellbildung

Sitzungsleiter: K.-D. Sommer, Technische Universität Ilmenau (Deutschland)

6.4.1 Separation of locally determined work piece deviations and measurement uncertainties for structured-light scanning of customized polymer gear wheels 527
A. M. Müller, S. Metzner, T. Hausotte, Institute of Manufacturing Metrology (FMT), Erlangen (Deutschland), D. Schubert, D. Drummer, Institute of Polymer Technology (LKT), Erlangen-Tennenlohe (Deutschland)

6.4.2 Entwicklung eines Systems für hochpräzise Mikroverzahnungsmessungen 535
S. Jantzen, M. Stein, K. Kniel, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland), D. Metz, A. Dietzel, Technische Universität Braunschweig (Deutschland)

6.4.3 Modelle der Oberflächenform mechanisch bearbeiteter Werkstücke für die Messunsicherheitsbestimmung durch Simulation bei taktilen Koordinatenmessungen 541
T. Kistner, D. Imkamp, ZEISS Industrial Quality Solutions, Oberkochen (Deutschland), T. Hausotte, Friedrich-Alexander-Universität Erlangen-Nürnberg (Deutschland)

6.4.4 Investigations on the sine fitting algorithm in the Planck-Balance 547
S. Lin, C. Rothleitner, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland), N. Rogge, Technische Universität Ilmenau (Deutschland)

6.4.5 Model based correction of motion deviations in the Planck-Balance 554
H. Bai, Tianjin University, Tianjin (China), N. Rogge, T. Fröhlich, Technische Universität Ilmenau (Deutschland), C. Rothleitner, Physikalisch-Technische Bundesanstalt, Braunschweig (Deutschland)

Postersession

P1: Sensoren

P1.1 **Performanceanalyse von zwei optimierten Messverfahren in der kapazitiven Füllstandmessung 561**
J. Matheis, A. Uhl, T. Föhn, F. Dayi, EBE Elektro-Bau-Elemente GmbH, Leinfelden-Echterdingen (Deutschland)

P1.2 **Packaging-Konzepte für Thermoelementen auf keramischer Basis für höchste Einsatztemperaturen**
P. Gierth, H.-P. Martin, B. Feng, L. Rebenklau, Fraunhofer IKTS, Dresden (Deutschland)
(Beitrag lag nicht vor)

P1.3 **Validierung von Hochtemperatur-Packaging-Konzepten für Sensoren mit Einsatztemperatur oberhalb 500 °C**
P. Gierth, H. Barth, L. Rebenklau, Fraunhofer IKTS Dresden (Deutschland)
(Beitrag lag nicht vor)

P1.4 **Optische Sensorik für die additive Fertigung**
J. Hartmann, K. Knopp, P. Lenski, D. Ochs, M. Zänglein, Hochschule für Angewandte Wissenschaften, Würzburg-Schweinfurt (Deutschland), M. Arduini, J. Manara, T. Stark, M. Zipf, Bayerisches Zentrum für Angewandte Energieforschung, Würzburg (Deutschland), M. Brunner, Netzsch Gerätebau GmbH, Selb (Deutschland), U. Krüger, F. Schmidt, Techno Team Bildverarbeitung GmbH, Ilmenau (Deutschland), E. Schreiber, KE Technologie, Stuttgart (Deutschland)
(Beitrag lag nicht vor)

P1.5 **Kostengünstige Nass/Nass-Differenzdruckmessung 566**
R. Werner, H. Schöpe, Analog Microelectronics GmbH, Mainz (Deutschland)

P1.6 **Mikroanalysesystem auf Basis eines ladungsempfindlichen Fluidsensors mit Rückseitenkontakten**
M. Steinmaßl, J. Boudaden, H.-E. Endres, D. Reiser, M. Heigl, K. Neumeier, I. Eisele, C. Kutter, Fraunhofer Einrichtung für Mikrosysteme und Festkörpertechnologien EMFT, München (Deutschland)
(Beitrag lag nicht vor)

P1.7 **Elaboration of reduced graphene oxide electrodes functionalized with RGO-iron oxide composite for nitrite detection in water 573**
A. Al-Hamry, A. Reddy, A. Bouhamed, O. Kanoun, Technische Universität Chemnitz (Deutschland), A. Brahem, S. B. Nasrallah, M. B. Ali, Tunisia und Universität von Sousse (Tunesien), L. G. Paterno, Universität von Brasilien (Brasilien)

P1.8 **Photo-electrochemical detection of nitrite ions on a SPE functionalized with TiO_2 nanoparticles for environmental monitoring 578**
A. Bouhamed, A. Al-Hamry, O. Kanoun, F. Wendler, Technische Universität Chemnitz (Deutschland), M. A. Elaguech, S. Ameur, M. B. Ali, University of Sousse (Tunesien)

P1.9 **Sensitivity of layer-by-layer deposited GO/PDAC to volatile organic compounds 584**
A. Al-Hamry, T. Zubkova, R. Baumann, O. Kanoun, Technische Universität Chemnitz (Deutschland), L. G. Paterno, Universidade de Brasilia (Brasilien), A. Errachid, Université de Lyon (Frankreich)

P1.10 **HTCC-multilayerbasierte ionensensitive Sensoren für die Wasseranalyse 588**
C. Feller, U. Partsch, S. Körner, Fraunhofer-Institut für Keramische Technologien und Systeme IKTS, Dresden (Deutschland)

P1.11 **LHyCON (Low hydrogen concentration measurement sensor) 594**
A. Balakrishnan, S. Akan, M. Völkel, LAMTEC Meß- und Regeltechnik für Feuerungen GmbH & Co. KG, Walldorf (Deutschland), S. Geiger, P. Rabenecker, Fraunhofer-Institut für Chemische Technologie ICT, Pfinztal (Deutschland)

P1.12 **Pulspolarisation: Einfluss der Polarisationsspannung auf die NO_X-Detektion mit dem System Pt|YSZ 601**
N. Donker, D. Schönauer-Kamin, R. Moos, Universität Bayreuth (Deutschland), A. Ruchets, J. Zosel, U. Guth, Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, Waldheim (Deutschland)

P1.13 **Innovatives Sensorsystem und Auswerteverfahren zum Monitoring bei der Lebensmittelüberwachung 606**
R. Seifert, H. B. Keller, Karlsruher Institut für Technologie, Eggenstein-Leopoldshafen (Deutschland), H. Kohler, Hochschule Karlsruhe Technik und Wirtschaft (Deutschland)

P1.14 **Fasergekoppelter konfokaler Sensor zur exakten Abstandsregelung für maskenlose Lithografieanwendungen 610**
J. Kirchner, R. Mastylo, U. Gerhardt, T. Sasiuk, L. Weidenfeller, M. Hofmann, M. Kühnel, S. Sinzinger, E. Manske, TU Ilmenau (Deutschland)

P1.15 **Mittels aerosolbasierter Kaltabscheidung bei Raumtemperatur hergestellte schichtförmige NTC-Thermistorbauelemente 617**
R. Moos, M. Schubert, C. Reichl, J. Kita, Universität Bayreuth (Deutschland), C. Münch, Vishay Electronic GmbH, Selb (Deutschland)

P1.16 **Vergleichende Betrachtung verschiedener Methoden zur Bestimmung der Inhomogenität von Thermoelementen 621**
P. Germanow, P. Mehring, H. Neumann, Reckmann GmbH, Hagen (Deutschland), S. Augustin, T. Fröhlich, H. Mammen, TU Ilmenau (Deutschland)

P1.17 **Herstellung von multipolaren, kunststoffgebundenen SmCo-Magneten für Sensoranwendungen**
A. Gardocki, Evitron sp. z o.o., Police (Polen)
(Beitrag lag nicht vor)

P1.18 **Objektrekonstruktion in der Mikrowellen-Defektoskopie 627**
T. Gagelmann, M. Nesterov, S. Wöckel, Institut für Automation und Kommunikation, Magdeburg (Deutschland), C. Li, S. Hantscher, J. Auge, Hochschule Magdeburg-Stendal (Deutschland)

P1.19 **Dynamisches haptisches Feedback für Multi-Touch-Interaktionen 634**
F. Edner, U. Steinmann, Otto-von-Guericke-Universität, Magdeburg (Deutschland)

P1.20 **Design und Modellierung eines 3D-Nanotasters 642**
J. Thiesler, G. Dai, Physikalisch Technische Bundesanstalt, Braunschweig (Deutschland), R. Tutsch, Technische Universität Braunschweig (Deutschland)

P1.21 **Simulation von thermischen Einflüssen auf MEMS-Mikrofone 647**
H. Ebbinghaus, G. Feiertag, Hochschule für angewandte Wissenschaften, München (Deutschland), S. Walser, TDK Electronics AG, München (Deutschland)

P1.22 **Dynamic measurement methods for solid electrolyte gas sensors 651**
D. Schönauer-Kamin, N. Donker, R. Moos, Universität Bayreuth (Deutschland), A. Ruchets, J. Zosel, U. Guth, Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, Waldheim (Deutschland)

P1.23 **CH_4-Sensitivity of thermoelectric gas sensors 654**
J. Wohlrab, G. Hagen, R. Moos, Universität Bayreuth (Deutschland), H. Kohler, Hochschule Karlsruhe (Deutschland)

P2: Messsysteme

P2.1 **Messgenauigkeit und Reproduzierbarkeit bei der Messung der Brillouin-Frequenzverschiebung für die faseroptische Bestimmung der Temperaturverteilung 656**
A. Wosniok, BAM Bundesanstalt für Materialforschung und -prüfung, Berlin (Deutschland)

P2.2 **Rotationsgeschwindigkeitsmessung mithilfe von "Self-Mixing"- Interferometrie 663**
Y. Liu, J. Liu, Q. Song, ChenYang Technologies GmbH & Co. KG., Finsing (Deutschland), R. Kennel, Technische Universität München (Deutschland)

P2.3 **Entwicklung eines akustischen Füllstandsensors für mobile Schüttgutsilos 671**
P. Sivasothy, H. Cavuldak, V. Rjasanzew, J. Seewig, Technische Universität Kaiserslautern (Deutschland)

P2.4 **ExtruSens - Low-Cost-Spektroskopie zur Inline-Prozesskontrolle in der Kunststoffverarbeitung 679**
N. Halmen, C. Kugler, B. Baudrit, T. Hochrein, M. Bastian, SKZ - Das Kunststoff-Zentrum, Würzburg (Deutschland), H.-P. Jungbauer, H. Nitsche, J & M Analytik AG, Essingen (Deutschland)

P2.5 **Detektion von Kunststoffen in Wasser mithilfe von Fluoreszenz 687**
M. Wohlschläger, M. Versen, TH Rosenheim (Deutschland)

P2.6 **Analysis of photoelastic properties of monocrystalline silicon 695**
M. Stöhr, S. Schönfelder, Hochschule für Technik, Wirtschaft und Kultur Leipzig (Deutschland), G. Gerlach, TU Dresden (Deutschland), T. Härtling, Fraunhofer-Institut für Keramische Technologien und Systeme, Dresden (Deutschland)

P2.7 **Study of interdigital electrode structures for the detection of water spray 700**
J. Döring, L. Tharmakularajah, K.-L. Krieger, ITEM Universität Bremen (Deutschland)

P2.8 **Simultaneous measurement of resistance and temperature changes in bridge circuits 708**
K. Gintner, Hochschule Karlsruhe (Deutschland)

P2.9 **Comparison of deep feature extraction techniques for varying-length time series from an industrial piercing press 715**
C. Thiel, BENTELER Steel/Tube GmbH, Paderborn (Deutschland), C. Steidl, B. Henning, Universität Paderborn (Deutschland)

P2.10 **Entwicklung eines taktilen Mikrotaster-Messsystems für Hochgeschwindigkeits-messung von Form, Rauheit und mechanischen Eigenschaften 720**
M. Fahrbach, E. Peiner, Technische Universität Braunschweig (Deutschland)

P2.11 **Formbestimmung von Si-Mikrotasterspitzen mittels Rechteckstrukturen unterschiedlicher Breite**
J. Langfahl-Klabes, T. Ahbe, S. Bütefisch, U. Brand, L. Koenders, PTB Braunschweig (Deutschland), R. Tutsch, TU Braunschweig (Deutschland)
(Beitrag lag nicht vor)

P2.12 **Tastsystem zur Bestimmung der instrumentierten Eindringhärte 726**
T. Frank, C. Maier, M. Kermann, A. Grün, A. Cyriax, J. Patel, T. Ortlepp, Forschungsinstitut für Mikrosensorik und Photovoltaik GmbH, Erfurt (Deutschland)

P2.13 **Korrelation von wiederholten Messungen in einer Modenverwirbelungskammer 730**
C. Cammin, R. Heynicke, G. Scholl, Universität der Bundeswehr Hamburg (Deutschland)

P2.14 **Intelligenter Zahnriemen 738**
D. Großkurth, G. Martin, Technische Universität Darmstadt (Deutschland)

P2.15 **Energieautark arbeitendes Indoor-Smart-Gardening-System mit drahtlosem Monitoring und automatisierter Bewässerung 744**
E. Mackensen, A. Rombach, A. Spitznagel, J. Klose, Hochschule Offenburg (Deutschland)

P2.16 **Construction of an universal blood collecting system with integrated analytics 751**
S. Schlegel, E. Krumnow, S. Zinn, A. H. Foitzik, Technical University of Applied Science Wildau (THWi) (Deutschland) M. Riester, midge medical GmbH, Berlin (Deutschland), M. Minieri, Università degli Studi di Roma “Tor Vergata”, Rom (Italien)

P2.17 **Optischer Messplatz zur präzisen und hochdynamischen Bestimmung des zweidemensionalen Kippwinkels von vektoriell positionierenden 2D-Mikrospiegeln**
T. Sardner, Fraunhofer-Institut for Photonic Microsystems (IPMS), Dresden (Deutschland)
(Beitrag lag nicht vor)

P2.18 **Characterization of the single point precision of steel gear wheel measurements using tactile coordinate measurement machines in scanning mode 756**
A. M. Müller, S. Metzner, T. Hausotte, Institute of Manufacturing Metrology (FMT), Erlangen (Deutschland)

P3: Anwendungen

P3.1 **Mit Sensoren und Blockchain ins Internet der Dinge!**
M. Haid, N. Berezowski, J. Biswas, I. Blanc, I. Boyaci, Hochschule Darmstadt (Deutschland)
(Beitrag lag nicht vor)

P3.2 **Präzisionsmessungen bei nicht geregelter Temperatur**
D. Sumin, R. Tutsch, Technische Universität Braunschweig (Deutschland)
(Beitrag lag nicht vor)

P3.3 **Modellbasierte Analyse von Stahlbau-Strukturen, am Beispiel Maßprüfung und Vollständigkeitsprüfung an Schienenfahrzeug-Wagenkästen mittels 3D-Scan und Software-Algorithmus 763**
U. Jurdeczka, ALSTOM Transport Deutschland GmbH, Salzgitter (Deutschland)

P3.4 **Dimensionale Messtechnik in der digitalen Produktion - Aufgaben, Vernetzung und Schnittstellen 769**
D. Imkamp, Carl Zeiss IMT GmbH, Oberkochen (Deutschland)

P3.6 **Untersuchung eines Prozessroboters als Messroboter in der Montage 774**
A. Blum, Universität des Saarlandes, Saarbrücken (Deutschland), R. Müller, M. Scholer, A. Kanso, ZeMA - Zentrum für Mechatronik und Automatisierungstechnik gemeinnützige GmbH, Saarbrücken (Deutschland)

P3.7 **Überprüfung und Bewertung des Schriftbildes sowie des Anlagenzustandes industrieller Signieranlagen mit maschinellem Lernen 781**
N. Immesberger, S. Klein, A. Schütze, Zentrum für Mechatronik und Automatisierungstechnik, Saarbrücken (Deutschland), M. Jochum, Aktien-Gesellschaft der Dillinger Hüttenwerke, Dillingen (Deutschland)

P3.8 **Hardware-in-the-loop test concept for an energy-optimized process control 789**
T. Kull, G. Fischerauer, Universität Bayreuth (Deutschland), B. Zeilmann, Richter R&W Steuerungstechnik GmbH, Ahorntal (Deutschland)

P3.9 **Systematische thermophysikalische Charakterisierung von Wärmedämmschichten**
K. Knopp, P. Lenski, D. Ochs, M. Zänglein, M. Müller, J. Hartmann, Hochschule für angewandte Wissenschaften Würzburg - Schweinfurt (Deutschland), M. Arduini, F. Hemberger, S. Vidi, J. Manara, Bayerisches Zentrum für Angewandte Energieforschung e.V., Würzburg (Deutschland), F. Möller, Rauschert Heinersdorf-Pressig GmbH, Pressig (Deutschland)
(Beitrag lag nicht vor)

P3.10 **Proof-of-concept trial of the portable electronic nose PEN3 for detection of formic acid concentration in the beehive 794**
C. Tiebe, M. Hofmann, Bundesanstalt für Materialforschung und -prüfung, Berlin (Deutschland), A. Genath, R. Einspanier, Freie Universität Berlin (Deutschland)

P3.11 **Recommendations for optimal sampling and filter rates in liquid chromatography 800**
D. T. Marehn, D. Wilhelm, AnaTox GmbH & Co. KG, Fürstenwalde (Deutschland), H. Pospisil, Technische Universtität Wildau (Deutschland), R. Pizzoferrato, Università degli Studi Di Roma ”Tor Vergata“, Rom (Italien)

P3.12 **Integration of Bragg grating sensors in components made of carbon fiber reinforced polymers 806**
T. Sauer, S. Kefer, W. Ruppert, R. Hellmann, M. Kaloudis, Technische Hochschule Aschaffenburg (Deutschland)

Key Words 811

DOI 10.5162/sensoren2019/Festvortrag

Für alle Zeiten und Culturen! Das revidierte Internationale System der Einheiten

Joachim Hermann Ullrich[1]
[1]*Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Deutschland*

Zusammenfassung

Der 20. Mai 2019 markiert eine Zäsur in der Entwicklung der physikalischen und technischen Maßeinheiten. An diesem Tag, der zugleich der Weltmetrologietag des Jahres 2019 ist, trat ein Paradigmenwechsel im Internationalen System der Einheiten (Système international d`unités, kurz: SI) in Kraft. Vorbei die Zeiten, in denen ein Metallzylinder definiert, was ein Kilogramm sein soll, oder eine spezifisch herzustellende Isotopenmischung von Wasser die Einheit der Temperatur festlegt. Naturkonstanten, darunter die Lichtgeschwindigkeit *c*, das Planck´sche Wirkungsquantum *h* und die Boltzmann-Konstante *k,* definieren ab diesem Zeitpunkt alle physikalischen Einheiten. Das so definierte Einheitensystem bildet eine Art „universelle Sprache“ und hat einzigartige, insbesondere auch technologisch relevante Eigenschaften: Die Naturkonstanten bevorzugen, anders beispielsweise als das Urkilogramm oder der Tripelpunkt von Wasser, keine speziellen Punkte auf der jeweiligen Einheitenskala, so dass viele Einheiten nun entlang der gesamten Skala realisiert werden können. Zum anderen ist das revidierte SI offen für technologische Innovationen bei der Realisierung der Einheiten: jede passende Gleichung der Physik kann zur Realisierung herangezogen werden und die Genauigkeit der Realisierung kann entsprechend des technischen Fortschritts gesteigert werden, ohne Begrenzung durch die Definition selbst.

Keywords: Metrologie, Internationales Einheitensystem, Meterkonvention, Naturkonstanten

Die Anfänge des Messens

Schon in den alten Hochkulturen wussten die Menschen, dass Eigenschaften von Körpern, Substanzen und Phänomenen gemessen werden können. Das Prinzip des Messens lautete damals wie heute: Um das Ergebnis einer Messung auszudrücken, braucht man eine Zahl und eine Einheit. Die Einheit ist typischerweise durch ein Artefakt oder verkörpertes Normal festgelegt, das einen speziellen Wert der Messgröße repräsentiert.
Das berühmteste (und letzte) Artefakt ist das Urkilogramm, das 1889 von der Meterkonvention als die Einheit für die Masse definiert worden war. Im November 2018 entschied die 26. Generalkonferenz für Maße und Gewichte (CGPM), das Urkilogramm abzulösen und gleichzeitig drei weitere Basiseinheiten neu zu definieren. Diese Entwicklung kann man als Konsequenz der Tatsache sehen, dass die Natur im Kleinsten „quantisiert“ ist und durch die Quantenmechanik beschrieben wird. Quanten sind die kleinsten natürlich vorkommenden Einheiten für bestimmte Größen. So kann man einen elektrischen Strom messen, indem man in einem Leiter fließende, einzelne Elektronen, also die kleineste Einheit der elektrischen Ladung, in einer vorgegebenen Zeit zählt und so „das neue Ampere“ realisiert.

Von Artefakten zu Quanten – mit diesen beiden Schlagworten lässt sich die generelle Entwicklung bei den Einheiten umreißen. Schon Maxwell und Planck forderten vor mehr als 100 Jahren, die Artefakte als Grundlage der Einheiten abzuschaffen. Insbesondere Planck schlug das ultimative Einheitensystem vor, das auf den – soweit heute experimentell nachgewiesen – unverrückbaren Naturkonstanten beruhen sollte. Diesen Weg haben die Metrologen, die Wissenschaftler, die sich mit der Messkunst beschäftigen, inzwischen höchst erfolgreich beschritten.

Die ersten einheitlichen Einheiten

Zunächst ein Blick zurück zur Französischen Revolution, die auch beim Thema Messen grundlegende Änderungen hervorbrachte. Das Ziel der Revolutionäre lautete, neue Einheiten zu schaffen, die für alle Zeiten und für alle Völker gelten sollten: à tous les temps, à tous les peuples. Die neuen Einheiten sollten nicht einem bestimmten Kaiser, einem bestimmten König oder einem bestimmten Land zugeordnet werden, sondern sie sollten allen Menschen auf der Welt gehören. Und sie sollten auf dem Dezimalsystem beruhen. Selbst der Tag wurde zunächst in zehn Stunden eingeteilt. Hier kehrte man allerdings nach kurzer

Zeit zum alten Zwölfersystem der Babylonier zurück, das bis heute bei der Zeitmessung gilt. Als Grundlage der Einheiten sollte keine menschliche Eigenschaft mehr gelten, sondern die Eigenschaften der Erde. Dazu bildete die Académie des Sciences im Jahr 1790 ein Gremium, dem unter anderem Joseph-Louis Lagrange (Mathematiker und Astronom) und Pierre-Simon Laplace (Mathematiker, Physiker und Astronom) angehörten.



Abb.1: Diese Briefmarke illustriert die große Bedeutung der „metrischen Idee“ und der neuen Einheit Meter, die von der Größe der Erde abgeleitet war. (Quelle: Musée de la Poste)

Dieses Gremium definierte die neue Längeneinheit als den vierzigmillionsten Teil eines Erdmeridians. Die Astronomen Jean-Baptiste Delambre und Pierre-François Méchain bekamen den Auftrag, einen Teil des Erdmeridians, der durch Paris lief, genau zu vermessen. Mitten in den Revolutionswirren triangulierten die beiden die Strecke von Dünkirchen bis Barcelona. Damit war die neue Längeneinheit Meter geschaffen. Da die Realisierung des Meters durch erneutes Vermessen zu aufwendig gewesen wäre, zog man sich wieder auf das Konzept der Maßverkörperung zurück und goss aus Platin ein Artefakt, das „Mètre des Archives“. Die neue Einheit sowie das ganze neue, dezimale Messsystem sollten allgemein und universal gültig werden. Tatsächlich gelang dies aber erst fast hundert Jahre später.

Die Meterkonvention

Im Jahre 1875 führte das Thema „Einheiten“ 17 Staaten, darunter die meisten der damals wichtigsten Industriestaaten, in einem der ersten internationalen Verträge zusammen: der Meterkonvention. Interessanterweise waren die USA dabei, das Vereinte Königreich jedoch nicht. Es hatte seine eigenen Einheiten und war (als Commonwealth) groß genug, um eigene Wege zu gehen. Heute ist es Mitglied der Meterkonvention.
Zur Meterkonvention gehören mehrere Organe: die „Generalkonferenz für Maße und Gewichte“ (CGPM) tagt alle vier Jahre und ist allein beschlussfähig. Die laufenden Geschäfte übernehmen das „Internationale Komitee für Maße und Gewichte“ (CIPM) sowie diverse Konsultativkomitees (CC).
Mit der Unterzeichnung des Vertrages von 1875 akzeptierten alle beteiligten Staaten die neue Einheit Meter und erhielten 1889 Kopien des sogenannten Urmeters oder Internationalen Meterprotoypen, eines Präzisions-Strichmaßes gefertigt aus einer Legierung von Platin und Iridium, mit der Länge des „Mètre des Archives“. Aus dem Meter abgeleitet wurde das Kilogramm als die Masse eines Kubikdezimeters Wasser bei 4 Grad Celsius definiert. Die Mitgliedsstaaten erhielten eine Kopie des Urkilogramm-Stückes, ebenfalls aus PtIr gefertigt. Der Urmeter und das Urkilogramm wurden in einem neugegründeten wissenschaftlichen Institut in Paris, dem „Internationalen Büro für Maße und Gewichte“ (BIPM), das neben der Generalkonferenz und dem internationalen Komitee das dritte Organ der Meterkonvention ist, aufbewahrt und waren fortan für die Mitglieder die gültigen Referenznormale für Länge und Masse.
Jetzt konnte man alle Längenmessmittel oder Gewichtsstücke direkt oder über mehrere Zwischenschritte mit den nationalen Normalen vergleichen. In der Metrologie wird dies „metrologische Rückführung“ genannt. Sie sorgt dafür, dass in allen Mitgliedsstaaten der Meterkonvention ein einheitliches Maßsystem verwendet wurde, und ist bis heute ein sehr erfolgreiches Prinzip.
Die Einheit der Zeit, die Sekunde, stand nicht unter der Verantwortung der Meterkonvention, sondern hier war bis zum Jahr 1967 die International Astronomical Union zuständig.
Alle drei Einheiten, Meter, Kilogramm und Sekunde, beruhten entsprechend der Vision der französischen Revolutionäre auf den Eigenschaften der Erde – entweder auf ihren Ab-

messungen oder ihrer Rotation. Allerdings hatten damals schon einige Vordenker weiter gedacht: Der Physiker James Clerk Maxwell hatte schon 1870 gemahnt, die so definierten Einheiten seien keine wahren Invarianten, denn „die Eigenschaften unseres Planeten können sich ändern, und es wäre immer noch unser Planet, aber wenn sich die Eigenschaften eines Atoms änderten, wäre es nicht mehr dasselbe Atom." Daher sei es besser, die Eigenschaften von Atomen oder Molekülen als Grundlage für die Definition der Einheiten zu nehmen.

Physikalisch-Technische Reichsanstalt

Etwa zur gleichen Zeit, 1887, wurde in Berlin das weltweit erste Metrologieinstitut gegründet: die Physikalisch-Technische Reichsanstalt (PTR), die direkte Vorgängerin der heutigen Physikalisch-Technischen Bundesanstalt (PTB). Werner von Siemens und Hermann von Helmholtz waren neben Wilhelm Förster, dem Direktor der Normal-Eichungs-Kommission, der obersten Eichbehörde des Deutschen Reichs, die beiden Gründungsväter. Helmholtz wurde der erste Präsident der PTR. Er galt als einer der prominentesten Wissenschaftler seiner Zeit, Werner von Siemens war einer der prominentesten Unternehmer. Und genau in diesem Spannungsfeld – zwischen Wissenschaft und Industrie – bewegt sich die PTB heute noch; sie betreibt Wissenschaft in der Messtechnik für Gesellschaft, Industrie und nicht zuletzt als Grundlagenforschung.



Abb. 2: Das Observatorium auf dem Gelände der ehemaligen Physikalisch-Technischen Reichsanstalt (PTR) in Berlin-Charlottenburg, dem heutigen Institut Berlin der PTB. (Quelle: PTB)

Das sogenannte Observatorium, das heute noch restauriert auf dem historischen Gelände in Berlin-Charlottenburg in der Nähe des Ernst-Reuter-Platzes steht, war das erste Laborgebäude der PTR und gleichzeitig das erste Großlabor im Deutschen Reich. Werner von Siemens war der festen Überzeugung, dass es ein staatliches Institut brauche, das Präzisionsmessungen insbesondere auch für die elektrischen Einheiten durchführt und ganz allgemein die Einheitlichkeit des Messwesens in Deutschland sicherstellt. Da die Politik aus seiner Sicht nicht schnell genug vorankam, unterstützte er die Gründung der PTR mit seinem privaten Geld und baute insbesondere das Observatorium mit auf.
In diesem Gebäude fanden damals Messungen statt, die letztlich zur Entdeckung der Quantenmechanik führten. An einem sogenannten Schwarzen Strahler untersuchten Otto Lummer und seine Kollegen die Zusammenhänge zwischen Farbe und Intensität des Lichts. Sie fanden kleinste, unerklärliche Abweichungen von der damaligen theoretischen Vorhersage. Lummer rief noch am selben Abend Max Planck an, der später PTR-Kurator wurde. Planck konnte das Problem nur lösen, indem er in einem „Akt der Verzweiflung" (wie er selber sagte) annahm, die Energie der Strahlung werde nicht kontinuierlich, sondern in kleinen Päckchen transportiert, die er Quanten nannte. Seine Formel wurde später weltberühmt; die Konstante, die er dafür einführte, heißt seitdem Planck-Konstante. Somit wurde aus einer zunächst nicht breit beachteten Untersuchung an der PTR letztlich die Geburt der Quantenmechanik, die heute mit all ihren verschiedenen Facetten zu etwa 60 Prozent der Weltwirtschaft beiträgt.

Natürliche Maßeinheiten

Max Planck, im Kuratorium der Reichsanstalt tätig und damit in Gedanken den Einheiten nahe, sinnierte 1900 in seiner ebenfalls berühmt gewordenen Veröffentlichung zur Strahlung des schwarzen Körpers in der Zeitschrift „Annalen der Physik": „Demgegenüber dürfte es nicht ohne Interesse sein zu bemerken, dass mit Zuhülfenahme der beiden auftretenden Constanten a und b (Anmerkung: eine war seine später sogenannte Planck-Konstante, die andere die Boltzmann-Konstante) die Möglichkeit gegeben ist, Einheiten für Länge, Masse, Zeit und Temperatur aufzustellen, welche, unabhängig von speciellen Körpern oder Substanzen, ihre Bedeutung für alle Zeiten und für alle, auch ausserirdische und aussermenschliche Culturen notwendig behalten und welche daher als ‚natürliche Maaßeinheiten' bezeichnet werden können." Das waren im Jahr 1900 äußerst visionäre Gedanken.
Genau diesen Weg haben die Metrologen nun, mehr als hundert Jahre später, vollendet. Allerdings verlief er nicht ganz so scharf und konsequent, wie es Planck für die Einheiten für Länge, Zeit, Masse und Temperatur gefordert

hat. Planck definierte sie mithilfe von Konstanten, von denen sich die meisten im gerade beschlossenen revidierten Einheitensystem wiederfinden, mit der Lichtgeschwindigkeit, seiner eigenen Planck-Konstanten, der Boltzmann-Konstanten sowie der (im heutigen SI nicht verwendeten) Gravitationskonstante. Doch die so definierten Einheiten sind extrem klein, die Planck-Zeit zum Beispiel beträgt nur etwa $5 \cdot 10^{-44}$ Sekunden. Eine so kleine Größe ist in der Praxis nicht gut verwendbar. Für theoretische und Hochenergie-Physiker sind die Planck-Einheiten dagegen sehr interessant und werden bis heute genutzt. Sie zeigen uns die Grenzen des heutigen Wissens auf, die genau bei dieser Planck-Zeit nach dem Urknall liegen: Was davor war, davon wissen wir so gut wie nichts, dort versagen unsere physikalischen Gesetze.

Das Internationale Einheitensystem

Das Internationale Einheitensystem SI (Système International d'unités) wurde 1960 von der Generalkonferenz für Maße und Gewichte eingeführt. Es beruhte bis zu seiner Revision am 20. Mai 2019 auf den sieben Basiseinheiten für Länge, Masse, Zeit, elektrische Stromstärke, Temperatur, Stoffmenge und Lichtstärke, von denen alle anderen Einheiten abgeleitet wurden, und bot einen Satz kohärenter Einheiten, um alles Erdenkliche zu messen, was von wissenschaftlicher, wirtschaftlicher, gesundheitlicher oder sonstiger Relevanz ist.
Das SI ist weltweit gültig. Denn inzwischen hat die Meterkonvention 102 Mitglieds- und assoziierte Länder, vier internationale Organisationen sowie 153 sogenannte Designierte Institute. Sie umfassen etwa 98 Prozent der gesamten Weltwirtschaft. Zusammengenommen ist dies eine globale Infrastruktur und damit die eigentliche Grundlage für globalen Handel. Weltweit sind etwa 65 000 Prüf- und Kalibrierlaboratorien nach der internationalen Norm ISO/IEC 17025 akkreditiert. Damit haben sie sich verpflichtet, alle Messungen auf das SI rückzuführen – eine wichtige Voraussetzung dafür, um auf verschiedenen Kontinenten zu produzieren und international Handel zu treiben.
Dieses System wurde in der Folge kontinuierlich an neueste wissenschaftliche Erkenntnisse und technische Entwicklungen angepasst und diese Anpassungen für die Sekunde, den Meter und später das Ampere zeigen die Richtung zur jetzigen grundlegenden Revision auf. Das soll ein Blick auf die „Vorreiter" Sekunde und Meter erläutern:
Bei jeder Uhr gibt es ein schwingendes Element, ein Pendel. Je kürzer ein Pendel, desto schneller schwingt es und desto genauer ist die Uhr. Bei einer Cäsium-Atomuhr ist das „Pendel" eine bestimmte elektronische Übergangsfrequenz des Cäsium-133-Atoms; es schwingt rund neun Milliarden Mal pro Sekunde. Exakt 9 192 631 770 Schwingungen dieses kleinen Pendels definieren seit 1967 eine Sekunde. Damit ist Maxwells Forderung, man solle Einheiten auf der Basis atomarer Eigenschaften beruhen lassen, erfüllt. Im Jahr 1967 konnte man die Sekunde auf 14 Stellen hinter dem Komma genau messen. Heute sind es 16 Stellen hinter dem Komma – und damit eine Sekunde Abweichung in 180 Millionen Jahren. Nötig ist diese enorme Präzision für die Satellitennavigation, das „global positioning system" (GPS) aber auch für die physikalische Grundlagenforschung. Interessant ist dabei, dass bei gleichbleibender Definition die Realisierung der Einheit inzwischen um den Faktor 100 gesteigert werden konnte. Das zeigt, dass diese Art der Definition offen für Innovationen ist.
Genauso ist es beim Meter. Im Jahr 1983 legte man die Lichtgeschwindigkeit im Vakuum zahlenmäßig fest. Mithilfe dieses Zahlenwertes und einer gemessenen Zeit wurde der Meter definiert. Heute messen wir kleinste Distanzen bis hinab zum Pikometer. Gleichzeitig sind größte Distanzen zugänglich, etwa die Entfernung zwischen Erde und Mond – und dies mit einer Messunsicherheit von wenigen Zentimetern. Wie bei der Sekunde ist der wissenschaftliche Fortschritt nur durch die technischen Möglichkeiten begrenzt, nicht durch die Definition selbst.
Anders sah es beim Kilogramm aus: Die Einheit der Masse war bis zum 20. Mai 2019 durch das Urkilogramm definiert – jenes Artefakt, das ursprünglich aus der Französischen Revolution kommt und 1889 in der Meterkonvention als internationaler Prototyp des Kilogramms festgelegt wurde.
1889 wurde das Urkilogramm als Platin-Iridium-Zylinder hergestellt und in eine Glasvitrine in einen Tresor gelegt. Bis heute wurde es nur dreimal herausgenommen. Denn es ist äußerst wertvoll – alle Masse-Messungen auf der Welt hingen bis zum 20. Mai 2019 an diesem Stück Metall. Bei diesen drei Gelegenheiten wurde es mit seinen besten Kopien verglichen: mit den offiziellen Kopien, die auch in Paris liegen, sowie mit den nationalen Normalen in den nationalen Metrologieinstituten überall auf der Welt. Bei diesen Vergleichsmessungen zeigte sich ein Trend. Zwar bleibt das Kilogrammstück – per definitionem – immer ein Kilogramm. Aber relativ dazu bewegen sich die anderen Normale, und zwar im Bereich von etwa 50 µg über 100 Jahre. Das ist sehr wenig, aber für Metrologen eine sehr unbefriedigende Situation.

Zudem gab es keine Innovationsoffenheit: Die Realisierung des Kilogramms konnte nie genauer werden, als dieses Artefakt und seine Eigenschaften es vorgaben.
Weil sich vermutlich auch das Urkilogramm selber verändert hat, ruht die Einheit der Masse auf unsicherer Basis. Und davon sind auch jene Einheiten betroffen, in deren Definition das Kilogramm eingeht: die Candela (die Einheit der Lichtstärke), das Mol (die Einheit der Stoffmenge) und das Ampere (die Einheit der Stromstärke).
Die bisherige Definition des Ampere beruht auf einer Kraft zwischen zwei stromdurchflossenen Leitern, die unendlich dünn und unendlich lang sind. Das lässt sich experimentell nur näherungsweise umsetzen. Daher wird bereits seit 1990 die Einheit der Stromstärke anders – und zudem sehr viel genauer – realisiert. Grundlage dafür sind Quanteneffekte für den elektrischen Widerstand und die elektrische Spannung sowie die dazugehörigen zwei Konstanten: die von-Klitzing-Konstante und die Josephson-Konstante. Beide beruhen auf Verhältnissen fundamentaler Konstanten, nämlich der Planck-Konstante und der Elektronenladung. Damit bewegte sich allerdings die Realisierung des Ampere außerhalb des SI und seiner Regeln.
Ähnlich war es bei der Temperatur. Sie war definiert durch den absoluten Nullpunkt und den sogenannten Tripelpunkt des Wassers; diejenige Temperatur, bei der Wasser gleichzeitig in festem, flüssigem und gasförmigem Zustand existieren kann. Bei vorgegebenem Luftdruck, hier dem Normaldruck der Atmosphäre, und bei Wasser möglichst ohne jede Verunreinigung ist dies nur bei einer einzigen Temperatur möglich. Auf diesen beiden Fixpunkten baute die Temperaturskala auf, indem die Differenz in genau hundert Teile, Grad Kelvin, unterteilt wird. Also war auch die Definition der Temperatur vollständig unabhängig von allen anderen Einheiten.
Insgesamt basierten also nur Sekunde und Meter auf den Zahlenwerten von Konstanten. Bei den fünf anderen Basiseinheiten (Kilogramm, Ampere, Kelvin, Mol und Candela) gab es eine nicht ideale Situation. Gelänge es, auch sie (und damit das ganze System) quantenbasiert zu machen, würden sie alle innovationsoffen werden. In der Tat ist es nun gelungen, ein kohärentes System zu schaffen, das einheitlich auf Quanten und fundamentalen Naturkonstanten begründet ist.

Naturkonstanten im SI

Im revidierten SI sind jetzt für alle bisherigen Basiseinheiten die Zahlenwerte von „definierenden Konstanten“ – unter ihnen Naturkonstanten – festgelegt worden, so wie schon zuvor bei der Sekunde und beim Meter. Auf dieser Basis werden mithilfe der Gleichungen der Physik, also mithilfe des jetzigen Wissens von den physikalischen Zusammenhängen, die Einheiten generiert. Das Ganze soll für alle Zeiten und Kulturen im gesamten Universum (soweit wir es denn kennen) gültig sein, wie Max Planck es gefordert hat.

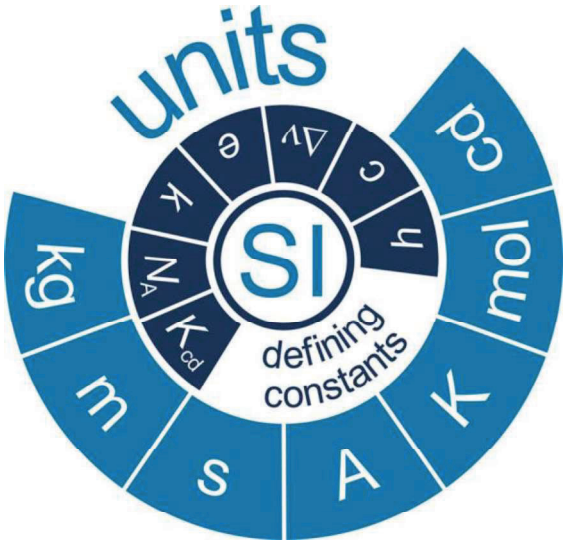


Abb. 5: Das revidierte SI beruht auf einem konsistenten und kohärenten Satz von sieben „definierenden Konstanten“.

Im alten Einheitensystem waren Einheiten über Artefakte festgelegt. Mithilfe dieser Einheiten wurden (mit einer gewissen Messunsicherheit) die Naturkonstanten gemessen. Im neuen SI ist es anders herum: Die Werte von Naturkonstanten sind zahlenmäßig festgelegt, und daraus werden die Einheiten abgeleitet. Die Sekunde ist über eine bestimmte Zahl der Schwingungen des „Pendels“ im Cäsium-Atom definiert. Dieselbe Konstante plus die Lichtgeschwindigkeit wird für die Definition des Meters gebraucht. Für das Kilogramm musste noch der numerische Wert einer dritten Konstante festgelegt werden: die Planck-Konstante.

Das Avogadro-Experiment

Dem Ziel, die Planck-Kontante so genau wie möglich zu ermitteln und schließlich festzulegen, dienten zwei fundamental unterschiedliche experimentelle Herangehensweisen, beides wissenschaftliche Großprojekte, die an verschiedenen nationalen Metrologieinstituten über mehrere Jahrzehnte hinweg betrieben wurden: das Avogadro- und das Wattwaagen-Experiment.
Im Avogadro-Projekt der PTB zusammen mit internationalen Partnern wurde die Planck-

Konstante mithilfe einer speziell gefertigten Siliziumkugel ermittelt. Sie besteht zu einem sehr hohen Anteil aus einem von drei natürlichen vorkommenden Silizium-Isotopen. Diese Isotopenreinheit garantiert, dass die Masse der Atome in der Kugel sehr genau gemessen werden kann. Kennt man die Masse eines Atoms, lässt sich die Masse der Kilogrammkugel über das Zählen ihrer Atome bestimmen.



Abb. 6: Das rundeste makroskopische Objekt der Welt: die Siliziumkugel der PTB im internationalen Avogadro-Projekt für ein „neues“ Kilogramm und ein „neues“ Mol.

Doch die 10^{25} Atome lassen sich nur mit indirekten Methoden zählen. Diese nutzen die Tatsache aus, dass Silizium regelmäßige Kristalle bildet. Misst man den Gitterparameter, das Volumen der gesamten Kugel und einige weitere Parameter, lässt sich die Zahl der Atome und die Planck-Konstante ermitteln.
Bei der Volumenmessung der in einer komplexen Fertigungskette hergestellten Si-Kugel mithilfe eines speziellen Kugel-Interferometers lassen sich Abweichungen zwischen Maximum und Minimum auf der Kugeloberfläche von 16 Nanometern ermitteln. Auf die Dimensionen der Erde übertragen, hieße das: Die größten Höhenunterschiede zwischen tiefstem Ozean und höchstem Berg betrügen nur zwei Meter. Und dies wird mit einer Genauigkeit von zwei oder drei Millimetern gemessen. Das liegt an der Grenze der heutigen Längenmesstechnik. Als nächster Schritt wird mit einer der besten Waagen der Welt die Masse der gesamten Kugel ermittelt. Die Masse der Kugel geteilt durch die Zahl der Atome ergibt die Masse eines einzelnen Atoms – und letztlich die Planck-Konstante.

Keiner dieser Werte lässt sich auf direktem Wege messen. Die gewünschten Konstanten (neben der Planck- auch die Avogadro-Konstante für die Einheit der Stoffmenge, Mol) werden rechnerisch aus den Messungen ermittelt. Wird, so wie im revidierten SI, der Zahlenwert der Planck- (Avogadro-) Konstanten festgelegt, realisiert man über das Abzählen der Atome in der Siliziumkugel die Einheit der Masse (Kilogramm) und der Stoffmenge (Mol).

Das Wattwaagen-Experiment

Eine zweite Methode auf dem Weg zu einem neuen Kilogramm ist die sogenannte Watt- oder Kibble-Waage. Das Prinzip ist relativ einfach: Auf der einen Seite der Waage liegt das zu messende Massestück. Die entsprechende Gewichtskraft wird auf der anderen Seite der Waage mit einer elektromagnetischen Kraft kompensiert, wobei die elektrischen Messgrößen mit elektrischen Quantennormale bestimmt werden. Als Ergebnis erhält man eine Gleichung, die den Zusammenhang zwischen Masse und Planck-Konstante herstellt. Legt man den Zahlenwert der Planck-Konstanten fest, so kann man mit einer solchen Waage das Kilogramm realisieren. Die Methode ist vollständig anders als in dem Siliziumkugel-Projekt. Über den Vergleich der beiden unterschiedlichen Methoden wird sichergestellt, dass die zugrundeliegenden physikalischen Gleichungen mit der geforderten Genauigkeit übereinstimmen.

Elektronen zählen

Elektrischer Strom besteht aus einzelnen Elektronen, die in einem Zeitintervall durch einen Leiter fließen. Gelingt es, einzelne Elektronen mit einer bestimmten Frequenz durch einen Leiter zu schicken und diese Frequenz zu messen, dann lässt sich über den festgelegten Zahlenwert der Elektronenladung der Strom definieren und messen. Konzeptionell einfach, jedoch in der Praxis sehr komplex. Zur Kontrolle der Elektronen verwendet man Quantenpunkte: Nanostrukturen, mit denen sich einzelne Elektronen einfangen und einzeln hindurchschieben lassen; sie werden auch „Elektronenpumpen“ genannt. Elektronen sind „Quantenteilchen“, bei denen man nie mit hundertprozentiger Genauigkeit wissen kann, ob genau ein Elektron „gepumpt“ wurde. Daher braucht man Detektoren, die die einzelnen Elektronen beim Verlassen der Pumpe nachweisen können. Schaltet man drei Pumpen und Nachweissysteme hintereinander, verzählt man sich bei zehn Millionen gepumpter Elektronen nur etwa zweimal und kann die Einheit Ampere auf sieben Stellen hinter dem Komma genau realisieren.

Genauso wie bei diesen vier Basiseinheiten (Sekunde, Meter, Kilogramm, Ampere) funktioniert es auch bei den drei übrigen: Durch Festlegung des Zahlenwertes einer zugehörigen Naturkonstante lassen sich Einheiten definieren. Beim Mol ist es die Avogadro-Kontante, beim Kelvin die Boltzmann-Konstante und bei der Candela das photometrische Strahlungsäquivalent.

Eigenschaften des SI

Dem revidierten SI liegt ein fundamental neues Konzept zugrunde, das viele Vorteile bietet: Es garantiert Langzeitstabilität. Die Konstanten behalten für alle Zeit den festgelegten Wert und ändern sich nicht mehr. Ein weiteres Charakteristikum ist, dass man im Allgemeinen mehr als eine Konstante braucht, um eine Einheit zu realisieren: Beim Ampere sind es zwei (Elektronenladung und Frequenz des Hyperfeinstrukturüberganges), beim Kilogramm drei (Frequenz des Hyperfeinstrukturüberganges, Lichtgeschwindigkeit und Planck-Konstante). Meist stehen zur Realisierung der Einheiten unterschiedliche Messmethoden zur Verfügung, die helfen, die Ergebnisse zu kontrollieren. Beim Kilogramm sind es das Abzählen von Atomen in der Siliziumkristallkugel und die Watt- oder Kibblewaage – anstelle des einen Urkilogrammstückes in Paris. Die Wissenschaftler sind vollständig unabhängig von bestimmten Orten: Jeder in einem Labor überall auf der Welt (oder auf einem anderen Stern) kann ein Experiment aufbauen, um eine Einheit zu realisieren. Eine Einheit ist nicht nur mehr an einem einzelnen Messpunkt definiert, also bei genau einem Kilogramm für die Masse oder durch die zwei Punkte für die Definition der Temperaturskala, den absoluten Nullpunkt und den Tripelpunkt des Wassers. Durch das Fertigen kleinerer Siliziumkugeln oder von Silizium-Nanokristallen sowie mithilfe von Wattwaagen lässt sich das Kilogramm über weite Bereiche der Masseskala direkt realisieren. Analog ist es insbesondere auch bei der Temperatur, wo man verschiedene Gleichungen heranziehen und so die ganze Temperaturskala abdecken kann. Schließlich sind die elektrischen Einheiten nicht mehr isoliert in ihrer eigenen Quantenwelt, sondern in das ganze System integriert. Alle Einheiten sind in der Quantenwelt angekommen, ihre Konsistenz gesichert.

Offen für Innovationen

Ein entscheidender Vorteil des revidierten Systems ist die Tatsache, dass sich durch verbesserte Experimente die Einheiten immer besser realisieren lassen, ohne die Definition zu ändern. Nicht mehr Artefakte, sondern nur noch die Quantennatur der Welt limitiert die Experimente. Während sich vorher eine Masse auf höchstens acht Stellen hinter dem Komma genau bestimmen ließ, weil der Kilogramm-Prototyp in dieser Größenordnung schwankte, ist diese Grenze mit den neuen Definitionen aufgehoben. Prinzipiell wäre sogar eine bis auf ein Atom genaue Massenbestimmung, also mit etwa 25 Stellen hinter dem Komma, möglich. Das heißt, die Definition lässt viel Freiheit für Verbesserungen.

Beim Ampere begrenzte die alte Definition (Kraft zwischen zwei Leiterbahnen) die Realisierung auf sieben Stellen hinter dem Komma. Jetzt ist die prinzipielle Grenze im neuen System durch die sehr kleine Ladung des Elektrons gegeben, bei etwa 10^{-19} Coulomb. So lässt sich im Prinzip, wenn Elektronen pro Sekunde abgezählt werden und man sich nur um ein Elektron verzählt, ein Ampere auf 19 Stellen hinter dem Komma genau messen.

Die Limitierung liegt nicht mehr in den Definitionen, sondern in den Gesetzen der Quantenphysik, in der Tatsache, dass offensichtlich die Messgrößen nicht unendlich klein werden können, sondern in der Natur in kleinsten Paketen, den Quanten vorkommen. Das ist eine bahnbrechende Verbesserung und in diesem Sinne eine enorme Veränderung.

Allerdings sollte einen Tag nach Inkrafttreten des revidierten SI ein Kilogramm, etwa bei Einkäufen im Supermarkt, noch genauso schwer sein wie am Tag davor. Es durfte im Alltagsleben keine Änderungen geben. Daher sind die jetzt festgelegten Werte der Konstanten so „krumm“. Für „gerade“ Werte der Konstanten hätten zwangsläufig die Länge eines Meters, die Masse eines Kilogrammes usw. verändert werden müssen. Zudem musste sichergestellt werden, dass sich diese Konstanten mit einer Genauigkeit messen lassen, die für die Hightech-Messindustrie gebraucht wird, um die Einheiten zu realisieren.

Es war die Arbeit vieler Metrologieinstitute und hunderter Metrologen auf der ganzen Welt. Was zum Teil einfach aussieht, beruht auf Hochpräzisionsexperimenten und jahrelangen Entwicklungen. Erst 2017, genau 117 Jahre nach der wegweisenden Idee von Max Planck, waren die Voraussetzungen erfüllt, um das letzte Artefakt loszuwerden. Im Herbst 2017 sprach ein Expertengremium, das Internationale Komitee für Maße und Gewichte (CIPM), die Empfehlung aus, das Einheitensystem grundlegend zu revidieren. Ein Jahr später, am 16. November 2018, beschloss die Generalkonferenz für Maße und Gewichte (CGPM) die Revision. Am 20. Mai 2019 trat mit dem revidierten SI ein neues System der Einheiten in Kraft, das Langzeitstabilität, universelle Realisierbar-

keit, Genauigkeit, die mit dem technologischen Fortschritt mitwächst, sowie Innovation in Wirtschaft und Technik garantiert.
An einer Stelle wurde nicht das gemacht, was Max Planck gefordert hatte, nämlich durchgehend Fundamentalkonstanten zu verwenden: Die Definition der Sekunde beruht weiterhin auf der Eigenschaft eines bestimmten Atoms. Dieses, nämlich Cäsium, wurde lediglich aus pragmatischen experimentellen Gründen gewählt. Diese Inkonsequenz hat gute Gründe: Die so definierte Sekunde lässt sich genauer als jede andere Einheit messen: auf 16 Stellen hinterm Komma. Und es gibt schon Uhren, die auf 18 Stellen hinterm Komma genau ticken. Allerdings beruhen sie auf Pendeln in anderen Atomen, die bis zu hunderttausendmal schneller schwingen als die im Cäsium-Atom. Im Jahr 2026 könnte mithilfe dieser sogenannten optischen Atomuhren die Sekunde neu definiert werden. Bereits jetzt werden sie in der Praxis verwendet. Hebt man eine solche Uhr im Gravitationsfeld der Erde um einen Zentimeter an, dann geht sie nach Einsteins Relativitätstheorie um 10^{-18} Sekunden anders. Damit ist eine optische Uhr auch ein Detektor für das Gravitationsfeld der Erde, mit dem sich unterirdische Lager oder Wasserströme detektieren lassen. Man kann mit ihnen im Detail sehen, wie das Eis in Grönland aufgrund der Erderwärmung schmilzt, oder beim Tunnelbau von zwei Seiten z. B. durch die Alpen sicherstellen, dass sich die beiden Bohrtrupps auf Zentimeter genau aufeinander zubewegen. Es gab also gute Gründe, diese Definition aufrechtzuerhalten. Sie ist praktikabel, hochpräzise und sehr gut – wenn auch nicht ganz im Sinne des Theoretikers.

Konstante Konstanten?

Sind diese Konstanten, die wir jetzt definiert und festgelegt haben, physikalisch gesehen wirklich konstant, oder ändern sie sich im Laufe der Zeit? Genau dies sagen einige physikalische Theorien voraus, die darauf zielen, eine vereinheitlichte Theorie der Physik zu entwickeln, also die elektroschwache, die starke sowie die gravitative Wechselwirkung zu vereinheitlichen.
Dazu gibt es Experimente, insbesondere zur sogenannten Sommerfeld'schen Feinstrukturkonstante, einer Kombination von Elektronenladung, Lichtgeschwindigkeit und Planck-Konstante. Die weltweit genauesten Experimente an der PTB mit den besten verfügbaren Uhren zeigen, dass die jährlichen Änderungen – wenn sie überhaupt vorhanden sind – höchstens an der 18. Stelle hinterm Komma liegen. Sie sind also auf jeden Fall so klein, dass sie die Präzision der Realisierung der Einheiten im revidierten System zumindest auf sehr lange Zeit nicht merkbar beeinflussen werden – nicht nur bei den Genauigkeiten, die im alltäglichen Leben gefragt sind, sondern auch für höchste Präzision, wie sie in Hightech-Bereichen der Industrie und in der Wissenschaft gefordert sind.
Die Menschheit hat einen jahrtausendelangen Weg zurückgelegt: von den ersten Artefakten bis hin zu einem System, das die wohl höchste erreichbare Abstraktionsstufe darstellt, um ein Einheitensystem zu generieren, das nach den Kriterien der Französischen Revolution universell ist – nicht nur auf der Erde, sondern auch im gesamten uns bekannten Universum, entsprechend der Vision von Max Planck. Für alle Zeiten, für alle Culturen!



Abb. 8: Bei der Generalkonferenz für Maße und Gewichte in Versailles beschlossen die Staaten der Meterkonvention am 16. November 2018 einstimmig das „neue" SI. (Quelle: BIPM)

Literatur

[1] Decisions of the 26th meeting of the General Conference of Weights and Measures, Versailles, 13-16 November 2018, www.bipm.org/en/measurement-units/rev-si/
[2] The International System of Units, 9th edition, Bureau international des poids et mesures www.bipm.org/en/publications/si-brochure/
[3] Das neue Internationale Einheitensystem (SI), PTB-Infoblatt, November 2017, www.ptb.de

Mess- und Sensortechnik in der digitalen Transformation

Ulrich Kaiser[1], Klaus-Dieter Sommer[2]
[1]Endress+Hauser AG, Reinach, Schweiz
[2]TU Ilmenau, Ilmenau, Deutschland

Zusammenfassung

Die digitale Transformation ist über einige Ihrer Handlungsfelder unmittelbar mit der Mess- und Sensortechnik verbunden. Dafür müssen Sensoren „smart“ sein und können so Teilnehmer eines IoT-Ökosystems werden. „Smarte“ Sensoren unterstützen Diagnose und vorausschauende Wartung und können Ihren eigenen Kalibrierungsbedarf optimieren. Die IoT-Ökosysteme sind eine gute Plattform für die Fusion vieler diverser Sensoren zu Soft-Sensoren mit neuen Messgrössen. Hierfür bietet die recht neue Methodik des maschinellen Lernens neue Möglichkeiten. Der Schutz von Informationen ist bei „smarten“ Sensoren in IoT-Ökosystemen sehr viel anspruchsvoller. Da braucht es immer Massnahmen der OT-Security.

Keywords: Ökosystem für Internet of Things, Soft-Sensoren, maschinelles Lernen, OT-Security

Die DigitaleTransformation

Eine digitale Transformation geht weit über die eigentliche Digitalisierung hinaus. Digitalisierung ist die Optimierung bestehender Prozesse mit digitalen Mitteln, wobei diese Prozesse technische, aber auch Geschäftsprozesse sein können. Darüber steht die digitale Transformation, welche zusätzlich die Veränderung bestehender und/oder den Aufbau neuer (Geschäfts)-Prozesse mit den Möglichkeiten digitaler Technologien zum Inhalt hat. Nach Peters /1/ findet digitale Transformation in sieben Handlungsfeldern statt /1/, von denen drei, nämlich „Cloud and Data“, „New Technologies“ und „Process Engineering“ unmittelbar mit der Mess- und Sensortechnik verbunden sind.
Dabei sind im Besonderen die digitalen Technologien, Internetzugang, hoch leistungsfähige Embedded Systeme, Funkkommunikation und künstliche Intelligenz für die Mess- und Sensortechnik von Relevanz und erlauben dieser einen wesentlichen Beitrag zur digitalen Transformation zu leisten.

Was braucht es nun, damit ein Sensor oder ein Messsystem einen solchen Beitrag leisten kann? Diese Frage wurde von der NAMUR-Organisation in ihrer Technologie Roadmap „Prozesssensoren 4.0“ /2/ beantwortet. Sensoren müssen „smart“ sein, um so Teil des Internet of Things (IoT) werden zu können. Das bedeutet Sensoren müssen über Konnektivität und Kommunikationsfähigkeit verfügen. Sie müssen Instandhaltung und Betriebsfunktionen integrieren. Dazu gehört die Selbstdiagnose und – angestrebt – die eigene Wartung wie die Selbstkalibration. Sie müssen Fähigkeiten für Rückführbarkeit zum SI-System (Traceability) und Compliance aufweisen. Sie besitzen eine virtuelle Beschreibung von sich selbst – auch digitaler Zwilling genannt - und sind gegebenenfalls in der Lage, untereinander zu interagieren (Abbildung 1).



Abb. 1: Anforderungen an „smarte“ Sensoren (aus /2/)

Das IoT-Ökosystem für „smarte“ Sensoren

Um „smarte“ Sensoren wirksam werden zu lassen, braucht es natürlich eine angepasste Architektur des Steuerungssystems und eine Plattform für die Ausführung der Methoden: das IoT-Ökosystem. Dabei ist eine der Architekturrealisierungen die Ergänzung des klassischen geschlossenen Steuerungssystems in der Struktur der Automatisierungspyramide um ein offenes IoT-Ökosystem (Abbildung 2). In dieser Variante wird nur ein Teil der Sensordaten und Informationen im Ökosystem verarbeitet und zum Beispiel kritische Daten bleiben nach wie vor im klassischen Steuerungssystem. Eine übliche Anwendung ist die Erweiterung eines bestehenden Messsystems um Verwaltungsfunktionen zur Diagnose und Wartung der Sensoren über das Ökosystem, während die eigentliche Verarbeitung von Messdaten noch

dem vorhandenen Steuerungssystem verbleibt. Eine Realisierung einer solchen Architektur ist zum Beispiel die Namur Open Architecture (NOA /3/). Für neu zu erstellende Messsysteme übernimmt das IoT-Ökosystem auch die Aufgaben der Steuerung (Abbildung 3).

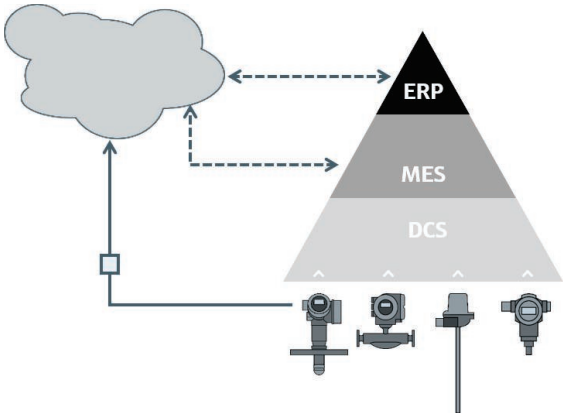


Abb. 2: Erweiterung der klassischen Automatisierungspyramide um ein IoT-Ökosystem

In einem IoT-Ökosystem werden alle Sensordaten und -Informationen in Echtzeit oder zeitnah in einer zentralen Datenbank abgebildet. Diese kann cloud-basiert oder lokal sein. Ein zentrales Rechnersystem, welches auch cloud-basiert sein kann, offeriert dann beliebige Dienste aus Daten und Informationen der Sensoren und anderen Quellen wie dem Internet. Auf diese Dienste kann dann über standardmässige Frontend-Technologien wie Mobile Apps zugegriffen werden.

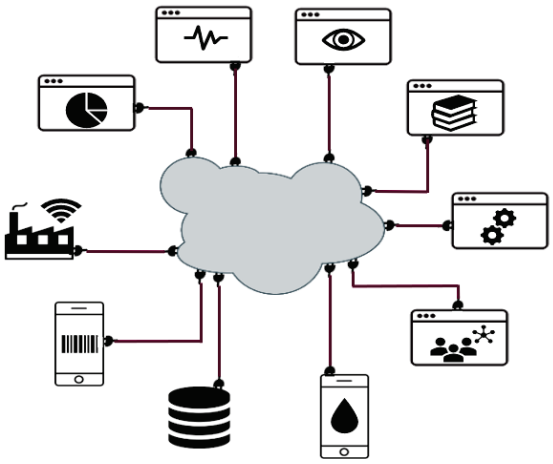


Abb. 3: Skizze eine IoT-Ökosystems

Diagnose und vorausschauende Wartung

Eine der am häufigsten diskutierten und schon oft realisierten Anwendungen auf IoT-Ökosystemen ist die Diagnose und die vorausschauende Wartung. Dabei sind Sensoren dafür erforderlich, um entweder Informationen zur Diagnose von anderen technischen Systemen oder Diagnose-Informationen für sich selbst liefern. Die Übermittlung von Diagnoseinformationen ist an sich nichts Neues; der besondere Nutzen besteht aber darin, dass in einem IoT-Ökosystem das Kollektiv aller Diagnose-informationen betrachtet werden kann, woraus dann übergreifende Informationen wie der Gesundheitszustand einer technischen Anlage und Trends ermittelt werden können.

Solche Trends können auch eine voraus-schauende Wartung ermöglichen, wenn sie behandlungsbedürftige Zustände vorhersagen können. Vorausschauende Wartung kann nach grundsätzlich zwei Prinzipien realisiert werden. Einmal über die direkte Messung eines Abnutzungsvorrats oder anderen Indikators für das Wartungsereignis. Oder aus einer längeren Beobachtungsphase heraus, wo dann reale Wartungsereignisse mit einer grossen Anzahl von Messdaten und anderen Informationen korreliert werden und so eventuell Indikatoren mittels statistischer Auswertung oder Methoden des maschinellen Lernens ermittelt werden können. Für Letzteres bietet unser IoT-Ökosystem eine geeignete Plattform. Als ein Beispiel für vorbeugende Wartung sei die frühzeitige Erkennung von Schaumbildung in Tanks genannt, die direkt über interne Messwerte von Radarfüllstandssensoren im Tank ermittelt werden kann.

Kalibrierung von Sensoren

Die Technik von Sensoren unterliegt zeitlichen Veränderungen, abhängig von der Belastung durch den Betrieb. Zur Sicherstellung der metrologischen Qualität ist daher eine regelmässige Kalibrierung und gegebenenfalls Justage der Sensoren notwendig. Die zeitliche Rate der notwendigen Kalibrierungen wird bestimmt durch das Driftverhalten des Sensors und die Kritikalität seines Einsatzes. Kalibrierungen dominieren häufig die Betriebskosten eines Sensors und deshalb ist eine Reduzierung der Rate von Kalibrierungen immer anzustreben. Ein erfolgreicher Ansatz dafür ist das regelmässige, explizite Über-prüfen von Komponenten, welche die zeitlichen Veränderungen bewirken. Diese Überprüfung passiert dabei gegen interne Referenznormalen wie zum Beispiel

(Quanten)-Normalen für die elektrische Spannung. Diese Sensorverifikation wird entweder mittels digitaler Technik im Sensor selbst oder über das IoT-Ökosystem realisiert. Über Letztere können dann auch die für eine Dokumentation notwendigen Zertifikate automatisch publiziert werden.
Die angestrebte Ideallösung ist, Referenznormale – möglichst mit direkter Rückführbarkeit auf das (aktuelle) SI – in Sensoren oder dem Messsystem zu integrieren. Dann könnte der Sensor ohne jegliche externe Infrastruktur und Aufwand sich selber kalibrieren. Das National Institute of Standards and Technology der USA (NIST) hat diese große technologische Herausforderung mit ihrem NIST-on-a-Chip-Projekt in Angriff genommen /4/. In diesem Projekt sollen Referenznormale mikrotechnisch konstruiert werden, die in Sensoren integriert werden können /5/.

Soft-Sensoren

Soft-Sensoren sind Messsysteme, wo schwierig zu messende Messgrössen aus einer Mehrzahl von einfach zu messenden Messgrössen errechnet werden /5/. Schwierig zu messende Messgrössen sind diejenigen, für die kein direktmessendes Messsystem zur Verfügung steht, zum Beispiel für bestimmte stoffliche Eigenschaften, wo der Ort der Messung nicht zugänglich ist oder wo eine Vorhersage eines zukünftigen Messwertes gefordert ist. Die Gesamtheit der Eingangs-Messgrössen muss natürlich die Ergebnis-Messgrösse repräsentieren können. In /6/ ist das Verfahren zur Entwicklung eines Soft-Sensors beschrieben, der aus den Informationen von 85 „einfachen“ Sensoren die Konzentration einer Substanz in einem chemischen verfahrenstechnischen Prozess berechnen kann. Als technische Plattform für solche Soft-Sensoren ist unser IoT-Ökosystem wieder sehr gut geeignet. Für die Berechnung der gewünschten Messgrösse aus den Eingangsdaten gibt es zwei grundsätzliche Verfahren, das modellgetriebene und das datengetriebene. Beim modellgetriebenen Verfahren muss der Zusammenhang zwischen den Eingangsgrössen und der Ausgangsgrösse - der gewünschte Messwert - bekannt und analytisch darstellbar sein. Bei den datengetriebenen Verfahren wird die Beziehung zwischen Eingangsgrösse und der gewünschten Messgrösse anhand von Trainingsdaten erlernt. Die Trainingsdaten der gewünschten Messgrösse müssen dann natürlich auf eine andere Art und Weise beschafft werden. In dem erwähnten Beispiel durch separate Labormessungen.

Für die datengetriebenen Verfahren gibt es eine grosse Anzahl von mathematischen Verfahren, von der einfachen Regressionsanalyse bis zum Deep Learning, Letzteres ist eine Methode des maschinellen Lernens, dieses wiederum eine Disziplin der künstlichen Intelligenz /7/. Moderne Cloud-Systeme von Amazon, Google, Microsoft, …, auf denen die IoT-Ökosysteme häufig aufgebaut sind, bieten eine Vielzahl von Tools für solche Berechnungen an.

Informationssicherheit für die Mess- und Sensortechnik

Werden Sensoren und Messsysteme Objekte des Internet of Things, so werden Sie auch verwundbar wie die anderen Objekte des Internets. Manipulationen von Sensordaten und Unterbrüche von Kommunikations-Verbindungen können gefährliche Auswirkungen haben. Ebenso kann das unautorisierte „Mithören“ von Sensordaten bei bestimmten Applikationen unerwünscht sein. Die OT-(Operational Technology)-Security befasst sich mit der Informationssicherheit von technischen Systemen /8/. Dabei finden die gleichen Methoden und Massnahmen der Informationssicherheit Anwendung wie Verschlüsselung, Authentifizierung, etc., um den für jede Kritikalität der Anwendungen benötigten Schutz zu gewährleisten.

Literaturnachweis

[1] Marc K. Peter (Hrsg.) KMU-Transformation: Als KMU die digitale Transformation erfolgreich umsetzen; *FHNW Hochschule für Wirtschaft*, Olten, Schweiz, (2017); https://kmu-transformation.ch/

[2] Technologie-Roadmap "Prozess-Sensoren 4.0", VDE/VDI, NAMUR, 2015

[3] NAMUR Open Architecture; *Interessengemein-schaft Automatisierungstechnik der Prozess-industrie e.V.;* https://www.namur.net/fokusthemen/namur-open-architecture.html

[4] NIST on a Chip; https://www.nist.gov/pml/productsservices/nist-chip-portal

[5] Th. Simmons, G. Gerlach, K.-D. Sommer: Linking Innovators in Sensors and Measurement Technology. NIST - Visitors' Seminar, Gaithersburg, 2019-04-10

[6] Marcin Budka et al. (2014) From Sensor Readings to Predictions: On the Process of Developing Practical Soft Sensors. In: Blockeel H., van Leeuwen M., Vinciotti V. (Hrsg) *Advances in Intelligent Data Analysis XIII. IDA 2014. Lecture Notes in Computer Science*, Band 8819. Springer,

[7] Wolfgang Ertel, Grundkurs Künstliche Intelligenz; Eine praxisorientierte Einführung; *Springer Vieweg*; 4. Auflage, (2013)

[8] IEC 62443-3 Security for industrial process measurement and control – Network and system security

Plenarvortrag: Sensoren und Messsysteme 2019 Tagung

Nürnberg, 26. Juni 2019

Schnellste Vorgänge im Mikrokosmos:

Von der Grundlagenforschung bis hin zur Krebsbekämpfung

Prof. Ferenc Krausz

Max-Planck-Institut für Quantenoptik, Garching, Deutschland
Ludwig-Maximilians-Universität München, Deutschland
attoworld.de, ferenckrausz.de

In unserem Leben spielen Elektronen eine elementare Rolle. Ihre Bewegungen steuern alle chemischen Prozesse in der Natur. Ebenso bestimmen Elektronen die Geschwindigkeit unserer Computer- und Informationstechnologien. Bewegungen der Elementarteilchen und Lichtwellen bedingen sich gegenseitig, auf der Zeitskala von Attosekunden, also Milliardstel von Milliardstel Sekunden.

Um zu erkunden, was in so kurzen Zeiträumen im Mikrokosmos passiert, benötigt man Lichtblitze, die ebenfalls nur Attosekunden lang dauern. Solche Blitze wurden in der Arbeitsgruppe von Professor Krausz in Wien im Jahr 2001 erstmals produziert. Diese Technologie ermöglicht es nun erstmals, in Echtzeit zu verfolgen, wie sich ultraschnelle Phänomene um die Kerne von Atomen herum abspielen.

In den letzten Jahren sind die Attosekunden-Technologien den Kinderschuhen entwachsen und zu verlässlichen Messtechniken gereift. Jetzt eröffnen sie neue Wege um etwa Krankheiten, wie Krebs, frühzeitig zu detektieren. Die Grundlage dafür bieten Messungen kleinster Veränderungen in der molekularen Zusammensetzung von Blut mittels direkten Abtastens ultraschnell schwingender elektrischer Felder, die durch angeregte Moleküle ausgesendet werden.

Sitzungen

DOI 10.5162/sensoren2019/1.1.1

Ein Framework zur Simulation optischer Sensoren

Steffen Sauer[1], Thomas Dunker[1], Michael Heizmann[2]
[1] Fraunhofer Institut für Fabrikbetrieb und –automatisierung IFF, Sandtorstraße 22, 39106 Magdeburg
[2] Karlsruher Institut für Technologie (KIT) Hertzstraße 16, 76187 Karlsruhe

Zusammenfassung

In dieser Arbeit wird ein neuartiges Framework vorgestellt, mit dem sich die Erzeugung von optischen Messdaten simulieren lässt. Im Vergleich zu bisherigen Arbeiten liegt hier der Fokus auf einer schnellen, realitätsnahen Approximation des Sensorverhaltens. Das Framework besteht aus Verfahren zur merkmalsbasierten Simulation von Kamerabildern, sowie zur 3D-Messdatensimulation von Laser-Lichtschnitt- und Phasenshift-Sensoren. Die Simulationen laufen vollständig auf der GPU mit enorm hohen Geschwindigkeiten. In diesem Beitrag wird die Funktionsweise der Verfahren erläutert und es werden Zeitabschätzungen für die Verfahren gegeben. Anhand zweier Beispiele wird schließlich die Anwendung der Simulationen zur Ansichtenplanung bei optischen Prüfaufgaben vorgestellt.

Keywords: Simulation, Ansichtenplanung, GPU, Phasenshift, Laserlichtschnitt

Motivation und Problemstellung

Mit Einsetzen der vierten industriellen Revolution befindet sich die Produktion von Waren in einem starken Wandel. Eine derzeit absehbare Tendenz ist, dass sich die Variantenvielfalt der Produkte kontinuierlich steigern wird, wobei Losgrößen bei der Herstellung durch eine zunehmende Individualisierung stark sinken [1]. Dies hat direkten Einfluss auf die Planung und den Betrieb optischer Inspektionssysteme. Einerseits sinken damit die zur Verfügung stehenden Ressourcen, um Inspektionen vorzubereiten und zu parametrisieren, andererseits stellt der Variantenreichtum eine enorme Hürde dar, um optische Prüfungen wirtschaftlich betreiben zu können.
Wir schlagen daher vor, in diesen Fällen modellbasierte Prüfungen durchzuführen. Die Grundidee besteht darin, anhand von CAD-Modelldaten der zu prüfenden Objekte und mathematischen Modellen der eingesetzten Sensoren Inspektionen bereits offline zu planen und während des Betriebs synthetische Messdaten als Vergleichsdaten zu erzeugen. Dazu spielt eine geeignete Simulation von Messdaten eine entscheidende Rolle, weshalb wir hier ein Framework vorstellen, mit dem dies für optische Sensoren realitätsnah durchgeführt werden kann.

Stand von Literatur und Technik

Die Simulation von optischen Sensoren ist ein sehr ausgiebig erforschtes Gebiet der Computergrafik. Für Kameras existieren eine Vielzahl von Ansätzen, mit denen künstliche Szenen generiert werden. Hierbei muss abgewogen werden, ob eine Szene möglichst realitätsnah oder möglichst schnell dargestellt werden soll. Für realitätsnahe Darstellungen sind physikalisch korrekte Simulationsumgebungen erforderlich wie beispielsweise in [2] beschrieben. Andere, zum Teil GPU-gestütze Verfahren arbeiten deutlich schneller, erzeugen naturgemäß nur vereinfachte Darstellungen [3].
Im Bereich der 2,5-dimensional messenden Sensoren existieren für das Laserlichtschnitt-Verfahren bereits eine Vielzahl an Lösungsansätzen: Sie teilen sich im Wesentlichen in Verfahren auf, die die erwarten Oberflächenpunkte explizit berechnen [4] oder durch Raytracing implizit simulieren [5]. Für Sensoren, die flächig 3D-Daten durch strukturierte Beleuchtung rekonstruieren, sind Simulationsverfahren bekannt, die im ersten Schritt virtuelle Kamerabilder erzeugen und die eigentliche 3D-Rekonstruktion dann auf diesen ausführen [6].

Modellierung von schiefen Sichtvolumina

Um das Verhalten optischer Systeme simulieren zu können, muss deren Abbildungsverhalten hinreichend genau approximiert werden. Das weit verbreitete Lochkameramodell [7] hat sich hier als praktisch erwiesen, deckt jedoch nicht alle Kamera-Objektiv-Anordnungen ab. Beispielsweise lässt sich damit die Scheimpfluganordnung, wie sie in großem

Umfang in Laser-Lichtschnittsensoren eingesetzt wird, nicht abbilden [8].
In der Computergrafik hingegen hat sich das Einheitssichtvolumen (Normalized Device Coordinates) mit den Eckpunkten mit den Eckkoordinaten $[-1,-1,-1]$ bis $[1,1,1]$ durchgesetzt [9]. Wir suchen daher nach einem universellen Lösungsweg, ein Sichtvolumen, das aus dem Datenblatt des Sensorherstellers entnommen oder durch Kamerakalibrierung berechnet wird, durch Projektion auf das Einheitssichtvolumen zu transformieren.
Gegeben seien zwei Punktemengen als Matrizen mit den homogenen Punktekoordinaten in den Spalten

$$X = \begin{bmatrix} x_1^1 & \cdots & x_1^k \\ \vdots & \ddots & \vdots \\ x_n^1 & \cdots & x_n^k \\ 1 & \cdots & 1 \end{bmatrix} \text{ und } Y = \begin{bmatrix} y_1^1 & \cdots & y_1^k \\ \vdots & \ddots & \vdots \\ y_m^1 & \cdots & y_m^k \\ 1 & \cdots & 1 \end{bmatrix},$$

und gesucht ist die $(m+1) \times (n+1)$ Matrix der homogene Abbildung A, die X auf Y bis auf Skalierung abbildet

$$AX = \begin{bmatrix} a_1^1 & \cdots & a_1^{n+1} \\ \vdots & \ddots & \vdots \\ a_{m+1}^1 & \cdots & a_{m+1}^{n+1} \end{bmatrix} X = Y \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_k \end{bmatrix}$$

In dieser Gleichung lassen sich beide Seiten transponieren und spaltenweise aufschreiben

$$X^T \begin{bmatrix} a_i^1 \\ \vdots \\ a_i^{n+1} \end{bmatrix} - \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_k \end{bmatrix} \begin{bmatrix} y_i^1 \\ \vdots \\ y_i^k \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} y_i^1 \\ \vdots \\ y_i^k \end{bmatrix}$$

mit $i = 1, \cdots, m+1$. Vertauschen der Rolle von $s_2, \cdots, s_k$ und $y_i^1, \cdots, y_i^k$ ergibt

$$X^T \begin{bmatrix} a_i^1 \\ \vdots \\ a_i^{n+1} \end{bmatrix} - \begin{bmatrix} 0 & \cdots & 0 \\ y_i^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & y_i^k \end{bmatrix} \begin{bmatrix} s_2 \\ \vdots \\ s_k \end{bmatrix} = \begin{bmatrix} y_i^1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Daraus kann ein Gleichungssystem für die Unbekannten $s_2, \cdots, s_k$ und $a_i^1, \cdots, a_i^k$ zusammengestellt werden

$$\begin{bmatrix} 0 & \cdots & 0 & & & \\ y_1^2 & \cdots & 0 & & & \\ \vdots & \ddots & \vdots & X^T & \cdots & 0 \\ 0 & \cdots & y_1^k & & & \\ & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & & & \\ 1 & \cdots & 0 & 0 & \cdots & X^T \\ \vdots & \ddots & \vdots & & & \\ 0 & \cdots & 1 & & & \end{bmatrix} \begin{bmatrix} s_2 \\ \vdots \\ s_k \\ a_1^1 \\ \vdots \\ a_1^{n+1} \\ \vdots \\ a_{m+1}^1 \\ \vdots \\ a_{m+1}^{n+1} \end{bmatrix} = \begin{bmatrix} y_1^1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

welches mit einem beliebigen Verfahren gelöst werden kann. In einigen unserer Beispiele war es sinnvoll die erste gefundene Lösung iterativ zu verbessern, siehe z.B. Kapitel 2.5 in [10].
Für die Modellierung der Sichtpyramide können in einem allgemeinen Fall Eckpunkte des Bildfeldes im Raum und das Projektionszentrum gegeben sein. Um diese mit Hilfe von Koordinaten zu beschreiben, wird im Projektionszentrum ein Koordinatensystem festgelegt, z.B. wie in Abb. 1 mit Y nach oben und Z von der Sichtpyramide weg.

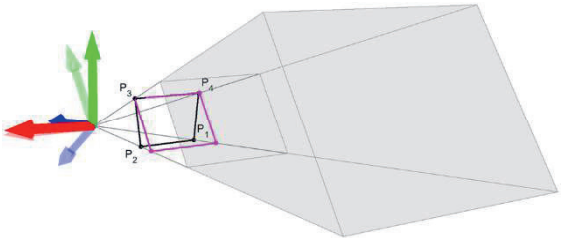


Abb. 1: Konstruktion der Sichtpyramide aus vier gegebenen Eckpunkten (schwarz) über ein Hilfskoordinatensystem (transparent) mit rechteckigem Sichtfeld (magenta).

Nun können die Eckpunkte, die z.B. auf einer Projektionsebene gegeben sind, auf die Ebene $z = -1$ projiziert werden. In der Skizze sind diese Punkte mit $P_1, \cdots, P_4$ bezeichnet. Sie sind folgendermaßen sortiert links unten, rechts unten, rechts oben und links oben. Im allgemeine Fall bilden sie kein Rechteck. Es muss folglich noch die Abbildung bestimmt werden, die diese Punkte auf ein Rechteck abbildet. Dazu verwenden wir obige Methode und ermitteln die 3×3 Matrix A für

$$X = \begin{bmatrix} p_1^1 & \cdots & p_1^4 \\ p_2^1 & \cdots & p_2^4 \\ -1 & \cdots & -1 \end{bmatrix} \text{ und } Y = \begin{bmatrix} -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}.$$

Diese Matrix beschreibt eine Drehung und eine Projektion $A = CD$, wobei die Projektion eine obere Dreiecksmatrix ist. Um diese Zerlegung mit verfügbaren Methoden zu gewinnen, verwenden wir die Householder QR-Zerlegung. Da die Reihenfolge Drehmatrix und obere Dreiecksmatrix nicht die gewünschte ist nutzen wir eine Hilfsmatrix

$$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

und zerlegen $M^T E^T = QR$. Daraus folgt $M = (E^T R^T E)(E^T Q^T)$ und somit $C = E^T R^T E$ und $D =$

$E^T Q^T$. Die Projektionsmatrix wird mit $C = C/c_3^3$ normalisiert. Eventuelle negative Elemente auf der Diagonale werden durch folgende Prozedur korrigiert. Falls $c_1^1 < 0$ gilt, werden die Matrizen folgendermaßen korrigiert $C = C\,\mathrm{diag}([-1,-1,1])$ und $D = \mathrm{diag}([-1,-1,1])\,D$. Falls $c_2^2 < 0$ gilt, werden $C = C\,\mathrm{diag}([1,-1,-1])$ und $D = \mathrm{diag}([1,-1,-1])\,D$ gesetzt. Damit hat C die typische Form einer Kameramatrix.
Nun können wir die Eckpunkte in das Koordinatensystem drehen, in welchem die Kameramatrix angewendet wird. In Abb. 1 ist dieses Koordinatensystem transparent dargestellt. In diesem Koordinatensystem können die Eckpunkte $X' = DX$ ebenfalls auf die Ebene $z = -1$ projiziert werden, die wir mit X'' bezeichnen und mit $X''' = D^T X''$ in das ursprüngliche Koordinatensystem zurück drehen. In der Skizze sind diese Punkte in magenta dargestellt und es ist erkennbar, dass sie ein Rechteck bilden. X''' kann nun mit den gewünschten Entfernungen skaliert werden, um die nahe und ferne Begrenzung der Sichtpyramide zu konstruieren.
Die Abbildung auf das Einheitssichtvolumen wird nun nach obiger Methode mit den Matrizen

$$\begin{bmatrix} s_{\mathrm{nah}} X''' & & s_{\mathrm{fern}} X''' \\ 1 \quad \cdots \quad 1 & & 1 \quad \cdots \quad 1 \end{bmatrix}$$

und

$$\begin{bmatrix} -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

gefunden.

Framework zur Simulation von Messdaten

Um die Frage zu beantworten, wo ein optischer Sensor platziert werden muss, um ein Merkmal ausreichend gut erfassen zu können, kann ein „Generate and Test"-Ansatz [11] verfolgt werden. Dabei wird der unendlich große Lösungsraum der Sensorposition disketisiert und für jede Position eine Messdatensimulation durchgeführt. Deren Ergebnis wird dann anhand einer Qualitätsfunktion bewertet. Dabei gilt im Allgemeinen, dass mehr potenzielle Sensorpositionen bewertet werden können, je schneller die Simulation durchgeführt werden kann.
Im Bereich der 2D-Bildverarbeitung werden in vielen Anwendungen die Kanten eines Objekts zur Lokalisierung und Bewertung herangezogen. Das hier vorgestellte Framework ist in der Lage, für Kameras, deren Abbildungseigenschaften nach dem obigen Verfahren definiert werden können, mögliche Objektkanten auf der GPU zu erzeugen. Dazu wird das in [12] vorgestellte Verfahren aufgegriffen, bei dem durch ein 2-Pass-Rendering synthetische Objektkanten erzeugt werden. Diesem Verfahren liegen folgende Annahmen zugrunde. Erstens: Flächen, die in unterschiedlichen Entfernungen zur Kamera liegen, erzeugen an dem Tiefensprung eine Kante. Zweitens: Flächen, die Unstetigkeiten in der Oberflächennormale aufweisen, erzeugen an diesen sich abbildende Kanten. Und drittens: Unterschiedliche Objekte, die aneinandergrenzen, erzeugen möglicherweise Kanten.

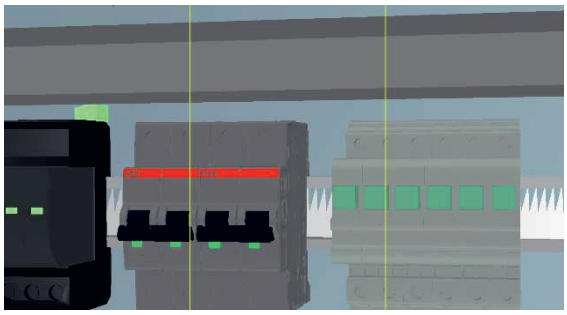


Abb. 2: CAD-Modell einer Baugruppe.



Abb. 3: Zwischenschritte zur Feature-Extraktion des Modells aus Abb. 2: links Tiefenpuffer, mittig Normalenpuffer, rechts Objekt-ID-Puffer.

Zu Umsetzung wird ein zweistufiges Render-Verfahren genutzt. Mit den erforderlichen Kameraparametern wird eine Szene im ersten Pass so gerendert, dass die Entfernung zur Kamera, die Oberflächennormalen und die Objekt-ID in drei Puffer geschrieben werden (Multi-Target-Rendering, s. Abb. 3). In einem zweiten Render-Pass wird jeder Puffer pixelweise mit einem Differenzoperator gefaltet und eine Schwellwertbildung durchgeführt (s. Abb. 4). Die Vereinigung dieser drei Kantenmengen ergibt dann das simulierte Ergebnisbild (Abb. 5).
Im Bereich der 3-dimensionalen Objekterfassung spielt die aktive Triangulation insbesondere mit dem Laser-Lichtschnitt- und der Streifenprojektion eine wichtige Rolle. Beim Laser-Lichtschnitt wird das Licht einer Laserdiode über eine geeignete Optik zu einer Laserebene aufgeweitet und auf das Zielobjekt

projiziert. Das dabei diffus reflektierte Licht wird in einer Kamera als Laserlinie abgebildet. Ist die Lage der Laserebene zur Kamera bekannt, lässt sich entlang der abgebildeten Linie über Triangulation die Oberfläche des Objekts diskret rekonstruieren.

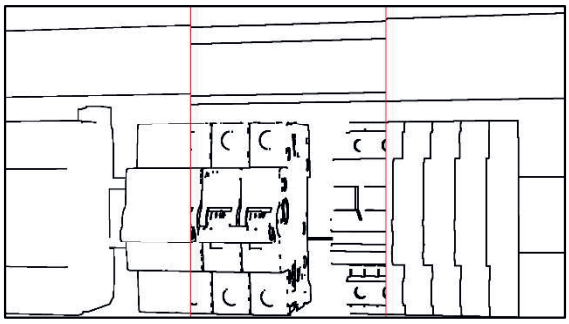


Abb. 4: Extrahierte Kanten aus den Zwischenpuffern.

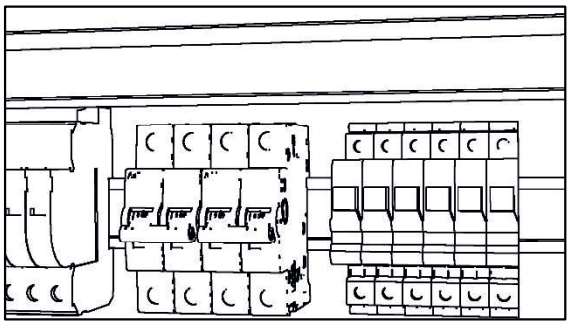


Abb. 5: Vereinigung aller Kanten aus den Zwischenpuffern.

Für eine realitätsnahe Simulation setzt das hier vorgestellte Framework das sogenannte Shadow-Mapping ein [13]. Mit diesem Verfahren werden in der Computergrafik Schatten durch Punkt und Richtungslichtquellen erzeugt. Zur Simulation des Laserlichtschnitts wird in einem Render-Verfahren mit zwei Pässen entgegen dem eigentlichen Strahlenverlauf die Szene zunächst aus Sicht der Kamera beleuchtet und anschließend aus Sicht des Lasers beobachtet. Im ersten Pass wird die virtuelle Kamera mit den Abbildungseigenschaften der realen Kamera parametrisiert und die Szene aus dieser Perspektive gerendert. Dabei werden lediglich die Tiefenwerte der Szene gespeichert. Für den zweiten Pass führen wir die sogenannte Laser-Kamera ein. Hierbei handelt es sich um eine als Lochkamera modellierte Zeilenkamera, deren horizontaler Öffnungswinkel dem der Laserebene entspricht. Die Laserkamera besitzt eine Auflösung von $1 \, x \, N$, wobei N die typischerweise angegebene Punktauflösung des Sensors ist. Aus Sicht des Lasers wird im zweiten Pass die Szene wiederholt gerendert und das Shadow-Mapping anhand der Tiefenwerte aus dem ersten Pass durchgeführt. Dadurch werden abgeschattete Bereiche gefiltert und es bleiben die Oberflächenpunkte, die sowohl von der Laserebene geschnitten als auch von der Kamera aus gesehen werden übrig. Die 3D-Koordinaten werden zusammen mit den Oberflächennormalen als Ergebnis ausgegeben (s. Abb. 6). Anhand der Oberflächennormale können anschließend Bewertungsfunktionen beispielsweise die Qualität der zu erwartenden Messpunkte vorherbestimmen.

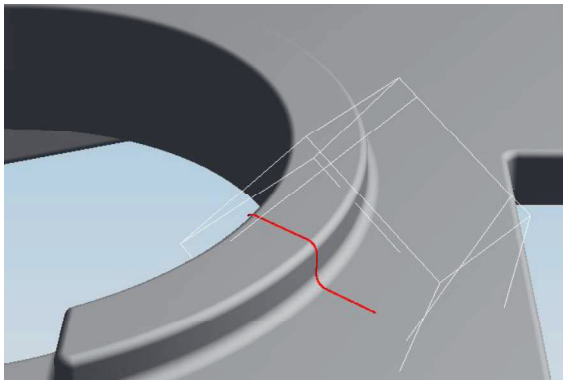


Abb. 6: Simulierter Laserlichtschnitt mit einer Auflösung von 1024 Punkten.

Soll ein Objekt flächig erfasst werden, bietet sich der Einsatz von Sensoren an, die nach dem Prinzip der strukturierten Beleuchtung arbeiten (Streifenlicht- bzw. Phasenshift). Hierbei wird statt einer Linie eine Serie von Bildern auf die Objektoberfläche projiziert. Die Bilderserie ermöglicht eine eindeutige Codierung mehrerer Lichtebenen und eine entsprechend großflächige 3D-Rekonstruktion.
In der Praxis werden dabei zwei unterschiedliche Ansätze verfolgt. Im ersten Fall wird das projizierte Muster genutzt, um der Objektoberfläche zusätzliche und robuste Merkmale hinzuzufügen. Es erfolgt dann eine 3D-Rekonstruktion über ein Kamerapaar durch eine Stereo-Korrespondenzsuche. Im zweiten Fall wird der Strahlengang des Projektors kalibriert. Dadurch kann eine Triangulation direkt zwischen Projektor und Kamera erfolgen. Hier finden sich in der Praxis Sensoren, die zusätzlich eine zweite Kamera – zum Teil mit unterschiedlichem Basisabstand – besitzen, um die Messdatenaufnahme genauer und robuster zu gestalten.
Die Simulation von Phasenshift-Sensoren ist in dem hier vorgestellten Framework ebenfalls auf der GPU umgesetzt worden. Die Simulation erfordert vier Render-Pässe und erzeugt analog zur Laser-Lichtschnitt-Simulation eine 3D-Punktewolke der zu erwartenden Messdaten. Für die Simulation werden das Sichtvolumen der linken und rechten Kamera sowie des

Projektors benötigt (s. Abb. 7). In den ersten beiden Render-Pässen wird die Szene aus der linken bzw. rechten Kamera gerendert und die Tiefenwerte gespeichert. Damit sind die Oberflächenpunkte bekannt, die theoretisch rekonstruiert werden können. Im dritten Pass wird die Szene aus Sicht des Projektors gerendert und die Tiefenwerte ebenfalls gespeichert. Damit sind die Oberflächenpunkte ermittelt, die vom Projektor bestrahlt werden.

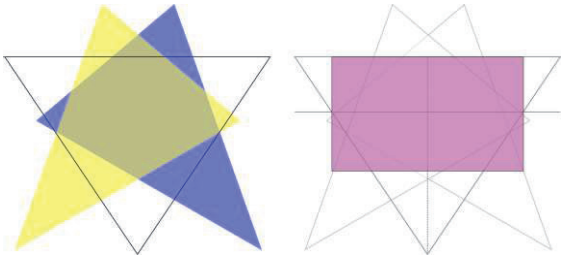


Abb. 7: Sichtvolumen während der Phasenshift-Simulation: links: Kameras und perspektivischer Projektor, rechts zusätzlich der orthogonale Projektor.

Im vierten Schritt findet die eigentliche 3D-Rekonstruktion statt. Hier wird die Szene orthogonal mit der physischen Auflösung des Phasenshift-Sensors aus Sicht des Projektors gerendert. Dies entspricht der 3D-Rekonstruktion der strukturierten Beleuchtung, die im Allgemeinen parallel zur optischen Achse des Projektionssystems durchgeführt wird. Während der Rekonstruktion wird das Shadow-Mapping durchgeführt, indem die Tiefenwerte der Kameraabbildungen aus den ersten beiden Render-Pässen mit den der perspektivischen Projektorabbildung aus dem dritten Pass verglichen wird. Für eine aktive Stereo-Rekonstruktion muss ein Punktkandidat in beiden Kameras sichtbar sein, für eine Projektor-Kamera-Rekonstruktion ist die Sichtbarkeit in einer Kamera bereits ausreichend

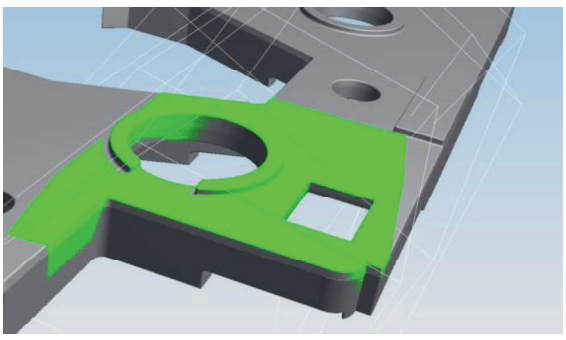


Abb. 8: Ergebnis der Phasenschift-Simulation auf einem Testobjekt.

Für nicht abgeschattete Punkte werden ebenfalls die 3D-Oberflächenkoordinaten und –normalen ausgegeben (s. Abb. 8).
Die Leisungsfähigkeit der Simulationen wurde mit synthetischen Tests ermittelt. Dazu wurden CAD-Modelle von Kugeln mit steigender Anzahl an Dreiecken simuliert und die benötigte Simulationszeit gemessen. Die Tests wurden auf einem Intel Core I7-8700K mit einer Consumer-Grafikkarte Nvidia GeForce GTX 1070 Ti durchgeführt. Die gemessenen Zeiten sind in Abb. 9 dargestellt.

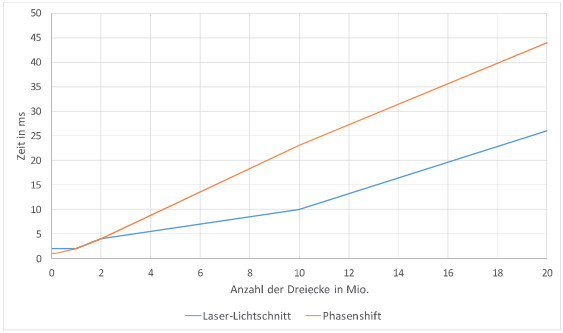


Abb. 9: Benötigte Zeit der 3D-Simulationen in Abhängigkeit von der Anzahl der Dreiecke in einer Szene.

Es zeigt sich, dass für CAD-Modelle mit moderater Komplexität von bis zu 10 Mio. Dreiecken selbst auf Mittelklassegrafikkarten die 3D-Simulationen noch mit Geschwindigkeiten berechnet werden, die eine interaktive Darstellung zulassen.
Das Framework steht unter Windows als C-Bibliothek bereit und kann dank der universellen Definition von Sichtvolumen für eine Vielzahl optischer Sensoren eingesetzt werden.

Anwendung der 2D-Simulation

Eine Inspektionsanlage soll automatisiert die korrekte Montage von Schaltschranktafeln überprüfen. Die Tafeln werden mit einer sehr hohen Variantenvielfalt hergestellt und unterscheiden sich daher zum Teil sehr in ihrer Bestückung. Zur Inspektion soll ein kamerabasiertes Prüfsystem mit einem Roboter über der Schaltschranktafel positioniert und dann eine bildbasierte Auswertung durchgeführt werden. Da die Schaltschranktafeln nur grob zum Roboter positioniert werden können, ist vor der eigentlichen Prüfung eine Referenzierung der Tafel zum Roboter erforderlich. Deren Ergebnis ist die Transformation von Schaltschranktafel- zum Roboterkoordinatensystem. Die Referenzierung soll rein kamerabasiert anhand von natürlichen Merkmalen durchgeführt werden.
Das Ziel besteht nun darin, mit Hilfe einer Messdatensimulation automatisiert geeignete

Sensorpositionen zu bestimmen, aus deren Perspektive der optische Sensor die Merkmale verdeckungsfrei erfassen kann.
Als natürliche Merkmale werden Bohrungen in der Schaltschranktafel, deren 3D-Koordinaten aus dem CAD-System bekannt sind, herangezogen. Als Kamerasensor wird ein Stereo-Kamera-Paar verwendet, dessen intrinsische und extrinsische Parameter zuvor ermittelt wurden. Zur Bestimmung der realen 3D-Bohrungskoordinaten werden diese mit dem Kamerapaar aufgenommen, wobei sie sich typischerweise als Ellipsen in den Kamerabildern abbilden. Ist eine Zielbohrung in beiden Kameras sichtbar, kann durch Vorwärtsschnitt der Ellipsenzentren deren 3D-Koordinate bestimmt werden. Nach [14] wird mit den korrespondierenden 3D-Koordinaten aus dem CAD-Modell die Transformation der Schaltschranktafel direkt berechnet. Durch die elektrischen Komponenten und Kabelkanäle kann es jedoch zu Verdeckungssituationen kommen, so dass keine pauschale Positionierung des Stereo-Sensors zu den Bohrungen vorgegeben werden kann. Eine geeignete Aufnahmeposition muss somit für jede Zielbohrung jeder Schaltschranktafel-variante individuell bestimmt werden.

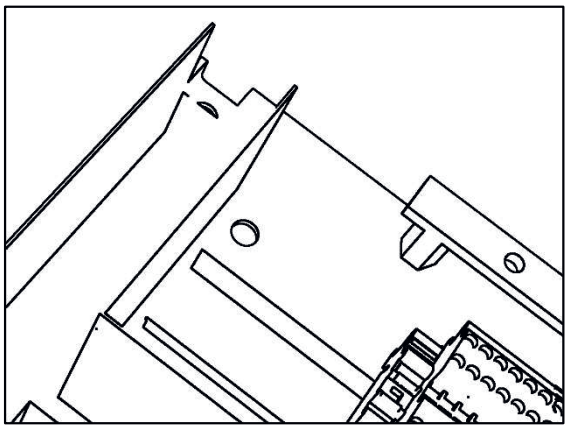


Abb. 10: Simulierte Objektkanten einer Schaltschranktafel. Links ein Kabelkanal, rechts unten Anschluss-klemmen, mittig die Zielbohrung.

Zur Lösung dieser Aufgabe wird eine feature-basierte Stereo-Kamerasimulation durch-geführt. Deren Ergebnis sind zwei Abbildungen, in denen die Kanten der gesamten Schalt-schranktafel dargestellt werden (s. Abb. 10).
Als prinzipiell geeignet werden Sensor-positionen bewertet, bei denen die Zielbohrung in beiden Kameras vollständig sichtbar ist. Die Qualität einer möglichen Position wird dadurch bestimmt, wie kreisförmig die Bohrung in den Kameras abgebildet wird.

Zur Abtastung des Lösungsraums wird der virtuelle Stereo-Sensor entlang einer Halbkugel oberhalb der Zielkoordinate platziert. Der Halbkugelradius ergibt sich aus dem optimalen Arbeitsabstand des Sensors. Zusätzlich wird bei jeder Position die Rotation des Sensors um seine Sichtachse variiert und die Position des Schaltschranktafelmodells innerhalb des Erwartungsbereichs verschoben. Zur Bewertung werden aus den simulierten Bildern Ellipsen nach [15] extrahiert und – wenn vollständig vorhanden – deren Verhältnis von langer zu kurzer Halbachse als Optimierungs-kriterium berechnet. Dabei werden Sensorpositionen besser bewertet, in denen die Bohrungsabbildung kreisförmiger ist, um eine möglichst senkrechte Aufnahmeposition zu erzielen.

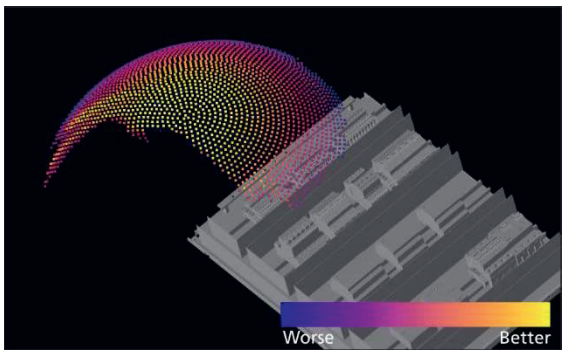


Abb. 11: Visualisierung des abgetasteten Lösungsraums. Farblich codiert die Qualität der zu erwartenden Messdaten aus der jeweiligen Position

Das Ergebnis ist eine Menge an potenziellen Sensortransformationen, bei denen sicher-gestellt ist, dass eine Zielbohrung aus beiden Kameras möglichst senkrecht und verdeckungsfrei erfasst werden kann.

Anwendung der 3D-Simulation

Ein weiteres Anwendungsbeispiel ist bei einem Sondermaschinenhersteller gegeben. Hier werden Guss- und Schweißteile durch CNC-Maschinen spanend bearbeitet. Charakteristisch ist dabei auch hier, dass die Teile mit einer sehr hohen Variantenvielfalt auftreten. Vor der Bearbeitung sollen die Rohteile einer Qualitätsprüfung unterzogen werden, bei der geometrische Abweichungen identifiziert werden. Die Geometrieprüfung soll für Großbauteile mit Abmessungen bis 5,4 x 1,7 x 0,8 m³ in einer separaten Prüfzelle punktuell stattfinden. Dazu wird ein Streifenlichtsensor eingesetzt, der über ein externes Tracking-System verfügt, mit dessen Hilfe lokale 3D-Sensormessdaten in ein globales

Prüfzellenkoordinatensystem transformiert werden können. Der Sensor ist handgeführt und kommt an kritischen Stellen der zu inspizierenden Teile zum Einsatz. Dort wird das Bauteils flächig erfasst und zur Bestimmung von Geometrieabweichungen eine Abstandsberechnung zum CAD-Modell vorgenommen.
Zur Prüfung wird das Bauteil meist auf einer Palette in der Mitte der Prüfzelle abgesetzt. Als erstes ist deshalb ein Einmessen des Bauteils im Koordinatensystem der Prüfzelle notwendig. Für das Einmessen genügt die 3D-Erfassung weniger Referenzflächen am Bauteil. In diese Messungen kann das 3D-Modell durch Minimierung der Punktabstände zur CAD-Oberfläche eingepasst werden (s. Abb. 12) und damit die Bauteilposition in der Prüfzelle bestimmt werden.

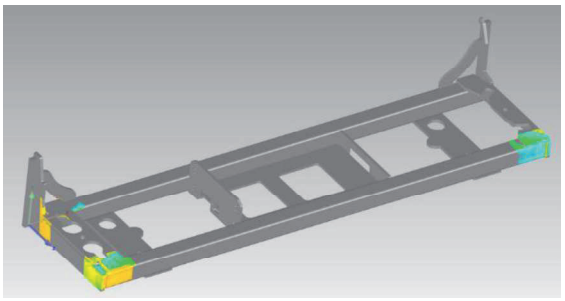


Abb. 12: Registrierung des CAD-Modells in 3D-Messdaten.

Welche Referenzflächen sich für das Einmessen besonders eignen, erfordert ein hohes Maß an Expertenwissen und wird daher von der Arbeitsvorbereitung manuell festgelegt. Im Allgemeinen orientieren sich die Flächen an der Aufspannung in der CNC-Maschine. Zur Markierung der jeweiligen Flächen steht den Mitarbeitern ein virtueller 3D-Pinsel zu Verfügung mit dem sie interaktiv im 3D-Modell Markierungen vornehmen können (s. Abb. 13).

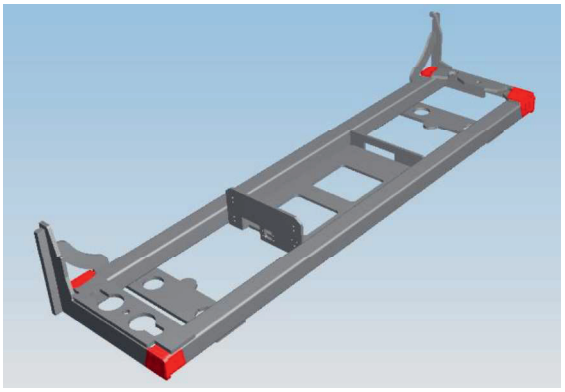


Abb. 13: Mit einem virtuellen 3D-Pinsel markierte Referenzflächen (rot) an einem Bauteil.

Eine geeignete Ausrichtung des Streifenlichtsensors zur Erfassung einer Referenzfläche kann mit Hilfe der zuvor beschriebene Simulationsverfahren automatisch ermittelt werden. Zum einen kann die Arbeitsvorbereitung eventuelle Probleme bei der Erfassung der markierten Referenzfläche erkennen (z.B. schlechte Erreichbarkeit) und zum anderen kann dem Bediener die Sensorausrichtung in einer 3D-Ansicht präsentiert werden, um ihn bei der Durchführung des Einmessens anzuleiten (s. Abb. 14).

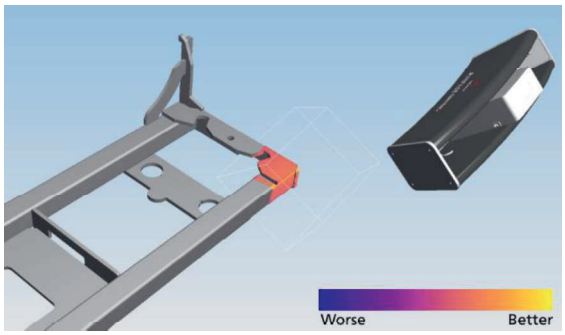


Abb. 14: Visualisierung einer einzunehmenden Sensorposition und der erwarteten 3D-Messdaten. Die Qualität der Messdaten ist farblich codiert.

Zur automatischen Bestimmung der Sensorausrichtung wird der virtuelle Streifenlichtsensor im optimalen Arbeitsabstand von 760mm kugelförmig um das Zentrum der markierten Referenzfläche positioniert. Empirisch wurde ermittelt, dass sich ca. 4000 Positionen auf der Kugel als ausreichend erweisen. Ein weiterer Freiheitsgrad des Sensors ist eine Drehung um die Sichtachse. Der ergonomisch sinnvolle Bereich wird hier durch Drehungen von -30° bis +30° in Schritten von 5° abgedeckt. Zur Bewertung der Sensorpositionen wird für jede Position des diskretisierten Lösungsraums eine 3D-Messdatensimulation durchgeführt und die Qualität der virtuellen 3D-Punkte evaluiert. Dazu werden im ersten Schritt zunächst alle Punkte verworfen, deren Oberflächennormale einen Grenzwinkel zu Kameras und Projektor des Sensors überschreitet. Für die typischerweise auftretenden metallischen Oberflächen in diesem Anwendungsbeispiel hat sich ein Schwellwert von 60° als gut herausgestellt. Zur weiteren Bewertung wird dann die Überlappung der Messdaten mit den markierten Flächen berechnet. Für Sensorpositionen, die eine gleich hohe Abdeckung erreichen, wird zusätzlich die Qualität der zu erwartenden Messdaten bewertet, indem der mittlere Oberflächenwinkel

aller Messpunkte berechnet wird. Dabei gilt, dass eine Sensorposition besser ist, je senkrechter die Flächen zum Sensor stehen. Aufgrund der recht matten Oberflächencharakteristik treten Totalreflexionen sehr selten auf.
Das Ergebnis dieser Simulationen sind Sensorposition, die der Bediener während der Bauteilreferenzierung ansteuern soll. Die Positionen werden als 3D-Darstellung visualisiert und dienen als intuitive Anleitung. So ist sichergestellt, dass Messdaten von genau den erforderlichen Referenzflächen aufgenommen werden können, anhand derer die Bauteilposition durch Punkt-zu-CAD-Modell-Registrierung automatisch berechnet werden kann.

Zusammenfassung

In diesem Beitrag wurde ein Framework vorgestellt, mit dem Messdaten von Kameras, Laserlichtschnitt- und Streifenlichtsensoren simuliert werden können. Die Simulation basiert auf den Abbildungs- und Projektionsparametern der Sensoren sowie auf den CAD-Modellen der zu erfassenden Objekte. Für eine möglichst universelle Beschreibung der Kameraabbildung wurde beschrieben, wie ein Mapping von Sichtpyramidenkoordinaten in Einheitskoordinaten stattfinden kann. Die Berechnung der Simulationsdaten zeichnet sich dank vollständiger Implementierung auf der GPU durch eine sehr hohe Geschwindigkeit aus. Dabei ist das Ergebnis der Simulation ausreichend realitätsnah, so dass die Daten beispielsweise für eine Ansichtenplanung verwendet werden können. Es wurden zwei Anwendungen beschrieben, bei denen die Simulation zur Planung von geeigneten Sensorpositionen in 2D und 3D praktisch zum Einsatz kommt.

Danksagung

Diese Arbeit ist im Projekt »3D-KOSYMA« entstanden. Das Projekt wird unter dem Förderkennzeichen 03ZZ0446H im Förderprogramm »Zwanzig20 – Partnerschaft für Innovation« teilweise durch das BMBF gefördert.

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Literaturnachweis

[1] Buchholz, B. et al.: Digitalisierung industrieller Wertschöpfung. Transformationsansätze für KMU, Berlin, 2017.
[2] Irgenfried, S. et al.: Image formation simulation for computer-aided inspection planning of machine vision systems. In (Beyerer, J.; Puente León, F. Hrsg.): Automated Visual Inspection and Machine Vision II. SPIE, 2017.
[3] Nilsson, J.; Ericsson, M.; Danielsson, F.: Virtual Machine Vision in Computer Aided Robotics: 2009 IEEE Conference on Emerging Technologies & Factory Automation Piscataway, NJ, 2009.
[4] Mahmud, M. et al.: 3D part inspection path planning of a laser scanner with control on the uncertainty. In Computer-Aided Design, 2011.
[5] Malik, A. et al.: A simulation of automatic 3D acquisition and post-processing pipeline. In (Niel, K. S.; Fofi, D. Hrsg.): Image Processing: Machine Vision Applications II. SPIE, 2009.
[6] Jørgensen, T. B. et al.: Simulation-based Optimization of Camera Placement in the Context of Industrial Pose Estimation. In (Imai, F. H.; Trémeau, A.; Braz, J. Hrsg.): Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018) - Volume 5: VISAPP, SciTePress, 2018.
[7] Xu, G.; Zhang, Z.: Epipolar Geometry in Stereo, Motion and Object Recognition. A Unified Approach. Springer, Dordrecht, 1996.
[8] Fasogbon, P. et al.: Intrinsic camera calibration equipped with Scheimpflug optical device. In (Meriaudeau, F.; Aubreton, O. Hrsg.): Twelfth International Conference on Quality Control by Artificial Vision 2015. SPIE, 2015.
[9] Kessenich, J. M.; Shreiner, D.; Sellers, G.: OpenGL programming guide. The official guide to learning OpenGL, version 4.5 with SPIR-V. Addison-Wesley, Boston, 2017.
[10] Press, W. H.: Numerical recipes in C++. The art of scientific computing. Cambridge Univ. Press, Cambridge, 2005.
[11] Xi, F.; Shu, C.: CAD-based path planning for 3-D line laser scanning. In Computer-Aided Design, 1999.
[12] Saito, T.; Takahashi, T.: Comprehensible rendering of 3-D shapes. In (Baskett, F. Hrsg.): SIGGRAPH '90 conference proceedings. ACM Press, New York, 1990.
[13] Brabec, S.; Annen, T.; Seidel, H.-P.: Practical Shadow Mapping. In Journal of Graphics Tools, 2002.
[14] Haralick, R. M. et al.: Pose estimation from corresponding point data. In IEEE Transactions on Systems, Man and Cybernetics, 1989.
[15] Teutsch, C. et al.: Real-time detection of elliptic shapes for automated object recognition and object tracking. In (Meriaudeau, F.; Niel, K. S. Hrsg.): Machine Vision Applications in Industrial Inspection XIV. International Society for Optics and Photonics, 2006.

DOI 10.5162/sensoren2019/1.1.2

HoloPort – 3D-Sensor for machine tools

Tobias Seyler, Johannes Engler
Fraunhofer Institute for Physical Measurement Techniques IPM,
Heidenhofstraße 8, 79110 Freiburg
tobias.seyler@ipm.fraunhofer.de

Abstract:

Manufacturing of high-precision components requires accuracies that even the most modern processing machines are often unable to deliver reliably. Slightly worn tools, incorrectly calibrated sensors or even different trajectories can lead to component integrities that do not comply with the desired specifications. Quality control is usually performed on randomized samples in special measuring rooms outside the machine tool. Closed-loop quality control is currently a cumbersome iterative process. We present the digital-holographic sensor system *HoloPort* that measures the two dimensional surface integrity of machined components in the sub-micrometer range directly inside machine tools: To our knowledge, *HoloPort* is the first wireless (digital holographic) sensor inside a machine tool worldwide. Being fully integrated, it features not only a multi wavelength interferometer but also a miniaturized GPU. This allows full data evaluation directly in the sensor and thus inside the machine tool. A single measurement is taken and processed within 3 s in wireless operation. *HoloPort* can be easily integrated into a variety of machine tools. This contribution includes detailed information about the sensor architecture. Experimental results on milled parts show the system's performance and illustrate possible inline applications as well as perspectives of the sensor.

Keywords: Digital Holography, Machine Tool, Inline Inspection, Parallel Computing, Wireless, GPU.

Introduction

In machining production, the measurement of component dimensions and parameters is a fundamental part of quality and efficiency assurance. Therefore, the demand for on-machine metrology is continuously increasing. It avoids time-consuming realignment operations and possible loss or damage during the transportation between different machining and measuring stations. Luo et al. for example consider “hybrid machining” as one of the main future developments in production and demand a measurement technology suitable for use directly at the processing location: *“However, currently there is no method to realize on-line measurement of surface integrity.”* [1] Here, we present a sensor that addresses this problem.

Separate coordinate measuring machines (CMM) as well as scanning tactile systems which can be exchanged as a “tool” are well-established to this day but are slow and cannot resolve the surface texture [2]. Competitive optical technologies allow contact-free operation but are either limited in their application (reflecting parts for deflectometry), slow (white-light interferometry, confocal, focus variation) or have insufficient accuracy (laser triangulation) [3]. In addition, none of these techniques is commercially available for in machine-use.

Featuring high measuring speed and accuracy at the same time, digital holography has become a versatile tool for the inline 100 % 3D quality control in industry [4]. The origins of digital holography date back to the 1950s [5]. However, only the development of electronic sensors [6] and increasing computation power such as provided through graphics processing units (GPU) enabled digital holography for industrial use [7–12]. With the introduction of small, powerful computing units from the field of autonomous driving, a fully integrated wireless measuring system becomes possible.

This contribution presents the first wireless interferometer that allows a two dimensional surface measurement of freshly machined components directly inside machine tools. In order to be capable of measuring optically rough technical surfaces, the sensor system works based on the principle of multi wavelength digital holography and achieves accuracies in the sub-micrometer range. It features a multi wavelength interferometer with three stabilized single mode, single frequency laser sources and a miniaturized GPU – all fully integrated in a compact housing and allowing wireless operation.

Sensor description

Integrating a surface-measuring optical sensor into a machine tool poses a number of challenges. In addition to the harsh environmental conditions a sensor has to withstand, tool holders are currently neither prepared for energy nor data transmission. The development of such standardized interfaces has been in progress for several years, but none is yet in use [13]. Thus, commercially available tactile sensors work wirelessly and are battery powered. In order to make the measuring system accessible for all machine tools, the standard interface according to DIN 69893 was chosen. Therefore, all electronics necessary for the operation of the sensor had to be integrated into the measuring system.

Fig. 1 depicts the system architecture: The sensor system *HoloPort* consists of the optical sensor head *HoloCut* and an integrated control system.

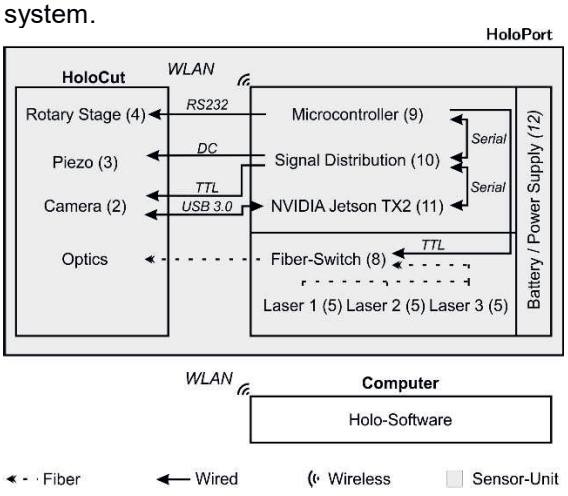


Fig. 1. System architecture: Equipped with the control module, the wired sensor head HoloCut [14] upgrades to the fully integrated measuring system HoloPort and operates completely wirelessly. Label numbers refer to *Fig. 2* and *Fig. 3*.

The optical design of HoloCut [15] – shown in Fig. 2 – is based on the industrially proven HoloTop sensor [4]. With a focus on the form factor, the height of a standard tool was set as the mechanical boundary condition for the mechanical design. Miniaturization of the sensor and mounting in the center of gravity (1) were achieved by a three-dimensionally folded beam path as shown in Fig. 2. The sensor head itself has a size of 235 × 140 × 215 mm³ at 7.5 kg. For further details regarding the optical and mechanical design, please refer to our publication [14].

At least three temporally phase-shifted interferograms [16] are recorded with a 9 Megapixel camera (2), corresponding to a lateral sampling rate of 7 µm × 7 µm and an (x, y) measurement area of 20 mm × 20 mm. Phase shifts are introduced by a piezo actuator (3). An integrated controlled rotary motor (4) allows the sensor to be adapted to different surface conditions.

For the reasons set out above, the control system has also been miniaturized and directly integrated onto the sensor head. The 19" control system (500 × 500 × 640 mm³) of *HoloTop* was reduced below the size of a shoe box (235 × 140 × 75 mm³).

Fig. 3 shows a computer-aided design (CAD) model of the complete measuring system *HoloPort*. Due to the sensor's field of application are technical surfaces, z measurement ranges of at least a couple micrometer are needed. To achieve this, multiple lasers are used to record interferograms at different wavelengths [17]. These are used to numerically generate interferograms at so-called *synthetic* wavelengths, corresponding to the resulting beat frequencies. The wavelength difference determines the unambiguous height measurement range: The smaller the difference, the larger the synthetic wavelength. Large measurement ranges result in less accurate measurements. The combination of more than two wavelengths allows cascading the measurement ranges, and thus increases the ratio of measurement range and accuracy [4].

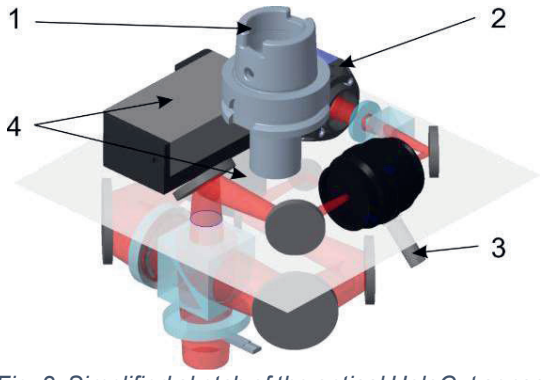


Fig. 2. Simplified sketch of the optical HoloCut sensor head with interfaces to the control unit: camera (2), piezo actuator (3) and rotary stage (4). The machine integration in the center of gravity of the sensor (1) is realized by a three-dimensional layout of the optical beam path.

The lasers are mounted onto an aluminum block (6) and insulated thermally (7). A fiber switch (8) sequentially couples these lasers into the sensor head. The system can be configured with different laser wavelengths. The results shown in this paper were achieved using synthetic wavelengths of 60 µm to 88 µm. The extension of the measurement range to several millimeters has already been shown by using

additional lasers [9] or specialized algorithms [18].

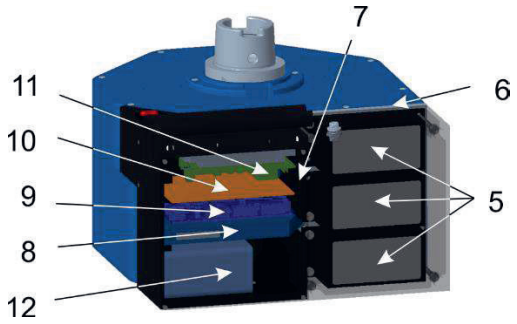


Fig. 3. CAD model of the integrated control system including a battery powered (12) embedded architecture (9-11), a fiber-switch (8) toggling a three-laser system (5).

The previously introduced components are controlled by a microcontroller (9). Energy supply and signal distribution is performed by a specially designed circuit board (10). Since the speed of the sensor system determines the economic efficiency and benefits in the machine tool decisively, a powerful CUDA capable computing unit has been integrated: A Jetson TX2 board (11) allows not only wireless data transfer according to the state of the art, but also allows data reduction inside the sensor itself. The resulting camera data of up to 270 MB per measurement can be pre-processed internally and thus reduced by half – a processing time of less than 3 s is achieved for 9 million sample points. The pure measuring time is less than 1 s.

In addition, the computing power can be used for in situ quality control of the acquired raw data. Interferograms disturbed by vibrations are internally evaluated for sufficient interference contrast [19] even before the wireless data transmission takes place. The power consumption of the whole system is in the range of 20 W. Powered by a 5.2 Ah battery; it lasts for up to 7 h and is charged via a magnetic connector within 30 minutes.

HoloPort is the first wireless (digital holographic) sensor inside a machine tool worldwide.

Applications

HoloPort has been developed for use in machine tools and quantitative full field analysis of manufactured surfaces. Those can be classified into two main categories: Engineered surfaces and non-engineered surfaces. Engineered surfaces such as bearings and seals are generally manufactured to provide functional properties or are subject to special safety requirements [20]. Fig. 4 shows a photograph of the sensor system inside a Hermle C32 U 5-axis milling machine at Fraunhofer IPM.



Fig. 4. Inspection of a typical milling surface by HoloPort inside a Hermle C32U 5-axis milling machine at Fraunhofer IPM.

In this Fig. 4, a typical non-engineered milling surface is inspected. *HoloPort* easily resolves the milling structure of the surface revealing height differences *h* between individual milling paths of up to 10 µm as shown in Fig. 5. A cross section of ten individual measurements at the same position illustrates the repeatability in the sub-µm range. With a measuring field of up to 20 mm × 20 mm at 9 million sample points, irregularities within the microstructure as well as machine faults quickly become apparent. In contrast to the tactile state of the art sensors, the system cannot only resolve long-wave components but also short-wave structures and defects of the milling operation. Due to the inherent morphological filtering of the probe head, tactile sensors in principle cannot detect these characteristics.

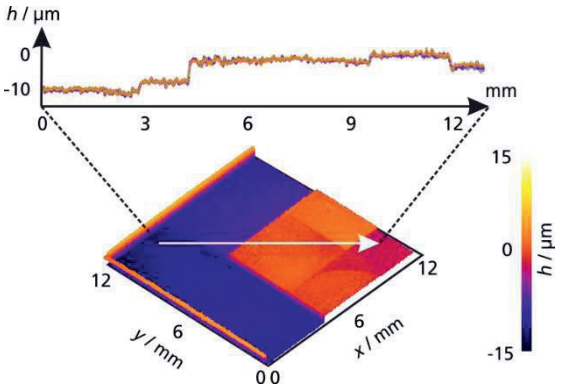


Fig. 5. Measured height differences h between milling paths of a typical milled part in the range of 10 µm. A cross section of ten individual measurements at the same position demonstrates the measuring repeatability.

For example, hidden machine faults could be observed like an incorrectly calibrated temperature sensor leading to unpredictable surfaces displacing whole milling paths by more than 10 µm [19]. For the human eye, this defect

is not visible on the surface and the machine was unable to register the error at all.

Fig. 6 and Fig. 7 show measurements of a surface machined with an end mill at 5.5 mm axial tool infeed, respectively.

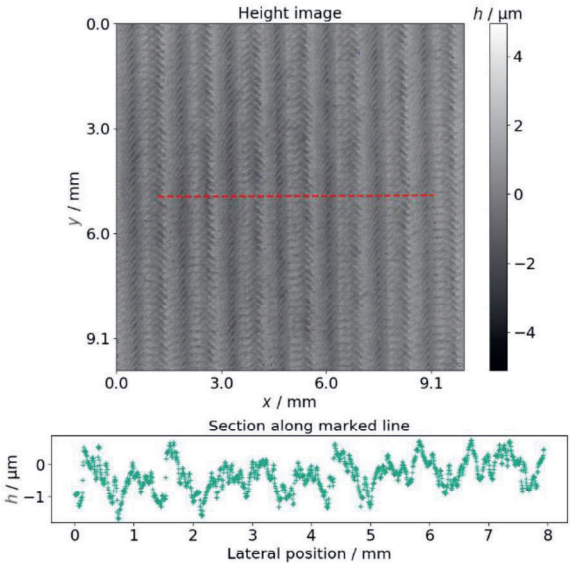


Fig. 6. Milling surface machined with an end mill under 5.5 mm axial infeed alternate milling.

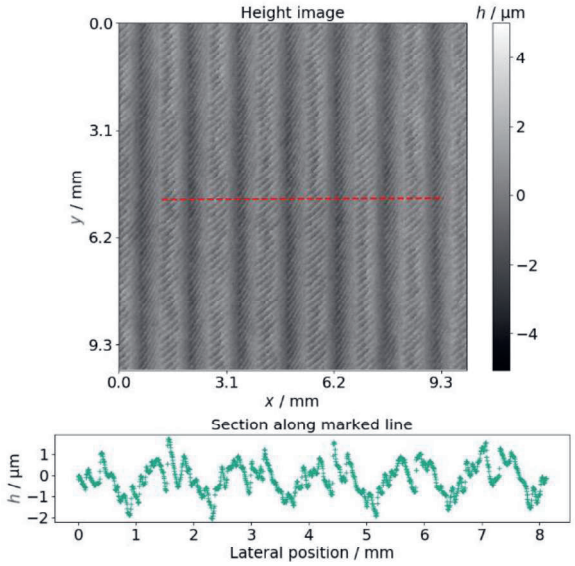


Fig. 7. Milling surface machined with an end mill under 5.5 mm axial infeed parallel milling.

Looking at the two-dimensional surface plot, a clear difference can be observed. However, since the surface roughness is within the range of only 3 µm, a tactile probe could not resolve these short-wave structures due to its inherent morphological filtering. Calculating the material ratio curve e.g., parameter optimization can be carried out directly inside the machine. In case of an engineered surface with requirements on the contact ratio, the high resolving two-dimensional plot is much more significant.

Fig. 8 shows the milling result with reduced axial infeed of only 0.5 mm using alternate milling. With a reduction of the axial infeed, a lower surface roughness is to be expected. In fact, the surface roughness decreases, but with a varying tilting of the individual toolpaths towards each other (manually marked in orange). This effect is beyond the machine accuracy and can be caused by the tolerances between tool, tool holder and spindle during the tool change itself. Further possible causes are insufficient cutting pressure or even a tilted or defective spindle.

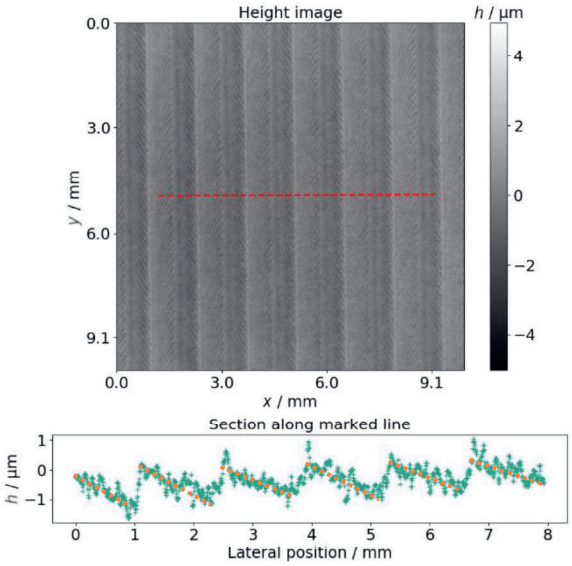


Fig. 8. Improved, less rough milling surface of an end mill with only 0.5 mm axial infeed. The varying tilting of the individual toolpaths towards each other is manually marked in orange.

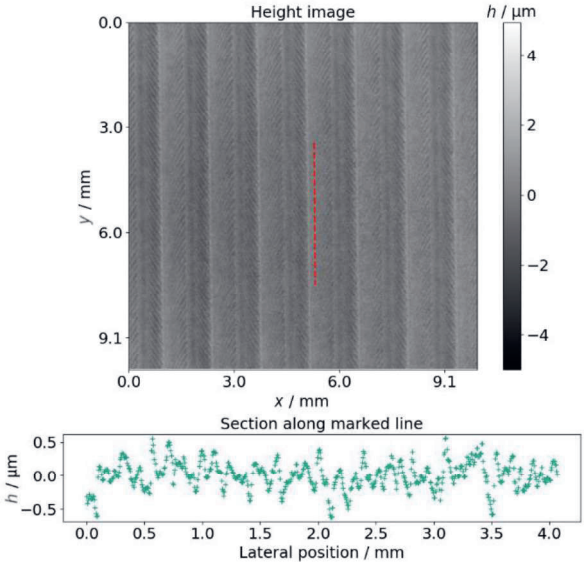


Fig. 9. Line section within a single milling path. Without the share being caused by the path tilts, the surface roughness ranges only within 1 µm.

Fig. 9 shows a single section in milling direction. The inclination of the individual toolpaths to each other is not visible within a single line cut. If one could eliminate the tilting of the milling paths by suitable measures, the surface roughness could be approximately halved. Especially in the production of injection molds with subsequent polishing, the surface quality and frequencies are of great importance. Certain spatial frequency components can be removed much better than others during polishing [21].

The benefit of the measuring system in the optimization of milling parameters quickly becomes obvious at this point.

For comparison to a white-light interferometric and confocal laboratory reference 3D measuring devices, surface parameters according to ISO 25178 were evaluated previously: All three measuring methods deliver comparable results and are able to detect both an increasing roughness of the surface due to wear of the milling cutter and a machine error [22].

Not only the analysis of surface parameters, but also the extensive measurement of different inspection parameters represent applications for quality monitoring [21]. Fig. 9 shows a milled laser heat sink with mapped surface measurement.

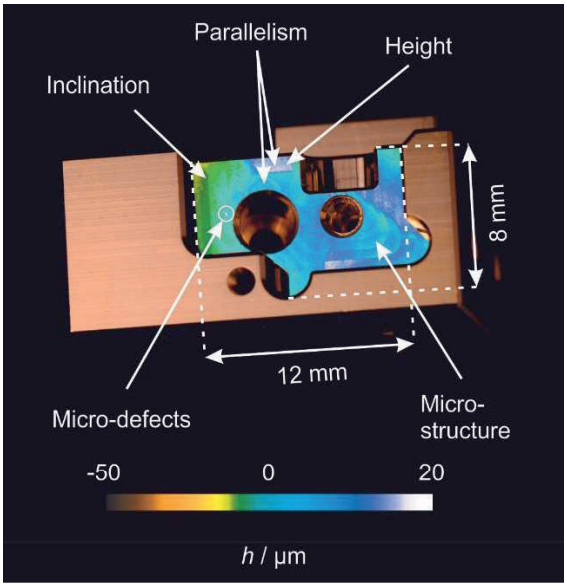


Fig. 10. HoloPort maps the component surface in the fraction of a second with micrometer accuracy directly in the machine tool (as in the example of a heat sink for laser assembly). Alongside the geometry, relevant surface parameters such as micro-defects and microstructures as well as the height, parallelism and inclination of surfaces can be measured simultaneously.

The full field measurement shows the surface quality and in addition allows the quantitative extraction of geometrical parameters: Micro-defects and microstructures as well as the height, parallelism and inclination of surfaces can be evaluated based on the gathered measurement data.

Using the dynamic numerical control (DNC) interface of the machine tool, axis positions can be read out and used for a CAD comparison – the system becomes a full-field CMM. Even complex work pieces can be inspected completely and directly within the machine for the first time [23]. The accuracy of a complete point cloud of the surface is limited by the accuracy of the machine tool axes, which is typically in the single micrometer range.

Further test measurements of a complex formed milling part have been carried out to demonstrate feasibility of quality inspection directly in the machine tool instead of using an additional CMM situated within a temperature controlled measurement room. The combination of more than 30 measurements for flatness determination has been carried out with an extension of 100 mm × 60 mm which corresponds to almost 300 million data points. Evaluation of the combined data set shows flatness information within the single micrometer range. The determination of lateral distances between reference marks located on the measured surface in the range of the lateral sampling of the sensor at 7 µm appears realistic.

Perspectives

An optical sensor measuring the surface integrity at a large field of view directly inside the machine tool opens up new possibilities for process control and optimization [24].

The feedback of the measurement results into the manufacturing process is a key point, especially in hybrid manufacturing – the automated batch size 1 production according to industry 4.0 becomes possible. Since the Jetson module was developed for autonomous driving purposes, the system is also prepared for machine learning approaches.

The system is also prepared for the use on CMM and robots – the possibilities and limits will be determined in future studies. In addition, the measuring system will be further miniaturized to be compatible with smaller size tool changers.

Digital holography has been proven to work even on moving objects at speeds of several mm/s [25]. With a resulting simplified handling, the upper limit for production cycles can be further decreased. The limits of a scanning

sensor in a multi-axis system will be analyzed in future experiments.

Summary

On-line measurements of surface integrity in machined production have been motivated. Digital holography meets the requirements for measuring accuracy in the µm range at high measuring speed at the same time.

A new digital-holographic sensor system for two-dimensional surface measurements inside a machine tool has been presented. The sub-µm precision of the system on an area of up to 20 mm x 20 mm was demonstrated with various milled samples. Due to the short processing time of 3 s (pure measuring time less than 1 s), closed-loop quality control directly inside the machine tool is feasible for the first time.

Acknowledgement

This project was financed by the Baden-Württemberg Stiftung gGmbH.

We appreciate the support and the good cooperation during the integration of the measuring system by Hermle AG.

References

[1] X. J. Jiang, F. Gao, H. Martin, J. Williamson, D. Li, On-Machine Metrology for Hybrid Machining. In: X. Luo, Y. Qin (eds.). *Hybrid Machining*: *Theory, Methods, and Case Studies*. San Diego: Elsevier Science & Technology; 2018.

[2] M. Weck, C. Brecher, *Werkzeugmaschinen 4*. Springer Berlin Heidelberg, Berlin, Heidelberg; 2006.

[3] M. C. Knauer, C. Richter, G. Häusler, 3D sensor zoo – Species and natural habitats, *Laser Technik Journal* 3, 33–37 (2006); doi: 10.1002/latj.200790081.

[4] M. Fratz, T. Beckmann, A. Schiller, T. Seyler, A. Bertz, D. Carl, K. Buse, Digital Holography: Evolution from a Research Topic to a Versatile Tool for the Inline 100% 3D Quality Control in Industry. In: *AMA Conferences 2017 - SENSOR 2017 and IRS2 2017*, 286–289; 2017.

[5] D. Gabór, A New Microscopic Principle, *Nature* 161, 777–778 (1948); doi: 10.1038/161777a0.

[6] J. W. Goodman, R. W. Lawrence, Digital Image Formation From Electronically Detected Holograms, *Journal of the Optical Society of America* 11, 77–79 (1967); doi: 10.1063/1.1755043.

[7] U. Schnars, W. Jüptner, Direct recording of holograms by a CCD target and numerical reconstruction, *Applied Optics* 33, 179–181 (1994); doi: 10.1364/AO.33.000179.

[8] S. Kuwamura, I. Yamaguchi, Wavelength scanning profilometry for real-time surface shape measurement, *Applied Optics* 36, 4473 (1997).

[9] D. Carl, M. Fratz, M. Pfeifer, D. M. Giel, H. Höfler, Multiwavelength digital holography with autocalibration of phase shifts and artificial wavelengths, *Applied Optics* 48, H1-H8 (2009).

[10] M. Fratz, D. Carl, Novel Industry Ready Sensors for Shape Measurement Based on Multi Wavelength Digital Holography. In: W. Osten (ed.). *Fringe 2013*: *7th International Workshop on Advanced Optical Imaging and Metrology*, 479–484. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014.

[11] J.V. M. Bove, W. J. Plesniak, T. Quentmeyer, J. Barabas, Real-time holographic video images with commodity PC hardware. In: A. J. Woods, J. O. Merritt, M. T. Bolas, I. E. McDowall (eds.). *Stereoscopic Displays and Virtual Reality Systems XII*, 255–262; 2005.

[12] N. Masuda, T. Ito, T. Tanaka, A. Shiraki, T. Sugie, Computer generated holography using a graphics processing unit, *Optics Express* 14, 603 (2006); doi: 10.1364/OPEX.14.000603.

[13] Komet Group GmbH, *Bauteilgerechte Maschinenkonfiguration in der Fertigung durch Cyber-Physische Zusatzmodule: BaZMod*; 08.11.2017.

[14] T. Seyler, M. Fratz, T. Beckmann, A. Bertz, D. Carl, Miniaturized multiwavelength digital holography sensor for extensive in-machine tool measurement. In: P. Lehmann, W. Osten, A. Albertazzi Gonçalves (eds.). *Optical Measurement Systems for Industrial Inspection X*: *26-29 June 2017, Munich, Germany*, 103290F. Bellingham, Washington, USA: SPIE; 2017.

[15] T. Seyler, M. Fratz, T. Beckmann, A. Bertz, D. Carl, 3D-Prüfung von Präzisionsoberflächen: Qualitätssicherung – submikrometergenau im Sekundentakt, *WOMAG* (2017).

[16] L. Z. Cai, Q. Liu, X. L. Yang, Generalized phase-shifting interferometry with arbitrary unknown phase steps for diffraction objects, *Optics Letters* 29, 183 (2004); doi: 10.1364/OL.29.000183.

[17] C. Wagner, W. Osten, S. Seebacher, Direct shape measurement by digital wavefront reconstruction and multiwavelength contouring, *Optical Engineering* 39, 79–85 (2000).

[18] T. Seyler, M. Fratz, T. Beckmann, A. Schiller, A. Bertz, D. Carl, Extending the

Depth of Field beyond Geometrical Imaging Limitations Using Phase Noise as a Focus Measure in Multiwavelength Digital Holography, *Applied Sciences* 8, 1042 (2018); doi: 10.3390/app8071042.

[19] T. Seyler, M. Fratz, T. Beckmann, A. Bertz, D. Carl, V. Grün, R. Börret, F. Ströer, J. Seewig, Extensive microstructural quality control inside a machine tool using multiwavelength digital holography. In: M. Kujawińska, L. Jaroszewicz (eds.). *SPECKLE 2018 - VII International Conference on Speckle Metrology*; 2018.

[20] K.J. Stout, L. Blunt, A contribution to the debate on surface classifications—random, systematic, unstructured, structured and engineered, *International Journal of Machine Tools and Manufacture* 41, 2039–2044 (2001); doi: 10.1016/S0890-6955(01)00069-4.

[21] M. Speich, R. Börret, W. Rimkus, D. K. Harrison, Material removal in polishing processes - finite element simulation and experiment. In: Glasgow Caledonian University (ed.). *10th International Conference on Manufacturing Research*; 2012.

[22] T. Seyler, M. Fratz, V. Grün, R. Börret, F. Ströer, J. Seewig, Messen am eingerichteten Werkstück, *Qualität und Zuverlässigkeit* 63, 52–54 (2018).

[23] T. Seyler, M. Fratz, V. Grün, R. Börret, Flächige 3D-Messung in der Werkzeugmaschine, *Maschinenmarkt*, 10–12 (2018).

[24] X. Luo, Y. Qin, eds., *Hybrid Machining: Theory, Methods, and Case Studies*. Elsevier Science & Technology, San Diego; 2018.

[25] A. Schiller, T. Beckmann, M. Fratz, D. Belzer, A. Bertz, D. Carl, K. Buse, Digital holography on moving objects: interference contrast as a function of velocity and aperture width, *Applied Optics* 56, 4622–4628 (2017); doi: 10.1364/AO.56.004622.

DOI 10.5162/sensoren2019/1.1.3

3-Komponenten-Laser-Doppler-Anemometer mit Selbst-Kalibrierung

Harald Müller[1], *Christian Gutsmuths*[1], *Michael Eggert*[1], *Helmut Többen*[1]
[1] *Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Deutschland*

Zusammenfassung

Auf die SI-Einheiten messtechnisch rückgeführte, präzise und rückwirkungsfreie Strömungsgeschwindigkeitsmessungen erfolgen in der Regel mit Laser Doppler Anemometern als Bezugsnormalen. Hierzu werden Laser-Doppler-Anemometer (LDA) kalibriert, wobei die Kalibrierung für jede einzelne, Komponente des Strömungsgeschwindigkeitsvektors unabhängig von der geometrischen Lage der Vektorkomponenten zueinander erfolgt. Daher ist für viele Messanordnungen zur Messung von mehr als einer Strömungsgeschwindigkeitskomponente eine messtechnisch rückgeführte Messung des resultierenden Strömungsgeschwindigkeitsvektors problematisch. Das in der PTB entwickelte 3-Komponenten-LDA mit Selbstkalibrierung hingegen sieht vor, durch eine während des Messvorgangs ausgemessene und nachgeregelte Messgeometrie, Strömungsgeschwindigkeitsvektoren rückgeführt zu messen. Der LDA-Arbeitsabstand im Bereich von 1,5 m und der angestrebte Geschwindigkeitsbereich von bis zu 100 m/s sollen insbesondere Möglichkeiten zu rückgeführten Messungen in größeren Windkanälen zur Analyse komplexer Strömungsvorgänge eröffnen.

Keywords: 3-K-Laser-Doppler-Anemometrie, Bezugsnormal, Strömungsanalyse, Doppler-Referenz-Verfahren, messtechnische Rückführung.

Einleitung

Laser-Doppler-Anemometer haben sich in den letzten Jahrzehnten in vielen Anwendungsbereichen der Strömungsmesstechnik als präzise Messsysteme insbesondere für rückwirkungsfreie, örtlich und zeitlich hochaufgelöste Strömungsgeschwindigkeitsmessungen etabliert. Während für die Kalibrierung von Anemometern in Windkanälen vorzugsweise kalibrierte, d.h. auf die SI-Einheiten rückgeführte, 1-K-LDA-Systeme als Bezugsnormale zur Bestimmung der Geschwindigkeitskomponente in Strömungsrichtung eingesetzt werden, ist für die Analyse komplexer Strömungsvorgänge die Messung aller drei Komponenten des Strömungsgeschwindigkeitsvektors erforderlich, wobei zunehmend auch hier die messtechnische Rückführung zur Sicherstellung der Messrichtigkeit gefragt ist. Die messtechnische Rückführung konventioneller 3-K-LDA-Systeme ist jedoch problematisch, da die meisten für größere Arbeitsabstände > 1 m eingesetzten Systeme [3, 5] auf dem Einsatz von zwei auf den gleichen Messort anwendungsspezifisch auszurichtenden Messköpfen basieren, so dass eine eindeutig definierte Gesamt-Messgeometrie in der Regel nicht vorliegt. Kompakte 3-K-LDA-Systeme mit einem einzigen Messkopf [1, 2, 4, 6] sind hingegen in der Regel nur für Arbeitsabstände bis zu ca. 1 m ausgelegt, wobei die Messgeometrie für die drei Komponenten derartiger Systeme nicht rückgeführt ist. Somit fehlt insbesondere für die messtechnisch rückführbare Messung von Strömungsgeschwindigkeitsvektoren in Großwindkanälen mit Arbeitsabständen > 1m eine geeignete Messtechnik. Um für größere Arbeitsabstände wie z.B. für Messaufgaben in Großwindkanälen die messtechnische Rückführung von 3-K-Strömungsgeschwindigkeitsvektormessungen zu ermöglichen, wird in der PTB ein Referenzstrahl-Laser-Doppler-Anemometer (LDA) entwickelt, das auf einer während des Messvorgangs selbst vermessenen und einjustierten Messgeometrie basiert.

Messprinzip

Das entwickelte 3-K-Referenzstrahl LDA nutzt einen Messtrahl und drei Beobachtungsstrahlen [7], wobei die messtechnische Rückführung auf die SI-Einheiten auf der Kenntnis der Laserwellenlänge, der Messung der Doppler-Frequenzverschiebung in den einzelnen Beobachtungsstrahlen und – als entscheidender Neuerung! – der durch Regelalgorithmen im Rahmen der Messunsicherheit eingehaltenen Messgeometrie beruht. Das Grundprinzip ist in Abbildung 1 dargestellt und gleicht dem Prinzip des in der PTB bereits entwickelten in der Validierungsphase befindlichen bistatischen Wind-Lidar-Systems [8].

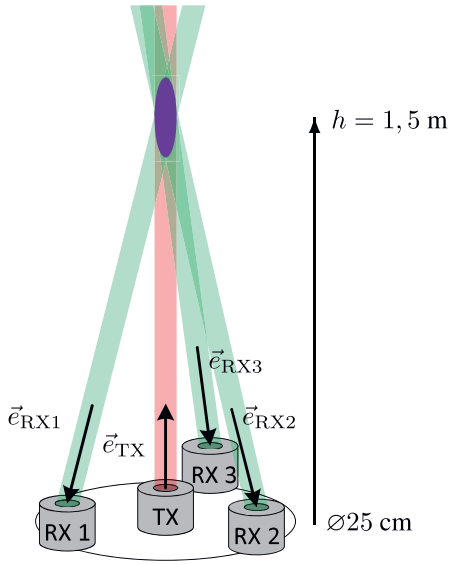


Abb. 1: Messgeometrie des 3-K-Referenzstrahl LDA-Systems mit einem Sendestrahl (rot) und drei Beobachtungsstrahlen (grün), alle Strahlen sind auf den Messort im Abstand von 1,5 m fokussiert und ausgerichtet.

Bei dem zugrunde liegenden Messprinzip wird der Strömungsgeschwindigkeitsvektor $\vec{v}$ aus den gemessenen Doppler-Frequenzverschiebungen f_{Di} (i = 1, 2, 3) des von den Beobachtungsstrahlen erfassten Streulichts von in der Strömung mitgeführten Partikeln nach Gl. 1 bestimmt:

$$\vec{v} = \lambda_{TX} \cdot M^{-1} \cdot \begin{pmatrix} f_{D1} - f_{D2} \\ f_{D1} - f_{D3} \\ f_{D1} + f_{D2} + f_{D3} \end{pmatrix} \quad (1)$$

mit λ_{TX}: Laserwellenlänge, M: Geometriematrix aus Gl. 2 - 5 und f_{Di} (i = 1, 2, 3) nach Gl. 6:

$$M = \begin{pmatrix} M_x & M_y & M_z \end{pmatrix} \quad (2)$$

$$M_x = \begin{pmatrix} e_{RX1_x} - e_{RX2_x} \\ e_{RX1_x} - e_{RX3_x} \\ e_{RX1_x} + e_{RX2_x} + e_{RX3_x} - 3 \cdot e_{TX_x} \end{pmatrix} \quad (3)$$

$$M_y = \begin{pmatrix} e_{RX1_y} - e_{RX2_y} \\ e_{RX1_y} - e_{RX3_y} \\ e_{RX1_y} + e_{RX2_y} + e_{RX3_y} - 3 \cdot e_{TX_y} \end{pmatrix} \quad (4)$$

$$M_z = \begin{pmatrix} e_{RX1_z} - e_{RX2_z} \\ e_{RX1_z} - e_{RX3_z} \\ e_{RX1_z} + e_{RX2_z} + e_{RX3_z} - 3 \cdot e_{TX_z} \end{pmatrix} \quad (5)$$

$$f_{Di} = \frac{\vec{v}}{\lambda_{TX}} \cdot (\vec{e}_{RXi} - \vec{e}_{TX}) \quad (6)$$

Zur Messung der Doppler-Frequenzen wird - dem Referenzstrahlverfahren entsprechend - jeweils ein vom Laser TX mittels Faserkoppler ausgekoppelter Strahlungsanteil mit dem detektierten Dopplerverschobenen Streulicht überlagert. Dabei erfolgt die Überlagerung über 90° Hybridkoppler auf balanced Photodetektor-Paaren zur Bereitstellung des Real- und Imaginärteils der auszuwertenden Schwebungssignale (s. Abb.2).

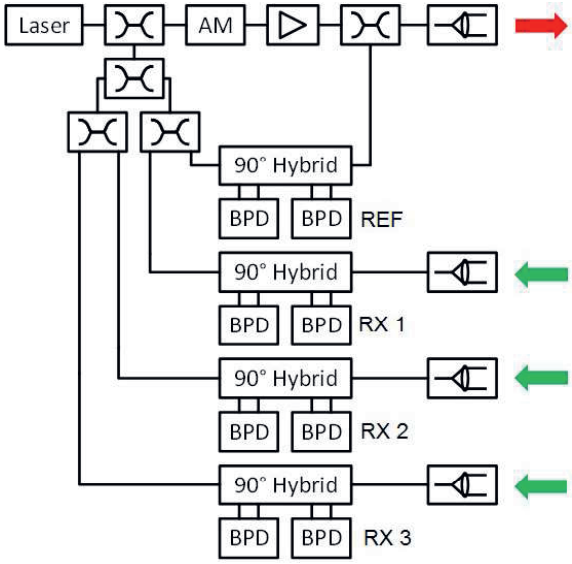


Abb. 2: Blockschaltbild zum Messprinzip des 3-K-Referenzstrahl LDA-Systems mit Kollimationsoptiken (rot: Messtrahl, grün: Beobachtungsstrahlen). Die Hybridkoppler liefern jeweils 90° in der Signalphase gegeneinander verschobene Gegentaktsignale für die „balanced Photo-Dioden" (BPD) zur Referenz- (REF) und Messignalerzeugung (RX 1, RX 2, RX 3).

Die Herausforderung bei dem 3-K-Referenzstrahl LDA besteht in der genauen Bestimmung der Messgeometrie und damit in der exakten Erfassung des gewählten Messabstands im Meterbereich, dessen Kenntnis eine notwendige Voraussetzung für die messtechnische Rückführbarkeit des Strömungsgeschwindigkeitsvektors darstellt. Die Abstandsbestimmung erfolgt über Signalphasen-Laufzeitdifferenzen (TX-RX$_i$) in allen drei Beobachtungsrichtungen und erfordert für die benötigte Auflösung die Verwendung zweier Lichtfrequenzen mit hinreichend großem Frequenzabstand innerhalb der Messsignalbandbreite. Dies wird durch den Amplitudenmodulator (AM) (s. Abb. 2) zur Erzeugung zweier optischer Frequenzen, der Laseremissionsfrequenz plus und minus 80 MHz, erreicht. Als Laser wird ein Faserlaser mit einer Wellenlänge von 1550 nm verwendet, dessen Ausgangssignal über einen Faserverstärker auf ca. 1 W (cw) verstärkt wird. Die eingesetzten Faserkoppler dienen der optischen Leistungsteilung für die Schwebungssignalbildung

der komplexen Signale für das Referenzsignal und die drei Messsignale. Die Abbildung 3 zeigt das Blockschaltbild zur Signalverarbeitung der einzelnen Dopplerfrequenzen und Signalphasen, aus denen die Geschwindigkeitskomponenten des Strömungsvektors (s. Matrix Gln. 2 - 5) und der Abstand zum Messort bestimmt werden.

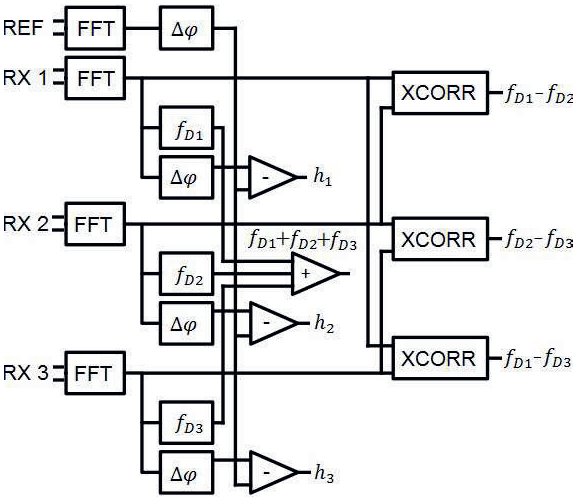


Abb. 3: *Blockschaltbild zur Signalverarbeitung 3-K-Referenzstrahl LDAs.*

Aus den Signalphasen der drei Beobachtungsstrahlenpfade werden die jeweiligen innerhalb der Mess- und Justierunsicherheiten übereinstimmenden Messabstände (h_1, h_2, h_3) bestimmt, deren Mittelwert den genauen Abstand zum Messort angibt. Die Positionierung der einzelnen Beobachtungsstrahlen erfolgt durch piezoelektrische Stellelemente für die Horizontal- und Vertikalrichtungen. Dabei werden für einen vorgegebenen Messabstand durch Scannen der Beobachtungsstrahlen, die optimale Ausrichtung auf den Messstrahl und durch Korrelation der drei Messignale eine optimale Überlagerung aller Strahlen am Messort realisiert. Durch die Korrelation der Messsignale ist sichergestellt, dass alle drei Beobachtungsstrahlen auf den gleichen Messort ausgerichtet sind und die Auswertung der Dopplersignale aus den unterschiedlichen Beobachtungsrichtungen sich auf die jeweils gleichen Streuteilchen in dem durch den Überlappungsbereich aller Strahlen definierten Messvolumen bezieht. Damit lässt sich der Strömungsgeschwindigkeitsvektor aus den Strömungsgeschwindigkeitsvektorkomponenten von einzelnen Streuteilchen innerhalb des Messvolumens eindeutig bestimmen und durch die Kenntnis der während des Messvorgangs eingeregelten und gemessenen Beobachtungsgeometrie messtechnisch auf die SI-Einheiten rückführen. Die Abbildung 4 zeigt den Gesamtaufbau des 3-K-LDA-Systems mit Ansteuereinheit und Messkopf ohne Abdeckplatte. Zu sehen sind die Kollimationsoptiken mit den justierbaren Spiegeloptiken in den Ecken des Messkopfs sowie die Sendeoptik in der Mitte des Messkopfs. Abbildung 5 zeigt den 3-K-Messkopf mit Abschlussdeckel und den vier planparallelen entspiegelten Fenstern für den Messstrahl und die drei Beobachtungsstrahlen.



Abb. 4: *3-K-Referenzstrahl-LDA-System*



Abb. 5: *3-K-Referenzstrahl-LDA-Messkopf*

Ausblick

Vorgestellt wurde ein 3-K-Referenzstrahl-LDA, das durch die mögliche Positionierung der Beobachtungsstrahlen, die Korrelation der drei Messsignale und die Auswertung der einzelnen Signallaufzeiten die Realisierung einer vom System selbst eingemessenen Geometrie zur Bestimmung des kompletten Strömungsgeschwindigkeitsvektors erlaubt. Auf diese Weise ist ein 3-K-Laser-Doppler-Anemometer realisierbar, dessen Regelalgorithmen erstmals die Selbstkalibrierung des Systems für einen vorgebbaren Arbeitsabstand ermöglicht. Das Ziel der im Rahmen einer Kooperation mit der Deutschen WindGuard Wind Tunnel Services GmbH durchgeführten Entwicklung ist der Einsatz des Messsystems für die rückführbare Messung aller Komponenten des Strömungsgeschwindigkeitsvektors in Großwindkanälen.

Literatur

[1] G. D. Byrne, S. W. James, R. P. Tatam, Meas. Sci. Technol. 15 (2004) 1 – 8; A single-headed fibre optic laser Doppler anemometer probe for the measurement of flow angles

[2] S. W. James, R. A. Lockey, D. Egan, R. P. Tatam, R. L. Elder: 3D Fibre Optic Laser Doppler Velocimetry, SPIE Vol. 2839/323 – 2839/334, 2016

[3] Dantec Dynamics: Integrated solutions in Laser Doppler Anemometry, Publication No.: 318_v1

[4] Dantec Dynamics: FiberFlow Probe, 112 mm diameter, 5 beams, Three velocity components with a single probe, Publication No.: PI_168_v10

[5] TSI Incorporated: Powersight Laser Doppler and Phase Doppler Measurement Systems, P/N 5001100 Rev C, 2014

[6] TSI Incorporated: 5-Beam Fiber Optic Probe Model TR360, P/N 1990677 Rev D, 2014

[7] C. Gutsmuths, M. Eggert, H. Müller: Konzeption eines Mehrkomponenten-Referenzstrahl-LDA in Rückstreurich-tung zur rückgeführten Messung von komplexen Strömungen in Großwind-kanälen, Proc. 25. GALA-Fachtagung, 4/1 – 4/7, 2017

[8] S. Oertel, M. Eggert, C. Gutsmuths, H. Müller, H. Többen: Three-component fiber laser based Lidar sensor for traceable highly resolved wind vector measurements, ITG-Fb. 281: Sensoren und Messsysteme, 2018

DOI 10.5162/sensoren2019/1.1.4

Optisches Lokalisierungssystem für autarke Sensorknoten in industriellen IoT Anwendungen

Leander B. Hörmann[1], Markus Pichler-Scheder[1], Peter Priller[2], Hans-Peter Bernhard[3], Andreas Springer[3]
[1]Linz Center of Mechatronics GmbH, Area: Sensors and Communication, Altenberger Straße 69, 4040 Linz, Österreich
[2]AVL List GmbH, ITS Global Research & Technology, Hans-List-Platz 1, 8020 Graz, Österreich
[3]Institut für Nachrichtentechnik und Hochfrequenzsysteme, Johannes Kepler Universität Linz, Altenberger Straße 69, 4040 Linz, Österreich

Zusammenfassung

Zur Überwachung von industriellen Prozessen und zur laufenden Messdatenerfassungen von Industrieanlagen und -Geräten werden immer mehr und mehr Sensoren benötigt. Diese Entwicklung wird vor allem durch industrielle Internet-of-Things (IoT) Anwendungen gleichzeitig ermöglicht und vorangetrieben. Die Verkabelung stellt einen erheblichen Aufwand und potenzielle Fehlerquelle dar. Daher sind autarke Sensorknoten, die ihren Leistungsbedarf mittels sogenanntem Energy Harvesting aus ihrer Umgebung decken und eine drahtlose Kommunikationstechnologie zur Datenübertragung verwenden, oft von Vorteil. Die Sensorknoten bilden ein drahtloses Sensornetzwerk, um die Daten effizient an den Zielort zu übertragen. Weiters sind autarke Sensorknoten in bestimmten Situationen notwendig, da dort eine kabelgebundene Messung nicht möglich ist. Für Anwendungen, bei welchen sich das Messobjekt oft ändert und daher die Sensoren neu angebracht werden müssen, verringert man durch die Verwendung von autarken Sensorknoten die Anzahl von Kabeln zu einer zentralen Messdatenerfassung. Die Zuordnung vom jeweiligen Sensorknoten zu einer bestimmten Messgröße ist jedoch aufwendig und muss meist händisch vorgenommen werden. Um dies zu vereinfachen präsentiert dieser Beitrag ein optisches Lokalisierungssystem für autarke Sensorknoten in industriellen IoT Anwendungen. Die Sensorknoten sind mit einer schaltbaren Lichtquelle ausgestattet, welche eine kodierte Blinksequenz erzeugen. Diese Blinksequenz wird von mindesten zwei Kameras erfasst und zur Identifizierung und gleichzeitigen Lokalisierung der Sensorknoten verwendet. Dadurch kann eine teilautonome Zuordnung eines Sensorknotens zu einer bestimmten Messgröße erfolgen.

Keywords: Optische Lokalisierung, autarke Sensorknoten, Sensornetzwerk, industrielles IoT.

Einleitung

Autarke Sensorknoten werden immer häufiger zur Erfassung industrieller Messgrößen verwendet. Einerseits wird diese Entwicklung durch die technologische Verfügbarkeit und Verbesserung von kostengünstigen und energieeffizienten Transceiver-Schaltungen und Erhöhung der Rechenleistung integrierter Schaltungen ermöglicht. Andererseits sind zur weiteren Optimierung technischer Prozesse zusätzliche Messdaten notwendig und werden daher von der Anwendung selbst gefordert. Um diese zusätzlichen Messdaten kosteneffektiv erfassen und meist an eine zentrale Einheit übertragen zu können, werden die Sensorknoten zu sogenannte drahtlose Sensornetzwerke („Wireless Sensor Networks“, WSN) zusammengefasst. Die Daten werden drahtlos unter Einhaltung eines Kommunikationsprotokolls zu einer Datensenke übertragen. Die Sensorknoten selbst erfassen die physikalische Messgröße und vorverarbeiten die Messdaten. Sensorknoten und Sensornetzwerke sind essenzielle Bestandteile des sogenannten industriellen Internet der Dinge („Industrial Internet-of-Things“, IIoT).

Durch die drahtlose Kommunikation wird eine erhebliche Zahl an Messleitungen eingespart. Um nun nicht auf kabelgebunden Energieversorgung angewiesen sein zu müssen, sind grundsätzlich zwei verschieden Prinzipien möglich [2]. Erstens kann die für die gesamte Betriebszeit des Sensorknotens nötige Energie bereits bei Installation in Form von Batterien mit eingebracht werden. Die maximale Betriebszeit ergibt sich im Wesentlich aus der durchschnittlichen Leistungsaufnahme des

Sensorknotens und der Selbstentladung der Batterien. Für bestimmte Anwendungsfälle kann dies die kostengünstigste Lösung darstellen, vor allem dann, wenn die Sensorknoten zu Wartungs- und Kalibrierungszwecken regelmäßig getauscht werden müssen.
Zweitens kann der Sensorknoten mittels der in der Umgebung verfügbaren Leistung versorgt werden. Dabei wird durch sogenanntes „Energy Harvesting“ diese Umgebungsleistung in elektrische Leistung umgewandelt. Dazu können zum Beispiel Thermogeneratoren oder Solarzellen eingesetzt werden. Die verfügbare Energie aus der Umgebung ist im Allgemeinen unbegrenzt, es kann jedoch meist nur eine geringe elektrische Leistung je nach Anwendungsort erzielt werden [3]. Weiters schwankt die umwandelbare Leistung mit der Zeit und ist im Detail nicht vorhersagbar [4]. Daher muss Energie in Superkondensatoren oder wiederaufladbaren Batterien zwischengespeichert werden, um einen kontinuierlichen Betrieb des Sensorknotens gewährleisten zu können.
Lokalisierungssysteme werden mit zunehmender Anzahl von Sensorknoten immer wichtiger. Eine automatisierte Erfassung der Position der Sensorknoten kann zur schnellen und zuverlässigeren Konfiguration von Sensornetzwerken verwendet werden. Vor allem bei statischen Anwendungen, bei welchen sich die Position der Sensoren bezogen auf das Messobjekts nicht ändert, ist dies sinnvoll. Optische Methoden haben für solche Anwendungen eine gute Genauigkeit. Kameras können die Sensorknoten mittels an ihnen angebrachten schaltbaren Lichtquellen erfassen. Wenn ein Sensor von mindestens zwei Kameras erfasst wird und die Position und Orientierung derer bekannt ist, kann daraus die Position des Sensorknotens berechnet werden. Bisher stellte der Energieverbrauch der schaltbaren Lichtquelle ein erhebliches Problem dar. Dieser Beitrag präsentiert ein optisches Lokalisierungssystem, welches einerseits eine hohe Genauigkeit im unteren Zentimeterbereich erreicht und andererseits den Energieverbrauch der schaltbaren Lichtquelle der Sensorknoten reduziert.
Der folgenden Abschnitte beschreiben das optische Lokalisierungssystem, das verwendete Kommunikationsprotokoll und die optimale Blinksequenz. Danach werden Messergebnisse und die erreichte Messgenauigkeit dargestellt. Schlussendlich fasst der letzte Abschnitt den Beitrag zusammen und gibt einen Ausblick über künftige Entwicklungen.

Optisches Lokalisierungssystem

Für die optische Lokalisierung der Sensorknoten in 3 Dimensionen werden wie in Abb. 1 dargestellt rund um den Messbereich Kameras (K) installiert und so platziert, dass sich jeder Sensorknoten (S) im Sichtbereich von mindestens zwei Kameras befindet. Um allenfalls Bereiche im Testaufbau abdecken zu können, die ansonsten von Teilen des Messaufbaus, Kabeln, etc. verdeckt wären, kann die Anzahl an Kameras gemäß den Erfordernissen im jeweiligen Szenario erhöht werden. Die Sensorknoten werden mit einer schaltbaren Lichtquelle, im Speziellen mit einer hellen Leuchtdiode (LED) ausgestattet, die in einem transparenten Teil der Sensorgehäuse so angebracht ist, dass sie aus möglichst vielen Raumrichtungen direkt sichtbar ist. Abb. 2 zeigt die aktuelle Ausgestaltung eines solchen Sensorknotens. Für die Datenübertragung sind die Kameras über Ethernet mit einer zentralen Steuerungs- und Signalverarbeitungs-Einheit verbunden, die Triggerung der Blinksequenzen in den Sensoren und der Aufzeichnung von Kamerabildern erfolgt über das im nachfolgenden Abschnitt beschriebene Kommunikationsprotokoll.

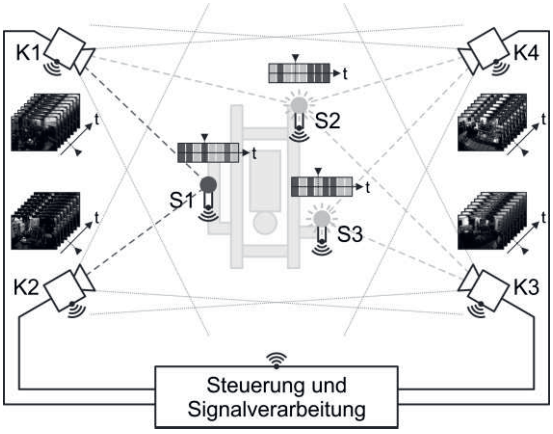
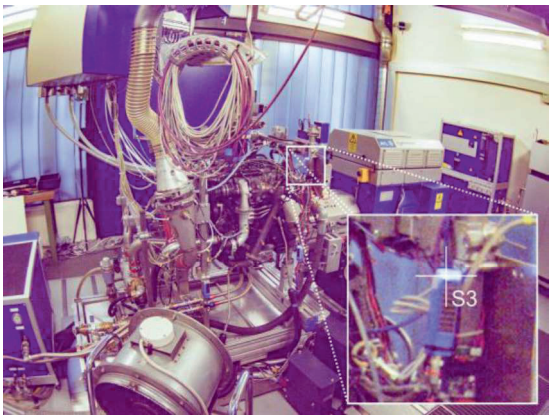


Abb. 1: Schematischer Aufbau der optischen Lokalisierung mit 4 Kameras (K1–K4) und 3 Sensoren (S1–S3) mit Darstellung der zugehörigen Blinksequenz.

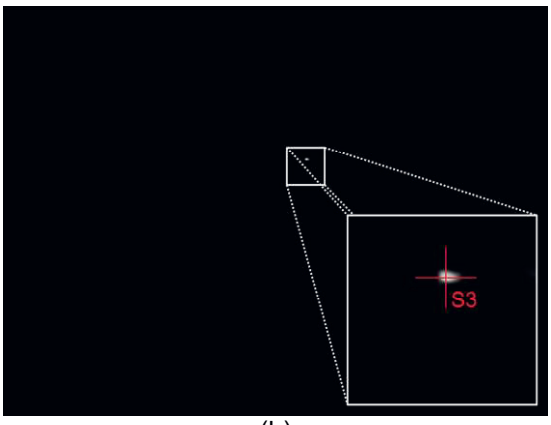


Abb. 2: Sensorknoten mit transparentem Gehäuseteil für optische Lokalisierung.

Zur Lokalisierung wird durch jede Kamera zunächst eine Sequenz von Bildern aufgezeichnet. Synchron zu den Aufnahmezeitpunkten wird durch Ein- und Ausschalten der LEDs in den Sensoren eine für jeden Knoten individuell codierte Blinksequenz durchlaufen, anhand derer die einzelnen Sensoren schließlich voneinander unterschieden werden können. Die Länge der Blinksequenz bestimmt den Umfang des Coderaums und damit die maximale Anzahl der lokalisierbaren Sensorknoten; gleichzeitig erhöht aber eine Verlängerung der Blinksequenz auch den Energieverbrauch für die Lokalisierung. Eine Analyse zur optimalen Blinksequenz ist in einem späteren Abschnitt gezeigt.
Während der Aufzeichnung sollte nach Möglichkeit ein statisches Szenario herrschen, sich in den Kamerabildern bis auf die unterschiedlichen Leuchtzustände der LEDs also möglichst wenig ändern. Das erleichtert die Detektion auch nur schwach leuchtender oder teilweise verdeckter LEDs in den Kamerabildern und verringert die Wahrscheinlichkeit von Fehldetektionen. In den typischen Szenarien, in denen diese Methode zum Einsatz kommen soll, ist dies in der Regel einfach sicherzustellen. Verbleibende unvermeidliche unerwünschte Änderungen in den Kamerabildern etwa durch Bildrauschen, Vibrationen, Beleuchtungsänderung etc. können durch robuste Schätzmethoden in der Signalverarbeitung kompensiert werden.



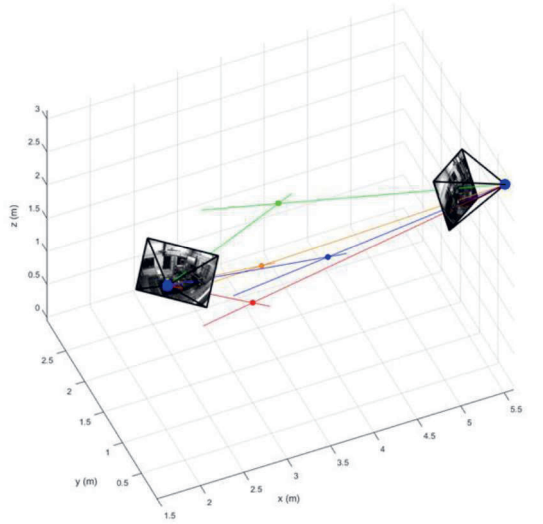
(a)



(b)

Abb. 3: Exemplarisches (a) Kamerabild und (b) zugehöriges Detektionsbild für die Detektion von Sensor S3.

In den von den Kameras aufgenommenen Bildsequenzen werden nun zunächst die einzelnen installierten Sensoren anhand der bekannten Blinksequenzen detektiert und in Bildkoordinaten jeder Kamera lokalisiert. Dabei sind zum einen die Sensorknoten voneinander zu unterscheiden, aber auch zu entscheiden, ob ein bestimmter Sensorknoten von einer Kameraposition aus überhaupt sichtbar ist. Hierzu wird aus den Einzelbildern jeder Kamera durch Addition bzw. Subtraktion gemäß dem Code jedes installierten Sensorknotens ein Detektionsbild errechnet, in dem anschließend das Maximum gesucht wird. Ein exemplarisches Kamerabild und ein zugehöriges Detektionsbild für Sensor S3 sind in Abb. 3 dargestellt.





(b)

Abb. 4: Kamerabild des zur Kalibrierung verwendeten Schachbrettmusters (a) vor und (b) nach Entzerrung.

Nach der Detektion stehen für jede Kamera die Koordinaten aller sichtbaren Sensoren in Kamerakoordinaten zur Verfügung. Für die Umrechnung in Raumkoordinaten ist eine Kalibration der Kameras erforderlich, um unvermeidliche Verzerrungen der Objektive zu entfernen. Eine gängige Methode hierfür ist die Aufnahme einer mit einem Schachbrettmuster versehenen Tafel aus verschiedenen Blickwinkeln, woraus die Parameter der Kameraoptik berechnet und die aufgenommenen Bilder entzerrt werden können, wie exemplarisch in Abb. 4 dargestellt. Sofern Position und Orientierung der Kameras im Raum bekannt sind, lässt sich das Bild jeder detektierten Leuchtdiode nun in einen Sichtstrahl im Raum umrechnen. Die Sichtstrahlen verlassen die Kameras unter durch die zuvor ermittelten Bildkoordinaten der Leuchtdioden gegebenen Winkel, weshalb es sich hierbei um ein „Angle-of-Arrival"-Messverfahren handelt. Wie in Abb. 5 schematisch dargestellt, schneiden sich die von verschiedenen Kameras ausgehenden Sichtstrahlen im Idealfall an der 3D-Position des entsprechenden Sensors. In der Realität schneiden sich die Strahlen wegen unvermeidlicher

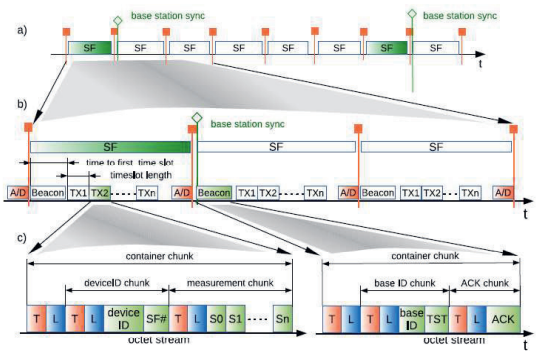


Abb. 6: a) Superframe-Taktstruktur, b) Superframe-Aufbau und c) Datenstruktur des EPhESOS Kommunikationsprotokolls.

Messfehler, nicht vollständig kompensierter optischer Verzerrungen sowie Unsicherheiten in Kamerapositionen und -orientierungen nicht in einem Punkt, sondern verlaufen windschief. Als Schätzergebnis wird deshalb derjenige Punkt ermittelt, der den quadratischen Abstand zu allen Sichtstrahlen minimiert.

Kommunikationsprotokoll

Im Folgenden wird das verwendete Kommunikationsprotokoll EPhESOS beschrieben. Dieses Protokoll ist speziell für drahtlose Sensornetzwerke entwickelt worden, die eine möglichst geringe durchschnittliche Leistungsaufnahme sowie eine geringe Latenzzeit benötigen. Der durchschnittliche Leistungsverbrauch wird durch die Verwendung eines TDMA-artigen Protokolls minimiert. Dazu müssen die Sendezyklen exakt aufeinander abgestimmt und die Sensorknoten untereinander genau synchronisiert sein. Dadurch kann die Sende- und Empfangszeiten der einzelnen Sensorknoten minimiert werden.
Die Frame-Struktur des EPhESOS-Kommunikationsprotokolls ist in Abb. 6 gezeigt. Es basiert auf einem TDMA MAC Verfahren, welches Zeitschlitze innerhalb eines Superframes (SF) vorgibt (Abb. 6a). Jeder SF besteht aus einem Zeitfenster für einen Beacon zur Downlink-(DL)-Kommunikation und einem Zeitfenster für jeden Sensorknoten zur Uplink-(UL)-Kommunikation (Abb. 6b). Je nach Anwendung können diese Zeitschlitze individual konfiguriert werden. Der Beacon wird zur Übertragung von erstens Synchronisationsnachrichten, zweitens Paketbestätigungsinformationen und drittens Netzwerkbefehle verwendet. Die Datenstruktur sowohl der DL- als auch UL-Pakete werden durch das EPhESOS Datengram strukturiert (Abb. 6c). Eine Struktureinheit besteht aus Type (T), Länge (L) und Wert (V) und wird TLV-Chunk bezeichnet.

Diese Struktureinheiten können ineinander verschachtelt sein, um Daten noch weiter zu strukturieren. In der Abb. 6c ist unter anderem der Aufbau eines Beacon-Chunks dargestellt, der die Nummer der aussendenden Basisstation (baseID) und einen Zeitstempel (TST) zur Synchronisation enthält.
Das Kommunikationsprotokoll unterstützt zwei verschiedene Übertragungsmodi, welche im Folgenden kurz erläutert werden: Erstens, der sporadischen Übertragungsmodus EPhESOS-S ist für die Netzwerkinbetriebnahme, die Integration neuer Knoten, die Initialsynchronisation, die Knotenkonfiguration und die Wartung zuständig. Bereits in diesem Modus wird die SF-Taktstruktur eingehalten, obwohl noch nicht im jedem SF ein Beacon zu den Knoten gesendet wird. Um einen neuen Knoten in diese Zeitstruktur zu integrieren, sendet dieser zu einem beliebigen Zeitpunkt eine Anfrage an die Basisstation und bekommt eine Antwort mit der Synchronisationsinformation und der Zeitschlitzinformation. Danach wird die SF-Taktstruktur auch von diesem Knoten eingehalten. Dadurch kann der drahtlose Sensorknoten, wenn keine Daten übertragen werden sollen, im energiesparenden Schlafmodus verbleiben [5].
Zweitens, der kontinuierlichen Übertragungsmodus EPhESOS-C kann nur aus EPhESOS-S erreicht werden, da die Synchronisationsinformation und der Zeitschlitzinformation bereits allen Knoten bekannt sein müssen. In diesem Modus sendet die Basisstation zu Beginn jedes SF einen Beacon, der unter anderem die Empfangsbestätigung von davor gesendeten UL-Paketen enthält. Ein Knoten muss allerdings nicht in jedem SF Daten senden, sondern kann auch Daten aggregieren und diese energiesparend in einem UL-Paket verpacken. Wird ein Paketverlust festgestellt, wird die fehlerhaft übertragenen Daten im folgenden SF erneut gesendet. Die Sensorknoten bleiben zwischen Empfangs- und Sendezeiten im Ruhezustand. Um einen reibungslosen Betrieb zu gewährleisten, ist eine präzise Synchronisation erforderlich, da sonst die Sendezeitpunkte driften würden. Zur Synchronisation im EPhESOS Protokoll wird eine Kombination aus einer Frequenzschätzmethode [6] und einem ΔΣ-Ansatz verwendet [7].
Die LEDs der Sensorknoten werden synchron zur SF-Taktstruktur je nach Codevorgabe aktiviert und deaktiviert. Dazu muss sich das Sensornetzwerk im EPhESOS-C-Modus befinden. Zusätzlich werden spezielle Triggerknoten, die an den Kameras angebracht sind, zur Auslösung jener verwendet. Die Codesequenz für diese Triggerknoten ist immer ident und besteht nur aus Aktivierungen, damit jede Kamera in jedem Blinkintervall auslöst.

Tab. 1: Parameter zu Berechnung des Energieverbrauchs während eines Lokalisierungsvorganges.

Parameter	Wert
Leistungsaufnahme LED	42 mW
Durchschn. Leistungsaufnahme Comm. & Sync.	400 µW
Leistungsaufnahme Ruhe	30 µW
Zeit für Comm. & Sync. pro Sekunde	20 ms
Zeit LED aktiv (1 Blink)	100 ms
Zeit zwischen 2 Blinks	100 ms

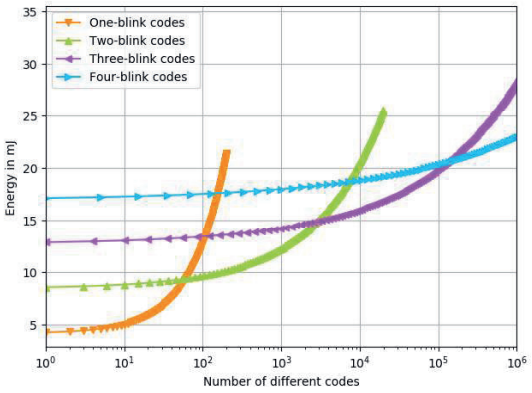


Abb. 7: Energieverbrauch eines Lokalisierungsvorganges pro Sensorknoten über die max. Anzahl an Sensorknoten für ein bis vier LED Aktivierungen.

Optimale Blinksequenz

Dieser Abschnitt zeigt die Anforderung an die optimale Blinksequenz. Als maßgebliches Optimierungskriterium für drahtlose Sensorknoten wurde der Gesamtenergieverbrauch eines Sensorknotens während eines Lokalisierungsvorganges ausgewählt. Da ein Lokalisierungsvorgang aufgrund der statischen Sensorknotenanordnung nur selten durchgeführt werden muss, ist die Gesamtdauer nicht entscheidend und wird hier nicht berücksichtigt. Dieser Gesamtenergieverbrauch setzt sich aus dem Energieverbrauch der LED zur Signalisierung der Blinksequenz, dem Energieverbrauch zur laufenden Kommunikation und Synchronisation und dem Energieverbrauch in den Ruhezeiten zusammen. Um den Energieverbrauch berechnen zu können wurden folgende Werte herangezogen (siehe Tab. 1).
Da die LED einen sehr hohen Leistungsverbrauch im Vergleich zu den anderen Komponenten aufweist, ist es sinnvoll diese nur so wenig wie möglich einzuschalten. Im einfachs-

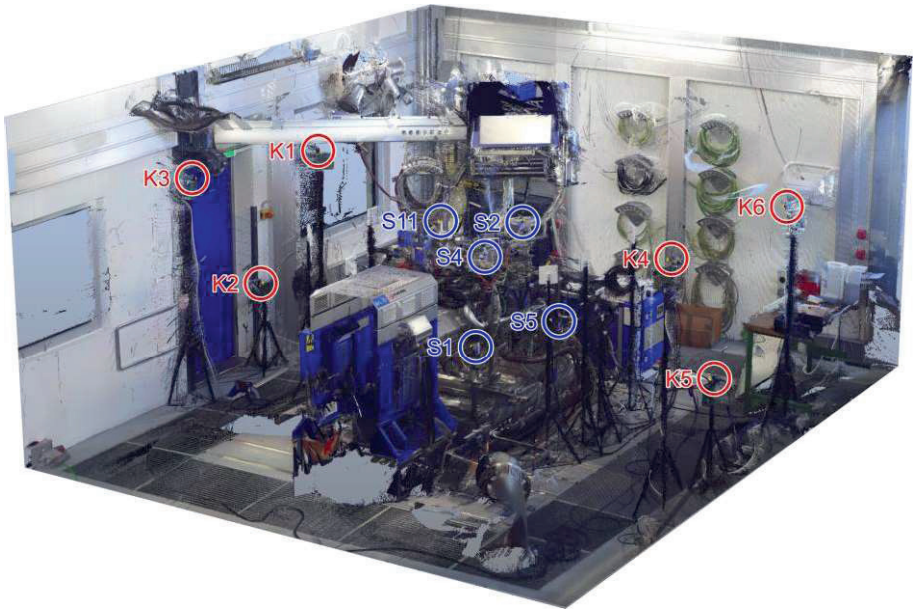


Abb. 8: Ansicht des Messaufbaus aus 3D-Laserscan mit eingezeichneten Kameras (K1–K6) und einigen sichtbaren Sensoren (S).

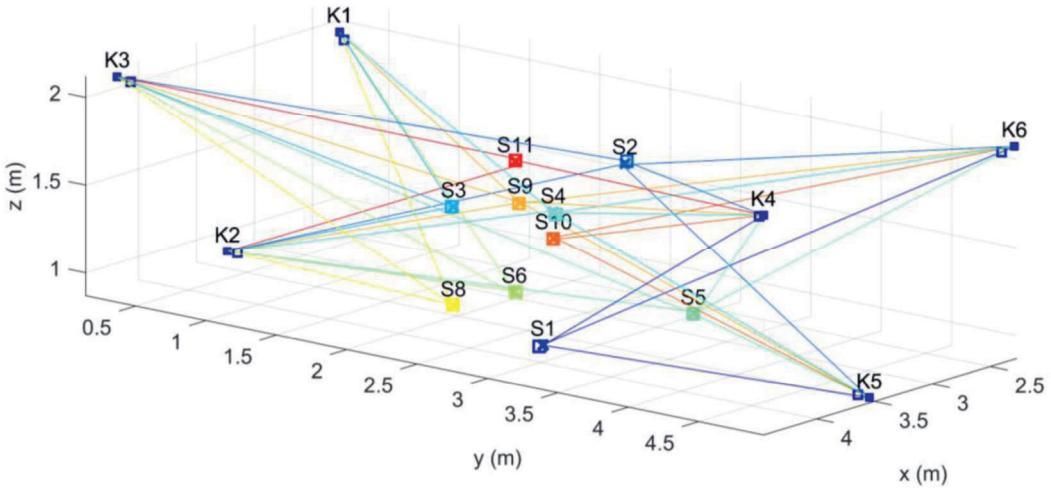


Abb. 9: Dreidimensionale Darstellung von Kamera- (K1–K6) und Sensorpositionen (S). Referenzpositionen aus dem 3D-Laserscan sind als Quadrate ohne Füllung dargestellt, geschätzte Positionen als Kreuze. Für von einer bestimmten Kamera aus sichtbare LEDs sind Sichtstrahlen in der Farbe des Sensors dargestellt.

ten Fall enthält jede Blinksequenz (oder auch Code genannt) nur eine einzelne LED-Aktivierung. Dadurch steigt natürlich die Länge der Codes linear mit der Anzahl der teilnehmenden Sensorknoten. Ab einer gewissen Anzahl wird jedoch der Energieverbrauch der LED geringer im Vergleich zum restlichen Energieverbrauch des Knotens. Daher ist es ab einer gewissen Anzahl von teilnehmenden Sensorknoten energetisch besser, pro Code zwei LEDs zu aktivieren. Dieses verhalten ist in Abb. 7 dargestellt. Sie zeigt den Gesamtenergieverbrauch eines Lokalisierungsvorganges eines Sensorknotens für ein bis vier LED Aktivierungen pro Code in Abhängigkeit von der maximalen Sensorknotenanzahl.

Messergebnisse und Genauigkeit

Die Lokalisierung wurde in einem Motorprüfstand unter realen Bedingungen getestet. In Abb. 8 ist ein 3D-Laserscan der Testumgebung dargestellt, der auch zur Ermittlung der Referenzpositionen von Sensoren und Kameras verwendet wurde. Im Zentrum des Testbereichs befindet sich ein an einem Metallrahmen montierter Motorblock mit den zugehörigen Zu- und Ableitungen von Kraftstoff, Schmiermitteln, Abgasen etc. An diesem Testaufbau wurden

10 Sensoren (S1–S11, ohne S7) an charakteristischen Positionen montiert. Auf beiden Seiten des Motorblocks wurden jeweils 3 Kameras (K1–K6) in zwei verschiedenen Höhen platziert, sodass jeder Sensor von mindestens 3 Kameras aus sichtbar war. Unter Verwendung einer Codelänge von 16 wurde eine Serie von Bildern aufgezeichnet und daraus die Sensorpositionen in 3D berechnet. Die Positionen der Kameras wurden hierbei vorgegeben, die Orientierungen wurden ebenso wie die Positionen der Sensoren durch das beschriebene Verfahren geschätzt. In Abb. 9 sind die aus dem Laserscan ermittelten Referenzpositionen (Quadrate), sowie die Sichtstrahlen der Kameras zu den sichtbaren Sensoren, sowie die daraus geschätzten Sensorpositionen (Kreuze) dargestellt. Die räumlichen Distanzen der Schätzungen zu den Referenzpositionen sind in Abb. 10 dargestellt, die Distanzen zwischen den Schätzungen und den einzelnen Sichtstrahlen, die sich in der realen Messsituation nicht perfekt in einem Punkt schneiden in Abb. 11. Man kann erkennen, dass die erzielten Ergebnisse zueinander konsistent sind und die erzielbaren Fehler in einem Bereich von ca. 10-20 mm liegen, der für die Anwendung jedenfalls ausreichend ist.

Zusammenfassung und Ausblick

Dieser Beitrag präsentiert ein optisches Lokalisierungssystem für autarke Sensorknoten in industriellen IoT Anwendungen. Es beschreibt den Anwendungsfall und den Messaufbau. Zusätzlich wird das verwendete Low-Power Kommunikationsprotokoll erklärt und auf die optimale Blinksequenz eingegangen. Dabei zeigt sich, dass je nach Anzahl an Sensorknoten, eine unterschiedliche Anzahl an LED Aktivierungen energieoptimal ist. Weiters sind die Messergebnisse zur Lokalisierung von Sensorknoten an einem Motorprüfstand gezeigt. Dabei ergibt sich ein Fehler der ermittelten Position der Sensorknoten von ca. 5-25 mm.
Zukünftige Entwicklungen zielen auf eine Verbesserung der Lokalisationsgenauigkeit sowie einer automatisierten Zuweisung der optimalen Blinksequenzen zu den jeweiligen Sensorknoten ab.

Danksagung

Diese Arbeit wurde vom LCM - K2 Center im Rahmen des österreichischen COMET-K2-Programms und durch die Österreichische Forschungsförderungsgesellschaft mbH (FFG) unter der Antragsnummer 853456 (FASAN: Flexible Autonome Sensorik in industriellen ANwendungen) unterstützt.

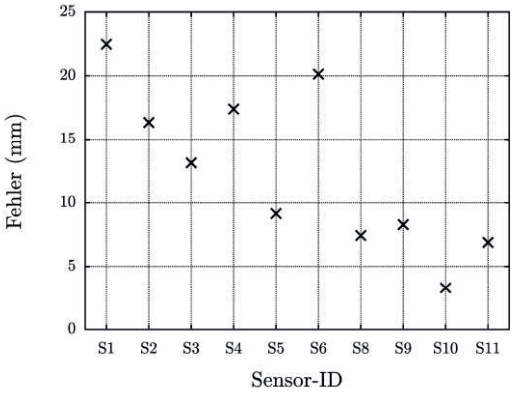


Abb. 10: Lokalisierungsfehler der 10 Sensoren als Abstand der Schätzung von der mit dem Laserscanner vermessenen Referenzposition.

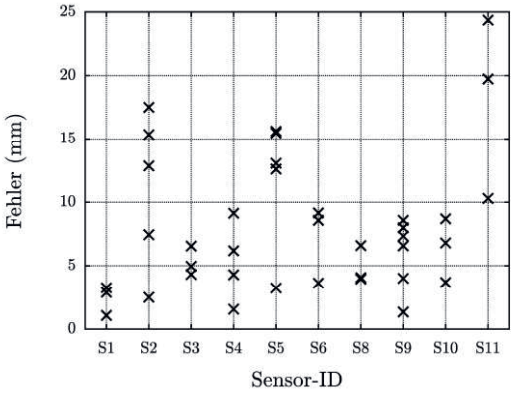


Abb. 11: Lokalisierungsfehler der 10 Sensoren als Abstand der Schätzung von den Sichtstrahlen aller Kameras mit Sichtverbindung zum jeweiligen Sensor.

Literaturnachweis

[1] M. Bal, “Industrial applications of collaborative wireless sensor networks: A survey,” in 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE), June 2014, 1463-1468; doi: 10.1109/ISIE.2014.6864830

[2] L. B. Hörmann, T. Buchegger, and C. Steger, “Optimizing the Energy Supply of Autonomous Wireless Sensor Nodes,” in Microelectronic Systems Symposium (MESS 2014), 2014, 1-6; doi: 10.1109/MESS.2014.7010257

[3] L. B. Hörmann, P. M. Glatz, K. B. Hein, M. Steinberger, C. Steger, and R. Weiss, “Towards an On-Site Characterization of Energy Harvesting Devices for Wireless Sensor Networks,” in IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops 2012), 2012, 415-418; doi: 10.1109/PerComW.2012.6197523

[4] A. Janek, C. Trummer, C. Steger, R. Weiss, J. Preishuber-Pfluegl, and M. Pistauer, “Simulation based Verification of Energy Storage Ar-

chitectures for Higher Class Tags Supported by Energy Harvesting Devices," Microprocessors and Microsystems, vol. 32, no. 5-6, 2008, 330-339; doi: 10.1016/j.micpro.2008.03.009

[5] H.-P. Bernhard, A. Springer, A. Berger, and P. Priller,"Life cycle of wireless sensor nodes in industrial environments," in 13th IEEE Int. Workshop Factory Commun. Sys., Trondheim, Norway, May 2017; doi: 10.1109/WFCS.2017.7991943

[6] H. P. Bernhard and A. Springer, "Linear complex iterative frequency estimation of sparse and pulse and point processes," in Europ. Sig. Process. Conf. (EUSIPCO), Kos, Greece, Aug. 2017; doi: 10.23919/EUSIPCO.2017.8081380

[7] H.-P. Bernhard, A. Berger, and A. Springer, "Analysis of delta-sigma-synchronization in wireless sensor nodes," in IEEE Int. Conf. Industrial Informatics (INDIN), Cambridge, UK, Jul. 2015, pp. 914–918; doi: 10.1109/INDIN.2015.7281857

DOI 10.5162/sensoren2019/1.2.1

Radialsymmetrischer Scherkraftaufnehmer als Transfernormal

Thomas Kleckers
Hottinger Baldwin Messtechnik GmbH
Im Tiefen See 45
64293 Darmstadt
+49 6151 803 570
Thomas.kleckers@hbm.com

Radialsymmetrische Scherkraftaufnehmer sind in der experimentellen Mechanik ein bewährtes und häufig genutztes Messmittel. Neueste Entwicklungen zeigen, dass das Potential dieser Technologie auch den Einsatz als präzises Transfernormal für Referenzkraftmessungen erlaubt- mit hoher Genauigkeit zu vertretbaren Kosten.

1. Funktionsweise von radialsymmetrischen Scherkraftaufnehmern

Das Funktionsprinzip von Dehnungsmessstreifen – basierte Kraftaufnehmer ist seit der ersten industriellen Produktion 1956 bewährt und wurde immer weiter entwickelt. Ein metalliischer Federkörper wird in den Kraftfluss eingebaut. Die Einwirkenden Kräfte erzeugen mechanische Spannungen, die iherseits nach dem Hookschen Gesetz Dehnungen nach sich ziehen. In zuvor bestimmten Zonen werden Dehnungsmessstreifen aufgeklebt, die diese Dehnungen in Widerstandsänderungen wandeln. Durch Verschaltung zur so genannten Wheatston´schen Brückenschaltung und externer Spannungsversorgung entsteht ein der eingeleiteten Kraft proportionales Spannungssignal.

Die Besonderheit der symmetrischen Scherkraftaufnehmern liegt in der Kombination der folgenden konstruktiven Details:

- Die eingesetzten Dehnungsmesstreifen basieren auf Chrom – Nickel – Legierungen, die einen höheren K – Faktor (eine höhere Empfindlichkeit) als die meist genutzten Konstantan – DMS auf. Die Empfindlichkeit ist um etwa 10 – 15 % höher.
- Das Dehnungsfeld im Bereich der DMS – Installation ist ohne Gradienten oder Dehnungsmaxima von einem gleichmäßigen Verlauf. Damit ist es möglich, den gesamten DMS mit dem technisch größten Dehnungsniveau zu belasten, was ein maximale Widerstandsänderung nach sich zieht.
- Die Konstruktion is so ausgelegt, dass die höchsten mechanischen Spannungen im Bereich der DMS vorzufinden sind, alle anderen Bereiche des Kraftaufnehmers sind nur geringen mechanischen Spannungen ausgesetzt.

Aus den oben genannten Konsstruktionsdetails ergeben sich eine Reihe von Vorteilen, die den Erfolg dieses Messprinzipes in Experiment und am Prüfstand erklärt:

- Bereits bei geringer mechanischer Belastung der DMS ein steht ein hohes Ausgangssignal zur Verfügung. Wird der Sensor auf die in der Kraftmesstechnik üblichen 2 mV/V kalibriert, ergeben sich sehr hohe mechanische Reserven. (Überlasttoleranz, Wechsellastfähigkeit)
- Umbekehrt ermöglicht dieses Messprinzip die Kalibrierung bei höheren Kräften, so dass das Ausgangssignal bis zu 4 mV/V beträgt. Dies entspricht dem doppelten Wert, den andere Prinzipien bei vergleichbarer Betriebsfestigkeit zur Verfügung stellen. In diesem Fall verkleinert sich die Tolernaz gegenüber Überlasten und die Wechsellastfähigkeit verglichen mit einer 2 mV/V Kalibrierung, jedoch erhöht sich die Genauigkeit beträchtlich.

Es zeigt sich, dass vor allem die dynamischen Anforderungen bei Betriebsfestigkeitsprüfungen und in Prüfständen von radialsymmetrische Scherkraftsensoren seht gut erfüllt wird.

2. Einsatz als Transfernormal

In Experiment und am Prüfstand steigen die Anforderungen an die Genauigkeit der eingesetzten Kraftsensoren und der angeschlossenen Elektronik stetig. Dies hat wirtschaftliche Gründe, da höhere Genauigkeit zu erweiterten Messbereichen und damit zu größerem Einsatzbereich des Prüfstandes führt.

Bei Transfernomalen, also Kraftaufnehmer die dazu dienen, Kalibriermaschinen zu vergleichen und existierende Maschinen über Vergleichsmessungen an eine höheres Normal anzuschließen, kommen zahlreiche Forderungen zu den Anforderungen aus Experiment und Prüfstand hinzu:

- Die Widerholpräzision in verschiedenen Einbaustellungen ist ein Merkmal der maßgeblichen ISO376. Hiermit wird garantiert dass der Einbau zur Kalibrierung ohne großen Einfluss auf das Ergebnis ist.
- Die Grenzen für die Umkehrspanne (Hysterese) sin deutlich enger gesetzt als für die industrielle Anwendung.
- Die Kennlinie des Sensors soll durch eine kubische Approximation sehr gut beschrieben werden können- in industriellen Anwendungen ist die Abweichung von einer Geraden (Linearitätsabweichung) entscheidend

Radialsymmetrische Scherkraftaufnehmer sind Sensoren, die Schraubverbindungen aufweisen, um den eigentlichen Federkörper mit dem unbedingt erforderlichen Fußadapter zu verbinden. Diese Flanschverschraubung ist ein Unterschied zu konventionellen Transfernormalen (Referenzkraftaufnehmer), die monolithisch konstruiert sind. Schraubverbindungen im Kraftfluss limitieren die Genauigkeit, da kleinste Verschiebungen der verbundenen Konstruktionselemente zueinander geänderte Dehnungszustände im Bereich der DMS verursachen- im Endergebnis sind diese als Hysterese nachweisbar.

3. Neues Transfernormal

Erste Referenzkraftaufnehmer nach dem Scherprinzip wurden in den frühen 2000er Jahren zur Marktreife gebracht.
Diese Kraftmessdosen sind für Zug- und Druckbelastungen ausgelegt und erreichen die Klasse 0,5 nach ISO376, das ist die zweitbeste Klassengenauigkeit, die dieser international gültige Standard vorsieht. Die Anfälligkeit des Prinzips gegen Einflüsse der Verschraubung verhinderte eine bessere Klasseneinstufung. Dies gilt bei Zugkräften, wenn auf die Verbindung die zu messende Kraft und die Vorspannung durch die Verschraubung selbst einwirken. Bei Messung von Druckkräften ist dies nicht von Bedeutung.

Letzte Entwicklung ist ein Druckkraftaufnehmer, der die höchste Genauigkeitsklasse 00 nach ISO376 erreicht. Um dieses Ziel konstruktiv zu erreichen, waren verschiedenen Maßnahmen notwendig:

- Die Biegesteifigkeit des Sensors musste zurückgenommen werden, um den Einfluss der Eigenschaften von Kalibriermaschinen auf das Ergebnis zu verringern
- Beschränkung auf Druckkräfte, um den Einfluss der Verschraubung zu minimieren.

Mit einer Neukonstruktion der Lasteinleitung konnten die Ziele erreicht werden Hierbei ist eine Nut in den Fußadapter eingebracht, so dass die Biegemomentsteifigkeit gezielt zurückgenommen wird. Eventuell während der Kalibrierung auftretende Momente werden somit nicht mehr in den Federkörper geleitet und sind ohne Einfluss auf das Ergebnis. Durch verbesserte Ebenheit konnte der Kraftschluss zwischen Fußadapter und Messkörper verbessert werden.



Transfernormal C15. Deutlich zu erkennen die Ringförmige Nut im Fußadapter, die Biegemomentsteifigkeit dieses Bauteils verringert. Der obere Messkörper ist durch Schrauben an dem Fußadapter montiert.

4. References

[1] Thomas Kleckers, Markus Graef, “New , IMEKO 22nd TC3, 12th TC5 and 3rd TC22 International Conferences; 3 to 5 February, 2014, Cape Town, Republic of South Africa
[2] Karl Hoffmann, “An introduction to Measurement using strain gauges”, Hottinger Baldwin Messtechnik GmbH, 1989
[3] Thomas Kleckers, “Radial Shear Beams used as high precision transfer TRANSDUCERS”, IMEKO 2010 TC3, TC5 and TC22 Conferences Metrology in Modern Context, November 22-25, 2010, Pattaya, Chonburi, Thailand

DOI 10.5162/sensoren2019/1.2.2

Design study for a multi-component transducer for wind turbine test benches

Jonas Gnauert[1], Georg Jacobs[1], Stefan Kock[1], Dennis Bosse[1], Benjamin Janik[1]
[1]Chair for Wind Power Drives, Campus-Boulevard 61, 52074 Aachen, Deutschland

Abstract

This paper covers the design study of a multi-component transducer (MCT) for wind turbine test benches. The MCT will cover the characteristics of wind turbines in the performance class up to 6 MW. The motivation to develop such a MCT is to provide a satisfying accuracy in measurement of loads and moments in all six degrees of freedom, to reduce the uncertainty in the traceability of the drive train behavior due to the applied loads. Therefore, the estimation of the measurement uncertainty is significant to evaluate the design of the MCT. At the beginning, the design process of the MCT is shortly introduced. Afterwards, the strain gauges based transducer design is investigated under operational conditions (e.g. torque, and multiaxial loads) by using FE simulations to determine the crosstalk effects. Finally, the measurement uncertainties of all measurements are determined based on these FE simulations according to the type B evaluation of the GUM [1], including metrological aspects (e. g. linearity deviation and hysteresis) and the crosstalk. It can be shown that the MCT provides a high potential to significantly improve the measurement uncertainty of the applied wind loads on a wind turbine test bench.

Keywords: mechanical measure, multi-component transducer, wind turbine test bench, measurement uncertainty, design study

Introduction

Wind industry is a major pillar of climate-friendly energy supply in Germany. The wind industry supplies 17.2% of the energy production, which sums up to 49.2% of the renewable resources (Germany 2018) [2]. To ensure that wind industry keeps expanding and becomes competitive to fossil and nuclear sources wind turbines must be further optimized concerning availability and reliability [3]. Wind turbine test benches (TB) are crucial to face these aims. A conventional TB comprises a Prime Mover for torque generation on the Main Shaft, Load Application System for the application of wind forces and bending moments, Device under Test and Electrical Power Supply Equipment, e.g. converter, transformer, and filters, see Fig. 1. The existing TB [4], compared to default field tests, provide the opportunity to investigate the behavior of the entire drive train and control system under free adjustable, controlled and reproducible operating conditions. Thereby it is possible to apply simultaneously multiaxial dynamic wind loads as well as highly fluctuating electrical grid loads on the wind turbine, to simulate a realistic field load situation. With the 4 MW RWTH nacelle test bench, it is possible to apply 3.4 MN·m torque under maximum bending moments of about 7 MN·m, 3.3 MN radial forces and 4.0 MN thrust [5].

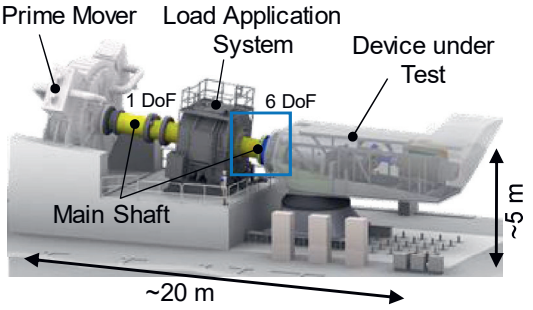


Fig. 1 4 MW RWTH Wind Turbine Test Bench [5] – Main components

The exact knowledge of the applied wind loads on a TB is thereby necessary for the [6]:

- Determination of the wind turbine efficiency,
- validation of simulation models as well as
- characterization of the drive train behavior under critical operation modes.

Today the applied wind loads in WT can be determined just imprecise, for example: the wind loads are calculated by the measurement of the hydraulic pressure in the hydraulic cylinders of the Load Application System. This method neglects for instance the speed-

dependent friction inside the cylinders. A multi-component transducer (MCT), which is directly implemented in the power transmission between the Load Application System and the Device under Test, would solve this issue. Therefore, a MCT, which is capable of a precise measuring of all six degrees of freedom simultaneously (DoF) under rotation, is required to improve the traceability of the applied wind loads. Until now, there are no commercial MCT for this dynamic and high load application available. This paper intends to introduce a design study of such a MCT for TB.

Design Study

A MCT between the load application system and the device under test on a TB shows a high potential to reduce the uncertainty in the traceability of the drive train behavior due to the applied loads. For this reason, a design study is carried out to identify a suitable transducer design. The design study is executed according to the design methodology of the VDI 2221 [7]. The design methodology divides the design process into four different steps:

- Classification of the Task,
- Conceptual Design,
- Embodiment Design and
- Detail Design.

This design study is focusing on the first three steps (see Fig. 2). First, the requirements for a MCT are identified and a various conceptual designs are developed. Based on a multi-criteria decision analysis the most suitable design concept is selected and a strain gauges based transducer is designed.

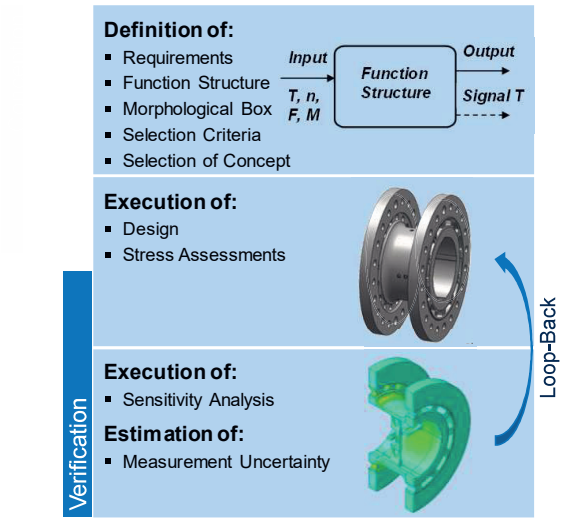


Fig. 2 Approach of the Design Study of the MCT

For this design a sensitivity analysis with Finite Element simulations are executed to improve the geometry concerning the load capacity, measurement range and to reduce crosstalk between the measuring points under operating loads (e.g. torque, and multiaxial loads). Finally, the measurement uncertainty for each measurement is estimated, taking the crosstalk and the metrology characteristics for strain gauges like e.g. linearity, hysteresis and repeatability into account.

Classification of the Task

For the design study, the characteristics of wind turbines in the performance class up to 6 MW are considered for the requirements. By consideration of this performance class, the following nominal measurement ranges arise:

- Axial force 1,5 MN
- Lateral force 2 MN
- Bending moments 5 MN·m
- Torque 5 MN·m

To achieve a precise measuring with a high sensitivity for all DoF, an individually design of the measurement ranges are necessary. Besides the individually design of the measurement ranges the crosstalk between the measuring points must be minimized. Additionally, the MCT must fulfill also these general requirements, which are:

- Each measurement uncertainty: < 1% referred to measured value,
- Operation conditions e.g. maximum rotational speed of 30 rpm,
- Ambient conditions: maximum temperature of 45 °C
- Dimensional conditions: max diameter of 4,5 m and max length of 2 m

Based on these requirements design concepts were identified in the next design phase.

Conceptual Design

In the Conceptual Design phase, a function structure of the MCT is developed. The function structure identifies all the necessary function, which must be fulfilled to reach the requirements. Starting from the main function “Measurement of all six DoF”, the structure is further divided into the main sub-functions of the measurement of each DoF. Every main sub-function is furthermore divided into the sub-function of:

- Separation of the measure
- Conversion of the measure
- Creating of the measuring signal
- Compensation of disturbance variables
- Transfer of measuring signal

For each sub-function various principles were developed, which combined are capable of solving the main function “Measurement of all six DoF”. With these design principles, different

design concepts were generated and with the help of a multi-criteria decision analysis (weighted sum model) compared. The most suitable design concept is selected (see Fig. 4) for further investigations.

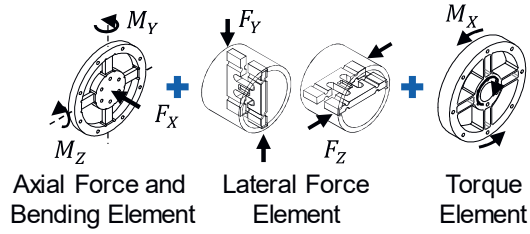


Fig. 3 Selected design concept for the MCT

Embodiment Design

Starting from the selected design concept the Embodiment Design of the MCT is designed. The Embodiment Design of the MCT can be seen in Fig. 5. The measuring body of the MCT consists of three main elements:

- Lateral Force Elements,
- Torque Element as well as the
- Axial Force and Bending element.

Strain gauges are installed on the elements to measure the strain resulting from the input loads. The strain gauges are connected in Wheatstone bridges. The separation of the single sub-functions in individual elements ensure the possibility to design the measurement ranges independent. This allows an optimization of the measuring points concerning to the sensitivity and crosstalk.

To achieve these optimizations a FE model of the MCT is build up. The strain gauges are modelled with the *Spring method* according to [8] and the Wheatstone bridges are considered afterwards with an analytical approach. With the help of this FE model, the geometry was optimized concerning the load capacity, measurement range and crosstalk reduction between the measuring points under operating loads: axial load, lateral load, bending moment and torque (see Fig. 6).

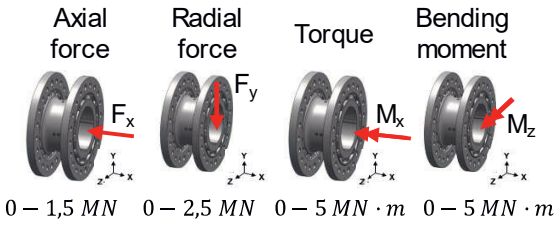


Fig. 5 Operational conditions

Therefore, the operating loads were applied on the MCT and the signal of each Wheatstone bridge was calculated depending on the simulated strains and the analytic behind each Wheatstone bridge. By this, it is possible to determine the nominal output signals depending on each nominal load and the amount of crosstalk from another load on an output signal. To reduce the crosstalk an iteration process (see Fig. 7) was used, which made it possible for example to reduce the crosstalk of bending moment to the lateral force element from 2.68% to 0.61%.

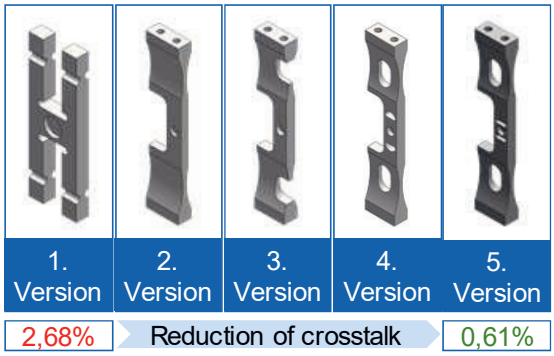


Fig. 6 Reduction of Crosstalk thru an iteration process (Loop Back)

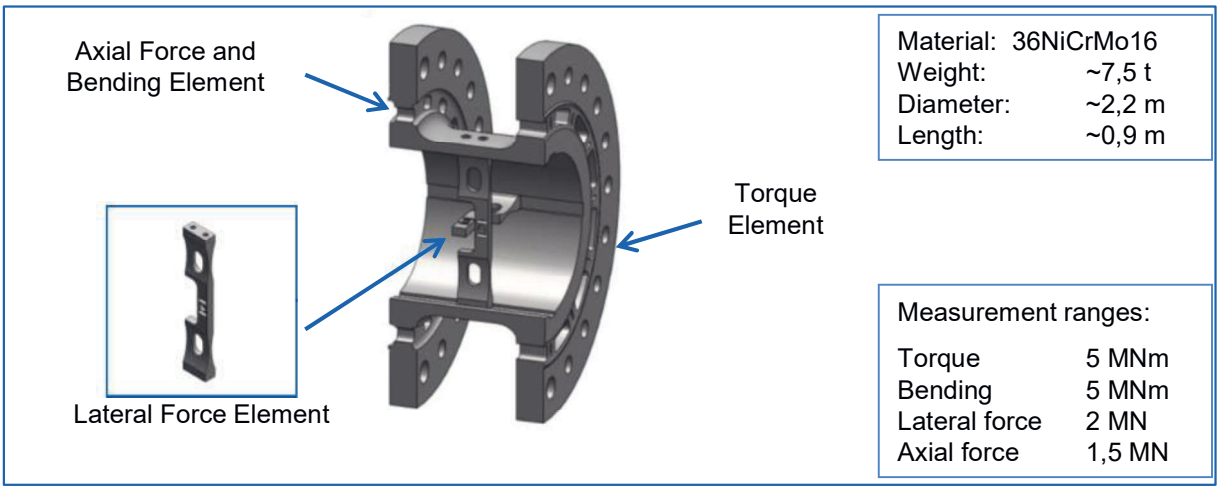


Fig. 4 Developed design of a MCT for TB

The remaining crosstalk in the MCT is summarized in Tab. 1. By knowing the crosstalk and the metrology characteristics for strain gauges like e.g. linearity, hysteresis and repeatability it is possible to estimate the measurement uncertainty of each measurement.

Tab. 1: Crosstalk of each nominal load $F_{nom,j}$ on the nominal output signal S_j of each measurement and the combined crosstalk d_{talk}

	S_{Fx} [%]	S_{Fy} [%]	S_{Mx} [%]	S_{Mz} [%]
F_x	-	0,052	0,000	0,002
F_y	0,004	-	0,001	0,447
M_x	0,016	0,608	0,002	-
M_z	0,003	0,003	-	0,011
d_{talk}	0,023	0,663	0,003	0,460

Estimation of the measurement uncertainty

The estimation of the measurement uncertainty for the MCT is performed according to the „Guide to the Expression of Uncertainty in Measurement“ (GUM) [1]. For the estimated measurement uncertainty, it is assumed that the MCT will be calibrated according to metrological standards.

To estimate the measurement uncertainty, a mathematical model for each measurement of the MCT is defined. The mathematical model of each measurement consists of the analytical formula (Equ. 1) of Wheatstone bridges:

$$S_i = \frac{k_{SG}}{4} \cdot (\varepsilon_1 - \varepsilon_2 + \varepsilon_3 - \varepsilon_4) = \frac{k_{SG}}{4} \cdot \bar{\varepsilon}_a \quad (1)$$

The k-factor k_{SG} is a strain gauge depending constant, which can vary for each strain gauge. The variation can be controlled by a calibration of the MCT. Nevertheless, the strain $\bar{\varepsilon}_a$ is effected by linearity deviation and hysteresis (ε_{lh}), repeatability (ε_{rp}), temperature influence (ε_{TK}) as well as the crosstalk (ε_{talk}):

$$\varepsilon_{entire} = \bar{\varepsilon}_a + \Delta\varepsilon_{lh} + \Delta\varepsilon_{rp} + \Delta\varepsilon_{TK} + \Delta\varepsilon_{talk} \quad (2)$$

For the metrological uncertainty components of strain gauges, values from literature [9] and for the crosstalk the simulated effects are considered for the estimation of the measurement uncertainty (see Tab. 2). Additionally, probability density functions (PDF) for the uncertainty components are assumed to a uniform distribution except for the repeatability, which is assumed to a normal distribution. By knowing the crosstalk and the uncertainties of strange gauges, it is possible to estimate the expanded measurement uncertainty according to GUM [1]. With the distributions, the absolute standard uncertainty $u_i$of each uncertainty component is calculated (Equ. (3)).

$$u_i = c_i \cdot d_i \cdot F_{nom,j} \quad (3)$$

With the PDF factor c_i, which is the standard uncertainty of the chosen distribution and the relative uncertainty d_i of the different uncertainty components.

Tab. 2: Relative uncertainty of strain gauges for linearity deviation and hysteresis (d_{lh}), repeatability (d_{rp}), temperature influence (d_{TK}) [9] as well as the crosstalk (d_{talk})

	Value [%]	Referred to	PDF-Factor
	d_i	$F_{nom,j}$	c_i
d_{lh}	0,1	Nominal load	0,58
d_{rp}	0,01	Nominal load	1
d_{TK}	0,05	Nominal load	0,58
d_{talk}	See Tab.1	Nominal load	0,58

Afterwards the combined uncertainty is calculated out of the absolute standard uncertainties of the uncertainty components in respect to the mathematical model and the law of propagation of uncertainty (Equ. 4).

$$u_{total} = \sqrt{\Sigma_{i=1}^{N} (\partial u_i / \partial i)^2} \quad (4)$$

By taking a coverage probability of 95 % (k=2) into account the absolute expanded measurement uncertainty is calculated after Equ. (5) (see Tab. 3).

$$U_{total} = k \cdot u_{total} \quad (5)$$

The relative expanded measurement uncertainty for each measurement is with 0.22% of the nominal load below the aimed requirement.

Tab. 3: Expanded measurement uncertainty for each measurement

	Nominal load	U_{total}	$U_{total,rel}$
F_x	1,5 MN	3,298 kN	≈ 0,22%
F_y	2 MN	4,421 kN	≈ 0,22%
M_x	5 MNm	10,933 kN	≈ 0,22%
M_z	5 MNm	11,022 kN	≈ 0,22%

Conclusion

This paper shows the design study of a MCT for wind turbine test benches . The developed MCT shows a high potential for a precise measurement of all six DoF in its own individual measurement ranges. A first estimation of the measurement uncertainty shows that the expanded uncertainties are below 0.5 % for all components. This includes the influence of crosstalk, temperature change as also the metrology characteristics like e.g. linearity, hysteresis and repeatability. In that way the MCT would provide a high potential to reduce the uncertainty in the traceability of the drive train behavior due to the applied loads.

Currently there is no possibility to calibrate such a MCT in the full measurement range, but with the upcoming 5 MN·m torque standard measuring device (expected 2021) at the National German Metrology Institute (PTB) the basis is laid for a calibration according to metrological standards [10].

Literature Reference

[1] JCGM, "Evaluation of measurement Evaluation of measurement data - Guide to the expression of uncertainty in measurement", (2010)
[2] AGEB - AG Energiebilanzen e.V., "Bruttostromerzeugung in Deutschland ab 1990 nach Energieträgern", https://ag-energiebilanzen.de/index.php?article_id=29&fileName=20181214_brd_stromerzeugung1990-2018.pdf, accessed April 11, 2019
[3] Bundesministerium für Wirtschaft und Energie, 7. Energieforschungsprogramm der Bundesregierung: Innovationen für die Energiewende, (2018)
[4] D. Bosse, G. Jacobs, T. Duda, Capabilities of Wind Turbine Ground Testing, World Congress and Exhibition on Wind & Renewable Energy, Berlin, Germany (2016)
[5] R. Averous, M. Stieneker, S. Kock, C. Andrei, A. Helmedag, W. De Doncker,K. Hameyer, G. Jacobs, A. Monti, Development of a 4 MW Full-Size, PEDG 2015 IEEE 6th International Symposium on Power Electronics for Distributed Generation Systems, Aachen (2015); doi: 10.1109/JESTPE.2017.2667399
[6] Stefan Kock, Georg Jacobs, Dennis Bosse, Gisa Foyer, "Influence on MN·m torque measurement in multi-MW nacelle test benches", DEWEK 2017 13th German Wind Energy Conference, Bremen (2017)
[7] VDI - Verein Deutscher Ingenieure, "Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte", Vol. 03.100.40, No. 2221, Beuth Verlag GmbH, Berlin (1993)
[8] S. Kock; G. Jacobs; D. Bosse; F. Strangfeld, XXII World Congress of the International Measurement Confederation, (2018); doi: 10.1088/1742-6596/1065/4/042014
[9] R. Schwart, Kraft, Masse, Drehmoment. In: Gevatter, H.-J.; Grünhaupt, U. (Hrsg.): Handbuch der Mess- und Automatisierungstechnik in der Produktion, 55–92 (2006)
[10] Dr. Prof. h. c. F. Härtig, Physikalisch-Technische Bundesanstalt, Abteilung 1: Mechanik und Akustik, Abteilungsbericht 2017, (2018)

DOI 10.5162/sensoren2019/1.2.4

Für internationale Vergleichsmessungen optimierter Drehmoment - Transferaufnehmer TN

Dr.- Ing. André Schäfer
Anschrift:
Hottinger Baldwin Messtechnik GmbH (HBM),
Darmstadt, Germany,
Tel.-Nr.: +49-6151-803224;
E-Mail-Adresse: andre.schaefer@hbm

Zusammenfassung

Drehmoment-Referenzaufnehmer, welche als Transferstandards für die Rückführung der Messgröße Drehmoment dienen, werden in der Regel auf der Basis von Dehnungsmessstreifen aufgebaut. Um von den Vorteilen eines komplett monolithischen Designs in metrologischen Anwendungen profitieren zu können, muss dieser Drehmomentsensor in Schaftform ausgeführt werden. Schaftsensoren wiederum können nur "offen" ausgeführt werden, anders als bei der üblicheren Ausführung als Flansch, wo die Dehnungsmessstreifen (DMS) hermetisch gekapselt werden können. Die Ausführungsform Schaft ist also eine besondere Herausforderung, da Folien-DMS sind ohne hermetische Kapselung recht stark von den Umgebungsbedingungen abhängig sind und so hierauf besonderes Augenmerk gelegt werden muss.

Da Drehmoment- Transferstandards für Vergleiche zwischen internationalen Metrologie-Laboren gedacht sind, ist deren Verwendbarkeit auch unter veränderten Bedingungen von großer Bedeutung. Die Gesamtzeit der Messungen aller Labore inkl. der Transferzeit soll kurz sein, die Ergebnisse sollen eine bestimmte Messunsicherheit bestätigen. Daher wurden Auswirkungen bestimmter Anwendungsbedingungen, und zwar ebenso für die benötigte Hochpräzisionselektronik, eingehend untersucht.

Keywords: Drehmoment-Referenzaufnehmer, Dehungsmessstreifen, DMS, Wheatstone-Brücke, Präzisionsmessverstärker, Präzisionsbrückenverstärker

Einführung

Seit dem Jahre 1958 offeriert HBM Drehmomentmesswellen. Die im Zusammenhang mit dieser Messgröße stehenden Geschäfte wuchsen vor allem in den letzten drei Jahrzehnten recht schnell. Im Jahre 1990 erlangte die am Stammhaus in Darmstadt aufgebaute 25 kNm - Bezugsnormalmesseinrichtung sogar die Rolle eines nationalen Normals. Diese Anlage war damals auf der Basis eines Schneidenlagers aufgebaut worden. Heute verwendet man vor allem Luftlager, die noch einmal deutlich reibungsärmer sind.



Abbildung 1: Wichtige Baureihen des HBM Drehmomentaufnehmer-Programms

Um die Forderungen der Automobilindustrie zu meistern, müssen Drehmomente gemäß den angegebenen und weiter reduzierten Unsicherheiten dargestellt und weitergegeben werden. Daher müssen Messketten, bestehend aus Referenzaufnehmern und Präzisionsinstrumenten, die z. B. für die Rückführung von Drehmomentschlüsseln in diesen Drehmomentlaboratorien vorgesehen sind, hohe Anforderungen erfüllen. Zugleich muss die Präzisionsmesskette portabel sein, um als stabiler "Travel Standards" für Vergleiche zwischen Labors verwendet werden zu können, wobei auf dem Weg vom Labor zum Labor eine bestimmte Zeit erforderlich sein wird, um sich an das Klima im Labor anzupassen und die Messung schließlich starten zu können.

Neben der oft gut dokumentierten Temperatur muss hier auch die Luftfeuchtigkeit sorgfältig, als mögliche Quelle von Messunsicherheiten, betrachtet werden. Hygroskopische Substanzen reichern sich mit Wasser an. Diese verdient vor allem im DMS-Aufnehmerbau großes Augenmerk, da bei "offenen" DMS-Konstruktionen deren Signal durch ein ggf. hygroskopisches Verhalten des DMS-Trägers beeinflusst wird.



Abbildung 2:
Bei Schaftaufnehmern erforderlicher Ringspalt: Um einen Kraftnebenschluss zu verhindern, muss der Rotor vollständig vom Rotor getrennt sein. Durch den erforderlichen Ringspalt kann Umgebungsluft in die DMS eindringen.

Verdeutlichen kann man sich dies, indem man sich einen Schwamm vorstellt, der das Wasser aus der Umgebung aufsaugt, und der sich daher vergrößert. Vier DMS sind in der Wheatstone-Brücke verschaltet, aber auch die externe Beschaltung, z.B. zur Temperaturkompensation, kann dies geschehen. Wenn ein solcher Drehmoment-Referenzaufnehmer auf DMS-Basis und in "offener Bauart" unterschiedlichen Feuchtigkeitswerten ausgesetzt ist, zeigt er eine Reaktion sowohl in der Empfindlichkeit als auch im Nullsignal.

Die Wirkung auf das Nullsignal ist zwar oft deutlich höher als die Empfindlichkeitsänderung, jedoch kann dies weitestgehend durch Tarieren vor dem Messvorgang kompensiert werden. Der Einfluss auf die Empfindlichkeit, also Sensitivität, schlägt jedoch voll durch auf das Ergebnis der Labore und eine Minimierung dieses ist für die Verwendbarkeit in Vergleichen zwischen Labors unabdingbar [1]; [2].

Der neue Referenzdrehmomentaufnehmer

Die Schaft-Grundform ist bereits seit vielen Jahren Basis für die Referenzdrehmomentaufnehmer-Serie „TN“ von HBM. Er wurde bereits weitestgehend auf den Einsatz als „Transfer standard“ in Kalibrierlaboratorien gemäß DIN 51309 z.B. in DAkkS-akkreditierten Stellen ausgelegt.

Von den Messbereichen her wird weiterhin für Nenndrehmomente 100, 200, 500 Nm und 1, 2, 5, 10 und 20 kNm angeboten. Spezielles Design, aber auch der Lieferumfang des Aufnehmers wurden hingegen weiter optimiert.

Nun wird auch die bisher oft gefragte Biegemomentoption Bestandteil des Lieferumfangs. Das bedeutet, dass der neue TN neben dem Drehmomentmesskanal nun standardmäßig mit zwei Biegemoment-Messkanälen (X-/Y-Richtung) ausgestattet ist, um mögliche Lage-Abweichungen zu detektieren.



Abbildung 3: TN-Referenz-drehmoment-aufnehmer in seiner neuen Ausprägung. Auf der Frontseite: neben dem Drehmoment-messkanal nun standardmäßig zwei Biegemoment-Messkanäle (X-/Y-Richtung)

Zusätzlich ist nun ein Tragekoffer im Standardlieferumfang enthalten, welcher mit einer speziellen Dichtlippe ausgestattet ist, die einen guten mechanischen Schutz, aber eben auch eine gute Feuchtigkeitsisolierung bietet. Aufgrund des erweiterten Standard-Lieferumfangs ist zukünftig nur noch eine Ausprägung pro Messbereich erforderlich, was auch die Lieferzeit verkürzt.



Abbildung 4:

Neuer, kompakter Transportkoffer mit spezieller, umlaufender Dichtlippe

Untersuchungen zur Empfindlichkeit gegenüber Feuchtigkeit

Die Feuchtigkeitsempfindlichkeit ist nicht nur ein Thema für den Drehmomentaufnehmer, sondern für die gesamte hochpräzise Messkette. Das dieser Tatsache ist trotz der Klimatisierung der Labore der Nationalen Metrologie-Institute (NMI) große Bedeutung beizumessen ist, zeigen Berichte von „Key Comparisions“ als auch speziell durchgeführte Untersuchungen [3]; [4].

Die generelle Einsicht, dass es angesichts deutlicher Fehlerbilder infolge Luftfeuchtigkeit sinnvoll ist, auf klimasichere Komponenten der Referenzkette, d.h. sowohl Aufnehmer als auch Elektronik hinzuarbeiten, ist seit langem klar [5]; [6].

Zunehmend stellte sich heraus, dass mit immer kleiner werdendem Messunsicherheits-budget der Einfluss der Luftfeuchtigkeit nicht mehr vernachlässigbar ist. In Bezug auf den Aufnehmer haben sich die NMIs bereits seit einigen Jahren mit dem Problem befasst [7]; [8]. Schließlich wurden auch für die Präzisionsinstrumente entsprechende Untersuchungen angeregt [9]; [10]; [11].

Die für die Anwendung „internationale Vergleichs-messung“ so wichtige Un-empfindlichkeit gegenüber Feuchtigkeit aus der Umgebungsluft konnte nun, im Ergebnis umfangreicher Untersuchungen erstmalig in der DMS-Messtechnik für die Messkette quantifiziert werden.

Daher konnten hier folgende Angaben zum Einfluss der relativen Feuchte von TN ermittelt werden:

Gerätefamilie	In den Aus-prägungen (Bestell-nummer)	Feuchte-empfindlichkeit - typische Werte
Transfer-normal TN (neue Aus-führung 2019/ 2020)	1-TN/100Nm 1-TN/200Nm 1-TN/500Nm 1-TN/1kNm 1-TN/2kNm 1-TN/5kNm 1-TN/10kNm 1-TN/20kNm	Typ. < 10 ppm/ % rH

Schlussfolgerungen

Obwohl bereits vielfach bewährt, konnte der Drehmoment-Transferaufnehmer TN in seiner Eignung für internationale Vergleichs-messungen noch einmal deutlich verbessert werden.

Für die gegebenen Konstruktionsoptionen bieten die Verbesserungen am Sensor in Verbindung mit dem neuen Tragekoffer nicht nur ausreichenden mechanischen Schutz bei trotzdem hoher Portabilität, sondern eben auch

eine höhere Immunität gegenüber verschiedensten Umgebungsbedingungen.

Dabei ermöglicht eine Verringerung der Feuchtigkeitsempfindlichkeit aufgrund der Verwendung eines deutlich weniger hygroskopischen DMS-Trägermaterials es auch die erforderliche Zeit, welche für eine ausreichende "Klimaanpassung" vor Messbeginn in jedem Labor erforderlich ist, entsprechend zu verkürzen.

Die Motivation hierfür ist, dass mit ihm vor allem primäre Anwendungen bedient werden sollen und daher kleinste Messunsicherheiten erforderlich sind. Dazu muss ein deutlich größerer Aufwand getrieben werden.

Je geringer das Messunsicherheitsbudget ist, desto mehr Parameter stellen sich für dessen Einhaltung als relevant heraus. Dafür betreffen die Maßnahmen hin zu einer noch besseren Feuchtigkeitsunempfindlichkeit den Aufbau des Aufnehmers selbst wie auch den neuen Tragekoffer, zumal dieser zusätzlich auch einen besseren mechanischen Schutz gegen unterschiedlichste Umgebungsbedingungen bietet.

Im Allgemeinen ermöglicht eine Verringerung der Feuchtigkeitsempfindlichkeit/ -aufnahme durch eine bessere Verpackung und die Verwendung weniger hygroskopischer DMS-Werkstoffe außerdem eine Verkürzung der oben erwähnten Zeit für die „Akklimatisation“ vor dem Start der Messung.

Für Ringvergleiche zwischen Labors und die kleinste Messunsicherheit kombiniert man am besten TN-Aufnehmer mit einem DMP41-T6 in einer Präzisionsmesskette [12].

Literaturnachweis

[1] Röske, D.; Mauersberger, D. On the stability of measuring devices for torque key comparisons, IMEKO XVIII World Congress & IV Brazilian Congress of Metrology, September 2006, Rio de Janeiro, Brazil.

[2] Brüge, A. Influence of humitidy on torque – estimation methods for calibration laboratories; XX IMEKO World Congr. 2012, Busan, Rep. of Korea

[3] Röske, D., Key comparisons in the field of torque measurement, 19th IMEKO TC3 on Force, Mass and Torque: "Theory and Application in Laboratories and Industries", 2005, Cairo, Egypt.

[4] Röske, D. Final report on the torque key comparison CCM.T-K1: Measurand torque: 100 N · m, 500 N · m, 1000 N · m, Metrologia, 46 (2009), Tech. Suppl., 07006.

[5] Ring, K. Klimasicherheit von elektronischen Baugruppen (engl. transl.: Climate safety of electronic assemblies), ZVE Fraunhofer Institut für Zuverlässigkeit & Mikrointegration, Berlin, 2007

[6] Käs, G. Qualität und Zuverlässigkeit elektronischer Bauelemente und Systeme, R. Oldenburg publishing, Vienna, 1983, ISBN-3486-2735-5

[7] Brüge, A. Simplified measurements of the humidity coefficient of torque transducers in calibration laboratories, ACTA IMEKO June 2014, Volume 3, Number 2, 32 – 38, www.imeko.org

[8] Khaled, K.M.; Röske, D. at al. Humidity and temperature effects on torque transducers, bridge calibration unit and amplifiers, Measurement 74 (2015) p. 31–42, Elsevier B.V. Publishing house, N.Y.; U.S.A.

[9] Drung, D. at al ; Improvements in the field of traceable high-accuracy measurement and generation of small electrical currents, Presentation of Helmholtz award winners, PTB WG 2.61, Bad Honnef, Germany, May 2018

[10] Schäfer, A. The Ultra-Precision Instrument DMP41 –First Experiences & Appropriate Filter settings, TC3 Conference, 2014, Cape Town, South Africa

[11] Schäfer, A. DMP41–A new Chapter of Ultra-Precision Instrument for Strain Gauge Transducers XX IMEKO World Congr. 2012, Busan, Rep. of Korea

[12] Schäfer, A. Notwendigkeiten und Gesetz-mäßigkeiten beim Aufbau von Präzisions-messketten auf Basis von Dehnungs-messstreifen. Tagung Innovation Messtechnik, Linz, Österreich, Mai 2019

DOI 10.5162/sensoren2019/1.3.1

Erfassung biomechanischer Belastungen und Asymmetrien bei Nachwuchs-Eliteläufern mithilfe mehrachsiger Inertialsensoren

Olaf Ueberschär[1], Julia Roediger[1], Daniel Fleckenstein[1,2], Frank Warschun[1], Nico Walter[1], Matthias W. Hoppe[3,4]
[1]Institut für Angewandte Trainingswissenschaft (IAT), Marschnerstraße 29, 04109 Leipzig
[2]Deutscher Leichtathletik-Verband, Alsfelder Str. 27, 64289 Darmstadt
[3]Institut für Sportwissenschaft, Universität Wuppertal, Fuhlrottstraße 10, 42119 Wuppertal
[4]Klinikum Osnabrück, Am Finkenhügel 1, 49076 Osnabrück

Zusammenfassung

Problemstellung: Trotz ihrer gesundheitsförderlicheren Aspekte bedingt die Sportart Laufen eine erhebliche muskuloskelettale Beanspruchung. Die Inzidenz von Überlastungsverletzungen fällt entsprechend hoch aus, wobei statistisch u. a. die Schienbeine als häufige Verletzungslokalisation in Erscheinung treten. Die Erfassung tibialer Beschleunigungen gilt als etabliertes indirektes Messverfahren zur Ermittlung von Stauchungskräften. Unter Kenntnis der individuell auftretenden tibialen Belastung kann der Trainer bzw. Therapeut die mit einem Trainingsprogramm einhergehende biomechanisch-orthopädische Beanspruchung besser auf den jeweiligen Sportler bzw. Patienten abstimmen.

Lösungsansatz und Methodik: Es wurden handelsübliche, funkbasierte 9-achsige Inertialsensoren (3D- Akzelerometer, 3D-Gyroskope und 3D-Magnetometer) an die distalen Schienbeine (2x), den Kreuzbeinbereich (1x) sowie die Schulterblätter (2x) von 45 gesunden Nachwuchs-Eliteläufern nichtinvasiv angebracht. Mit ihnen wurden die tibialen, sakralen und skapularen Stauchungsbelastungen sowie deren Dämpfung und seitliche Asymmetrien bei Laufgeschwindigkeiten von 14 bis 16 km h^{-1} rückwirkungsfrei auf einem motorisierten Laufband gemessen.

Ergebnisse: Es wurden mittlere tibiale Spitzenbeschleunigungen von 14±3 bis 16±3 *g* bei 14 bis 16 km h^{-1} ermittelt. Die entsprechenden Spitzenwerte an Kreuzbein und den Schultern betrugen 4±1 *g* bis 5±1 *g* (d. h. 32±8 % der Schienbeinbelastung) bzw. 4±1 *g* (27±6 %). Die beobachteten lateralen Asymmetrien der tibialen und skapularen Beschleunigungen beliefen sich im Absolutwert-Mittel auf 9±8 % (95. Perzentile: 24 %) bzw. 9±10 % (32 %).

Schlussfolgerungen: Inertialsensoren in Form von vollintegrierten Smart Systems stellen einen vielversprechenden Ansatz für die Erfassung biomechanischer Belastungen und Asymmetrien für orthopädische und bewegungswissenschaftliche Fragestellungen unter Feldbedingungen dar. Mithilfe der Orientierungswerte, wie sie in der vorliegenden Arbeit präsentiert werden, könnten Inertialsensoren zukünftig zur Verletzungsprophylaxe und Laufeffizienzoptimierung eingesetzt werden.

Schlagworte: Akzelerometer; tibiale, sakrale, skapulare Beschleunigung; Stoßdämpfung, Asymmetrie

1 Einleitung

Der (Langstrecken-)Lauf gehört als olympische Kernsportart zu den weltweit populärsten Freizeitsportarten, sicherlich auch aufgrund des niederschwelligen und kostengünstigen Zugangs, des schnell erlebbaren Fitnesszuwachses und des enormen gesundheitsförderlichen Aspekts [1-3]. Aus allgemeinmedizinischer Sicht trägt die Sportart Laufen durch die einhergehende ganzkörperliche Belastung und einen mittleren bis hohen kardiovaskulären Reiz wesentlich zur Gesunderhaltung bei. Die weitläufige Annahme eines vergleichsweise geringen Verletzungsrisikos bestätigt sich statistisch jedoch nicht [1, 2, 4]. Tatsächlich geht aus epidemiologischen Studien hervor, dass Laufen ein nicht unerhebliches orthopädisches Verletzungsrisiko infolge der erheblichen biomechanischen Beanspruchung des muskuloskelettalen

Systems birgt [1, 2, 4, 5]. Vor allem in Trainingsphasen hoher Laufumfänge erweisen sich die unteren Extremitäten als anfällig für orthopädische Leiden. Jenseits eines „kritischen“ (individuellen) wöchentlichen Trainingspensums von ca. 60 km steigt das Risiko einer Überlastungsverletzung nachweislich. In der Wettkampfvorbereitung werden von Spitzenathleten jedoch bis zu 150–200 Wochenkilometer zurückgelegt [5, 6]. Verschiedene epidemiologischen Arbeiten haben ergeben, dass zwischen 27 % bis 70 % der Freizeit- und Wettkampfläufer im Jahresverlauf eine Überlastungsverletzung erleiden [6, 7]. Dabei gehören das Kniegelenk, die Achillessehne und das Schienbein zu den häufigsten Verletzungslokalisationen [2, 6, 7].

Die an den Schienbeinen auftretenden Beschleunigungen werden von Sportmedizinern und Bewegungswissenschaftlern als valides Maß für die auftretenden biomechanischen (Stauchungs-)Belastungen angesehen [8]. Obwohl die Messung jener tibialen Beschleunigungen letztlich nur eine indirekte Abschätzung der auf Knochen und Bindegewebe wirkenden Kräfte ermöglicht, liefert sie dennoch eine direktere Aussage über die Schienbeinbelastung als der konventionelle Ansatz mittels dynamometrisch erfasster Bodenreaktionskräfte [9].

Je nach Kontext werden entweder der Betrag $a \equiv |\vec{a}|$ des dreiachsigen Beschleunigungsvektors $\vec{a} \equiv (a_x, a_{ML}, a_{AP})^T$ oder seine drei Komponenten entlang der ausgewiesenen anatomischen Achsen betrachtet, d. h. die axiale (a_x), die medio-laterale (a_{ML}) und die anterior-posteriore (a_{AP}) Komponente. Entsprechend gilt die Beziehung $a = \sqrt{a_x^2 + a_{ML}^2 + a_{AP}^2}$ [8].

In Hinblick auf die Stauchungsweiterleitung im Körper liefern die im Moment des ersten Bodenkontaktes im Schrittzyklus (Impact-Moment) an Becken und Schultern gemessenen „Rest-Beschleunigungen“ in Relation zur Schienbeinbeschleunigung wertvolle biomechanische Einblicke: Anhand dieser „Stoßdämpfungskaskade“ von den Füßen bis zu den Schultern können belastbare Aussagen über die Steifigkeit der unteren Extremitäten unter Beteiligung von Sprung-, Knie- und Hüftgelenk, über die Rumpfstabilität, die sportliche Leistungsfähigkeit und auch über den körperlichen Erschöpfungszustand des Läufers abgeleitet werden. Die Rolle der Stoßdämpfung wird daher auch bei laufbedingten Verletzungen diskutiert [9, 10].

Zur Messung der am Körper auftretenden Beschleunigungen bieten sich seit jüngerer Zeit mehrachsige Inertialsensoren an (engl.: *inertial measurement units, IMUs*), wobei diese in der Regel aus ein- bis dreiachsigen Beschleunigungssensoren, Gyroskopen und Magnetometern bestehen. Trotz zahlreicher Anwendungsbeispiele lagen bislang nur unzureichende Studienergebnisse vor, ob handelsübliche, kostengünstige IMUs auch tatsächlich in der Sport- und Bewegungspraxis für die Beurteilung biomechanischer Belastungen und Asymmetrien im (Langstrecken-)Lauf geeignet sind. Für die Bewegungs-, Sport- und Rehabilitationswissenschaften nehmen entsprechend valide Informationen jedoch eine Schlüsselrolle ein, da nur mit ihrer Hilfe die biomechanische Belastung und insbesondere die muskuloskelettale Beanspruchung einer Trainingseinheit unter Feldbedingungen abgeschätzt werden können. Aus dem Laborbetrieb bekannte Technologien wie Kraftmessplatten oder Dynamometer sind im Feld-Einsatz weder verfügbar noch sinnvoll einsetzbar.

Insgesamt besitzen Inertialsensoren das Potenzial, sowohl im leistungssportlichen als auch im rehabilitativem Kontext zu einer besseren Belastungsüberwachung und Trainingssteuerung in der Sportart Laufen beizutragen. Diese Arbeit geht dabei der Frage nach, ob, und falls ja wie dies mit kommerziell erhältlichen Inertialsensoren gelingen kann. Insbesondere stehen dabei die an Schienbeinen, Kreuzbein und Schultern von Nachwuchs-Eliteläufern auftretenden Impact-Beschleunigungen, ihre seitlichen Asymmetrien sowie die Stoßdämpfung im Mittelpunkt.

2 Methodik

2.1 Probanden

Fünfundvierzig Nachwuchs-Eliteläufer aus dem Langstreckenbereich der deutschen Nationalmannschaft wurden in diese Studie einbezogen (27 Männer: 18,5±1,8 Jahre, 182±5 cm, 67,8±6,9 kg, BMI 20,3±1,5 kg m^{-2}; 18 Frauen: 18,2±2,7 Jahre, 170±5 cm, 52,8±6,1 kg, BMI 18,3±1,6 kg m^{-2}). Ihre durchschnittliche maximale Sauerstoffaufnahme betrug 67,0±4,8 ml min^{-1} kg^{-1} bzw. 57,7±2,7 ml min^{-1} kg^{-1}, während der wöchentliche Trainingsumfang zum Zeitpunkt der Untersuchung bei 50,2±19,5 km lag.

2.2 Inertialsensoren

Für die Beschleunigungsmessungen wurden funkbasierte, nicht-invasive, dreiachsige Inertialsensoren eingesetzt. Weltweit bieten verschiedene Hersteller derartige vollintegrierte Sensorsysteme zur Erfassung von Beschleunigungen, Winkelgeschwindigkeiten und (teilweise) Magnetfeldstärken für sportliche und orthopädische Zwecke an. In dieser Studie wurde das validiert System MTw Awinda von einem der führenden Anbieter (Xsens Technologies B.V., Enschede, Niederlande) verwendet, ins-

besondere um die Vergleichbarkeit zu vorherigen Studien sicherzustellen. Die eingesetzten Sensoren sind klein (4,7 cm · 3,0 cm · 1,3 cm), leicht (16 Gramm) und haben eine ausreichende Akkulaufzeit im Dauerbetrieb von ca. vier Stunden. Die Datenrate der per Sensor-Fusion verrechneten Daten wurde auf 120 Hz gesetzt, während die interne Abtastrate der Beschleunigungsachsen bei 1000 Hz lag [11]. Jede der drei Akzelerometerachsen besaß einen Messbereich von ±160 m s^{-2} und eine rauschbedingte Messunsicherheit von lediglich ±0,02 m s^{-2}. Ausgehend von den drei gemessenen Komponenten a_x, a_y und a_z des Beschleunigungsvektors $\vec{a} = (a_x, a_y, a_z)^T$ im mitbewegten Sensorkoordinatensystem ergibt sich der Impactbetrag zu

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}\,. \qquad (1)$$

Da die Sensorachsen und die äußere Beschleunigungsrichtung in der Regel nicht zusammenfallen, erhöht sich der effektiv verfügbar Messbereich für den Betrag dabei von ±16 *g* der einzelnen Sensorachsen auf etwa 18–22 *g* $\left(\approx \sqrt{2} \cdot 16\,g\right)$ bis bestenfalls (Raumdiagonale) $28\,g\left(\approx \sqrt{3} \cdot 16\,g\right)$.

Im Allgemeinen unterliegen aktuelle mikroelektronische Beschleunigungssensoren (und Gyroskope) einer Sensordrift, d. h. einer nicht vorhersehbaren, systematischen Langzeit-Messunsicherheit für kleine, quasikonstante Beschleunigungen $a \ll g$ (bzw. kleine Winkelgeschwindigkeiten). Für die in dieser Studie verwendeten Akzelerometer ist die angegebene Offset-Stabilität jedoch hoch und entspricht einer systematischen Abweichung von lediglich 0,0001 *g* [11]. Darüber hinaus sind Beschleunigungsmessungen innerhalb eines Schrittzyklus durch mehrere intrazyklische, kurzfristige Vorzeichenänderungen gekennzeichnet, wobei die Beschleunigungsbeträge $a \gtrsim g$ jeweils deutlich größer als die Fallbeschleunigung sind. Beide Aspekte minimieren den Störeinfluss der Akzelerometer-Drift bei Lauf-Anwendungen, insbesondere in Hinblick auf die für die vorliegende Arbeit relevanten Impact-Spitzenwertmessungen. Messfehler durch Sensordrift können in dieser Studie daher als vernachlässigbar klein betrachtet werden. Hinsichtlich der Achsenlage beträgt der statische Ausrichtungsfehler zwischen Gehäuse- und Sensorachsen ferner maximal 0,1° [11] und ändert sich während der Messungen nicht. Insgesamt ist somit auch für Langzeitmessungen bei laufspezifischen Anwendungen keine Kalibrierung oder Neuausrichtung der Akzelerometer erforderlich. Da nur Akzelerometriedaten verwendet werden, ist auch eine etwaige Gyroskop-Drift gleichfalls irrelevant für diese Studie.

2.3 Sensoranbringung

In der vorliegenden Arbeit wurden bei jedem Sportler und jeder Sportlerin zwei Inertialsensoren nicht-invasiv am distalen anteromedialen Ende des rechten und linken Schienbeins mittels enganliegender Klettbänder derart befestigt, dass ihre Unterkanten etwa 8 cm über der Malleolus-Ebene lagen (Abb. 1 (c)). An jener Position ist die Relativbewegung zwischen Schienbein und Sensor während des Laufens minimal, da sich kein voluminöses Muskel- oder Weichteilgewebe zwischen Sensor und Knochen befindet. Darüber hinaus wurde jeweils ein Sensor auf jedem Schulterblatt und dem Kreuzbein angebracht, wobei hierzu entweder spezielle Funktionshemden (Abb. 1 (a)) oder eine Kombination aus Standard-Leukotape™ und Klettbändern (Abb. 1 (b)) eingesetzt wurden.

Aufgrund der geringen Bauhöhe des eigentlichen mikroelektronischen Sensorchips im Sensorgehäuse ist ein systematischer Messfehler infolge eines künstlich verlängerten Hebelarmes für die betrachteten tibialen, sakralen und skapularen Messpunkte vernachlässigbar: Während der zusätzliche Hebelarm des Sensors als Abstand zwischen Sensorchip und Haut maximal 6 mm beträgt, ist der Abstand vom Sensorsystem zur jeweiligen anatomischen Drehachse mit ≳150 mm über eine Größenordnung größer. Der resultierende maximale Messfehler liegt damit unter 5 %.

Ungeachtet dessen muss kritisch bemerkt werden, dass ein gewisser Dämpfungs- bzw. Verzerrungseffekt durch Weichteil- und/oder Hautbewegungen nicht vollständig ausgeschlossen werden kann. Diese Störeinflüsse können für die vorliegende Studie jedoch als ausreichend klein angesehen werden, insbesondere angesichts der genutzten anatomischen Messpunkte und den vorliegenden läufertypischen schlanken Somatypen der Sportler.

Um weiterhin im Kontext der Symmetrieanalysen etwaige Fehler aufgrund von systematischen Messunsicherheiten der einzelnen Sensoren auszuschließen, wurden diese nach einer Testreihe in zufälliger Reihenfolge zwischen den Messpunkten ausgetauscht. Ferner wurden vor jedem Test alle Sensoren synchronisiert und auf ihre ordnungsgemäße Funktionalität überprüft.

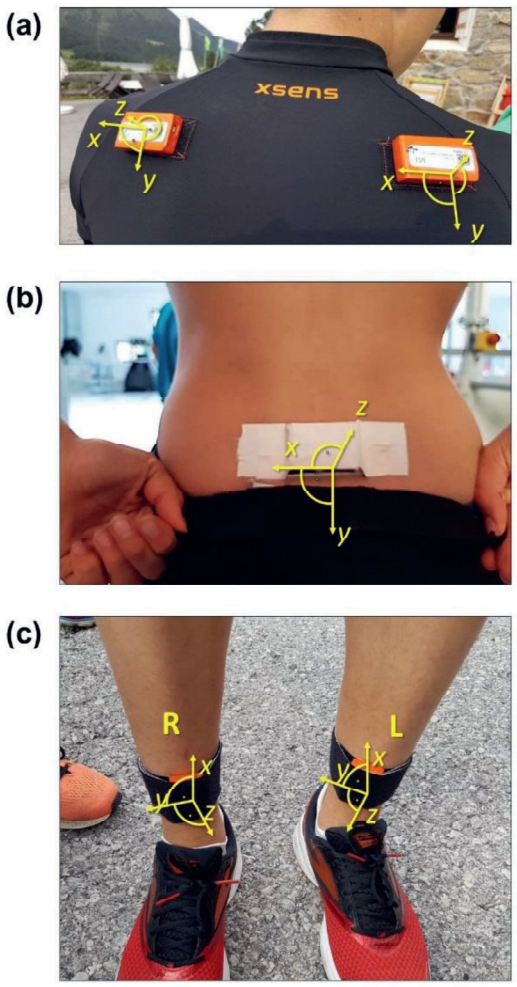


Abb. 1 Sensorplatzierung an Schienbeinen, Schulterblättern und Kreuzbein: (a) Schultersensoren. (b) Kreuzbein-Sensor, (c) Schienbein-Sensoren. Die gelben Pfeile geben die Lage der Sensorkoordinatensysteme wider.

2.4 Versuchsdurchführung

Die Untersuchungen fanden im Rahmen der regulären sog. komplexen Leistungsdiagnostik der Nachwuchs-Eliteläufer am Institut für Angewandte Trainingswissenschaft (IAT) in Leipzig statt, dem nationalen Forschungsinstitut für den deutschen Spitzen- und Nachwuchsleistungssport. Jene Leistungsdiagnostik beinhaltete dabei u. a. einen submaximalen Laufbandtest, bestehend aus vier Stufen von 2000–3000 m auf einem motorisierten Laufband ohne Neigung, wobei die genaue Distanz und Geschwindigkeit individuell auf die Athleten in Abhängigkeit ihrer Wettkampfdistanz und ihres Leistungsvermögens festgesetzt wurde. Zwischen den einzelnen Stufen hatten die Athleten eine Pause von einer Minute, nach der jeweils die Geschwindigkeit um 0,9 km h^{-1} erhöht wurde. Die Anfangsgeschwindigkeit wurde von einem erfahrenen Trainer auf der Grundlage des aktuellen Leistungsniveaus des Athleten gewählt und betrug im Mittel 13,9±1,4 km h^{-1}.

Von den 45 Athleten absolvierten alle die Stufen 1 und 2, während die Stufen 3 und 4 noch von 44 bzw. 37 Athleten erfolgreich beendet wurden. Die durchschnittliche Zeit für die Durchführung aller vier Stufen einschließlich der Pausen betrug 34±4 min. Alle Athleten trugen beim Laufen ihre eigenen individuellen Laufschuhe. Während der Läufe wurden die Beschleunigungen an Schienbeinen, Kreuzbein und Schulterblättern wie oben beschrieben gemessen. Es war keine Kalibrierung oder Neukalibrierung der verwendeten mikroelektronischen Beschleunigungssensoren erforderlich, wie zuvor erläutert.

2.5 Datenauswertung

Zunächst wurden die erhobenen Binär-Rohdaten der Beschleunigungsmessungen mit einer herstellerseitigen Software (MT Software Suite MTw™ AwindaTM 4.8, Xsens Technologies B.V., Enschede, Niederlande) aufbereitet und in ASCII-Daten umgewandelt. Die anschließenden Analysen wurden mithilfe einer selbstprogrammierten Software durchgeführt (LabVIEW 2016, National Instruments, Austin, USA). Diese ermittelte die höchsten Impact-Werte jedes Schrittes und des kompletten Laufes in der aufgenommenen Zeitreihe des Beschleunigungsbetrags. Anschließend wurden die Mittelwerte für die Spitzenbeschleunigungen und Standardabweichungen (SD) für das linke und rechte Schienbein, für die Schulterblätter und das Kreuzbein für repräsentative Zeiträume berechnet. Zusätzlich zu diesen Daten für den Vektorbetrag $a \equiv |\vec{a}|$ wurde ferner die einachsige Beschleunigungskomponente a_x entlang der Schienbeine („axiale Komponente“) analog für jedes Bein deskriptiv statistisch ausgewertet. Die intra- und interindividuellen Statistiken wurden dabei mit Microsoft Excel 2016 (Microsoft Corporation, Redmond, USA) und IBM SPSS Statistics 23 (IBM, Armonk, USA) durchgeführt. Als Fehlerbalken wurde die einfachere SD angesetzt.

Zur Untersuchung von etwaigen Asymmetrien zwischen linkem und rechtem Bein wurden der Asymmetrie-Index Δ_a für den Beschleunigungsbetrag (Schienbeine und Schultern) sowie Δ_{a_x} für die axiale Beschleunigung (nur tibial) definiert:

$$\Delta_a \equiv \frac{a^{(L)} - a^{(R)}}{\frac{1}{2}\left(a^{(L)} + a^{(R)}\right)}; \qquad \Delta_{a_x} \equiv \frac{a_x^{(L)} - a_x^{(R)}}{\frac{1}{2}\left(a_x^{(L)} + a_x^{(R)}\right)}, \quad (2)$$

wobei $a^{(L)}$, $a_x^{(L)}$ und $a^{(R)}$, $a_x^{(R)}$ die entsprechenden mittleren Impact-Spitzenbeschleunigungen (Betrag und axial) des linken (L) und rechten (R)

Schienbeins bzw. Schulterblatts bezeichnen. Anhand Gl. (2) ist ersichtlich, dass eine völlig symmetrische Beschleunigungsverteilung zu Indexwerten von $\Delta_a = 0$ und $\Delta_{a_x} = 0$ führen würde. Bei einem dominanten linken Bein würden stattdessen Δ_a und/oder Δ_{a_x} positiv werden, bei einem dominanten rechten Bein hingegen negativ.

3 Ergebnisse

Die wichtigsten Ergebnisse sind in Abb. 3 zusammengefasst. Das Diagramm (a) zeigt die maximalen tibialen, sakralen und skapularen Beschleunigungen in Form von Mittelwerten ± SD bei unterschiedlichen Laufgeschwindigkeiten, während das Diagramm (b) die relative Stoßdämpfungskaskade darstellt, die sich aus ersteren Größen ableitet. Für die beiden Schienbein- und Schulter-Sensoren wird das arithmetische Mittel beider Sensoren (d. h. linkes und rechtes Bein; linke und rechte Schulter) verwendet.

Die jeweiligen Beschleunigungen nehmen mit der Laufgeschwindigkeit signifikant zu (p<0,004) und in Richtung Kopf signifikant ab (p<0,001, zweifaktorielle ANOVA mit Messwiederholung). Die geschwindigkeitsbedingte relative Erhöhung von Stufe 1 zu Stufe 4 beträgt im Mittel 15 % an den Schienbeinen, 9 % am Kreuzbein und 13 % an den Schulterblättern. Von der erfassten Impact-Spitzenbeschleunigung der Schienbeine ist noch ein Anteil von 32±9 % am Kreuzbein messbar. Diese verbleibende Impact-Beschleunigung wird weiter auf 27±6 % reduziert, wenn sie an den Schulterblättern gemessen wird.

Zu beachten ist hierbei, dass sich die am Kreuzbein und an den Schulterblättern gemessenen Beschleunigungswerte im Allgemeinen aus einer Überlagerung aus den (gedämpft) übertragenen Impact-Kräfte der Beine und zusätzlichen mit- und/oder gegenläufigen (Ausgleichs-) Bewegungen des Beckens und der Schultern ergeben. Dies trifft insbesondere auf die laufbildtypische sog. „Kreuzkoordination“ von Armen und Beinen zu. Da die Impact-Spitzenbeschleunigungen am Kreuzbein und an den Schultern in dieser Arbeit jedoch in unmittelbarem zeitlichem Zusammenhang mit den vorausgehenden tibialen Impact-Ereignissen ermittelt werden, kann dieser systematische Messfehler als vernachlässigbar klein eingestuft werden: Kontrollierte Ausgleichsbewegungen des Oberkörpers fallen in ihrem Maximum in der Regel nicht zeitlich mit tibialen Belastungsspitzen zusammen.

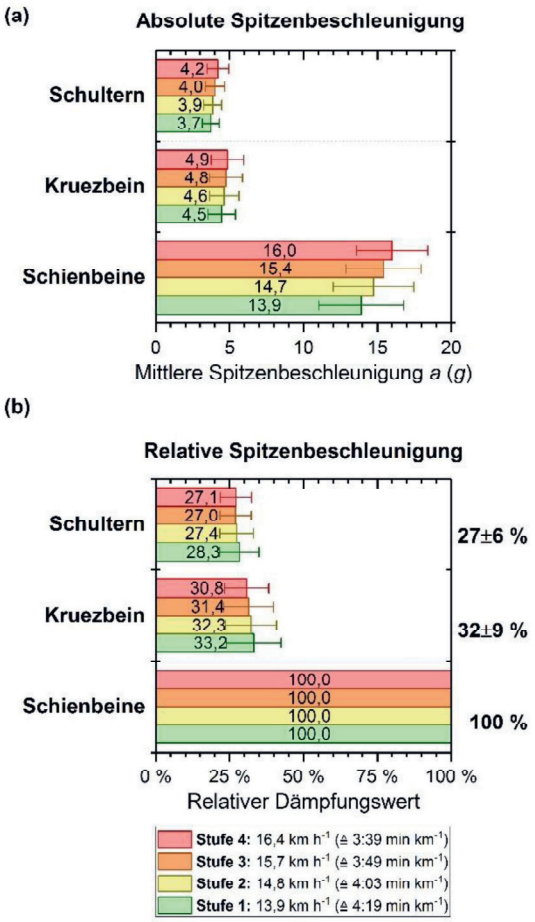


Abb. 2. Stauchungsbelastungen von Schienbeinen, Kreuzbein und Schultern bei Nachwuchs-Eliteläufern bei verschiedenen Laufgeschwindigkeiten. (a) Mittlere Impact-Spitzenbeschleunigungen für Schienbeine, Kreuzbein und Schultern. (b) Verbleibende Stauchungsbelastung in Bezug auf den Schienbeinwert in Prozent.

Bezüglich der Seitensymmetrie zeigen die ermittelten Asymmetrie-Indizes der Schienbeine keine relevante Seitigkeit, weder für den Beschleunigungsbetrag (Δ_a) noch für die axiale Beschleunigung (Δ_{a_x}). Das arithmetische Mittel für alle Probanden und alle Stufen beträgt dabei $\langle \Delta_a \rangle$=-2,5±11,8 % und entspricht somit einer geringfügigen Dominanz des rechten Beins. Axial liefert der Mittelwert $\langle \Delta_{a_x} \rangle$=+0,4±14,8 %, sodass praktisch keine Seitendominanz vorliegt. Für die Schulterblätter beträgt der Gesamtmittelwert des Asymmetrieindex $\langle \Delta_a \rangle$ =+4,7±12,8 %, was einer leichten Dominanz der

linken Schulter entspricht und somit die erwähnte Kreuzkoordination von Beinen und Armen geringfügig widerspiegelt.
Bemerkenswert ist, dass zwar alle drei Mittelwerte nahe bei null liegen, die Standardabweichungen jedoch deutlich größere Werte zwischen 11 % und 15 % aufweisen. Dies deutet auf ein beträchtliches Spektrum der individuellen Ausprägung von lateralen Asymmetrien hin. Die Absolutwert-Mittelwerte der Asymmetrie-Indizes betragen $\langle|\Delta_a|\rangle$ =9,0±8,1 % (95. Perzentil: 24,3 %), $\langle|\Delta_{a_x}|\rangle$ =11,6±9,2 % (29,3 %) und $\langle|\Delta_a|\rangle$ =9,2±9,9 % (31,7 %) für den Betrag von Schienbeinbeschleunigungen, ihre axialen Komponenten bzw. den Schulterbeschleunigungen. Sie sind praktisch unabhängig von der Laufgeschwindigkeit.

4 Ausblick

In dieser Arbeit konnte gezeigt werden, dass Inertialsensoren eine vielversprechende, praxistaugliche Messoption für die Sport-, Bewegungs- und Rehabilitationswissenschaften unter Feldbedingungen darstellen. Durch ihre geringen Kosten, einfache Bedienbarkeit und hohen Tragekomfort besitzen Inertialsensoren ein großes Potenzial, vermehrt als rückwirkungsfreie Messsysteme „am Sportler, im Feld und rund um die Uhr“ Anwendung zu finden. Gegebenenfalls eröffnen Inertialsensoren somit gegenwärtig in ähnlicher Weise eine neue Ära der nutzerfreundlichen Messwerterfassung für Sportler und Patienten, wie es tragbare Herfrequenzmesser vor 40 Jahren oder die ersten GPS-fähigen Sportuhren vor 20 Jahren taten. Mit ihrer Hilfe könnte in Zukunft eine verbesserte Bewertung der Laufeffizienz im Leistungssport stattfinden. Außerdem könnte das individuelle Risiko von Überlastungsschäden im Leistung- wie auch im Breitensport frühzeitig erkannt und verringert werden. Letzten Endes, so hoffen die Autoren, könnte diese Studie einen kleinen Beitrag dazu leisten, zukünftig das Verletzungsrisiko in der beliebten „Volkssportart“ Laufen weiter zu reduzieren.

Hinweis

Der vorliegende Beitrag stellt einen Auszug einer deutschsprachigen Kurzfassung des Fachartikels [12] dar.

Literaturnachweis

1. B. M. Nigg, *Impact forces in running.* Current Opinion in Orthopaedics **8**(6) (1997) 43-47.
2. J. E. Taunton, M. B. Ryan, D. B. Clement, D. C. McKenzie, D. R. Lloyd-Smith, B. D. Zumbo, *A retrospective case-control analysis of 2002 running injuries.* British Journal of Sports Medicine **36**(2) (2002) 95-101.
3. M. Van Middelkoop, J. Kolkman, J. Van Ochten, S. M. Bierma-Zeinstra, B. W. Koes, *Risk factors for lower extremity injuries among male marathon runners.* Scand J Med Sci Sports **18**(6) (2008) 691-697.
4. A. Hreljac, *Etiology, prevention, and early intervention of overuse injuries in runners: a biomechanical perspective.* Phys Med Rehabil Clin N Am **16**(3) (2005) 651-667, vi.
5. L. I. Tjelta, *The training of international level distance runners.* International Journal of Sports Science & Coaching **11**(1) (2016) 122-134.
6. W. van Mechelen, *Running injuries. A review of the epidemiological literature.* Sports Med **14**(5) (1992) 320-335.
7. R. Ferber, A. Hreljac, K. D. Kendall, *Suspected mechanisms in the cause of overuse running injuries: a clinical review.* Sports health **1**(3) (2009) 242-246.
8. K. R. Sheerin, D. Reid, T. F. Besier, *The measurement of tibial acceleration in runners—A review of the factors that can affect tibial acceleration during running and evidence-based guidelines for its use.* Gait & Posture **67**(2019) 12-24.
9. C. E. Milner, R. Ferber, C. D. Pollard, J. Hamill, I. S. Davis, *Biomechanical factors associated with tibial stress fracture in female runners.* Med Sci Sports Exerc **38**(2) (2006) 323-328.
10. J. Mizrahi, O. Verbitsky, E. Isakov, *Fatigue-induced changes in decline running.* Clin Biomech (Bristol, Avon) **16**(3) (2001) 207-212.
11. M. Paulich, M. Schepers, N. Rudigkeit, G. Bellusci, *Xsens MTw Awinda: Miniature WirelessInertial-Magnetic Motion Tracker for Highly Accurate 3D Kinematic Applications.* Xsens Technologies B.V.: Enschede, the Netherlands. (2018) 1-9.
12. O. Ueberschär, D. Fleckenstein, F. Warschun, S. Kränzer, N. Walter, M. Hoppe, *Measuring biomechanical loads and asymmetries in junior elite long-distance runners through multi-axis inertial sensors.* Sports Orthopeadics and Traumatology (under review) (2019).

DOI 10.5162/sensoren2019/1.3.2

A novel non-invasive, non-conductive method for measuring respiration

Jan Ringkamp[1], Philipp Radler[1], Philipp Lebhardt[1], Jens Langejürgen[1]
[1]Fraunhofer IPA, Theodor-Kutzer-Ufer 1-3 (Cubex41), 68167 Mannheim, Germany

Abstract

We present a novel method for measuring dynamic changes in thoracic parameters due to breathing based on the coupling of two UHF antennas. The phase shift between the antennas as well as the resonance frequency of the coupled antenna pair change with the air volume in a lung phantom. For evaluation, we built a dynamic 3D printed phantom encasing two compartments separated by an elastic diaphragm. By filling this artificial lung with air the effective permittivity in the compartment changes resulting in a detectable phase and frequency shift. We use a vector network analyzer to show that there is a strong linear dependency between the volume of air in the artificial lung and the frequency at which the S_{21} parameter peaks (R^2=0.98). Related to this, we observe a phase shift ($\Delta\varphi$) when looking at a single frequency close to the resonance frequency of the system. This shift is also linearly dependent on the volume of air in the lung (R^2=0.998). The presented method is a promising candidate to overcome some of the most important technical burdens of measuring respiratory parameters and might be used as trigger for patient–ventilator synchronization in infants and neonates.

Keywords: plethysmography; respiration, antenna detuning, noncontact monitoring, electromagnetic measurement

Introduction

Respiratory parameters like breathing frequency, tidal volume or inhalation onset are among the most crucial parameters in intensive care. Medical respirators need current values to adjust timing and flow rates in order to ensure an optimal and lung protecting ventilation. This is especially important when it comes to the ventilation of infants or neonates. Here, low tidal volumes (as low as 3 ml), the comparably higher breathing frequency, and the sensitive skin hamper precise measurement and control of respiratory parameters. When ventilating premature infants, it is common practice to use uncuffed endotracheal tubes to prevent the development of ventilator induced lung injury, like volu- or barotrauma [1]. This poses an additional challenge for respiration sensors, because the unknown leakage due to the uncuffed endotrachael tube makes simple pressure and airflow measurement highly unreliable.

Electrical and electromagnetical measurements are a promising class of sensors, because they are leakage independent and several embodiments can measure without direct contact preventing damage to the infant's sensitive skin.

Uesawa et al. use an electromyogram in [2] to detect muscle activity necessary for diaphragm movement. In this work, an infant was placed on a mattress with integrated sheet-type electrodes. The electrodes function as capacitor plates deriving an ECG and an EMG. The sensing electrodes are connected to a boost-trap buffer to provide a high-impedance input leading to high sensitive. The EMG signal of Uesawa et al. was closely correlated (r = 0.96) to the readings of a commercially available belt transducer.

A different approach used in several works is using the patient as the dielectric medium of a plate capacitor:

Teichmann et al. present a sensor in [3], where they used the detuning of a Colpitts oscillator to measure the respiration. Two methods were described. The first system uses a coil as sensing element. Therefore, the patient's respiration changes the inductances of the oscillator by inductive coupling. The second system uses two capacitor plates as sensing elements. In this case, the patient's respiration changes the permittivity between the plates, therefore changing the capacitance and detuning the oscillator. A similar system is described by Kundu et al. in [4]. Kundu et al. use two capacitor plates as well, whereas the plates are made of a conductive textile and integrated in a T-shirt. Both system come with the disadvantage of omnidirectional sensing, especially in the approach of Teichmann et al.

Because the capacitor plates are mounted in the same plane in [3], there is no distinction between an object in front or behind the plates. Kundu et al. partly overcome this issue by placing the patient between the capacitor plates. Still, in the approach of Kundu et al. a body with high permittivity, e.g. an arm, forms a parasitic capacitance to ground and disturbs the respiration measurement.
In this paper, we present a new kind of sensor system in the class of electromagnetic respiration sensors by using two coupled antennas as sensor element. The issue of omnidirectional sensing can be resolved by coupling the antennas and by carefully choosing their radiation pattern. However, in this paper we only investigate the feasibility of a coupled antenna pair as sensing element under ideals conditions. The radiation pattern of the antennas is not investigated to form a well-defined region of interest, yet.

Theory

The theory of operation of our sensor system is based on the change of the propagation speed of an electromagnetic signal when passing through a medium [5]. The propagation speed in the medium mainly depends on the mediums permeability and permittivity. The propagation speed in a medium can be calculated as follows [5]:

$$v = \mathrm{c}_0/\sqrt{\mu_r \varepsilon_r} \tag{1}$$

where c_0 is the speed of light in a vacuum, ε_r and μ_r are the relative permittivity and the relative permeability of the medium.
Because the wavelength λ and the frequency f of the electromagnetic signal are hardly linked with the propagation speed in the medium, one can derive from Eq. (1) the wavelength in the medium[5]:

$$\lambda = \mathrm{v}/\mathrm{f} = c_0/(f\sqrt{\mu_r \varepsilon_r}) \tag{2}$$

This well-known relation is used, e.g. to reduce the mechanical length of antennas [6] at fixed frequencies.
We use this change in wavelength when passing through a dielectric medium to measure changes in dielectric properties. To do so, we detect the phase shift of an electromagnetic wave transmitted by a first antenna when received by a second antenna. By increasing the permittivity of the medium between the two antennas, the wavelength at a constant frequency decreases and the phase difference between transmitted and received signal increases.
A second effect we have to consider when using antennas that are in close connection to a dielectric material is the change in resonance frequency with the dielectric properties. By solving Eq. (2) for the frequency f, one gets:

$$f = c_0/(\lambda\sqrt{\mu_r \varepsilon_r}) \tag{3}$$

The effect arising from Eq. (3) is well-known and can be rather undesired because in common practice it leads to antenna detuning. This is especially the case for body-worn antennas in wearables [7,8], because the body has a rather high permittivity.
In this paper, we investigate the effects described by Eq. (2) and Eq. (3) as two independent means to measure respiration. When considering the human respiration the permittivity of the thorax strongly changes while breathing, because the lung with its rather high permittivity is flooded with air, which has a low permittivity. In the case of Eq. (2), the frequency is fixed and all variables except for the relative permittivity can be considered constant in the application as respiration sensor. The speed of light c_0 is a constant, the frequency is fixed by the user and the permeability of air and most biological material is the same [9]. In case of Eq. (3), the antennas dimension fixes the wavelength λ and we vary the frequency to find the resonance frequency of the coupled antenna pair. So Eq. (3) can be simplified to:

$$f = \mathrm{k}/\sqrt{\varepsilon_r} \tag{4}$$

Where $k = c_0/(\lambda\sqrt{\mu_r})$.
As mentioned earlier, we used a coupled antenna pair as sensor element. Coupled in this case means that the impedance of the first antenna is not only dependent on its self-impedance but also on the impedance of the second antenna, both antennas are linked by a mutual impedance Z_{12}[10]. The coupling of two antennas depends on their distance, where two antennas far from each other can be consider loosely coupled [10]. The mutual impedance Z_{12} of two small electric dipoles of the same length can be modeled in the following way [11]:

$$Z_{12} = \frac{l^2}{4\pi r^2}\left[\underbrace{\frac{1}{j\omega\cdot\varepsilon\cdot r}}_{C} + R_0 + \underbrace{j\omega\cdot\mu\cdot r}_{L}\right] e^{-j2\pi r/\lambda} \tag{5}$$

Where l is the dipole length, r the distance between the dipoles and R_0 the wave resistance in free space.
From Eq. (5) an equivalent circuit consisting of a capacitor, a resistor and an inductor can be derived to model the mutual impedance Z_{12}[11]. Where the capacitor dominates in the antennas near-field and the inductor in the antennas far-field [11]. Because the focus of our sensor systems lies in the detection of permittivity change, it is obvious from the equivalent circuit of the mutual impedance to bring the antennas

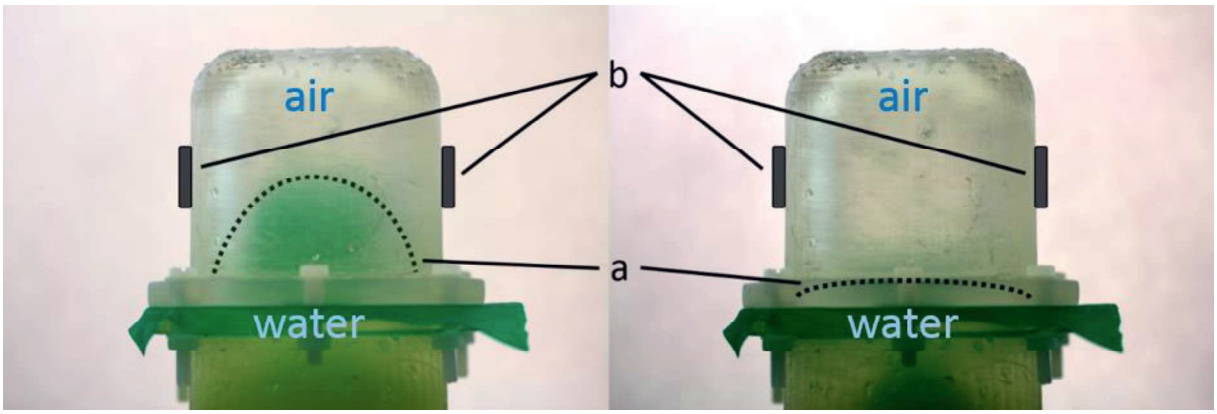


Fig. 1: The lung phantom consists of two compartments. The lower compartment is filled with water and contains a balloon. By inflating the balloon, the water is pushed in the second compartment and an elastic diaphragm (a, dotted line) between the two compartments dilates. This case is shown in the left picture. This models an exhaled state of the lung. By deflating the balloon, the diaphragm contracts. This models an inhaled state and is shown in the right picture. The gray blocks (b) indicate the position at which the antennas were placed.

close together and operate them in or close to their respective near-field.

Material and Methods

As sensor elements, we used two identical UHF antennas (Molex 105262-0002). The antennas have a transmission band starting at 863 MHz, specified by the manufacturer. The antenna is a slightly meandered dipole with a physical length of 79 mm and a width of 10mm. We measured scattering parameters, most importantly the S_{21} parameter describing the transmission between the two antennas using a vector network analyzer. The measured frequency range was 10 MHz to 1200 MHz with a step size of 100 kHz.

As model for an infant's lung, we fabricated a plastic phantom by 3D printing. We printed two compartments using a Form 2 3D-printer with clear resin from Formlabs Inc. The outer dimensions of the phantom are 135x76x65mm with a wall thickness of 2.4mm.

An elastic diaphragm separates the two compartments. The lower compartment contains water to simulate the electrical properties of an infant's thoracic tissue and a single balloon. The inflated balloon pushes the water into the upper compartment by dilating the diaphragm. The upper compartment contains air and in conjunction with the diaphragm serves as the model for the infant's lung, where a dilated diaphragm models an exhaled state of the lung and a diminished diaphragm an inhaled state, see Fig. 1.

The balloon was inflated using a 20 mL syringe held by two clamps mounted on a rod. The second clamp holding the syringe's piston was movable and was used to inflate the balloon with up to 20 mL with a step size of 1 mL.

Results

Before we investigated the effect of the lung volume on the antenna pair, we conducted a series of experiments to characterize the antennas. Here, the emphasis was to evaluate the effect of the coupling between the two antennas in free space. The results are shown in Fig. 2. We characterized each antenna using its S_{11} parameter. In a first setup, we measured S_{11} of each antenna without the other antenna in its vicinity. In this case, the antennas are loosely coupled or uncoupled. The two resonances of the each transmission band were clearly visible in the S_{11} parameter. The two resonances were slightly higher than the center frequencies stated by the manufacturer, probably due to different mounting conditions in our setup. In a second setup both antennas were placed in close vicinity to each other. In this case the antennas are coupled. Due to the coupling the resonances in S_{11} of the first antenna shifted to slightly lower frequencies, but more importantly a third resonance was visible around 290 MHz at a much lower frequency than the other two resonances. This resonance arising from the coupling is especially prominent when looking at the S_{21} parameter of the antenna pair. We focused on the lower frequency resonance, because it arose from coupling the two antennas and is related to the mutual impedance. In addition, as shown in equation (5), the mutual impedance has a retardation term depending on the ratio of antenna distance to wavelength. Therefore, a higher wavelength (i.e. lower frequency) leads

to a more dominant mutual impedance and a higher sensitivity on permittivity changes.

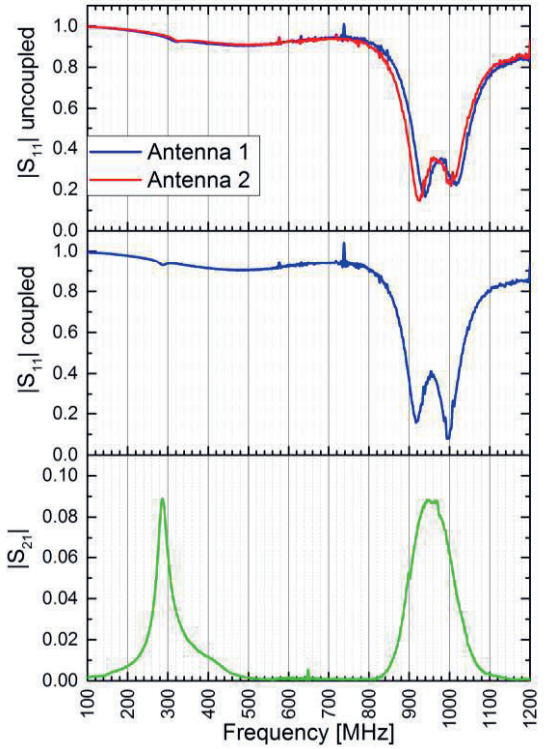


Fig. 2: Coupling of two antennas in free space. The upper diagram shows the S_{11} parameter of the two antennas without the other antenna in its vicinity, respectively. In this case both antennas are uncoupled. Clearly visible are the two resonance frequencies for the antennas two transmission bands between 900 MHz and 1100 MHz. The middle diagram shows the S_{11} parameter of the first antenna, when the second antenna is in its close vicinity. In this case, both antennas are closely coupled. It can be seen that the resonance frequencies above 900 MHz shift to slightly lower frequencies compared to the uncoupled case. Also, a very small resonance around 290 MHz is visible in the coupled case. The lower diagram shows the S_{21} parameter of the coupled antenna pair. The small resonance in S_{11} around 290 MHz is far more prominent when looking at the S_{21} parameter.

The low frequency resonance shifted to even lower frequencies around 240 MHz to 250 MHz, when the antenna pair was mounted on the phantom. The high permittivity of the phantom's water-filled compartment detuned the antenna pairs mutual resonance, as predicted by Eq. (4).

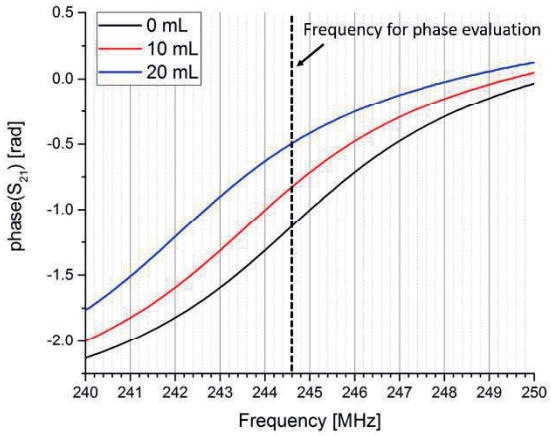


Fig. 3: Phase response for three different water volumes pushed in the upper compartment. The increase of permittivity by pushing water in the upper compartment significantly increases the phase when the evaluation frequency (dotted line) is fixed.

Fig. 3 depicts the phase response for varying water volumes in the upper compartment in the frequency range around the lower resonance frequency. For the evaluation of the phase, we used the resonance frequency when no water was pushed in the upper compartment, as indicated by the dotted line in Fig. 3. It is clearly visible that the permittivity increase leads to a phase increase as predicted by Eq. (2).

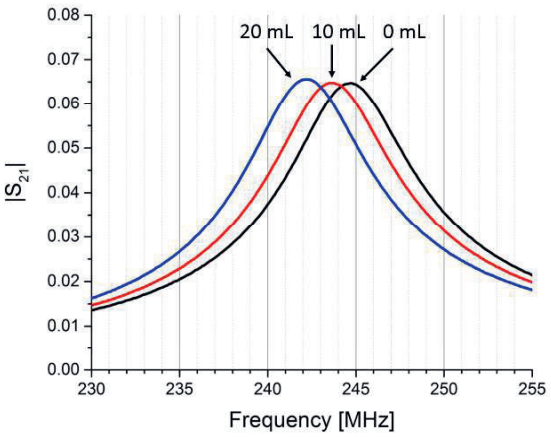


Fig. 4: Resonance in the lower frequency range for three different water volumes pushed in the upper compartments. Clearly visible is the shift of the resonance frequency to lower frequency, when the water volume is increased.

Fig. 4 depicts the resonance in the lower frequency range for varying water volumes in

the upper compartments. It is clearly visible that the resonance frequency shifts to lower frequencies by increasing the water volume. This matches the prediction made by Eq. (4).
The phase at a fixed frequency and the shifted resonant frequency versus the volume of water pushed in the upper compartments are shown in Fig. 5 and Fig. 6, respectively.

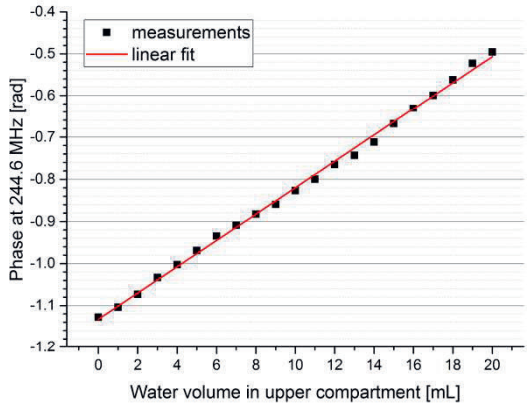


Fig. 5: Phase at constant frequency vs. water volume pushed in the upper compartment. The data was fitted with a linear fit (red curve) yielding R^2=0.9978.

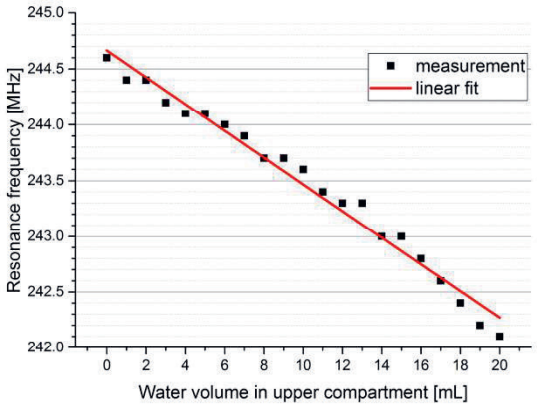


Fig. 6: Resonance frequency vs. water volume pushed in the upper compartment. The data was fitted with a linear fit (red curve) yielding R^2=0.98.

Discussion

We fitted the data shown in Fig. 5 and Fig. 6 linearly to evaluate our results. The linear fits are shown in the figures as red lines. The linear fit of the phase at constant frequency yielded a coefficient of determination of 0.9978. The root-mean-square error is 0.00887 rad, the slope of the linear fit is 0.0312 rad per mL. The root-mean-square error and the slope therefore yield a resolution of 0.28 mL.

The linear fit of the resonance frequency versus the water volume yielded a coefficient of determination of 0.98. The root-mean-square error of the resonance frequency fit is 0.104 MHz, the slope of the fit is -0.12 MHz per mL. The root-mean-square error and slope of the resonance frequency fit therefore yield a resolution of 0.87 mL.
The phase at a constant frequency shows a significantly better resolution compared to the resonance frequency, as can be seen in Fig. 5 and Fig. 6. Using the phase instead of the resonance frequency greatly simplifies a later implementation of our sensor system. The resonance frequency as measurement signal would need a smaller step size of the frequency sweep or a fit algorithm to pinpoint the center frequency. Whereas a higher integration time, i.e. more periods, can simply improve the resolution when using the phase as measurement signal. When using the phase only an initial frequency sweep would be necessary to find a suited measuring frequency.

Conclusion

We presented a new kind of electromagnetic respiration sensor using coupled antennas as sensor element. The sensor was characterized using a 3D printed lung phantom and showed a high sensitivity and linear behavior in the volume range of an infant's lung.
When the sensor is used to assess the absolute volume of a lung, a proper calibration routine probably will be necessary, but this has to be addressed in clinical or pre-clinical trials. However, the proposed sensor might be a viable choice as a trigger for patient–ventilator synchronization in infants and neonates as it can measure dynamic changes in volume with a high temporal and volume resolution. The sensor shows a high sensitivity even at low volume changes and might be well suited to derive breathing patterns in real-time, while being insensitive to leakage. This could strongly reduce the ventilation-induced lung injuries like baro- or volutrauma in the smallest of patients.
Future work will include improvement of the antennas as well as non-invasive measurements on adult humans and small rodents.

References

[1] R. Mahmoud, H. Proquitté, S. Hadhood, G. Schmalisch, Effect of endotracheal tube leakage on respiratory function monitoring: Comparison of three neonatal ventilators, *Journal of Pediatric Intensive Care* 01, 61–69 (2015); doi: 10.3233/PIC-2012-012.
[2] H. Uesawa, T. Takehara, A. Ueno, Non-contact measurements of diaphragm electromyogram, electrocardiogram and respiratory variations

with sheet-type fabric electrodes for neonatal monitoring. In: *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, 25–28. IEEE; 04.03.2018 - 07.03.2018.

[3] D. Teichmann, J. Foussier, J. Jia, S. Leonhardt, M. Walter, Noncontact monitoring of cardiorespiratory activity by electromagnetic coupling, *IEEE Transactions on Biomedical Engineering* 60, 2142–2152 (2013); doi: 10.1109/TBME.2013.2248732.

[4] S. K. Kundu, S. Kumagai, M. Sasaki, A Wearable Capacitive Sensor for Monitoring Human Respiratory Rate, *Japanese Journal of Applied Physics* 52, 04CL05 (2013); doi: 10.7567/JJAP.52.04CL05.

[5] C.-C. Timmermann, *Einführung in Leitungen, Vierpole, Transistormodelle und Simulation mit numerischen und symbolischen CAD/CAE-Systemen: Mit 23 Tabellen*. PROFUND-Verl., Plankstadt; 1997.

[6] K. Mizuno, Y. Daiku, S. Ono, Design of Printed Resonant Antennas for Monolithic-Diode Detectors, *IEEE Transactions on Microwave Theory and Techniques* 25, 470–472 (1977); doi: 10.1109/TMTT.1977.1129136.

[7] J. Buckley, K. G. McCarthy, B. O'Flynn, C. O'Mathuna, The detuning effects of a wrist-worn antenna and design of a custom antenna measurement system. In: *2010 European Wireless Technology Conference (EuWiT 2010)*: *Paris, France, 27 - 28 September 2010 ; [part of the 13th European Microwave Week 2010 (EuMW 2010)*, 261–264. Piscataway, NJ: IEEE; 2010.

[8] P. J. Soh, H. A. Sulaiman, M. A. Othman, Abd. Aziz, Mohamad Zoinol Abidin, M. F. Abd Malek, K. Kamardin, M. K. A. Rahim, N. A. Samsuri, M. E. Jalil, N. A. Ahmad, Planar Textile Antennas Performance Under Wearable and Body Centric Measurements. In: P. J. Soh, H. A. Sulaiman, M. A. Othman, Abd. Aziz, Mohamad Zoinol Abidin, M. F. Abd Malek (eds.). *Theory and Applications of Applied Electromagnetics*, 81–95. Cham: Springer International Publishing; 2016.

[9] R. Pethig, D. B. Kell, The passive electrical properties of biological systems: their significance in physiology, biophysics and biotechnology, *Physics in Medicine and Biology* 32, 933–970 (1987); doi: 10.1088/0031-9155/32/8/001.

[10] R. King, Coupled Antennas and Transmission Lines, *Proceedings of the IRE* 31, 626–640 (1943); doi: 10.1109/JRPROC.1943.233357.

[11] H. Wheeler, The Radiansphere around a Small Antenna, *Proceedings of the IRE* 47, 1325–1331 (1959); doi: 10.1109/JRPROC.1959.287198.

DOI 10.5162/sensoren2019/1.3.3

Hochminiaturisiertes nicht-invasives Messsystem zur Erfassung von Vitalparametern bei Kleinstlebewesen mit drahtloser RFID-/NFC-Ausleseschnittstelle

Patrick Moser[1], Elke Mackensen[1], Fabian Rank[2]
[1]Hochschule Offenburg; Fakultät Elektrotechnik, Medizintechnik und Informatik; Badstraße 24; 77652 Offenburg
[2]VEGA Grieshaber KG; Am Hohenstein 113; 77761 Schiltach

Zusammenfassung

Im Rahmen dieses Beitrags wird ein quasi energieautarkes, nicht-invasives Messsystem für Kleinstlebewesen vorgestellt, das Vitalparameter erfasst und diese in einem FRAM-Speicher bis zum Auslesen abspeichert. Durch eine drahtlose RFID-/NFC-Ausleseschnittstelle kann die erfasste Körpertemperatur und der Puls der letzten Wochen ausgelesen werden. Alle Einstellungen des Messsystems können durch einen geeigneten RFID-Reader für Laptops oder durch Smartphones über die NFC-Schnittstelle geändert werden. Das vollständige Aufladen des nur 3 g leichten und 15 mm x 25 mm großen Messsystems erfolgt durch eine selbstgedruckte RFID-Reader-Antenne in Verbindung mit einem RFID-Reader und benötigt hierzu weniger als 21 Stunden. Bei vollständig aufgeladenem Energiespeicher ist ein Betrieb von 47 Tagen möglich. Dies wird durch ein speziell für das Messsystem konzipiertes Lade- und Powermanagement erreicht. Neben der Auswahl von energiesparenden Komponenten für die Hardware und deren bestmöglichen Nutzung, wurde die Software so optimiert, dass das Programm schnell und stromsparend abgearbeitet wird. Die Erweiterbarkeit und Anpassung wird durch das modulare Konzept auch in anderen Bereichen gewährleistet.

Keywords: nicht-invasives Messsystem, Vitalparameter, Kleinstlebewesen, RFID, gedruckte Elektronik

Einleitung, Stand der Technik und Problemstellung

Die Messung von Vitalparameter, insbesondere Puls und Temperatur, an Kleinstlebewesen wie Tieren [1], [2] aber auch von Frühgeborenen [3] ist von großer Bedeutung. Eine Messung von Vitalparametern kann zu unter-schiedlichsten Zwecken notwendig sein, z.B. im Rahmen von pharmazeutischen Untersuchungen an Mäusen/Ratten [1] oder der Untersuchung der Auswirkungen der Klimaveränderungen auf bestimmte Tierarten [4]. Vergleichbare invasive Messsysteme setzen bei subkutaner Implantation ein Mindestgewicht des Kleinstlebewesens von 19,5 g voraus [5].

Problemstellung

Die Herausforderungen und Anforderungen bei der Entwicklung von Messsystemen zur Erfassung von Vitalparametern an Kleinstlebewesen sind recht vielfältig. Zumeist stehen für das Messsystem ein geringes Gewicht und ein quasi energieautarkes Arbeiten über einen gewissen Zeitraum neben der eigentlichen messtechnischen Funktionalität im Vordergrund. Hinzu kommt oftmals die Forderung, dass das Messsystem ein drahtloses Auslesen der gemessenen Parameter ermöglichen soll. Sowohl in der Forschung als auch bereits kommerziell erhältliche Systeme verfolgen unterschiedliche Ansätze. Insbesondere bei den kommerziell erhältlichen Systemen handelt es sich zumeist um Systeme, die den Tieren implantiert werden müssen, was auf Grund des operativen Eingriffs für kleine Tieren nicht optimal ist. Ein weiterer Nachteil der in der Literatur beschriebenen Systeme ist, dass die Abmessungen und das Gewicht der Messsysteme nicht für Kleinstlebewesen geeignet sind.

Die Anforderungen an das in diesem Beitrag vorgestellte entwickelte Messsystem sind wie folgt: Das Gewicht des Messsystems soll weniger als 3 g betragen, die Abmessung kleiner als 15 mm x 25 mm. Es sollen Körpertemperaturen im Bereich von 0° C und

40° C und ein Puls mit 5 bis maximal 600 Schläge pro Minute gemessen werden können. Das System soll mindestens einmal pro Stunde die Temperatur und den Puls erfassen. Die gemessenen Daten sollen mindesten einmal am Tag drahtlos ausgelesen werden können. Die Erfassung von Messwerten über einen längeren Zeitraum ohne den Wechsel/das Aufladen eines Energiespeichers ist wünschenswert.

Im Folgenden wird nun das Systemkonzept und die Funktionalität des hochminiaturisierten nicht-invasives Messsystems zur Erfassung von Puls und Temperatur vorgestellt, das insbesondere für kleine Vögel/Küken im Ruhezustand [4] konzipiert ist. Das Messsystem eignet sich jedoch auch für andere Kleinstlebewesen.

Systemkonzept und Funktionalität

In Abb. 1 ist das Systemkonzept des nicht-invasiven Messsystems zur Erfassung von Puls und Temperatur bei Kleinstlebewesen dargestellt. Auf Grund der energetischen und platzbedingten Anforderungen wurde als drahtlose Kommunikationsschnittstelle eine aktive Funktechnologie wie Bluetooth ausgeschlossen. Ein vielversprechender Ansatz ist der Einsatz einer RFID- bzw. NFC-Schnittstelle zur drahtlosen Kommunikation. In Betracht gezogen wurde hier zunächst ein hochintegrierter SoC (System on Chip) für biomedizinische Anwendungen [6]. Auf Grund seiner recht großen Chipabmessungen (ohne Gehäuse 1,52 mm x 3,2 mm) und einer Leistungsaufnahme von ca. 600 µW wurde dieser jedoch nicht für diese Anwendung eingesetzt. Nach ersten Designstudien [7] wurde stattdessen ein kommerziell erhältlicher RFID-/NFC-Chip mit Abmessungen von 3 mm x 3 mm ausgewählt, der zudem auch als ungehäuster Chip verfügbar ist. Neben der drahtlosen Kommunikation ermöglicht der RFID-/NFC-Chip zudem das Harvesting von Energie, sobald der Chip sich in einem entsprechenden magnetischen Feld befindet. Sollte die vom RFID-/NFC-Chip bereitgestellte Spannung zu stark belastet werden und somit unzulässig weit absinken, ist keine Kommunikation mehr möglich. Zur Aufbereitung und Begrenzung der vom RFID-/NFC-Chip bereit gestellten Energie wird ein Harvester-Chip eingesetzt. Durch das Festlegen des MPP (Maximum Power Points), wird sichergestellt, dass die 53 kbit/s schnelle Kommunikation während des Ladens nicht gestört wird, und die maximal zur Verfügung stehende Leistung geerntet werden kann. Bei der Auswahl der sonstigen Komponenten der Messelektronik wurde insbesondere auf eine geringe Leistungsaufnahme im aktiven Betrieb geachtet. Erreicht wird dies beispielsweise durch einen Mikrocontroller mit FRAM-Speicher (Ferroelectric Random Access Memory). Die Pulsmessung basiert auf einem Sensor mit integriertem Infrarotemitter und Photodiode, so dass nach dem Prinzip der Photoplethysmographie der Puls erfasst werden kann. Der zum Einsatz kommende Temperatursensor, mit lediglich zwei elektronischen Anschlüssen, gibt bei Aktivierung entsprechend der Temperatur Strom-Pulse aus. Das Messsystem wird über einen wieder aufladbaren Energiespeicher mit Energie versorgt. Dieser kann durch die Energy-Harvesting-Funktion des RFID-/NFC-Chips wieder aufgeladen werden, sobald sich das System in einem magnetischen Feld befindet.

Abb. 2 zeigt eine beispielhafte Nutzung des Messsystems an einem Vogelküken. Das Nest des Vogels wird über der RFID-Reader-Antenne positioniert. Wie in [7] beschrieben, bewegt sich das Küken kaum, weshalb der 3 g leichte Sensor auf dem Rücken des selbst nur 15 g leichten Vogels positioniert werden kann. Die mit einem flexiblen Leitungsteil abgesetzten Temperatur- und Pulssensoren können an den entsprechenden Positionen des Kleinstlebewesens angebracht werden.

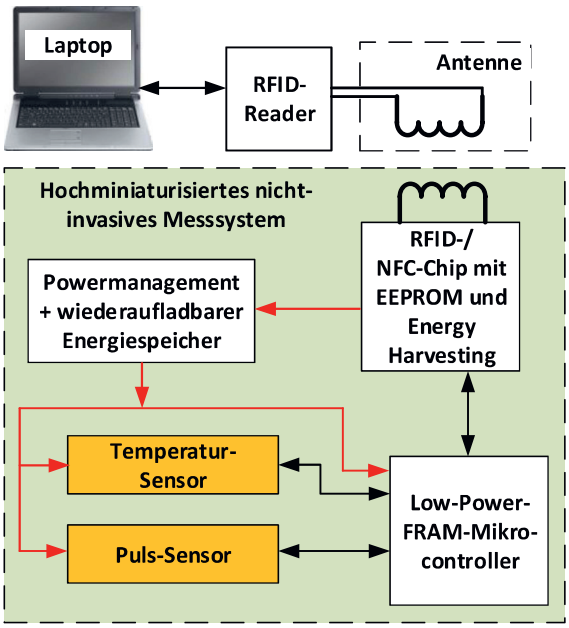


Abb. 1: Blockschaltbild des nicht invasiven Messsystems zur Erfassung von Puls und Temperatur an Kleinstlebewesen

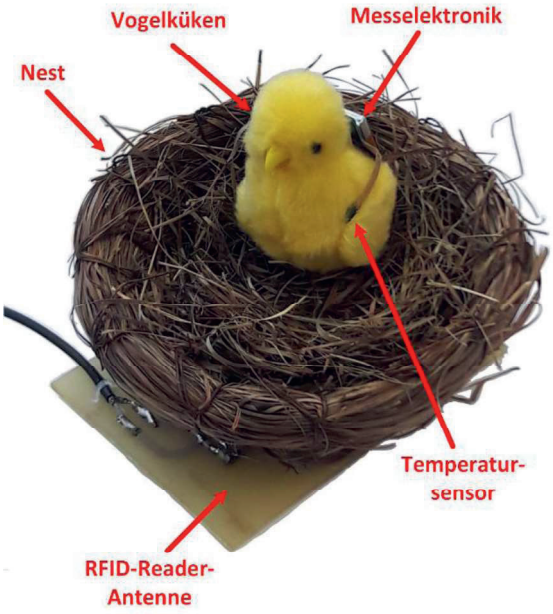


Abb. 2: Beispielhafte Nutzung des nicht invasiven Messsystems bei einem Vogelküken

RFID-Reader mit gedruckter Antenne

Mit zum Gesamtsystem gehört eine selbst entworfene Antenneneinheit, die unter dem Vogel/Tier platziert ist. Diese Antenneneinheit kann mit einem entsprechenden RFID-/NFC-Lesegerät über eine SMA-Steckverbindung verbunden werden, um die Messdaten drahtlos per Laptop auszulesen, oder eine Aufladung des Energiespeichers durchzuführen. Alternativ kann in Betracht gezogen werden, die Daten per Smartphone mit der NFC-Schnittstelle auszulesen. Als RFID-Reader wurde der „ID ISC.MR102-USB“ von Feig Electronic ausgewählt [8], welcher durch eine selbst entworfene grafische Benutzeroberfläche gesteuert wird. Mit dieser Benutzeroberfläche kann das Messsystem konfiguriert werden. Dies betrifft die Auswahl des gewünschten Messintervalls, welche Messgröße gemessen werden soll, sowie verschiedene Einstellungen des Pulssensors.

Eine eigene Entwicklung der Antennen war notwendig, da keine passende, kommerziell erhältliche Antenne nutzbar war. Abb. 3 zeigt die selbstentwickelte Antenne, mit einer Abmessung von 110 mm x 90 mm. Bei der Antenne handelt es sich um eine mit dem Voltera V-One-Platinendrucker (siehe https://www.voltera.io/) gedruckte Antenne auf 2 mm dickem FR4-Platinenmaterial. Dies ermöglicht eine schnelle und kostengünstige Herstellung von wenigen Antennen, welche an die Anforderungen angepasst sind. Mithilfe eines Schutzlackes werden die Leiterbahnen und Komponenten vor Umwelteinflüssen, wie Feuchtigkeit geschützt. Durch Verwendung von Folien oder dünnem FR4-Platinenmaterial kann mit dem Platinendrucker auch eine flexible Antenne produziert werden, welche sich optimal an die räumlichen Gegebenheiten anpasst.

Die in Abb. 4 gezeigte Messaufnahme der RFID-Reader-Antenne zeigt deren Impedanzverlauf von 13 MHz bis 14 MHz. Der angestrebte Impedanzwert von 50 Ohm bei der Frequenz von 13,56 MHz wurde mit 51 Ohm nicht ganz erreicht. Abb. 5 zeigt die Reflektion der Antenne im Bereich von 13 MHz bis 14 MHz. Je geringer der Wert ist, desto besser überträgt die Antenne die Energie zum Messsystem. Auch hier wurde der beste Wert nicht bei der gewünschten Resonanzfrequenz von 13,56 MHz, sondern bei 13,59 MHz erreicht. Zu beachten ist, dass durch metallische Strukturen, wie es beispielsweise das Messsystem besitzt, die ermittelten Werte der RFID-Reader-Antenne verändert werden. Wird das Messsystem über der RFID-Reader-Antenne positioniert, verschiebt sich die Impedanz und die Resonanz auf die gewünschten Werte, wodurch das System optimal arbeitet. Weitere in der Umgebung vorhandene Messsysteme verändern die Parameter kaum, wodurch theoretisch mehr als ein Messsystem gleichzeitig durch einen RFID-Readern ausgelesen und geladen werden könnten.

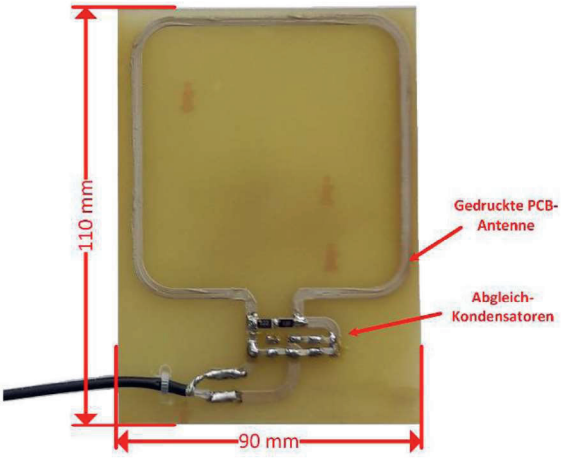


Abb. 3: RFID-Reader mit gedruckter Antenne

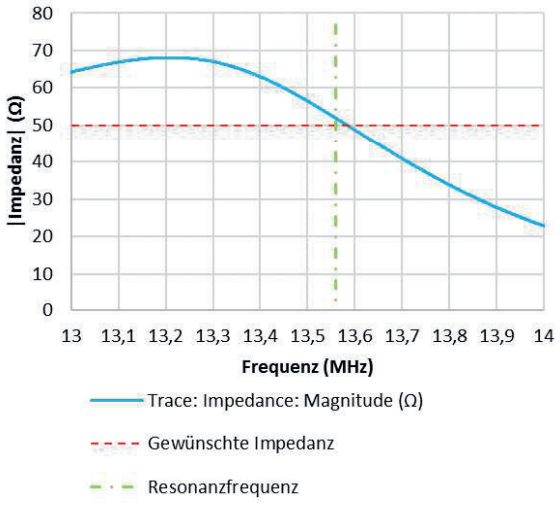


Abb. 4: Impedanz der RFID-Reader-Antenne im Frequenzbereich 13 bis 14 MHz

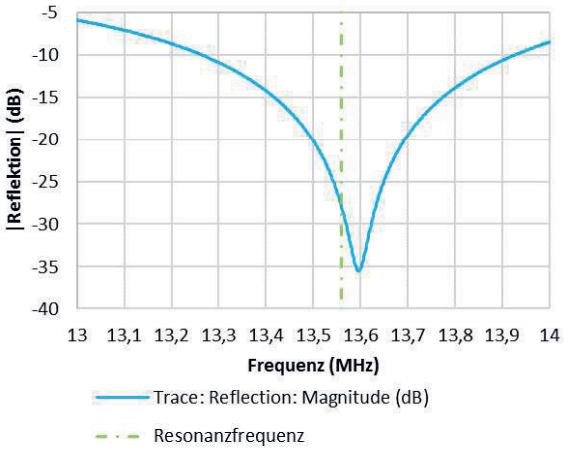


Abb. 5: Reflektion der RFID-Reader-Antenne im Frequenzbereich 13 bis 14 MHz

Hardware des nicht-invasiven Messsystems und Powermanagement-Strategien

Abb. 6 zeigt die Hardware des nicht-invasiven Messsystems, die so konstruiert ist, dass der RFID-/NFC-Chip, der Harvester-Chip mit dem Powermanagement, der Energiespeicher und der Mikrocontroller auf einem 4-lagigen Platinenteil der Abmessung 15 mm x 25 mm platziert sind. Dieser 1 mm dicke Platinenteil wird am Beispiel des Vogels auf dessen Rücken platziert (siehe Abb. 2). Aus diesem Platinenteil ragen seitlich flexible Leitungsteile heraus, an dessen Ende sich jeweils der Temperatur- und Pulssensor befinden. Durch die flexiblen Leitungsteile können die Sensoren an die entsprechenden Körperstellen geführt werden, wo die Messwerte erfasst werden müssen. Die Platine enthält für die Datenübertragung per RFID/NFC eine integrierte PCB-Antenne. Durch die Integration in die Platine wird kein externes Bauteil benötigt, wodurch Gewicht und Platz eingespart wird. Die Antenne besteht aus vier Windungen, welche sich auf alle Lagen der Platine aufteilen. Die Elektronik auf der Platine wird durch einen Schutzlack vor Feuchtigkeit geschützt.

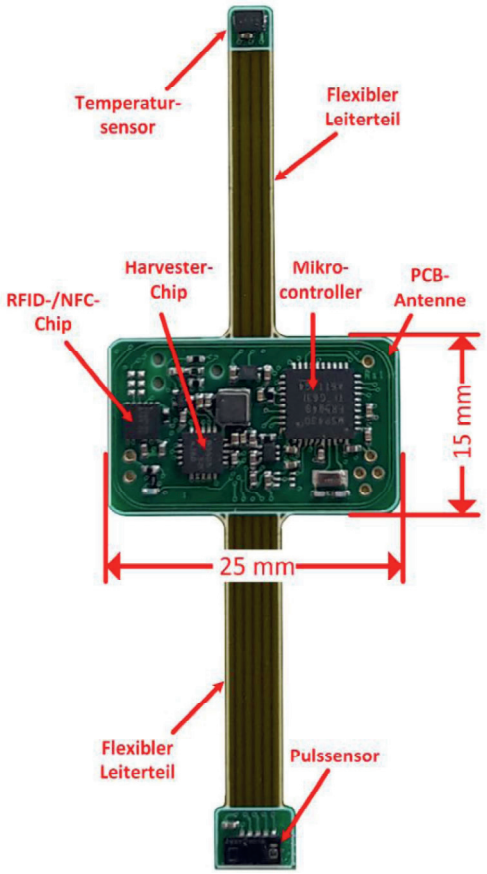


Abb. 6: Aufbau der Messelektronik mit flexiblen Leiteranschlüssen für die Sensoren

Die Hardware und Arbeitsweise des Messsystems sind so konzipiert, dass das Gesamtsystem extrem stromsparend arbeitet. Details hierzu werden im Folgenden beschrieben:

- Durch die Verwendung eines Low-Power-FRAM-Mikrocontrollers, der die überwiegende Zeit im Sleep-Mode arbeitet, kann der Ruhestrom des Systems auf wenige µA reduziert werden.
- Durch das Powermanagement, welches durch den Mikrocontroller gesteuert wird, werden alle nicht benötigten Komponenten von der Spannungsversorgung getrennt. Hierdurch entfallen eventuell auftretende Leckströme, was zu einer weiteren Reduzierung des Ruhestromes und somit auch der durchschnittlichen Stromaufnahme führt.

- Trotz der am RFID-/NFC-Chip abgeschalteten Versorgungsspannung, wird durch ihn am Mikrocontroller ein Interrupt ausgelöst, sobald sich das Messsystem in einem entsprechenden RFID-/NFC-Feld befindet. Anschließend kann der RFID/NFC-Chip durch den Mikrocontroller aktiviert werden, wodurch er seine volle Funktionalität bezüglich Kommunikation und Bereitstellen von Energie erlangt. Dies ermöglicht das sofortige Auslesen der ermittelten Messdaten, die im nichtflüchtigen Speicher des Mikrocontrollers abgelegt werden. Bei einem Messintervall von einer Messung pro Stunde, können für einen Zeitraum von über einer Woche Messwerte gespeichert werden.
- Um das System für den genannten Zeitraum mit Energie zu versorgen, wird die vom RFID-/NFC-Chip bereitgestellte Energie in einen 259 mWh großen Lithium-Polymer-Akkumulator gespeichert. Dies geschieht durch einen Harvester-Chip, welcher die benötigte Ladespannung für den Energiespeicher erzeugt und zusätzlich das Laden überwacht.
- Zur Ermittlung der Temperatur des Lebewesens wird die Hauttemperatur gemessen. Hierzu muss der Sensor direkten Kontakt zur Haut besitzen. Der Temperatursensor kommuniziert über Strom-Pulse mit dem Mikrocontroller. Die hierbei entstehenden Analogenwerte werden in einem Komparator mit Schmitt-Trigger digitalisiert und im Mikrocontroller gezählt. Dieser Vorgang wurde so konzipiert, dass keine Rechenzeit benötigt wird, wodurch Energie eingespart wird.
- Um den Puls des Lebewesens nicht-invasiv zu ermitteln, wird ein Pulssensor eingesetzt, der nach dem Photo-plethaysmographie-Verfahren arbeitet. Bei diesem Verfahren leuchtet ein Infrarotemitter auf eine Arterie. Durch die sich periodisch ändernde Absorptionsrate, welche bei einem Herzzyklus entsteht, kann auf den Puls rückgeschlossen werden. Der Pulssensor liefert bereits digitale Werte, welche im Mikrocontroller zwischen der Abtastrate des Sensors verarbeitet werden. Hierdurch entfällt die Notwendigkeit eines großen Speichers im Mikrocontroller, um die ermittelten Daten zwischen zu speichern. Um eine präzise Aussage über den Puls treffen zu können, wird dieser eine Minute überwacht. Der Sensor kann für die jeweiligen physikalischen Gegebenheiten, wie Dicke der Haut oder des Gewebes, sowie die zu erwartenden Herzfrequenz des Lebewesens angepasst werden. Hierzu lassen sich die Anzahl der Messungen pro Sekunde, Leuchtintensität und Modulationsfrequenz des integrierten Infrarotemitters einstellen.
- Die durch die Messung berechneten Vitalparameter werden im FRAM-Speicher des Mikrocontrollers abgelegt, bis diese dann per RFID/NFC ausgelesen werden.

Energetische Betrachtung des Messsystems

In Abb. 7 ist die Leistungsaufnahme des Messsystems für eine Messung des Pulses und der Temperatur in einem Messintervall von einer Stunde nicht maßstabsgetreu grafisch dargestellt. Das Messsystem hat bei aktivem Betrieb und der geforderten Messdaten-erfassung einmal pro Stunde, eine durchschnittliche Leistungsaufnahme von ca. 115 µW. Mit dem zum Einsatz kommenden hochminiaturisierten, wieder aufladbarem Energiespeicher, der über eine Kapazität von 259 mWh verfügt, und dem man ca. 50 % seiner Energie bis zur Wiederaufladung entnehmen kann, kann das System ca. 47 Tage die Datenerfassung ohne Notwendigkeit der Energieaufladung erfolgen. Dies ist möglich, da sich das System die meiste Zeit im Sleep-Mode befindet. In diesem Zustand hat das Messsystem lediglich eine Leistungsaufnahme von 37 µW. Für Berechnungen und Initialisierungen arbeitet der Mikrocontroller mit einem Takt von 8 MHz. Hierbei wird für eine kurze Zeit vor und nach den Messungen eine Leistung von 3256 µW benötigt. Für die Messung der Temperatur sind es 370 µW, was die Kommunikation des Temperatursensors mit dem Mikrocontroller bereits beinhaltet. Der Mikrocontroller selbst benötigt während der Kommunikation kaum Energie, da er sich im Sleep-Mode befindet. Der Pulssensor benötigt bei sieben Messungen pro Sekunde, wobei der Infrarotemitter für kurze Zeit eine Stromaufnahme von 100 mA besitzt, was einer Leistungsaufnahme von 1184 µW entspricht. Bei dieser Messung verarbeitet der Mikrocontroller die erhaltenen Messwerte sofort, was zu einer gesamten Leistungsaufnahme von 4440 µW führt. Wird die Zeit zwischen den Temperatur- und Pulsmessungen vergrößert, sinkt die durchschnittliche Leistungsaufnahme, wodurch auch eine deutlich größere Betriebsdauer erreicht wird.

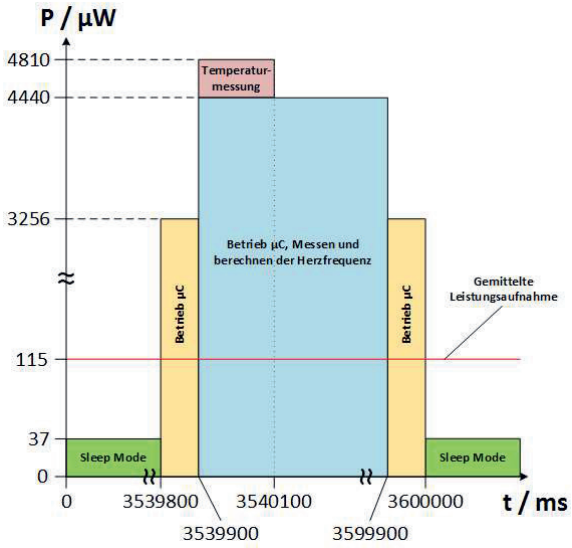


Abb. 7: Leistungsaufnahme des Messsystems für eine Messung des Pulses und der Temperatur in einem Messintervall von einer Stunde

Der Energiespeicher der Messelektronik wird durch das von der RFID-Reader-Antenne erzeugte magnetische Feld aufgeladen, da der im Messsystem eingesetzte RFID-/NFC-Chip das Energy-Harvesting der Energie aus dem magnetischen Feld ermöglicht. In Abb. 8 ist aufgezeigt, welche Leistung aus dem Magnetfeld in Abhängigkeit der Ausgangsspannung am RFID-/NFC-Chip für unterschiedliche Abstände zwischen der RFID-Reader-Antenne und der PCB-Messsystemantenne geerntet werden kann. Die PCB-Empfangsantenne ist parallel und mittig über der selbstgedruckten RFID-Reader-Antenne ausgerichtet. Zu beachten ist, dass bei steigendem Abstand die mögliche Leistung, mit der der Energiespeicher geladen werden kann, abnimmt. Ab einem Abstand der Antennen von mehr als 100 mm ist das Laden, sowie die Kommunikation nicht mehr möglich. Die vom RFID-Reader bereitgestellte Leistung kann nicht eingestellt werden und beträgt laut Datenblatt [8] maximal 1,2 W. Die Leistung ist nur dann maximal, wenn die beiden Antennen parallel zueinander ausgerichtet sind. Mit steigendem Winkel reduziert sich die übertragene Leistung. Neben metallischen Materialen können auch weitere Materialien einen Einfluss auf die Leistung haben.

Zur optimalen Leistungsentnahme aus dem RFID-Feld wird der MPP des Harvester-Chips auf 65 % der möglichen Leerlaufspannung gesetzt. Dies entspricht circa 2,3 V, was bei den beiden in Abb. 8 gezeigten Abstände der Antenne dem Leistungsmaximum entspricht. Hierdurch kann unabhängig der Abstände der 259 mWh große Energiespeicher schnellstmöglich geladen werden. Bei einem Abstand der Antennen von 30 mm entspricht dies 14800 µW. Wird von einem Wirkungsgrad des Harvester-Chips von 85 % ausgegangen, kann der Energiespeicher in unter 21 Stunden komplett geladen werden. Dies sollte vor dem Anbringen an das Kleinstlebewesen geschehen. Der Energiespeicher kann beim Auslesen der Messwerte nachgeladen werden. Bei einem täglichen Nachladen reicht – bei optimaler Ausrichtung der Antennen zueinander – eine Zeit von weniger als 15 Minuten.

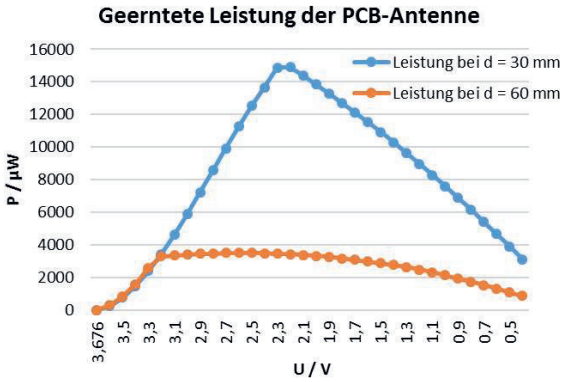


Abb. 8: Geerntete Leistung an der PCB-Empfangsantenne des Messsystems bei unterschiedlichen Abständen in der Mitte der selbstgedruckten RFID-Reader-Antenne

Zusammenfassung und Ausblick

Das in diesem Beitrag vorgestellte nicht-invasive Messsystem ermöglicht es, die Vitalparameter von Kleinstlebewesen, wie z.B. Vogelküken für einen gewissen Zeitraum energieautark zu überwachen. Die wichtigsten Kenndaten des Messsystems sind im Folgenden zusammenfassend aufgeführt:

- Mit dem Messsystem lassen sich Körpertemperaturen im Bereich von 0° C bis 40° C überwachen.
- Die optische Ermittlung des Pulses nach dem Photoplethaysmographie-Verfahren ermöglicht eine Pulsmessung von 5 bis maximal 600 Schläge pro Minute.
- Die Hauptplatine des Messsystems besitzt eine Abmessungen 15 mm x 25 mm und einen über RFID/NFC wieder aufladbaren Energiespeicher. Per RFID/NFC können die gemessenen Vitalparameter drahtlos mit 53 Kbit/s ausgelesen werden. Dies kann über einen speziellen RFID-Reader oder ein Smartphone mit NFC-Schnittstelle geschehen.

- Die genutzten Sensoren sind durch flexible Leitungsteilen von der Hauptplatine abgesetzt, um diese an die entsprechenden Körperstellen positionieren zu können.
- Das System ist momentan dafür ausgelegt mindestens einmal pro Stunde die Temperatur und den Puls zu erfassen. Für das Messintervall von einer Stunde und mit dem gewählten miniaturisierten Energiespeicher beträgt die Leistungsaufnahme des Messsystems 115 µW, wodurch ein energieautarkes Arbeiten von 47 Tagen erreicht wird.
- Bei täglichem Laden des Energiespeichers von nur 15 Minuten über das Magnetfeld eines RFID-Readers, kann das Messsystem 24 Stunden energieautark betrieben werden.
- Das Messsystem, welchen am Kleinstlebewesen angebracht wird, wiegt lediglich 3 g.

Das hier vorgestellte Messsystem kann prinzipiell auch für andere Anwendungsbereiche eingesetzt und angepasst werden. Neben der Überwachung von Kleinstlebewesen wie Tieren, könnte z.B. durch leichte Anpassung der Messelektronik ein Wearable zur Überwachung von Säuglingen entstehen. Durch den modularen und somit flexiblen Aufbau der Messelektronik könnte zudem das System durch andere Sensoren ausgestattet werden, wodurch andere Messgrößen erfasst werden könnten. Beispielsweise denkbar wäre die Überwachung von Temperatur und Luftfeuchtigkeit beim Transport von Waren, um eine unterbrochene Kühlkette oder falsche Lagerung von Waren zu erkennen.

Literaturnachweis

[1] Volk, T.: Semipassives Sensorimplantat mit RFID ähnlicher Schnittstelle zur Überwachung der Vitalfunktionen von Kleintieren. Dissertation, Hochschule Offenburg & Universität Freiburg, 2014.

[2] McCafferty, D.J; Gallon, S.; Nord, A. (2015): Challenges of measuring body temperatures of free-ranging birds and mammals. In. Animal Biotelemetry (2015). pages 1-10. DOI 10.1186/s40317-015-0075-2

[3] Flügel, M.: Kontinuierliche Messung der Körperkerntemperatur bei Frühgeborenen mittels eines nicht-invasiven Sensors und Korrelation mit der durch Nah-Infrarot-Spektroskopie gemessenen regionalen zerebralen Sauerstoffsättigung. Dissertation, Universität Lübeck, 2013.

[4] Kuepper, N.D.; et. al.: Facultative hypothermia as a survival strategy during snowstorm induced food shortages in Antarctic storm-petrel chicks. In: Comparative Biochemistry and Physiology Part A. Molecular & Integrative Physiology, Volume 224, October 2018, Pages 76-83; DOI 10.1016/j.cbpa.2018.06.018

[5] DSI: Guide to DSI´s - Telemetry Devices, Extra-Small Implants, 2019

[6] Bhattacharyya, M.; Dusch, B.; Jansen, D.; Mackensen, E.: Design and Verification of a Mixed-Signal SoC for Biomedical Applications. In: Proceeding of the 54. MPC-Workshop. Ulm, Juli 2015. Pages 43-48. ISSN 1868-9221

[7] Volk, F.: Entwicklung einer drahtlosen hochminiaturisierten Elektronik zur Untersuchung von antarktischen Sturmschwalbenküken. Masterthesis, Hochschule Offenburg, 2016.

[8] Feig Electronic: HF Mid Range Reader ID ISC.MR102, Datasheet, Dezember 2017

DOI 10.5162/sensoren2019/1.3.4

Study of Safe Sensor Configurations in Human Robot Collaboration

W.-J. Baek[1], S. Huang[1], C. Ledermann[1], R. Kirsten[2], T. Kroeger[1]

[1]Karlsruhe Institute of Technology (KIT), Institute of Anthropomatics and Robotics – Intelligent Process Automation and Robotics (IAR-IPR), Engler-Bunte-Ring 8, 76131 Karlsruhe, Germany
[2]ABB AG Forschungszentrum, Wallstadter Straße 59, 68526 Ladenburg, Germany

Abstract

In this contribution, we present two approaches yielding an automatic generation of the optimal sensor configuration in a defined human robot collaboration scenario with respect to the compliance of the human operator's safety and the minimum floor space consumption in the working cell. Based on a simulation setup created with the software Gazebo 7, we develop two algorithms providing the combination and position information of such devices. Following the derived procedures, we focus not only on the placement of safety elements, namely the laser scanner, fences and safety mats but also the influence of the corresponding parameters on the obtained results.

Keywords: Sensor Placement, Safety, Human Robot Collaboration.

I Introduction

The deployment of robots in industrial applications has led to promising results in recent years, in particular in manufacturing facilities. Typically, robots are made responsible for simple tasks including welding, picking and placing or even packaging and labeling products. One of the most important aspects for human operators involved in such processes is the compliance of safety, which is basically defined by norms provided by the International Organization of Standardization (ISO). In order to realize a safe working environment, manufacturing facilities make use of safety elements like mechanical barrier elements or sensors. Doing so, the working cell will be monitored and human operators are supposed to be alerted in time when entering a critical zone around the robot. Robots on the other hand are supposed to slow down or stop their movement accordingly.
One crucial issue in this context is the optimal placement of safety devices in the human robot working cell. Within the scope of this study, two approaches yielding the automatic generation of localization points for safety elements have been developed. Based on the software Gazebo, a simulation environment representing a human robot collaboration scenario has been established. By defining a certain area around the industrial robot, we investigate how sensors or mechanical barrier elements have to be positioned in order to slow down, stop the robot movement or prevent the human being from entering the working area of the robot, thus preventing the human being from hazards. The distinction of critical and safe situations is achieved on norms defined by the International Organization for Standardization (ISO).

II State of the Art

In previous works, such as (Andreas Krause, 2006), (Emily Clark, 2018), (Zainab M. Ismail, 2018), the sensor placement issue has been dealt in the context of different applications as structural health monitoring and computer communication networks. In particular, (Andreas Krause, 2006) presents an algorithm providing sensor localization points based on probabilistic models aiming cost-effective and informative selections.
Our contribution focuses on the determination of a combination of devices yielding the compliance of safety for the human operator. To date, there exist different norms and measures defined by the International Organization of Standardization, which are usually considered as criteria for safety. Apart from this fact the movement of the robot in human robot cooperation, is distinguished in three categories according by IEC 50204-1. ISO norm 13855 has been stated in 2010 and equals a simple relation which allows the determination of a safety distance S between

the actuators for a known velocity of the approaching human operator v and different influences addressed by the constant c:

$$S = vt + c. \quad (1)$$

A more specified formalization is presented by norm ISO/TS 15066 and a publication Marvel et al., 2016 (Jeremy A. Marvel, 2017), where v_H refers to the velocity of the human operator, v_R to such of the robot and v_s considers the velocity of sensors:

$$S(t_0) \geq \int_{\tau=0}^{\tau=t_0+T_R+T_S} v_H(\tau)d\tau + \int_{\tau=t_0}^{\tau=t_0+T_R} v_R(\tau)d\tau$$

$$+ \int_{\tau=t_0+T_R}^{\tau=t_0+T_R+T_S} v_S(\tau)d\tau + C + Z_S + Z_R \quad (2)$$

Here, T_R denotes the time of robot manipulator needed to respond, T_S the time required to bring the robot to a stop, C the intrusion distance safety margin based on the expected reach of the human, Z_s the sensor position uncertainty and Z_R the robot's position uncertainty. In this study, sensors are assumed to be positioned and not moved afterwards. In addition, uncertainties are not part of investigation. Thus, the last term vanishes. Furthermore, the approaches apply to cases where the robot and the human operator move with constant velocities, i.e. it is assumed that the actuators are not subject to any forces which might occur in the setup. Consequently, equation (2) becomes equation (1).
In order to distinguish safe from critical situations, we thus refer to equation (1) in the presented approaches.

III Simulation procedure

Investigations within the scope of this contribution have been performed based on simulations conducted with the open source software ROS (Robot Operating System) kinetic and Gazebo 7. Combining ROS with Gazebo enabled the consideration of dynamic states of the actuators. In the scope of this study, ROS kinetic has been applied on the simulation software Gazebo. The considered setup including two tables is shown in figure 1.

Simulation steps

The focus of the simulations performed lies in the determination of the optimal sensor placement with respect to two constraints in the given order:

1) Safety according to norm ISO 15066

2) Minimum consumed floor space by devices

In the following, the cycle of one simulation run will be explained.

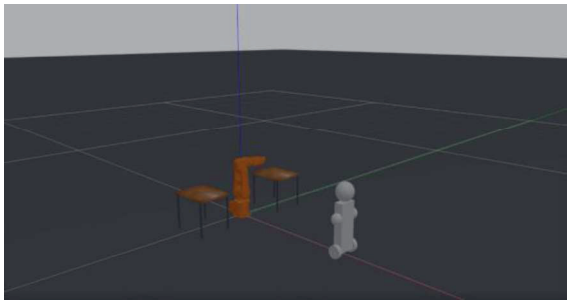


Figure 1: *Screenshot of the simulation scenario in Gazebo 7. The simulation setup consists of two tables and the robot model ABB-IRB120 moving its arms with constant velocity. The human model performs random trajectories within the working cell.*

a) Placement of safety element:
The first step in the simulation procedure is given by the placement of the safety device. In the first loop, the element type and its position are selected randomly. Subsequent steps depend on the approach applied. Within the frame of this contribution, two methods have been considered: While the first method relies on static parameters of the safety devices, the second algorithm allows the variation of the measures. Both approaches will be explained in more detail in IV.

b) Classification of situations in safe and critical: Within the simulation procedure, different scenarios are expected to occur which can be classified in *safe* and *critical*. In first place, this classification is based on the different radii around the robot model representing different zones with respect to the human operator's safety as shown in figure 2. Once the human operator approaches the robot and enters one of the circles, the situation is categorized as critical. In our contribution, we assume the outermost circle as the measure for safety. Doing so, we argue that the worst case scenarios are covered in our investigations.

c) Generation of sensor localization points:
The result obtained in the previous step equals a binary output distinguishing safe from critical situations. In case the binary output links to the occurrence of a critical situation, the second

algorithm is initialized. This algorithm is responsible for the generation of the sensor position. Once generated, the information is forwarded to the simulation interface in order to visualize the sensor type and its location.

d) Termination of simulation: If the binary output of the external observer in c) categorizes the situation as safe, the simulation is terminated. The location information as well as the type of the sensors placed in the simulation and the corresponding floor space consumption is stored in an array.

e) Selection of the combination yielding the minimum floor space:
After termination, the combination of safety elements including the corresponding consumed floor space is stored in an array. The entries of this array are sorted with respect to the consumed floor space values for each combination of devices.

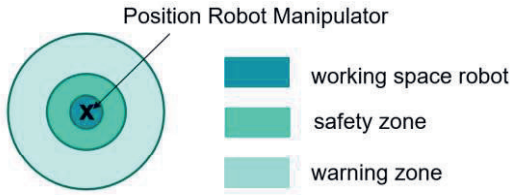


***Figure 2:** Depiction of the warning zones around the robot. Considering the position of the robot in the center, the zones around it can be split into three different zones as shown above.*

Considered Devices
As mentioned above, three different safety devices have been subject to investigations. Besides laser scanners and safety mats, we make use of one mechanical barrier element to investigate its influence on the result. If not stated specifically, the characteristics of such are set according to the following table.

Table 1: *Overview of considered devices*

	Laser scanner	**Safety Mat**	**Fence Element**
Measures	Aperture angle: 120 °	Width: 0.5 m	Thickness: 0.1 m
	Range: 7.0 m	Length: 0.5 m	Length: 1.0 m

IV Approaches

In both algorithms presented below, above devices have been taken into account. However, the methods are distinguished by the fact that one relies only on the localization issue taking the measures of the devices as fixed values while the second algorithm considers different values for each parameter, thereby allowing for the investigation of the measures' influences on the recommended combination of devices.

Customized Approach from scratch
The first approach has been developed from scratch based on three main assumptions, which are defined in the following.

Assumption 1: Sensor characteristics are given and set according to table 1. The specifications are known and do not have to be evaluated. Therefore, device parameters are treated as constant parameters in this approach.

Assumption 2: Once a sensor or barrier element is placed, the corresponding area is covered immediately, i.e. the latency time of the sensor $t_{s,latency}$ is neglected assuming:

$$t_{s,latency} = 0s. \tag{3}$$

Assumption 3: If a sensor or barrier element is placed, it covers and detects the relevant area. The direction of the coverage is therefore assumed to be positioned ideally.

The classification of safe and critical situations is executed by an external code, the so-called classification code, which evaluates the situation as follows: By means of the sensor characteristics, which are basically the radius of the range R_S covered by it and its aperture angle φ_S, the covered area A_S can be determined with the following relation:

$$A_S = \int_0^{R_S} \int_0^{\varphi_S} r d\varphi dr. \tag{4}$$

On the other hand, the area corresponding to the outermost zone of the robot manipulator A_R can be simply calculated via

$$A_R = \pi \cdot r_{w,\max}^2 \tag{5}$$

where $r_{w,max}$ denotes the radius of the warning zone. Here, the worst case is taken into account, i.e. the widest range which can be

reached by the robot and therefore the largest zone is considered.
Hence, the warning zone corresponding to the robot can be assumed to be static.
As the exact position of the sensor is known, it is possible to estimate if there is a cut set of A_S and A_R by means of equation (6).

$$A_S \cap A_R \neq 0 \quad (6)$$

i.e. if the detection area of the sensor covers the outer circle of the robot's warning zone as depicted in figure 2.
In order to simplify this issue, it is possible to only consider the circle of the warning zone. Doing so, the classification still yields an output corresponding to the worst case scenario. Hence, the prevention of the human operator from hazards due to the robot is assumed to be fulfilled with a high probability.
One possible approach to face this issue is the representation of the warning zone by means of equidistant discrete points. By verifying how many of such points are covered by the detection area of the sensor, it can be estimated if the placement of additional safety element is required, that is, if the loop will be terminated. The stepwise addition of sensors is carried out until the coverage of the whole circle is reached. Once all points have been covered by sensors, the sensor placement loop is terminated.
In contrast to sensors, mechanical barrier elements do not offer a certain aperture angle. Hence, such barrier elements are treated as lines, projecting the endpoint to the circle and thus determining the area covered of the warning zone- In terms of the location of the sensors to be placed, a method which contains a strict structure is applied in this approach. In general, the type of the safety element is chosen randomly at the beginning of each loop. Also, the position of the first safety element is generated randomly. If the corresponding area A_S does not cover all points, i.e. if it does not fulfill the conditions of safety, the start of an additional loop is initialized. The second safety element then is localized by reflecting the position of the first element on a line.

Before placing the safety device on the calculated position, it is checked if a previous element already has been positioned on the same point. If this applies, two further lines which are perpendicular to each other are defined and act as mirror lines.
In case the safety requirements are not fulfilled after having completed the reflection process on each line, additional sensors are placed by calculating half of the distance to the neighboring element.
Once a simulation run is terminated, the corresponding data, namely the number and types of safety elements is stored in an array. In addition, the floor space covered by all sensors is calculated which can be completed because sensor parameters are fully known. Doing so, the combination yielding the minimal floor space can be determined by comparing the calculated values.
The comparison of the obtained results with regard to the consumed floor space can be performed either automatically or manually by the user, depending on the number of sensors to be considered.
In the presented approach, the number of considered sensor types is limited.
Thus, there exist a fixed number of possible solutions. Presented approaches therefore are recommended for applications where only small number of safety elements is taken into account.

Approach based on MatLab
In contrast to the previous method, the second approach relies on the usage of the software MatLab. Using this tool, the workspace of the human operator and the robot was split in different grids of the size of 0.2 m x 0.2 m. By considering different assumptions, the device parameters, such as the aperture angle, width and lengths have been varied to determine the most reasonable combination of safety elements.
Apart from that, comparisons between the results obtained in approach 1 allows for the determination of the dimensions of employed devices with respect to the minimum floor space coverage. The second assumption referring to a vanishing latency time of safety elements was taken accordingly. The main difference between the two approaches therefore lies in the fact that the previous method takes sensor specifications as constant values while this approach considers different measures.

V Results and Discussion

Figure 3 presents the consumed floor space required for certain sets of sensors fulfilling the safety constraint after applying the first approach presented above. The x-axis shows the space area taken into account by a certain combination of safety elements. The number of the recommended elements is shown on the y-axis. Such a representation is possible and

representative due to the fact that the combination of safety elements is not ambiguous with regard to the floor space consumed, that is, to each value of consumed floor space one combination has been selected by the algorithm. Figure 3 thus allows to obtain the recommended combination of devices for a given floor space. For example, the requirement of safety and minimum floor space consumption in consideration of the given safety elements, can be achieved by making use of two laser scanners which result in a consumed floor space of

$$A_{opt} = 0.0575m^2.$$

Apart from this fact, it can be seen that the amount of floor space is highly correlated to the number of safety mats. In contrast, the number of laser scanners and fence elements do not show clearly distinguishable courses over the area. This observation matches with the theoretical expectancy since safety mats correspond to the largest area among the three devices considered in this study. In addition, it is noted that merely using the safety mats cannot yield a safe configuration. In case 13 mats are placed in the working cell, at least one laser scanner is required to guarantee the human operator's safety. This result can be explained by the fact that the area covered by the mats equals to the floor space taken up. Since the dimensions of the safety mats show the same value for the length as well as for the width, i.e. are considered as quadratic elements, the consumed floor space increases quadratically while the covered area of the circle representing the warning zone arises with linear behavior. However, based on the obtained results, it is recommended to employ two laser scanners at two opposite corners of the working cell. Such a scenario is depicted in figure 4 where the blue area corresponds to the space detected by the laser scanners.

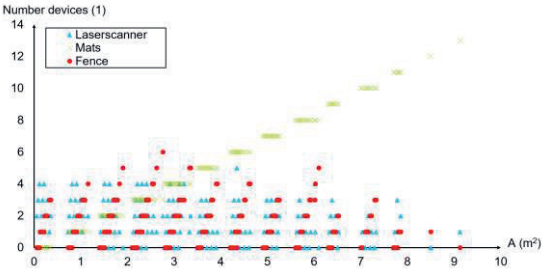


***Figure 3**: Evaluation of different combinations of safety elements over the consumed floor space based on the elements parameters as presented by table 1. For a given floor space, the optimal combination of safety elements can be determined by this figure.*

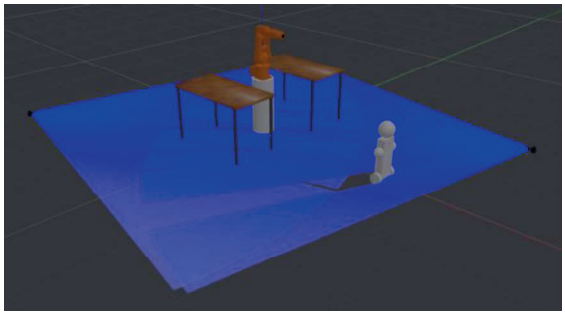


***Figure 4:** Setup depicted in Gazebo matching safety requirements and minimum floor space consumption constraints. This configuration includes two laser scanners which are positioned at two opposite corners. The blue area depicts the detection area of the devices.*

Variation of the parameters

By making use of the second approach in MatLab, parameters of the considered devices have been varied. Doing so, it was investigated how such variations affect the overall results. Hence, this treatment refers to an optimization method in contrast to the previous approach. In the following, the results will be discussed in more detail. It is noted that only one parameter was subject to variation to each point of time while holding the parameters of the remaining elements constant in order to investigate the influence respectively.

a) Variation of the Range (Laser Scanner)

One characteristic of the laser scanner lies in the fact that two parameters can be varied. In first place, the variation of its range has been performed.

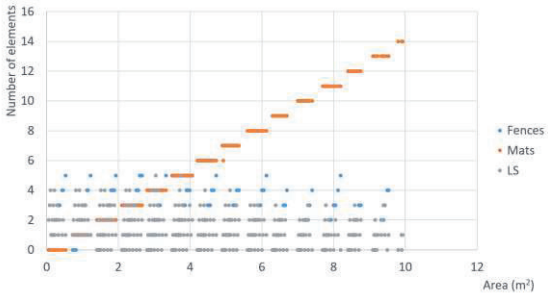


***Figure 5:** Number of safety elements over the consumed floor space for a laser scanner with a range of 5 m and an aperture angle of 120°.*

As figure 5 shows, the number of required devices increases with decreasing range values. At the same time the recommended number of fence elements shows a different behavior than in figure 3.

Comparing this observation with the observation in figure 3 allows the conclusion that in above case the usage of a higher number of fence elements corresponds to a

lower floor space consumption. According to figure 5, the number of recommended number of mats in contrast is not affected by the variation of the laser scanner range.

b) Variation of the Aperture Angle (Laser Scanner)
In the subsequent step, the variation of the aperture angle has been performed. Considering three different values for this quantity enables for the determination of its impact on the overall result. As shown in Appendix A., the deployment of fence elements with the dimensions of 0.1 m x 1.0 results in a smaller floor space consumption when considering smaller aperture angles of the laser scanner.

c) Variation of the Fence Dimensions
Varying the dimensions of the fence leads to the finding that the effect of the laser scanner's parameters is strongly correlated with the number of recommended fences to achieve a minimum floor space consumption. This observation can be explained by the fact that safety mats show a quadratic behavior in context to the floor space consumption. Thus, the developed algorithms prefer the deployment of fences which result in the coverage of the warning zone. However, the correlation between the laser scanner and fence element is suggested to be investigated in more detail.

d) Variation of the Safety Mat dimension
Changing the values of the mat's measures, where the width and length have been varied simultaneously have led to the conclusion that making use of mats is inefficient with respect to the achievement of a solution with the minimum floor space consumption.
The largest dimension corresponds to a mat with dimensions of 1.5m for each, width and height. However, the fulfillment of the safety constraint requires the employment of further devices. In addition, the amount of the required floor space increases rapidly with the number of employed safety mats. As a consequence, the usage of this device is not recommended at first glance.
However, presented approaches are based on the procedure of solely covering a circle corresponding to the outermost warning zone around a robot. In order to provide a further developed result, a more specified application scenario is needed. The variation of the safety mat parameters however do not lead to a different result with respect to the recommended combination regarding the minimum consumed floor space value.

Considering the simplified scenario described in figure 1, the reduction of the laser scanners aperture angle or its range do not affect the results obtained through the algorithms. Therefore, it can be stated that the employment of laser scanners is recommended independent from the variation of the safety elements parameters considered in our approach.

VI Conclusion and Outlook

In this contribution, we investigated the sensor placement issue in a human robot collaboration scenario by means of a simulation setup. It was aimed to fulfill two constraints, namely the human operator's safety defined by norm ISO 15066 in the first and the minimum floor space consumption in second place. We developed two approaches, leading to the same result for the considered scenario. Investigations concerning the influence of the safety devices specifications on the results showed that the deployment of two laser scanners with an aperture angle of 120° and a range of 7 m is recommended for the considered scenario. The results have been obtained by means of two different methods. The first approach is based on an algorithm clearly defining the characteristics and certain measures before starting the simulation. In contrast, the focus of the second method was laid on the selection of the devices with most appropriate measures with regard to floor space consumption and safety. In both cases, safety was considered based on the norm ISO15066 taking into account the distance between robot and human.
However, in order to derive a result applicable for real world applications, the human robot collaboration scenario has to be specified in more detail in order to meet the appropriate requirements. This would allow for the consideration of further additional aspects such as the throughput and economical costs. The result obtained by employing our algorithms yield the recommendation of using two laser scanners which matches the theoretical expectancy for the considered scenario. Merely considering the circle around the outer warning zone simplifies the issue such that velocity measures of the human operator are neglected. Therefore, it is highly recommended to develop further approaches, taking into account the priority of constraints, for example by means of weight factors or functions.
Apart from this fact, the movement of the human operator could be part of further investigations, in terms of the direction. Body parts could be specified to distinguish the

possible consequences of hazards on the human being. Similar to the consideration of different constraints in a certain order, one possible approach to treat this issue could be the use of an optimization method with weights.

Appendix

A. Variation of the Aperture Angle (Laser Scanner)

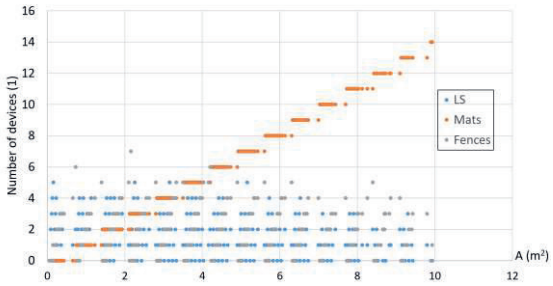


Figure 6: *Recommended combination of safety devices for a laser scanner with an aperture angle of 45°. The recommended combination of safety devices is depicted over the required floor space.*

Literature References

Andreas Krause, C. G. (2006). Near-optimal Sensor Placements: Maximizing Information while Minimizing Communication Cost. *IPSN* , p. 10.

Emily Clark, T. A. (2018). Greedy Sensor Placement with Cost Constraints. *ArXiv Submission*, p. 13.

Jeremy A. Marvel, R. N. (2017, August). Implementing speed and separation monitoring in collaborative robot. *Robotics and Computer-Integrated Manufacturing* , pp. 144-155.

Zainab M. Ismail, M. A. (2018, June). The application of Genetic Algorithm for sensor placement of PZT wafers towards the application in structural health monitoring. *European Conference on Non-Destructive Testing (ECNDT 2018)*, p. 9.

DOI 10.5162/sensoren2019/1.4.1

Optische Sauberkeitskontrolle in der Industriellen Bauteilreinigung – Lösungsansätze für problematische Messaufgaben

Markus Windisch[1]
[1] Fraunhofer-Institut für Verfahrenstechnik und Verpackung IVV, Institutsteil Verarbeitungstechnik
Heidelberger Straße 20, 01189 Dresden, Deutschland

Zusammenfassung

Die Restschmutzmessung als Qualitätskontrolle in der industriellen Bauteilreinigung erfolgt derzeit vorwiegend fluoreszenzbasiert. In einem von der Sächsischen Aufbaubank SAB und dem Europäischen Sozialfonds ESF geförderten Forschungsprojekt wurde anhand industrieller Musterapplikationen die Eignung von IR-Spektroskopie und Optisch Stimulierter Elektronenemission (OSEE) als alternative optische Verfahren für mittels Fluoreszenz nicht lösbare Messaufgaben untersucht. Das Ergebnis sind im Labor erprobte Lösungsansätze für die Produktentwicklung neuer Prozessmesstechnik zur Sauberkeitskontrolle.

Keywords: Restschmutzmessung, Sauberkeitskontrolle, Infrarotspektroskopie, Optisch stimulierte Elektronenemission (OSEE), Industrielle Bauteilreinigung

Praxisproblem Sauberkeitskontrolle

Die Teilereinigung ist ein integraler Bestandteil vieler industrieller Wertschöpfungsketten in der Bauteilfertigung. Ihre Aufgabe ist es, die für das Erreichen der geforderten Produktqualität hinreichende Sauberkeit von Teileoberflächen zu erzeugen. Zu hohe Restschmutzmengen auf Bauteilen verursachen Störungen in Nachfolgeprozessen wie Lackieren, Beschichten, Galvanisieren, Kleben, Schweißen oder der Montage. Neben der Abreinigung unerwünschter Stoffe werden teilweise auch technische wirksame Filme, wie Fügehilfsstoffe für Folgeprozesse oder temporäre Korrosionsschutzschichten für Lagerung und Transport, aufgebracht.

Stand der Technik zum Überwachen von Filmen auf Bauteiloberlfächen sind inline oder atline Messgeräte auf Basis der Fluoreszenzmesstechnik. Sie ermöglichen beispielsweise die Sauberkeitskontrolle bezüglich fluoreszierender Stoffe auf Metalloberflächen [1]. Für nicht hinreichend fluoreszierende Verunreinigungen (z. B. bestimmte Silikon-, Stanz- und Ziehöle, Paraffine und Korrosionsinhibitoren wie Aminoethanole) oder die Sauberkeitskontrolle fluoreszierender Substrate (z. B. Faserverbundwerkstoffe) ist fluoreszenzbasierte Messtechnik ungeeignet. Beispielhaft zeigt Abb. 1 hierzu das Fluoreszenzverhalten unterschiedlicher Ziehöle. Während die Öle A und B in unterschiedlichen Detektionsbereichen gut messbar sind, zeigte Öl C weder bei der im Diagramm dargestellten Anregung mit 365 nm noch bei anderen Anregungswellenlängen eine hinreichende Fluoreszenz und ist deshalb mit diesem Messprinzip nicht detektierbar.

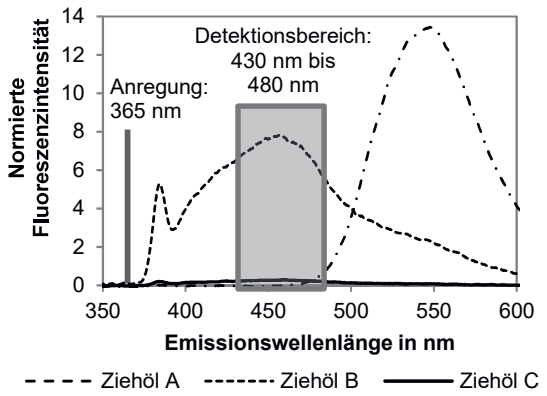


Abb. 1: Fluoreszenzspektren unterschiedliche Ziehöle

Für die vorgenannten Problemfelder werden derzeit aufwändige und fehleranfällige manuelle Kontrollverfahren, wie beispielsweise Testtinten oder Laboranalysen, eingesetzt. Zum Realisieren einer zielgerichteten Qualitätslenkung und der Integration in die Echtzeit-Informationsflüsse der Industrie 4.0 ist eine neue, technisch und wirtschaftlich auf diese bisher ungelösten Überwachungsaufgaben zugeschnittene Prozessmesstechnik zur Sauberkeitskontrolle notwendig. Hierzu wurden anhand diverser Musterapplikationen aus realen Fertigungsprozessen alternative optische Verfahren untersucht.

Musterapplikationen zur Analyse alternativer optischer Messverfahren

Für die Eignungsanalyse wurden die in Tab. 1 angeführten Verschmutzungen ausgewählt. Kriterien waren hierbei neben der technisch-wirtschaftlichen Relevanz die breite Repräsentation unterschiedlicher Stoffgruppen, Schichtdicken und Industriezweige.

Tab. 1: Analysierte filmische Verschmutzungen

Produkt	Stoffgruppe	Anwendung	Merkmale
SurTec 531	Aminoethanole	temporärer Korrosionsschutz	Schichtdicke: 10^0 µm
Fuchs Ecocut FE	Isoalkane	Korrosionsschutz, Schmierstoff	teils flüchtige Bestandteile; Schichtdicke: 10^1 µm
SurTec 089	Tenside	Einpress-Hilfsstoff	Schichtdicke: 10^0…10^1 µm
Elbesil B100	Silikone	Gleit- und Schmiermittel	Schichtdicke: 10^1…10^2 µm
Henkel Frekote	Silikone	Trennmittel	Schichtdicke: 10^0 µm

Probenvorbereitung

Durch Dip-Coating in nach Hersteller- bzw. Prozessvorgaben verdünnten Lösungen wurden Testbleche (Stahl) mit den in Tab. 1 genannten Produkten beschichtet. Das Trennmittel *Frekote* wurde nach dem gleichen Verfahren auf Kohlefaserverbundplatten (CFK) aufgetragen. Detaillierte Angaben zu den Dip-Coating Lösungen sind in Tab. 2 zusammengefasst.

Tab. 2: Dip-Coating Lösungen

Produkt	Lösemittel	Konzentrationen [vol.%]
SurTec 531	Wasser	0,5; 1; 2
Fuchs Ecocut FE	Butanol	10; 50
SurTec 089	Wasser	0,3; 0,5; 0,8
Elbesil B100	Butanol	10; 50
Henkel Frekote	Heptan	1; 3; 10

Infrarotspektroskopie

Die Spektren der Ausgangsprodukte sowie der hergestellten Lösungen wurden in Küvetten mittels Transmissionsmessung aufgenommen. Anschließend wurden die oben beschriebenen Proben hergestellt und mit den in Tab. 3 aufgeführten Spektroskopieverfahren im NIR- und MIR-Bereich analysiert [2], [3].

Tab. 3: Eingesetzte Spektroskopieverfahren

Spektroskop	Verfahren
Bruker Vertex 80 NIR: 1…2,5 nm	Diffuse Reflexion (10°…80°)
Bruker Equinox 55 2,5…17 µm	ATR und µATR
	IRRAS (45°…80°)
	Diffuse Reflexion

Die Oberflächenrauheit der genutzten Testbleche und CFK-Substrate entsprach der Rauheit realer, durch Umform- oder Zerspanungsprozesse erzeugter Bauteiloberflächen. Deshalb waren Ebenheitstoleranzen und die Oberflächenrauheit größer als die Dicke der filmischen Verschmutzung. Aufgrund dessen konnte mit ATR der für eine gute Signalqualität notwendige durchgängige Kontakt zur filmischen Verschmutzung nicht zuverlässig realisiert werden. Wegen der wesentlich kleineren Kontaktflache ist der Einfluss der Oberflächengeometrie und –textur bei µATR signifikant geringer, so dass mit diesem Verfahren bei allen Proben eine gute Qualität (hohe Intensität / SNR) des IR-Signals erreicht wurde. Für den Praxiseinsatz zur Inline-Messung ist jedoch die unvermeidliche Kontamination des Messkopfes bei der berührenden Messung nachteilig. Die Oberflächenrauheit bewirkt ferner – insbesondere bei flachen Einfallswinkeln – eine geringe Reflektivität. Deshalb lässt sich der prinzipbedingte Vorteil des langen Strahlgangs in dünnen Schichten bei IRRAS-Messungen mit sehr flachen Einfallswinkeln (> 80° zur Oberflächennormalen) für reale Bauteiloberflächen nicht nutzen. Im übrigen Winkelbereich bestanden bezüglich der Signalqualität keine signifikanten Unterschiede zwischen IRRAS und diffuser Reflexion.

Für die Restschmutzmessung in der Bauteilreinigung sind µATR und IRRAS bzw. diffuse Reflexion potenziell einsetzbare Spektroskopieverfahren. Die individuellen Nachteile der berührenden Messung bei µATR und der technisch tendenziell aufwändigeren Fremdlichtabschirmung bei den berührungslosen Verfahren sind applikationsspezifisch abzuwägen. Alle nachfolgend präsentierten Ergebnisse für MIR basieren auf IRRAS-Messungen mit einem Einfallswinkel von 45°.

Alle untersuchten Produkte weisen charakteristische spektrale Merkmale im NIR und MIR Bereich auf, die prinzipiell eine Unterscheidung vom Substrat sowie von Restfeuchte als wesentliche Störgröße in wässrigen Prozessen ermöglichen. Die Abbildungen 2 und 3 zeigen

hierzu beispielhaft per Transmission gemessene Spektren unterschiedlicher Produkte in verschiedenen Wellenlängenbereichen.

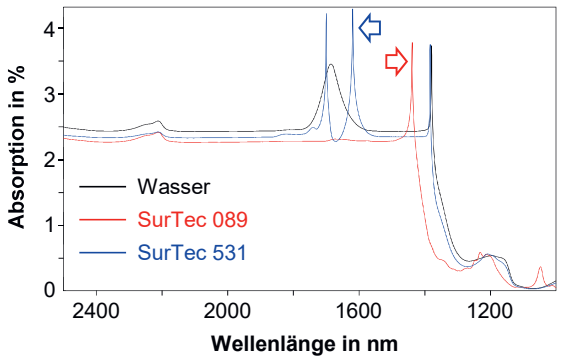


Abb. 2: NIR-Spektren von Einpresstensid SurTec 089 und Korrosionsschutz SurTec 531 (enthalten jeweils Wasser im Ausgangsprodukt) mit charakteristischen Absorptionen zur Unterscheidung von Wasser

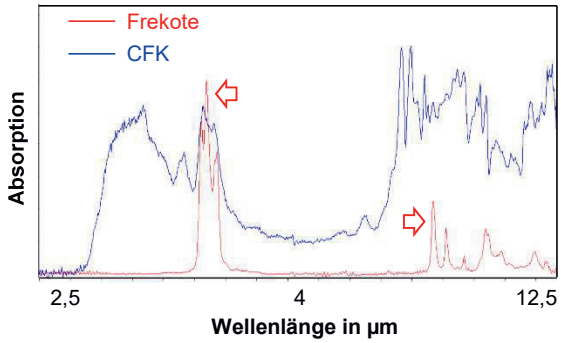


Abb. 3: MIR-Spektrum des Trennmittels (Frekote) mit charakteristischen Absorptionen zur Unterscheidung vom CFK-Substrat (Absorption nicht normiert)

Für die meisten untersuchten Messaufgaben waren bei den applikationsrelevanten Schichtdicken keine charakteristischen, für die Sauberkeitskontrolle nutzbaren Spektrenmerkmale messbar, da aufgrund der sehr kurzen Weglänge durch die dünnen Schichten keine hinreichende Absorption stattfand. Ausnahmen hiervon sind:

- Messung der maximalen Schichtdicke (Dip-Coating in 0,8 vol.%) des Einpresstensids SurTec 089 mit IRRAS (45°); siehe Abb. 4
- Messung dicker Schichten Schmierstoffe Fuchs Ecocut FE und Elbesil B100 mit IRRAS (45°); siehe Abb. 5

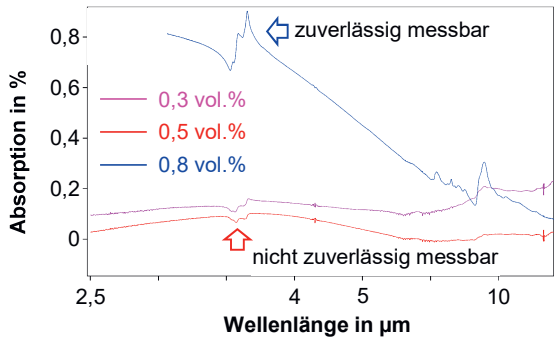


Abb. 4: MIR-Spektren verschiedener Schichtdicken (durch Dip-Coating in unterschiedlich konzentrierten Lösungen hergestellt) des Einpresstensids SurTec 089 auf Stahlblechen

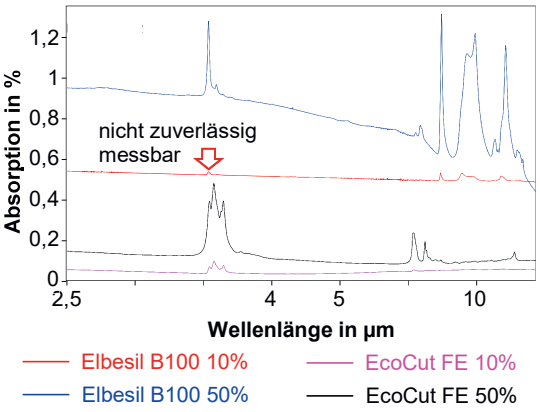


Abb. 5: MIR-Spektren verschiedener Schichtdicken (durch Dip-Coating in unterschiedlich konzentrierten Lösungen hergestellt) der Schmierstoffe Elbesil B100 und EcoCut FE auf Stahlblechen

Optisch stimulierte Elektronenemission (OSEE)

Aufgrund der Nutzung optischer Effekte und der berührungslosen Messung wurde OSEE im Projektkontext den optischen Verfahren zugeordnet und hinsichtlich seiner Eignung zur Messung filmischer Verschmutzungen analysiert. Bei OSEE werden durch UV-Licht Photoelektronen aus einem leitfähigen Substrat gelöst und von einer darüber befindlichen Kathode als Stromfluss erfasst. Auf dem Substrat befindliche filmische Verschmutzungen reduzieren den als Messsignal erfassten Photoelektronenstrom [4].

Die hergestellten Proben filmischer Verschmutzungen wurden durch Abtastung mit einem OSEE-Sensor (Surface Quality Monitor von Photo Emission Tech Inc.) flächig vermessen. Ein Messpunkt entspricht einem Flächenausschnitt von ca. 3 x 3 mm. Der Sensor wurde im

Modus für maximale Sensitivität betrieben und der Photoelektronenstrom *I* für die gewählte Verstärkereinstellung auf $I_n \in [0\ \%; 100\ \%]$ normiert.
Der Referenzmesswert auf einem sauberen Material hängt von der Austrittsenergie der Elektronen des Grundwerkstoffs ab. Für eine im Sauberraum mehrfach mit Aceton, Isopropanol und DI-Wasser gereinigte Aluminiumprobe wurde über die abgetastete Fläche von 100 x 100 mm ein Mittelwert von $\bar{I}_n = 80{,}3\ (\pm 28{,}9)\%$ ermittelt und durch Wiederholungsmessungen bestätigt. Die verhältnismäßig große Schwankungsbreite verdeutlich die hohe Sensitivität von OSEE bezüglich kleinster Restschmutzmengen [5]. Dies erschwert die Entwicklung praktikabler Lösungen zur Kalibrierung entsprechender Sensoren unter industriellen Einsatzbedingen.
Neben Metallsubstraten wurde auch CFK mit OSEE untersucht. In diesem Verbundwerkstoff sind elektrisch leitende Kohlefasern in nichtleitendes Harz eingebettet. Die theoretischen Überlegungen zum Messprinzip lassen keinen Photoelektronenstrom erwarten, da das Harz analog einer filmischen Verschmutzung wirken und diesen unterbrechen müsste. In der Versuchsreihe wurde jedoch bei allen gereinigten CFK-Proben eine vergleichsweise hohe Stromstärke gemessen (siehe Abb. 6). Zur Ursachenanalyse dieses Effektes sind weiterführende Arbeiten notwendig. Derzeitige Hypothesen gehen davon aus, dass entweder durch das Anschleifen eine hinreichende Anzahl Kohlefasern freigelegt wird, das Harz in geringem, aber hinreichendem Umfang leitet oder aber durch die Anregung mit UV-Licht leitfähig wird.

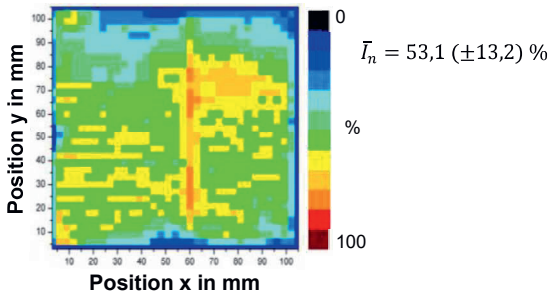


Abb. 6: Flächige OSEE-Messung einer sauberen, angeschliffenen CFK Probe (Strom normiert)

Ausgehend von den Ergebnissen mit gereinigten Oberflächen wurden im nächsten Schritt mit den in Tab. 1 genannten Produkten beschichtete Substrate analysiert. Auf Aluminium- und Stahlblechen führten alle aufgebrachten Filme bereits in der jeweils geringsten Schichtdicke zur vollständigen Unterbrechung des Photoelektronenstroms. Dies ermöglicht eine Unterscheidung zwischen sauberen und kontaminierten Oberflächen; nicht jedoch eine Differenzierung unterschiedlich dicker Filme. Einzig auf den CFK-Proben konnten verschieden dicke Filme unterschieden werden (siehe Abb. 7).

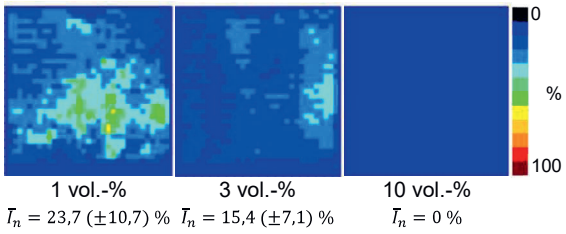


Abb. 7: Flächige OSEE-Messungen des Trennmittels Frekote auf CFK-Proben (100 x 100 mm; Strom normiert)

Zusammenfassung und Schlussfolgerungen zur Messtechnikentwicklung für die Industrielle Bauteilreinigung

Als potenzielle Entwicklungsansätze für nicht durch fluoreszenzbasierte Messtechnik lösbare Überwachungs- und Messaufgaben der Industriellen Bauteilreinigung wurden verschiedene IR-Spektroskopieverfahren sowie OSEE anhand ausgewählter Musterapplikationen experimentell untersucht.
Für vergleichsweise dicke Schichten nichtfluoreszierender Schmierstoffe bietet die MIR-Spektroskopie in den technischen Ausprägungen µATR, IRRAS oder diffuse Reflexion metrologische Lösungsansätze. Technische Herausforderungen bestehen aufgrund teils berührender Messungen bzw. der notwendigen anspruchsvollen Fremdlichtabschirmung bezüglich der Prozessintegration.
Das Vorhandensein filmischer Verunreinigungen auf Metallteilen lässt sich mit sehr hoher Sensitivität mit OSEE messen. Da hierbei jedoch keine Unterscheidung unterschiedlicher Schichtdicken möglich ist, werden die meisten praktischen Messaufgaben nicht gelöst. Beispielsweise ist es nicht möglich, die Einhaltung zulässiger Restschmutzmengen für die bedarfsgerechte und dadurch wirtschaftliche Führung industrieller Reinigungsprozesse zu überwachen. Für technisch erwünschte Filme, wie beispielsweise Einpresstenside, können vorgegebene Untergrenzen der zulässigen Schichtdicke nicht gemessen werden. Einzig auf CFK-Teilen lassen sich mit OSEE technisch relevante Trennmittelfilme quantifizieren und somit eine praxisgerechte Restschmutzmessung realisieren.
Aus Sicht der Messtechnikentwicklung betreffen die beschriebenen Lösungsansätze ver-

gleichsweise schmale Teilbereiche der Sauberkeitskontrolle. Für das wichtige Problemfeld dünner, nichtfluoreszierender Verschmutzungen auf Metallteilen wurden keine praktikablen Ansätze gefunden. Ausgehend von den Ergebnissen zur technisch-metrologischen Eignung und den Einsatzgrenzen der untersuchten Verfahren ist in weiterführenden Arbeiten eine wirtschaftliche Bewertung notwendig. Hierbei ist zu untersuchen, ob die Produktentwicklung von Prozessmesstechnik vor dem Hintergrund der damit adressierbaren Anwendungsfelder ökonomisch zielführend ist.

Danksagungen

Der Autor dankt der SITA Messtechnik GmbH für die Zusammenarbeit im Förderprojekt und die in diese Publikation eingebrachten Forschungsdaten. Weiterer Dank gilt den Kollegen der Arbeitsgruppe Qualitätssicherung am Fraunhofer IFAM für die durchgeführten Messungen sowie die aufschlussreichen Gespräche zu deren Interpretation.

Literaturnachweis

[1] Hofmann, A.: Bildgebende Fluoreszenz-Messtechnik zur Inline-Kontrolle von Oberflächen. *Tagungsband EFDS Workshop Nasschemische Reinigung für das Beschichten*, Dresden, 2015

[2] Ge, M., et al.: Effect of Lube Oil Film Thickness on Spray/Wall Impingement with Diesel, M20 and E20 Fuels. *SAE Technical Paper*, 2017.

[3] Mehta, Narinder K., et al. Development of an in situ spectroscopic method for cleaning validation using mid-IR fiber-optics. *BIOPHARM-EUGENE- 15.5*, 2002, 36-43

[4] Brune, K., et al.: Pre-bond quality assurance of CFRP surfaces using optically stimulated electron emission. *Proceedings of the 3rd International Conference of Engineering Against Failure. ICEAF*, Kos, Greece, 2013

[5] Perey, D. F.: A portable surface contamination monitor based on the principle of optically stimulated electron emission (OSEE). *NASA Report*, 1996

DOI 10.5162/sensoren2019/1.4.3

Concepts for Accurate Electrical Conductivity Measurement of Liquids in Industrial Process Analytics

Michael Vogt[1,2], Santiago Hidalgo[2], Malte Mallach[3], Theresia Lange[3], Jan Förster[3], Thomas Musch[2]
[1]KROHNE Messtechnik GmbH, Ludwig-Krohne-Str. 5, 47058 Duisburg, Germany
[2]Institute of Electronic Circuits, Ruhr University Bochum, 44780 Bochum, Germany
[3]KROHNE Innovation GmbH, Ludwig-Krohne-Str. 5, 47058 Duisburg, Germany

Summary

The electrical conductivity of liquids is of interest in many industrial processes, for example in water desalination, cleaning, mixing of different liquids, analysis of acids and bases, etc. In this contribution, specifically two-electrode conductivity sensors are discussed. A problem with this type of sensor is that a double layer (DL) is given at the interfaces between each of the electrodes and the liquid, resulting in varying and unpredictable diffusion capacitances. For this reason, the electrical resistance of the liquid is not directly accessible from measurements of the sensor input impedance at low frequencies. For accurate determination of the conductivity, the parasitic and disturbing diffusion capacitances, as well as the unknown capacitance of the liquid and the parasitic sensor capacitance have to be eliminated from the measured impedance. The approach in electrical impedance spectroscopy (EIS) is to measure over a very large frequency range for explicitly determining and eliminating the unknown network elements. The motivation behind the work presented here is to allow for precise conductivity measurement at only low frequencies in the range of some kHz. Very advantageously with this, low-cost and low-demanding sensor electronics can be utilized for the realization of according industrial sensors. An equivalent circuit network model of the sensor has been analyzed in detail for this purpose. Approaches for estimating the resistance of the liquid from low frequency measurements of the sensor input impedance are presented and discussed. Finally, the resulting systematic measurement error has been evaluated based on the network model and experimentally.

Keywords: Conductivity measurement, constant phase element, electrical impedance spectroscopy, equivalent circuit, coaxial sensor.

Introduction

Liquids containing dissociated ions, which are referred to as *electrolytes*, have an electrical conductivity that increases with increasing concentration of the ions. Electrical *conductivity sensors* are widely used for easy and continuous in-line monitoring and control of various processes in chemistry, water treatment, and many other applications.

Electrical impedance spectroscopy (EIS) can be understood as an extension of conductivity sensors to additionally also perform measurements of the electrical permittivity and to measure over a much wider frequency range. Numerous works have already been published on using EIS, covering various applications and the analysis of very different substances [1-6].

In this contribution, a model-based approach for eliminating the parasitic network elements from the measured sensor input impedance at relatively *low* frequencies is presented and evaluated. The underlying concept can be understood as an *extrapolation* to measurements at high frequencies, which are in EIS *explicitly* performed. The advantage is that a low-cost digital-to-analog converter (DAC) and analog-to-digital converter (ADC) with a low sampling frequency can be utilized for direct signal generation and digitization, respectively. Accordingly, a microcontroller with low complexity and low power consumption can be used for the sensor electronics and for realizing inexpensive industrial sensors. Low power is very interesting if for example a 4 to 20 mA current loop should be used as field bus interface and also as power supply of the sensor. The available computing power is very limited in this case, so that complicated parameter estimation approaches like the ones used in EIS cannot be implemented.

It will be shown below that the sensor input impedance at low frequencies is dominated by the series connection of the diffusion

capacitances and the resistance of the liquid. On the other hand, the parallel connection of the resistance with the capacitances of the liquid and the sensor dominates at high frequencies. We have defined two characteristic frequency points for a more detailed quantitative analysis of the sensor input impedance in the transition region between the two border cases at medium frequencies.

Electrical Conductivity Sensor

In Fig. 1 a), the setup of a *two-electrode coaxial conductivity sensor* is illustrated, and Fig. 1 b) shows a practical realization of such a sensor.

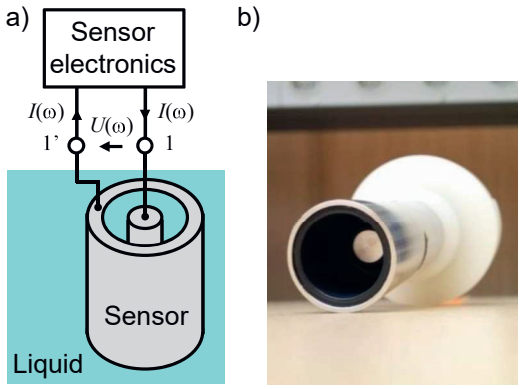


Fig. 1: Two-electrode coaxial electrical conductivity sensor: a) Illustration, b) photo of realized sensor.

This type of sensor is preferably utilized for measurements of liquids with small electrical conductivity. The sensor is directly immersed into the liquid to be analyzed. Using sensor electronics connected to the sensor input port 1/1', see Fig. 1 a), the impedance into the sensor, with the liquid under test between the two electrodes, is measured.
The complex *sensor input impedance* $Z(\omega) = U(\omega) / I(\omega) = R(\omega) + \mathrm{j} \cdot X(\omega)$ over the (angular) frequency ω is derived from the measured voltage $U(\omega)$ across the sensor and the current $I(\omega)$ flowing into the sensor.

Equivalent Circuit Network Model

In Fig. 2, an equivalent circuit network model of the two-electrode conductivity sensor immersed into the liquid is shown.
The *conductance* $G_\mathrm{L} = 1 / R_\mathrm{L} = \sigma_\mathrm{L} / k_\mathrm{cell}$ of the liquid is proportional to its *electrical conductivity* σ_L, with the *resistance* R_L of the liquid and the so-called '*cell constant*' k_cell of the sensor. The latter parameter depends on the geometry (size, distance) of the two electrodes.

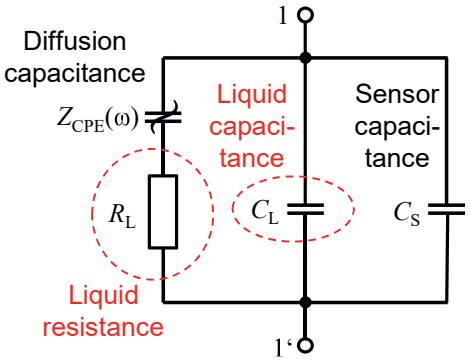


Fig. 2: Equivalent circuit network model: Resistance R_L and capacitance C_L of the liquid, parasitic constant phase element (CPE) $Z_\mathrm{CPE}(\omega)$ (double layers), parasitic sensor capacitance C_S.

At the interfaces between each of the two electrodes and the liquid (being an electrolyte), a *double layer* (DL) builds up. The resulting series connection of these two DLs can be summarized in one network element. A so-called *constant phase element* (CPE) with an impedance $Z_\mathrm{CPE}(\omega)$ described as follows is a good model for the DLs (with $0 < n \leq 1$):

$$Z_\mathrm{CPE}(\omega) = \frac{1}{Q_0 \cdot (\mathrm{j}\omega)^n} = \frac{\cos\left(n \cdot \frac{\pi}{2}\right) - \mathrm{j} \sin\left(n \frac{\pi}{2}\right)}{Q_0 \cdot \omega^n} \qquad (1)$$

The impedance $Z_\mathrm{CPE}(\omega)$ is defined by the two parameters n and Q_0, and it has a *constant phase* over the (angular) frequency ω. For $n = 1$, the CPE is equal to a pure *capacitance* Q_0. In the general case $0 < n < 1$, the CPE is a *lossy* capacitance, whereby the resistance and the reactance of the CPE impedance (having a constant phase) *both* change over frequency, see (1). With the frequency going to infinity, the impedance $Z_\mathrm{CPE}(\omega)$ goes in any case to zero.
The capacitance $C_\mathrm{L} = \varepsilon_\mathrm{L} / k_\mathrm{cell,\varepsilon}$ of the liquid is proportional to its permittivity $\varepsilon_\mathrm{L} = \varepsilon_\mathrm{r,L} \cdot \varepsilon_0$ (relative permittivity $\varepsilon_\mathrm{r,L}$, vacuum permittivity ε_0), with a cell constant $k_\mathrm{cell,\varepsilon}$. In the case of metallic electrodes with no additional dielectric coating etc., $k_\mathrm{cell,\varepsilon} \approx k_\mathrm{cell}$ is given. The parallel capacitance C_S in Fig. 2 summarizes any given additional parasitic input capacitances at the port 1/1' of the sensor. This might include the capacity of the cable connection to the electrodes.

Sensor Input Impedance

In the following, the task of deriving the *electrical conductivity* σ_L of the *liquid* from the measured impedance $Z(\omega)$ data at the port 1/1', see Fig. 2, is discussed:

$$Z(\omega) = \frac{R_L + Z_{CPE}(\omega)}{1 + j\omega \cdot (C_L + C_S) \cdot (R_L + Z_{CPE}(\omega))} \qquad (2)$$

The resistance R_L is not directly accessible from this measurement. Accordingly, the challenge is to eliminate the unknown parallel capacitance $C_L + C_S$ of the liquid and the sensor, and also the CPE impedance $Z_{CPE}(\omega)$ from the measurement. The *parasitic* CPE is very unpredictable and unknown, and it can largely vary over time and over different samples of the same type of sensor. The CPE parameters n and Q_0, see (1), depend largely on various properties (electrodes material, surface roughness, and surface area of the electrodes, type and conductivity of the liquid, etc.).

In principle, the sensor input impedance $Z(\omega)$ also gives access to the *electrical permittivity* ε_L of the liquid, as long as C_L can be separated from C_S in the parallel connection, see Fig. 2. In *industrial* applications, however, the electrical *conductivity* σ of the liquid is the primarily interesting parameter to be measured. Taking the measured temperature into account, the *concentration of dissolved ions* in the liquid can be derived from the measured conductivity.

In the general case, the permittivity of the liquid is unknown and might largely vary (e.g. $\varepsilon_{r,L} = 2.3$ for gasoline, $\varepsilon_{r,L} = 80$ for water). For this reason, the liquid's capacitance C_L has been explicitly taken into consideration in Fig. 2 as an unknown *parasitic* network element. Accordingly, C_L has to be eliminated from the measured $Z(\omega)$. The same applies for the sensor capacitance C_S, although it might be largely compensated by a *calibration* of the sensor. Despite of this, a possibly remaining residual capacitance after calibration is also explicitly taken into account by means of C_S.

Characteristic Frequency Points

In EIS, parametric *Nyquist plots* of the (conjugate-complex) sensor input impedance $Z^*(\omega) = R(\omega) - j\,X(\omega)$ over frequency are quite commonly used to visualize measurement results etc. [2-5]. Nyquist plots are also used in the following to discuss strategies for deriving the resistance R_L of the liquid from $Z(\omega)$ in more detail. The plots also give an easy understanding and access to given systematic measurement errors.

In Fig. 3, a *normalized* Nyquist plot of $Z^*(\omega)/R_L$ is illustrated (under some limitations and for varying parameter n).

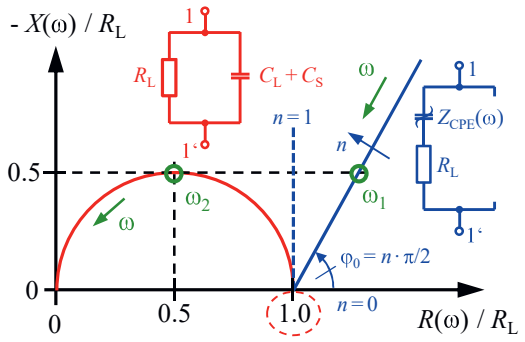


Fig. 3: Nyquist plot (illustration): Conjugate-complex normalized sensor input impedance $Z^*(\omega)/R_L$ for border cases of solely series and parallel connections of parasitic elements with the resistance R_L of the liquid; definition of characteristic frequency points ω_1 and ω_2.

The *blue* curve in Fig. 3 shows the *border case* of *negligibly small* liquid and sensor capacitances (i.e. $C_S = C_L = 0$). Under this condition, a *linear* function over the (angular) frequency ω is given, see (1). The slope of the according curve depends on the CPE parameter n (see the angle $\varphi_0 = n \cdot \pi/2$ in Fig. 3).

Because the reactive current through C_S and C_L, see Fig. 2, only plays a significant role at *higher* frequencies, the blue plot is a good approximation of the sensor input impedance at (relatively) *low* frequencies. A similar consideration can be made for the *border case* of having a *negligibly small* CPE impedance (i.e. for $Z_{CPE}(\omega) = 0$) at (relatively) *high* frequencies. In this case, a half-circular curve is given in the Nyquist plot, see the *red* curve in Fig. 3. This follows from calculating the reciprocal (*Möbius transformation*) of the sensor input *admittance* $Y(\omega)$ in this case. The latter has a constant real part and an imaginary part proportional to the frequency as follows, see (2) with $Z_{CPE}(\omega) = 0$:

$$Y(\omega) = \frac{1}{Z(\omega)} = R_L^{-1} + j\,\omega \cdot (C_L + C_S) \qquad (3)$$

Obviously, for the general case of the sensor according to Fig. 2, the CPE and the two parallel capacitances *all* have to be taken into account at the same time. Especially at *medium* frequencies, a Nyquist plot *in-between* the blue and the red curves in Fig. 3 is expected. However, at *low* and *high* frequencies, the two border cases discussed above will still be a good approximation. For testing the validity of this assumption, we have defined the two *characteristic frequency points*

ω_1 and ω_2 as shown in Fig. 3. Using (1) and (3), they are given as follows:

$$\omega_1 = \left(\frac{2\cdot\sin(n\cdot\pi/2)}{R_L\cdot Q_0}\right)^{1/n}, \quad \omega_2 = \frac{1}{R_L\cdot(C_L+C_S)} \qquad (4)$$

The blue and red curves in Fig. 3 will only be good approximations in the case that $\omega_2 >> \omega_1$ is given. For a more detailed analysis, the *ratio* $\omega_{rel} = \omega_2 / \omega_1$ of the two characteristic frequency points is used below as an important parameter:

$$\omega_{rel} = \frac{\omega_2}{\omega_1} = \frac{R_L^{\left(\frac{1}{n}-1\right)}}{(C_L+C_S)} \cdot \left(\frac{Q_0}{2\cdot\sin(n\cdot\pi/2)}\right)^{1/n} \qquad (5)$$

As given in (5), the Nyquist plot at *medium* frequencies depends on the given *relationship* between the parasitic elements Q_0, n, C_L+C_S, and relative to R_L.

Numerical Simulations

The detailed sensor input impedance over a large frequency range has been analyzed based on numerical simulations using the complete equivalent circuit network model in Fig. 2. Using (1), (4), and (5), the normalized sensor input impedance according to (2) can be expressed as follows with the *normalized frequency* $\omega_n = \omega / \omega_1$:

$$\frac{Z(\omega)}{R_L} = \frac{1+\frac{1}{2\cdot\sin(n\cdot\pi/2)\cdot(j\omega_n)^n}}{1+j\omega_n/\omega_{rel}\cdot\left(1+\frac{1}{2\cdot\sin(n\cdot\pi/2)\cdot(j\omega_n)^n}\right)} \qquad (6)$$

The parasitic network elements in (6) are now completely described by ω_{rel} and the CPE parameter n.

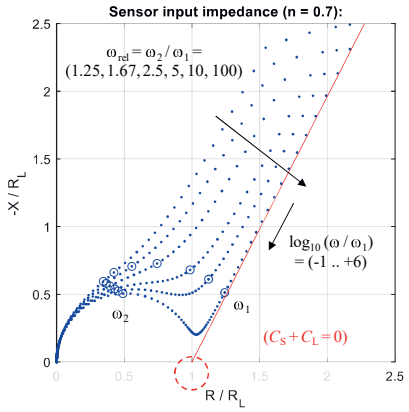


Fig. 4: Nyquist plot: Calculated normalized sensor input impedance $Z^*(\omega)/R_L$ over normalized frequency $\omega_n = \omega/\omega_1$ for parameterization with $n = 0.7$ and varying ratio $\omega_{rel} = \omega_2/\omega_1$ of the characteristic frequencies ω_1 and ω_2.

As an example, the blue dots in Fig. 4 show the Nyquist curves of the calculated normalized sensor input impedance according to (6) for $n = 0.7$ and for varying $\omega_{rel} = \omega_2/\omega_1$ (in the range from 1.25 to 100) at discrete (exponentially increasing) normalized frequencies $\omega_n = \omega/\omega_1$ (in the range from 10^{-1} to 10^6). As expected and discussed above, the linear and circular curves in Fig. (3) are only good approximations for a large ratio of the two characteristic frequency points.

Based on the results in Fig. 4 the resulting *systematic measurement error* in case of using one of the two approximations can be analyzed. The red curve is the sensor input impedance for the low-frequency approximation with $C_S = C_L = 0$. The blue circles in Fig. 4 indicate the characteristic frequency points ω_1 and ω_2 in each of the Nyquist plots. This allows for choosing reasonable measurement frequencies for a sufficiently small measurement error in the case of using the approximations.

Fig. 5 shows another example of the calculated normalized sensor input impedance for $n = 0.3$.

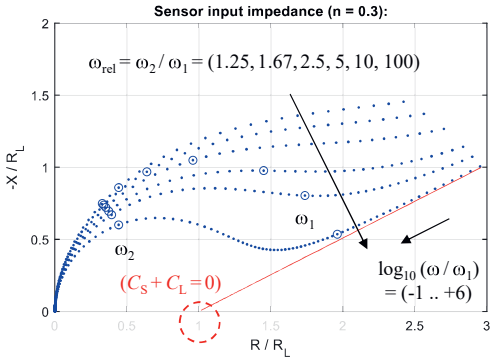


Fig. 5: Nyquist plot: Calculated normalized sensor input impedance for $n = 0.3$.

It can be seen that the relative deviations from the approximations are now (for $n = 0.3$) much larger as compared to the example in Fig. 4 (with $n = 0.7$). Consequently, the actually given parameters n and ω_{rel}, describing the parasitic elements, have a significant effect.

Measurement Concepts

The usual approach in EIS is to measure the sensor input impedance over a very *large frequency range* and to explicitly determine the parameters of an equivalent circuit network model. This is done by fitting the parametric network model to the *wideband* measurement data [1].

It has already been discussed above that our goal is to only utilize measurements at *low* frequencies for estimating the resistance R_L of the liquid. Accordingly, we aim at using the

border case and approximation given by neglecting the liquid and sensor capacitances ($C_S = C_L = 0$) discussed above, see the blue curve in Fig. 3.
A first concept is to use the measured impedances at only *two* low *measurement frequencies* for estimating R_L. Referring to the red curves in Figs. 4 and 5, this can be understood as a *linear extrapolation* from these two points to the intersection with the abscissa in the Nyquist plot, see Fig. 6 a).

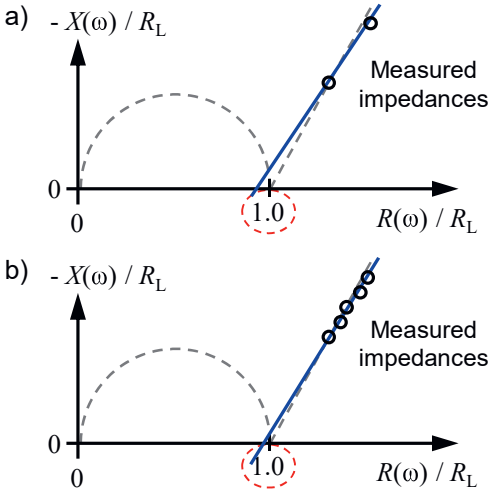


Fig. 6: Nyquist plot (illustration): a) Linear extrapolation using impedance measurements at only two low measurement frequencies, b) linear regression fit using multiple measurement frequencies; the same frequency interval is covered with both concepts in this example.

The characteristic frequency points ω_1 and ω_2 in (4) allow to assess the applicability of this approach in more detail. First, $\omega_2 >> \omega_1$ (i.e. $\omega_{rel} >> 1$) should be fulfilled so that the linear case is a sufficiently good approximation at low frequencies for the actually given situation, see Figs. 4 and 5. Second, the utilized measurement frequencies should be much smaller than ω_1. This criterion is to guarantee that the measurements are actually done in the linear part of the curve. Accordingly, *systematic* measurement errors given by applying the linear approximation are minimized. As a further criterion, the distance between the two frequency points should be as large as possible. The goal behind this is to be as less sensitive as possible against *random* measurement errors (given by the noise of the sensor electronics). However, the linear approximation has of course still to be valid at the lower frequency.
A reasonable second concept is to extend the linear extrapolation to impedance measurements at *multiple* low frequencies by using a *linear regression fit*, see Fig. 6 b). The interval of utilized frequencies should of course comply with the same criteria formulated for the first approach.

Experimental Evaluation

The findings and approaches discussed above have also been evaluated by means of experimental measurements. The coaxial conductivity sensor shown in Fig. 7 a) (see also Fig. 1 b), stainless steel electrodes, cell constant $k_{cell} = 1.0\ \mathrm{cm}^{-1}$) has been used for all measurements.

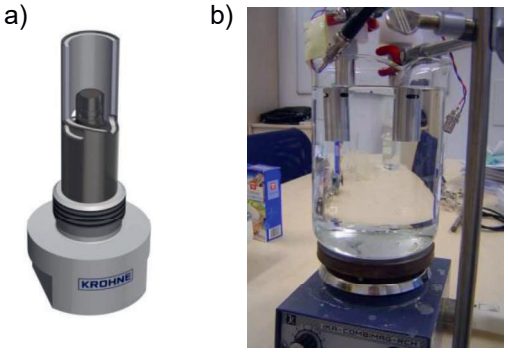


Fig. 7: a) Coaxial measurement sensor (CAD-model), b) realized measurement setup with a magnetic stirrer.

A *vector network analyzer* (VNA, model Bode 100; OMICRON electronics GmbH, Klaus, Austria) has been utilized for measuring the complex sensor input impedance. The VNA has been connected to the sensor using a coaxial cable ($Z_L = 50\,\Omega$ line impedance). Prior to the measurements, the VNA was calibrated at the sensor port using a *TOSM* (through, open, short, $50\,\Omega$ match) *calibration* scheme.
Salt-water with varying salinity (0 to $100\,\mathrm{g/l}$ of NaCl) has been tested. A magnetic stirrer was used to guarantee for a good dissolution of the salt, see Fig. 7 b). The advantage of using *water* (with its very large relative permittivity $\varepsilon_{r,L} \approx 80$) is, that the *worst-case* with regard to introducing disturbances to the conductivity measurement by the capacitance C_L of the liquid is analyzed.

Measurement Results

In Fig. 8, Nyquist plots of the measured (absolute) sensor input impedance $Z(\omega)$ over a large frequency range from $10\,\mathrm{Hz}$ to $150\,\mathrm{kHz}$ are shown for varying salinity.
It can be seen that the CPE model (with a constant phase of the impedance over frequency) discussed above fits well to the measured impedance. Additionally, in the blue curve (tap water with no salt added at all), the

expected influence of the parallel capacitances (liquid and sensor) can clearly be seen.

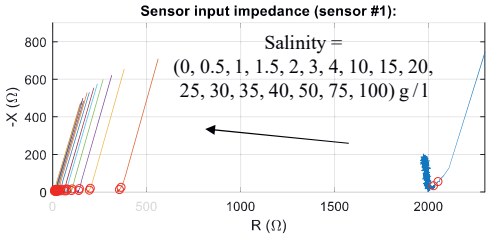


Fig. 8: Nyquist plots: Measured sensor input impedance $Z^*(\omega)$ (sensor #1) over a large frequency range ($10\,\mathrm{Hz}$ to $150\,\mathrm{kHz}$) for varying salinity, measurements at $1\,\mathrm{kHz}$ and $3\,\mathrm{kHz}$ marked with red circles.

With increasing salinity (i.e. with increasing conductivity and decreasing resistance of the liquid), however, no large influence of the capacitances can be identified any more.

In order to estimate the *actually given* ('true') resistance R_L of the liquid for each different salinity, we have used the complex sensor input admittance $Y(\omega) = 1/Z(\omega)$ at the *highest* measurement frequency of $150\,\mathrm{kHz}$. The underlying assumption is that the impedance $Z_\mathrm{CPE}(\omega)$ of the CPE at this high frequency is negligibly small against R_L. According to (3), R_L has been estimated as the inverse of the *real part* of the admittance at that frequency. For the given salinities in the range from 0 to $100\,\mathrm{g/l}$, the accordingly determined R_L is in the range from about $2\,\mathrm{k\Omega}$ down to $8\,\Omega$. Using the nominal cell constant $k_\mathrm{cell} = 1.0\,\mathrm{cm^{-1}}$ of the utilized sensor (which has been optimized by means of electrostatic field simulations), the corresponding conductivity σ is in the range from about $0.5\,\mathrm{mS/cm}$ to $125\,\mathrm{mS/cm}$.

Fig. 8 shows the results obtained with a *first sample* sensor #1 of the utilized sensor type. The same measurements have also been performed using a second sample sensor #2 of the same sensor type. As can be seen in Fig. 7 b), the experiments with the two sensors have been made parallel to each other under exactly the same conditions. The motivation behind this was to investigate deviations from one sample to another in more detail, especially deviations of the CPEs.

In Fig. 9, the Nyquist plots for sensor#2 are shown. The actually given R_L has again been estimated for each varying salinity as described before using the measured impedance at the highest frequency of $150\,\mathrm{kHz}$. The estimates obtained from the measurements with the two sensors #1 and #2 are in a good agreement with each other (maximum deviation smaller $\pm 1.5\%$).

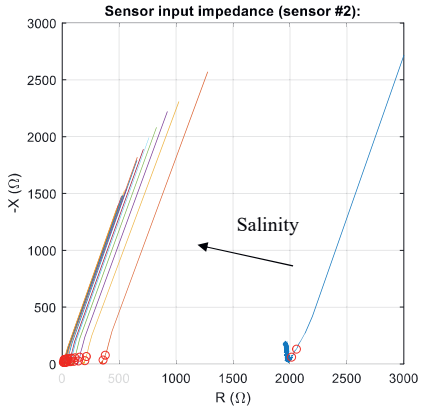


Fig. 9: Nyquist plots: Measured sensor input impedance $Z^*(\omega)$ (sensor #2).

Taking the different scaling of the axes in Figs. 8 and 9 into account, it can be seen that the reactance of the CPE of sensor #2 is much larger as compared to sensor #1 (at the same measurement frequencies). It can be concluded from this that the DL diffusion capacitance (and so the CPE parameter Q_0) is much smaller for sensor #2 compared to sensor #1. A reasonable explanation behind this could be that the two sensors have been used with different intensities and different types of liquids prior to the experiments described here. Accordingly, it is possible that different microscopic surface structures and deposits have been given at the electrodes of the sensors, thus changing the properties of the DLs of each sensor.

For a more detailed analysis, the *normalized* sensor input impedance $Z^*(\omega)/R_\mathrm{L}$ has been calculated (normalized to the estimates for the actually given R_L as discussed above) for each varying salinity and for each of the two sensors.

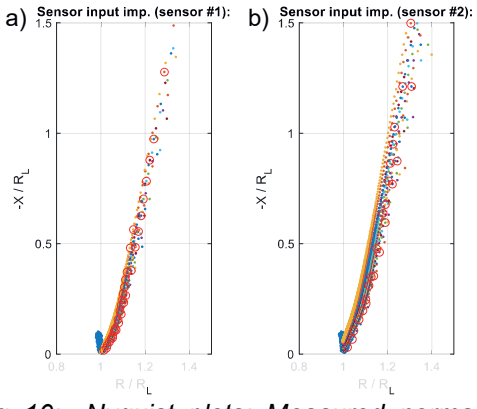


Fig. 10: Nyquist plots: Measured normalized sensor input impedance $Z^*(\omega)/R_\mathrm{L}$ over a large frequency range ($10\,\mathrm{Hz}$ to $150\,\mathrm{kHz}$), measurements at $1\,\mathrm{kHz}$ and $3\,\mathrm{kHz}$ marked with red circles: a) Sensor #1, b) sensor #2.

In Fig. 10, the normalized impedances at the individual frequencies (960 equally spaced measurement frequency points) have been plotted by using *dots*. (For a better visualization, the corresponding consecutive points have been connected with each other by using a *line* in the plots in Fig. 9.) In Figs. 8 and 9, the same colors have been used for each same salinity, and the red circles in Figs. 8 to 10 indicate the two measurement frequencies of $1\,\mathrm{kHz}$ and $3\,\mathrm{kHz}$ for each individual curve.

The variations of the slopes of the curves for each individual sensor #1 and #2 in Figs. 10 a) and b), as well as the variation between the two sensors are very small. This shows that the CPE parameter n remains largely constant over all measurements. It can be seen now the definition of the characteristic frequency point ω_1 in (4) was very reasonable. As discussed above, the CPE parameter Q_0 is smaller for sensor #2 compared to #1. Accordingly, ω_1 is larger for sensor #2 and the normalized impedances with this sensor are more far away from the abscissa as compared to sensor #1 (at the same measurement frequencies). This can be seen very clearly from comparing the positions of the red circles in Figs. 10 a) and b), respectively.

Using this measurement data, both, the *two measurement frequencies* concept as well as the *multiple measurement frequencies* concept discussed above, have been evaluated. As the parallel capacitances play no significant role here (except for the smallest salinity), relatively high measurement frequencies of $1\,\mathrm{kHz}$ and $3\,\mathrm{kHz}$ have been used for the first concept. The second concept has been tested by using 14 equally spaced frequencies in a corresponding interval from $1\,\mathrm{kHz}$ to $3\,\mathrm{kHz}$.

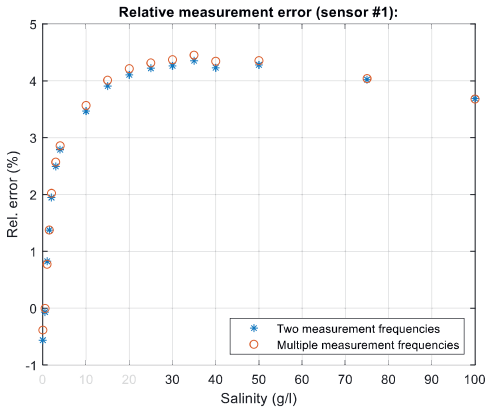


Fig. 11: Relative conductivity measurement errors for varying salinity (sensor #1): Two measurement frequencies and multiple frequencies concepts.

In Fig. 11, the resulting relative measurement errors for the data of sensor #1 are shown (errors referring to the estimated actually given R_L in each case, as described above). It can be seen that, although only measurements at very low frequencies are utilized, the relative measurement error (referring to measurements at a very high frequency of $150\,\mathrm{kHz}$) are quite small. The given errors are smaller than $4.3\,\%$, what is acceptably small. The systematic measurement errors at small salinities (i.e. for small conductivities of the liquid) are even much smaller. The latter is a plausible result, because it confirms again that the CPE is getting less relevant with increasing resistance R_L. The comparison of the errors in Fig. 11 obtained with the two measurement concepts shows no significant difference. Based on the expectation that the second concept, i.e. using a linear regression fit, is less sensitive against electronic noise, it can be concluded that the given signal-to-noise ratio (SNR) with the VNA was sufficiently large and was not a limiting factor here. The deviations given in Fig. 11 are for this reason mainly systematic measurement errors, which are introduced by using the linear approximation.

In Fig. 12, the corresponding measurement errors obtained by using sensor #2 are shown.



Fig. 12: Relative conductivity measurement errors for varying salinity (sensor #2).

Again, the systematic measurement error at small salinities is very low, and an only slightly larger maximum error of $4.9\,\%$ is given. With sensor #2, the errors at larger salinities are much smaller as compared to sensor #1, see Figs. 11 and 12. It is known from (4) that ω_1 generally increases with decreasing R_L (i.e. with increasing salinity) and also with decreasing CPE parameter Q_0. A plausible explanation for the finding here now is that the *increase* of ω_1 *with increasing salinity* is larger for sensor #2, because of its larger Q_0, as compared to sensor #1. Finally, the larger the characteristic frequency ω_1 is, the better the assumption of *linear* curves is actually valid (for the same given low frequencies), see Figs. 4 and 5.

Summary and Conclusions

In this contribution, the applicability of estimating the electrical conductivity of liquids by using measurements of the sensor input impedance at only low frequencies has been analyzed and validated. The motivation behind this was to allow for direct signal generation and digitization using inexpensive and low-complex sensor electronics. This is in some contradiction to the usual approach in EIS of using measurements over a large frequency range for explicitly estimating both, the interesting network elements as well as the parasitic ones.

An equivalent circuit network model of the sensor has been discussed, wherein the DL at the electrodes is modeled as a CPE. The problem with this element is that its properties can largely vary from one sample of the same type of sensor to another. Numerical simulations with the network model have been performed. The goal in this context was to analyze systematic measurement errors, which are introduced by using border case approximations for describing the network model. Accordingly, the applicability of using measurements at low frequencies only has been evaluated. Two different measurement concepts have been investigated. The first approach uses the impedances at only two measurement frequencies for linear extrapolation. In the second concept, linear regression is performed with the measured impedances at multiple discrete frequencies.

Finally, experimental measurements have been performed with salt-water of varying salinity. The presented results show that the systematic measurement, which is introduced by measuring at only small frequencies in the range from $1\,\mathrm{kHz}$ to $3\,\mathrm{kHz}$, is acceptably small.

References

[1] O. Kanoun, U. Tröltzsch: Application of Parameter Extraction Techniques for Impedance Spectroscopy. Instrumentation and Measurement Conf., 2281-2286 (2005); doi: 10.1109/IMTC.2005.1604583

[2] U. Tröltzsch, R. Gruder, O. Kanoun, A. Buchholz, V. Beck: Anwendungspotenziale der Impedanzspektroskopie für die Waschlaugensensorik. GMA/ITG-Fachtagung Sensoren und Messsysteme, 650-661 (2012); doi: 10.5162/sensoren2012/6.2.4

[3] R. Gruder, O. Kanoun: Water Quality Assessment by Combining Impedance Spectroscopy Measurement with Cyclic Voltammetry. AMA Conf. – Sensors, Opto and IRS2, 164-169 (2013); doi: 10.5162/sensor2013/A8.1

[4] A. Abdelkafi, P. Büschel, A. Fendri, O. Kanoun: Impedance Investigation of Milk Dilution. AMA Conf. – Sensors and IRS2, 156-159 (2015); doi: 10.5162/sensor2015/A7.2

[5] A. Fendri, P. Büschel, A. Abelkafi, O. Kanoun, A. Buchholz: AdBlue Quality Control using Impedance Spectroscopy. AMA Conf. – Sensors and IRS2, 830-832 (2015); doi: 10.5162/sensor2015/P7.3

[6] A. Fendri, H. Ghariani, O. Kanoun: Dielectric Spectroscopy for Assessment of Water Content in Edible Oils. Int. Multi-Conf. on Systems, Signals & Devices, 728-732 (2017); doi: 10.1109/SSD.2017.8167027

DOI 10.5162/sensoren2019/2.1.1

Kontinuierliche Feinstaubpartikelklassierung – von Emissionsüberwachung zu Prozessanwendungen

Dr. rer. nat. Alexander Schladitz
SICK Engineering GmbH, Bergener Ring 27, 01458 Ottendorf-Okrilla

Zusammenfassung

Der Vortrag beschreibt ein sich in der Entwicklung befindliches Messgerät (Funktionsmuster) zur kontinuierlichen Feinstaubpartikelklassierung. Es überwindet die in der behördlichen Emissionsmessung notwendige Kalibrierung gegenüber einem Standardreferenzverfahren durch die direkte Ausgabe der Staubmassenkonzentration und darüber hinaus ihrer Feinstaubanteile $PM_{2.5}$ und PM_{10}. Das beschriebene Verfahren verhält sich unabhängig von Staubart und Größe der Partikel und kann im Bereich der Prozessüberwachung als auch in einer verkleinerten Version im Bereich der Arbeitsschutzüberwachung, z.B. in Montagehallen, Anwendung finden.

Keywords: Feinstaub, Prozessüberwachung, PM_{10}, Emission, optische Messverfahren

Problemstellung Emission

Durch den Einsatz hocheffizienter Filtertechniken, z.B. Elektrofilter bei Braunkohlekraftwerken und Gewebefilter bei Müllverbrennungsanlagen sinkt die Gesamtstaubmassenkonzentration industrieller Anlagen immer weiter. Gleichzeitig bestehen die internationalen, europäischen und nationalen Interessensgemeinschaften auf die Durchsetzung strengerer Emissionsgrenzwerte. So betrug 2007 der Emissionsgrenzwert für Feuerungsanlagen für feste Brennstoffe nach 13. BImSchV. noch 20 mg/m³, während er mittlerweile auf 10 mg/m³ abgesenkt wurde. Bei den Abfallverbrennungsanlagen nach 17. BImSchV. wurde der Grenzwert im gleichen Zeitraum ebenfalls um die Hälfte auf mittlerweile 5 mg/m³ abgesenkt. Es ist daher absehbar, dass die von den Anlagenbetreibern einzuhaltenden Emissionsgrenzwerte weiter sinken. Hinzu kommt sinkende Margen aufgrund der Energiewende und damit steigender Kostendruck auf die Anlagenbetreiber. Staubmessgeräte für Emissionsanwendungen sind nach dem heutigen Stand der Technik noch nicht in der Lage ein Staubmassenkonzentrationswert direkt zu messen. Stattdessen wird die Lichtschwächung im Kamin (Transmissometer, Opazimeter) durch Vergleich der Intensität des gesendeten und empfangenen Lichtstrahls bestimmt (siehe Abb. 1).

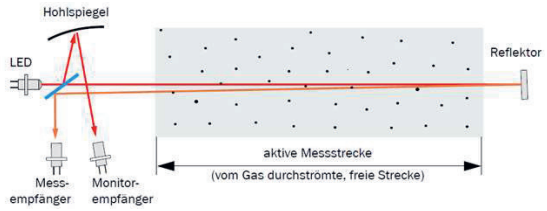


Abb. 1: Prinzipskizze des Transmissions-verfahrens

Alternativ wird die Streuung von Laserlicht an Partikeln mit einem Detektor in Vorwärts- oder Rückwärtsrichtung erfasst (siehe Abb. 2).

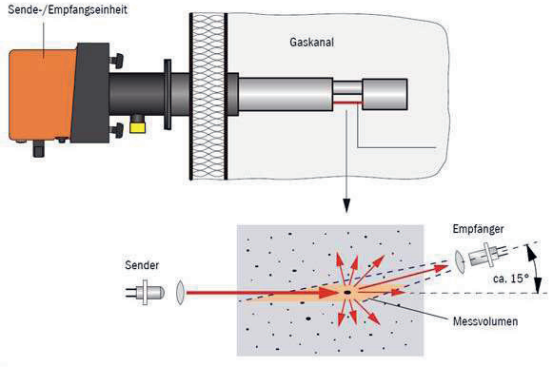


Abb. 2: Prinzipskizze des Streulichtverfahrens

Beide Verfahren geben nur ein relatives Signal (meist in mA) aus, die der Menge des Staubgehaltes im Kamin entspricht. In der Praxis ist es daher notwendig diese sogenannten automatischen Messeinrichtungen (AMS) gegenüber dem gravimetrischen Standardreferenzverfahren (SRM) nach [1] zu kalibrieren (siehe Abb. 3). Aufgrund der Staubkonzentrationen hinter der Abgasreinigungsanlage in der Nähe von Null ein schwieriges und langwieriges und damit teures Unterfangen für den Anlagenbetreiber. Dabei werden enorme Anstrengungen unternommen, z.B. durch das Eindüsen von zusätzlichem Staub oder durch Verringerung der Effizienz der Abgasreinigungsanlage, die Staubkonzentration künstlich zu erhöhen, um so den SRM-Messwert über die Bestimmungsgrenze zu bringen. Alternativ erlaubt der Gesetzgeber die Erhöhung der Probenahmedauer des SRM zur besseren Quantifizierung sehr geringer Staubmassenkonzentrationen oder die Verwendung sogenannter Ersatzstoffe deren Partikelgrößenverteilung ähnlich dem des Abgases entspricht. Dies sind unserer Ansicht nach Übergangslösungen für einen neuen Technologieschritt, da sie nicht Lösungen für das eigentliche Problem liefern sondern im Gegenteil die Kosten für die Anlagenbetreiber für die Erstkalibrierung oder jährliche Funktionsüberprüfung nach [2] unnötig in die Höhe treiben.

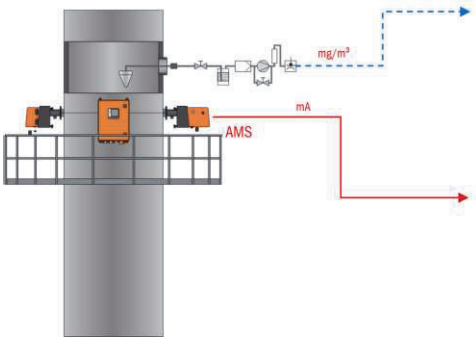


Abb. 3: Kalibrierung des AMS gegenüber dem SRM am Abgaskamin einer Anlage.

Studien von [3] zeigen das beim Großteil der untersuchten Anlagen, darunter zählen Kraftwerke auf Basis von Kohle, Öl und Holz, Zementindustrie, Schmelzen, Glasproduktion, Chemiebetriebe und Anlagen zur Abfallverbrennung, der Anteil von PM_{10} am Gesamtstaub 90% oder höher liegt (siehe Abb 4). Für $PM_{2.5}$ (siehe Abb. 5) liegt der Anteil am Gesamtstaub bei 50 – 90%. Der entscheidende Unterschied am Anteil zwischen den Anlagen wird durch den Prozess selbst bestimmt, d.h. wurde etwas thermisch verbrannt, so sind die Partikel kleiner oder wenn es sich um einen mechanischen Prozess oder eine Mischung aus beidem handelt, ruft dies größere Partikel hervor. Eine zweite Studie von [4] durchgeführt an einem finnischen Kohlekraftwerk mit Rauchgasentschwefelung und Taschenfilter weist einen PM_1 Anteil von 62% auf, wobei das Maximum der Volumengrößenverteilung bei 0.68 µm liegt. Dabei wurde eine PM_{10} Massenkonzentration im Mittel von gerade einmal 0.12 mg/Nm³ gemessen.

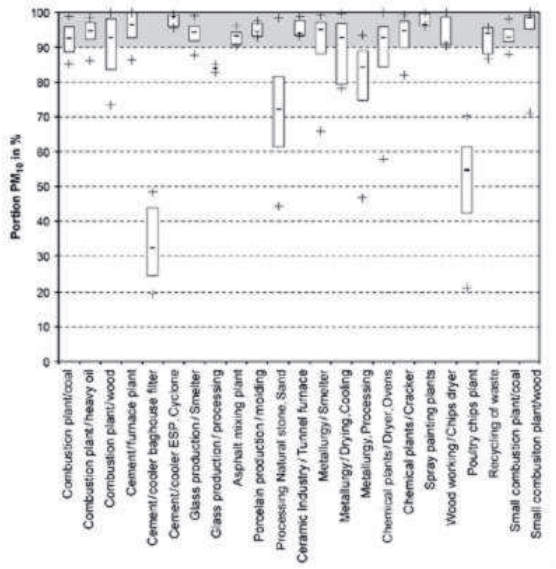


Abb. 4: Anteil von PM_{10} am Gesamtstaub für untersuchte Anlagen (entnommmen aus [3]).

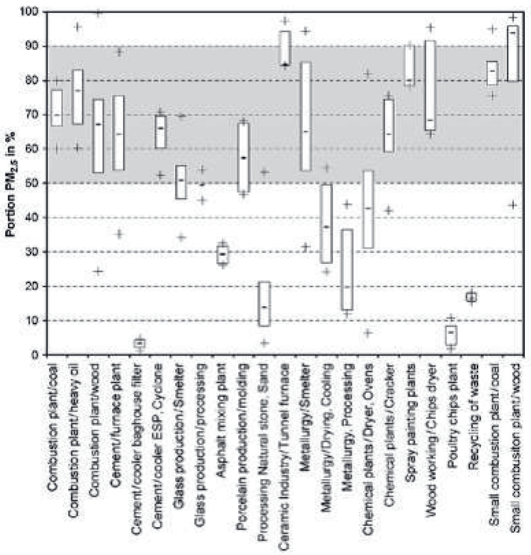


Abb. 5: Anteil von PM2.5 am Gesamtstaub für untersuchte Anlagen (entnommen aus [3]).

Dies verdeutlicht wie wichtig es ist, neben der Gesamtstaubemission, zusätzlich die Fraktionen PM_{10}, $PM_{2.5}$ oder gar PM_1 zu bestimmen.

Darüber hinaus gibt es Anfragen von Kunden und Anlagenbetreibern neben der Gesamtstaubmassenkonzentration ebenfalls kontinuierlich die Partikelgrößenverteilung, z.B. in den Fraktionen PM_{10}, $PM_{2.5}$ und PM_1 messen zu wollen, um die Produktions- und Verbrennungsprozesse bzw. die Abgasreinigung weiter zu optimieren. Dies trägt auch zu einer Reduzierung der

Immissionsbelastung in den umliegenden Gebieten dieser Anlagen bei. Zur Ermittlung der PM_{10}- und $PM_{2.5}$-Fraktionen in den Emissionen stationärer Quellen nach [5] oder im Rahmen der Ursachenermittlung für Immissionen von PM_{10} und $PM_{2.5}$ ist es notwendig, Feinstäube auch in den Abgasen industrieller Anlagen zu messen. Letztendlich leistet die kontinuierliche Überwachung der PM-Fraktionen einen wertvollen Beitrag für die Gesundheit des Menschen und damit einen individuellen Mehrwert. Es ist eine Frage der Zeit, wann sich auch in der Emissionsmessung Grenzwerte zu den genannten Fraktionen durchsetzen werden, sofern es eine Technologie gibt, die diese einfach und verlässlich bestimmen kann.

Problemstellung Prozessüberwachung und andere Anwendungen

Bei Veränderung der Brennstoffart, Filterbruch oder Fahrweise der Anlagen ändert sich meist die Partikelgrößenverteilung mit. Aufgrund der oben beschriebenen Kalibrierung ist jedoch das automatische Messsystem mit der Kalibrierkurve an eine bestimmte Brennstoffart, Fahrweise der Anlage und somit Partikelgrößenverteilung gekoppelt. Das Problem ist exemplarisch in Abb. 6 dargestellt.

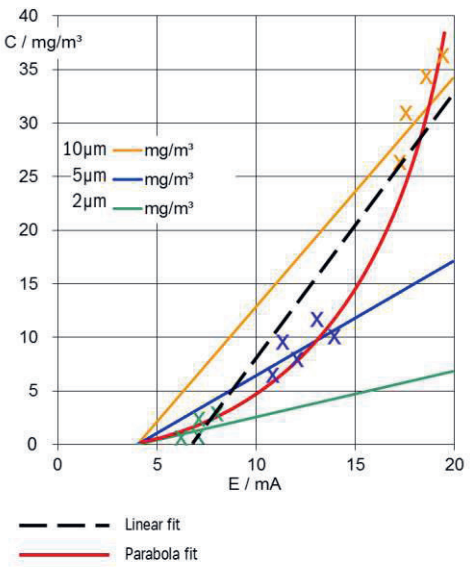


Abb.6: Verschiedene Kalibriergeraden zwischen gravimetrischen Verfahren (y-Achse) und automatischen Messsystem (x-Achse) je nach Abhängigkeit der Partikelgröße (2 µm, 5 µm und 10 µm).

Ändert sich bei einem Filterbruch die mittlere Partikelgröße von 2 µm auf 10 µm und die Kalibriergerade wurde bei 2 µm mittlerer Partikelgröße bestimmt, kommt es zu einer Unterschätzung der wahren Staubmasse, da das automatische Messsystem diesen Zuwachs in der Partikelgröße nicht erfassen kann.

Weiterhin existieren verbindliche Grenzwerte für Staub im Bereich Arbeitsschutz oder in der Halbleiterindustrie die überwacht werden müssen. Durch ebenfalls herabgesetzte Grenzwerte, z.B. im Bereich des Arbeitsschutzes von 3 mg/m³ auf 1.25 mg/m³ für aveolengängigen Staub (A-Staub) [6] oder der Ausweitung von zu messenden Partikelklassen in der Halbleiterindustrie [7], ergeben sich neue Anforderungen an die Messverfahren. Beispielsweise ist die Bestimmungsgrenze für optische Größenspektrometer aufgrund der Verwendung von Laserdioden im sichtbaren Bereich bei ~0.2 µm Partikelgröße begrenzt. Im Bereich des Arbeitsschutzes werden z.T. heute noch tragbare diskontinuierliche Messgeräte eingesetzt, die auf der Trägheitsabscheidung von Partikeln auf Sammelplatten und anschließender Wägung basieren. Eine Messmethode die zum einen kontinuierlich arbeitet und zum anderen auch Partikelgrößen bis <0.1 µm analysiert würde in diesem Bereich ein Fortschritt erzielt.

Lösungsvorschlag und wissenschaftlichtechnischer Fortschritt

Die Fa. SICK verfügt mit dem PSD-Prototyp bereits über einen kontinuierlich messenden Feinstaubpartikelanalysator, der o.g. messtechnische Anforderungen erfüllt. Es handelt sich hierbei um ein extraktives System mit genauer Absaug-Volumenstrommessung. Dies ist insbesondere im Bereich der Emissionsmessung notwendig, da hier eine isokinetische Probenahme entsprechend der geltenden Richtlinien gefordert wird. Nach erfolgter Probenahme wird das Messgas mit den enthaltenen Partikeln in eine optische Messkammer hindurchgeführt bei der eine optische Detektion der Partikel erfolgt (siehe Abb. 7).

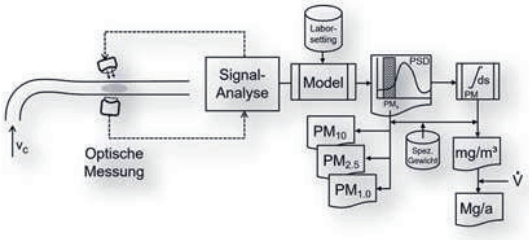


Abb. 7: Systembeschreibung eines kontinuierlich messenden Feinstaub-partikelanalysators

In der Messkammer wird die winkel- und wellenlängenabhängige Partikelstreuung mittels geringer Anzahl von Detektoren gemessen. Eine Signalanalyse wird im Industrie-PC des Prototypen durchgeführt und transformiert die gemessenen Detektorwerte mittels Laborsetting in ein physikalisches Messsignal. Ein exaktes mathematisches Modell berechnet daraus periodisch das Teilchenvolumen über der Partikelgröße (sogenannte PSD-Kurve). Aus der PSD-Kurve werden durch partielle Integration im Bereich von 0.1 bis 30 µm Partikeldurchmesser unter Annahme einer Partikeldichte (spezifisches Partikelgewicht) beliebige Partikelmassenkonzentrationen berechnet. Ein Fortschritt der PSD-Technologie gegenüber derzeitig auf dem Markt befindliche Technik ist, dass es ohne initiale Kalibrierung an der Anlage auskommt. Das hat zum Vorteil, dass die vom Feinstaubpartikelanalysator generierten Messergebnisse unabhängig von der verwendeten Staubart sind. Die auf dem Markt befindlichen optischen Messgeräte, z.B. [8], werden hingegen meist mit einer Staubart kalibriert. Im praktischen Einsatz liegen aber meist andere Staubarten als bei der Kalibrierung vor, mit der Konsequenz dass dies einen Einfluss auf die Genauigkeit der gewonnenen Messwerte hat.

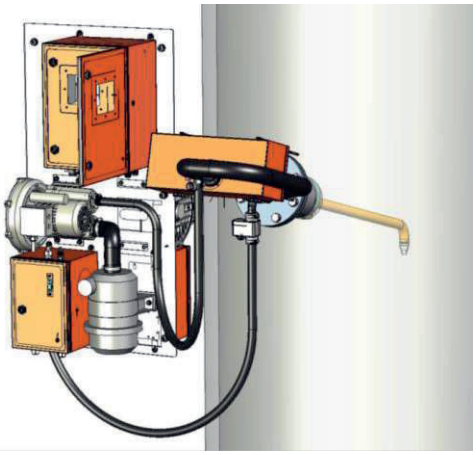


Abb. 8: Prototyp des kontinuierlich messenden Feinstaubpartikelanalysators für Emissionsanwendungen.

Abbildung 8 zeigt das Funktionsmuster eines kontinuierlich messenden Feinstaubpartikelanalysators für Emissionsanwendungen mit seinen Komponenten. Dabei handelt es sich um ein Probenahmerohr mit der das Messgas mit einem nachgeschalteten Ejektor aus dem Kamin gefördert wird. Um den nötigen Unterdruck am Ejektor zu erzeugen, existiert eine Gebläsetafel, die gefilterte Luft ansaugt und dem Ejektor zu Verfügung stellt. Die angesaugte Luft wird gemeinsam mit der Abgasluft mithilfe eines Rückführschlauches in den Kamin zurückgeleitet. Eine zweite Gebläsetafel an der Rückseite stellt zweifach gefilterte und beheizte Spülluft zur Verfügung und sorgt für die Vermeidung der Kontamination der optischen Flächen mit Staub und der Vermeidung von Kondensation des heißen Prozessgases (> 160°C) in der Ansaugstrecke bis zur Sensoreinheit. In der Kontroll- und Auswerteeinheit erfolgt hauptschlich die Steuerung der Gerätekomponenten und die Auswertung der Detektorsignale zu einer Größenverteilung auf einem Industrie-PC.

Im September 2018 fand eine Vergleichsmessung des Funktionsmusters an der Emissionssimulationsanlage der HLNUG in Kassel statt mit der folgenden Zielsetzung:

- quantitativer Vergleich für die Messgröße Gesamtstaub, PM_{10} sowie $PM_{2.5}$ gegenüber dem gravimetrischen Standardreferenzverfahren
- Untersuchung des Einflusses verschiedener Partikelgrößen auf das Messergebnis
- Verifikation der Messmethode im typischen Konzentrationsbereich von Anlagen zwischen 0.5 und 10 mg/m³

Für die Untersuchung des Einflusses verschiedener Partikelgrößen auf das Messergebnis wurden zwei verschiedene Stäube aus dem Material Calciumcarbonat mit einer mittleren Partikelgröße von 3 µm (eskal300, siehe auch Abb. 11) und 5 µm (eskal500) ausgewählt.

Abbildung 9 zeigt die Gegenüberstellung für Gesamtstaub (TSP). Die Linearität im Bereich von 1 bis 8 mg/m³ wurde erbracht, ebenso die Unabhängigkeit von der Partikelgröße. Das Funktionsmuster zeigt jedoch einen Minderbefund von 39%, der vermutlich auf Partikelablagerungen zurückzuführen ist. Dieser Faktor muss beim späteren Seriengerät mit berücksichtigt werden.

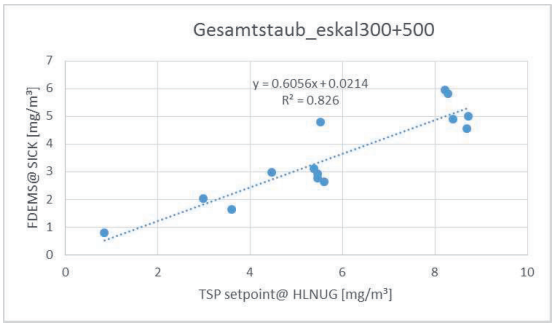


Abb. 9: Gegenüberstellung aller Messwerte des kontinuierlichen Feinstaubmessgerätes und des gravimetrischen Referenzverfahrens für die Messgröße Gesamtstaub.

Analog wird in Abbildung 10 die Linearität der Messgröße PM_{10} gegenüber dem Standardreferenzverfahren [9] gezeigt, ebenfalls mit einem leichten Minderbefund von 26%.

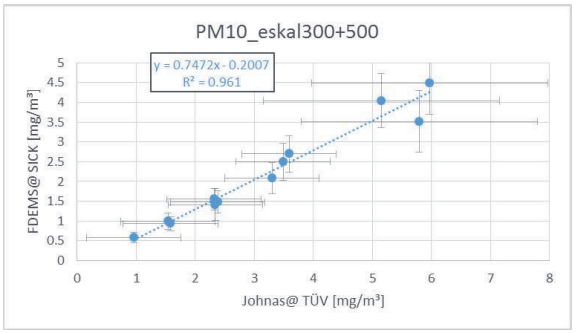


Abb. 10: Gegenüberstellung aller Messwerte des kontinuierlichen Feinstaub-messgerätes und des gravimetrischen Referenzverfahrens für die Messgröße PM_{10}.

Abschließend wird in Abbildung 11 eine Gegenüberstellung der gemessenen Größenverteilung mit dem Datenblatt für den verwendeten Staub eskal300 [10] gezeigt.

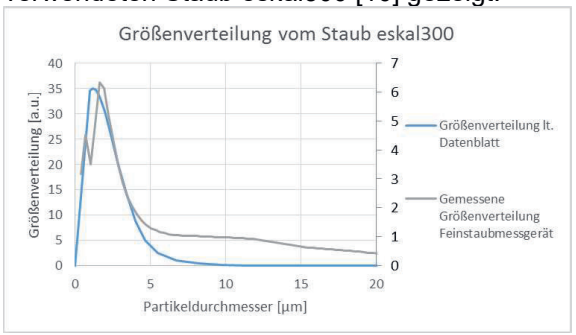


Abb. 11: Vergleich der Größenverteilungen für die Staubart eskal300.

Literaturnachweis

[1] EN13284-2:2015: Emissionen aus stationären Quellen–Ermittlung der Staubmassenkonzentration bei geringen Staubkonzentrationen–Teil 2: Automatische Messeinrichtungen.

[2] EN14181:2014: Emissionen aus stationären Quellen–Qualitätssicherung für automatische Messeinrichtungen

[3] Ehrlich, C., Noll, G., Kalkoff, W.-D., Baumbach, G., and Dreiseidler, A., „PM10, PM2.5 and PM1.0-Emissions from industrial plants-Results from measurement programmes in Germany.“, Atmos. Environ. 41(29), 6236-6254 (2007)

[4] Saarnio, K., Frey, A., Niemi, J. V., Timonen, H., Rönkkö, T., Karjalainen, P., Vestenius, M., Teinilä, K., Pirjola, L., Niemelä, V., et al., “Chemical composition and size of particles in emissions of a coal-fired power plant with flue gas desulfurization.” J. Aerosol Sci. 73, 14-26 (2014)

[5] EN ISO 23210:2009: Emissionen aus stationären Quellen–Ermittlung der Massenkonzentration von PM10/PM2,5 im Abgas–Messung bei niedrigen Konzentrationen mit Impaktoren

[6] Technische Regeln für Gefahrstofffe (TRGS 900) – Arbeitsplatzgrenzwerte, Fassung vom 14.05.2018.

[7] Reinräume und zugehörige Reinraumbereiche – Teil 12: Klassifizierung der Luftreinheit anhand der Nanopartikelkonzentration (ISO/DIS 14644-12:2013); Deutsche Fassung prEN ISO 14644-12:2013

[8] Europäische Patentschrift EP 0 391 256 B1: Vorrichtung und Verfahren zum Bestimmen der Korngrössenverteilung und der Gesamtkonzentration von Partikeln in einem Gas, insbesondere in Luft vom 05.01.1994.

[9] Verein Deutscher Ingenieure (VDI 2066-Blatt 10): Messung der Emissionen von PM10 und PM2.5 an geführten Quellen nach dem Impaktionsverfahren vom Oktober 2004.

[10] eskal300: KSL Staubtechnik GmbH, Westendstraße 11, 89415 Lauingen.

DOI 10.5162/sensoren2019/2.1.2

Microwave based Electrical Read-Out of Optical Sensors

Ulrich Nordmeyer[1], Niels Neumann[1], Xiaozhou Wang[1], Torsten Thiel[2], Konstantin Kojucharow[3], Dirk Plettemeier[1]
[1]Chair for RF and Photonics Engineering, TU Dresden, 01062 Dresden, Germany
[2]AOS GmbH, Overbeckstr. 39a, 01139 Dresden
[2]KMDC, Zur Bleiche 15, 01279 Dresden
ulrich.nordmeyer@tu-dresden.de

Abstract

Optical fiber sensors cover a wide range of applications. They offer versatile advantages including resilience against electromagnetic interference, biocompatibility and chemical resistivity. Even in environments with restricted accessibility, integration difficulties can be overcome by using Radio-over-Fiber (RoF) technology. Conventionally, the evaluation of optical fiber sensors is carried out optically by analysing the amplitude or spectrum of either the transmitted or the reflected light. A novel approach is to feed an RF modulated laser into the optical sensor and carry out a full electrical analysis of the resulting RF signal, which is changed by the sensors characteristics. This method will be investigated in the present paper along with different sensors and their applicability for this scheme will be discussed.

Keywords: optical fiber sensors, RF modulated laser, electrical analysis, wireless, radio-over-fiber

Introduction

Optical sensors are available for various quantities, e.g. temperature, strain, pressure, pH, humidity and even electrical or magnetic field. Besides that, optical sensors offer some key advantages over competing sensing principles. Their main benefit is the resilience against electromagnetic interference. Furthermore, optical sensors provide immunity to corrosion, biocompatibility, high chemical resistivity and high sensitivity. Hence, they are suitable for measurement purposes in a wide range of applications, especially for those with harsh environmental conditions. These include the monitoring of buildings, healthcare, chemical processing and many more. [1, 2, 3]
Different measuring methods based on optical sensors are well-known and share similar fundamental principles: The sensor, e.g. a fiber grating, is interfaced directly with fiber optics attached to a light source and the evaluation module. The light fed from the source into the optical sensor is evaluated after having passed the (passive) sensor. With photo detectors inside the evaluation module, the optical power, wavelength or phase of the transmitted or the reflected light may be measured.
While another key advantage of measurement setups based on optical sensors generally is the possibility of transmitting the signal over long distances without signal impairments, often times industrial environments do not allow to place the evaluation units remotely. In addition, setups with many sensors require many evaluation units which not only increases costs but also the complexity of cabling.
In environments with restricted accessibility or the need of many sensors, integration difficulties can be overcome by using Radio-over-Fiber (RoF) technology [4]. In such an RoF setup each sensor is connected to a simple transducer that converts the signal from the optical to the electrical domain and transmits it wirelessly to a central station, which again carries out the signal evaluation for all sensors together. The optical-to-wireless converter basically consists of a photodiode, an amplifier and an antenna. For this principle to work along with such a simplified transducer, it is necessary to modulate the light source with an RF signal according to the desired carrier frequency of the radio transmission.
As explained above, conventionally the evaluation of optical fiber sensors is carried out optically by analysing the amplitude or spectrum of either the transmitted or the reflected light. Since the signal in RoF setups is RF modulated, an alternative and novel approach is to carry out a full electrical analysis of the received signal, which is changed by the sensor's characteristics. This method provides several benefits and is especially qualified for use in RoF applications since an RF modulation is present in those already, but can also be adapted to non-RoF applications by simply applying an RF modulation to the existing setup.

Sensors

A typical application scenario for optical sensors is the measurement of temperature with fiber gratings. Their working principle is based on a periodic variation of the refractive index in the core of a single mode fiber. The grating acts as an interference filter with its specific resonance wavelength λ_R according to

$$\lambda_R = \Lambda \, \Delta n_{eff} \qquad (1)$$

with the grating period Λ and the difference of the effective refractive indices of the involved modes Δn_{eff} [5]. Fiber gratings having a period of some hundred nanometres belong to the sub-category of Fiber Bragg Gratings (FBG). The optical transmission characteristic of FBGs corresponds to a small band-stop filter around its resonance wavelength. Wider stop bands can be achieved by sweeping the grating period continuously from one end to another. Those gratings with a non-constant but consistently modulated grating period are called chirped FBGs. Depending on a change of temperature, the grating period increases or shrinks slightly. These changes again affect the resonance wavelength and the whole filter characteristic in form of a shift in wavelength. A measured change in the ratio of transmitted to inserted optical power of a stabilized laser source can be transferred back to a specific change in temperature provided that the filter characteristic is known. As controlling the temperature of the sensor leads to a more complex experimental setup, it will be kept constant while the shift of the filter characteristic relative to the inserted laser wavelength is simulated by changing the latter.

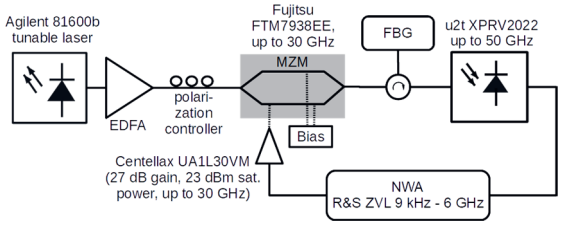


Figure 1: Schematic of the experimental setup

Setup

The approach of electrical read-out is tested with two different FBGs whereof one is chirped and the other one is not. Their characteristic parameters are shown in table 1. The setup is identical for both sensors: A laser source with tuneable wavelength (TLS) is connected to an erbium-doped fiber amplifier (EDFA) to achieve sufficient optical power. The amplifier's output is connected to a Mach-Zehnder modulator (MZM) with a polarisation controller in between. The first port of the sensor under test is connected to the output of the modulator, to the second port of the sensor a photodiode (PD) with sufficient RF bandwidth is connected. For the electrical read-out a 2-port Vector Network Analyser (VNA) is used whereas port one is connected to the MZM and port two to the PD. While the described setup integrates the sensor in transmission configuration, in the reflection variant a circulator is inserted between the MZM and the PD. One port of the sensor is connected to the circulator instead and the other one left open.

At first, the filter characteristics of both gratings have to be determined. By scanning the sensor with incident light in a wide but small stepped range of wavelengths, the sensor is analysed using an optical power meter. The resulting filter characteristic serves as reference for the outputs of the following electrical analysis.

The very same procedure is repeated by conducting full electrical S21 measurements on the VNA, that is amplitude and phase, in a frequency band from 9 kHz to 6 GHz. The electrical analysis is performed for each wavelength the TLS is set to. Beforehand the whole setup is calibrated electrically by means of forward normalization at a wavelength with minimum attenuation. For the transmission configuration that is a wavelength beyond the filter notch, for reflection measurements a wavelength at the peak attenuation shown by the optical determined reference characteristic. This not only allows to calibrate the insertion loss of the whole setup including cables, electro-optical conversion, optical loss and opto-electrical conversion but also provides a reference of the group delay of the aforementioned setup. As a consequence, the measured S21 values are independent from the used electro-optical setup.

Table 1: Parameters of sensors under test

Sensor	Stop-band	Notes
FBG L~25 mm	1537,7 nm ± 0,075 nm	apodized cos-shape
Chirped FBG L~5 mm 10 nm/cm	1536,9 nm ± 2,5 nm	apodized tanh-shape

Results

For the transmission configuration, the measured electrical amplitude characteristics of the unchirped sensor are shown at selected wavelengths in Fig. 2. Those characteristics were acquired directly with the VNA in the form of S21 measurements. From these amplitude characteristics the filter spectrum of the sensor was calculated.

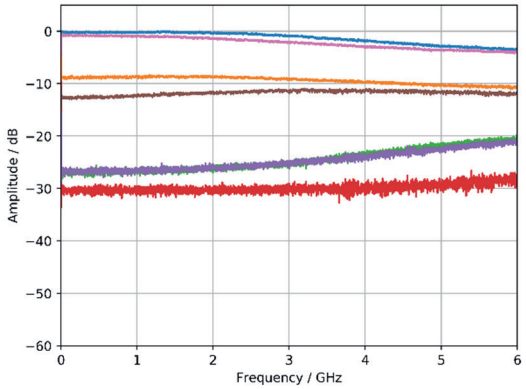


Figure 2: Electrical amplitude of S21 measurements of unchirped FBG in transmission configuration

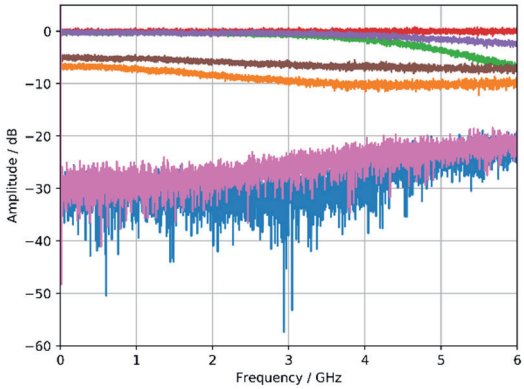


Figure 5: Electrical amplitude of S21 measurements of unchirped FBG in reflection configuration

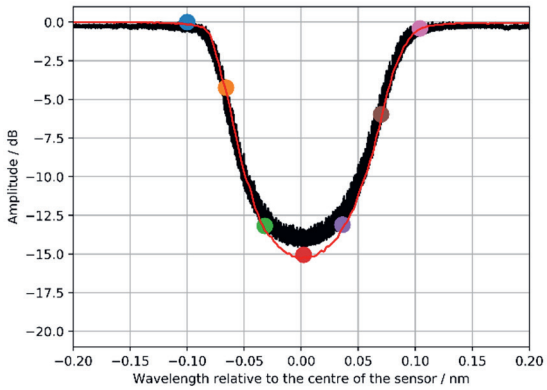


Figure 3: Electrically (red) vs. optically derived filter characteristic of unchirped sensor in transmission configuration

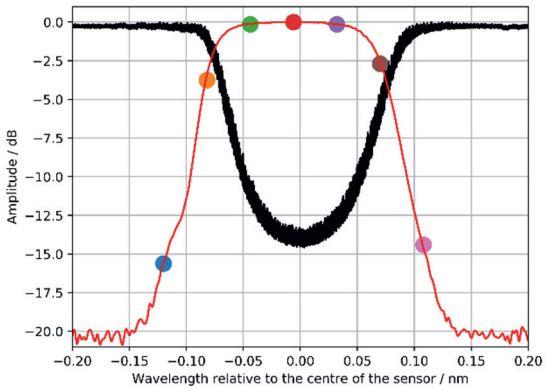


Figure 6: Electrically (red) derived filter characteristic of unchirped sensor in reflection configuration vs. optically measured characteristic in transmission configuration

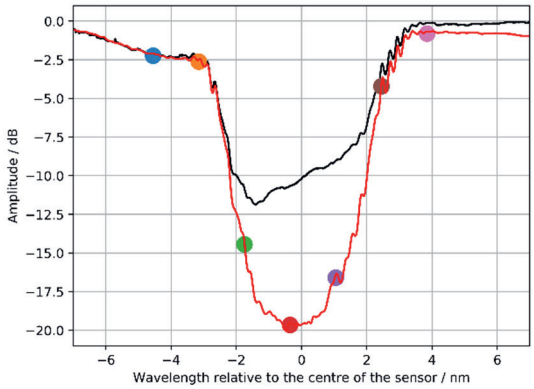


Figure 4: Electrically (red) vs. optically derived filter characteristic of chirped sensor in transmission configuration

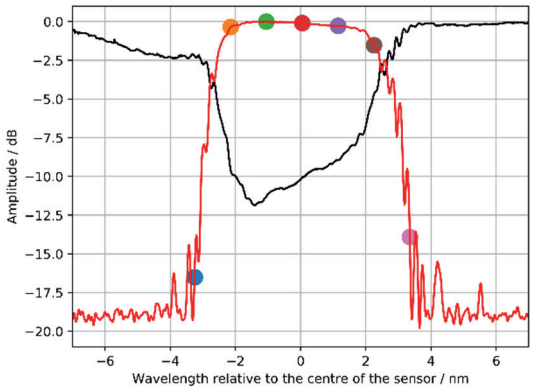


Figure 7: Electrically (red) derived filter characteristic of chirped sensor in reflection configuration vs. optically measured characteristic in transmission configuration

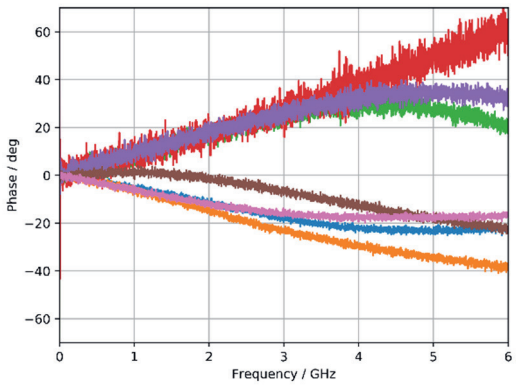


Figure 8: Electrical phase of S21 measurement of unchirped sensor in transmission configuration

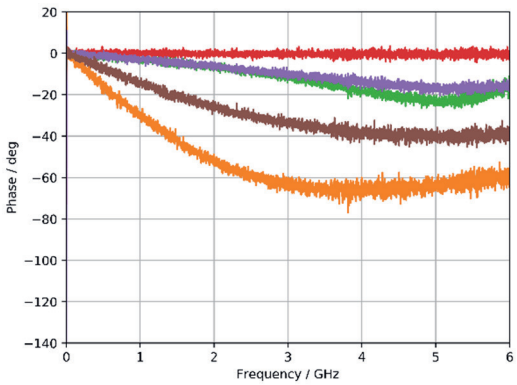


Figure 12: Electrical phase of S21 measurement of unchirped sensor in reflection configuration

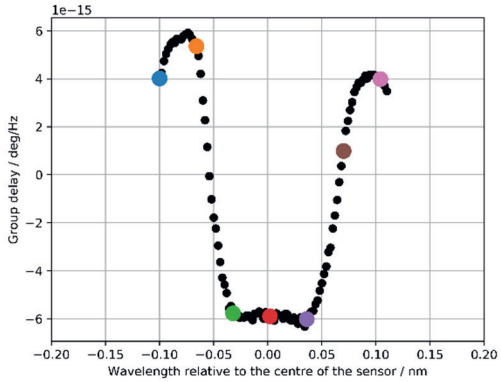


Figure 9: Group delay derived from el. Phase of unchirped sensor in transmission configuration

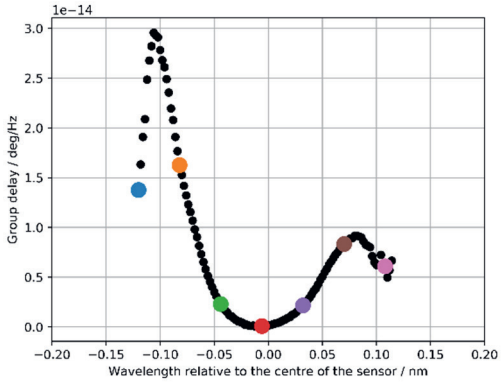


Figure 13: Group delay derived from el. Phase of unchirped sensor in reflection configuration

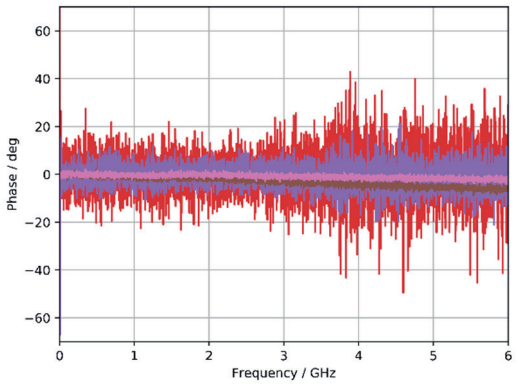


Figure 10: Electrical phase of S21 measurement of chirped sensor in transmission configuration

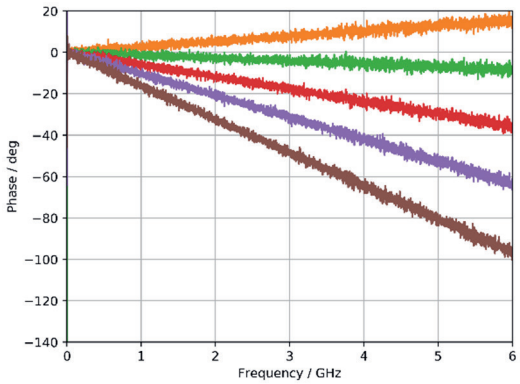


Figure 14: Electrical phase of S21 measurement of chirped sensor in reflection configuration

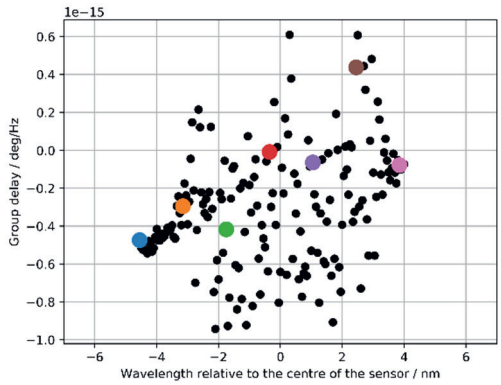


Figure 11: Group delay derived from el. Phase of chirped sensor in transmission configuration

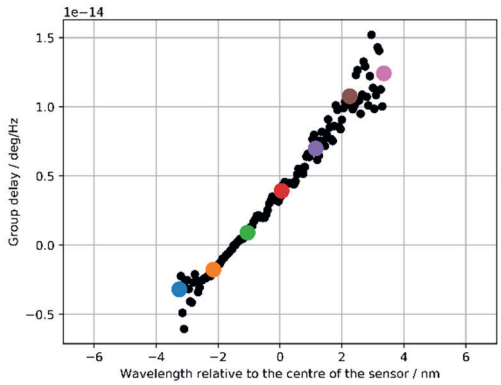


Figure 15: Group delay derived from el. Phase of chirped sensor in reflection configuration

Figure 3 shows the back traced amplitude vs. wavelength curve in comparison to the reference acquired with the optical power meter. The calculation was implemented by taking the arithmetic mean of the measured amplitude in the frequency range of 0.5 to 2.0 GHz for each measurement. To visualize the reconstruction better, Fig. 3 contains selected coloured dots that correspond to the depicted curves in Fig. 2. The same data is shown in Figs. 5 and 6 for the measurements in reflection configuration. It can be observed that the filter characteristic is inverted compared to the transmission setup.
Additionally the analogously back traced filter characteristics for the transmission and reflection measurements of the chirped sensor are shown in Figs. 4 and 7. In comparison to the unchirped sensor, the filter characteristic spans over a wider range of wavelengths.
Because the S21 measurements performed with the VNA produce complex values for the amplitude, the phase information is included in the data. The derived phase characteristics are sloped according to the expected group delay variation. For all wavelengths outside of the stop-band in the transmission setup, the group delay equals the group delay of the optical fiber whereas within the stop-band it turns zero (ideally no transmission = photonic bandgap). This behaviour can be observed for both evaluated sensors in both setup configurations. The phase measurements are shown in Figs. 8 and 9, the group delay characteristics in Figs. 10 and 11. Again, the colours highlight the affiliation to the respective amplitude graphs. Based on the phase measurements the group delay was calculated according to

$$\tau_g = -\frac{\partial \Phi}{\partial \omega}. \qquad (2)$$

Discussion

The filter characteristics calculated from the electrical read-out show reasonable results compared to the conventionally determined graphs (Figs. 4-7). For the transmission experiment with the unchirped sensor, the electrically acquired curve has almost no deviation from the reference curve (Fig. 4). Highly remarkable is the much better resolution of the transmission filter characteristic of the chirped FBG (Fig. 6). In comparison to the conventional optical read-out, the new approach is affected less by the relative intensity noise (RIN) of the laser source, the ASE noise of the optical amplifier as well as by the noise currents in the photodiode, which leads to a better signal to noise ratio (SNR).

The stop-band attenuation of the grating is so high that even with the better sensitivity of the electrical measurement especially the phase measurements are still very noisy (Figs. 7 and 12). For the same reason the derived results for the group delay are unrepresentative. Besides, the positive slopes of some phase measurements and their corresponding negative values for the group delay that can be seen, may be attributed to the calibration procedure. A normalisation to the steepest positively sloped phase measurement would correct that phenomenon but in this paper all results are evaluated qualitatively only. In contrast to the transmission setup, the reflection configuration shows very smooth phase and group delay curves within the wavelength range of the stop bands. This matches with the previously explained observations: Because in transmission configuration the stop band attenuates the signal strongly, high noise ratios are resulting, while in reflection configuration the band-stop characteristic is flipped to a band-pass characteristic with accordingly low attenuation and therefore the noise contribution remains very low in the same band.
Another key property of the electrical evaluation can be studied in Figures 2 and 3. While the positively sloped phase measurements are artefacts of the calibration process, the positive and negative slopes of the amplitude curves depend on another effect. The power of an optically modulated signal is split between the carrier and both sidebands. Considering the given measurement setup, the light of the TLS acts as the carrier while the sidebands are located symmetrically around it and their distance to the carrier is defined by the frequency of the RF-signal that is fed to the MZM. Higher modulation frequencies mean wider spectral separation of the sidebands. On the non-linear filter characteristic this can lead to different attenuations of each sideband. Considering the coloured markers in Figure 4 that correspond to the equally coloured amplitude graphs in Figure 2, that means e.g. for the measurement at the wavelength highlighted in orange: With increasing modulation frequency the sidebands move on the characteristic curve to each side of the carrier. The lower side band (LSB) moves to the right while the upper side band (USB) moves to the left. For small frequencies both sidebands remain in the almost linear slope of the curve and the lower attenuation of the USB compensates the higher attenuation of the LSB. So the impact of the rising frequency is negligible until one sideband leaves the linear range. Then the USB stays constantly at a low attenuation while the LSB is attenuated further

until it reaches the peak attenuation. This leads to a negative slope which can be confirmed by Figure 2. For the measurement at the blue highlighted wavelength the effect is even stronger as the USB stays constantly at the lowest attenuation from the beginning. Since the photodiode only yields the mixing product of both sidebands with the carrier, the results for wavelengths on the right slope of the characteristic curve do not differ systematically from the measurements on the left slope, except from asymmetries immanent to the sensors filter curve. For all wavelengths near to the concave part of the filter curve, the presented effect produces a positive slope in amplitude. The same phenomenon can clearly be identified for the studies in reflection configuration (Fig. 3 vs. 5). This is an interesting property that on the one hand has to be taken into account when choosing the RF modulation frequency but can also be used for determining the spectral position of the edges of the grating-based sensor characteristic.

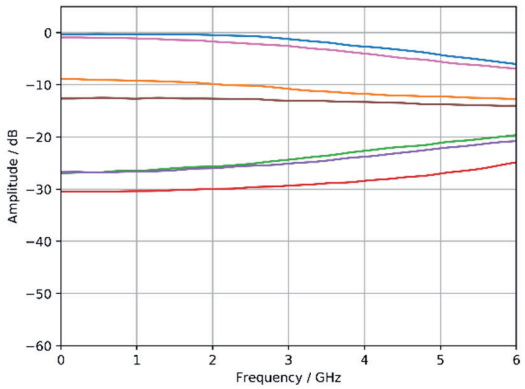


Figure 16: Analytically modelled amplitude for unchirped sensor in transmission configuration

To validate the stated theory, the electrical amplitude curves were analytically modelled according to the expected principle based on the measured filter characteristic. As it was calculated from a small band of low frequencies (equivalent to a DC measurement), it is not affected by the phenomenon beforehand. All calculations were performed using a modulation index of m=1 which equals a total power of 33% in both sidebands together. The results are depicted in Figure 16 and indicate the courses of curves as expected. Regarding higher frequencies the deviations from the measurements increase, but the synthetically generated data essentially matches the measured data. Further investigations should be carried out to identify the influence of different values for the modulation index and the non-constant group delay characteristic regarding the slopes of the amplitude measurements.

Conclusion

In this paper, the electrical read-out of optical sensors has been studied using Fiber Bragg Gratings and an RF-modulated laser source.
An introduction to optical sensors, their manifold measurement abilities and the conventional optical read-out has been given. Also, the context to Radio-over-Fiber applications has been pointed out where the novel approach originates from.
The results have shown new possibilities regarding the evaluation of optical sensors like the evaluation of the electrical phase and the group delay. In this context, the potentials and limitations of measurements in transmission versus reflection configuration have been discussed. Furthermore, it could be revealed, that the filter characteristic of an optical sensor does affect the shape of the measured electrical amplitude curves. This effect has to be taken into account when reconstructing filter characteristics by electrical read-out to avoid distortion as an impact of asymmetrical side band power distribution. Therefore, the precise reconstruction of filter characteristics from electrical S21 amplitude measurements should be conducted in small ranges of low frequencies only.
Future investigations will cover the influence of the modulation index and the non-constant group delay regarding the electrical read-out method.

Acknowledgement

This work has been funded by the EU and the Free State of Saxony in the project "OSMAT" within the framework of ESF-InnoTeam (SAB-Nr. 100329864).

References

[1] Schuster, T.; Herschel, R.; Neumann, N. & Schäffer, C. G. Miniaturized Long-Period Fiber Grating Assisted Surface Plasmon Resonance Sensor. *Journal of Lightwave Technology*, IEEE, 2012, 30, 1003-1008
[2] James, S. W., & Tatam, R. P. (2003). Optical fibre long-period grating sensors: characteristics and application. *Measurement science and technology*, 14(5), R49.
[3] Neumann, N.; Schuster, T.; Häntzsche, E.; Nocke, A.; Plettemeier, D. Wireless Read-Out of Optical Fiber Sensors. *Sensoren Und Messsysteme*: Beiträge Der 18. ITG/GMA-Fachtagung 30. Mai - 01. Juni 2017 in Nürnberg. Berlin: VDE Verlag; 2017:197-201
[4] Neumann, N.; Schuster, T.; Plettemeier, D. Novel approach for simultaneous wireless transmission and evaluation of optical sensors. *Photonics Asia*, 2014
[5] Erdogan, T., Fiber grating spectra. *Journal of Lightwave Technology* 15(8), 1277-1294 (1997)

DOI 10.5162/sensoren2019/2.1.3

Aktorisches Miniaturspektrometer für die Gassensorik

Martin Schädel[1], *Julia Baldauf*[1], *Nicole Thronicke*[1], *Dennis Mitrenga*[1], *Dominik Karolewski*[1], *Hans-Georg Ortlepp*[1], *Adrian Grewe*[2], *Stefan Sinzinger*[2], *Thomas Ortlepp*[1]

[1] CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Deutschland

[2]Technische Universität Ilmenau, Institut für Mikro- und Nanotechnologien MacroNano® Fachgebiet Technische Optik, Am Helmholtzring 1, 98693 Ilmenau, Deutschland

Zusammenfassung

Der Beitrag zeigt den bisherigen Zwischenstand der Entwicklung eines neuartigen MEMS-basierten Infrarot-Spektrometers für die Gassensorik. Dieses Spektrometer besteht aus diffraktiven optischen Elementen (DOE) mit einer gezielt hohen chromatischen Aberration, welche auf einer Aktormembran angeordnet sind, einem Blendensystem und Thermopilesensoren. Die DOE in Kombination mit den Blenden ermöglichen eine gezielte Auswahl des Spektralbereiches.
DOE-Design und Fertigung, sowie Membrandesign und Fertigung werden in diesem Beitrag detailliert erläutert.

Keywords: Gassensorik, optische Messverfahren

Ziel der Entwicklung

Das Ziel ist die Entwicklung eines Sensors für die Konzentrationsanalyse von Gasen und Gasgemischen auf dem Prinzip der Infrarotabsorption. Die im Fokus stehende Detektoreinheit nutzt eine innovative Methode für die Gewinnung eines kontinuierlichen Absorptionsspektrums. Diese Methode basiert auf der Kombination von diffraktiven optischen Elementen (DOE) mit gezielt hoher chromatischer Aberration, welche sich auf einem Aktor befinden, durch den der Abstand zwischen den DOE und dahinter befindlichen Blenden eingestellt werden kann. Je nach Abstand zwischen Blende und DOE passiert ein anderer Wellenlängenbereich die Blende und wird vom Thermopile erfasst. Durch die Kombination von Spule und Permanentmagnet kann der Abstand zwischen DOE und Blende mit Hilfe einer Aktuatormembran stufenlos eingestellt werden. Auf diese Weise kann ein kontinuierliches Absorptionsspektrum aufgenommen werden. Die Erweiterung auf mehrere Kanäle erlaubt das Erfassen eines mehrere Mikrometer breiten Spektralbereichs.
Abb. 1 zeigt einen schematischen Schnitt des geplanten Gasmesssystems und der geplanten Infrarotdetektoreinheit mit durchstimmbarer spektraler Empfindlichkeit. Die infrarote Strahlung, durchläuft ein Gasgemisch und wird durch die beiden verschiedenen DOEs auf die darunter angebrachten Blenden fokussiert.

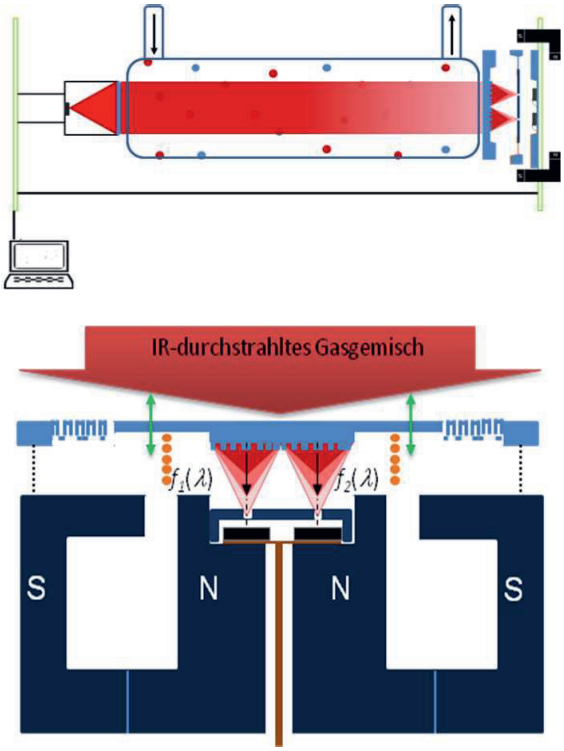


Abb. 1: Oben: Schematischer Schnitt der geplanten IR-Detektoreinheit mit Infrarotquelle, Gasmesszelle und Detektoreinheit, Unten: Detektoreinheit mit durchstimmbarer spektraler Empfindlichkeit.

Design und Fertigung der DOEs

DOEs mit einer hohen chromatischen Aberration werden häufig für Messsysteme, welche nach dem chromatisch-konfokalen Prinzip zur Abstandsmessung arbeiten, eingesetzt [1]. Die für das chromatisch-konfokale Messprinzip zu beachtenden Design-Kriterien (Arbeitsabstand, Messbereich, axiale und laterale Auflösung, erlaubte Verkippung des Testobjektes) [1] unterscheiden sich von den hier betrachteten Design-Kriterien für die DOEs und das Gesamtsystem.

Ausgehend von den Design-Parametern für die DOEs (Linsendurchmesser ≥ 1mm, Brennweite etwa 1 mm, zentrale Wellenlänge 4,3 µm, Verstimmbereich ±250 µm, Substratdicke 300 µm) wurden optische Simulationen auf Basis der entsprechenden DOE und Blenden mit Hilfe kommerzieller Software (Zemax) durchgeführt. Dabei wurde der Blendendurchmesser variiert, um den Einfluss verschiedener Systemparameter auf die Gesamtintensität und die Auflösung des Systems zu untersuchen.

Es wurden drei DOE (siehe Tab. 1) mit Linsendurchmessern von 1,5 mm simuliert.

Tab. 1: Parameter der simulierten DOEs

	DOE1	DOE2	DOE3
Brennweite (λ=4,3 µm)	0,8 mm	1,0 mm	1,6 mm
Wellenlängenbereich	3,59 µm - 5,16 µm	3,64 µm - 5,14 µm	3,79 µm - 4,95 µm
Numerische Apertur	0,68	0,60	0,42

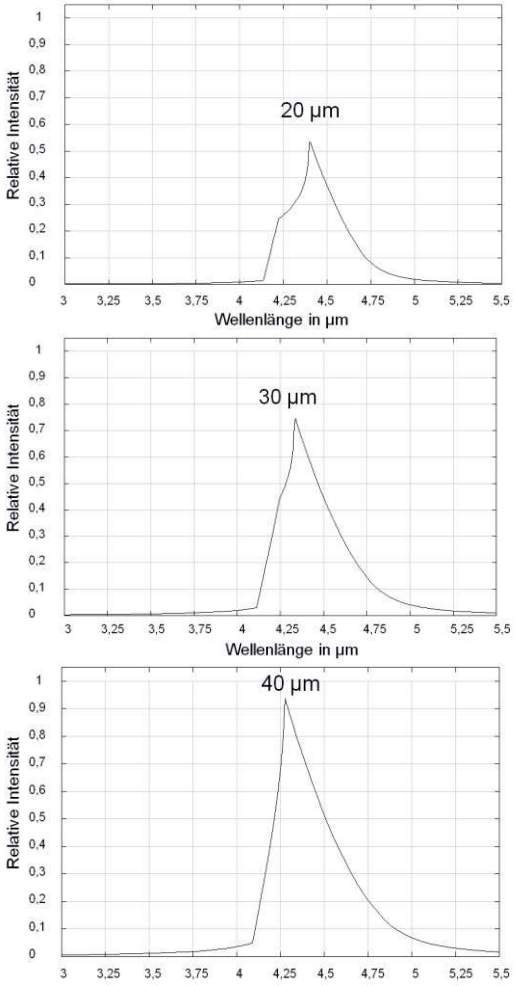


Abb. 2: Simulierte relative Intensität für DOE1 in Kombination mit verschiedenen Blenden.

Anhand der Simulationsergebnisse für verschiedene Blendendurchmesser wurden geeignete Blendendurchmesser evaluiert.

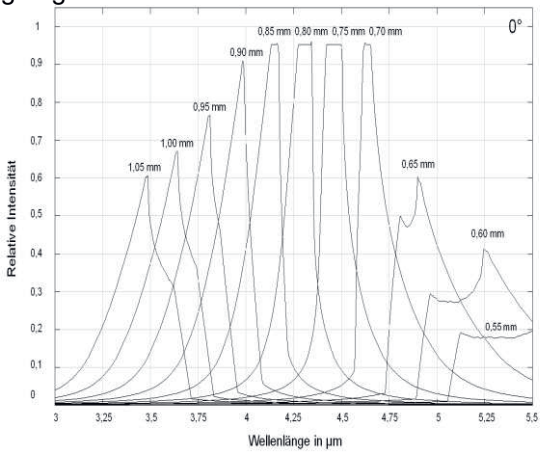


Abb. 3: Simulierte relative Intensität in Abhängigkeit von der Wellenlänge für ein System aus DOE1 und einer Blende von 30 µm Durchmesser für verschiedene Abstände mit Lichteinfall unter einem objektseitigen Winkel von 0°.

In Abb. 3 ist beispielhaft die simulierte relative Intensität in Abhängigkeit von der Wellenlänge für DOE1 und einer Blende von 30 µm Durchmesser für verschiedene Abstände zwischen DOE und Blende aufgetragen. Anhand dieser Simulation ist zu erkennen, dass es theoretisch möglich ist die angestrebte Sensorfunktionalität auf Basis der Systemaufbaues zu realisieren.
Es wurde außerdem simuliert wie sich eine Änderung des objektseitigen Winkelbereiches auf die relative Intensität in Abhängigkeit von der Wellenlänge auswirkt.

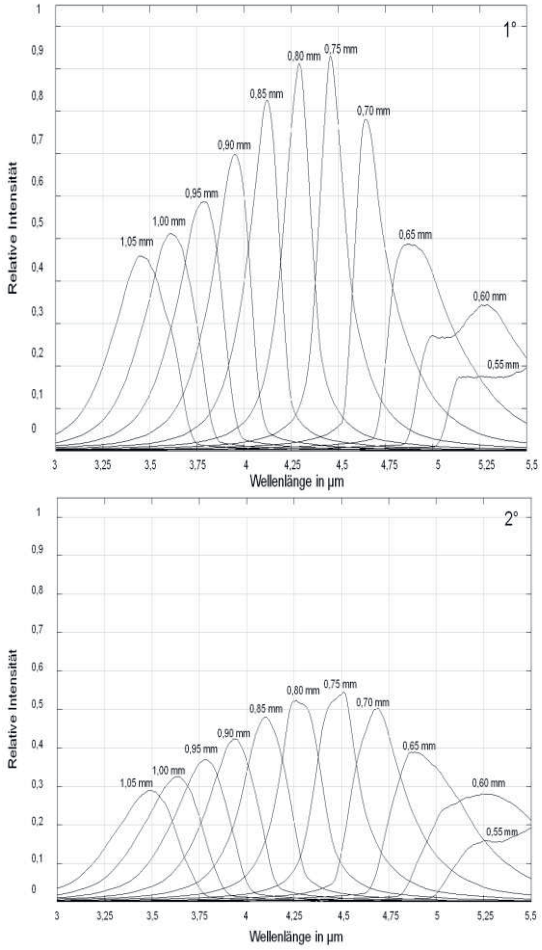


Abb. 4: Simulierte relative Intensität in Abhängigkeit von der Wellenlänge für ein System aus DOE1 und einer Blende von 30 µm Durchmesser für verschiedene Abstände mit Lichteinfall unter objektseitigen Winkeln von 1° und 2°.

Die zuvor simulierten DOEs wurden mit Hilfe von Photolithographie und Inductively Coupled Plasmaätzprozess in einem zweistufigen Verfahren (siehe Abb. 5) erzeugt. Dazu wurde Photolack aufgebracht, dieser mit Hilfe einer ersten Maske strukturiert und die entsprechenden Strukturen durch einen ersten Ätzschritt ins Silizium übertragen. Daran schloss sich eine zweite Beschichtung mit Photolack, Belichtung mit einer zweiten Maske und ein zweiter Ätzschritt mit etwa der doppelten Tiefe statt. Die durch diesen Ätzprozess gefertigten DOEs wurden mittels Rasterelektronenmikroskop und Weißlichtinterferometer hinsichtlich ihrer Struktur untersucht (siehe Abb. 6).

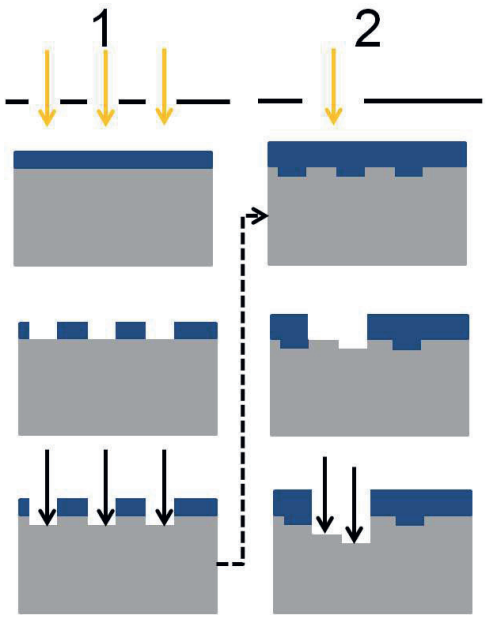
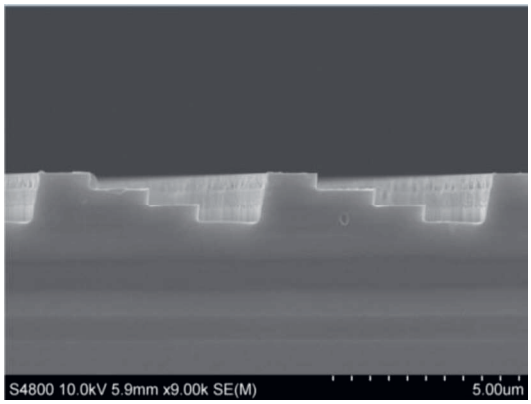


Abb. 5: Schematische Darstellung der Belichtung und der Ätzprozessfolge.



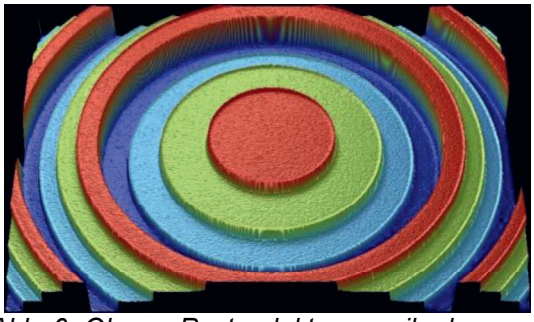


Abb. 6: Oben: Rasterelektronenmikrokop-aufnahme eines vierstufigen DOE, Unten: Weißlichtinterferometer-aufnahme eines vierstufigen DOE.

Design und Fertigung der Aktormembranen

Ähnlich aufgebaute Gesamtsysteme mit aktorischen Membrandesigns wurden in [2] für die Erzeugung eines Mikrolinsenscanners verwendet. Die hier verwendeten Aktormembranfedern wurden im Design an die in [3] simulierten Federstrukturen angelehnt.
Abb. 7 zeigt Siliziummembranen mit Federstrukturen, welche im System als Membranen für den Aktor dienen.

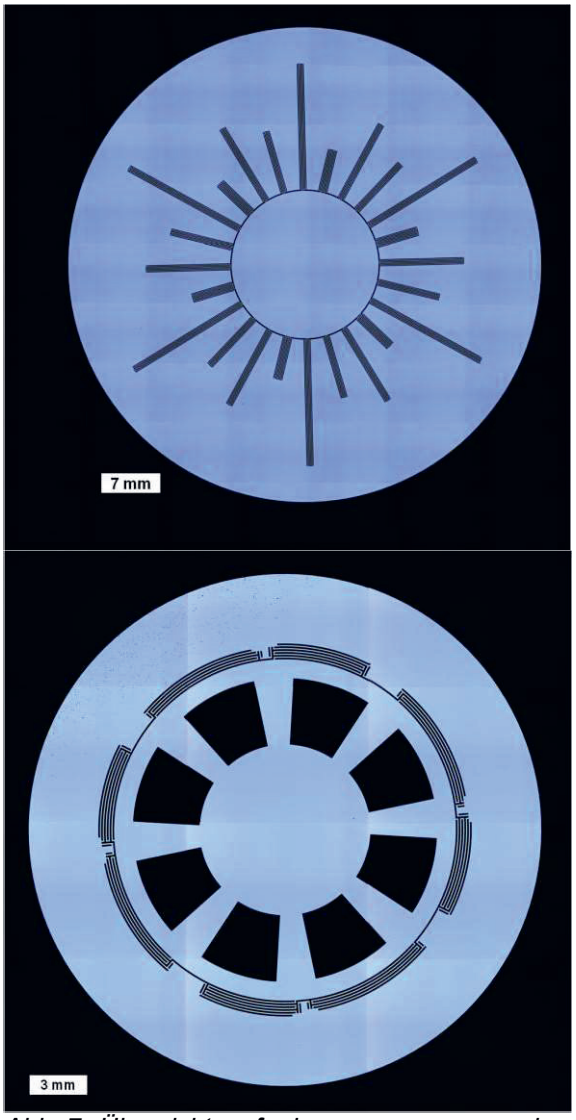


Abb. 7: Übersichtsaufnahmen zweier Aktormembranen.

Detaillierte Aufnahmen der Federstrukturen der beiden Aktormembranen sind in Abb. 8 dargestellt.

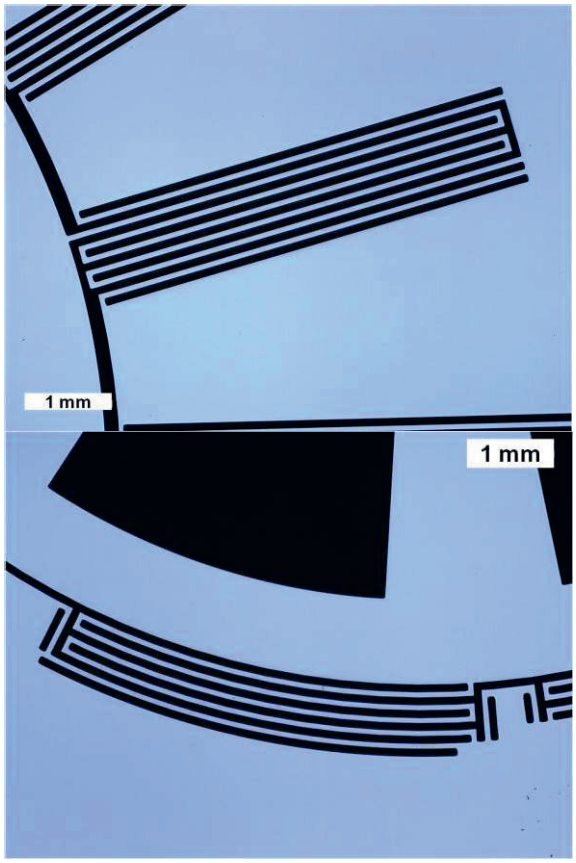


Abb. 8: Aufnahmen der Federstrukturen der Aktormembranen.

Auf Basis von mechanischen Simulationen konnten Abschätzungen zu erreichbaren Auslenkungen verschiedener Aktormembranen durchgeführt werden (siehe Abb. 9).

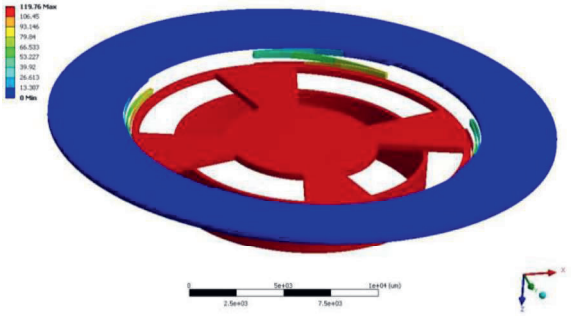


Abb. 9: Mechanische Simulation der Auslenkung einer Aktormembran mit Federstruktur.

Die Fertigung der Aktormembranen erfolgte mittels Inductively Coupled Plasmaätzprozess. Rasterelektronenmikroskopaufnahmen der Federstrukturen sind in Abb.10 gezeigt.

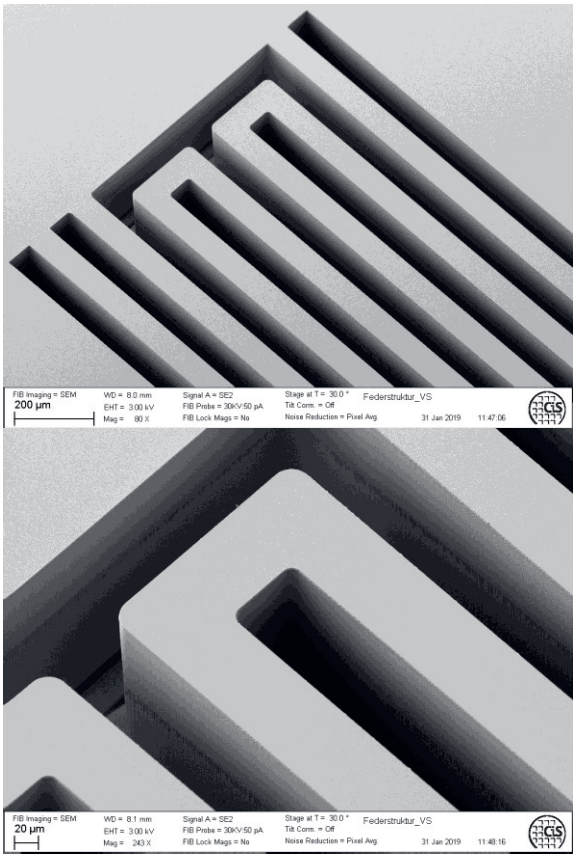


Abb. 10: Rasterelektronenmikroskopaufnahmen von Federstrukturen.

Danksagung

Die Forschungs- und Entwicklungsarbeiten für das Aktorisches Miniaturspektrometer für die Gassensorik wurden im Projekt „AMIGA“ gefördert durch das Bundesministerium für Wirtschaft und Energie, InnoKom Förderkennzeichen MF170005

Literaturnachweis

[1] M. Hillenbrand, B. Mitschunas, C. Wenzel, A. Grewe , X. Ma, P. Feßer , M. Bichra, and S. Sinzinger, Adv. Opt. Techn. 1, 187–194 (2012); doi: 10.1515/aot-2012-0017

[2] S. Bargiel, C. Gorecki, T. Verdot, K. Laszczyk, J. Albero, and L. El Fissi, Procedia Engineering 5, 762–765 (2010); doi:10.1016/j.proeng.2010.09.220

[3] G. Barillaro, A. Molfese, A. Nannini, and F. Pieri, Journal of Micromechanics and Microengineering 15, 736–746 (2005); doi: 10.1088/0960-1317/15/4/010

DOI 10.5162/sensoren2019/2.1.4

Quantum cascade laser-based tunable laser absorption spectroscopy for the detection of stable isotopes of CO_2

P. Nitzsche[1], C. Dinc[1], K.Schmitt[1] and J. Wöllenstein[1, 2]
[1] Department of Microsystems Engineering (IMTEK), University of Freiburg, Georges-Köhler-Allee 102, 79110 Freiburg, Germany
[2] Fraunhofer Institute for Physical Measurement Techniques (IPM), 79110 Freiburg, Germany
Tel.-Nr.: +49 761 203-96808, E-Mail: ponkanok.nitzsche@imtek.uni-freiburg.de

Abstract

We present a conceptual design and results on the determination of three stable carbon dioxide (CO_2) isotopologues. The isotopic analyses have been performed using a quantum cascade laser (QCL)-based tunable laser absorption spectroscopy (TLAS) setup. A mid-infrared (MIR), tunable and narrow linewidth QCL is used as a light source to probe absorption lines at around 4.3 µm. Tuning the QCL in this selected region permits the simultaneous detection of three absorption lines of CO_2 isotopologues including $^{16}O^{12}C^{16}O$, $^{16}O^{13}C^{16}O$ and $^{16}O^{12}C^{18}O$. The result shows that the wavelength conversion method using the reference gas cell is well-suited for this technique.

Keywords: Gas Sensors, Optical Measuring System, Absorption Spectroscopy, Quantum Cascade Laser, Isotopic Composition Analysis.

Motivation

Measuring and monitoring of stable isotopic compositions of carbon dioxide (CO_2) play a crucial role in many research and industry fields such as in ecology [1], atmospheric sciences [2], breathe gas analysis [3] and combustion processes [4]. Since every photosynthesis and respiration processes of any biological systems in this planet influence unique changes in isotopic compositions, analysis of the stable isotopic fractionation of CO_2 in atmosphere provides information about this biosphere-atmosphere exchange process. This could help us to understand more about the photosynthesis process and the discrimination against the heavier isotopes such as ^{13}C- or ^{18}O-isotope by all plants [5].
Due mainly to anthropogenic activities including land use change, burning of biomass and fossil fuel, and modern industrial structure, large amount of greenhouse gases has been emitted into the atmosphere in last few decade [6]. This causes a shift in the balance of biogeochemical processes which are closely associated with an effect of global warming [7]. The study of stable isotope ratios of carbon in CO_2 enable the investigation of the climate in the past and understanding the effects of the external forcing to the climatic system [8], thus allow an improvement of the climate model and prediction of future climate conditions [9], [10]. Moreover, studying of currently change of the natural isotopic ratio can help to identify and quantify sources and sinks for atmospheric trace gases [8].

In this contribution, an optical spectroscopic sensing technology based on the so-called tunable laser absorption spectroscopy (TLAS) [11] using a quantum cascade laser (QCL) as a light source is proposed in order to measure the isotopic ratio of atmospheric CO_2. The underlying theoretical principle is presented. The criteria and methods considering for the line selection will be discussed. Based on the selected absorption lines the mid-infrared, tunable QCL around 4.3 µm is chosen. The experimental setup for the detection of CO_2-isotopic compositions is built and characterized. At last the detected signal and evaluated results are discussed.

Theoretical background

In convention, the expression of an isotopic signature or a ratio of stable isotopes is given in parts per thousand (per mil, ‰). The delta δ notation is often used to describe the isotopic abundance obtaining by comparison of isotope ratio of a sample (R_{sample}) and a standard or reference material (R_{standard}) [12]. As example, the abundance of ^{13}C-isotope in CO_2 gas yields as:

$$\delta^{13}\mathrm{C}(\mathrm{CO_2}) = \left(\frac{^{13}R(\mathrm{CO_2})_{\mathrm{sample}}}{^{13}R(\mathrm{CO_2})_{\mathrm{standard}}} - 1\right) \cdot 1000‰ \quad (1),$$

where

$$^{13C}R(\mathrm{CO_2}) = \frac{[^{13}\mathrm{CO_2}]}{[^{12}\mathrm{CO_2}]} = \frac{\text{abundance of rare isotope}}{\text{abundance of abundant isotope}} \quad (2).$$

The reference material or isotopic standard is referred to VPBD-(Vienna Pee Dee Belemnite) scale. The standard value of ^{13}C-isotopic ratio is given as $^{13}R(\mathrm{CO_2})_{\mathrm{standard\ (VPDB)}}$ = 0.01180, as determined from the measurement of NBS 19 calcite [13].

Due to an individual structure and composition of each isotopologue of any molecules their absorption spectrum shows a specific profile. Since the variation of isotopic ratio in atmospheric surrounding is very small, mid-infrared laser absorption spectroscopy is an appropriate technique used to determine the stable isotope abundances of atmospheric CO_2. By ramping the injected laser current with a sawtooth or triangle signal, the laser's wavelength is probed over one or more absorption lines of interest, thus adequate absorption lines of CO_2 isotopologues must be carefully selected. The transmitted laser intensity is detected with a photodetector after passing through the measuring gas as a signal. The Lambert-Beer law, Eq. (3) describes the relationship between the transmitted light, $I_T(\lambda)$ after propagating through the medium with an optical path length of L and the material properties including the absorption coefficient, $\alpha(\lambda)$

$$I_T(\lambda) = I_0(\lambda) \cdot e^{-\alpha(\lambda)\cdot L} \quad (3).$$

Here, λ denotes the wavelength and $I_0(\lambda)$ the initial light intensity at particular wavelength. Since the measurement sensitivity depends on the absorption path length, a long gas cell or a multi-pass cell can be used to improve the limit of detection of the setup. However, the well-known limitation of this TLAS-technique is optical interference fringes rising from the Fabry-Pérot etalons between reflecting and scattering surfaces of the optical components [14].

Lines selection

Since we would like to spectroscopically measure the ratio of CO_2-isotopologues, well-separated absorption lines of each isotope must be selected. In the experimental design the selection of absorption lines detected by TLAS is typically based on two main criteria: (1) high absorption line strength and (2) a minimal interference with other gas lines in the measuring environment. In mid-infrared the absorption line strength of CO_2 is almost two orders of magnitude larger than in near-infrared range, as illustrated in Figure 1. The abundance weighted line intensity of four CO_2 isotopologues from 1.5 µm to 5 µm is simulated applying an appropriated line list from HITRAN2016-database [15] and the simulation is depicted in Figure 1.

In order to reach a sufficient absorption line strength, especially for the less abundant isotopologues, a mid-infrared light source in 4.3 µm range is used in this TLAS-measurement. Moreover, there is sufficiently low interference with bother atmospheric gas spectra in this wavelength range, since the measurement should be performed in the atmospheric surrounding.

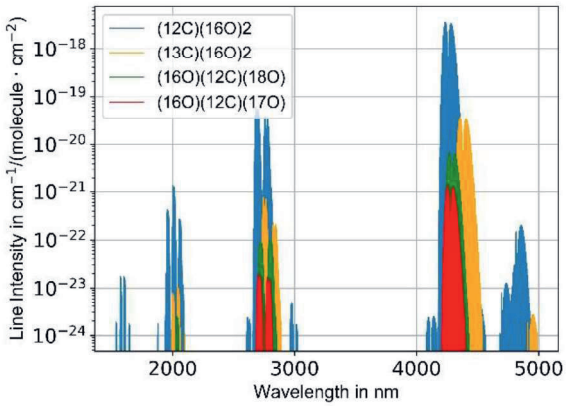


Fig. 2: The simulated line strength S_0 from 1.5 µm to 5 µm of four different isotopic compositions of CO_2 using HITRAN2016-database [15]. Blue, yellow, green and red lines represent the weighted line strength of $^{16}O^{12}C^{16}O$, $^{16}O^{13}C^{16}O$, $^{16}O^{12}C^{18}O$ and $^{16}O^{12}C^{17}O$, respectively.

However, further criteria must be considered when selecting the absorption lines for the spectroscopic detection of the isotopologues. In one tuning range of the laser's wavelength, there must be at least one absorption line for each isotopologue of interest so that one can simultaneously quantify their concentration and thus the ratio. As the weighted line strength of the most abundance isotope, Figure 2: blue line, is considerably higher than those with minority abundance, the absorption line of all measuring isotopologues should possess a similar line strength to avoid any detector saturation.

A suitable tuning range have been found, in which the spectrum of at least three different CO_2-isotopologues with comparable absorption line strength are located. The absorption spectra of four CO_2-isotopologues from 4328.1 nm to 4329.3 nm with 407 ppm CO_2 concentration which comprises 98.4204% $^{16}O^{12}C^{16}O$ (blue), 1.1057% $^{16}O^{13}C^{16}O$ (yellow), 0.39471% $^{16}O^{12}C^{18}O$ (green) and 0.0734% $^{16}O^{12}C^{17}O$ (red) are simulated and illustrated in Figure 2. According to HITRAN-database [15], this composition is referring to the natural abundance of CO_2 isotope in atmosphere.

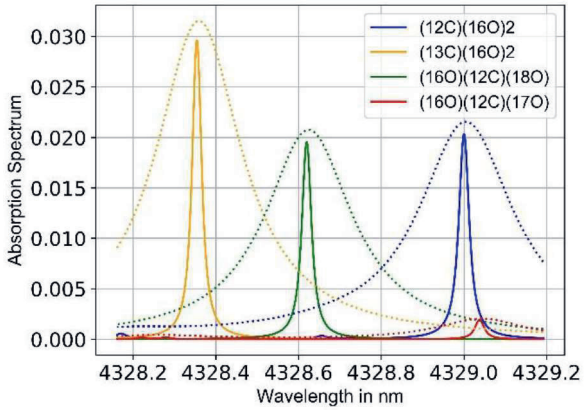


Fig. 2: The simulated absorption lines of four CO_2-isotopologues from 4328.1 nm to 4329.3 nm. The solid lines and dashed lines represent the simulation with gas pressure of 100 mbar and 1 bar, respectively [15].

The absorption path length is set to 100 cm and a temperature of 298 K. The absorption spectra are simulated with two gas pressures 100 mbar and 1 bar represented by solid and dashed lines, respectively. To minimize problems caused by pressure broadening and h the absorption line overlap, the measurement must be performed in a negative pressure chamber of 100 mbar. The absorption line of $^{16}O^{12}C^{17}O$ (red line) is however completely overlapped with the line of $^{12}CO_2$, thus the spectrum of $^{16}O^{12}C^{17}O$ in the measurement is not expected to be detected under this condition. In the selected spectral region, one can simultaneously measure three absorption lines with line center positions at 4328.354 nm, 4328.619 nm and 4329.000 nm corresponding to isotopologues $^{13}CO_2$, $^{16}O^{12}C^{18}O$ and $^{12}CO_2$, where $^{13}CO_2$ and $^{12}CO_2$ are $^{13}C^{16}O_2$ and $^{12}C^{16}O_2$, respectively.

Experimental Setup

Figure 3 (Top) depicts a scheme of experimental setup consisting of an optical system, control modules, signal acquisition module and gas flow systems. The image of the experimental setup is shown in Figure 3 (Bottom).

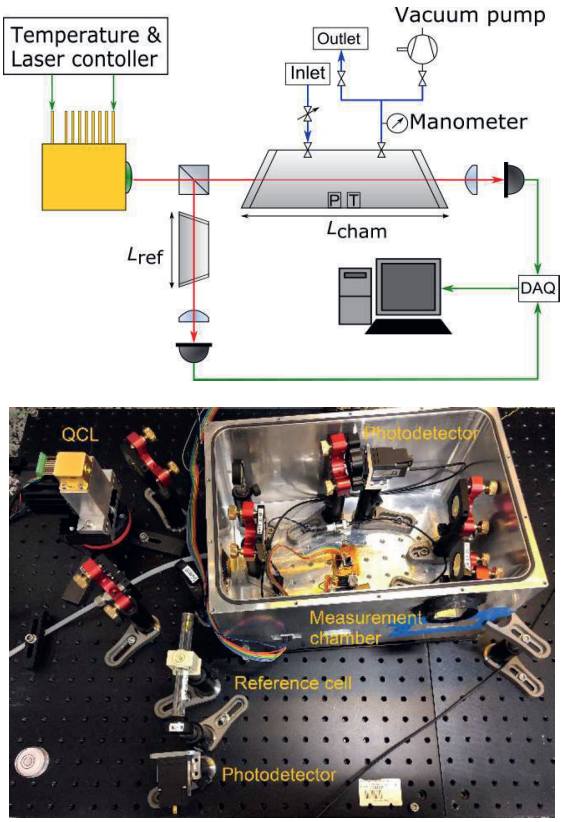


Fig. 3: Top: a schematic sketch of the experimental set-up for the QCL-based TLAS measurement. Bottom: photograph of the optical configuration of the setup.

The QCL (HHL680) from *Alpes Lasers* in HHL-package with a collimation lens and a TEC-cooling unit is controlled by a laser driver (QCL1000LAB) and a temperature controller (PTC5K-CH) from *Wavelength Electronics*. Connecting a function generator (WaveStation2012) from *LeCroy* with the laser driver enables the laser's wavelength to be tuned around 4.32 µm with a sawtooth signal and a ramp frequency of f_{ramp} = 1 Hz. The beam splitter (BSW510) from *Thorlabs* splits the laser beam in two beams, where the first beam propagates through a hermetically sealed spectroscopic gas cell filled with 98.4 % $^{16}O^{12}C^{16}O$, 1.1 % $^{16}O^{13}C^{16}O$ and 0.5 % N_2 with a total pressure of p_{ref} = 10 Torr or 13.332 mbar which acts as a reference cell. The concentration of the filled gases and the condition in the reference cell are specified from the manufacturer, *Wavelength References*. The cell has an optical path length of 10 cm. The reference signal is detected using a photodetector (PVI-3TE-5) from *Vigo System*

with a neutral density filter from *Thorlabs* (NDIR20B).
The second beam propagates through the measured gas chamber, which tilted in two dimensions reducing the interference fringes. With three gold coated mirrors (PF10-03-M01) and a concave mirror (CM254-150-M01) with focal point of 150 mm from *Thorlabs* the laser beam is directed and reflected back and forth inside the chamber to obtain a total optical path of L_{cham} = 107±2.5 cm. The absorption signal is likewise detected with Vigo-photodetector and a neutral density filter (NDIR30B). To reduce the gas pressure inside the chamber to $p_{cham} \approx 100$ mbar, a diaphragm vacuum pump (N840.3 FT.18) from *KNF* with a turn valve at the gas outlet and a fine-metering valve at the gas inlet are connected with the chamber to regulate the gas flow and pressure. Inside the measuring chamber a temperature sensor (SHT31-ARP-B) from *Sensirion* and pressure sensor (HSCMAND001BAAA5) from *Honeywell* are integrated. The detected signals including applied laser current and temperature, gas pressure and temperature in the chamber, photodetector signals and the modulation signal from the function generator are acquired using *National Instrument* DAQ cards (NI9205 and NI9239) with a sampling rate of 1 k and 25 k samples per second, respectively.

Results and discussion

According to the QCL's data sheet, the laser's temperature should be set to - 15 °C in order to reach the selected tuning range from 4328.1 nm to 4329.3 nm. In the experiment, we use the function generator to periodically change the laser's current at the chosen temperature. To ensure that the laser's wavelength alters in the desired range, we firstly observe the signal from the reference cell with an oscilloscope and compare it with the simulated spectrum to define the accordant voltage range for the sawtooth ramp. The frequency of the ramp is set to 1 Hz.

The reference signal is used not only for the determination of the set-modulation voltage but also for the calibration of laser's emitted wavelength. However, the reference signal must be recorded synchronously with the TLAS-signal so that the frequency calibration is likewise valid for the TLAS-signal. The reference signal as a function of counts is illustrated in Figure 4. The blue line is an averaged signal using a set of 20 time-series raw data corresponding to 20 s. The red curve refers to the curve fit using a self-written algorithm in Python-program.

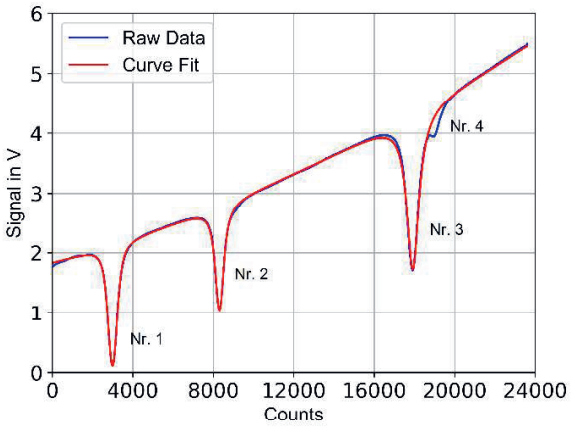


Fig. 4: The measured reference signal as a function of counts. The blue and red curves represent the average raw data and the curve-fit of it, respectively.

For the comparison, the reference signal has been simulated. The gas mixture inside the reference cell consisting of 98.4 % $^{16}O^{12}C^{16}O$, 1.1 % $^{16}O^{13}C^{16}O$, as told in the data sheet and the natural abundance for another two isotopologues is applied for the simulation. The simulated absorption spectrum is shown in Figure 5. The temperature is set to the averaged room temperature measured during the experiment T = 298.09±0.01 K and the pressure is set to p_{ref} =10 Torr.

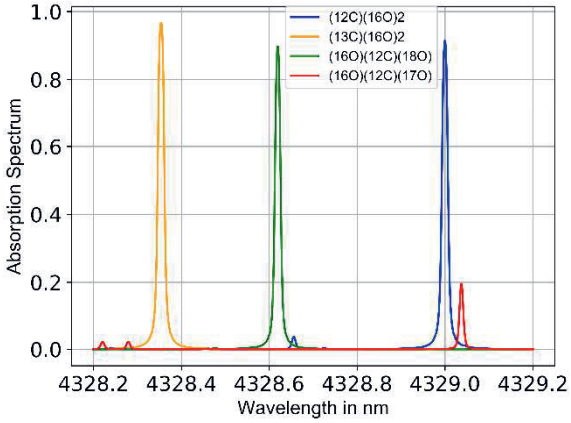


Fig. 5: The simulated reference signal of four isotopes using the HITRAN-database [10]. The temperature is set to the averaged room temperature T = 298.09±0.01 K.

Comparing the measured reference signal and those from the simulation, four absorption spectra of four different isotopologues can be identified. The spectrum of $^{12}CO_2$ and $^{13}CO_2$ masked as Nr.1 and Nr.3 can be clearly detected. The spectrum of $^{16}O^{12}C^{18}O$ (Nr.2) and $^{16}O^{12}C^{17}O$ (Nr.4) are not expected, however both spectra are detected in the measured signal. Even though the manufacturer can only confirm the filling of two isotopes with specific

concentrations, using this spectroscopic method another isotope of CO_2 in this reference gas cell can be identified.

Three center positions from the fitting parameters of the Nr.1 to Nr.3-spectrum have been taken for the calibration of the laser's wavelength. The center positions of the same spectra are likewise taken from the simulation in the unit of nm. Employing these two sets of data, a conversion-function is calculated as:

$$\lambda = a \cdot \text{counts}^2 + b \cdot \text{counts} + c \qquad (4).$$

Here, λ is the wavelength in nm and a, b, c-coefficients are - 7.0058×10^{-13}, 5.7956×10^{-5} and 4.3282×10^{3}, respectively. To gain a higher accuracy in frequency conversion a polynomial with higher degree can be used. For this application a quadratic polynomial is sufficient.

In order to measure the isotopes of atmospheric CO_2, the pressure in the measuring chamber is reduced to about 100 mbar using the diaphragm vacuum pump. In this measurement, we use the laboratory-air as a test gas. The parallel TLAS-signal is detected with the sampling rate of 25 k samples per second. To calculate the averaged signal, 20 periodical data-series are employed and the averaged signal against the number of counts is illustrated in Figure 6.

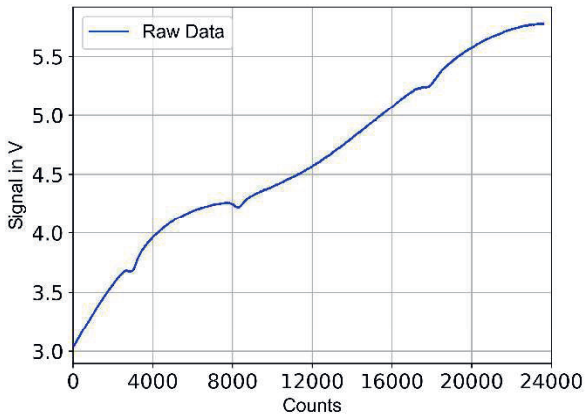


Fig. 6: The averaging of 20 data-series of TLAS-signal as a function of counts. The gas pressure and temperature in the measuring chamber is measured as p_{cham} = 99.5±0.3 mbar and T = 298.09±0.01 K.

The pressure and temperature inside the chamber is monitored using the integrated sensors during the measurement. The average values of the pressure and temperature over the 20 s is determined as p_{cham} = 99.5±0.3 mbar and T = 298.09±0.01 K, respectively.

The signal shows three absorption spectra of the isotopes as expected, however the baseline comprises strong interference fringes, which causing a difficulty in determination of the baseline. The signal around the position of absorption lines has been masked in order to determine a best-fit baseline function with a high-order polynomial function. Divided the averaged signal with this baseline function to remove the baseline from the signal, we obtained the transmittance. The negative natural logarithm of the transmittance, $-\ln(I/I_0)$ is calculated and shown in Figure 7 as a blue line. Employing the conversion-function Eq. (5) to convert the counts number to wavelength, the transmittance can be plotted as a function of wavelength.

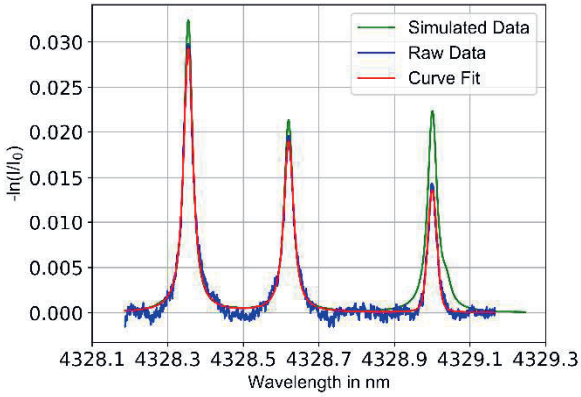


Fig. 7: The negative natural logarithm of the transmittance as a function of wavelength. Green, blue and red lines represent the simulated, raw data and the curve fit of the raw data.

A best fit-function of the transmittance spectrum has been computed using Voigt-profiles and illustrated as red line in Figure 7. The simulation of the negative logarithm of transmittance is likewise depicted as the green curve. The results show that the wavelength conversion method using information from the reference spectrum and simulation is sufficiently accurate for the calibration of the laser's frequency.
The first two spectra of the curve-fit corresponding to $^{13}CO_2$ and $^{16}O^{12}C^{18}O$ are in a well agreement with the simulation spectrum. However, the spectrum of $^{12}CO_2$ shows a deviation from the simulation. The calculated concentration obtained from this spectrum is thus not matching with the expected value, likewise the isotopic ratio. We suppose that the non-linearity of the detector could be a reason for this deviation. Further investigation of the system is required.

Summary and outlook

In this letter, the conceptual design and construction of an analysis of CO_2-isotopic

compositions using the so-called QCL-based tunable absorption spectroscopy have been presented. Three absorption lines of stable isotopologues, $^{13}CO_2$, $^{16}O^{12}C^{18}O$ and $^{12}CO_2$ in 4.3 µm range is selected. To reduce the pressure broadening effect, the measurement is performed in an under-pressure gas chamber. For the calibration of the laser's frequency, the reference gas cell is employed. The result confirms that the wavelength conversion for the TLAS-signal using the reference cell is well-matched with the simulation.
However, further works must be undertaken. These include the calculation of the CO_2-isotopic ratio, the evaluation of the long-term stability such as the thermal and pressure stabilities, and a measurement with a well-known CO_2-concentration and its isotopic compositions.

Acknowledgements

This work is partly funded by Georg H. Endress foundation.

References

[1] J. Balesdent, A. Mariotti, and B. Guillet, "Natural^{13}C abundance as a tracer for studies of soil organic matter dynamics," *Soil Biol. Biochem.*, vol. 19, no. 1, pp. 25–30, 1987.
[2] D. R. Bowling, S. D. Sargent, B. D. Tanner, and J. R. Ehleringer, "Tunable diode laser absorption spectroscopy for stable isotope studies of ecosystem-atmosphere CO_2 exchange," *Agric. For. Meteorol.*, vol. 118, no. 1–2, pp. 1–19, 2003.
[3] L. Cocola, A. Ghetti, G. Tondello, G. Galzerano, and L. Poletto, "Performance evaluation of a TDLAS system for carbon dioxide isotopic ratio measurement in human breath," *Opt. Sens. Detect. V*, no. May 2018, p. 31, 2018.
[4] M. Schumacher *et al.*, "Oxygen isotopic signature of CO_2 from combustion processes," *Atmos. Chem. Phys.*, vol. 11, no. 4, pp. 1473–1490, 2011.
[5] D. R. Bowling, D. D. Baldocchi, and R. K. Monson, "Dynamics of isotopic exchange of carbon dioxide in a Tennessee deciduous forest," *Global Biogeochem. Cycles*, vol. 13, no. 4, pp. 903–922, 1999.
[6] IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G. K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
[7] E. A. G. Schuur *et al.*, "Climate change and the permafrost carbon feedback," *Nature*, vol. 520, no. 7546, pp. 171–179, 2015.
[8] P. Ghosh and W. A. Brand, "Stable isotope ratio mass spectrometry in global climate change research," vol. 228, pp. 1–33, 2003.
[9] G. B. Bonan and S. C. Doney, "Climate, ecosystems, and planetary futures: The challenge to predict life in Earth system models," *Science*, vol. 359, no. 6375. 2018.
[10] A. Lourantou *et al.*, "Constraint of the CO_2 rise by new atmospheric carbon isotopic measurements during the last deglaciation," vol. 24, no. C, pp. 1–15, 2010.
[11] D. D. Nelson, J. B. McManus, S. C. Herndon, M. S. Zahniser, B. Tuzson, and L. Emmenegger, "New method for isotopic ratio measurements of atmospheric carbon dioxide using a 4.3 µm pulsed quantum cascade laser," *Appl. Phys. B Lasers Opt.*, vol. 90, no. 2, pp. 301–309, 2008.
[12] W. A. Brand, T. B. Coplen, J. Vogl, M. Rosner, and T. Prohaska, "Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report)," *Pure Appl. Chem.*, vol. 86, no. 3, pp. 425–467, 2014.
[13] Q. Zhang and W. Li, "A calibrated measurement of the atomic weight of carbon," *Chinese Sci. Bull*, vol. 35, no. 4, pp. 290–296, 1990.
[14] J. Hodgkinson and R. P. Tatam, "Optical gas sensing: a review," *Meas. Sci. Technol.*, vol. 24, no. 1, p. 12004, 2013.
[15] R. V. Kochanov, I. E. Gordon, L. S. Rothman, P. Wcisło, C. Hill, and J. S. Wilzewski, "HITRAN Application Programming Interface (HAPI): A comprehensive approach to working with spectroscopic data," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 177, pp. 15–30, 2016

DOI 10.5162/sensoren2019/2.2.1

Einfluss von DC- und AC-Versorgungsspannungen auf Dünnschicht-Drucksensoren aus ZrO2

M. Mathis[1], D. Vollberg[2], G. Schultes[2]
[1]FITT gGmbH, Saaruferstraße 16, 66117 Saarbrücken, Germany
[2]Hochschule für Technik und Wirtschaft des Saarlandes, Goebenstraße 40, 66117 Saarbrücken, Germany

Zusammenfassung

ZrO2 (3 mol% Y2O3 stabilisiert) eignet sich aufgrund seiner mechanischen Eigenschaften wie geringem E-Modul, hoher Biegefestigkeit und hoher Bruchzähigkeit hervorragend als Membranmaterial für Drucksensoren. In Kombination mit Dünnschichten kann ein leistungsfähiges, ausfallsicheres und kostengünstiges Sensorkonzept ohne Isolationsschicht realisiert werden. Nach diesem Konzept aufgebaute Sensoren wurden bis zu Temperaturen von 250 °C charakterisiert und zeigten ein stark von der Versorgungsspannungsart abhängiges Verhalten. Während mit Wechselspannung betriebene Sensoren keinerlei Auffälligkeiten zeigen, kommt es bei mit Gleichspannung betriebenen Sensoren zu systematischen Ausfällen. Bei diesen Sensoren zeigen sich lokale Abplatzungen der Dünnschicht, die eine starke Drift des Nullsignals bewirken. Hervorgerufen wird dieser Effekt durch die Sauerstoffionenleitfähigkeit von stabilisiertem ZrO2 und ein vorliegendes Gleichspannungsfeld.

Keywords: Drucksensorik, piezoresistive Dünnschicht, Zirkonoxid, Sauerstoffionenleitfähigkeit, Trägerfrequenzmesstechnik

ZrO2 als Federwerkstoff

Eine Standard-Drucksensorvariante zur präzisen Messung hoher Drücke besteht aus einer Stahlmembran mit Isolationsschicht und einer piezoresisitiven Dünnschicht. Die Herstellung der notwendigen Isolationsschicht ist aufwendig und teuer. Bei der Verwendung von elektrisch isolierenden, keramischen Membranmaterialien kann hingegen auf eine Isolationsschicht verzichtet werden. Standardmäßig werden daher kostengünstige Drucksensoren aus Al2O3-Membranen in Kombination mit Dickschichten eingesetzt. Diese Sensoren sind jedoch ungenauer und instabiler als Dünnschichtsensoren. Anstelle von Al2O3 kann mit 3 mol% Y2O3-stabilisiertes ZrO2 (3Y-TZP) als Membranmaterial eingesetzt werden [1]. 3Y-TZP hat im Vergleich zu Al2O3 ein um 40 % geringeres E-Modul, was bei gleicher Materialbelastung zu höheren Dehnungen und folglich einem höheren Sensorsignal führt. Gleichzeitig sind die Biegespannungsfestigkeit und Bruchzähigkeit von 3Y-TZP um ein Vielfaches höher als bei Al2O3, wodurch (bei gleicher Membranstärke) eine höhere Überlastsicherheit erreicht wird. Durch die geringe mittlere Kristallitgröße von ca. 0,4 µm kann 3Y-TZP sehr gut in Kombination mit piezoresistiven Dünnschichten eingesetzt werden. Feinkristallines 3Y-TZP wird aufgrund seiner guten Sauerstoffionenleitfähigkeit [2] ab 300 °C u.a. als Werkstoff für Lambdasonden, Brennstoffzellen und Sauerstoffpumpen eingesetzt.

Sensoraufbau

Der Aufbau der Drucksensoren mit 3Y-TZP Membranen erfolgte mittels einer Keramik-Metall-Klemmverbindung, wie in Abb. 1 dargestellt. Die durch Ringbuckelschweißen der Edelstahlkörper realisierte Verbindung dient, sowohl zur Fixierung der Keramik-membran, als auch zum Abdichten des Sensors. Wird der Sensor mit Druck beaufschlagt, führt dies zu einer Verformung der Membran. Mittels einer piezoresistiven Sensorschicht wird diese Verformung in ein elektrisches Signal gewandelt. Die Abfolge der Herstellungschritte der Sensoren wird im Folgenden kurz erläutert.
Bei Verwendung von Dünnschichten ist eine möglichst geringe Oberflächenrauheit von Vorteil. Daher wurden die 3Y-TZP-Membranen zunächst geläppt und poliert. Anschließend wurden piezoresistive Dünnschichten auf die Membranen gesputtert. Als Sensorschichten wurden hochempfindliche, kohlenstoffhaltige Schichten [3, 4], auf Chrom basierende Schichten und Platin verwendet. Die gesputterten Schichtdicken lagen zwischen 100 und 200 nm. Mittels eines

Pikosekundenlasers wurden Wheatstone-Messbrücken, gemäß Abb. 2 in die Dünnschichten strukturiert. Diese strukturierten Membranen wurden anschließend getempert und durch Ringbuckelschweißen zweier Edelstahlkörper zum oben beschriebenen Sensor komplettiert. Die Kontaktierung der strukturierten Messbrücken erfolgte mittels Aluminiumbonddrähten.

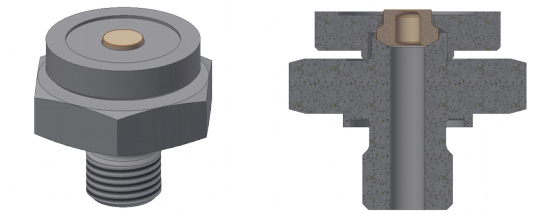


Abb. 1: links: CAD-Modell der Drucksensoren, bestehend aus einem Edelstahlgrundkörper mit Rohrgewinde, einer ZrO2-Membran und einem an den Grundkörper geschweißten Edelstahlniederhalter, rechts: Darstellung des Sensoraufbaus im Schnitt

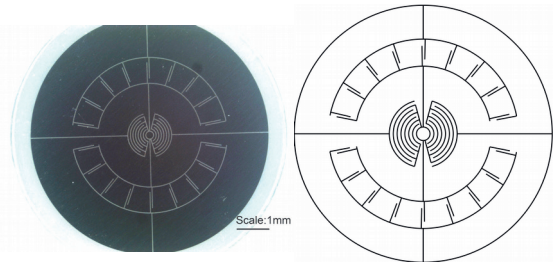


Abb. 2: links: Beschichtete und mittels Pikosekundenlaser strukturierte 3Y-TZP-Membran, rechts: Strukturierungslayout für eine Wheatstone-Messbrücke

Versuchsdurchführung

Die Charakterisierung der Sensoren wurde an einem pneumatischen Druckprüfstand mit Stickstoff als Medium durchgeführt. Der Anschluss der Messbrücke erfolgte in 6-Leitertechnik mittels eines Q.bloxx EC A106-Moduls für Messbrücken der Fa. Gantner Instruments. Die Speisung der Messbrücke kann hier sowohl mit Gleichspannung, als auch mit Wechselspannung (Trägerfrequenz) erfolgen. Die Trägerfrequenz kann auf 600 Hz oder 4,8 kHz eingestellt werden. Um die Frequenz der Versorgungsspannung variieren zu können kam eine zweite Messanordnung, bestehend aus einem Oszillator (Modell: 33120A) der Fa. Agilent, einem Lock-In-Verstärker (LIA-MVD-200-L) der Fa. Femto und einem Q.bloxx EC A106-Modul für Messbrücken der Fa. Gantner, zum Einsatz.

Der am Sensor anliegende Druck wurde mit Hilfe eines hochgenauen CPC8000 Druckcontrollers der Fa. Wika geregelt. Die Sensoren wurden zunächst bei Temperaturen von 30, 85 und 120 °C hinsichtlich Nullpunkt, Empfindlichkeit, Linearität, Hysterese und Kriechen charakterisiert. Die Speisung der Messbrücke erfolgte hierbei stets mit 5 VDC. Weiterhin wurde die Stabilität des Nullsignals der Sensoren bei 250 °C untersucht, wobei die Messbrücken mit 2,5 VDC, 5 VDC sowie 5 VAC gespeist wurden. Für den Betrieb mit Wechselspannung wurden die Frequenzen 600 Hz, 1 kHz, 4,8 kHz und 10 kHz verwendet. Ein Sensor wurde bei 200 °C hinsichtlich Nullpunkt, Empfindlichkeit, Linearität, Hysterese und Kriechen charakterisiert. Die Versorgung der Brücke erfolgte mit 5 VAC und einer Frequenz von 600 Hz.

Ergebnisse

Bei den Charakterisierungsmessungen der Drucksensoren bis 120 °C kam es zu keinen Auffälligkeiten. Es wurden Linearitätsfehler von 0,33-1,6 % bei Empfindlichkeiten von bis zu 12 mV/V (Nenndruck: 250 bar) erreicht. Bei den Stabilitätsmessungen des Nullsignals zeigte sich ein großer Unterschied zwischen den mit AC und den mit DC betriebenen Drucksensoren.

Ab Temperaturen von 150 °C wurde bei den mit DC betriebenen Sensoren eine starke Drift des Nullsignals von mehreren mV/ V, sowie starkes Rauschen des Nullsignals festgestellt. Diese Effekte zeigten sich völlig unabhängig von der verwendeten Dünnschicht. Die Ausfälle der Sensoren wurden auch von optischen Veränderungen der Dünnschicht begleitet. In allen Fällen kam es zu einer optischen Veränderung der Dünnschichtbereiche und der Kontaktpads mit dem höchsten elektrischen Potential. In Abb. 3 ist beispielhaft der betroffene Bereich der Dünnschicht dargestellt. Die hier erkennbaren Verfärbungen reichen nur in die Anfangsbereiche der strukturierten Widerstände hinein. In Bereichen mit niedrigerem Potential liegen keine Verfärbungen vor. Zur genauen Analyse der Dünnschicht wurden die ausgefallenen Sensoren mittels eines Rasterelektronen-mikroskops (REM) untersucht. Dabei konnten die optisch veränderten Bereiche der Dünnschicht, wie in Abb. 4 zu sehen, als

Bläschen und damit einhergehende Schichtabplatzungen identifiziert werden.

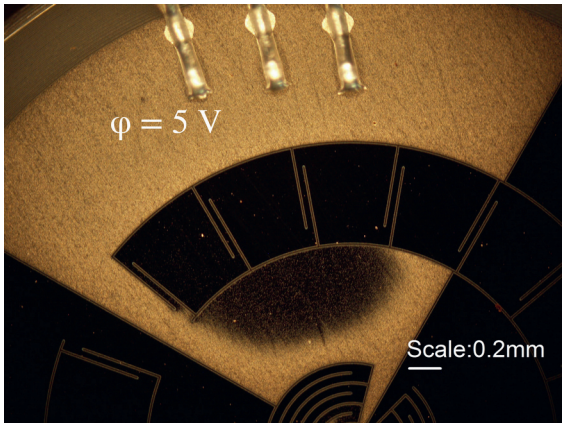


Abb.3: Mittels Dunkelfeldmikroskopie dargestelltes Kontaktpad (5 VDC) eines bei 250 °C mit 5 VDC betriebenen 3Y-TZP-Drucksensors mit kohlenstoffhaltiger Dünnschicht. Die Bereiche des Kontaktpads, die auf Versorgungsspannung liegen sind deutlich verfärbt.

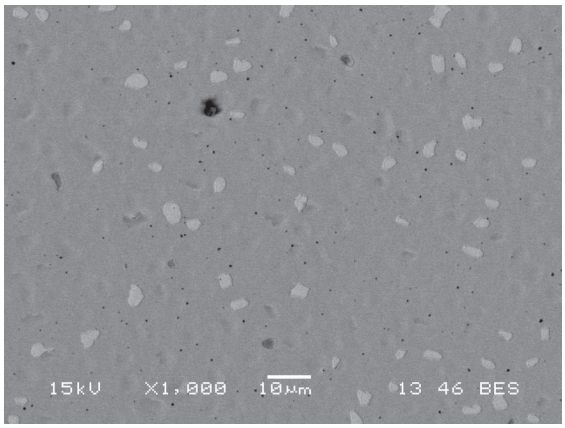


Abb.4: Mittels REM dargestelltes Kontaktpad (5 VDC) eines bei 250 °C mit 5 VDC betriebenen 3Y-TZP-Drucksensors mit kohlenstoffhaltiger Dünnschicht. Die optisch veränderten Bereiche der Dünnschicht können klar als Bläschen und Abplatzungen der Schicht identifiziert werden.

Im Gegensatz zu den geschilderten Beobachtungen bei DC-Speisung, zeigten sich bei den mit AC betriebenen Drucksensoren weder Driften oder Rauschen des Nullsignals, noch optische Veränderungen der Dünnschicht. Das Nullsignal über 90 h eines mit 2,5 VDC betriebenen und eines mit 5 VAC (Trägerfrequenz 4,8 kHz) betriebenen Sensors bei 250 °C sind in Abb. 5 dargestellt. Die Unterschiede in der Performance der Sensoren sind klar zu erkennen. Das Nullsignal der mit 2,5 VDC versorgten Messbrücke ändert sich in den 90 h um bis zu 3,5 mV/V. Bei der mit 5 VAC versorgten Messbrücke spielen sich die Änderungen des Nullsignals hingegen im µV/V-Bereich ab. Bei den Signalsprüngen dieses Sensors am Anfang und am Ende der Messung handelt es sich um den Temperaturkoeffizienten des Nullpunkts, den sogenannten TK0. Der TK0 gibt die Änderung des Nullsignals pro Änderung der Temperatur an und liegt hier bei -2,38 µV/V/K.

In weiteren Messungen zeigte sich, dass das Nullsignal der mit Wechselspannung betriebenen Sensoren auch bei geringeren Frequenzen (600 Hz und 1 kHz) und höheren Frequenzen (10 kHz) äußerst stabil bleibt. Es konnte ebenfalls festgestellt werden, dass mit 5 VDC betriebene Sensoren bei hohen Temperaturen deutlich stärkere Nullsignaländerungen zeigen als mit 2,5 VDC betriebene Sensoren.

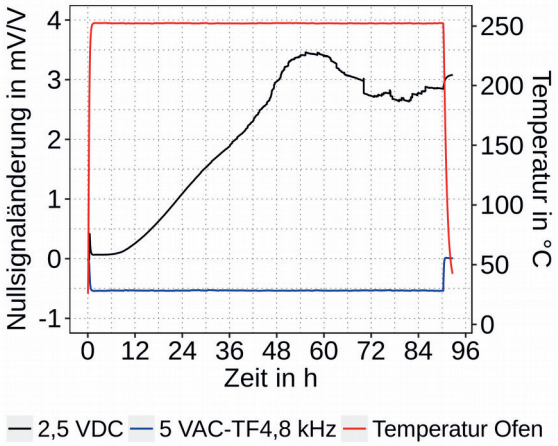


Abb.5: Nullsignaländerung über 90 h eines mit 2,5 VDC und es eines mit 5VAC (4,8 kHz Trägerfrequenz) betriebenen 3Y-TZP-Drucksensors bei 250 °C. Das Nullsignal des mit AC betriebenen Sensors bleibt im Gegensatz zum Nullsignal des mit DC betriebenen Sensors auch bei 250 °C stabil.

Um auch die Sensoreigenschaften der mit Wechselspannung betriebenen Sensoren bei hohen Temperaturen und Druck zu überprüfen wurde ein Sensor im pneumatischen Druckprüfstand bei 205 °C bezüglich Empfindlichkeit, Linearität und Kriechen charakterisiert. Bei diesem Sensor wurde zuvor über mehrere Tage bei 250 °C die

Stabilität des Nullsignals überprüft. Als Sensorschicht wurde in diesem Fall eine kohlenstoffhaltige, hochempfindliche Dünnschicht verwendet. Die ermittelten Kenngrößen des Sensors sind in Tab. 1 aufgeführt. Diese liegen auch bei 205 °C immer noch in einem akzeptablen Bereich. Die Kennlinie des Sensors ist in Abb. 6 dargestellt.

Tab. 1: Kenngrößen eines mit 5 VAC (600 Hz) betriebenen 3Y-TZP-Sensors (Nenndruck: 250 bar) bei 205 °C

Kenngröße	**3Y-TZP Sensor**
Empfindlichkeit	7,52 mV/V
Linearitätsfehler	0,15 %FS
Hysteresefehler	0,27 %FS
Kriechfehler (10 min)	0,38 %FS
Nullpunktsrückkehr	0,08 %FS

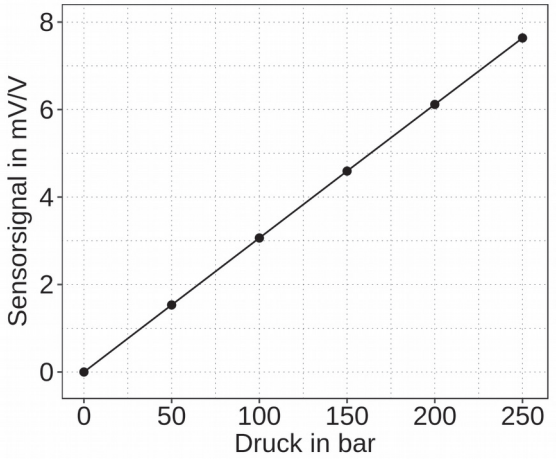


Abb.6: Kennlinie eines Drucksensors mit 3Y-TZP-Membran bei 205 °C. Der Sensor wurde mit 5 VAC (600 Hz) betrieben.

Diskussion

Die durchgeführten Versuche führten zu folgenden Erkenntnissen:
Die Art der Versorgungsspannung hat bei Drucksensoren mit ZrO2-Membran einen erheblichen Einfluss auf die Performance der Sensoren bei Temperaturen ab 150 °C. Mit DC betriebene Sensoren zeigen Ausfallerscheinungen, mit AC betriebene Sensoren hingegen nicht. Die Ausfallerscheinungen verschlimmern sich bei höheren DC-Spannungen, also stärkeren elektrischen Feldern. Daher ist es naheliegend, dass diese Effekte durch die Sauerstoffionenleitfähigkeit von stabilisiertem ZrO2 hervorgerufen werden. Dotiertes ZrO2 ermöglicht durch Gitterfehlstellen die Wanderung von Sauerstoffionen durch das Material. Ein an der Wheatstone-Messbrücke anliegendes elektrisches Feld übt Kräfte auf die zweifach negativ geladenen Sauerstoffionen aus und die Ionen wandern zum Pluspol. Da die Sauerstoffionen nicht durch die dichte Dünnschicht dringen können, kommt es zum Druckaufbau (Bläschenbildung) und schließlich zur Ablösung der Schicht. Je höher die Amplitude der Versorgungsspannung ist, desto stärker werden die Ionen zum Pluspol gezogen. Der Sensor zeigt schnellere und stärkere Ausfallerscheinungen. Wird statt der Gleichspannung eine Wechselspannung ausreichend hoher Frequenz als Versorgungsspannung der Brücke verwendet, können die Sauerstoffionen dem Feld nicht mehr folgen. Der Sensor bleibt intakt.

Das für Drucksensoren mechanisch vorteilhafte Membranmaterial 3Y-TZP kann also auch bei Temperaturen von 150 °C - 250 °C eingesetzt werden, solange als Versorgungsspannung des Sensors eine Wechselspannung mit ausreichend hoher Frequenz verwendet wird. Der hier ausgenutzte Mechanismus könnte ebenso bei der Diffusion von anderen Ionen (z.B. Wasserstoff) durch Membranmaterialien eingesetzt werden.

Literaturnachweis

[1] G. Schultes, M. Cerino, M. Langosch, T. Kuberczyk, D. Vollberg, D. Göttel, O. Freitag-Weber, A.C. Probst, tm - Technisches Messen, 83(3): 147-156 (2016); doi: 10.1515/teme-2015-0096
[2] R. Ramamoorthy, D. Sundararaman, S.Ramasamy, Solid State Ionics 123 (1999) 271–278; doi: 10.1016/S0167-2738(99)00103-4
[3] R. Koppert, et al., Diam. Relat. Mater. (2012), doi: 10.1016/j.diamond.2012.01.031
[4] G. Schultes, H. Schmid-Engel, S. Schwebke, U. Werner, J. Sens. Sens. Syst., 7, 1-11, 2018, doi: 10.5194/jsss-7-1-2018

DOI 10.5162/sensoren2019/2.2.2

Uncertainty evaluation of a traceable dynamic force transfer standard

Saher Hassan[1], Ch. Schlegel[1], R. Kumme[1], R. Tutsch[2]
[1]Physikalisch-Technische Bundesanstalt,Bundesallee 100, 38116 Braunschweig, Germany
[2]TU Braunschweig, Institut für Produktionsmesstechnik, Braunschweig, Germany
saher.r.hasan.ext@ptb.de

Abstract

A dynamic force transfer standard is designed at PTB to provide the traceability to dynamic testing metrology field. This contribution describes a method of evaluation the uncertainty of the primary dynamic force calibration of transfer standards, including the description of the uncertainty models of the sensitivity, which is the main measurand. Some influences have been reduced like the effect of the misalignment and telting effects. Other new contributions have been introduced like the contribution of the acceleration distribution correction due to elastic deformation of load masses, Sine fit, and vibration isolation of the laser vibrometer. A comparison between the estimation of uncertainty using classical GUM technique and Monte Carlo simulation has been performed.

Keywords: Dynamic force, Uncertainty, Dynamic force transfer standard.

Introduction

Dynamic force measurement plays a critical role in many industrial applications like automotive, aerospace, off shore energy, healthcare and construction. There is a need to convey traceability to this application in terms of dynamic force measurement. PTB has been developed a dynamic force transducer to act as a transfer standard for dynamic force metrology to provide the traceabililty to the dynamic force testing field. This transfer standard has been dynamically calibrated and metrologically characterised. The main tool of the metrological characterisation is the uncertainty the detailed dewscription of the dynamic calibration method has been mentioned in [1–3]. The error sources in dynamic force measurement and calibration are mentioned in [4–10].

Mathematical model

The dynamic sensitivity of a force measuring device can be expressed in terms of measurands as the ratio of the output signal of the force measuring device U_V in $\mathrm{mV/V}$ and the acting dynamic force F_{dyn} in N with a magnitude of:

$$S_f = \frac{U_V}{F_{dyn}} \tag{1}$$

While the phase of sensitivity is determined as the phase difference between laser vibrometer, which corresponds to the reference force, and the output signal of the force measuring device.

$$\Delta\varphi_{S_f} = \varphi_{F_{dyn}} - \varphi_{U_V} \tag{2}$$

The output signal of the force measuring device U_V is given by the ratio of the output signal of the bridge amplifier and the bridge excitation voltage
Using the excitation of an electrodynamic shaker, the dynamic force acting on the force transducer:

$$F_{dyn} = \int_V \rho \cdot a(x,t) \cdot dV \tag{3}$$

The distribution in acceleration $a(x,t)$ can be approximated by multiplying the acceleration a_t by a frequency-dependent correction factor, K_o.

$$K_o = \frac{sin\left(\sqrt{k_m\omega^2}\cdot l\right)}{\sqrt{k_m\omega^2}\cdot l}, \quad k_m = \frac{\rho}{E} \tag{4}$$

The acting dynamic force is then the product of all acting masses and their corresponding

accelerations. In case of measurement of acceleration on the surface of the top mass, a_t, the acting dynamic force is calculated from.

$$F_{dyn} = a_t \cdot (m_t + m_i)K_o \quad (5)$$

The internal mass m_i is determined using the equation

$$S_f = \frac{U_V}{a_t \cdot (m_t + m_i)} \quad (6)$$

That represents the dynamic sensitivity of the force proving instrument, to get the static sensitivity; $S_{f=0}$ we can rewrite the above equation like

$$\left(\frac{U_V}{a_t}\right)_{f=0} = S_{f=0} \cdot (m_t + m_i) \quad (7)$$

This lead to a linear equation between the ration of the transducer output to the top acceleration at one side and the unknown internal mass at the other side

$$\left(\frac{U_V}{a_t}\right)_{f=0} = S_{f=0} \cdot m_t + S_{f=0} \cdot m_i \quad (8)$$

With an offset equals $S_{f=0} \cdot m_t$ slope of $S_{f=0} \cdot m_i$. To get the variance of the internal mass as an influence to the uncertainty budget of the dynamic sensitivity

To calculate the dynamic sensitivity of the dynamic force measuring instrument, sinusoidal excitations at 17 different frequencies from 25 to 2000.25 Hz are performed. sine fitting is carried out by applying a least square fitting function of Matlab software. The uncertainty contribution of the fitting function is estimated using a Monte Carlo simulation that will be deployed in details in the next sections. The sampling rate was selected to be 25.625 kHz that is more than ten times the maximum excitation frequency, which satisfies the Nyquist-Shannon sampling theorem. At each frequency, a number of 2^{12} (4096) points are stored for each channel to avoid leakage error in the estimation of the sine function parameters. A number of a minimum of 10 cycles of data were recorded for each test frequency.

Elimination of tilting and misalignment effects

The acceleration of the mass measured by the laser vibrometer depends on the position of its laser beam on the top of the load mass. In this setup, 24 points are selected to make three octagons around each other. All three octagons are concentric to the load mass center as shown in Figure 1 above. Figure 1 below shows the acceleration distribution on top of the load mass at every octagon with different excitation frequencies. The more distance between measuring point and the load mass center, the higher variance in acceleration values exists. This is a combined resultant of the misalignment effect and the tilting modes. The variation of acceleration distribution on top of the load mass increases after resonance -around 1 kHz of the applied load masses- due to the increasing effect of rocking modes within higher frequencies. To eliminate the effect of tilting and misalignment effects on the resulted uncertainty, an averaging of every two points against each other has been performed. The advantage of that averaging, that if the variation of the acceleration distribution is only because of the misalignment or even a tilting vibration mode of the load mass around its center of gravity, this will lead to a standard deviation of zero. Figure 1 below shows a case vibration where the variation of the acceleration distribution on top of the load mass comes only from the tilting of the load mass around its center of gravity. The averaging of each two acceleration vector the same r coordinate and a shift of θ coordinate by π.

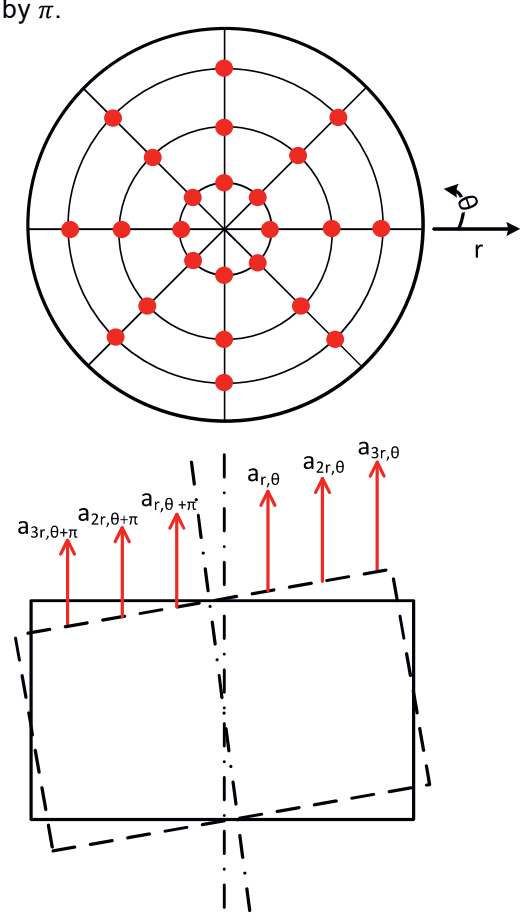


Figure 1 : Schematic of acceleration measurement on top of the load mass shows the 24 measurement points that are distributed to make three concentric octagons. top shows a top view and bottom shows section view with a dashed explanation of a mode shape that represents the rotation of the load mass around its center.

Uncertainty Analysis

The uncertainty of the sensitivity is determined as

$$u_{S_f}^2 = u_{U_V}^2 \frac{\partial S_f}{\partial U_V}^2 + u_{F_{dyn}}^2 \frac{\partial S_f}{\partial F_{dyn}}^2 \quad (9)$$

The quantity u_{rep} is the uncertainty contribution due to the error of reproducibility of the force measuring device in addition to the effects of the misalignment of between the sensing axis of the force transducer and the principal axis of motion. The different uncertainty contributions to the dynamic sensitivity of force measuring devices are shown in Figure 2, the red contributions are neglected.

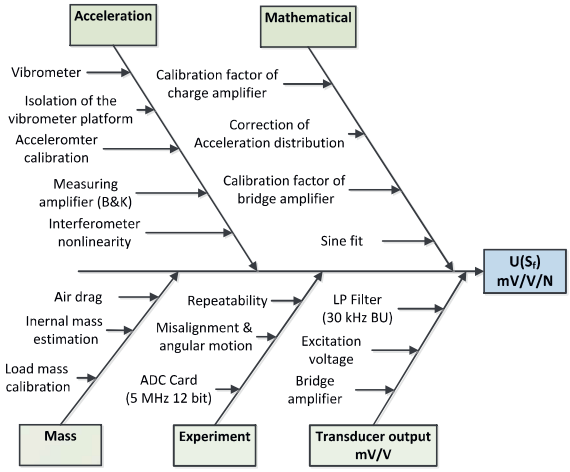


Figure 2: Ishikawa diagram of the uncertainty contributions to the dynamic sensitivity of force measuring devices.

Each suspected contribution to possible significant uncertainty is evaluated by measurement, estimation, by calculation or estimation by performing Monte Carlo MC simulation. The different predicted uncertainty contribution and their estimation method are listed in Table 1.

Table 1: Different uncertainty sources and corresponding estimation methods

Uncertainty Contribution	*Evaluation Method*
Vibrometer	*Calibration certificate*
Bridge amplifier	*Calibration certificate*
Charge amplifier	*Calibration certificate*
Air Drag	*Calculation of empirical formula*
Sine fit	*MC simulation*
ADC Card	*MC simulation*
Repeatability	*Measurement*
Correction of acceleration distribution	*MC simulation*
Isolation of the vibration platform	*Measurement*
Load mass	*Measurement*

Uncertainty of voltage output

The calibration of the bridge amplifier is performed with the precise multimeter HP3458A to calibrate the whole transfer function of the output signal of the force measuring device in $\mathrm{mV/V}$ with a certain uncertainty, $u_{U_{cal}}$. The uncertainty of the output signal of the force measuring device is determined as

$$u_{U_V}^2 = u_{U_{cal}}^2 + u_{filt}^2 + u_{rep}^2 + u_{sine}^2 \quad (10)$$

while u_{filt} is the uncertainty contribution of the used low pass filter, a Butterworth low pass filter has been inserted with a 30 kHz cut-off frequency.

Uncertainty of low-pass filter

A Butterworth low-pass filter with a 30 kHz and 100 kHz cut-off frequency has been applied to the bridge amplifier and charge amplifier respectively. The sampling rate of the data aquisition is set to 25.6 kHz, this offer a maximum attenuation at the Nyquist frequency. The Butterworth filter is a special type of so-called Chebyshev filter with the amplitude frequency response $G_n(\omega)$ as

$$G_n(\omega) = |H_n(j\omega)| = \frac{1}{\sqrt{1 + \varepsilon^2 T_n^2\left(\frac{\omega}{\omega_0}\right)}} \quad (11)$$

Where ε is the ripple factor, ω_0 is the cut-off frequency and T_n is a Chebyshev polynomial of the n order. With the application of the Butterworth filter, a ripple is set to 0 % which leads to a zero amplitude loss at the maximum test frequency of 2 kHz.

Calibration factor of charge and bridge amplifiers

The calibration the bridge and charge amplifiers of that have been used to measure force and acceleration signals respectively is performed with the same excitation frequencies that have been applied to calibrate the force transducer. The correction of the amplitude and phase frequency response of both amplifiers is performed point by point without applying curve fitting.

Uncertainty of sine-fit approximation

A Least-square fit is used to estimate the magnitude and phase of the measurands. The error of the estimate is affected by the resolution and sampling time of the ADC card in addition to the residual error of the sine approximation itself. To estimate the uncertainty of the sine approximation, a Monte Carlo simulation was set to run $\mathrm{M} = 1 \times 10^5$ trials of the model:

$$\mathrm{A} = \mathrm{A}_0 \sin(\omega \mathrm{t} + \varphi) \tag{12}$$

Uncertainty of acting dynamic force

This has an uncertainty of

$$\begin{aligned} u_{F_{dyn}}^2 = {} & u_{a_t}^2 \frac{\partial F_{dyn}}{\partial a_t}^2 + u_{m_t}^2 \frac{\partial F_{dyn}}{\partial m_t}^2 \\ & + u_{m_i}^2 \frac{\partial F_{dyn}}{\partial m_i}^2 \\ & + u_{K_o}^2 \frac{\partial F_{dyn}}{\partial K_{to}}^2 \\ & + u_{drag}^2 \frac{\partial F_{dyn}}{\partial F_{drag}}^2 \end{aligned} \tag{13}$$

The acting load masses include the top mass $\mathrm{m_t}$ and the effective internal mass of the force transducer $\mathrm{m_i}$ that applies a dynamic force on the sensing transducer element. In the case of measurement of acceleration in the force transducer close to the transducer sensing element, $\mathrm{a_b}$, the acting dynamic force will be

$$F_{dyn} = a_t \cdot m_t K_o + a_b \cdot m_i. \tag{14}$$

With uncertainty of

$$\begin{aligned} u_{F_{dyn}}^2 = {} & u_{a_t}^2 \frac{\partial F_{dyn}}{\partial a_t}^2 + u_{a_b}^2 \frac{\partial F_{dyn}}{\partial a_b}^2 \\ & + u_{m_t}^2 \frac{\partial F_{dyn}}{\partial m_t}^2 \\ & + u_{m_i}^2 \frac{\partial F_{dyn}}{\partial m_i}^2 \\ & + u_{K_o}^2 \frac{\partial F_{dyn}}{\partial K_{to}}^2 \\ & + u_{drag}^2 \frac{\partial F_{dyn}}{\partial F_{drag}}^2 \end{aligned} \tag{15}$$

Uncertainty of acceleration distribution factor, $\mathbf{K_o}$

The uncertainty of the correction factor of acceleration distribution $\mathrm{K_o}$ is dominated by uncertainty in the material properties, which were taken from tabulated data. A Monte Carlo simulation has been used to calculate the uncertainty using the simulation of the variation of material properties.

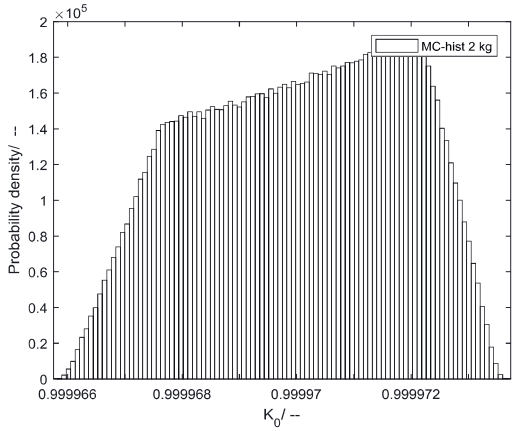


*Figure 3: Output PDFs of the acceleration distribution factor K0 for 2 load masses with variation of 3*10-4 95% confidence level.*

Uncertainty of air drag

As the load mass moves, a parasitic drag force acts on its surface due to the resistance of air. This force can be estimated from the load mass geometry, the maximum velocity of the load mass, V.

$$\mathrm{D} = 0.5\ \mathrm{C_d}\ \rho\ \mathrm{V}^2 \mathrm{A}\,. \tag{16}$$

Where D is the drag force, ρ is the density of air, A is the area of the load mass that subjected to air resistance, and $\mathrm{C_d}$ is the drag coefficient based on empirical data [11]. At the highest velocities (approximately 1 m $\mathrm{s^{-1}}$), the total drag force is estimated to be below 0.5 mN. The drag force is not corrected for, but rather is treated as an uncertainty of magnitude equal to the maximum expected drag of 0.5 mN. A uniform distribution is assumed for this quantity.

Uncertainty of load and internal masses

The mass value of the load mass was determined with a relative uncertainty of $5 \cdot 10^{-3}$ %. To estimate the uncertainty of the internal mass determination, an MC simulation is deployed to simulate the linear equation taking

into consideration a uniform distribution of all fit error in the estimation of the parameters of the above equation.

Uncertainty of vibration isolation

As the shaker is the source of the required vibration, its vibration might be transmitted to the vibrometer platform. An active and passive damping base is installed to the base of the laser head to isolate vibration that comes from the shaker or even from earth. A vibration transmission coefficient has been experimentally measured. The transmission function was measured using accelerometers mounted in both locations; on top of the shaker armature and the frame of the laser vibrometer. Figure 4**Error! Reference source not found.** shows the measurements of the transmission coefficient with the dimpling facilities of the vibrometer platform. The results show the neglected effect of the shaker vibration behind 1 kHz. The uncertainty contribution after 1 kHz was taken to be 0.06%.

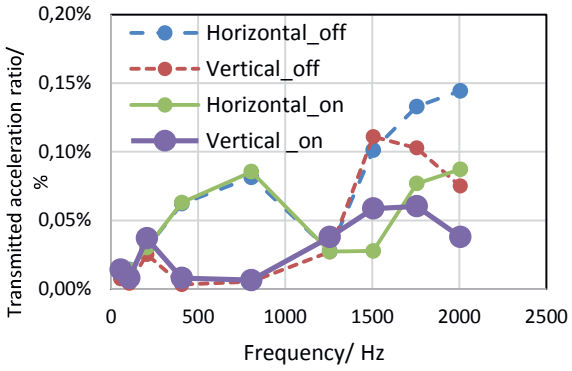


Figure 4: : The vibration transmission ration between the acceleration measured on top of the shaker armature and the acceleration of the frame of the laser vibrometer.

Uncertainty of Phase measurement

The uncertainty U_{phase} in the measured transducer sensitivity phase is calculated according to

$$u_{phase}^2 = u_{ph,acc}^2 + u_{ph,amp}^2 + u_{ph,rep}^2. \qquad (17)$$

The quantity $U_{ph,acc}$ is the uncertainty due to the phase delay introduced by the acceleration measurement which results mainly from the sampling rate of the data acquisition card of the laser vibrometer and was estimated using the calibration of the laser vibrometer with an expanded uncertainty of ±0.3°. The uncertainty $U_{ph,amp}$ accounts for the uncertainty contribution of the bridge amplifier which contributes with an expanded uncertainty of ±0.2°. The uncertainty $U_{ph,rep}$ describes the repeatability of the phase measurement of the force transducer itself. A summary of the sensitivity magnitude and phase measurement uncertainties is presented in Table 2 and Table 3.

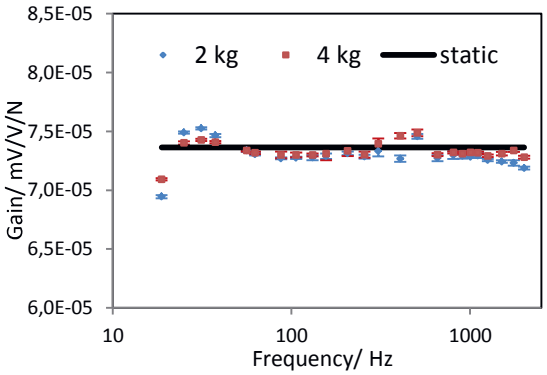


Figure 5: The measured sensitivity and the corresponding relative expanded uncertainty using different load masses in comparison to the static sensitivity

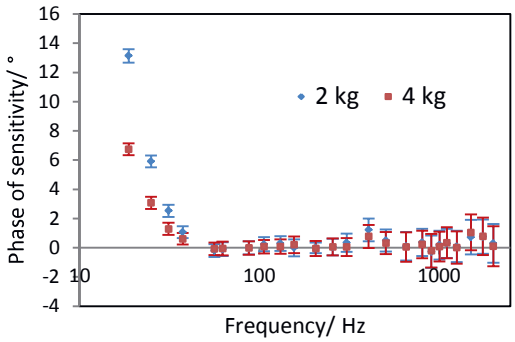


Figure 6: Phase of the sensitivity of the transducer and the corresponding expanded uncertainty (k=2)

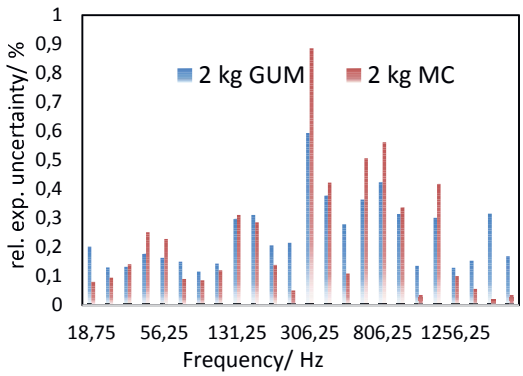


Figure 7: Comparison between the relative expanded uncertainty using classical GUM approach and the 95% coverage interval of the output sensitivity using the MC approach.

Table 2: Summary of uncertainties associated with force transducer sensitivity

Source	Uncertainty contribution [mV/V/N]	100× Index
1.Uncertainty of applied force, $u_{F_{dyn}}$		
Vibrometer calibration, $\boldsymbol{u_{a_t ref}}$	$3{,}06\times10^{-08}$-$3{,}51\times10^{-08}$	9-40
Vibrometer repeatability, $\boldsymbol{u_{a_t rep}}$	$1{,}54\times10^{-09}$-$2{,}66\times10^{-07}$	2-68
Vibration isolation, $\boldsymbol{u_{a_t iso}}$	0-$2{,}42\times10^{-08}$	0-22
Accelerometer calibration, $\boldsymbol{u_{a_b ref}}$	$2{,}46\times10^{-09}$-$1{,}03\times10^{-08}$	1-10
Accelerometer repeatability, $\boldsymbol{u_{a_b rep}}$	$1{,}99\times10^{-10}$-$5{,}17\times10^{-09}$	0.2-2
Top mass, $\boldsymbol{u_{m_t}}$	$3{,}06\times10^{-09}$-$3{,}56\times10^{-09}$	1-4
Sine approximation, $\boldsymbol{u_{sine}}$	$2{,}11\times10^{-13}$-$2{,}46\times10^{-13}$	< 0.02
Internal mass, $\boldsymbol{u_{m_i}}$	$3{,}23\times10^{-11}$-$1{,}76\times10^{-10}$	0-0.2
Air drag, $\boldsymbol{u_{drag}}$	$3{,}64\times10^{-12}$-$9{,}64\times10^{-11}$	0-0.1
Acceleration distribution, $\boldsymbol{u_{K_o}}$	$5{,}45\times10^{-14}$-$2{,}40\times10^{-13}$	< 0.02
2.Uncertainty of Voltage output, u_{U_V}		
Bridge amplifier, $\boldsymbol{u_{U_V ref}}$	$1{,}74\times10^{-08}$-$1{,}88\times10^{-08}$	5-22
Transducer Repeatability, $\boldsymbol{u_{U_V rep}}$	$6{,}24\times10^{-09}$-$2{,}13\times10^{-07}$	7-71
Rel. combined standard uncertainty, $u_{S_{Vf}}$	5,76×10-04-3,79×10-03	
Rel. expanded uncertainty, $U_{S_{Vf}}(k=2)$	1,15×10-03-7,57×10-03	

Table 3: Summary of uncertainties associated with phase measurement

Source	Uncertainty contribution [°]	100×Index
Acceleration, $u_{ph,acc}$	0.3	22-58
Bridge amplifier, $u_{ph,amp}$	0.2	15-39
Repeatability, $u_{ph,rep}$	0.03-2	3-64
Rel. combined standard uncertainty, $\boldsymbol{u_{phase}}$	0,18-0,36	
Rel. expanded uncertainty, $\boldsymbol{U_{phase}(k=2)}$	0,68-1,36	

Conclusion

This paper describes a detailed uncertainty model for the metrological characterization of dynamic force transfer standards in their calibration against dynamic force primary standard. Many uncertainty sources have been reduced like tilting and misalignment effects. Other contributions have been developed like the correction of the acceleration distribution over load masses, and sine-fit approximation. A good agreement has been observed between the two uncertainty estimation approaches; Classical GUM approach and MC approach.

Acknowledgement

The authors gratefully acknowledge support of the Braunschweig International Graduate School of Metrology (B-IGSM).

References

[1] Schlegel C, Kieckenap G, Glöckner B, Buß A and Kumme R 2012 Traceable periodic force calibration *Metrologia* **49** 224–35

[2] Vlajic N and Chijioke A 2016 Traceable dynamic calibration of force transducers by primary means *Metrologia* **53** S136–48

[3] Medina N and de Vicente J 2014 Force sensor characterization under sinusoidal excitations *Sensors (Switzerland)* **14** 18454–73

[4] Sawla A 1991 Measurment of dynamic forces and compensation of errors in fatigue testing *Proc. 12th IMEKO World Congress* (Beijing, China) pp 5–10

[5] Elster C and Link A 2008 Uncertainty evaluation for dynamic measurements modelled by a linear time-invariant system *Metrologia* **45** 464–73

[6] Kumme R 1994 Error sources in dynamic force calibration *Proc. 13th IMEKO World Congress* (Torino, Italy) pp 259–64

[7] Ch. Schlegel, G. Kiekenap and R K 2014 Uncertainty contributions in sinusoidal force measurement *IMEKO 22nd TC3, 12th TC5 and 3rd TC 22 International Conference* (Cape Town, Republic of South Africa) pp 1–4

[8] Zhang L, Wang Y, Yin X, Schlegel C and Kumme R 2018 Comparison results between PTB and CIMM on the force transducer calibration under sinusoidal loading *J. Phys. Conf. Ser.* **1065** 042019
[9] Schlegel C, Kiekenap G and Kumme R 2015 Uncertainty evaluation of sinusoidal force measurement *Proc. 21st IMEKO World Congress* (Prague)
[10] Kumme R 1997 The main influences on the dynamic properties of force measuring devices *Proc. 14th IMEKO World Congress* (Tampere, Finland) pp 102–7
[11] Hoener S F 1965 *Fluid Dynamics Drag: Practical Information on Aerodynamic Drag and Hydrodynamic Resistance*

DOI 10.5162/sensoren2019/2.2.3

Absolut messender Winkelsensor auf Basis optischer Interferometrie

Erik Oertel[1], Vinzenz Ullmann[1], Eberhard Manske[1]
[1]Institut für Prozessmess- und Sensortechnik, Technische Universität Ilmenau, PF 100565, D-98694 Ilmenau

Zusammenfassung

Zur absoluten, hochauflösenden Winkelmessung reflektierender Oberflächen werden nach dem derzeitigen Stand der Technik Autokollimationsfernrohre eingesetzt. Die mit diesem Prinzip erreichbaren Eigenschaften in Bezug auf die Sensorgröße, Auflösung und Messunsicherheit sind für Anwendungen mit hohen Anforderungen nicht mehr ausreichend. Aus diesem Grund wurde ein neuartiges Verfahren zur absoluten Winkelmessung entwickelt. Es basiert auf einem Kösters-Prisma und optischer Interferometrie. Das Interferenzsignal wird über eine CMOS-Matrix erfasst und durch einen Algorithmus ausgewertet. Als Signal für die Winkellage dient die Neigung und Streifenbreite der Interferenzstreifen. Bei ersten messtechnischen Untersuchungen konnte mit diesem Verfahren bereits eine Auflösung von unter 5 nrad und ein Messbereich von mehr als 5 mrad erreicht werden. Im Rahmen dieses Beitrages wird das neue Sensorkonzept vorgestellt. Das beinhaltet den optischen Sensoraufbau, die Strategie zur Signalauswertung und erste messtechnische Untersuchungen.

Keywords: Winkelsensor, Interferometer, CMOS-Matrix, digitale Bildverarbeitung, Kösters-Prisma

1 Einleitung

Zur hochauflösenden, absoluten Winkelmessung von verspiegelten Oberflächen kommen nach dem derzeitigen Stand der Technik primär Autokollimationsfernrohre (AKFs) zum Einsatz . Bei diesen Messgeräten wird die Empfindlichkeit – und damit die erreichbare Auflösung – maßgeblich durch die Brennweite des Kollimators bestimmt. Die mit diesem Prinzip erreichbaren Eigenschaften in Bezug auf die Sensorgröße, Auflösung und Messunsicherheit sind für Anwendungen mit hohen Anforderungen nicht mehr ausreichend. Als Alternative werden nach dem derzeitigen Stand mehrstrahlige Interferometer eingesetzt. Diese Geräte führen die Winkelmessung auf eine Differenzlängenmessung zurück und erreichen die Anforderungen in den genannten Eigenschaften. Im Gegensatz zu AKFs ist mit mehrstrahligen Interferometern aber nur eine relative Winkelmessung möglich. Daher können diese Interferometer die Funktion von AKFs nicht vollständig ersetzen.
Am Institut für Prozessmess- und Sensortechnik der TU Ilmenau wurde deshalb ein neuartiges interferometrisches Verfahren zur absoluten Winkelmessung entwickelt und untersucht [1] , [2] (Abbildung 1).

In diesem wird ein Kösters-Prisma durch einen divergenten, kurz-kohärenten Laserstrahl beleuchtet und an dem halbdurchlässigen Strahlteiler aufgeteilt. Beide Strahlen werden an dem zu vermessenden Spiegel reflektiert und interferieren anschließend auf der gegenüberliegenden Seite. Das dadurch entstehende Interferenzbild wird durch eine CMOS-Matrix erfasst. Als Signal für den Spiegelneigungswinkel φ_x wird der Winkel der Interferenzstreifen Ψ genutzt. Dieser bietet eine absolute Referenz, weil jedem Winkel Ψ exakt ein Spiegelneigungswinkel φ_x zugeordnet werden kann (Abbildung 2).

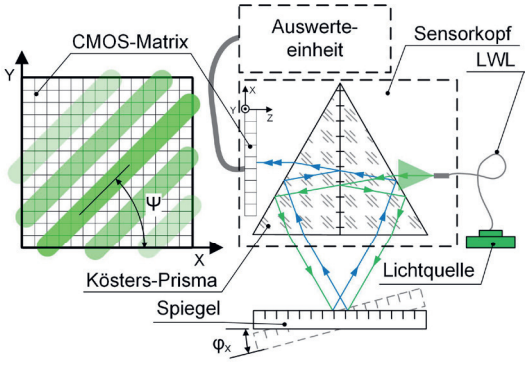


Abb. 1: Das in [1] und [2] vorgestellte Sensorkonzept zur Winkelmessung

Anhand eines experimentellen Aufbaus konnte in [1] und [2] nachgewiesen werden, dass mit diesem Verfahren hochauflösende absolute Winkelmessungen möglich sind. Die Kennlinie wurde durch eine Vergleichsmessung mit einem AKF ermittelt und kann durch Gl. (1) beschrieben werden.

$$\varphi_x = b \tan(\Psi) \quad (1)$$

Der Faktor b wird von Fertigungsabweichungen der einzelnen Kösters-Prismen beeinflusst und muss dementsprechend für jeden Sensor separat kalibriert werden. In experimentellen Untersuchungen wurden darüber hinaus weitere Einflussfaktoren auf den Parameter b ermittelt. Das betrifft beispielsweise den Abstand des Spiegels vom Kösters-Prisma. Weil der quantitative Wert von b durch die zusätzlichen Einflussgrößen – ohne weitere Maßnahmen – mit einer hohen Unsicherheit verbunden ist, ist eine Anwendung des vorgestellten Sensors lediglich als Nullpunktindikator denkbar. Für diesen Anwendungsfall ist nur die Auflösung und Stabilität von dem Winkel Ψ in der Nullage für die erreichbare Messunsicherheit limitierend.

Im Bezug auf die Auflösung und Stabilität konnten gegenüber AKFs mit Brennweiten von 300 mm bereits hervorragende Eigenschaften realisiert werden. Um das Prinzip für absolute Messungen außerhalb der Nullage zu Nutzen, ist eine Reduktion der limitierenden Einflussgrößen notwendig.
Mit dieser Ziellstellung wurde das Verfahren weiterentwickelt und untersucht. Im Rahmen dieses Beitrags werden die wesentlichen Erkenntnisse und Ergebnisse dieser Untersuchungen vorgestellt. Das betrifft Änderungen am optischen Sensoraufbau und der Signalauswertung.

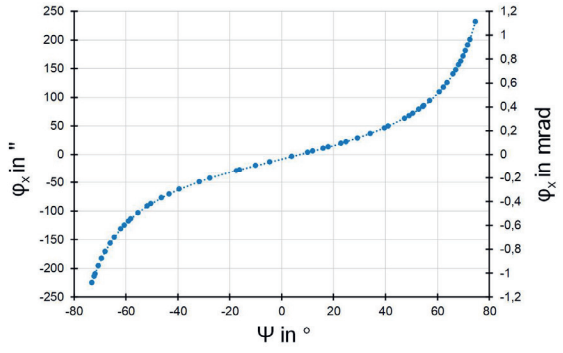


Abb. 2: Abhängigkeit zwischen dem Neigungswinkel des Spiegels und der Interferenzstreifen

2 Grundlagen

Die Abstandsabhängigkeit des in [1] und [2] vorgestellten Sensors wird primär durch die Nutzung von divergentem Licht verursacht und kann durch den Übergang zu kollimiertem Licht reduziert werden. In der bisherigen Konfiguration des Strahlengangs konnte in diesem Fall allerdings keine Neigungsänderung der Interferenzstreifen beobachtet werden.
Unter der Annahme von zwei ideal kollimierten Strahlen (ebene Wellenfronten), kann der Zusammenhang zwischen dem Winkelversatz der Strahlen (Abbildung 3) und dem resultierende Interferenzbild (Abbildung 4) durch Gl. (2), Gl. (3) und Gl. (4) beschrieben werden [3],[4].

$$\Psi_x = \frac{\lambda}{\lambda'} \sin(\Psi) \quad (2)$$

$$\Psi_y = \frac{\lambda}{\lambda'} \cos(\Psi) \quad (3)$$

$$\Psi = \arctan\left(\frac{\Psi_x}{\Psi_y}\right) \quad (4)$$

Die Gleichungen gelten nur näherungsweise für einen geringen Winkelversatz zwischen den Strahlen (Ψ_x und Ψ_y), weil ihnen die Linearisierung trigonometrischer Funktionen zugrunde liegt. Der Parameter λ entspricht der Wellenlänge des Laserlichts.
Bei dem vorgestellten Sensorprinzip aus Abbildung 1 wird der Spiegelneigungswinkel φ_x auf einen Winkelversatz zwischen den Strahlen Ψ_x in der X-Z-Ebene übertragen. Aufgrund der Reflexion beider Strahlen am Spiegel gilt der Zusammenhang aus Gl. (5) [5].

$$\Psi_x = 4\varphi_x \quad (5)$$

Somit ist es ausgehend von Gl. (2) und Gl. (5) möglich die absolute Winkelmessung für kleine Winkel auf eine Erfassung der Streifenbreite λ' und des Neigungswinkels Ψ der Interferenzstreifen zurückzuführen. Dafür ist eine entsprechende Signalauswertung notwendig, welche im nächsten Abschnitt näher diskutiert wird.

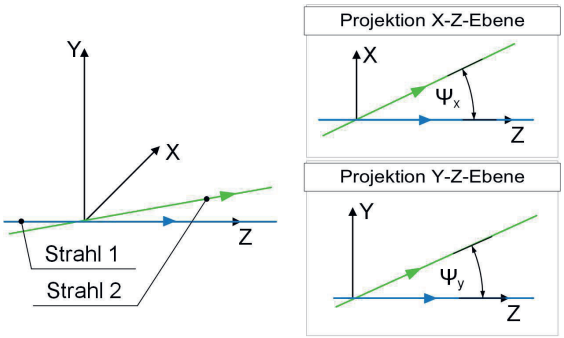


Abb. 3: Winkelversatz zwischen den interferierenden Interferenzstreifen

3 Strategie zur Signalauswertung

Eine verhältnismäßig einfache Auswertung ist möglich, wenn anstatt der Streifenbreite λ‘ die auf die Spalten und Zeilen projizierten Streifenbreiten λ'_y und λ'_x erfasst werden. In diesem Fall kann auf die aufwendigen Methoden der digitalen Bildverarbeitung verzichtet werden, indem jedes Spalten- und Zeilensignal einzeln ausgewertet wird. Die dafür geltenden Zusammenhänge ergeben sich aus den Gl. (2), Gl. (3) und Abbildung 4.

$$\Psi_x = \frac{\lambda}{\lambda'_y}\tan(\Psi) = \frac{\lambda}{\lambda'_x} \tag{6}$$

$$\Psi_y = \frac{\lambda}{\lambda'_x}\frac{1}{\tan(\Psi)} = \frac{\lambda}{\lambda'_y} \tag{7}$$

Unter Berücksichtigung von Gl. (5) und Gl. (6) kann der Spiegelneigungswinkel damit durch Gl. (8) ermittelt werden.

$$\varphi_x = \frac{1}{4}\frac{\lambda}{\lambda'_y}\tan(\Psi) = \frac{\lambda}{\lambda'_x} \tag{8}$$

Für die Auswertung kommen somit zwei grundlegende Strategien infrage. Zum einen ist es möglich, den Betrag des Spiegelneigungswinkels durch die Auswertung des Amplitudenspektrums von einem oder mehrerer Zeilensignale zu ermitteln (Gl. (8) rechter Ausdruck). Diese Strategie wurde in der Literatur anhand anderer Interferometerkonfigurationen bereits verfolgt [6], [7]. Mit diesem Ansatz entfällt allerdings die Richtungsinformation der Spiegelneigung, weil nur der Betrag der Streifenbreite λ'_x ermittelt werden kann. Weiterhin ist auf Basis der bisherigen Untersuchungen davon auszugehen, dass mit diesem Ansatz nicht die gewünschte Auflösung erreicht werden kann.
Die zweite Möglichkeit besteht in der Auswertung der Spaltensignale. Die Streifenbreite λ'_y kann dabei über das Spektrum und der Winkel Ψ über die Phasenverschiebung der einzelnen sinusförmigen Signale ermittelt werden.

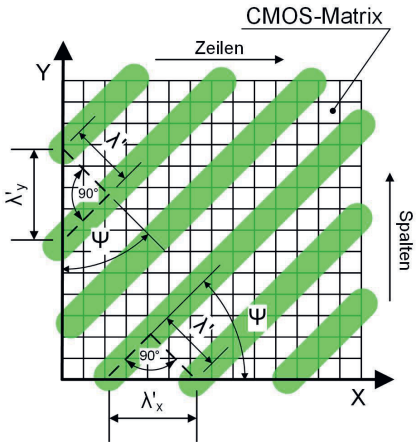


Abb. 4: Merkmale des Interferenzbildes

4 Algorithmus

Aufgrund der genannten Vorteile basiert der Algorithmus für die Auswertung des Interferenzbildes auf der Analyse der Spaltensignale (Abbildung 5).
Dazu wird von jeder Spalte der CMOS-Matrix das Amplituden- und Phasenspektrum mithilfe der schnellen Fourier-Transformation (FFT) ermittelt.
Das Amplitudenspektrum dieser Signale enthält außerhalb des Gleichanteils einen Peak. Über diesen wird die projizierte Streifenbreite λ'_y erfasst und über alle Spalten gemittelt.
Anschließend wird aus dem Phasenspektrum die zugehörige Phasenverschiebung ε für jede Spalte ermittelt. Die in dem Datensatz auftretenden 2π-Phasensprünge werden anschließend korrigiert. Das ist möglich, weil bekannt ist, dass die Interferenzstreifen keine Unstetigkeitsstellen aufweisen. Mit der bereits ermittelten Wellenlänge λ'_y wird daraufhin die jeweilige Phasenverschiebung [ε_1, …, ε_n] der Sinussignale in eine Längenverschiebung [Y_1, …, Y_n] umgerechnet. In den so erhaltenen Datensatz wird eine Gerade durch eine Lineare Regression angepasst. Aus dem Anstieg der Geraden wird anschließend der gesuchte Winkel Ψ berechnet.
Die eingesetzte Kamera besitzt eine Auflösung von 1024x1280 Pixel. Der für die ersten Versuche eingesetzte PC benötigte für die Auswertung des gesamten Interferenzbildes etwa 100 ms. Eine Reduktion der benötigten Zeit für die Auswertung ist unter Einsatz von FPGAs oder Grafikarten möglich.

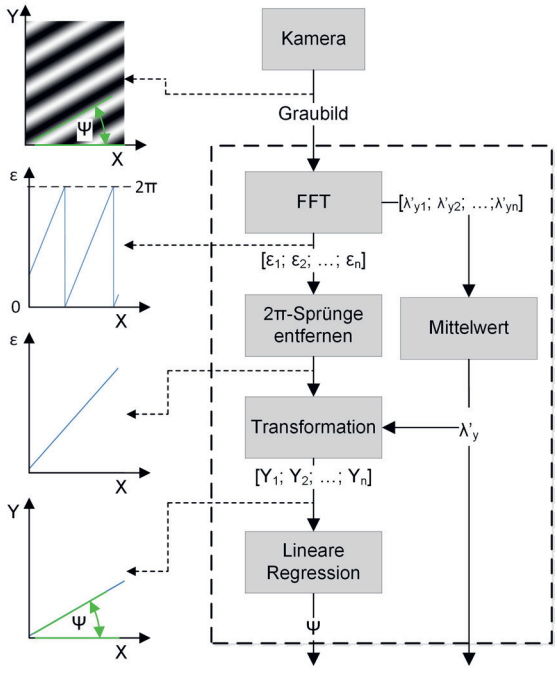


Abb. 5: Algorithmus zur Signalauswertung

5 Optischer Sensoraufbau

Die Gültigkeit von Gl. (8) beschränkt sich auf kollimiertes Licht und kann deshalb nicht auf den bisherigen Sensor angewandt werden (Abbildung 1). Bei diesem Sensor wird der Strahl senkrecht zum Strahlteiler in das Kösters-Prisma eingekoppelt (β=30°, Abbildung 6). Damit wird bewirkt, dass die Strahlen nach der Reflexion am Spiegel die jeweils gegenüberliegende Hälfte des Prismas durchlaufen. Unter Nutzung einer kurz-kohärenten Strahlungsquelle konnte bei fünf getesteten Prismen nur in dieser Konfiguration eine Neigungsänderung der Interferenzstreifen beobachtet werden.

In Abbildung 7 sind die zu beobachtenden Interferenzbilder (Strahlenverlauf A) unter dem Einsatz einer kohärenten Strahlungsquelle (He-Ne Laser) dargestellt. Bei kollimiertem Licht ist eine Neigungsänderung der Interferenzstreifen nicht zu beobachten und es treten in der Nullstellung des Spiegels (d.h. Ψ_x=0) keine sichtbaren Interferenzstreifen auf. Nach Gl. (7) kann das auf einen zu geringen Winkelversatz von Ψ_y zurückgeführt werden, weil dieser in direktem Zusammenhang mit der Streifenbreite λ'_y steht. Dieser Winkelversatz kann nur durch Abweichungen von der Idealform des Prismas erzeugt werden und ist eine notwendige Voraussetzung für die Realisierung des Sensorprinzips. Der Strahlenverlauf A weist eine hohe Symmetrie zwischen den Strahlen auf, wodurch dieser Winkelversatz minimiert wird.

Das Kösters-Prisma kann jedoch durch Variation des Winkels β noch unter einem weiteren Strahlenverlauf genutzt werden (Strahlenverlauf B). Bei diesem werden die Strahlen nach der Reflexion am Spiegel in die Prismenhälfte reflektiert, aus der sie vorher ausgetreten sind. Dadurch werden die beiden Strahlen gegenüber dem Strahlenverlauf A jeweils nur an einer Prismenhälfte totalreflektiert und gebrochen. Weil davon auszugehen ist, dass beide Prismenhälften unterschiedliche Fertigungsabweichungen aufweisen, ist bei diesem Strahlenverlauf mit einem größerem Winkelversatz Ψy zu rechnen. Diese Überlegung bestätigt die Beobachtung in Abbildung 7. Der Strahlenverlauf ist sowohl mit divergentem, als auch kollimiertem Licht nutzbar.

Gegenüber dem Strahlenverlauf A ist dabei allerdings eine ausreichend kohärente Lichtquelle notwendig, denn durch die unterschiedliche Auslenkung der Strahlen ergeben sich geringfügig unterschiedliche Weglängen. Dadurch ist bei der Neigungsänderung des Spiegels neben der Rotation eine Translation der Interferenzstreifen zu beobachten. Dieser Effekt ist für die Funktionsweise des Sensorprinzips nicht relevant, wenn eine Lichtquelle mit einer ausreichenden kohärenzlänge eingesetzt wird.

Bei dem modifizierten Sensor wird der Strahlenverlauf B verwendet (Abbildung 8). Der Laserstrahl wird senkrecht zur Prismen-oberfläche eingekoppelt (β = 90°), weil dadurch die Position des Interferenzbildes unabhängig vom Abstand des Spiegels ortsfest bleibt. In dieser Anordnung beeinflusst der Spiegel-abstand lediglich den erreichbaren Mess-bereich. Limitierend ist dabei der laterale Versatz der Lichtstrahlen durch die Reflexion an dem geneigten Spiegel (Abbildung 8). Eine Messung ist nicht mehr möglich, wenn dadurch die Lichtintensität auf der CMOS-Matrix zu gering wird. Dieser Effekt limitiert auch den Messbereich von AKFs.

Die Ausdehnung des Sensors wird primär durch das verwendete Prisma und den eingesetzten Kollimator vorgegeben.

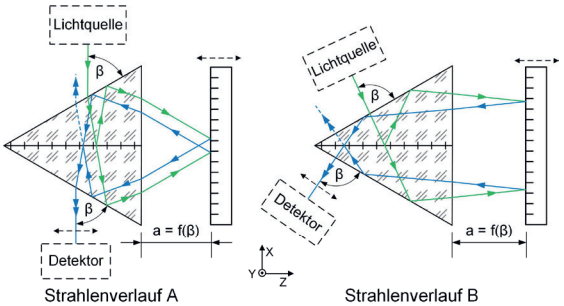


Abb. 6: Mögliche Strahlenverläufe

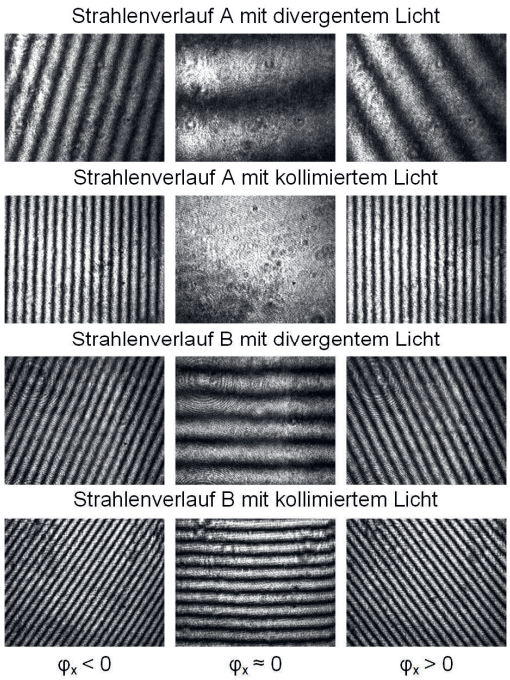


Abb. 7: Zu beobachtende Interferenzbilder

Dabei wird die Größe des Kollimators durch den notwendigen Strahldurchmesser bestimmt. Dieser wiederum ergibt sich aus der Ausdehnung der aktiven Sensorfläche. An einem realisierten Prototyp kommt eine CMOS-Matrix mit einer aktiven Sensorfläche von 6,66 mm x 5,32 mm zum Einsatz. Der Kollimator hat eine Brennweite von 50 mm. Die lange Kathete des verwendeten Kösters-Prismas beträgt 17 mm. Damit ist die Baugröße des Sensors bereits deutlich kleiner als vergleichbare AKFs mit Brennweiten von 300 mm bis 1000 mm. Eine weitere Reduzierung des benötigten Bauraums ist durch die Verwendung einer kleineren CMOS-Matrix möglich. Derartige Sensoren sind gegenwärtig bereits auf dem Markt erhältlich. Somit stellen die hier genannten Werte noch nicht die mit diesem Prinzip realisierbare Grenze dar [3].

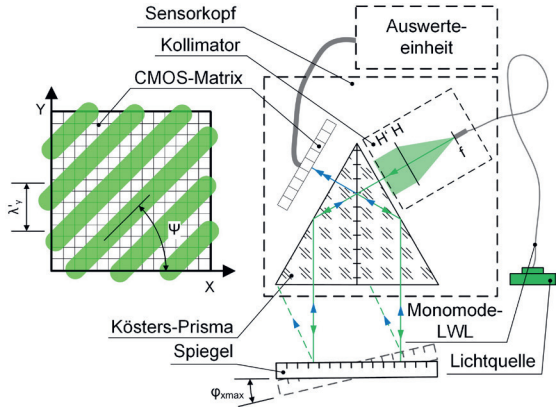


Abb. 8: Das überarbeitete Sensorkonzept zur absoluten Winkelmessung

6 Messtechnische Untersuchungen

Zur Überprüfung des Konzepts wurden erste Untersuchungen an einem einfachen Versuchsaufbau durchgeführt (Abbildung 9). Dabei wurde das neue Sensorkonzept mit einem kommerziell verfügbaren AKF verglichen. Das AKF besitzt eine Brennweite von 300 mm.

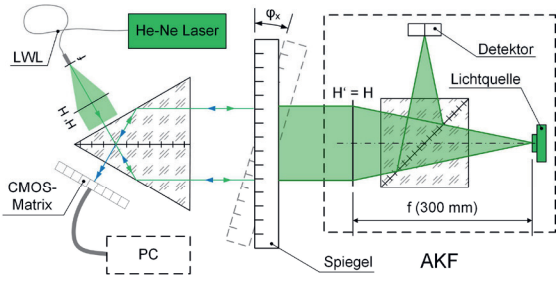


Abb. 9: Versuchsaufbau (nicht maßstabsgerecht)

Die eingesetzte Lichtquelle (He-Ne Laser) emittiert Licht mit einer Wellenlänge von 623,8 nm. Zur Aufnahme einer Kennlinie wurde der Spiegel schrittweise manuell bewegt. Die Messwerte wurden jeweils zwischen zwei Schritten im stationären Zustand aufgenommen.
Gegenüber dem bisherigen Sensorkonzept wird zusätzlich die Streifenbreite erfasst. Ausgehend von Gl. (2), Gl. (3) und Gl. (4) kann die Abhängigkeit der Streifenbreite λ‘ von dem Spiegelneigungswinkel und dem Neigungswinkel der Interferenzstreifen durch Gl. (9) und Gl. (10) beschrieben werden [3].

$$\lambda' = \frac{\lambda}{\sqrt{\Psi_y^2+\Psi_x^2}} = \frac{\lambda}{\sqrt{\Psi_y^2+(4\varphi_x)^2}} \tag{9}$$

$$\lambda' = \frac{\lambda}{\Psi_y\sqrt{1+tan^2\Psi}} \tag{10}$$

In Abbildung 10 und 11 sind die gemessenen Abhängigkeiten dargestellt. Der Winkel φ_x entspricht dem Referenzwert des AKFs. Zusätzlich enthalten beide Graphen die eigentlich gemessene Streifenbreite λ_y‘. Diese korreliert nach Gl. (7) mit dem Winkelversatz Ψ_y und weist keine Abhängigkeit zu dem Neigungswinkel der Interferenzstreifen Ψ und dem Spiegelneigungswinkel φ_x auf. In den Gleichungen (7), (9) und (10) ist der Winkelversatz Ψ_y der einzige unbekannte Parameter. Anhand dieser Gleichungen und den Messwerten wurde der tatsächlich vorhandene Winkelversatz durch eine Regressionrechnung quantifiziert. Als Optimierungskriterium kam dabei die kleinste Summe der Fehlerquadrate zum Einsatz. Die Ergebnisse sind in Tabelle 1 enthalten.

Tab. 1: Die jeweiligen Ergebnisse der nichtlinearen Regression

	Gl. (7)	Gl. (9)	Gl. (10)
Ψ_y in ‘‘	263	263	282
Ψ_y in mrad	1,28	1,28	1,37

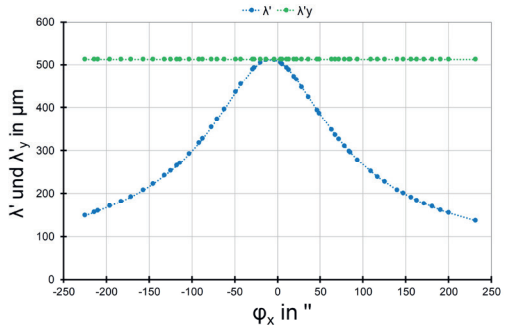


Abb. 10: Abhängigkeit zwischen der Streifenbreite der Interferenzstreifen und dem Spiegelneigungswinkel

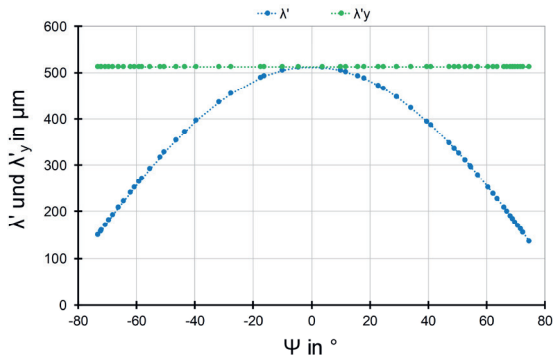


Abb. 11: Abhängigkeit zwischen der Streifenbreite und der Neigung der Interferenzstreifen

Für die eigentliche Messung des Winkels φ_x wird Gl. (8) angewandt. In Abbildung 12 ist die Differenz zwischen dem so ermittelten Spiegelneigungswinkel und dem Referenzwinkel des AKFs dargestellt. Es zeigt sich eine lineare Abweichung zwischen beiden Sensoren. Dafür können mehrere Gründe in Betracht gezogen werden. Ein wesentlicher Beitrag wird von dem Versuchsaufbau verursacht. Die Justierung entlang der optischen Achse beider Sensoren war nur bedingt möglich. Dadurch messen beide Sensoren nicht exakt den gleichen Winkel. Ein weiterer Beitrag kann der unzureichenden Information über die CMOS-Matrix zugeordnet werden. Diese bildet die Referenz für die Streifenbreite und den Neigungswinkel der Interferenzstreifen. Deshalb ist die Kenntnis über die Anordnung und Ausdehnung der einzelnen Pixel erforderlich. Im Rahmen dieser Untersuchungen wurden dafür die Angaben des Herstellers verwendet. Dieser gibt die Größe der Pixel mit 5,2 µm an. Diesem Wert ist jedoch keine Unsicherheit zugeordnet und somit unvollständig.
Für eine erste Einschätzung der Vergleichsmessung wurden deshalb die linearen Abweichungen korrigiert (Abbildung 13).

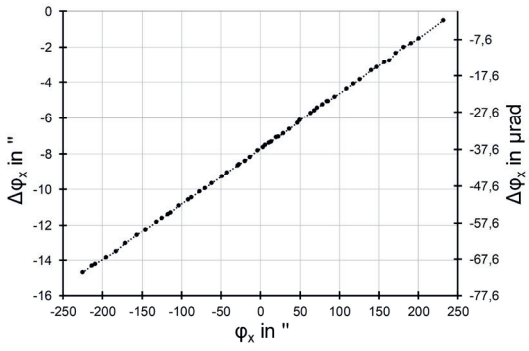


Abb. 12: Differenz zwischen dem Spiegelneigungswinkel des Sensorkonzepts und einem AKF.

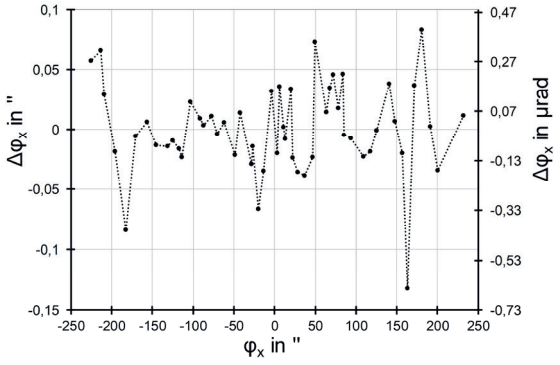


Abb. 13: Die Messwerte aus Abbildung 12 ohne lineare Abweichungen.

Ohne die linearen Abweichungen zeigen sich keine weiteren systematischen Abweichungen gegenüber Gl. (8). Zwischen den Messwerten wurde der Versuchsaufbau durch die manuelle Änderung der Spiegelposition negativ beeinflusst. Für weiterführende Untersuchungen sind dementsprechend bessere Versuchsbedingungen notwendig.
In [2] wurde bereits eine Auflösung von unter 5 nrad (0,001‘‘) mit diesem Prinzip erreicht. Im Unterschied zu den Ergebnissen in [2] wurde in den ersten Versuchen mit dem modifizierten Aufbau zusätzlich ein Messbereich von 5 mrad (1000‘‘) abgedeckt.

7 Zusammenfassung

Die Funktionsfähigkeit des Sensorkonzepts konnte anhand der ersten Messungen gezeigt werden. Für weiterführende Untersuchungen wurde ein Prototyp entworfen [3].
Gegenüber AKFs ermöglicht das Prinzip eine deutlich kleinere Bauweise bei gleichbleibender bzw. höherer Sensorauflösung. Über die Messunsicherheit kann zum gegenwärtigen Zeitpunkt noch keine Aussage getroffen werden.

Literaturnachweis

[1] V. Ullmann, E. Manske, *DGaO-Proceedings*, 118 (2017)
[2] V. Ullmann, E. Oertel, E. Manske, *Meas. Sci. Technol.* 29 (2018); doi:10.1088/1361-6501/aab252
[3] E. Oertel, *TU Ilmenau*, Masterarbeit (2018)
[4] M. Welter, *TU Ilmenau*, Dissertation (2005)
[5] J. Sindram, *TU Ilmenau*, Masterarbeit (2016)
[6] S. Strube, G. Molnar, H. Danzebrink, *Meas. Sci. Technol.* 22 (2011); doi:10.1088/0957-0233/22/9/094026
[7] G. Molnar, S. Strube, P. Köchert, H. Danzebrink, J. Flügge, *Meas. Sci. Technol.* 27 (2016); doi: 10.1088/0957-0233/27/8/084011

DOI 10.5162/sensoren2019/2.2.4

Hybridintegration von Mikrodehnungssensoren

Th. Frank[1], A. Grün[1], M. Kermann[1], A. Cyriax[1], A. Steinke[1], Th. Ortlepp[1], G. Reschke[2]
[1]Forschungsinstitut für Mikrosensorik, Konrad-Zuse-Straße 14, D-99099 Erfurt, Deutschland, tfrank@cismst.de
[2]Technische Universität Dresden, Fakultät Maschinenwesen, Institut für Fertigungstechnik, George-Bähr-Straße 3c, 01069 Dresden, Deutschland

Zusammenfassung

Silizium-Dehnungsmesssensoren oder Siliziumdehnmessstreifen (Si-DMS) substituieren in einige Anwendungen, insbesondere im Bereich Druckmesstechnik, die piezoresistiven Dünnschicht- und die Dickschichtwiderstände. Der höhere Koppelfaktor und die deutlich höhere Langzeitstabilität machen diese Sensoren zunehmend interessant für weitere Anwendungen. Besonders hohe Anforderungen an die Montage bereitet die zunehmende geometrische Größe der Prüfkörper. Die Prozesse, wie sie für kleine Prüfkörper üblich sind, z.B. Glasfritt- und Drahtbonden, stoßen schnell an ihre Grenzen. Thermische Prozesse sind aufgrund der Masse des Prüfkörpers unwirtschaftlich oder nicht gewünscht. Die Drahtbondprozesse an den großen Prüfkörpern sind sehr aufwendig und sollen vermieden werden. Die Prüfkörper sind lange Rohre, Druckbehälter, Maschinenschrauben oder Federkörper für Kraftsensoren. Für die Anwendung sollen die Sensoren direkt montiert und so hybrid integriert werden. Der Beitrag beschreibt verschiedene Strategien zum Aufbau makroskopischer Prüfkörper mit hybrid integrierten Siliziumdehnmessstreifen.

Keywords: Silizium, Dehnmessstreifen, mechanische Spannung, Zustandsüberwachung, piezoresistiv

Einleitung

Silizium-Dehnungsmesssensoren ersetzen insbesondere im Bereich der Druckmesstechnik, die piezoresistiven Dünnschicht- und die Dickschichtwiderstände. Durch die geringe geometrische Größe des Prüfkörpers sind die notwendigen Prozesse wie, Drahtbonden und thermische Fügeprozesse sehr einfach auszuführen sind. Somit sind sie unter anderem als Drucksensoren prädestiniert. Der höhere Koppelfaktor und die deutlich höhere Langzeitstabilität machen diese Sensoren zunehmend interessant für weitere Anwendungen. Besonders hohe Anforderungen an die Montage bereitet die zunehmende geometrische Größe der Prüfkörper. Die Prozesse, wie sie für kleine Prüfkörper üblich sind, z.B. Glasfritt- und Drahtbonden, stoßen schnell an ihre Grenzen. Thermische Prozesse sind aufgrund der Masse des Prüfkörpers unwirtschaftlich oder nicht gewünscht. Die Drahtbondprozesse an den großen Prüfkörpern sind sehr aufwendig und sollen vermieden werden, da sowohl die Temperaturbelastung nicht gewünscht ist, als auch die vollständige Erwärmung dieser Bauteile hohe Kosten verursacht. Die Prüfkörper sind oft lange Rohre, Druckbehälter, Maschinenschrauben oder Federkörper für Kraftsensoren. Im Folgenden werden verschiedene Strategien zum Aufbau makroskopischer Prüfkörper mit hybrid integrierten Siliziumdehnmessstreifen untersucht.

Bei der am häufigsten verwendeten Dopingvariationen, werden die resistiven Elemente im n-Typ-Basismaterial durch Dotierung typischerweise mit Bor realisiert. In diesem Fall sind die piezoresistiven Koeffizienten, quer (π_T) oder längs (π_L) in Bezug auf die Stromrichtung vom Betrag nahezu gleich, aber entgegengesetzt gerichtet. Für einen einachsigen Spannungszustand kann daher eine Messbrücke aufgebaut werden, indem zwei Widerstände parallel und zwei Widerstände senkrecht zur angelegten mechanischen Spannung oder Belastung angebracht sind. Dies ermöglicht ein hohes Maß an Miniaturisierung. Eine Fläche von 100 µm mal 100 µm reicht für eine Wheatstone-Brücke aus.

Mit Hilfe von MEMS-Technologien werden die Si-DMS hergestellt. In Abb. 1 ist das Layout und die Schaltung dargestellt. Die Implantationsparameter sind für einen Flächenwiderstand des piezoresistiven Widerstandes von ca. 200 Ω mit einem Temperaturkoeffizienten von ca. 0,8 ‰ / K ausgelegt. Die piezoresistiven Widerstände sind durch hoch implantierte Leitungen mit den Bondpads verbunden.

Die Chipfläche des Si-DMS beträgt 500 x 500 µm². Für die verschiedenen Anwendungen können die Sensoren bis zu 5 µm abgedünnt werden. Diese Variante ist sehr gut geeignet,

die mechanischen Spannungen in einem homogenen Spanungsfeld zu bestimmen.

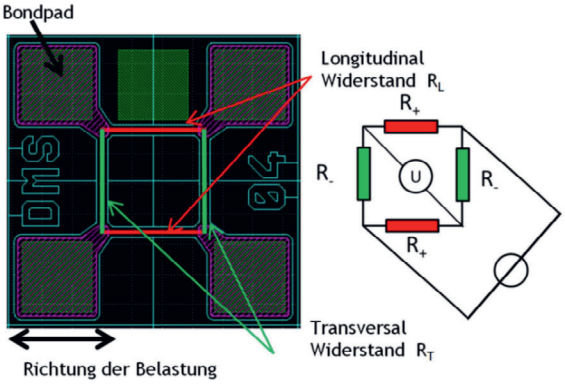


Abb. 1: Si-DMS mit kompletter Messbrücke mit zwei Widerständen für den Längs- und den Quereffekt

Durch die Anordnung der Messwiderstände ist der Si-DMS in der Ebene in alle Richtungen gleich empfindlich. Die durch unterschiedlichen Temperaturausdehnungskoeffizienten induzierten mechanischen Spannungen können so leicht kompensiert werden. Für die Positionierung des Sensors ist aber eine gute Kenntnis der mechanischen Belastung notwendig. Eine optimale Position weißt eine möglichst hohe Anisotropie der mechanischen Spannungen auf.

Montage Si-DMS aus Stahl

Die Si-DMS werden auf verschiedenen Federkörpern mittels Glasfritte montiert, die Verarbeitungstemperatur beträgt ca. 450 °C.
Aufgrund der hohen Unterschiede in der thermischen Ausdehnung zwischen metallischen Federkörpern und dem Silizium steht der DMS unter starken lateralen Druckspannungen. Diese beeinflussen natürlich auch die Messwiderstände, die zu einer Messbrücke verbunden sind. Die Messbrücke kompensiert die Belastung nahezu vollständig, wenn die Bedingungen an allen Messwiderständen identisch sind. Dies wird durch ein völlig symmetrisches Design erreicht.
An die Verbindung werden sehr hohe Anforderungen gestellt. Diese muss insbesondere langfristig stabil sein. Dies bedeutet, dass die mechanische Belastung vom Federkörper auf den Si-DMS ohne Dehnungsverluste übertragen wird. Temperaturschwankungen dürfen keine Hysterese verursachen. Das Kriechen aufgrund der hohen mechanischen Belastung ist auszuschließen. Als Werkstoff für den Federkörper eignen sich Aluminium, Titan und Stahl. Eine sichere Verbindung, welche die Anforderungen erfüllt, ist eine Verbindung mittels Glasfritte [1], [2].

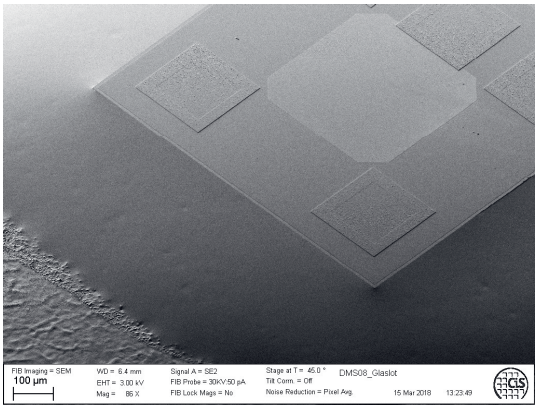


Abb. 2: Auf Stahl (1.4542) mit Glasfritte gefügter Si-DMS

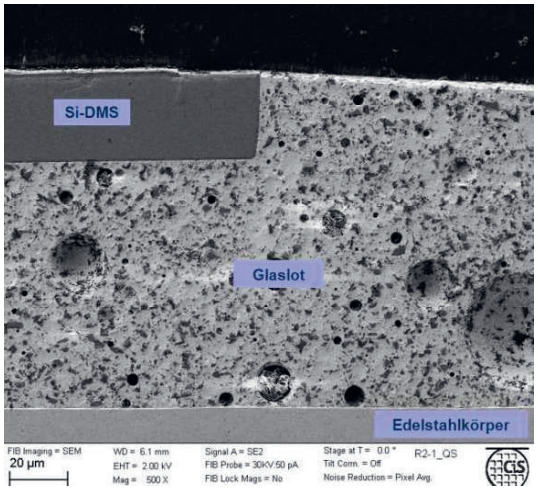


Abb. 3: Schnitt durch die Verbindung

Abb. 2 und Abb. 3 zeigen eine REM Aufnahme einer Glaslotverbindung mit einem bleihaltigen Glaslot. In [3] wurde bereits das Ergebnis einer Langzeituntersuchung gezeigt, nach 24 h betrug die Abweichung weniger als, 0,01% der Messspanne.

Simulation der Belastung

Als Beispiel dient eine Schraube als Prüfkörper bzw. Federelement.

Die Schraube als Verbindungselement

Sollen die Belastung eines komplexen Bauelementes, welche durch äußere und innere Kräfte wirken, gemessen und auch aufgezeichnet werden, eignen sich hierfür besonders die Verbindungselemente, da sich die Belastungen an diesen konzentriert. Als Beispiel sollen die Maschinenschrauben dienen.
Eine Schraube ist ein Maschinenelement zur kraftschlüssigen Verbindung von zwei oder mehr Bauteilen. Durch das Anzugsmoment wird zwischen Schraubenkopf und Muttergewinde die Montagevorspannkraft aufgebaut.

Die maximale zulässige Vorspannkraft wird durch den Bolzendurchmesser und die Festigkeitsklasse der Schraube definiert. Für eine dauerhaft sichere und funktionsfähige Verbindung ist zu gewährleisten, dass die Vorspannkraft über die gesamte Lebensdauer in einem definierten Toleranzbereich liegt.
Für Standardanwendungen mit bekannten Belastungen, ist die Einhaltung des Anzugsmomentes und der Richtlinie VDI 2230 ausreichend. Besonders sicherheitsrelevante Verbindungen bedürfen jedoch regelmäßiger Kontrollen. Die nachträgliche Kontrolle der Montagevorspannkraft ist ohne verschiedene Zusatzmaßahmen nicht möglich. Standardgemäß wird das Anzugsmoment manuell überprüft. Es ist zudem auszuschließen, dass Korrosion und Verschmutzung durch Erhöhung der Reibung ein korrektes Anzugsmoment vortäuschen. Vorteilhaft ist das Monitoring der Montagevorspannkraft. Die Vorspannkraft dehnt den Schraubenschaft, staucht die zu fügenden Elemente und deformiert sowohl den Schraubenkopf als auch das Muttergewindes. In der Deformation des Schraubenkopfes, bzw. der Mutter sind damit alle Informationen über die Vorspannkraft enthalten. Für die Messung haben sich Dehnmessstreifen als vorteilhaft erwiesen. Sie bilden die Deformation präzise ab.

Belastung der Schraube

Zwischen dem Schraubenkopf und dem Muttergewinde wird durch das Anzugsmoment die Montagevorspannkraft aufgebaut. Besonders an der Oberfläche des Schraubenkopfes ist die resultierende mechanische Dehnung, bzw. mechanische Spannung für die Si-DMS leicht zugänglich. In Abb. 4 ist die Belastung des Schraubenkopfes dargestellt. Insbesondere die mechanischen Spannungen diagonal über den Kopf sind von Interesse.
Die Abb. 5 zeigt die mechanischen Spannungen in Richtung x (blau) und y (grün). Die optimale Position liegt bei der betragsmäßig größten Differenz beider Spannungen, die lokalen Minima der roten Linie, hier ist die Anisotropie am größten. Dieser Bereich liegt ungefähr an der Stelle wo der Schraubenbolzen den Kopf durchdringt. Dies würde bedeuten, dass sich für jede Schraubenabmessung die Position ändert. Wünschenswert ist eine Montage des Sensors in der Mitte der Schraube. Dazu ist es erforderlich die auftretenden mechanischen Spannungen in einer Richtung vom Sensor fernzuhalten. Im Beispiel sind dies die Spannungen in Richtung y.

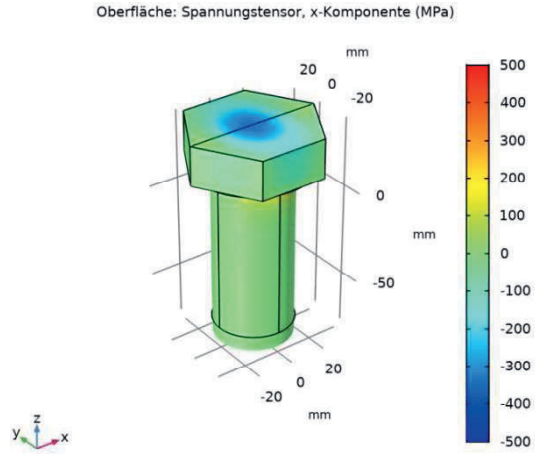


Abb. 4: FEM-Simulation Schraubenkopf

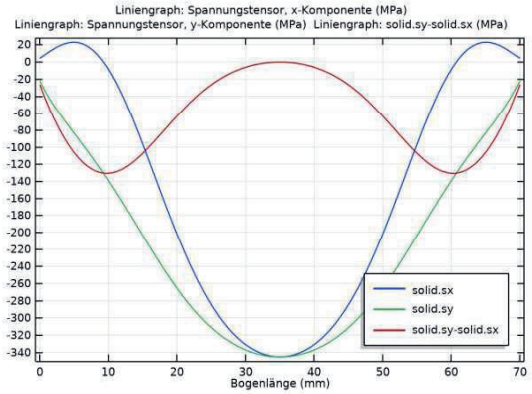


Abb. 5: Simulierte mechanische Spannung über dem Schraubenkopf, M42 nach DIN 931

Die Kompensation der Querspannung

Die Kompensation der Querspannung kann leicht erreicht werden, wenn der gefügte Si-DMS auf dem Federkörper links und rechts in Richtung y durch eine Nut vom Kraftfluss getrennt wird. Die Nut kann direkt in den Federköper eingearbeitet oder über einen Montageträger gefügt werden. Die Abb. 6 und Abb. 7 zeigen die Verteilung der mechanischen Spannung. Gut zu erkennen ist die Filterwirkung des Stegs, bzw. der Nuten.
Abb. 8 und Abb. 9 zeigen in einem Diagramm den Verlauf beider Spannungen. Der Unterschied besteht darin, dass in Abb. 8 ein Montageträger verwendet wurde, hingegen sind in Abb. 9 die Nuten direkt in den Schraubenkopf geschnitten. Die Funktion des Montageträges ist in [3] und [4] bereits erläutert. Durch die Verwendung des Montageträgers wird die übertragene mechanische Spannung geringfügig. Die direkte Applizierung der Nuten in den Federkörper, ist jedoch insbesondere für größere Bauteile sehr aufwendig und sollte vermieden werden.

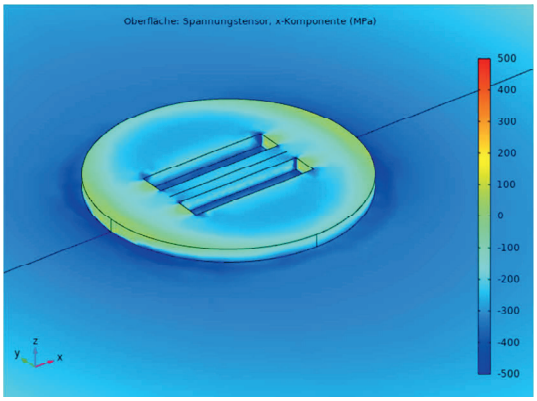


Abb. 6: FEM-Simulation der geschlitzten Montageträger, hohe mechanische Spannung in Richtung x, Belastung wie Abb. 4

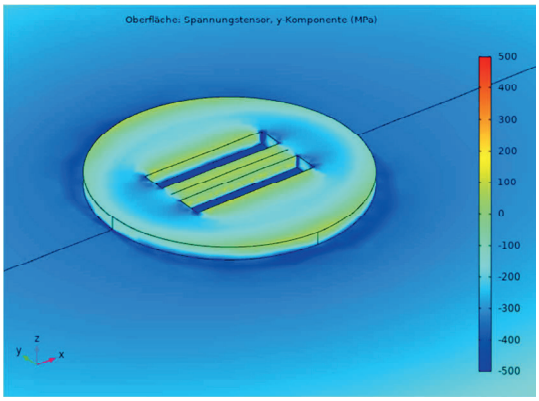


Abb. 7: FEM-Simulation der geschlitzten Montageträger, geringe mechanische Spannung in Richtung y

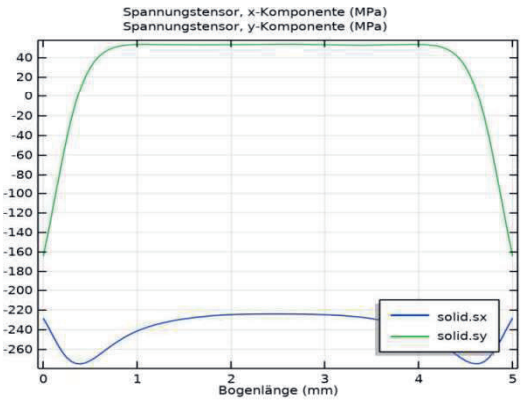


Abb. 8: FEM-Simulation der geschlitzten Montageträger, mechanische Spannung in Richtung x und y

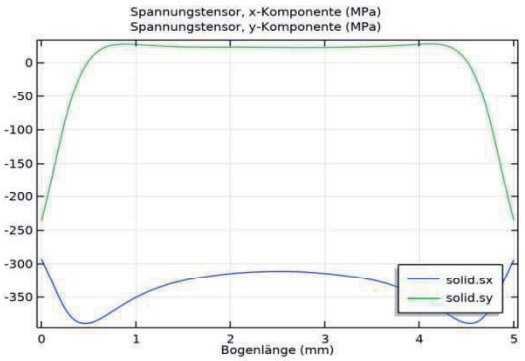


Abb. 9: FEM-Simulation, Nut direkt im Schraubenkopf, ohne Montageträger, mechanische Spannung in Richtung x und y

Beispielanwendung am Schraubenkopf

Am Beispiel einer Schraube wird die Funktion gezeigt. Dazu werden Sechskantschrauben M16*1,5 verwendet. Der Steg, bzw. die Nuten wurden in den Schraubenkopf eingearbeitet. Dies erwies sich für die geringe Stückzahl zweckmäßiger. Die Abb. 10 zeigt die Struktur. Soll zusätzlich die Richtung der Belastung bestimmt werden können 3 Si-DMS um 120° auf dem Kopf angeordnet werden, Abb. 10c. Die Si-DMS werden mit einer Glasfritte auf den Steg gefügt und mittels Drahtbondtechnik elektrisch kontaktiert.



a) b) c)

Abb. 10: Schraubenkopf mit eingearbeiteten Steg

Die folgende Darstellung widmet sich der erreichten Ergebnisse. Es wurden zwei verschiedene Prüfvorrichtungen verwendet. Die in der Abb. 11 dargestellte Vorrichtung dient zum hydraulischen Verspannen der Schrauben. Durch die Druckzylinder wird die Schraube belastet, der Referenzsensor, ein Ringkraftsensor befindet sich unter dem Schraubenkopf.

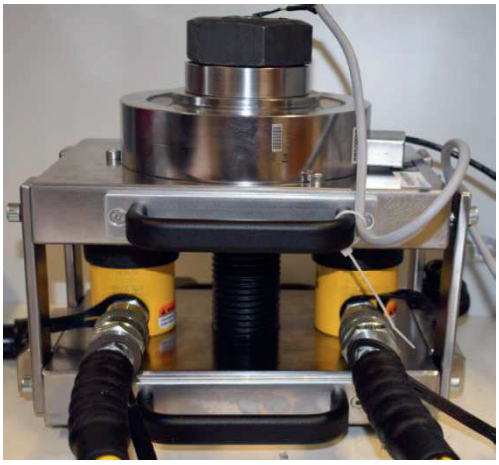


Abb. 11: Versuchsaufbau, hydraulisch verspannt, M48 8.8 als Beispiel

Zur Überprüfung des Einflusses des Anzugsmomentes wurde eine ähnliche Vorrichtung verwendet, die Vorspannkraft wurde mittels Drehmoment über die Schraubenmutter aufgebracht. Zur Darstellung der Wirksamkeit der Kompensation der Quereinflüsse ist in Abb. 12 das Ergebnis der Bestimmung der Vorspannkraft ohne Steg, planer Schraubenkopf, gezeigt. Die Messspanne ist durchaus ausreichend, resultiert aber aus Ungenauigkeiten in der Positionierung des Si-DMS. Die Schraube wurde 10 mal bis auf ca. 130 kN belastet und vollständig entlastet. Die Kompensation der Quereinflüsse sollte einen Zuwachs der Messspannung um das 20-fache ergeben.

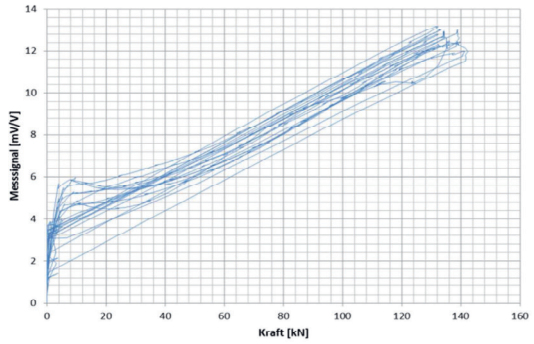


Abb. 12: Messung auf dem Schraubenkopf ohne Steg

In Abb. 13 sind die Untersuchungen der Schrauben mit der Kompensation dargestellt. Die Schraube wurde 10 mal be- und entlastet, sowohl hydraulisch als auch mittels Drehmoment.
Die Abb. 14 zeigt den Versuchsablauf. Sehr gut zu sehen ist aus beiden Abbildungen 13 und 14 die Unabhängigkeit des Messverfahrens von dem Anzugsverfahren. Eine wesentliche Fehlerquelle stellen die Referenzsensoren dar, es handelte sich um zwei verschiedene Ringkraftsensoren und die Wiederholgenauigkeit der Auflagepunkte der Schrauben auf dem Sensor konnte nicht überprüft werden.

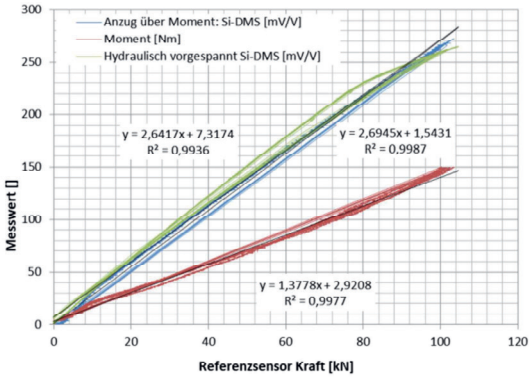


Abb. 13: Messwerte, Anzugsverfahren Drehmoment, hydraulisch verspannt über Referenzsensor (kN)

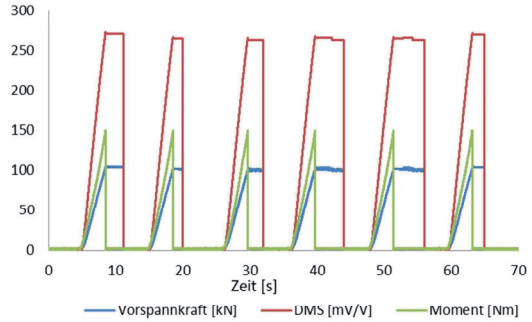


Abb. 14: Messwerte, Anzugsverfahren Drehmoment

Die nächste Frage stellt sich in Bezug auf die Kalibrierung. Wünschenswert ist eine Einpunktkalibration, das bedeutet, dass die Steigung in einem Toleranzbereich identisch ist, sich aber der Nullpunkt fertigungsbedingt variiert. Die Abb. 15 zeigt 10 Messgraden von je 3 verschiedenen Schrauben. Da bei diesem Anwendungsfall nur die oberen 20% des Messbereiches von Interesse sind ist die gezeigte Messgenauigkeit ausreichend. Weiter erhöht werden kann die Genauigkeit durch eine maschinelle Fertigung. Die wesentlichen Abweichungen betreffen die Position und die Dicke aller Materialien unter dem Si-DMS.

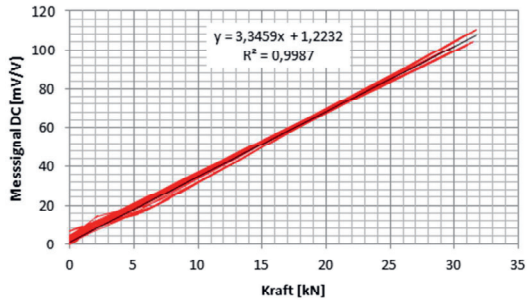


Abb. 15: Messgraden an 3 Schrauben

Beispielanwendung am Rohr

Eine ähnliche Konstellation ergibt sich bei der Aufgabe den Druck in einem Rohr mittels eines Si-DMS zu bestimmen. Ein langgestrecktes Rohr weißt durch die wesentlich größere Länge im Verhältnis zum Durchmesser eine ausgeprägte Anisotropie der mechanischen Spannungen auf.
Die Abb. 16 verdeutlicht diesen Sachverhalt. Zur besseren Montage des Si-DMS und zur Erhöhung der Sensitivität wurde das Rohr abgeflacht. Das Rohr ist zusätzlich noch verstärkt um die Fläche einfacher zu gestallten. Die Abb. 17 zeigt die Verteilung der mechanischen Spannungen. Ausgewertet wird die Komponente in Richtung x.

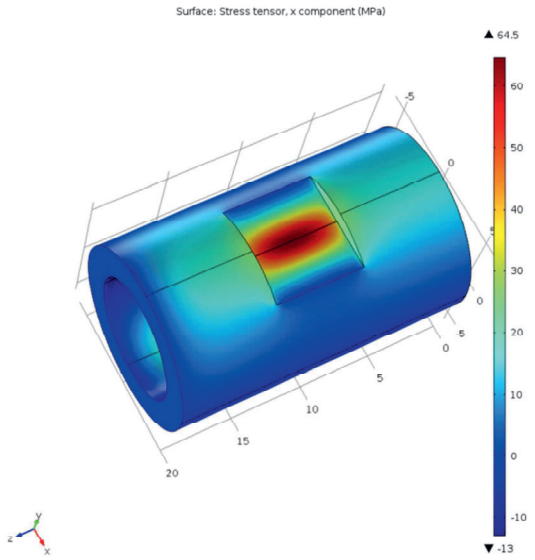


Abb. 16: Einfache Druckmessung im Rohr, Simulation 100 bar

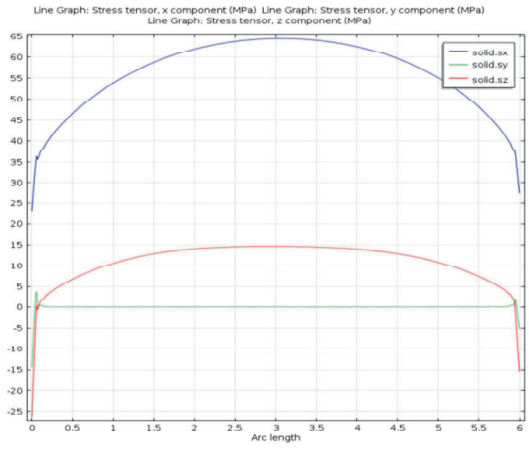


Abb. 17: Spannungsverteilung der einfachen Druckmessung im Rohr, Simulation 100 bar

Fazit

Es konnte gezeigt werden, dass mit der Anordnung von Sensoren auf dem Schraubenkopf die Vorspannkraft verlässlich bestimmt werden kann. Ein weitgehend kalibrierfreier Aufbau kann realisiert werden, wenn die Position der Sensoren auf dem Schraubenkopf exakt eingestellt werden. In diesem Fall wird es ausreichen nur den Nullpunkt zu kalibrieren, die Empfindlichkeit der untersuchten Sensoren lag in einer sehr engen Bandbreite. Die Langzeitstabilität konnte nur in Stunden untersucht werden, erforderlich sind aber Jahre, dies wird noch weitere Zeit benötigen.

Literatur

[1] J. Rausch, *Entwicklung und Anwendung miniaturisierter piezoresistiver Dehnungsmesselemente,* (2012)

[2] Th. Frank, M. Khatri, M. Fiedler, A. Gruen, M. Kermann, A. Cyriax, Th. Ortlepp, *Miniaturized silicon strain gauge elements to accurately measure mechanical quantities,* AMA Conferences 2015, Germany, (19.– 21.05.2015)

[3] Th. Frank, A. Grün, M. Kermann, A. Cyriax, A. Steinke, Th. Ortlepp, *Diagnoseprüfung und Monitoring an Verbindungselementen mit Hilfe von Silizium-Dehnmessstreifen,* 19. ITG/GMA-Fachtagung - Sensoren und Messsysteme 2018, Nürnberg, Deutschland, (26.-27.06.2018)

[4] Th. Frank, A. Grün, M. Kermann, A. Cyriax, A. Steinke, Th. Ortlepp, *Monitoring von Schraubverbindungen Verbindungselementen mit Hilfe von Silizium-Dehnmessstreifen, Technische Sicherheit,* Bd. 8, Nr. 9 (2018); 01.09.2018

Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Energie (BMWi), dem Bundesministeriums für Bildung und Forschung (BMBF) und der Europäischen Kommission für die finanzielle Unterstützung der Projekte AC-Brücke (MF150187), Bavi (49MF180012), HyPros-D (03ZZ0724D), Mikado (49MF180163), Screw Fixing Monitoring (ZF4001703HB8), SiStGAP (MF150091) und Smarter Si (ICT-02-2014).

DOI 10.5162/sensoren2019/2.3.1

Characteristics of vibrations in domestic environments as sources for kinetic energy harvesters

Mario Mösch, Gerhard Fischerauer
Chair of Measurement and Control Systems and Center of Energy Technology (ZET),
Universität Bayreuth, 95440 Bayreuth, Germany
mrt@uni-bayreuth.de

Abstract

Kinetic energy harvesters convert vibration energy to usable electrical energy. They work best when adapted to their environment. In particular, the resonance frequency of a resonant harvester should match the vibration frequency; otherwise, the harvested power decreases sharply. It is therefore important to know the vibration characteristics of the environment as well as possible. This work concerns itself with vibrations in domestic environments such as passive vibrations of solid structures like walls and ceilings or active vibrations of household devices. Particular attention is given to time-variant vibrations, i. e., to vibrations the characteristics of which vary with time. The results lead to guidelines for the mounting of kinetic energy harvesters and for useful harvester designs. We also present a sub-10-µW wireless sensor node for domestic temperature measurement. The small power consumption means that vibrations with small acceleration amplitudes suffice to power it via kinetic energy harvesting.

Keywords: Energy harvesting, vibration sources, characteristics, domestic environments, wireless sensor nodes.

Introduction

Wireless sensor networks (WSN) are advertised as basis for flexible condition monitoring or home automation tasks [1]. To fully exploit their potential, the WSN nodes must be operated cableless and without batteries. A potential approach to solve this problem is energy harvesting. Kinetic energy harvesters convert energy from the mechanical to the electrical domain by, e. g., electromagnetic [2] or piezoelectric principles [3]. The resonance frequency of a resonant harvester should match the vibration frequency; otherwise, the harvested power decreases sharply and may not suffice to power a WSN node.

Several solutions for this issue have been investigated in the past. They can be categorized according to their strategy. The first strategy is to widen the harvester's bandwidth, e.g., by building arrays with various resonance frequencies [4] or by nonlinear systems [5]. The second strategy is to change the resonance frequency with an adaptation mechanism, which can increase the output power unless the adaption process consumes too much energy itself [6], [7].

In either case, it is crucial to know the characteristics of the environmental vibrations so that a harvester with optimum response to these characteristics can be designed. The open literature reports on at least partial vibration measurements. Ref. [8] lists the resonance frequency and the acceleration amplitude of various sources such as a car hood, a microwave oven, and a tree. Ref. [9] does the same with household devices, e. g., a freezer and a washing machine, but does not discuss the influence of different operating points (device settings). Ref. [10] treats similar vibration sources as [8] and also goes into the details of time-variant spectra.

In our work, we characterized vibrations in domestic environments in an effort to derive rules for the design and the mounting of kinetic energy harvesters. We did not only consider the strengths and frequencies of the vibrations, but also determined how the vibration characteristics change as a function of time and the operating points of the respective vibration sources.

From this, consequences for vibration energy harvesting have been derived. We also demonstrate with an ultra-low-power WSN node how low-acceleration vibration sources can be exploited to power kinetic energy harvesters.

Minimum usable acceleration amplitude

We used acceleration measurements to characterize vibration sources. Assuming a pure-mode (monofrequent) vibration for simplicity's sake, the minimum acceleration amplitude required to power a WSN node is estimated as follows.

First note that the available electrical power P_{el} of a kinetic energy harvester at resonance is

$$P_{el} = m^2 A^2 c_{load} / 2c_t^2 \qquad (1)$$

with the mass m, the acceleration amplitude A, and the load and total damping coefficients c_{load} and c_t [11]. The prototype of Ref. [2] harvests 46 µW of power with $m = 0.66\,\mathrm{g}$ and a vibration with amplitude $A = 0.59\,\mathrm{ms}^{-2}$ at $f = 52\,\mathrm{Hz}$. This prototype would harvest 10 µW with $m = 1.82\,\mathrm{g}$ and $A = 0.1\,\mathrm{ms}^{-2}$ and otherwise unchanged conditions (the effective spring stiffness k needs to be modified to keep the resonant frequency $f_r = \sqrt{k/m}\,/\,(2\pi)$ constant). This value of 10 µW suffices to power a WSN node as is demonstrated below.

This leads us to believe that an acceleration amplitude of $A_{min} = 0.1\,\mathrm{ms}^{-2}$ is a reasonable estimate for the minimum vibration strength required to permanently supply a WSN node with its power.

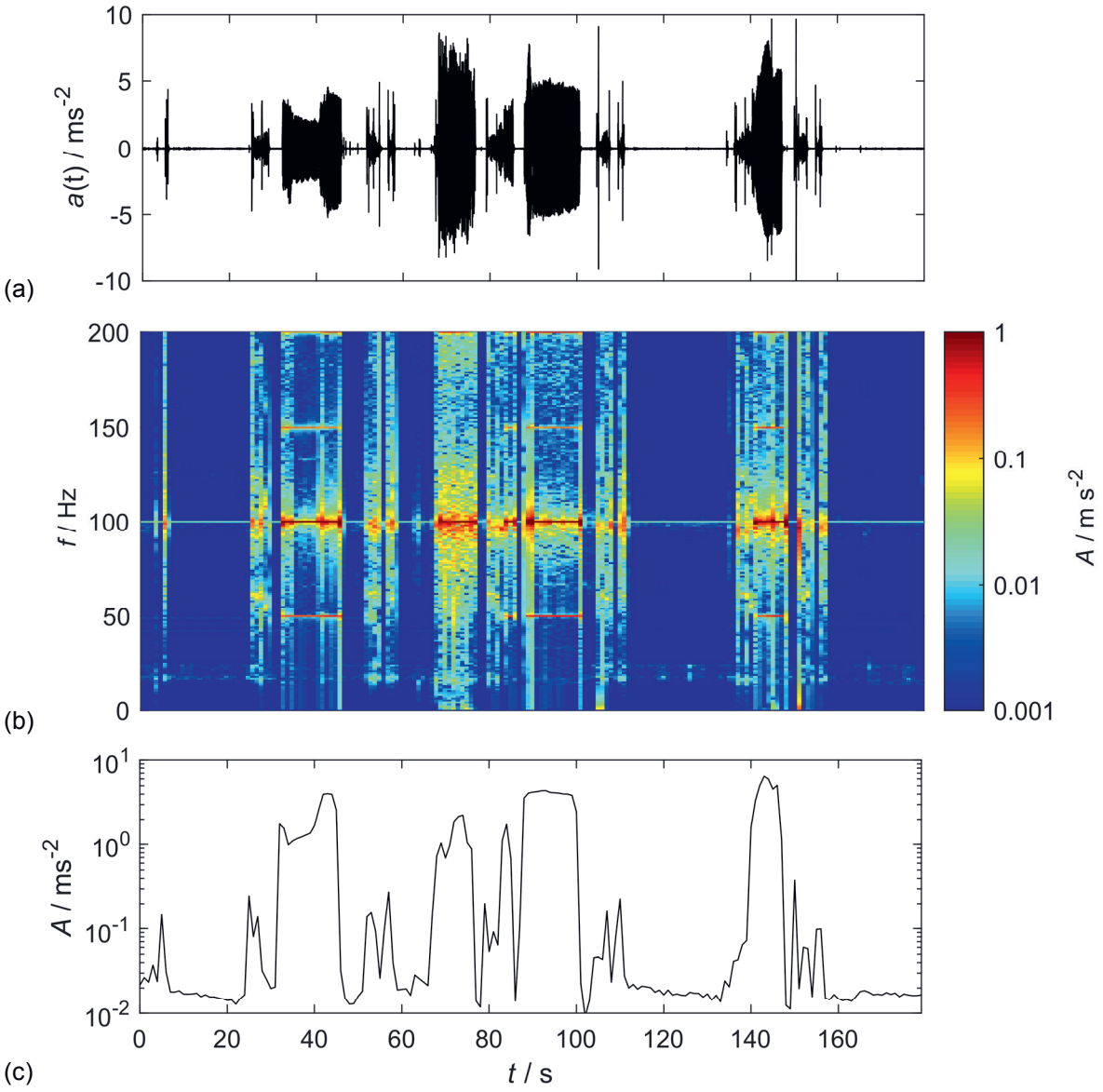


Fig. 1: Vibration of an automatic coffee machine during brewing. (a) Measured time response. (b) Spectrogram by short-time Fourier-transformation. (c) Time response of the 100-Hz vibration component. (See text on next page.)

Measurement approach

We investigated the vibrations on the floor, the ceiling, and the walls of a detached house and an office building and on devices operated therein, among them a microwave oven, a refrigerator, a kitchen hood, and an oil heating.

For the measurement, we mounted a high-sensitive acceleration sensor (PCB 356B18) on the vibration source to be investigated. Its output signal was amplified and registered by mobile measurement equipment (PCB480B21, NI-9239). The acceleration time series was transformed to the frequency domain by Fourier transformation. To cope with time-variant signals, we used the short-time Fourier transform (STFT).

By way of an example, Fig. 1 on the preceding page shows the results for an automatic coffee machine during brewing. One notes the time-variance of the vibration.

Measurement results

Figure 2 shows the amplitude spectra of vibrations on the floor, on a window frame, and on a door frame in the same room of a detached house. A resonance at 49.5 Hz, which is no mains interference, is present in every spectrum, while additional resonances at 24, 39, and 90.4 Hz occur only occasionally. The acceleration amplitudes of less than $3 \cdot 10^{-4}$ ms^{-2} are far below A_{min}. Such results are typical for passive vibrations of solid structures or switched-off devices.

More promising energy harvesting sources are shock vibrations. They occur nearly everywhere with peak accelerations up to several tens of ms^{-2} in domestic environments. The strength depends on the exact source of the shock, especially when created by human interaction, and the distance from the source. Furthermore, the typical duration of shock vibrations is in the seconds range or below.

Figure 3 shows an example, a 200-ms shock vibration with a peak acceleration amplitude of 0.5 ms^{-2}. The corresponding amplitude spectrum reveals that the dominant frequencies are about 100 Hz and 160 Hz, but the total spectrum is rather broadband. Such a broadband and high-peak vibration calls for special harvester geometries, e. g., piezoelectric membranes [12].

A drawback of shock vibrations is, of course, that they result from events that are difficult to predict. Consequently, shock vibrations are not suitable as energy sources for continuously operated WSN nodes, but can be used to trigger event-driven actions.

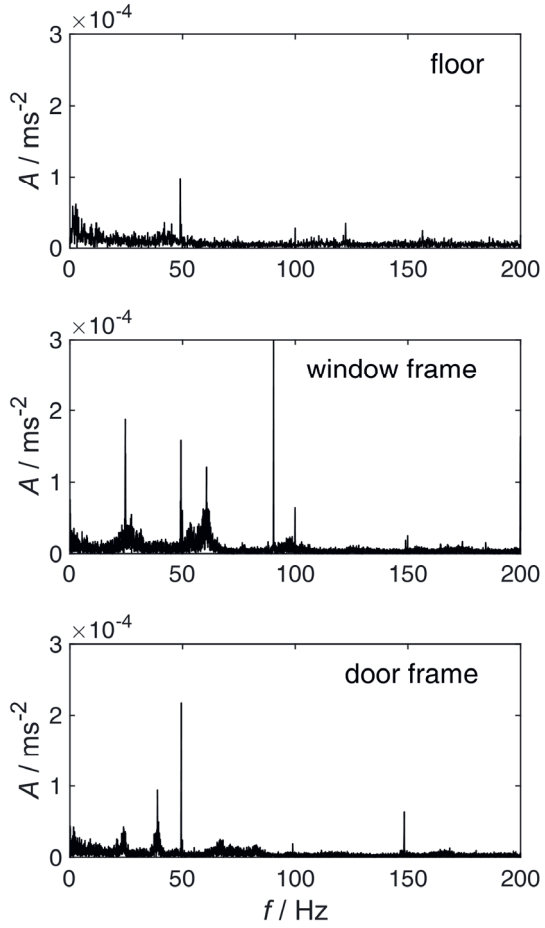


Fig. 2: Typical amplitude spectra of solid-structure vibrations in domestic environments. The resonances are too weak to be of practical use.

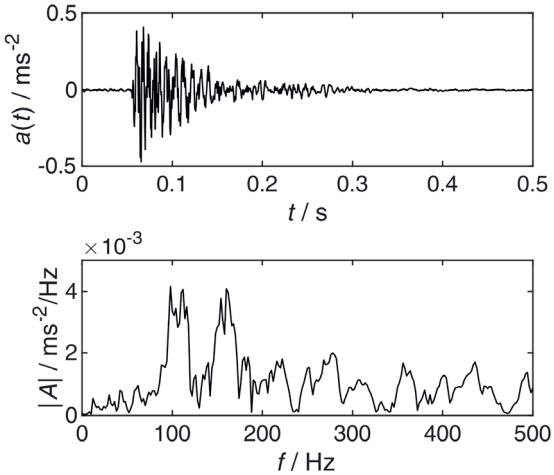


Fig. 3: Time response and associated amplitude spectrum of a shock vibration.

Household devices as potential vibration sources can be classified into devices with moving inner parts (e. g., washing machines) and other devices (e. g., TV sets). The latter are in the

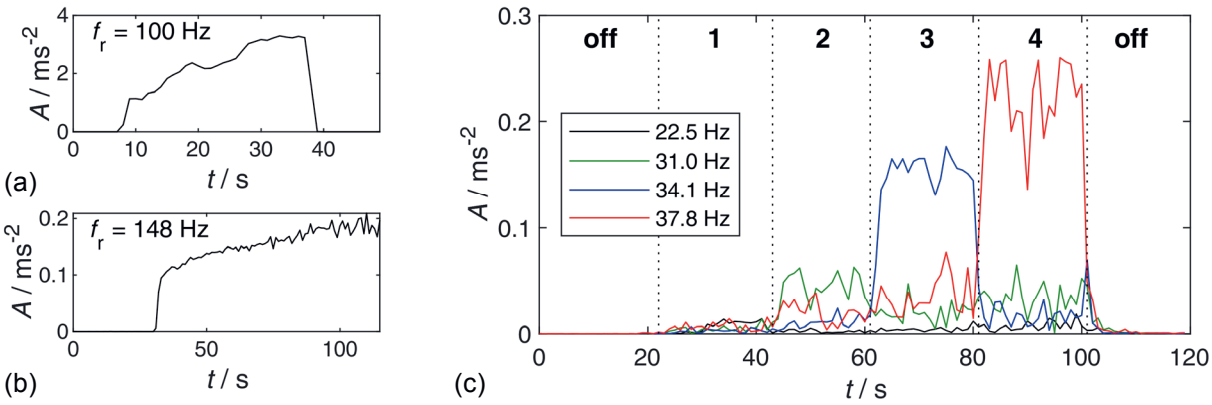


Fig. 4: Measured characteristics of some household devices. (a) Time-variant acceleration amplitude A at the main resonance frequency for a microwave oven. (b) As (a), but for a refrigerator. (c) Acceleration amplitudes at various frequencies for a kitchen hood over time. Different power stages (off, 1–4) obviously result in different main resonances. 1: 22.5 Hz; 2: 31.0 Hz; 3: 34.1 Hz; 4: 37.8 Hz).

same category as solid structures and switched-off devices. It is self-evident that only active vibrating devices are potential energy sources.

Based on their characteristics, these vibrating household devices can be subdivided into two classes V1, V2. The first class V1 comprises devices with a constant dominant resonance frequency. The coffee machine described by the results shown in Figure 1 belongs to this group of devices, even though the amplitude spectrum is broadened every so often because of the coffee grinding process. The maximum acceleration amplitude of the 100-Hz mode exceeds the 1 ms^{-2} limit several times and can reach up to 4.1 ms^{-2}. At other times, the acceleration amplitude stays below A_{min}.

Microwave ovens and refrigerators share similar characteristics with coffee machines in that their vibrations are dominated by single-mode resonances, cf. Figs. 4(a) and (b). The dominant 100-Hz mode of the microwave oven in Fig. 4(a) vibrates with acceleration amplitudes between 1 and 3 ms^{-2} whereas the 148-Hz mode of the refrigerator in Fig. 4(b) vibrates with amplitudes between 0.1 and 0.2 ms^{-2}. The time-variance of the vibration amplitudes has to be taken into account for the harvester design for two reasons. First, the maximum geometrical deflection of the harvester is a design parameter. Second, the expected harvested power influences how the overall system is to operated, e. g., how often a WSN node can transmit data.

The second class V2 consists of devices with variable-speed drives. Higher rotation speeds amount to higher resonance frequencies. Such effects are observed, e. g., with washing machines and kitchen hoods.

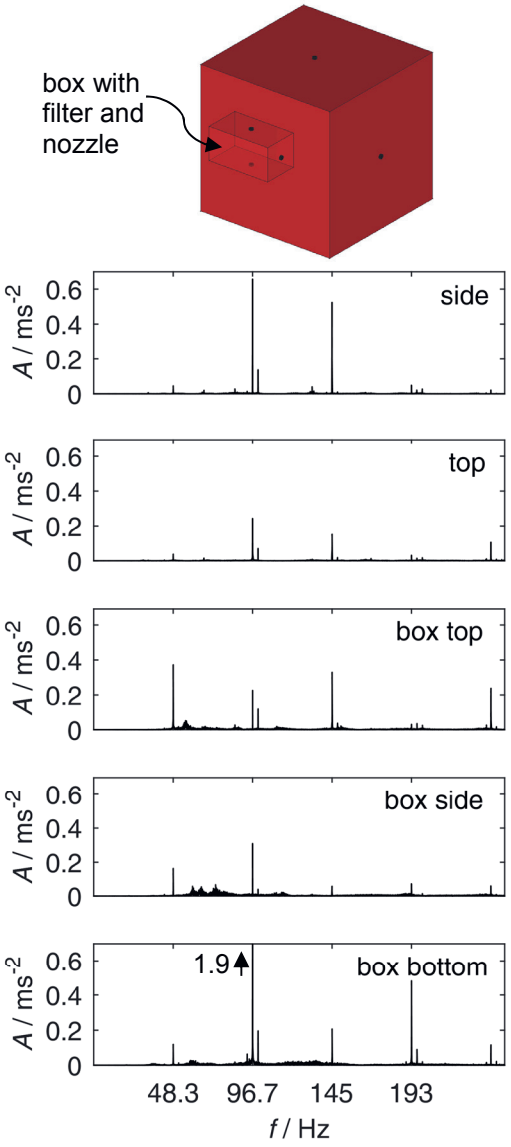


Fig. 5: Amplitude spectra of the vibrations on the surface of an oil heating.

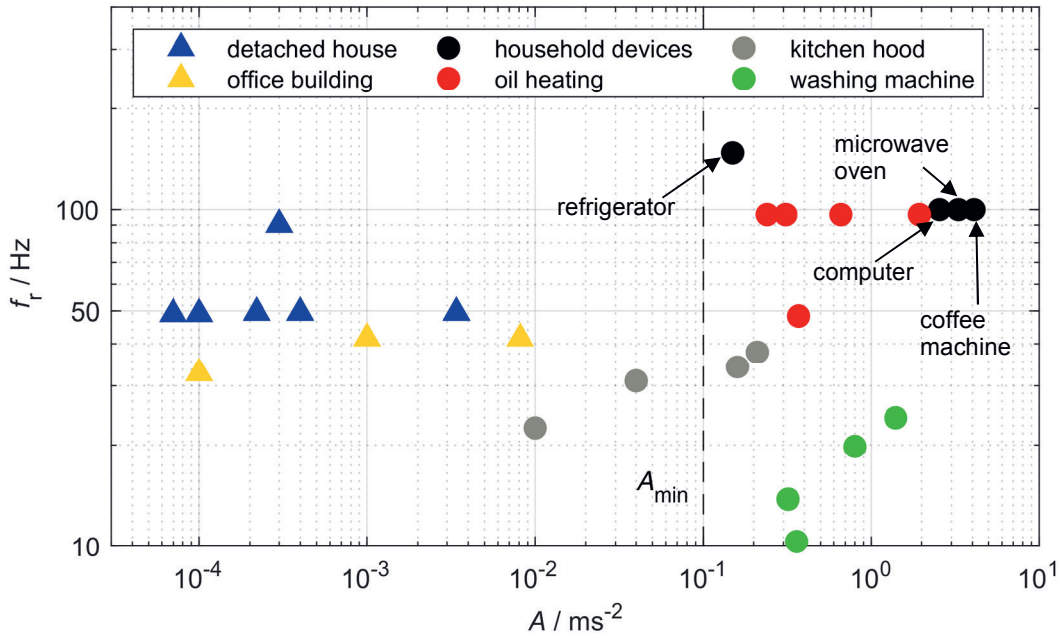


Fig. 6: Dominant resonance frequencies and associated acceleration amplitudes A of domestic vibration sources. Triangles and circles respectively mark passively and actively vibrating sources.

A washing machine tested by us offered various programs with four different speeds (600, 800, 1200, and 1600 rpm). The associated dominant resonance frequencies were about 10, 13, 20, and 27 Hz. The acceleration amplitudes varied from 0.3 to 1.4 ms^{-2}. An additional 100-Hz vibration could be tracked to the water pump and had amplitudes between 0.1 and 0.4 ms^{-2}.

A tested kitchen hood had four operating points with increasing suction due to increasing rotational speed of the fan motor. This leads to time-variant dominant vibration frequencies and amplitudes, cf. Fig. 4(c). It goes without saying that such time-variant behavior requires adaptive energy harvesters if the kinetic energy is to be harvested with maximum efficiency at each operating point.

The influence of the mounting position of an energy harvester was studied on the basis of an oil heating located in the basement of a detached house. The results are visualized in Fig. 5 on the preceding page. One observes a fundamental resonance at 48.3 Hz and its harmonics. Depending on the position of the acceleration sensor, quite different amplitudes were measured. The use of the 96.7-Hz signal appears to be the best overall choice as this signal is always present and is always above A_{min}. The general lesson from this is that the mounting position of an energy harvester on a vibration source well-suited in principle actually determines whether the results will be useful or useless.

Consequences for energy harvesting

Figure 6 shows the observed frequencies and acceleration amplitudes of the strongest vibrations in various cases. The acceleration amplitudes cover about five orders of magnitude. As a general rule, the low amplitudes occur for passively vibrating sources such as walls, ceilings, and floors, whereas the higher amplitudes above the useful limit of A_{min} = 0.1 m/s² involve actively vibrating sources. The highest acceleration amplitude (4 m/s²) was measured on the surface of a coffee machine.

Most resonances of actively vibrating devices in the household provide strong enough amplitudes to power WSN nodes. The resonances usually occur at the mains frequency of 50 Hz or multiples thereof. In these cases, linear vibration energy harvesters with fixed resonance frequencies of 50 or 100 Hz are sufficient. Only time-variant sources with variable-speed drives such as washing machines or kitchen hoods call for self-adapting harvesters [7].

Ultra-low-power WSN node for temperature measurement

We have assumed above that an acceleration amplitude of A_{min} = 0.1 m/s², which leads to a harvested power of 10 µW, may suffice for practical applications. This needs to be demonstrated. To this end, we considered a WSN node for home automation which measures the air temperature and transmits the result to a base station.

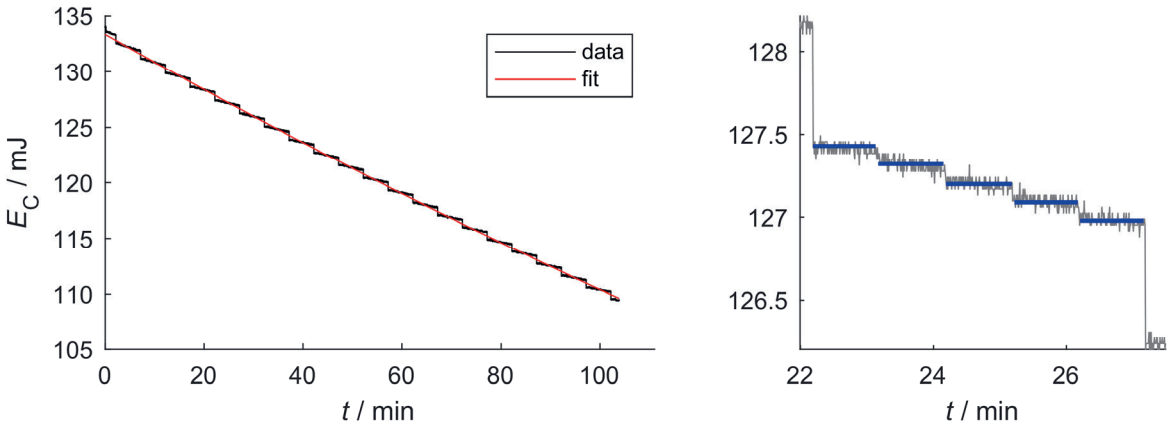


Fig. 7: Energy stored by a 47-mF capacitor connected to a WSN node with temperature sensor. The node transmits measurement values at intervals of $T_{Tx} = 300$ s. (a) Long-term measurement (—) and quadratic fit (—). (b) Detail of (a). The sensor averages $N = 5$ measured temperatures to obtain the final result. The horizontal blue lines mark deep-sleep periods.

To reduce the average power consumption of the WSN node, it is operated with alternating short measurement and data transmission cycles and long low-power (sleep) cycles. The WSN node consisted of a microcontroller (MCU) on a development board (STM32L452), a radio-frequency (RF) frontend for communication in the ISM band at 868 MHz (X-NUCLEO-IDS01A4) and an acceleration sensor with built-in temperature measurement capability (ADXL362). The MCU communicates with the RF frontend and the sensor via SPI bus.

After the cyclic wake-up from the deep-sleep state, the MCU initializes the sensor and receives several temperature values. The MCU enters a medium-sleep mode when waiting for a new value to save energy and is wakened by a pin interrupt. After having received several temperature values, the average temperature is transmitted to the basestation by RF communication at a transmit power of −10 dBm. Afterwards, the whole sensor node enters the deep-sleep mode with everything switched off except the real-time clock. The latter acts as an alarm clock which periodically wakes up the MCU. To reduce the power consumption, we also tested an alternative scenario in which the average temperature is stored by the MCu, but only transmitted every fifth time.

The sensor node is powered by a 47-mF capacitor. The system is operated such that the capacitor voltage stays between 1.8 and 2.8 V. Fig. 7 shows an example of the measured energy stored by the capacitor as a function of time. The instances at which the energy drops coincide with temperature measurement and data transmission processes.

T_{Tx} / s	N	P_{av} / µW
10	1	67
60	1	11.5
300	1 / 5	2.9 / 3.6

Table 1: Average power P_{av} required by the WSN node at a capacitor voltage of 2.2 V. The node transmits N averaged temperature values at time intervals of T_{Tx} (in other words, the deep-sleep time between two successive measurements is T_{Tx} / N).

Other operation scenarios than the one shown in Fig. 7 have been investigated. Since the power consumption of the MCU depends on its supply voltage, we aimed at making the results comparable by converting all measured capacitor energies to the values they would have had at a voltage of 2.2 V. This was done by quadratic fit functions such as the one shown in Fig. 7.

The influence of various operation scenarios on the WSN node power consumption is detailed in Table 1. Data transmission every 60 s and every 300 s respectively leads to an average power consumption of 11.5 µW and 3 to 4 µW. This justifies our prior assumption of a useful acceleration amplitude limit of $A_{min} = 0.1$ m/s².

Conclusions

We have experimentally characterized potential sources for vibration energy harvesting in domestic environments. It turns out that passively vibrating solid structures cannot power a WSN node via vibration energy harvesting. An exception are shock vibrations for event-driven systems.

The vibrations of several household devices, on the other hand, are strong enough to be useful for energy harvesting. We have encountered both stable narrow-band resonances and time-variant broad-band vibrations, which influences the optimum harvester design and its mode of operation. We have also shown that the mounting position of an energy harvester on the surface of a vibrating device is of critical importance.

Finally, we have shown by means of a demonstrator that an acceleration amplitude of 0.1 m/s², corresponding to an harvested power of 10 µW, suffices to operate an ultra-low-power WSN node.

Funding

This work was partially funded by the Bavarian State Ministry of Education, Science and the Arts within the framework of Technologie-AllianzOberfranken (TAO).

References

[1] M. Li, H.-J. Lin, *IEEE Trans. Indust. Electron.* 62, 4430–4442 (2015); doi: 10.1109/TIE.2014.2379586.
[2] S. P. Beeby et al., *J. Micromech. Microeng.* 17, 1257–65 (2007); doi: 10.1088/0960-1317/17/7/007.
[3] A. Khalid, A. K. Redhewal, *Mat. Sci. Appl.* 6, 818–827 (2015); doi: 10.4236/msa.2015.69084.
[4] D. Zhu, M.J. Tudor, S. P. Beeby, *Meas. Sci. Technol.* 21, 022001 (2010); doi: 10.1088/0957-0233/21/2/022001.
[5] M. F. Daqaq et al., *Appl. Mech. Rev.* 66, 040801 (2014); doi: 10.1115/1.4026278.
[6] D. Hoffmann et al., *Smart Mater. Struct.* 25, 035013 (2016); doi: 10.1088/0964-1726/25/3/035013.
[7] M. Mösch, G. Fischerauer, *Micromachines* 10, 44 (2019); doi: 10.3390/mi10010044.
[8] N. E. duToit, B. L. Wardle, S.-G. Kim, *Integr. Ferroelectr.* 71, 121–160 (2005); doi: 10.1080/10584580590964574.
[9] J. W. Matiko et al., *Meas. Sci. Technol.* 25, 012002 (2014); doi: 10.1088/0957-0233/25/1/012002.
[10] S. P. Beeby et al., *Smart Mater. Struct.* 22, 075022 (2013); doi: 10.1088/0964-726/22/7/075022.
[11] N. G. Stephen, *J. Sound Vibr.* 293, 409–25 (2006); doi: 10.1016/j.jsv.2005.10.003.
[12] D. Alghisi et al., *Sensors Actuat. A* 233, 569–81 (2015); doi: 10.1016/j.sna.2015.07.020.

DOI 10.5162/sensoren2019/2.3.2

Zutrittskontrolle mit Energy Harvesting aus der Türklinken-bewegung

Prof. Dr. Juan-Mario Gruber, BSc. Patrick Toma
Institute of Embedded Systems, ZHAW, Winterthur, Switzerland

Zusammenfassung

Dieser Beitrag stellt ein energieautarkes System für die Zugangskontrolle mittels Energy Harvesting vor. Das System ist für den Einbau in konventionelle Türen konzipiert. Die Einbaumasse entsprechen einem Standardtürschloss. Das vorgestellte energieautarke System kann herkömmliche Batteriebetriebene elektrische Schlösser ersetzen. Ziel des energieautarken Systems ist es, einen NFC-Tag anhand Umgebungsenergie auszulesen, um eine anschließende Authentifizierung durchzuführen. Bei erfolgreicher Authentifizierung wird die Tür mit einem Motor entriegelt. Es wurde ein funktionierender Prototyp erstellt und damit die Machbarkeit nachgewiesen.

Keywords: Energy Harvesting, Zugangskontrolle, NFC

Einführung

Das Einsatzgebiet von energieautarken Systemen hat sich in den letzten Jahren stark erweitert. Nicht zuletzt dank der energieeffizienteren Elektronik und durch Optimierung der Energy Harvester. Moderne Mikrocontroller und Sensoren benötigen immer weniger Energie, und Energy Harvesting Technologien konvertieren grössere Mengen an Umgebungsenergie. Dieser Trend sorgt für praxisrelevantere Lösungen. Energieautarke Systeme finden immer häufiger Anwendung in verkabelungstechnisch schwer zugänglichen Umgebungen oder als Ersatz für Batterien.

Anwendung und Funktion

In diesem Beitrag wird ein energieautarkes System zur Eintrittsidentifikation an Türen vorgestellt. Als Grundlage wird hier ein elektronisches Schloss verwendet. Handelsübliche elektronische Schlösser besitzen ein Auslesegerät für die Authentifikation und können die Tür entriegeln. Mit solch einem System kann sichergestellt werden, dass nur autorisierte Personen Zugang zu abgesperrten Bereichen erhalten. Für ein Hotelzimmer erhält jeder Gast eine Karte, mit der er eindeutig identifiziert werden kann. Mit dieser Karte ist es möglich das gemietete Zimmer zu betreten, jedoch kein anderes. Solche elektrischen Schlösser benötigen Zugang zu einem Stromnetz oder Batterien.
Das hier entwickelte energieautarke System sammelt die benötigte Energie aus der Drehbewegung der Türklinke und nutzt diese, um die Authentifikation mit anschließender Entriegelung der Tür durchzuführen.

Der Standardablauf zur Entriegelung der Tür beinhaltet folgende Schritte: Karte an das Lesegerät halten, die Türklinke betätigen, Karte auslesen, Tür entriegeln und anschliessen wieder verriegeln. Die einzelnen Schritte und die dazugehörige schematische Energiekurve sind in Abbildung 1 dargestellt. Zur Öffnung der Tür wird die Karte an das Kartenlesegerät gehalten. Beim Betätigen der Türklinke wird die mechanische Energie gesammelt und in elektrische Energie konvertiert. Damit wird die Karte ausgelesen, bei erfolgreicher Authentifizierung die Tür entriegelt und anschliessen wieder verriegelt.

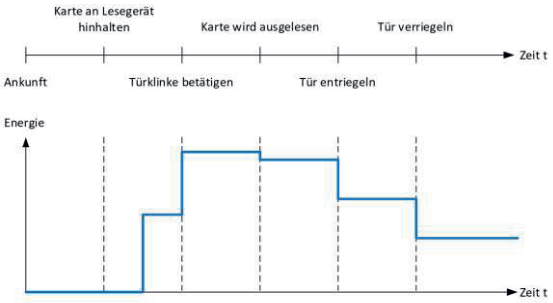


Abb. 1 Standardablauf mit Energiekurve

Systemkonzept und -aufbau

Das energieautarke System besteht aus sechs Komponenten.
Ein Motor mit einem zweistufigen Getriebe dient als Energy Harvester. Ein Spannungsregler wandelt die Ausgangsspannung des Motors in nutzbare 3.3 V um. Diese wir in einem Superkondensator gespeichert und über eine Powermanagementeinheit (PME) verwaltet. Zur Identifizierung dient ein Kartenlesegerät für NFC-Karten (Near Field Communication). Ein

Mikrokontroller übernimmt die Steuerung des Systems. Das Ver- und Entriegeln der Tür erfolgt durch Solenoid Aktuatoren. Abbildung 2 zeigt die einzelnen Komponenten im Zusammenspiel.

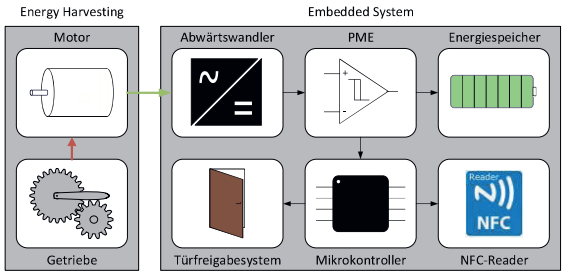


Abb. 2 Systemübersicht

Verriegelungssystem

Das Freigabesystem des Prototyps besteht aus zwei Solenoid Aktuatoren (Abbildung 4). Die Ver- und Entriegelung erfolgt durch lineares Verschieben des Türriegels. Der mechanische Aufbau der Solenoid Aktuatoren erzwingt eine lineare Bewegung des Kerns. Sobald ein Strom durch die Zylinderspule des Aktuators fließt, wird ein elektromagnetisches Feld aufgebaut und stößt den ferromagnetischen Kern in die oder aus der Spule. Die Richtung des Stroms bestimmt die Nord- und Südlage des Feldes (Abbildung 3).

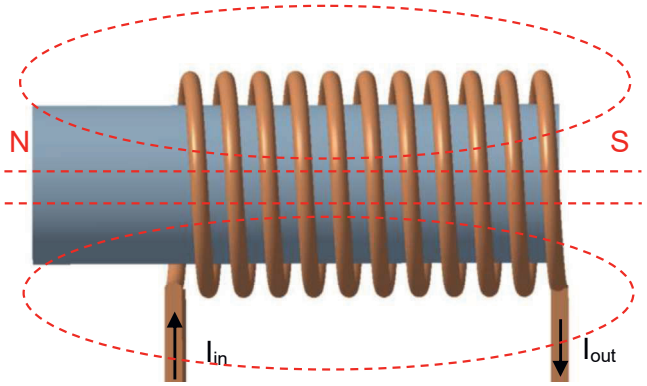


Abb. 3 Zusammenhang zwischen Pollage und Stromrichtung

Mit dieser Lösung wird keine Linearumsetzung benötigt, wie es bei einem regulären DC-Motor der Fall wäre. Alternativen wie Piezoaktuatoren benötigen meist eine hohe Antriebsspannung, oder haben einen geringeren Wirkungsgrad. Der rot umrandete Aktuator in Abbildung 4 wird verwendet, um die Tür zu entriegeln. Er muss nicht nur die Haft- und Gleitreibung des Türriegels, sondern auch die Feder- und Gewichtskraft des Kerns vom zweiten Aktuator überwinden können. Sobald der Türriegel ganz eingefahren ist, fällt der Kern des zweiten Aktuators automatisch in die Kerbe und hält den Türriegel im eingefahrenen Zustand. Um die Tür wieder zu verriegeln muss der zweite Aktuator betätigt werden und die Federkraft stößt den Türriegel in die entsprechende Öffnung im Türrahmen und verriegelt somit das Türschloss.

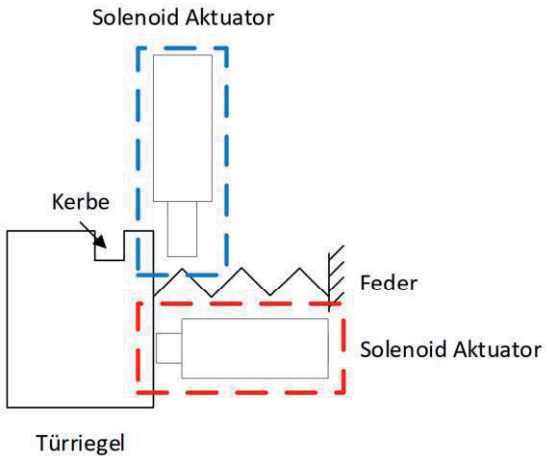


Abb. 4 Der Türriegel und das Entriegelungssystem

Energiequelle und Harvester

Die mechanische Rotationsenergie, die durch das Betätigen der Türklinke anfällt, wird mit einem Motor in elektrische Energie umgewandelt. In [6] wurden viele verschiedene Typen von Motoren getestet. Die Ergebnisse aus der Arbeit haben ergeben, dass sich Schrittmotoren aufgrund der zweiten Spule am Stator besser als Rotationsenergiekonverter eignen als DC-Motoren Zusätzlich wurde ein linearer Zusammenhang zwischen Rotationsgeschwindigkeit des Rotors und der konvertierten Energiemenge gefunden. Je schneller der Rotor dreht, desto mehr Energie wird konvertiert. Mit dem hier verwendeten Schrittmotor konnte ein Wirkungsgrad von bis zu $\eta = 0.45$ erreicht werden.

Der Schrittmotor ist über ein zweistufiges Getriebe mit der Türklinke verbunden, um den Aktionsradius der Türklinke zu erweitern und mehr Energie konvertieren zu können. Das Übersetzungsverhältnis beeinflusst die Kraft, welche angewendet werden muss, um die Türklinke zu betätigen. Handelsübliche Türklinken benötigen ein Drehmoment von etwa 1.25 Nm. Damit die Klassifizierung EN 12209 [7] eingehalten werden kann, darf das Getriebe im Verbund mit dem Schrittmotor ein maximales Übersetzungsverhältnis von 14.75:1 nicht überschreiten. Für den Prototyp wurde ein Übersetzungsverhältnis von 12:1 gewählt. Damit wird aus den anfänglichen 0.125 Um-

drehung an der Türklinke 1.5 Umdrehungen am Rotor des Schrittmotors.

Spannungswandler - Abwärtswandler

Die Ausgangsspannung des Schrittmotors beträgt bis zu 50 VAC. Für das Embedded System wird 3.3 VDC benötigt. Damit das Spannungsniveau angepasst werden kann, wurde ein Abwärtswandler mit vorgesetztem Brückengleichrichter eingesetzt (Abbildung 5 und Abbildung 6).

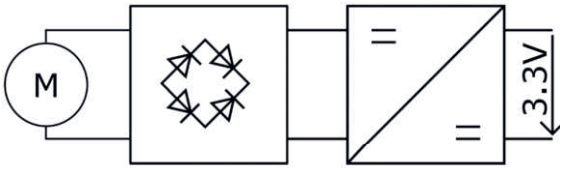


Abb. 5 Blockschaltbild der Versorgungsspannung

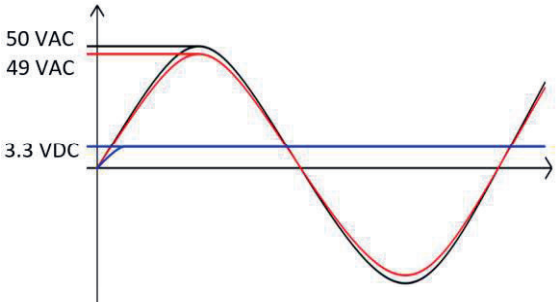


Abb. 6 Spannungskurven der Spannungsversorgung

Als Spannungswandler wurde ein LT3991 von *Linear Technology* ausgewählt aufgrund des sehr niedrigen Fliessstromes (2.1 µA) und des Eingangsspannungsbereiches von 4.3 VDC bis zu 55 VDC geeignet.

Energiespeicher

Für die hier betrachtete Anwendung ist der Bauraum stark begrenzt. Gleichzeitig kann der Strom durch die Spulen der verwendeten Solenoid Aktuatoren bis zu 1 A betragen. Daher wird ein Superkondensator verwendet. Der BZ015A104ZSB hat einen äquivalenten Serienwiderstand von 80 mΩ. Bei einer Spannung von 3.3 VDC können Ströme von bis zu 41.25 A fliessen.

Powermanagementeinheit

Die Powermanagementeinheit (Abbildung 7) ist auf Effizienz ausgelegt. Für die Spannungsüberwachung wird ein Komparator verwendet, dessen Ausgang zum invertierenden Eingang zurückgeführt wird und somit als Schmitt-Trigger geschaltet ist. So liegt am Ausgang des Komparators entweder V_{CC} oder GND an. Die Widerstände Rs3 und Rs4 erzeugen eine Hysterese am Ausgang.

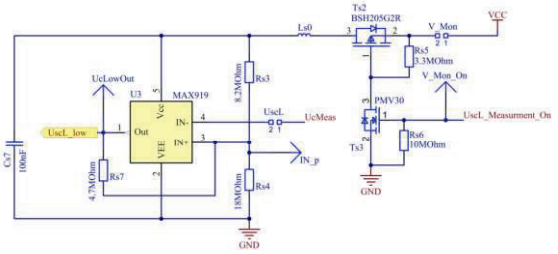


Abb. 7 Schema der Powermanagementeinheit

Prototyp des Gesamtsystems

Aus den betrachteten Komponenten wurde das Gesamtsystem prototypisch aufgebaut. Abbildung 8 zeigt ein CAD-Modell des Prototyps.
Die Grundstruktur sorgt für die mechanische Stabilität. Sie besteht aus Aluminium und trägt die übrigen Komponenten des Türschlosses.
Das zweistufige Getriebe mit Schrittmotor dient als Energy Harvester. Die Anordnung der Zahnräder des zweistufigen Getriebes (3) wurde so ausgelegt, dass die Löcher der Tür mit jenen des Türschlosses übereinstimmen. Der Schrittmotor kann von außen montiert werden, sobald das Türschloss in der Tür eingebaut ist. So ist es möglich das Türschloss zu montieren, ohne eine handelsübliche Tür zu bearbeiten.
Das energieautarke Embedded-System enthält das Energiemanagement und steuert den Ablauf. Das Freigabesystem ist mechanisch mit dem Türriegel gekoppelt.

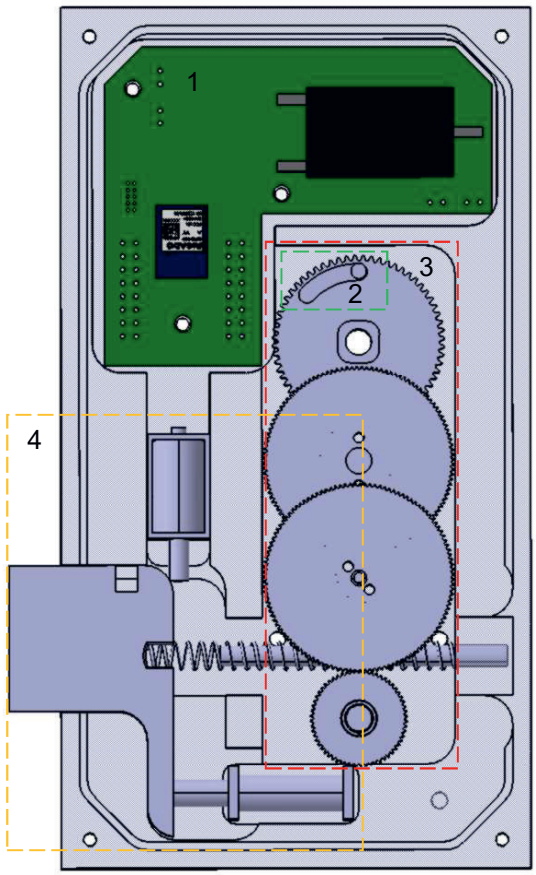


Abb. 8 CAD-Modell des Prototyps

Energiebilanz

Sobald die Türklinke betätigt wird, stehen dem energieautarken System 125 mJ [8] in Form von elektrischer Energie zur Verfügung. Dem gegenüber steht ein Energiebedarf von 58 mJ, der sich aus den einzelnen Schritten zusammensetzt (Abbildung 9).

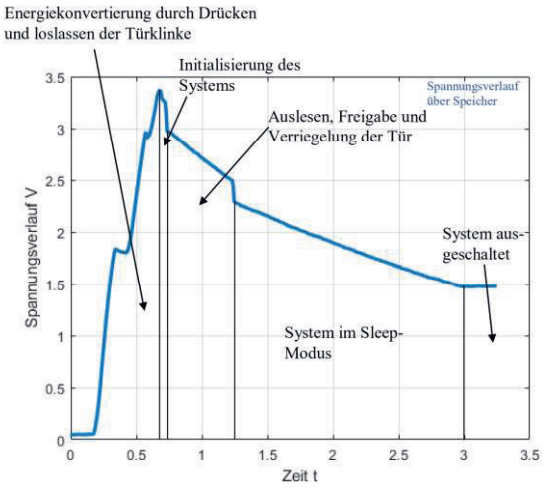


Abb. 9 Verlauf der Spannung am Kondensator

Das Starten und Initialisieren des Mikrokontrollers benötigt rund 1.1 mJ und das Auslesen der Karte 6.8 mJ. Mit den Low-Power-Solenoid-Aktuatoren kann die Tür mit 50 mJ entriegelt und anschliessend wieder verriegelt werden. Damit steht eine Reserve von 67 mJ zur Verfügung.

Fazit

Der in diesem Beitrag vorgestellte Prototyp zeigt, dass Energy Harvesting auch für komplexe Systeme mit mechanischen Aktuatoren geeignet ist. Die gesammelte Energie aus der Betätigung der Türklinke reicht aus, um das System zuverlässig zu betreiben. Im aktuellen Stand wird eine Energiereserve von über 50% erreicht. Optimierungen am Energiebedarf ermöglichen eine Anpassung des Übersetzungsverhältnisses des Getriebes, um den Kraftaufwand zur Betätigung der Türklinke zu minimieren.

Literaturverzeichnis

[1] STMicroelectronics, „https://www.st.com,“ 12 Juni 2017. [Online]. Available: https://www.st.com/resource/en/datasheet/cr95hf.pdf. [Zugriff am 08 Mai 2019].

[2] Adafruit Industries LLC, „https://cdn-shop.adafruit.com,“ 2017. [Online]. Available: https://cdn-shop.adafruit.com/product-files/324/C140-A+datasheet.jpg. [Zugriff am 08 Mai 2019].

[3] Linear Technology, „https://www.analog.com“ 1. November 2009. [Online]. Available: https://www.analog.com/media/en/technical-documentation/data-sheets/3991fa.pdf. [Zugriff am 8 Mai 2019].

[4] AVX Corporation, „http://datasheets.avx.com“ 2018. [Online]. Available: http://datasheets.avx.com/BestCap.pdf. [Zugriff am 8 Mai 2019].

[5] „https://nordicsemi.com,“ 10 Oktober 2017. [Online]. Available: https://infocenter.nordicsemi.com/pdf/nRF52832_PS_v1.4.pdf. [Zugriff am 8 Mai 2019].

[6] ZHAW - InES, „Energy Harvesting aus Türklinke - Evaluation,“ Winterthur, 2017.

[7] VSSB, „http://www.vssb.ch,“ 1 September 2005. [Online]. Available: http://www.vssb.ch/?inc=normen/normen.asp. [Zugriff am 8 Mai 2019].

[8] ZHAW - InES, „Energy Harvesting aus Türklinke“ Winterthur, 2018.

DOI 10.5162/sensoren2019/2.3.3

SMERS – Energieautarkes drahtloses Sensornetzwerk zur Detektion von Katastrophen und zur Lagebilderfassung in U-Bahnsystemen

Jonah Vincke[1], Clemens Horch[1], Frank Schäfer[1]
[1]Fraunhofer Institut für Kurzzeitdynamik, Ernst-Mach-Institut, EMI, Abteilung Systemlösungen
Ernst-Zermelo-Straße 4, 79104 Freiburg, Deutschland

Zusammenfassung

Bei Katastrophen in Tunnelsystemen von U-Bahnen stehen den Rettungskräften, wie der Feuerwehr, nur wenige Informationen zur Verfügung. Dies erschwert und verzögert die Arbeit erheblich. Um sich diesem Problem anzunehmen, wurde das drahtlose Sensornetzwerk „SMERS“ (security management and ermergency response system) entwickelt. Es detektiert Feuer, Wassereinbruch und Explosionen. Um die Robustheit zu erhöhen befindet sich alle 50-100 m ein Sensorknotenpaar, welches energieautark betrieben wird. Die Datenübertragung und Sensorik wurden auf einen besonders energiearmen Betrieb hin optimiert. So wird bei einer maximalen Folge von 20 Knotenpaaren für den Betrieb und die Datenübertragung 316 µW verbraucht und für die Sensorik je nach Knotentyp zwischen 13 µW und 36 µW. Zur Stromerzeugung wird ein piezoelektrischer Vibrationsharvester verwendet. Dieser liefert für die in einem Tunnel der Berliner S-Bahn gemessenen Vibrationen eine Leistung von bis zu 107 µW bei einer Zugdurchfahrt alle fünf Minuten. Bei einem Feldtest in einem Tunnel der Berliner U-Bahn wurden auf einer Strecke von ungefähr 750 m von 27 Knoten über eine Million Nachrichten aufgezeichnet.

Keywords: Drahtloses Sensornetzwerk, Vibrationsharvester, U-Bahntunnel, Lagebilderfassung

Einleitung

Aufgrund der zunehmenden Urbanisierung spielt der öffentliche Nahverkehr und insbesondere U-Bahnsysteme zur Bewältigung der Personenströme eine immer größere Rolle. Um die bestehenden Sicherheitssysteme weiter zu verbessern wurde „SMERS“ (security management and ermergency response system) entwickelt [1]. Im Fokus stehen die frühzeitige Erkennung von Naturkatastrophen und Anschlägen, sowie die verbesserte Lagebilderfassung speziell für die Rettungskräfte der Feuerwehr.
Bei SMERS handelt es sich um ein drahtloses Sensornetzwerk, bei dem ca. alle 70 m ein Sensorknotenpaar im Tunnel installiert ist [2]. Jedes Paar detektiert und misst dabei Wassereinbruch, Explosionen und Feuer (Rauch und Temperatur) [3]. Die Knoten sind energieautark um größtmögliche Unabhängigkeit zu gewährleisten. Hierfür wurde die Versorgung entweder mit einem an den Schienen befestigten Vibrationsharvester oder mit einem Windharvester in Kombination mit einer Backup-Batterie untersucht. Gleichzeitig wurden sämtliche Komponenten dahingehend ausgesucht oder entwickelt, dass sie möglichst wenig Energie benötigen. In einem Feldversuch in einem Tunnelabschnitt der Berliner U-Bahn wurde ein System bestehend aus insgesamt 27 Knoten und 2 Senken erprobt (siehe Abb. 1).

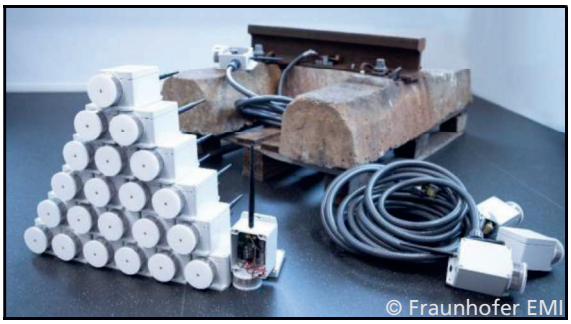


Abb. 1: Sensorknoten, teils mit Vibrationsharvester

Stand der Technik

Vergleichbare drahtlose Sensornetzwerke für Tunnelsysteme wurden in verschiedenen Arbeiten untersucht. Dabei wurden meist Autotunnel [5,6,7] oder Minentunnel [8,9] betrachtet. Weniger im Fokus liegen Zug- [10,11] und U-Bahn-Tunnel [12,13].
Bennet et al. [12] installierten ein drahtloses Sensornetzwerk in der Londoner und der Prager U-Bahn. Dabei betrug der untersuchte Abschnitt lediglich 170, bzw. 115m.

Cheung et. al [13] untersuchten ein drahtloses Sensornetzwerk zur Gasdetektion in U-Bahntunneln im selben Frequenzband wie die vorliegende Arbeit. Neben der anderen Anwendung wurden nur drei Knoten eingesetzt und ein energiesparender Betrieb wurde nicht berücksichtigt.
Harvester für Tunnelumgebungen wurden in verschiedenen Arbeiten untersucht [14,15,16]. Cleante et. al [14] untersuchten ebenfalls die Dimensionierung von Vibrationsharvestern an Schienen in England. Degrande et. al [15] charakterisierte die erzeugten Vibrationen in den unterschiedlichen Bauteilen der Zugstrecke (Boden, Gleis, Schwelle und Wand) in einer Londoner Teststrecke in Abhängigkeit von der Zuggeschwindigkeit. Neben der kurzen Erwähnung von Vibrationsharvestern untersuchten Wang et. al [16] die Simulation und Charakterisierung von durch einen Zug erzeugten Wind und die optimale Platzierung eines Windharvesters.

Drahtloses Sensornetzwerk

Bedingt durch den Einsatz in U-Bahn-Tunneln handelt es sich um ein Sensornetzwerk mit einer linienförmigen Topologie (siehe Abb. 2 und [2]). In jedem Bahnhof befindet sich eine Senke, welche die gesammelten Daten an einen zentralen Server weiterleitet. Zwischen den zwei Senken befinden sich die Sensorknotenpaare in einem Abstand von 50 m bis 100 m. Aufgrund der Anforderungen durch die Sensorik und zur Steigerung der Ausfallsicherheit befinden sich an jeder Position zwei Sensorknoten, ein oberer an der Decke und ein unterer an der Wand (siehe Abb. 3). Die einzelnen Knoten basieren auf dem Mikrocontroller CC2650 von Texas Instruments mit integriertem Funkmodul und sind modular aufgebaut. So wird je nach Typ zusätzlich noch ein Sensorboard und Harvesterboard verwendet. Als Betriebssystem wird Contiki OS [17] und zur Datenübertragung eine modifizierte Version des enthaltenen Protokollstapels RIME verwendet.

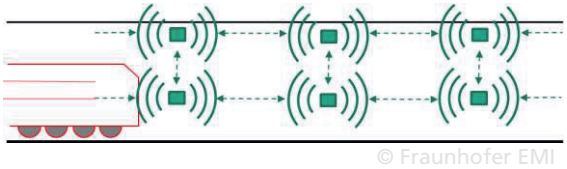


Abb. 2: Schematische Darstellung der Netzwerktopologie

Sensorik

Ziel des Systems ist die Detektion von Feuer, Wassereinbruch und Explosionen. Hierfür werden verschiedene Sensoren im oberen und unteren Knoten eingesetzt. Der Detektor für Explosionen befindet sich im oberen Knoten. Feuer wird jeweils in beiden Knoten detektiert. Der Wasserstand wird im unteren Knoten ermittelt (siehe Abb. 3) und lediglich dessen Messgenauigkeit durch einen Hilfswert des oberen Knotens verbessert.
Zur Detektion von Explosionen wird im oberen Knoten ein passiver Ultraschallsensor verwendet. Um den Stromverbrauch zu reduzieren wird das Signal nicht direkt gemessen. Stattdessen wird das Überschreiten von drei Schwellwerten mit Hilfe einer analogen Schaltung erkannt und die Dauer des Überschreitens ermittelt. Aus diesen Messungen lässt sich der Spitzendruck [MPa] sowie die Stoßwellenenergie [MJ] abschätzen.
Zur Detektion von Wassereinbruch und der Bestimmung des Wasserstandes wird im unteren Knoten ein Differenzdrucksensor verwendet, welcher über einen Schlauch den Druckunterschied zwischen dem Tunnelboden und dem Sensorknoten misst. Über den vom Wasser induzierten hydrostatischen Druck lässt sich der Wasserstand bestimmen. Mit Hilfe des atmosphärischen Drucks wird die Kompression der Luft im Schlauch kompensiert und so die Messgenauigkeit verbessert [3,4]. Dieser wird mit einem Absolutdrucksensor im oberen Sensorknoten ermittelt.
Zur Früherkennung von Rauch wird ein photoelektrischer Rauchmelder an der Tunneldecke verwendet. Zusätzlich wird zur Erkennung von Feuer und zur Lagebilderfassung die Temperatur mit dem im Drucksensor integrierten Temperatursensor bestimmt. Im unteren Sensorknoten wird der gleiche Rauchmelder verwendet, welcher aber auf Kopfhöhe im Tunnel angebracht ist. Hierdurch kann die Begehbarkeit des Tunnelabschnittes für Personen ohne Atemschutz besser abgeschätzt werden.

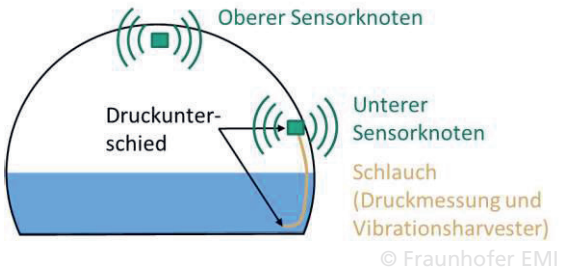


Abb. 3: Schematische Darstellung des Sensorkonzeptes zur Detektion von Wassereinbruch

Vibrationsharvester

Zur Energieversorgung der unteren Knoten wird ein Vibrationsharvester verwendet (siehe Abb. 4 und 6). Dieser besteht aus einem

einseitig eingespannten Biegebalken und einer am frei schwingenden Ende angebrachten Kopfmasse. Der Balken besteht aus einem Trägermaterial und einer beidseitig aufgebrachten Piezokeramik. Hierbei wurde für die Versuche der PPA-4011 der Firma Mide Technology genutzt. Die Auswahl des passenden Biegebalkentyps wurde auf Grundlage eines mathematischen Models in Anlehnung an die Arbeiten von Erturk A. und Inman D. [18,19,20] getroffen [21].



Abb. 4: Schematische Darstellung des Vibrationsharvesters

Bei einer Verformung d es Balkens wird eine Spannung erzeugt. Durch die Wahl des Gewichtes und der freien Länge des Balkens können die Eigenfrequenz des Harvesters festgelegt werden (siehe Abb. 5). Durch eine Abstimmung der Eigenfrequenz auf die vom Zug induzierten Frequenzen mit den größten Amplituden lässt sich die gewonnene Energie maximieren.

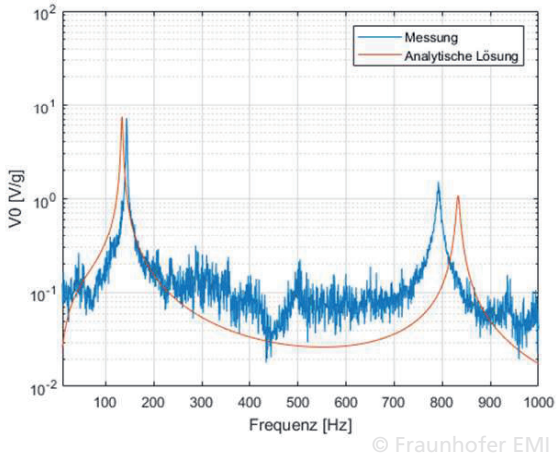


Abb. 5: Frequenzantwort eines Vibrationsharvesters vom Typ PPA-1011 mit einer freien Länge von 48mm und ohne Gewicht auf eine Beschleunigung von 1 g [21]

Zur Befestigung des Harvester an der Schiene wird ein entsprechend der Neigung des Schienenfußes ausgeführter Adapter am Schienenfuß mit Epoxidharz-Kleber befestigt. Auf diesen Adapter wird der Harvester per Schraubverbindung befestigt. Die Klemmvorrichtung für den Biegebalken ist ebenfalls als Schraubverbindung ausgeführt um eine leichte Anpassung der Eigenfrequenz zu ermöglichen.

Experimente

Neben der Bestimmung der Messgenauigkeit und des Energieverbrauchs der Knoten und dessen Teilkomponenten (siehe [2,3,4]) unter Laborbedingung wurde das Gesamtsystem in einem Feldtest in der Berliner U-Bahn untersucht. Hierfür wurden 8 Knoten im nicht befahrenen Übungstunnel Jungfernheide und 19 Knoten zwischen der Haltestelle Jungfernheide und dem Jakob-Kaiser-Platz installiert. Die oberen Knoten wurden aufgrund logistischer Gegebenheiten nicht an der Tunneldecke sondern in 3 m Höhe an der Wand befestigt. Die unteren Knoten wurden wie geplant auf Kopfhöhe in ca. 1,6 m befestigt. Zusätzlich befand sich in beiden Haltestellen jeweils eine Senke, welche die Daten lokal speicherten und per Mobilfunk (UMTS) an den zentralen Server weiterleiteten. Die Knoten im inaktiven Tunnel waren 123 Tage und die Knoten im aktiven Tunnel 101 Tage lang installiert. Insgesamt wurden über eine Million Übertragungen ausgewertet.

Zur Auslegung des Vibrationsharvesters wurden Vibrationsmessungen in einem Tunnel der Berliner S-Bahn und in der Berliner U-Bahn durchgeführt. Hierzu wurden an der Schwelle und auf dem Schienenfuß jeweils ein Beschleunigungssensor von Kistler (K-SHEAR Type 8704B500) mit Hilfe eines Adapters befestigt. Die Beschleunigungssensoren wurden über eine Kistler Type 4134A Kuppler an das mobile Oszilloskop PicoScope 4424 angeschlossen. Die Messungen wurden automatisch über eine Dauer von 24 Stunden durchgeführt. Zu den unterschiedlichen Messpositionen gehörte je ein Abschnitt in dem der Zug beschleunigt, bremst und bei dem er die maximale Geschwindigkeit erreicht hat.

Die Vibrationsmessungen wurden genutzt, um die optimale Auslegung des Harvesters zu ermitteln und die damit maximal generierbare Energie zu bestimmen. Hierfür wurden die Signale ausgewertet und die Eigenfrequenz des Harvesters entsprechend eingestellt. Anschließend wurde der Harvester auf einem Shaker befestigt und die aufgenommenen Beschleunigungsverläufe reproduziert (siehe Abb. 6). Über die Messung der erzeugten Spannung an einem Lastwiderstand von 5,4 kΩ, welche ungefähr dem Blindwiderstand des Harvesters entspricht, wurde die generierte Energie ermittelt.

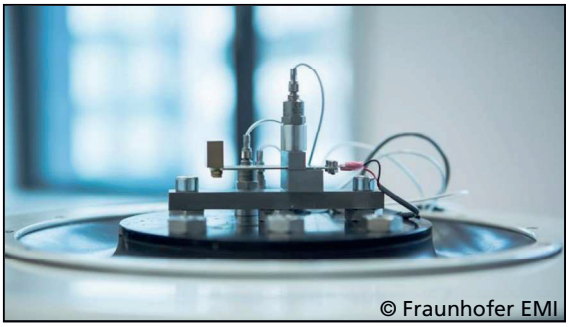


Abb. 6: Versuchsaufbau Vibrationsharvester

Versuchsergebnisse

Bei dem Feldversuch in der Berliner U-Bahn wurden insgesamt 1.013.065 Nachrichten im aktiven Tunnel und 187.808 im inaktiven Tunnel aufgezeichnet. Obwohl durch das Rime Protokoll verhindert werden soll, dass Nachrichten mehrfach einen identischen Knoten passiert, kam es bei 4060 Nachrichten zu einer solchen Schleife. Aufgrund der sehr geringen Häufigkeit konnte das Verhalten nicht gezielt reproduziert werden und die Ursache nicht ermittelt werden. Von den Nachrichten im aktiven Tunnel waren 385.051 Nachrichten an einen Knoten gesendet, welcher kein direkter Nachbar war. Hierzu zählt auch der andere Knoten eines Knotenpaares. Bei jeder Übertragung wurde als Kenngröße für die Signalstärke der RSSI (Received Signal Strength Indicator) bestimmt. Bei einem Freifeldversuch wurden RSSIs zwischen -38 dBm (ohne Abstand) und -95 dBm (200 m Abstand) gemessen. Erst ab 200 m kam es zu Paketverlusten. Hier betrug der RSSI zwischen -89 dBm und -95 dBm. In Tab. 1 sind die durchschnittlichen RSSIs in Abhängigkeit von der Lage des Senders und Empfängers dargestellt.

Der Gesamtenergieverbrauch des Systems hängt maßgeblich von der Last durch die Funkübertragung ab (siehe [2]). So beträgt der Stromverbrauch bei einer Channel-Check-Rate von 4 Hz und für einen Tunnelabschnitt mit bis zu 20 Knotenpaaren 316 µW, falls alle Knoten gleichverteilt Nachrichten übertragen. Um den Verbrauch für die vom Vibrationsharvester versorgten, unteren Knoten zu reduzieren, wurde ebenso ein Schema betrachtet, bei welchem nur die oberen Knoten Nachrichten weiterleiten und die unteren Knoten lediglich Nachrichten senden. Hierdurch erhöht sich die Leistungsaufnahme der oberen Knoten auf 482 µW pro Knoten, während die benötigte Leistung der unteren Knoten auf 23 µW reduziert wird. Zusätzlich zur benötigten Leistung des Betriebs und der Datenübertragung kommt noch die Sensorik, welche bei den oberen Knoten 13 µW und den unteren Knoten 36 µW benötigt [3].

Tab. 1: Durchschnittlicher RSSI [dBm] des Empfängers (Zeile) in Abhängigkeit des Knotentyps des Senders (Spalte)

	Obere Knoten	Untere Knoten
Obere Knoten	-75,77	-75,63
Untere Knoten	-76.31	-78,25
Direkte obere Nachbarn	-71,95	-73,72
Direkte untere Nachbarn	-76,29	-77,2

Tab. 2: Standardabweichung des durchschnittlichen RSSI [dBm] des Empfängers (Zeile) in Abhängigkeit des Knotentyps des Senders (Spalte)

	Obere Knoten	Untere Knoten
Obere Knoten	3,31	4,04
Untere Knoten	2,82	3,63
Direkte obere Nachbarn	5,14	6,8
Direkte untere Nachbarn	8,27	9,17

Bei der Messkampagne in dem Berliner S-Bahntunnel wurde an drei unterschiedlichen Positionen die in den Schienen erzeugten Schwingungen gemessen. Um den Harvester optimal auszulegen, wurden die Signale mittels schneller Fourier-Transformation vom Zeit- in den Frequenzbereich transformiert. Anschließend wurde der Mittelwert der Amplituden aller Messungen gebildet. Diese sind in Abb. 7-9 dargestellt. Hierbei sind beim Hochgeschwindigkeitsbereich zwei ausgeprägte Maxima bei 70,9 Hz mit 0,009 und bei 442,1 Hz mit 0,016 zu erkennen. Bei den beschleunigten Bereichen sind keine Maxima besonders ausgeprägt und die Amplitude liegt bei maximal 0,003. Außerdem wurde zur Ermittlung der einzustellenden Eigenfrequenz des Harvesters die Wahrscheinlichkeit berechnet, dass die Amplitude im Frequenzbereich bei einer Durchfahrt des Zuges einen festgelegten Wert überschreitet. Ein Beispiel für einen Grenzwert von 0,004 ist in Abb. 10 dargestellt. So ist auch hier die größte Wahrscheinlichkeit bei den vorher ermittelten Maxima zu finden.

Anhand der untersuchten Signale wurde als optimale Eigenfrequenzen des Harvesters eine Frequenz von ca. 71 Hz gewählt. Zur Ermittlung der erzeugten Energie wurde zudem ein Lastwiderstand von 5,4 kΩ verwendet. Unter diesen Bedingungen generiert der Harvester bei dem Hochgeschwindigkeitsabschnitt durchschnittlich 32,1 mJ und zwischen 0,9 mJ

und 1,6 mJ bei den beschleunigt befahrenen Abschnitten. Dies entspricht einer Leistung von 107 µW für den Hochgeschwindigkeitsabschnitt, sofern alle fünf Minuten ein Zug vorbeifährt. Für Abschnitte, in denen der Zug beschleunigt, entspricht dies bei gleicher Taktrate einer Leistung von 3 µW bis 5,3 µW.

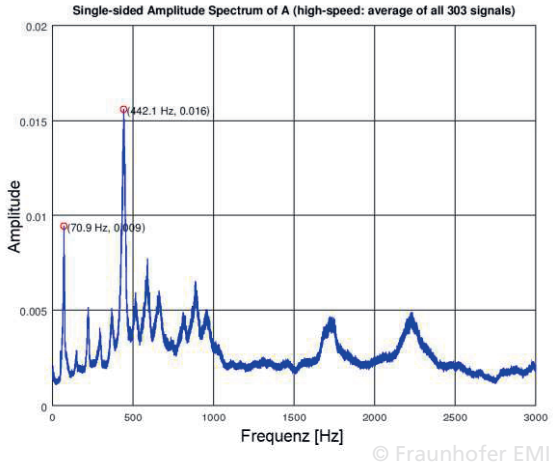


Abb. 7: Amplitudenspektrum der in den Schienen gemessenen Vibrationen bei maximaler Geschwindigkeit des Zuges im Berliner S-Bahntunnel

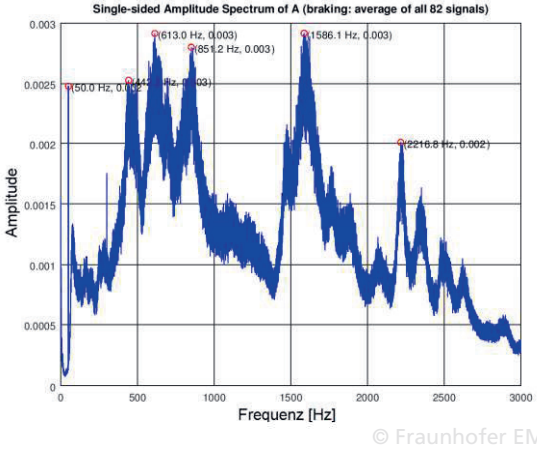


Abb. 8: Amplitudenspektrum der in den Schienen gemessenen Vibrationen beim Bremsen des Zuges im Berliner S-Bahntunnel

Diskussion

Durch die Laborversuche und den Feldtest wurde die Realisierbarkeit des Systems demonstriert. Die Sensorik erfüllt die Anforderungen an Messgenauigkeit, Wertebereich und geringem Energieverbrauch. Der Gesamtenergieverbrauch wird maßgeblich von der Übertragung bestimmt.

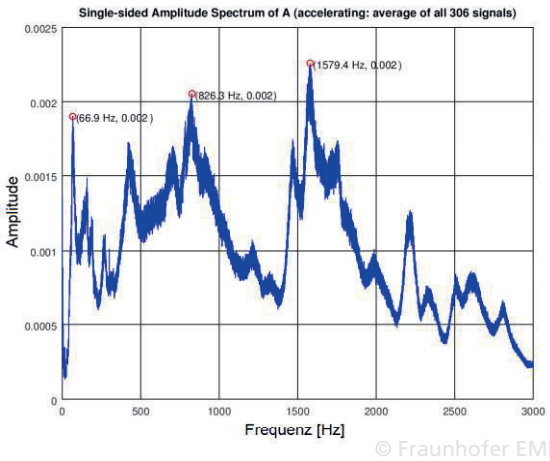


Abb. 9: Amplitudenspektrum der in den Schienen gemessenen Vibrationen beim Beschleunigen des Zuges im Berliner S-Bahntunnel

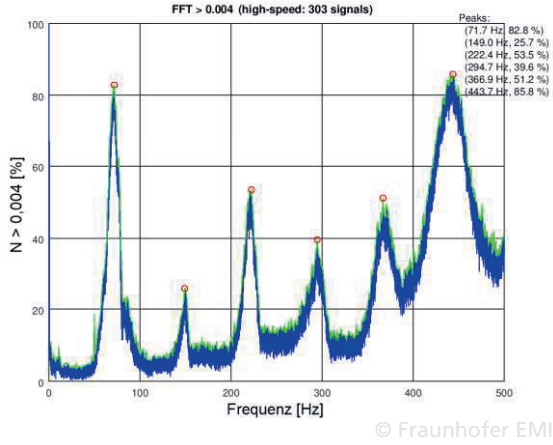


Abb. 10: Wahrscheinlichkeit in Abhängigkeit von der Frequenz, dass bei einer Zugdurchfahrt im Hochgeschwindigkeitsabschnitt die Amplitude von 0,004 im Frequenzbereich überschritten wird

Der Vibrationsharvester liefert unter optimalen Bedingungen ausreichend Energie für Knoten, welche keine fremden Nachrichten weiterleiten. Hierfür müsste das in [2] beschriebene Konzept Anwendung finden, bei welchem ausschließlich die oberen Knoten Nachrichten weiterleiten. In diesem Fall würde der erzeugten Leistung von 107 µW ein Gesamtenergieverbrauch von 59 µW gegenüberstehen.
Die erzeugte Energie in Abschnitten, in denen der Zug beschleunigt, reicht dagegen nicht aus. Hier müssen die Knoten anders versorgt werden. Da es sich hierbei meist um die Ein- und Ausfahrt des Zuges in Bahnhofsnähe handelt, wäre die Verschiebung der Senke weiter in den Tunnel hinein denkbar. Durch eine Erweiterung dieser um die Sensorik der unteren

und oberen Knoten würde die Senke den ersten Knoten im beschleunigten Bereich ersetzen.
Die Werte der RSSI Messung zeigen eine leichte Tendenz, dass die oberen Knoten einen besseren Empfang haben. Allerdings ist bei dem durchschnittlichen RSSI aller Nachrichten eine Beurteilung schwierig, da hier auch Übertragungen über mehr als einen Knoten hinaus enthalten sind, welche aufgrund der größeren Entfernung zu geringeren RSSIs führen. Besser zu beurteilen ist die Übertragung zwischen direkten Nachbarn. Hier ist die Tendenz deutlicher ausgeprägt (-71,95 dBm und -73,72 dBm zu -76,29 dBm und -77,2 dBm). Für eine ausschließliche Übertragung der Nachrichten über die oberen Knoten des in [2] beschriebenen Schemas ist dies von Vorteil. Aufgrund der großen Standardabweichung lässt sich dieser Effekt allerdings nur eingeschränkt bewerten. Außerdem lässt sich nicht bestimmen, welcher Einfluss für diesen Effekt sorgt.

Ausblick

Zur Erreichung des geringen Verbrauchs wurde eine geringe Channel-Check-Rate gewählt. Im Ereignisfall erhöht sich die Last erheblich, wodurch eine höhere Rate nötig wäre. Daher wäre es sinnvoll, eine automatische Änderung der Channel-Check-Rate im Ereignisfall zu untersuchen.
Bei der Verwendung des in [2] beschriebenen Konzepts, bei dem ausschließlich die oberen Knoten Nachrichten weiterleiten, fehlt die passive Ausfallsicherheit durch ein gleichwertiges Knotenpaar. Inwiefern bei einem Ausfall eines oberen Knotens der entsprechende untere Knoten automatisch umschalten kann, um den Ausfall auszugleichen, sollte ebenfalls untersucht werden.

Danksagung

Das diesem Bericht zugrundeliegende Forschungsprojekt “Sensor based Security and Emergency management system for underground Metro systems during disaster events” (SenSE4Metro) wurde mit Mitteln des Bundesministeriums für Bildung und Forschung durch das gemeinsame Programm „Kooperation in der zivilen Sicherheitsforschung: Kooperation zwischen Deutschland und Indien“ unter dem Förderkennzeichen 13N13039 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Literaturnachweis

[1] S. Kempf, et al. “SenSE4Metro: a bi-national multi-disciplinary project for monitoring underground metro environments in disaster events.” *11th Future Security Conference*, Berlin, 13-14 September, 2016
[2] Kempf, S., et al., „Design and Layout of an Energy Autarkic Wireless Sensor Network in Underground Metro Tunnels“, *Proceedings Sensor 2017*, 389-394, Nürnberg, 30 May – 01 June, 2017; DOI 10.5162/sensor2017/C6.3
[3] Vincke, J., et al. „A Concept for an Ultra-Low Power Sensor Network – Detecting and Monitoring Disaster Events in Underground Metro Systems“, *Proceedings of the 6th International Conference on Sensor Networks (SENSORNETS 2017)*, 150-155, Porto, 19-21 Feb. 2017; DOI: 10.5220/0006186901500155
[4] Vincke, J., et al., „Ultra-Low Power Sensor System for Disaster Event Detection in Metro Tunnel Systems“, *Sensors & Transducers 212*, Issue 5, 2017, 15-22.
[5] M. Ceriotti, M. Corrà, L. D. Orazio, R. Doriguzzi, D. Facchin, G. P. Jesi, R. L. Cigno, L. Mottola, A. L. Murphy, M. Pescalli, et al. “Is there light at the ends of the tunnel? Wireless sensor networks for adaptive lighting in road tunnels.” In *Information Processing in Sensor Networks (IPSN)*, 2011, pp. 187-198
[6] L. Mottola, G. P. Picco, M. Ceriotti, Ş. Gună, and A. L. Murphy “Not All Wireless Sensor Networks Are Created Equal: A Comparative Study on Tunnels. In *ACM Trans Sen Netw*, vol. 7, no. 2, 2010, p. 15:1–15:33
[7] Z. Sun and I. F. Akyildiz, 2008. “Not All Wireless Sensor Networks Are Created Equal: A Comparative Study on Tunnels. Channel Modeling of Wireless Networks in Tunnels.”, In *IEEE GLOBECOM 2008 - Global Telecommunications Conference*, 2008, pp. 1–5.
[8] D. Wu, L. Bao, and R. Li, „A holistic approach to wireless sensor network routing in underground tunnel environments“, In *Comput. Commun.*, vol. 33, no. 13, 2009, pp. 1566–1573
[9] H. Jiang, L. Chen, J. Wu, S. Chen, and H. Leung, „A Reliable and High-Bandwidth Multihop Wireless Sensor Network for Mine Tunnel Monitoring“, In *IEEE Sensors Journal*, vol. 9, no. 11, 2010, pp. 1511–15173
[10] M. Wischke, M. Masur and P. Woias, „Vibration harvesting in traffic tunnels to power wireless sensor nodes“, In *Smart Materials and Structures*, 2011, 20(8).
[11] U. Raza, A. Bogliolo, V. Freschi, E. Lattanzi, and A. L. Murphy, „A Two-Prong Approach to Energy-Efficient WSNs: Wake-Up Receivers plus Dedicated, Model-Based Sensing“, In arXiv:1601.06664, 2016
[12] P. J. Bennett, K. Soga, I. Wassell, P. Fidler, K. Abe, Y. Kobayashi, and M. Vanicek, „Wireless sensor networks for underground railway applications: case studies in Prague and London“ In *Smart Structures and Systems*, vol. 6, no. 5–6, 2010, pp. 619–639.
[13] Weng-Fong Cheung, Tzu-Hsuan Lin and Yu-Cheng Lin. “A Real-Time Construction Safety

Monitoring System for Hazardous Gas Integrating Wireless Sensor Network and Building Information Modeling Technologies". *Sensors 2018*, Volume 18, Issue 2

[14] Cleante, V.G. „Effects of Railway Track Vibration Induced by Passing Trains on an Energy Harvesting Device", São Paulo State University. Ilha Solteira, 2015

[15] Degrande, G., et al., "Vibrations due to a test train at variable speeds in a deep bored tunnel embedded in London clay", In *Journal of Sound and Vibration*, 293, 2006; Elsevier. https://doi.org/10.1016/j.jsv.2005.08.039

[16] Wang, P., Wang, Y, Gao, M and Wang, Y, „Energy Harvesting of Track-borne Wind Turbine by Train Induced Wind", In *Journal of Vibroengineering*, 19, 2017

[17] Contiki, 2019, "Contiki: The Open Source Operating System for the Internet of Things", At http://www.contiki-os.org/

[18] Erturk, A., Inman, D. J., "On Mechanical Modeling of Cantilevered Piezoelectric Vibration Energy Harvesters". *Journal of Intelligent Material Systems and Structures 19*, Issue 11, 1311–1325, 2008.

[19] Erturk, A., Inman, D. J., "A Distributed Parameter Electromechanical Model for Cantilevered Piezoelectric Energy Harvesters". *Journal of Vibration and Acoustics 130*, Issue 4, 41002, 2008.

[20] Erturk, A., Inman, D. J., "Piezoelectric energy harvesting". Wiley, Chichester, 2011

[21] Grittmann, J., „Modelbasierte Entwicklung und Aufbau eines Vibrations-Piezo-Harvesters im Gleisbett", *Master Thesis*, Fraunhofer Institut für Kurzzeitdynamik, Ernst-Mach-Institut, EMI, Freiburg, Deutschland, 2018.

DOI 10.5162/sensoren2019/2.3.4

REALISIERUNG EINES BISTABILEN FEUCHTESENSORSCHALTERS MIT SCHALTHYSTERESE

Nikolai Gulnizkij[1], Gerald Gerlach[1]
[1]TU Dresden, Institut für Festkörperelektronik, Mommsenstraße 15, 01062 Dresden, Deutschland

Zusammenfassung

Hydrogel-basierte Sensorschalter können zur Inline-Überwachung der Luftfeuchte in Gebäuden, Containern, der Lebensmittelindustrie etc. eingesetzt werden. Der wesentliche Unterschied leistungsloser Feuchtesensoren zu gängigen Feuchtesensoren ist, dass für den Sensorbetrieb selbst keine Betriebsspannung benötigt wird. Auf diese Weise ermöglichen solche Sensorschalter eine deutliche Vereinfachung des Aufbaus. Das Prinzip basiert darauf, dass das Schalten des Mikroschalters beim Überschreiten eines Schwellwertes durch die Luftfeuchte selbst ausgelöst wird, so dass dafür keine elektrische Energie benötigt wird. Dies erfolgt durch einen Bimorphbalken mit einer Hydrogelschicht, die durch Änderungen der Luftfeuchtigkeit quillt bzw. entquillt. Um ein sicheres Schalten zu gewährleisten, muss die mechanische Struktur eine Schalthysterese aufweisen. Dabei kann die Hysterese durch ein belastungsabhängiges Ausbeulen einer axial belasteten Struktur hervorgerufen werden. In diesem Beitrag soll ein solcher Hydrogel-basierter Sensorschalter mit Schalthysterese für die Feuchte vorgestellt werden.

Keywords: Hydrogel, Hysterese, Mikroschalter, leistungslos, Bimorph-Effekt

Motivation

Die wachsende informationstechnische Vernetzung in der industriellen Produktion führt zu einer zunehmenden Nachfrage nach Sensorsystemen, die die Messwerterfassung, die Signalaufbereitung und die Signalverarbeitung vereinen. Solche Sensorsysteme erzielten in der Sensor- und Messtechnikindustrie in den vergangenen Jahrzehnten ein überdurchschnittliches Wachstum von über 6% pro Jahr [1]. Dabei sind die wichtigsten Messgrößen die Temperatur und die Feuchtigkeit. Allerdings werden 70% aller Sensoren in der Prozesskontrolle und mehr als 90% in der Gebäudeüberwachung als Schwellwertschalter eingesetzt. In [2] wurde ein neuartiger Ansatz für einen leistungslosen Feuchteschwellwertschalter vorgeschlagen, der auf dem BIZEPS-Prinzip (binary zero-power sensor) basiert [3]. Bei diesem Prinzip wird die Schaltenergie direkt aus der Messgröße entnommen. In dem hier vorgestellten Beispiel wird dazu das feuchteabhängige Quellen eines Hydrogels genutzt, das Teil einer Bimorph-Biegeplatte ist. In dieser Arbeit soll ein vereinfachtes Sensorkonzept zur Beschreibung des Schaltvorgangs entwickelt und entsprechend ein Sensorschalter mit einer Schalthysterese realisiert werden.

Funktionsweise

Abb. 1 zeigt beispielhaft den Aufbau und das Funktionsprinzip eines solchen Feuchteschwellwertschalters [3][4].

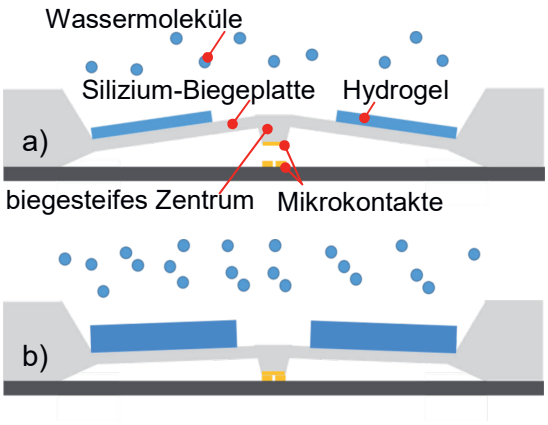


Abb. 1: Sensorschalter zur Feuchtemessung, bestehend aus einer Bimorph-Biegeplatte, einem biegesteifen Zentrum, einer strukturierten Hydrogelschicht und Mikrokontakten, a) geöffnet und b) geschlossen.

Er besitzt eine Biegeplatte, in deren Mitte sich ein biegesteifes Zentrum, das zur lokalen Versteifung der Biegeplatte führt, befindet. Es gehört zum schließenden bzw. zum öffnenden

Kontaktpaar und stellt sicher, dass der Mikrokontakt auch bei Auslenkung der Biegeplatte eben bleibt und damit eine sichere Kontaktierung der Mikrokontakte erfolgt. Das biegesteife Zentrum der Biegeplatte wird durch die feuchteabhängige Quellung des Hydrogels auf der Hydrogel-beschichteten Siliziumplatte ausgelenkt. Der Quellzustand des Hydrogels hängt von der Absorption der Wassermoleküle aus der Luft, d.h. der Luftfeuchte der Umgebungsluft, ab.

Beim Quellen leistet das Hydrogel mechanische Arbeit und deformiert so die Biegeplatte. Das bedeutet, dass die kinetische Energie aus der Umgebung entnommen und der Sensor damit autonom und ohne elektrische Energieversorgung betrieben wird. Da es sich bei der Quellung und Entquellung des Hydrogels um einen reversiblen Prozess handelt, kann der Sensorschalter die Mikrokontakte sowohl öffnen als auch schließen.

Um ein sicheres Schalten des elektrischen Mikrokontakts zu gewährleisten, müssen ein oszillierendes Öffnen bzw. Schließen der Kontakte im Schaltpunkt und die Ausbildung von Lichtbögen vermieden werden. Dies kann durch eine Schalthysterese im Sensorschalter, der zwischen zwei stabilen Zuständen schaltet, erreicht werden [5]. In der hier vorgestellten Lösung wird dazu als Verformungskörper ein Biegebalken verwendet, auf den eine axiale Druckkraft aufgeprägt wird. (2)

Modell zur Berechnung der Balkendurchbiegung in Abhängigkeit vom Beschichtungsgrad

Für das bessere physikalische Verständnis soll im Folgenden statt der Platte von Abb. 1 ein zweiseitig eingespannter Biegebalken betrachtete werden. Zur Berechnung der Durchbiegung des Biegebalkens in Abhängigkeit vom Beschichtungsgrad (Verhältnis der beschichteten zur gesamten Fläche) wird ein einfaches mechanisches Modell verwendet, das auf der Balkentheorie für monomorphe und bimorphe Balken basiert [6][7]. Abbildung 2 zeigt einen beidseitig eingespannten, am Balkenrand mit Hydrogel beschichteten Balkens. Der Balken aus Stahlblech besteht aus dem mit Hydrogel beschichteten Teil mit der längenabhängigen Nachgiebigkeit n_B und dem unbeschichteten Teil mit der ebenfalls längenabhängigen Nachgiebigkeit n_P. Dabei entspricht l_P der Länge des Biegebalkens, h der Balkendicke und b der Balkenbreite. Da das Hydrogel sehr weich ist, wird die jeweilige Nachgiebigkeit durch die Elastizität des Blechs bestimmt:

$$n_B = 12\,\frac{l_B}{Ebh^3}\,, \qquad (1)$$

$$n_P = 12\,\frac{l_P}{Ebh^3}\,. \qquad (2)$$

Der Bimorph aus Hydrogel und Blech wirkt bei Quellung des Hydrogels als Momentenquelle:

$$M_{r.F.} = \varepsilon_{r.F.}\,\frac{bE_2h_2(h_1+h_2)}{2(1+\frac{E_2h_2}{E_1h_1})} \qquad (3)$$

Dabei ist $\varepsilon_{r.F.}$ die feuchteabhängige Dehnung des Hydrogels.

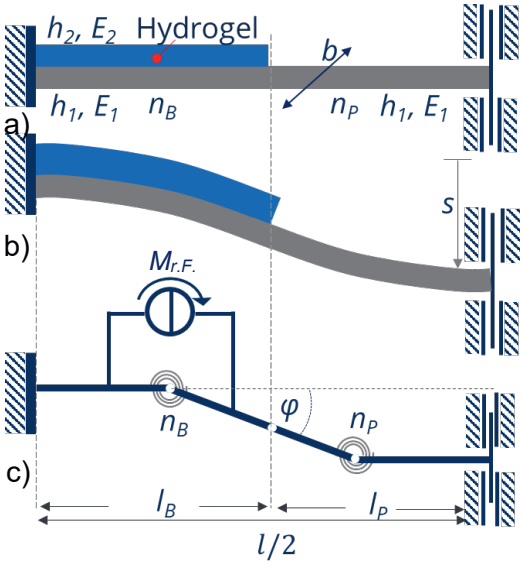


Abb. 2: Modell eines balkenförmigen Sensorschalters gemäß Abb. 1. a) schematische Darstellung des halben Sensorschalters, b) ausgelenkter Sensorschalter bei Hydrogelquellung und c) korrespondierendes mechanisches Modell. n_B, n_P längenabhängige Nachgiebigkeiten der Biegebalken mit und ohne Hydrogelschicht, l/2 Balkenlänge, l_B Länge des mit Hydrogel beschichteten Balkens, s Auslenkung, φ Auslenkungswinkel, $M_{r.F.}$ feuchteinduzierte Momentenquelle, b Breite, h Dicke, Indizes: 1 Substrat, 2 Hydrogel, P unbeschichtet, B Hydrogel bzw. mit Hydrogel beschichtet.

Mit der Annahme, dass die Hydrogelschicht und der Biegebalken vergleichbar dick sind ($h_1 \approx h_2$), ergibt sich aus Gl. (3):

$$M_{r.F.} = bE_2h_2\varepsilon_{r.F.} \qquad (4)$$

und für die Auslenkung des Biegebalkens:

$$s = \frac{l}{2}\varphi = \frac{l}{2} n M_{r.F.}, \qquad (5)$$

wobei die Gesamtnachgiebigkeit

$$n = \frac{n_B n_P}{n_B + n_P} \qquad (6)$$

ist.
Mit den Gleichungen (1) und (2) folgt für die Auslenkung:

$$s = 6\frac{E_2 l_B}{E_1 h}\frac{l_B}{l}\left(1 - \frac{l_B}{l}\right)\varepsilon_{r.F.}. \qquad (7)$$

Damit hängt die Auslenkung maßgeblich vom E-Modul E_2 des Hydrogels, von den geometrischen Abmessungen der Hydrogelschicht, von der feuchteabhängigen Quelldehnung $\varepsilon_{r.F.}$ und vom Beschichtungsgrad l_B/l ab.
Aus Gleichung (7) folgt, dass eine große Durchbiegung

- durch ein großes Verhältnis E_2/E_1 der E-Moduln von Hydrogel und Substrat (Biegebalken: Stahl),
- durch dünne Biegebalken und
- durch große Quelldehnungen $\varepsilon_{r.F.}$ bzw. Ausdehnungskoeffizienten $\alpha_{r.F.}$ in Folge der relativen Feuchte

erreicht werde kann.
Eine maximale Durchbiegung s_{max} wird für dieses vereinfachte Modell erzielt, wenn das Verhältnis l_B/l dem Wert 0,5 entspricht. Aus den Gleichungen (6) und (7) kann für komplizierte Anordnungen aber auch abgeleitet werden, dass die Auslenkung dann maximal wird, wenn die Nachgiebigkeiten n_B und n_P des beschichteten und unbeschichteten Teils gleich groß sind.

Balkenmodell mit Axialkraft

Für technische Anwendungen muss der Schaltvorgang zum Öffnen und Schließen von Mikrokontakten sehr schnell erfolgen. Dies wird hier durch eine axiale Krafteinwirkung, die zu einer Vorauslenkung des Balkens führt, realisiert. Abbildung 3 zeigt das mechanische Modell eines vorausgelenkten Balkens mit axialer Krafteinwirkung. Der Vorauslenkungswinkel φ_0 hängt hierbei von der längenabhängigen Nachgiebigkeit n aus Gl. (6) und dem Betrag der Kraft F_A ab.

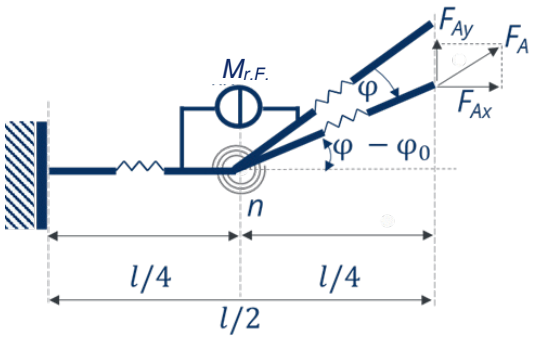


Abb. 3: Mechanisches Modell zur Beschreibung eines durch die Axialkraft F_A vorausgelenkten Balkens mit der Länge l/2, der Nachgiebigkeit n und dem Auslenkungswinkel φ_0. Dargestellt ist eine Balkenhälfte. Das Biegemoment $M_{r.F.}$ entspricht hierbei einer externen Momentenquelle. F_{Ax} und F_{Ay} sind die vektoriellen Komponenten von F_A.

Die Kraft F_A wird hierbei vektoriell in zwei Kraftkomponenten F_{Ax} und F_{Ay} zerlegt. Die Auslenkung des Balkens durch die externe Momentenquelle wird durch den Auslenkungswinkel φ und die Gesamtauslenkung durch den Winkel $\varphi = \varphi + \varphi_0$ beschrieben. Damit lässt sich für kleine Auslenkungswinkel die Momentengleichung herleiten:

$$M_{r.F.} - F_A\frac{l}{4}(\varphi - \varphi_0) + F_A(\varphi - \varphi_0)^3 = 0, \qquad (8)$$

sodass die Auslenkung des Balkens von der Balkengeometrie, den Materialparametern und der einwirkenden Axialkraft abhängt. In Gl. (8) weist die Beziehung zwischen Winkel φ und damit Auslenkung s am rechten Balkenende in Abb. 3 und dem zum Hyrogelquellen proportionalen Biegemoment $M_{r.F.}$ neben dem linearen Glied auch ein kubisches Glied auf. Dieses führt zu der in Abb. 4 dargestellten S-Kurve. Steigt die Feuchte an und übersteigt dabei $\varphi_{r.F.oben}$, ändert sich die Auslenkung plötzlich von s_1 zu s_2. Sinkt die Feuchte soweit, dass sie $\varphi_{r.F.unten}$ unterschreitet, ändert sich die Auslenkung plötzlich von s_3 auf s_4 und der Schalter öffnet sich.

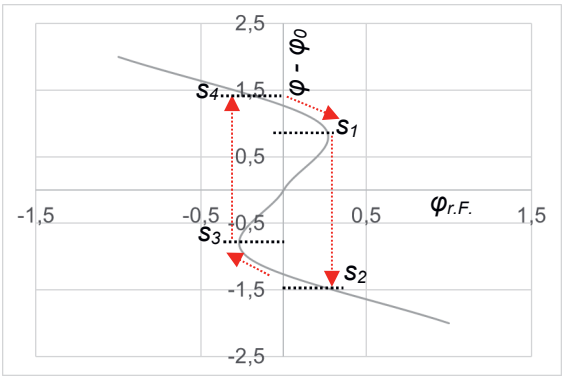


Abb. 4: Berechnete S-Kurve aus Gl. (8). Die Schaltpunkte sind mit den vertikal verlaufenden Pfeilen gekennzeichnet.

Aufbau und Herstellung

Um die Funktionsweise des hier vorgestellten Feuchtesensorschalters zu demonstrieren, wird der Aufbau von Abb. 5 verwendet, bei dem auf den Balken (n_P) eine axiale Kraft F_A aufgeprägt und damit die Schalthysterese erreicht wird.

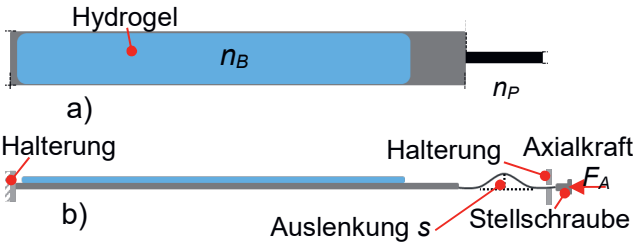


Abb. 5: Balkenbasierter Sensorschalter mit Schalthysterese: a) Draufsicht, b) Querschnitt.

Der Balken besteht aus einem breiten und dickeren, Hydrogel-beschichteten Teil mit der Nachgiebigkeit n_B und einem schmaleren Teil mit der Nachgiebigkeit n_P (Material- und Geometriedaten siehe Tabelle 1).

Tab. 1: Geometrie- und Materialparameter.

Parameter	Blech n_B | n_P	Hydrogel
Länge [mm]	55 | 7	50
Breite [mm]	10 | 1	10
Dicke [mm]	0,03 | 0,01	0,03
E-Modul [N/mm²]	210000	3500 (5%r.F.) 300 (90%r.F.)

Der Balken aus Stahlblech (DC01) wird dabei durch Zuschneiden gefertigt.
Für eine ausreichende Haftung der 0,03 mm dicken Hydrogelschicht wurde die Blechoberfläche chemisch gereinigt, plasmaaktiviert und mit einem Haftvermittler vorbehandelt. Zu Beginn wurde der Biegebalken für 15 Minuten in einer Lösung aus destilliertem H_2O, 25% Ammoniak und 25% Peroxid (2:1:1) bei 85 °C gereinigt. Danach wurde die Oberfläche mit Sauerstoffplasma (VTD: DREVA 450) bei einem Druck von 15 Pa bei 400 W für 5 min aktiviert. Anschließend wurden 3 ml einer Haftvermittlerlösung, bestehend aus 0,9 ml Haftvermittler (3-Aminopropyl)Triethoxysilan (APTES) (Carl Roth) und 99 ml H_2O auf die plasmabehandelte Oberfläche aufgebracht und 10 Minuten im Ofen bei 100 °C geheizt [8]. Für die Hydrogelstrukturierung wurden 5 ml einer Polymerlösung aus 15 wt% Polyvinylalkohol (PVA, Molekularmasse *Mw* = 89,000 … 98,000) und 7.5 wt% Polyacrylsäure (PAS, *Mw* = 450,000) in einem Massenverhältnis von 4:1 auf die vorbehandelte Oberfläche pipettiert und für 6 h bei 22 °C getrocknet. Anschließend wurde die strukturierte Polymerschicht in einem Ofen bei 120 °C für 20 Minuten thermisch vernetzt [9] [10].
Beide Biegebalken n_B und n_P wurden mit Sekundenkleber miteinander verklebt. Die Blechstruktur wurde in einer Halterung fixiert und anschließend mit einer Stellschraube vorausgelenkt (Abb. 5).

Messung der Schalthysterese am makroskopischen Sensorschalter

Die Vorauslenkung wurde durch einen axialen Krafteintrag von 0,02 N mittels Verstellen einer Stellschraube erreicht. Zur Messung der Axialkraft wurde auf der gegenüberliegenden Seite der Befestigung ein Kraftmessgerät (SAUTER FC 10, FC Compact) eingesetzt.
Die Feuchte in der Feuchtekammer wurde dabei mittels eines Blasen-Systems variiert. Die Änderung der relativen Feuchte wurde mittels eines kapazitiven Feuchtesensors überwacht. Zum Nachweis der feuchteabhängigen Auslenkung der eingespannten Blechstruktur wurde eine CMOS-Kamera (HIKVISON USB3.0, Vision SXGA, monochrom) verwendet (Abb. 6).

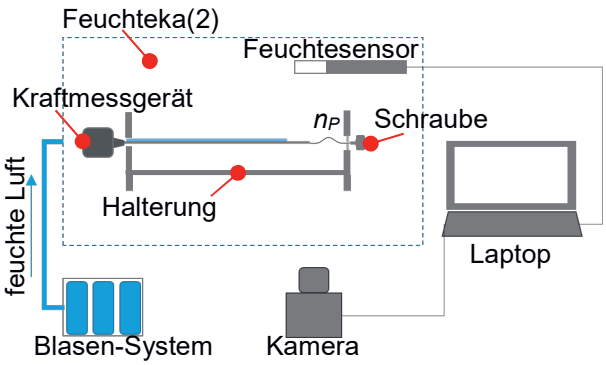


Abb. 6: Schematische Darstellung des Messaufbaus, bestehend aus Feuchtekammer, Feuchtesensor, Blasen-system, Kamera und Sensorschalter.

Messergebnisse und Diskussion

Nachdem das Hydrogel auf einer Biegebalkenhälfte strukturiert wurde, wurden n_B und n_P am Balkenrand miteinander verklebt und in einer speziell für diese Anwendung hergestellten Halterung fixiert. Anschließend wurde durch das Verstellen der Stellschraube die Balkenstruktur einseitig axial belastet. Auf der gegenüberliegenden Seite der Stellschraube wurde die Axialkraft durch ein Kraftmessgerät überwacht. Nach Erreichen der Axialkraft von 0,02 N wurde die feuchteabhängige Auslenkung des Balkens bestimmt. Abbildung 7 zeigt die Ergebnisse der Messung der feuchteabhängigen Auslenkung mit der resultierenden Schalthysterese und ihren zwei Schaltpunkten bei 58 %r.F. und bei 75 %r.F.

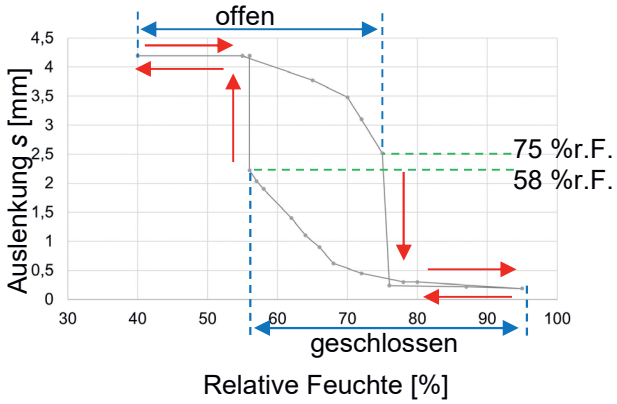


Abb. 7: Auslenkung s der Biegefeder n_P in Abhängigkeit von der relativen Feuchte mit den zwei Schaltpunkten bei 58 %r.F. und 75 %r.F.

Zusammenfassung

Zur Überwachung der Feuchte wurde ein leistungsloser Sensorschalter mit Hydrogel entwickelt. Es besteht aus einem Biegebalken, der teilweise mit Hydrogel beschichtet ist und der axial mit einer Druckkraft beaufschlagt wird, um die gewünschte Schalthysterese zu erreichen. Mit einem einfachen Mono-/Bimorph-Balkenmodell wurde gezeigt, dass die erreichbare Auslenkung maximal wird, wenn die Nachgiebigkeit des beschichteten und des unbeschichteten Balkenteils gleich groß sind bzw. der Bedeckungsgrad etwa 50% beträgt. Mit der Axialkraft wird eine Vorauslenkung erzielt, sodass beim Quellen des Hydrogels ein Durchschlagen des Balkens und damit ein schnelles Schalten erreicht wird. Die Schalthysterese resultiert dabei aus der S-Kurve, die durch den kubischen Funktionsanteil der Auslenkungs-Feuchtekurve verursacht wird. Anhand eines Demonstrators konnte die prinzipielle Funktionsweise der Schalthysterese nachgewiesen werden.

Literaturnachweis

[1] AMA: Sector Information: Quarterreports 2017, http://www.ama-sensorik.de/en/association/sector-information/quarterreports-2017/ – (data from 08/14/2018)

[2] C. Bellmann, R. Sarwar, A. Steinke, T. Frank, H. F. Schlaak and G. Gerlach, *Development of a humidity micro switch based on humidity sensitive polymers*, 23rd Micromechanics and Microsystems Europe Workshop (MME), Ilmenau, Germany, (2012)

[3] T. Frank, A. Steinke, and G. Gerlach, *Binary ZeroPower Sensors: an alternative solution for power-free energy-autonomous sensor systems*, Microsystem Technologies 18(7-8), 1225-1231 (2012)

[4] C. Bellmann, *Leistungslose Feuchteschwellwertschalter*, Dissertation, TU Dresden, (2016)

[5] L.T. Chen, C. Y. Lee, W. H. Cheng, *MEMS-based humidity sensor with integrated temperature compesation mechanism*, Sensors and Actors A:147, 522-528, (2008)

[6] U. Marschner, G. Gerlach, E. Starke, A. Lenk, *Equivalent circuit models of two-layer flexure beams with excitation by temperature, humidity, pressure, piezoelectric or piezomagnetic interactions*, Vol.3, Page 187-211, JSSS (2014)

[7] H. Balke, *Einführung in die Technische Mechanik*, 3. Auflage, Springer-Verlag Berlin Heidelberg 2010

[8] N. Gartmann, C. Schütze, H. Ritter, D. Brühwiler, *The effect of water on the functionalization of mesoporous silica with 3-aminopropyltriethoxysilane*, Journal of Physical Chemistry Letters 1, 379-382, (2010)

[9] J. Jose, F. Shehzad, M. A. Al-Harthi, *Preparation method and physical, mechanical, thermal characterization of poly(vinyl alcohol)/poly(acrylic acid) blends*, Polymer Bulletin 71(11), 2787-2802, (2014)

[10] K. F. Arndt, A. Richter, S. Ludwig, J. Zimmermann, J. Kessler, D. Kuckling and H. J. Adler, *Poly (vinyl alcohol)/poly (acrylic acid) hydrogels, FT-IR spectroscopic characterization of crosslinking reaction and work at transition point*, Acta Polymerica 50(11-12), 383– 390, (1999)

DOI 10.5162/sensoren2019/2.4.2

Berührungslose Temperaturmessung an Verbrennungsgasen bei hohen Temperaturen und hohen Drücken

Matthias Zipf[1], Jochen Manara[1], Thomas Stark[1], Mariacarla Arduini[1], Hans-Peter Ebert[1], Jürgen Hartmann[2]
[1]Bayerisches Zentrum für angewandte Energieforschung e.V., Magdalene-Schoch-Straße 3, 97074 Würzburg, Deutschland
[2]Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Ignaz-Schön-Straße 11, 97421 Schweinfurt, Deutschland

Zusammenfassung

Stationäre Gasturbinen sind von großer Bedeutung für die heutige Energieversorgung. Der Wirkungsgrad einer Gasturbine steigt mit zunehmender Heißgastemperatur an. Turbinenhersteller bzw. Kraftwerksbetreiber sind daher bestrebt, Turbinen bei der höchsten materialtechnisch möglichen Heißgastemperatur einzusetzen. Eine entsprechende Prozessoptimierung des Turbinenbetriebs setzt somit die exakte Kenntnis der Gastemperaturen während des Betriebs und damit eine verlässliche Messung derselben voraus.
Zur Messung der Gastemperatur werden derzeit in der Regel Thermoelemente unmittelbar im Abgasstrom platziert. Aufgrund der dort vorherrschenden extremen Bedingungen degradieren diese Sensoren allerdings sehr schnell. Ein alternativer Ansatz sieht die Entwicklung eines berührungslosen Messverfahrens auf der Grundlage von Strahlungsthermometern vor.
Für die Umsetzung dieses Vorhabens ist die genaue Kenntnis des Verhaltens der infrarot-optischen Emissions- und Transmissionsspektren der im Abgasstrom enthaltenen Gase bei hohen Temperaturen und Drücken eine wesentliche Voraussetzung. Aus diesem Grund wurde am ZAE Bayern eine Hochtemperatur-Hochdruck-Gaszelle entwickelt, die es in Verbindung mit einem FTIR-Spektrometer erlaubt, Gase und Gasgemische hinsichtlich dieser Gesichtspunkte zu charakterisieren.
In dieser Arbeit wird die neue Messapparatur vorgestellt und Gasgemische, die für die Turbinenanwendungen relevant sind, werden analysiert. Zur Identifizierung eines geeigneten Wellenlängenbereichs für die geplante berührungslose Temperaturmessung wurden erste Messungen durchgeführt, auf deren Grundlage ein adäquater Wellenlängenbereich ermittelt werden konnte.

Keywords: Gas-Transmissions-Spektren, stationäre Gasturbinen, Strahlungs-Thermometrie, berührungslose Gas-Temperaturmessung, FTIR-Spektrometrie.

Einführung

Eine der dringendsten Aufgaben, mit der wir derzeit konfrontiert sind, ist die Weiterentwicklung von Techniken und Methoden zur Elektrizitätsversorgung. Noch immer wird der größte Teil des weltweiten Bedarfs an Elektrizität aus fossilen Energieträgern – allen voran Kohle – gedeckt [1]. Unter ökologischen Gesichtspunkten sind diese jedoch problematisch, da z. B. Kohle-Kraftwerke im Zuge des Energieumwandlungsprozesses große Mengen an Kohlenstoffdioxid ausstoßen, welches zur Beschleunigung des Treibhauseffektes und somit des Klimawandels beiträgt. Daher müssen diese Energieträger im Zuge der Energiewende vollständig durch erneuerbare und kohlenstoffdioxidarme Energieträger wie Solar- und Windenergie ersetzt werden. Doch bisher sind die Entwicklung und der Ausbau der erneuerbaren Energieträger noch nicht weit genug fortgeschritten, um die konventionellen Energieträger bereits vollständig ersetzen zu können.
Verglichen mit Kohlekraftwerken erreichen moderne GuD-Kraftwerke einen deutlich höheren Wirkungsgrad [2], [3]. Außerdem entspricht der Ausstoß von Kohlenstoffdioxid je abgegebener Kilowattstunde an elektrischer Energie bei Gas-Kraftwerken nur etwa der Hälfte des CO_2-Ausstoßes von Kohlekraftwerken [1]. Bis kohlenstoffbasierte Energieträger vollständig ersetzt werden können, können Gasturbienen daher als Brückentechnologie dienen, und den weltweiten CO_2-Ausstoß bereits stark reduzieren.
Darüber hinaus haben Gasturbinen kurze Anlaufzeiten von nur wenigen Minuten und sind somit flexibel einsetzbar [4]. Daher können Gasturbinen auch bei einem weiteren Ausbau der erneuerbaren Energien dazu beitragen Bedarfslücken zwischen der Stromnachfrage und der schwankenden Erzeugung zu schließen,

die beispielsweise bei Windstille und gleichzeitig geringer solarer Einstrahlung entstehen können.
Eine Gasturbine stellt eine Realisierung des Brayton-Prozesses (auch Joule-Prozess) dar [4]. Der Wirkungsgrad einer Gasturbine kann somit erhöht werden, indem die Differenz zwischen den Gastemperaturen am Turbineneinund -auslass erhöht wird. In der Praxis bedeutet dies die Erhöhung der Verbrennungsgastemperatur [5].
Jedoch setzen die Materialparameter dieser Erhöhung Grenzen. Eine Überschreitung der materialtechnisch vorgegebenen Maximaltemperaturen reduziert die Lebensdauer der Turbinenkomponenten – vor allem der Turbinenschaufeln – erheblich. Im Extremfall können diese zerstört werden [4].
Daher besteht die Notwendigkeit einer verlässlichen Messung der Temperatur der Turbinenkomponenten und des Brenngases am Turbineneingang im Betrieb. Wegen der Rotation der Turbine im Betrieb ist ein berührungsloses Verfahren zur Messung der Bauteiltemperatur von Vorteil, wie es in [6] und [7] vorgestellt wird.
Die heute gebräuchliche Methode zur Messung der Brenngastemperatur besteht in der Platzierung von Thermoelementen direkt im Abgasstrom. Thermoelemente bieten hohe Messgenauigkeiten und Auflösungen. Allerdings stellt sich der wirkliche Messwert stets erst mit einem gewissen Verzug ein, was in Turbinenbetriebsarten mit schnellen Temperaturwechseln ein Nachteil ist. Außerdem reduziert sich die Lebenszeit von Thermoelementen bei hohen Temperaturen erheblich. Bei den extremen Bedingungen im Turbineninneren können daher Teile der Thermoelemente abbrechen und in der Turbine Schaden verursachen. Deshalb wird bisher die Gastemperatur am Turbineneingang in vielen Fällen nicht gemessen [4], [8].

Ein neuer Ansatz mittels Strahlungsthermometrie

Die Entwicklung einer berührungslosen Messtechnik stellt einen neuen Ansatz zur Gastemperaturmessung dar, welcher es unter Verwendung von Infrarotstrahlung ermöglicht, die Brenngastemperatur auch am Turbineneinlass zu messen. Mittels Pyrometrie soll dabei die vom heißen Gas emittierte Strahlung detektiert werden. Das Pyrometer selbst muss dazu nicht im Turbineninneren angebracht und daher nicht den dortigen Bedingungen ausgesetzt werden. Für die Temperaturmessung genügt ein infrarottransparentes Fenster im Turbinengehäuse, das die IR-Strahlung aus dem Inneren zum Pyrometer leitet.
Zur Entwicklung eines Pyrometers für diesen Zweck muss zuerst ein für die Brenngase passender Wellenlängenbereich bestimmt werden. Das Brenngasgemisch einer Gasturbine besteht im Allgemeinen hauptsächlich aus den in Tab. 1 genannten Gasarten. Sauerstoff, Stickstoff, Kohlenstoffdioxid und Wasserdampf [4]. Stickstoff und Sauerstoff bilden davon wiederum die größten Anteile. Als zweiatomige Moleküle aus zwei gleichartigen Atomen besitzen sie allerdings kein anregbares Dipolmoment und sind somit infrarotinaktiv, weshalb sie nicht nutzbar für eine Gastemperaturmessung mittels IR-Strahlung sind [9]. Wasserdampf und Kohlenstoffdioxid sind dagegen infrarotaktiv und somit prinzipiell für die geplanten Temperaturmessungen nutzbar [10].

Tab. 1: Typische Volumen-Anteile der Hauptbestandteile von Turbinen-Abgasen [4].

Gas	Volumen-Anteil / %
O_2	13
N_2	74
CO_2	4
H_2O	8

Bei der Gastemperaturmessung im Brenngasstrom einer Gasturbine kann es durch die Temperaturunterschiede zwischen dem sehr heißen Turbineninneren und dem gekühlten Gehäuse zur Bildung von Gasschichten unterschiedlicher Temperaturen kommen [4]. Eine Gasschicht mit niedriger Temperatur im Randbereich des Gasstroms kann das Messsignal erheblich beeinflussen [7]. Um diesem Problem zu begegnen und verlässliche Messergebnisse zu erhalten, ist es nötig, Temperaturinformationen aus verschiedenen Gasschichten im Gasstrom zu sammeln.
Abb. 1 zeigt das typische Verhalten einer Absorptionsbande eines infrarotaktiven Gases in Abhängigkeit von der Temperatur. Die Tiefe der Bande (die mit der Höhe des Absorptionsgrades korreliert) steigt mit steigendem Druck. Bei einem konstanten Druck, bei welchem bereits Sättigung vorliegt, steigt die Breite der Bande mit steigender Temperatur in Folge der Dopplerverbreiterung [11]. Aufgrund dieser Umstände können drei Wellenlängenintervalle im Bereich einer Absorptionsbande benannt werden, wo eine Temperaturmessung jeweils an einer bestimmten Position im Gasstrom möglich ist.
Im Zentrum einer Absorptionsbande (vgl. Abb. 1, grüne Fläche) geht der Transmissionsgrad gegen Null. Strahlung aus dem Turbineninneren wird in dieser Wellenlängenregion vollständig absorbiert, ehe sie einen Sensor außerhalb des Gasstroms erreicht. Im Wellenbereich des Zentrums einer Absorptionsbande kann so-

mit die Gastemperatur am Rand des Gasstroms bestimmt werden, da in diesem Wellenlängenbereich nur von dort Strahlung den Sensor erreicht.
Am Rand der Bande (vgl. Abb. 1, blaue Flächen) hängt der Transmissionsgrad von der Gastemperatur ab. Da somit auch ein gewisser Anteil der Strahlung aus dem Turbineninneren die äußeren Gasschichten passiert und zum Sensor gelangt, können in diesem Wellenlängenbereich Aussagen über die Gastemperatur im Inneren getroffen werden. Abseits einer Absorptionsbande (vgl. Abb. 1, purpurne Flächen) geht der Transmissionsgrad gegen eins – unabhängig von der Gastemperatur. Hier bleibt die Strahlung unbeeinflusst vom Brenngasstrom, wodurch hier Temperaturmessungen an Bauteilen im Turbineninneren möglich sind.

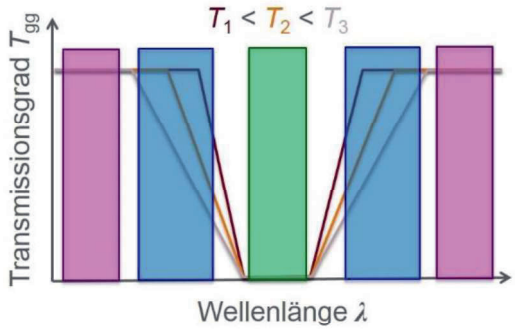


Abb. 1: Schematische Darstellung der Temperaturabhängigkeit des Verlaufs des spektralen Transmissionsgrades im Wellenlängenbereich einer Gas-Absorptionsbande bei konstantem Druck. Markierte Bereiche: 1) Zentrum der Bande, Absorption befindet sich in der Sättigung (grün); 2) Rand der Bande, Transmission ist temperaturabhängig (blau); 3) Region abseits der Bande, keine Absorption vorhanden (purpurn).

Die Positionen der Absorptionsbanden von Kohlenstoffdioxid und Wasserdampf sind seit langem und in hoher Genauigkeit aus der Literatur bekannt, z. B. [12], [13], [14], [15]. Auch können entsprechende Spektren simuliert werden [16]. Abb. 2 zeigt die HITRAN-Simulation des spektralen Transmissionsgrades eines typischen Brenngasgemisches einer stationären Gasturbine (vgl. Tab. 1) mit den entsprechenden Molekülschwingungen (vgl. Tab. 2). Jedoch muss das Verhalten dieser Banden unter Bedingungen wie im Inneren einer Gasturbine noch weitergehend experimentell untersucht werden. Das heißt, die Absorptionsspektren von Kohlenstoffdioxid und Wasserdampf müssen unter entsprechenden Bedingungen verlässlich gemessen werden.

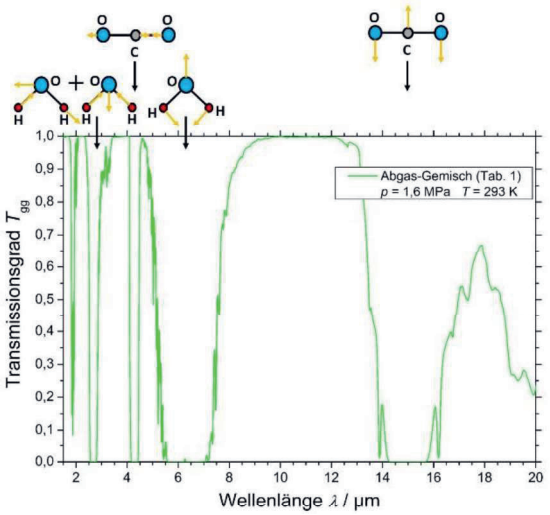


Abb. 2: HITRAN-Simulation [16] des spektralen Transmissionsgrades des Abgas-Gemisches aus Tab. 1 bei Raumtemperatur bei einem Druck von 1600 kPa im Wellenlängenbereich von 1,5 µm bis 20 µm. Den Absorptionsbanden sind die entsprechenden Molekülschwingungen zugeordnet.

Tab. 2: Erwartete Positionen der Molekülschwingungen von CO_2 und H_2O im untersuchten Wellenlängenbereich [10], [17], [18].

Gas	Bande	Position in µm	Bemerkung
CO_2	ν_1	7,78/7,20	IR-inaktiv
	ν_2	14,99	IR-aktiv
	ν_3	4,26	IR-aktiv
H_2O	ν_1	2,73	IR-aktiv
	ν_2	6,27	IR-aktiv
	ν_3	2,66	IR-aktiv

In dieser Arbeit wurde als ersten Schritt das Transmissionsspektrum von Kohlenstoffdioxid im mittleren Infrarotbereich hochaufgelöst unter Turbinenbedingungen gemessen, um so eine Grundlage für die Wahl einer Messwellenlänge für die Gastemperaturmessung zu schaffen.

Messaufbau

In Abb. 3 ist schematisch der hier zur Bestimmung der Gastransmissionsspektren verwendete Messaufbau dargestellt, der aus einem Fourier-Transformations-Infrarotspektrometer (FTIR) und einer Hochtemperatur-Hochdruck-Gaszelle besteht. Im FTIR-Spektrometer wird von einem Globar Infrarotstrahlung emittiert und in ein Michelson-Interferometer geleitet. Dort wird der Strahl geteilt. Ein Teilstrahl fällt auf einen festen, der zweite auf einen beweglichen

Spiegel, wodurch sie jeweils zurück zum Strahlteiler reflektiert werden, wo sie miteinander interferieren. Das resultierende Interferenzbild hängt von der Weglängendifferenz der Teilstrahlen ab [19]. Das Interferenzsignal wird von dort aus durch die Probenkammer und zum Detektor geleitet.
Der Detektor nimmt das Interferenzsignal in Abhängigkeit der momentanen Position des beweglichen Spiegels auf. Das Spektrometer ermittelt aus diesen Daten ein Interferogramm, welches dann durch eine Fourier-Transformation in den Frequenzraum konvertiert wird, woraus die spektrale Intensität resultiert. Durch dieses Verfahren wird zudem jegliche Störstrahlung (auch die Wärmeabstrahlung des Heißgases selbst), die zwischen Interferometer und Detektor in den Strahlengang eingekoppelt wird, eliminiert.

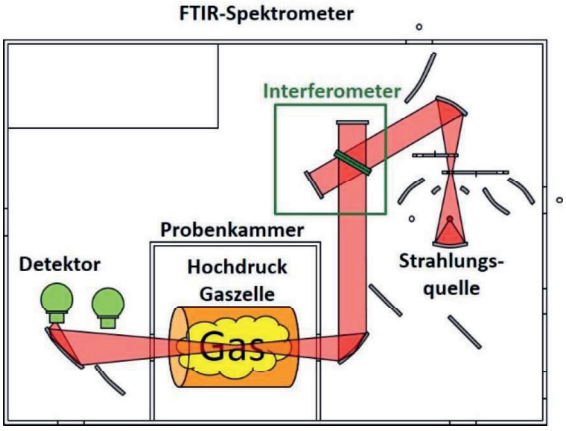


Abb. 3: Schematische Darstellung des Messaufbaus zur Bestimmung des Transmissionsgrades von heißen Gasen bestehend aus FTIR-Spektrometer und Hochdruck-Gaszelle

Zur Bestimmung des spektralen gerichtet-gerichteten Transmissionsgrades wird zunächst als Referenzmessung die spektrale Intensität bei leerem Strahlengang gemessen. Danach wird die zu analysierende Probe bzw. Gasmischung in der Probenkammer platziert und die spektrale Intensität erneut gemessen. Der spektrale gerichtet-gerichtete Transmissionsgrad wird nun als Quotient dieser beiden Einzelmessungen berechnet.
Um das in dieser Arbeit untersuchte gasförmige Kohlenstoffdioxid im Strahlengang zu platzieren, wird eine eigens entwickelte Gaszelle (vgl. Abb. 4) verwendet, welche eine den Turbinenanwendungen ähnliche geometrische Weglänge der Strahlung im Gas von 16 cm aufweist. Die erzielten Ergebnisse können daher gut übertragen werden.
Die Gaszelle kann in der Spektrometer-Probenkammer platziert werden. Sie ermöglicht die Untersuchung von Gasgemischen mit Drücken zwischen 0 Pa und 1 MPa und kann außerdem elektrisch auf Temperaturen bis 1100 K geheizt werden. Die Zelle kann abhängig von zu untersuchendem Gas und Wellenlängenbereich mit verschiedenen infrarotdurchlässigen Fenstern versehen werden. In dieser Arbeit wurden Fenster aus Kaliumbromid verwendet, welches einen weiten Bereich hohen Transmissionsgrades im NIR- und MIR-Bereich aufweist [20]. Durch eine Temperaturstabilisierung mittels Thermostat werden die Fenster einerseits vor thermischen Schäden geschützt und andererseits wird der mögliche Temperatureinfluss auf die optischen Eigenschaften der Fenster minimiert. Im in dieser Arbeit untersuchten Druckbereich kann die Druckabhängigkeit des Brechungsindexes von Kohlenstoffdioxid sowie dessen Einfluss auf den Reflexionsgrad an der Grenzschicht und somit auf die optischen Eigenschaften der Fenster vernachlässigt werden [21].

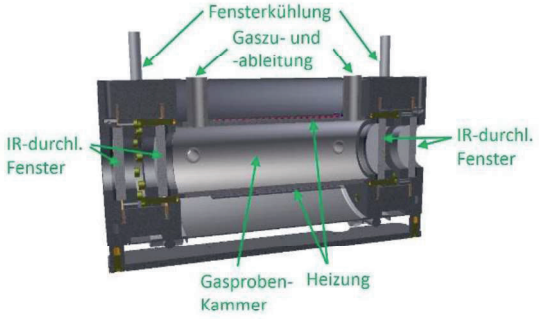


Abb. 4: Schnittdarstellung der elektrisch heizbaren Hochtemperatur-Hochdruck-Gaszelle zur Positionierung von Gas-Proben im FTIR-Spektrometer.

Messergebnisse und Diskussion

Ziel des ersten Versuchs in dieser Arbeit ist es, durch Analyse der Druckabhängigkeit des Verlaufs der einzelnen Absorptionsbanden von Kohlenstoffdioxid Banden zu identifizieren, die unter den Bedingungen beim Turbinenbetrieb in Sättigung sind. Dazu wird der spektrale gerichtet-gerichtete Transmissionsgrad von Kohlenstoffdioxid bei Raumtemperatur und verschiedenen Drücken gemessen.
Zur Charakterisierung von Kohlenstoffdioxid wird die Gaszelle in der Probenkammer des FTIR-Spektrometers platziert und der gesamte Strahlengang evakuiert. Zur Bestimmung eines jeden Transmissionsspektrums werden jeweils drei Einzelmessungen durchgeführt. Als erste davon wird stets bei evakuierter Zelle eine Referenzmessung ausgeführt. Anschließend wird die Zelle bis zu dem gewünschten Druck mit Kohlenstoffdioxid gefüllt und sogleich die spektrale Intensität gemessen. Nun wird die Zelle erneut evakuiert und es folgt eine weitere Refe-

renzmessung. Aus den beiden Referenzmessungen wird der Mittelwert gebildet, um so mögliche zeitabhängige Drifteffekte im Fenstermaterial zu minimieren. Der spektrale gerichtet-gerichtete Transmissionsgrad resultiert nun schließlich aus dem Quotienten aus der Probenmessung und der gemittelten Referenzmessung

In Abb. 5 ist der spektrale Transmissionsgrad von Kohlenstoffdioxid bei Raumtemperatur und Drücken zwischen 5 kPa und 600 kPa im Wellenlängenbereich von 1,5 µm bis 20 µm dargestellt. Wie erwartet, kann dort jede infrarotaktive Absorptionsbande (vgl. Tab. 2) von Kohlenstoffdioxid beobachtet werden. Darüber hinaus sind in Abb. 5 jedoch noch weitere, in Tab. 2 nicht genannte Banden zu sehen. Dabei handelt es sich um Kombinationsbanden, die aus einer Kombination zweier oder mehrerer fundamentaler Molekülschwingungen hervorgerufen werden. Ein Beispiel hierfür ist die Bande bei 2,7 µm, die aus den fundamentalen Schwingungen in den Kombinationen $\nu_3 + \nu_1$ und $\nu_3 + 2\nu_2$ (vgl. Tab. 2) besteht [10], [22].

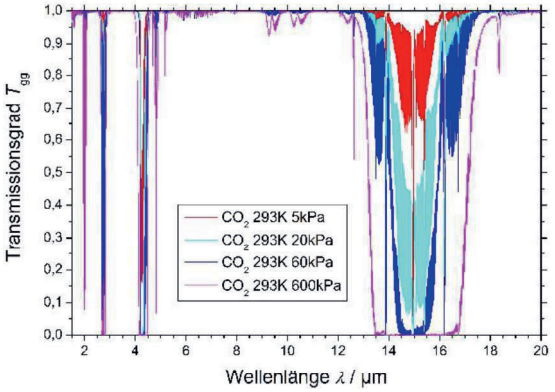


Abb. 5: Spektraler Transmissionsgrad von CO_2 bei Raumtemperatur und bei Drücken von 5 kPa bis 600 kPa im Wellenlängenbereich zwischen 1,5 µm und 20 µm.

Wie weiterhin erwartet, nimmt die Tiefe der Absorptionsbanden mit steigendem Druck zu. Im in dieser Arbeit untersuchten Druckbereich erreichen die drei Banden bei 2,7 µm, 4,26 µm und 14,99 µm Sättigung und erfüllen damit ein wesentliches Kriterium für die Nutzbarkeit für die spätere berührungslose Temperaturmessung.

In Abb. 6, Abb. 7 und Abb. 8 werden die spektralen Bereiche um diese drei Banden unter denselben Bedingungen wie zuvor, jedoch vergrößert gezeigt, um deren Verhalten bei verschiedenen Drücken näher zu betrachten. Abb. 6 zeigt die Bande bei 2,7 µm vergrößert. Es zeigt sich, dass die Absorptionsbande nur in einigen Bereichen überhaupt Sättigung erreicht. Selbst bei einem Druck von 600 kPa bleiben Bereiche partieller Transmission zurück.

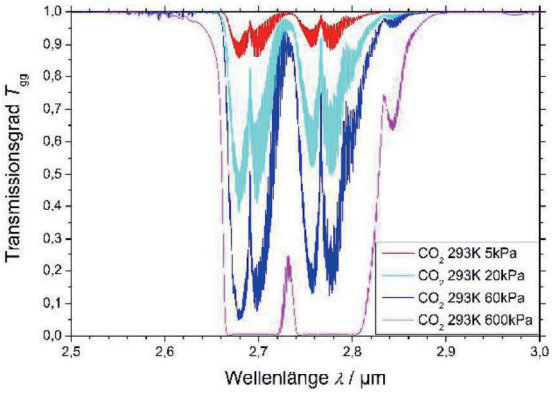


Abb. 6: Spektraler Transmissionsgrad von CO_2 im Wellenlängenbereich von 2,5 µm bis 3,0 µm bei Drücken zwischen 5 kPa und 600 kPa bei Raumtemperatur.

Abb. 7 zeigt den spektralen Transmissionsgrad von Kohlenstoffdioxid im Wellenlängenbereich zwischen 3,9 µm und 4,6 µm mit der Absorptionsbande bei 4,26 µm bei Raumtemperatur bei verschiedenen Drücken. Diese Bande ist bereits bei niedrigen Drücken sehr stark und vergrößert sich mit steigendem Druck noch deutlich. Ab einem Druck von 20 kPa ist sie bereits in Sättigung.

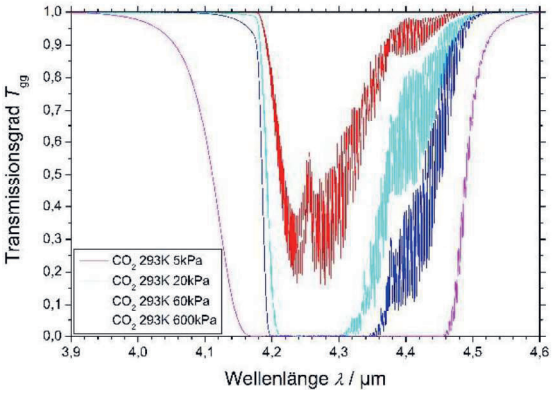


Abb. 7: Spektraler Transmissionsgrad von CO_2 bei Raumtemperatur und Drücken zwischen 5 kPa und 600 kPa im Wellenlängenbereich von 3,9 µm bis 4,6 µm.

In Abb. 8 ist der spektrale Transmissionsgrad von Kohlenstoffdioxid im Bereich der Absorptionsbande bei 14,99 µm bei Raumtemperatur und bei Drücken zwischen 5 kPa und 600 kPa dargestellt. Auch hier kann ein Anstieg der Tiefe der Bande mit steigendem Druck beobachtet werden. Der Anstieg ist stärker als bei der Bande bei 2,7 µm aber schwächer als bei der Bande bei 4,26 µm. Sättigung erreicht diese

Bande erst ab einem Kohlenstoffdioxiddruck von 60 kPa.

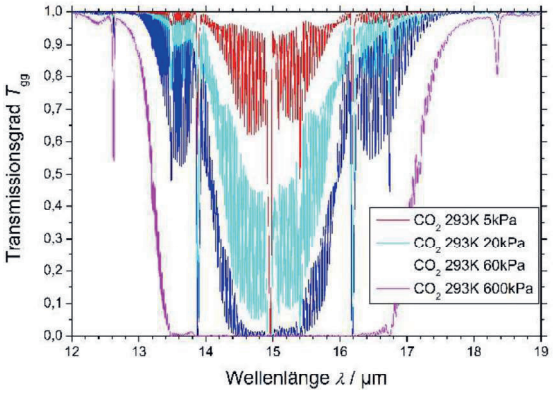


Abb. 8: Spektraler Transmissionsgrad von CO_2 im Wellenlängenbereich von 12 µm bis 19,5 µm bei Raumtemperatur und Drücken zwischen 5 kPa und 600 kPa.

Bei Raumtemperatur erscheint daher die Kohlenstoffdioxidbande bei 4,26 µm von den Untersuchten am passendsten. Bei geometrischen Weglängen in der Endanwendung von mindestens der der Gaszelle kann auch die Bande bei 14,99 µm genügen.

Neben der Druckabhängigkeit beeinflusst jedoch auch das Verhalten einer Bande bei hohen Temperaturen ihre Nutzbarkeit für berührungslose Temperaturmessungen. Deshalb wird im nächsten Schritt die Temperaturabhängigkeit des spektralen gerichtet-gerichteten Transmissionsgrades von Kohlenstoffdioxid bei dem in Turbinenbrenngasen üblichen Partialdruck von 60 kPa untersucht.

In Abb. 9 wird der spektrale gerichtet-gerichtete Transmissionsgrad von Kohlenstoffdioxid im Wellenlängenbereich zwischen 1,5 µm und 20 µm bei einem Druck von 60 kPa und bei Temperaturen zwischen Raumtemperatur und 1073 K gezeigt. Wie erwartet, nimmt die Breite der Banden mit steigender Temperatur zu. Zur näheren Betrachtung sind die Banden bei 4,26 µm und 14,99 µm in Abb. 10 und Abb. 11 erneut vergrößert dargestellt.

In Abb. 10 ist der spektrale Transmissionsgrad von Kohlenstoffdioxid im Wellenlängenbereich zwischen 4,1 µm und 4,8 µm dargestellt. Die Absorptionsbande bei 4,26 µm ist hier vergrößert dargestellt bei einem Druck von 60 kPa und bei Temperaturen zwischen Raumtemperatur und 1073 K. Die Breite der Bande zeigt eine starke Zunahme mit steigender Temperatur. Sie ist bei diesem Druck bei allen untersuchten Temperaturen klar in Sättigung, was hinsichtlich der Temperaturmessung von Vorteil ist.

Abb. 11 zeigt die Kohlenstoffdioxidbande bei 14,99 µm vergrößert im Wellenlängenbereich zwischen 12 µm und 19 µm. Wie zuvor ist der spektrale Transmissionsgrad bei einem Druck von 60 kPa und bei Temperaturen zwischen Raumtemperatur und 1073 K dargestellt. Auch hier zeigt sich eine Zunahme der Breite der Bande mit steigender Temperatur. Allerdings ist die Zunahme weniger deutlich als bei der Bande bei 4,26 µm (vgl. Abb. 10).

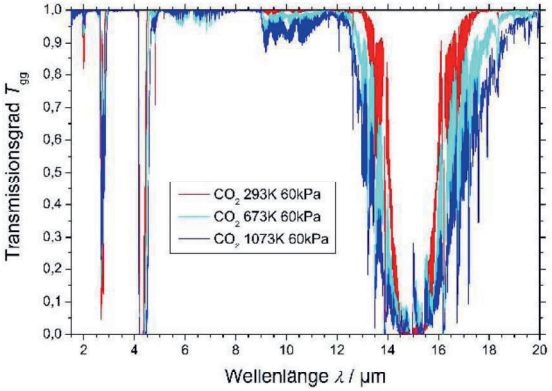


Abb. 9: Spektraler Transmissionsgrad von CO_2 bei einem Druck von 60 kPa und Temperaturen zwischen Raumtemperatur und 1073 K im Wellenlängenbereich von 1,5 µm bis 20 µm.

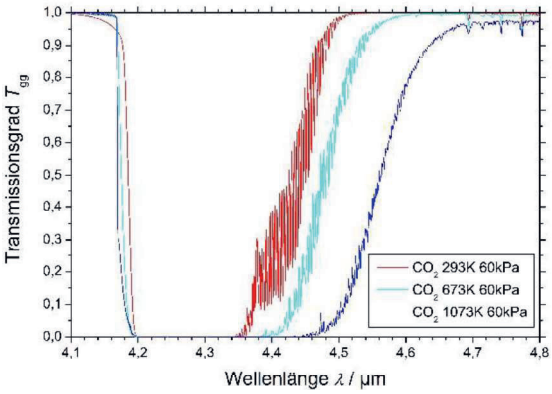


Abb. 10: Spektraler Transmissionsgrad von CO_2 bei Temperaturen zwischen Raumtemperatur und 1073 K, bei einem Druck von 60 kPa im Wellenlängenbereich von 4,1 µm bis 4,8 µm.

Die Kohlenstoffdioxidbande bei 4,26 µm ist daher unter den hier untersuchten Banden die für die berührungslose Temperaturmessung geeignetste. Das Zentrum der Bande ist bereits bei niedrigen Partialdrücken in Sättigung und die Breite der Bande zeigt einen starken Anstieg mit steigender Temperatur. Wie bereits erwähnt sind diese beiden Eigenschaften besonders bedeutend für die geplante Temperaurmessmethode.

Ein weiterer Vorteil wird in Abb. 12 deutlich, wo diese Absorptionsbande bei Temperaturen zwischen Raumtemperatur und 1073 K und bei ihrem niedrigsten Sättigungsdruck – 20 kPa –

gezeigt wird. Selbst bei diesem niedrigen Druck zeigt die Breite der Bande nahezu den gleichen Anstieg mit steigender Temperatur. Es existiert somit ein weiter Bereich von Kohlenstoffdioxidpartialdrücken und von geometrischen Weglängen, in welchem Temperaturmessungen möglich erscheinen.

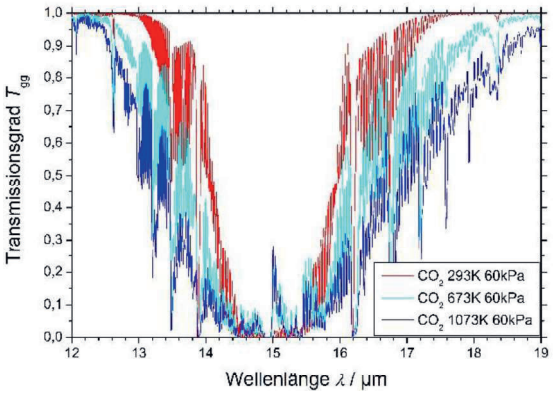


Abb. 11: Spektraler Transmissionsgrad von CO_2 im Wellenlängenbereich von 12 µm bis 19 µm bei einem Druck von 60 kPa und Temperaturen zwischen Raumtemperatur und 1073 K.

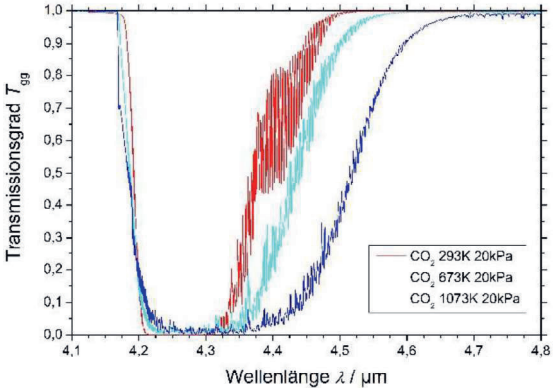


Abb. 12: Spektraler Transmissionsgrad von CO_2 bei Temperaturen zwischen Raumtemperatur und 1073 K, bei einem Druck von 20 kPa im Wellenlängenbereich von 4,1 µm bis 4,8 µm.

Die letztliche Wahl der exakten Messwellenlängenbereiche hängt von Parametern wie geometrischer Weglänge und Kohlenstoffdioxidpartialdruck im konkreten Anwendungsfall ab. Wie oben erläutert, wird für die Gastemperaturmessung in der äußeren Schicht eines Gasstroms ein Wellenlängenbereich benötigt, in welchem die Absorptionsbande in Sättigung ist. Für diesen Zweck kann somit ein Wellenlängenbereich etwa in der Region zwischen 4,20 µm und 4,35 µm gewählt werden. Für die Gastemperaturmessung an inneren Gasschichten des Brenngasstroms wird dagegen ein Wellenlängenbereich benötigt, in welchem der Verlauf des spektralen Transmissionsgrades von der Gastemperatur abhängt. Für diese Art der Messung kann daher ein Wellenlängenbereich in der Region des langwelligen Randes der Bande, etwa zwischen 4,35 µm und 4,65 µm gewählt werden.

Um die erzielten Ergebnisse zu verifizieren, wurden diese schließlich noch mit entsprechenden HITRAN-Simulationen verglichen. Die HITRAN-Datenbank beinhaltet eine Sammlung von hochaufgelösten linienbasierten Spektren bzgl. der Absorptionskoeffizienten einer großen Menge verschiedener Gase [23], [24], [25]. Die hier gezeigten Simulationen des spektralen Transmissionsgrades wurden mittels des Systems „HITRAN on the web“ erstellt [16]. Dieses simuliert das Transmissionsspektrum eines bestimmten Gases für ein zuvor gewähltes Parameterset, bestehend z. B. aus Temperatur, Druck und geometrischer Weglänge auf Grundlage der HITRAN-Datenbank [26].

Abb. 13 zeigt den Vergleich des gemessenen spektralen Transmissionsgrades von Kohlenstoffdioxid bei Raumtemperatur und bei einem Druck von 60 kPa mit einer entsprechenden HITRAN-Simulation im Wellenlängenbereich von 4,1 µm bis 4,8 µm. Das gemessene Spektrum ist als blaue Linie und die Simulation in grün dargestellt. Im gesamten untersuchten Wellenlängenbereich kann eine gute Übereinstimmung beobachtet werden. Abweichungen zwischen Simulation und Messung bei höheren Temperaturen zeigen jedoch die Notwendigkeit der Erhebung von experimentellen Daten.

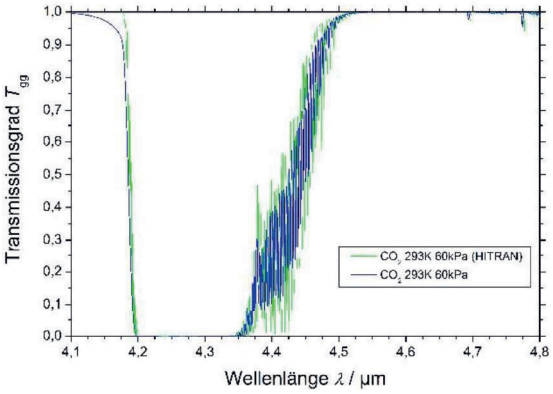


Abb. 13: Vergleich einer HITRAN-Simulation [16] (grün) mit der entsprechenden Messung des spektralen Transmissionsgrades (blau) von CO_2 im Wellenlängenbereich zwischen 4,1 µm und 4,8 µm bei Raumtemperatur und einem Druck von 60 kPa.

Zusammenfassung und Ausblick

Für die Entwicklung einer neuen berührungslosen Methode der Brenngastemperaturmessung in arbeitenden stationären Gasturbinen sind

einige Vorarbeiten nötig. Das Ziel dieser Arbeit war es daher, eine für diese Messung geeignete Absorptionsbande des Kohlenstoffdioxids zu identifizieren.
Um entsprechende Wellenlängenregionen auszuwählen, wurde der spektrale gerichtet-gerichtete Transmissionsgrad von Kohlenstoffdioxid im Wellenlängenbereich zwischen 1,5 µm und 20 µm gemessen. Dabei wurden Kohlenstoffdioxiddrücke zwischen 5 kPa und 600 kPa und Gastemperaturen zwischen Raumtemperatur und 1073 K untersucht. Die Messergebnisse wurden durch entsprechende HITRAN-Simulationen verifiziert. Als Folge dieser Experimente konnte die Kohlenstoffdioxidbande bei 4,26 µm für künftige Gastemperaturmessungen gewählt werden. Insbesondere ist die Wellenlängenregion zwischen 4,20 µm und 4,35 µm im Zentrum der Bande für Temperaturmessungen in äußeren Schichten eines Brenngasstroms geeignet während die Region zwischen 4,35 µm und 4,65 µm für Messungen im Inneren des Gasstroms genutzt werden können.
Die nächsten Schritte zur Weiterentwicklung der neuen berührungslosen Methode zur Brenngastemperaturmessung werden der Entwurf eines Pyrometers für die oben definierten Wellenlängenbereiche und die Durchführung von Testmessungen im Labor sowie danach in einer arbeitenden Gasturbine sein.
Ein weiteres wichtiges Ziel künftiger Arbeiten in diesem Bereich besteht im Aufnehmen zusätzlicher experimenteller Transmissionsspektren z. B. von Wasserdampf, dem zweiten infrarotaktiven Bestandteil der Turbinenbrenngase, zur Identifizierung möglicher weiterer nutzbarer Messwellenlängenbereiche. Darüber hinaus sollen auch Transmissionsspektren beider infrarotaktiver Gase (Wasserdampf und Kohlenstoffdioxid) in Kombination mit den übrigen Verbrennungsgasen (vgl. Tab. 1) untersucht werden, um die Transmissionsgrade unter den realen Anwendungsbedingungen zu ermitteln und um mögliche Einflüsse des Gesamtdrucks und der Gesamtteilchenzahl auf die Absorptionsbanden zu erfassen.

Literaturnachweis

[1] International Energy Agency, *World Energy Outlook 2017*, Paris: IEA Publications (2017)
[2] Siemens, https://www.siemens.com/press/de/pressemitteilungen/2016/power-gas/pr2016010135pgde.htm (18.10.2018)
[3] N. Armaroli, V. Balzani, *Energy & Environmental Science* 4, 3193-3222 (2011); doi: 10.1039/C1EE01249E
[4] C. Lechner, J. Seume, *Stationäre Gasturbinen*, Berlin: Springer Verlag (2010); doi: 10.1007/978-3-540-92788-4
[5] J. H. Perepezko, *Science* 326, 1068-1069 (2009); doi: 10.1126/science.1179327
[6] J. Manara, M. Zipf, T. Stark, M. Arduini, H.-P. Ebert, A. Tutschke, A. Hallam, J. Hanspal, M. Langley, D. Hodge, J. Hartmann, *Infrared Physics & Technology* 80, 120-130 (2017); doi: 10.1016/j.infrared.2016.11.014
[7] M. Zipf, J. Manara, T. Stark, M. Arduini, H.-P. Ebert, J. Hartmann, *High Temperatures High Pressures* 47, 3-21 (2018)
[8] F. Bernhard, *Handbuch der technischen Temperaturmessung* 2nd Edition, Berlin: Springer Verlag (2014); doi: 10.1007/978-3-642-24506-0
[9] H. Haken, H. C. Wolf, *Molekülphysik und Quantenchemie* 5th Edition, Berlin: Springer Verlag (2006); doi: 10.1007/3-540-30315-4
[10] H. Günzler, H.-U. Gremlich, *IR-Spektroskopie – Eine Einführung* 4th Edition, Weinheim: Wiley-VCH (2003); doi: 10.1002/9783527662852
[11] W. Demtröder, *Experimentalphysik 3 – Atome, Moleküle und Festkörper* 5th Edition, Berlin: Springer Verlag (2010); doi: 10.1007/978-3-642-03911-9
[12] J. M. Flaud, C. Camy-Peyret, *Molecular Physics* 26, 811-823 (1973); doi: 10.1080/00268977300102121
[13] P. Varanasi, *Journal of Quantitative Spectrocopy & Radiative Transfer* 40, 169-157 (1988); doi: 10.1016/0022-4073(88)90110-0
[14] C. P. Rinsland et. al., *NASA Technical Memorandum* 85764, 1-467 (1984)
[15] R. Paso, J. Kauppinen, R. Anttila, *Journal of Molecular Spectroscopy* 79, 236-253 (1980); doi: 10.1016/0022-2852(80)90304-5
[16] Harvard-Smithsonian Center for Astrophysics (CFA), Cambridge, MA, USA, V.E. Zuev Insitute of Atmosperic Optics (IAO), Tomsk, Russia, http://hitran.iao.ru/ (30.01.2019)
[17] J. T. Houghton, *Quarterly Journal of the Royal Meteorological Society* 117, 29-58 (2013); doi: 10.1002/qj.49709540302
[18] J. Tennyson et. al., *Journal of Quantitative Spectrocopy & Radiative Transfer* 117, 29-58 (2013); doi: 10.1016/j.jqsrt.2012.10.002
[19] P. R. Griffiths, J. A. de Haseth, *Fourier Transform Infrared Spectrometry* 2nd Edition, Hoboken: John Wiley & Sons (2007); doi: 10.1002/047010631X
[20] Korth Kristalle GmbH, http://www.korth.de/index.php/ material-detailansicht/items/15.html (06.02.2019)
[21] A. Michels, J. Hamers C.s.s.R., *Physica* 4, 995-1006 (1937); doi: 10.1016/S0031-8914(37)80197-4
[22] P. E. Martin, E. F. Barker, *Physical Review* 41, 291-303 (1932); doi: 10.1103/PhysRev.41.291
[23] R. A. McClatchey et. al., *AFCRL Environmental Research Papers* 434 (1973)
[24] L. S. Rothman et. al., *Applied Optics* 26, 4058-4097 (1987); doi: 10.1364/AO.26.004058
[25] I. E. Gordon et. al., *Journal of Quantitative Spectrocopy & Radiative Transfer* 203, 3-69 (2017); doi: 10.1016/j.jqsrt.2017.06.038
[26] Harvard-Smithsonian Center for Astrophysics (CFA), Cambridge, MA, USA, V.E. Zuev Insitute of Atmosperic Optics (IAO), Tomsk, Russia., http://hitran.iao.ru/home.sim-theory (30.01.2019)

DOI 10.5162/sensoren2019/2.4.3

Modulare Sensor-Testplattform für Hochtemperaturanwendungen bis 300 °C

Björn Bieske, Tom Reinhold, Marco Reinhard
IMMS Institut für Mikroelektronik- und Mechatronik-Systeme gemeinnützige GmbH (IMMS GmbH), Ehrenbergstr. 27, 98693 ILMENAU, Deutschland

Zusammenfassung

Die Dauertemperatureinsatzbereiche von Halbleiterbauelementen sind bis 125 °C als Standards definiert. Für spezielle elektronische Schaltungen in den Bereichen Automotive und Sensorik werden aber Einsatztemperaturen bis zu 300 °C verlangt. Zur Evaluierung dieser Schaltungen wird eine innovative Hochtemperatur-Testplattform vorgestellt, die am IMMS entwickelt wurde und als Kernstück eine Schaltkreisfassung mit Heizung und Temperaturregelung beinhaltet. Als Demonstrator wird ein Drucksensor-ASIC damit charakterisiert.

Keywords: Hochtemperaturtest; Testmethodik; IC-Fassung; Evaluierung, Sensorik.

Einleitung

In vielen Industrieanlagen und im Automobilbereich [1] werden Sensoren eingesetzt, die über einen großen Temperaturbereich korrekte Messwerte unterschiedlichster Messgrößen liefern müssen. Die Sensoren und deren Sensorelektroniken zur Aufbereitung der Messwerte sollen in ein gemeinsames Gehäuse integriert werden. Dadurch entstehen Anforderungen an Schaltkreise, welche die bisher definierten Temperaturbereiche weit übersteigen.
Als Standard wurden vier Einsatztemperaturbereiche für den Dauerbetrieb elektronischer (Halbleiter)-Bauelemente definiert:

- Kommerzieller Temperaturbereich:
  0 °C bis +70 °C:
  Unterhaltungs- und „Consumer-Elektronik“, Radio, TV, Computer, Smart Home
- Industrieller Temperaturbereich:
  –40 °C bis +85 °C:
  Industrielle Elektronik, Steuerungen, Regelungen, Sensorik, Maschinen
- Automobil-Temperaturbereich:
  –40 °C bis +125 °C:
  Motornahe Elektronik, Steuergeräte, Klimageräte, Sensoren, Abgasüberwachung
- Militärischer Temperaturbereich:
  –55 °C bis +125 °C:
  Luftfahrtindustrie, Funkgeräte, Messgeräte

Für spezielle Anwendungen liegen die Anforderungen jenseits dieser Temperaturbereiche. So werden in der Nähe von Hitzequellen integrierte Schaltungen benötigt, die im Temperaturbereich bis 300 °C zuverlässig funktionieren müssen, z.B. an Verbrennungsmaschinen oder an Bohrköpfen für Tiefenbohrungen. Derartige Schaltungen funktionieren bei diesen Temperaturen aufgrund physikalischer Grenzen nicht mehr in Standard-CMOS-Technologien [2].

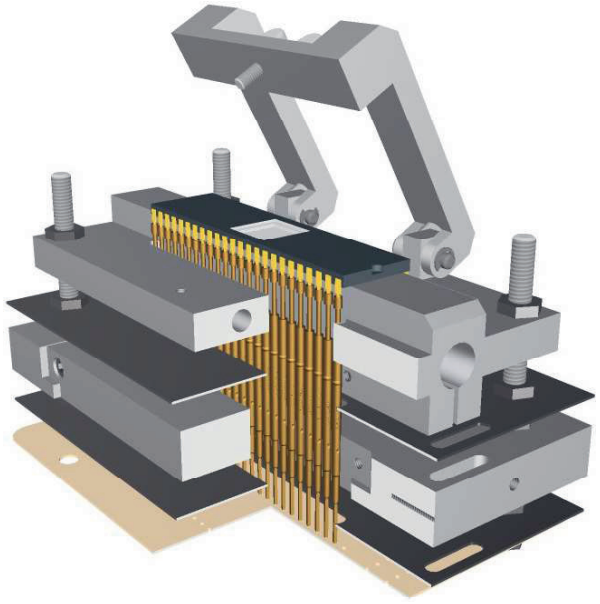


Abb. 1: Design der Hochtemperatur- Schaltkreis-Fassung für DIL-Gehäuse

Hochtemperatur-Schaltungen

Als Alternative werden spezielle Silicon-On-Insulator-Technologien (SOI) verwendet, die im Temperaturbereich bis 300 °C einsetzbar sind [3]. Voraussetzung für die Realisierung komplexer integrierter Strukturen und Schaltungen sind genaue Modelle der verwendeten

Elemente. Grundlage für eine genaue Modellierung sind genaue Messwerte über den kompletten Temperaturbereich. Entsprechend wird eine Hochtemperatur-Testplattform benötigt, die es ermöglicht, das Testobjekt („Device under Test“ - DUT) bis 300 °C reproduzierbar zu temperieren. Die umgebende Testschaltung mit externen Komponenten darf in vielen Fällen nur in den Standard-Temperaturbereichen betrieben werden und muss vom DUT separiert werden.
Zur Realisierung, Evaluierung und Charakterisierung von Hochtemperaturschaltungen gibt es folgende Möglichkeiten:

- Wafermessung, nötige externe Bauelemente werden auf einer Probecard platziert [4]
- Hybridschaltung auf Keramik mit externen Komponenten (Dickschichttechnologie)
- IC im Keramik-Gehäuse und externe Komponenten auf abgesetzter Platine

Die Keramik-Gehäuse zur Aufnahme der Hochtemperatur-Schaltung [5] sind bis 300 °C einsetzbar (Abb. 2). Die abgesetzte Platine zur weiteren Signalverarbeitung mit den externen Bauelementen wird nicht den hohen Temperaturen ausgesetzt und kann damit als konventionelle Lösung auf herkömmlichen Substratmaterialien realisiert werden [7]. Dieser Ansatz stellt einen guten Kompromiss zwischen Aufwand und Nutzen dar und wird mit einer Hochtemperaturfassung für DIL-Gehäuse (Dual in Line) realisiert, was im Folgenden detailliert beschrieben wird.

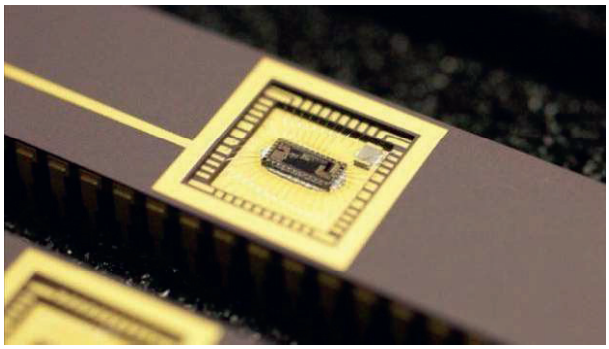


Abb. 2: ASIC D2040A und Temperatursensor im Keramikgehäuse DIL-40

Im Gegensatz dazu benötigt man für die Wafermessung einen Waferprober mit Thermochuck und Probecard bzw. Einzelnadeln zur Kontaktierung des DUT [4], was einen erheblichen technischen Aufwand darstellt.
Die Variante mit einer Hybridschaltung auf Keramiksubstrat ist eine sehr kompakte Lösung, die aber zusätzlichen Aufwand an Aufbau- und Verbindungstechnik voraussetzt. Die umgebende Applikationsschaltung mit den externen Komponenten ist dann aber fest definiert und nicht flexibel an die jeweiligen Bedingungen anpassbar.
Bei entsprechender Größe könnte diese aber auch in ein passendes DIL-Gehäuse zur Kontaktierung mit der vorgestellten Hochtemperaturfassung eingebettet werden.
Für konfigurierbare Tests zur Evaluierung und Charakterisierung von Sensor-Schaltkreisen wurde eine portable Testplattform entwickelt, welche Messungen an ASICs (Application Specific Integrated Circuit) unter Verwendung einer applikationsspezifischen Schaltung bis 300°C ermöglicht. Diese Testplattform wurde modular gestaltet, um Schaltungselemente flexibel an verschiedene Testanforderungen anpassen zu können. Das zentrale Element der Plattform bildet die spezielle Hochtemperaturfassung, die als Aufnahme für das Testobjekt (DUT) dient (Abb. 1).

Hochtemperatur-Fassung mit Heizung

Die am IMMS entwickelte Hochtemperaturfassung für DIL-Gehäuse (Abb. 1) wurde im Rahmen der Entwicklung der Testplattform für Evaluierungs- und Demonstrationszwecke von Sensorschaltungen konstruiert. Dabei spielt die optische und mechanische Erreichbarkeit des Testobjekts eine wichtige Rolle, was sich in den Eigenschaften der Fassung widerspiegeln muss.
Durch die deutlich höheren Temperaturen im Vergleich zu den aufgeführten Standard-Temperaturbereichen ergeben sich neue Anforderungen, die beim herkömmlichen Test von Bauelementen keine Rolle spielen. Die Auswahl der passenden Materialen [5] ist eine wesentliche Voraussetzung für eine stabile und robuste Testplattform. Kommerziell verfügbare Hochtemperaturfassungen [6,9] waren für den geplanten Demonstratoraufbau nicht einsetzbar, da sie zur Temperierung des Testobjekts (DUT) eine externe Wärmequelle oder Klimakammer benötigten und die Kontaktierung des DUT nicht in gewünschter Weise realisierbar war.
Folgende Anforderungen sollen von der neu entwickelten Hochtemperatur-Testfassung erfüllt werden:

- Elektrische Kontaktierung:
  DUT ist einfach wechselbar
- Temperierung des DUT:
  Aufbau ohne eine zusätzliche externe Wärmequelle z.B. Ofen oder Thermostream®
- Messung der DUT-Temperatur für Regelung:
  Temperatursensor oder Infrarotthermometer
- Testlösung als Demonstrator bzw. Referenz-Applikationsschaltung
- Kompakter und transportabler Aufbau

Daraus ergibt sich, dass direkt in die Hochtemperaturfassung ein Heizelement und ein Temperatursensor integriert sein müssen. Die Heizleistung muss ausreichend groß sein, um ein schnelles Aufheizen zu ermöglichen. Die Abkühlung soll durch einfache Konvektion erfolgen. Erste Tests mit verfügbaren Heizelementen ergaben, dass mindestens 100 W erforderlich sind, wenn von einer Aufheizzeit von maximal zehn Minuten ausgegangen wird. Die Temperaturverläufe beim Aufheizen sind in Abb. 3 dargestellt.
Die Heizelemente müssen mit Gleichspannung unter 60 V betrieben werden, da ansonsten der Grenzwert für Kleinspannungen überschritten wird und die erforderlichen Maßnahmen für Isolation und Berührungsschutz einen praktikablen Betrieb des Testsystems verhindern würden. Außerdem werden durch den DC-Betrieb Störquellen vermieden, die durch kapazitive Einkopplung zu einer Beeinträchtigung der Messergebnisse führen können.

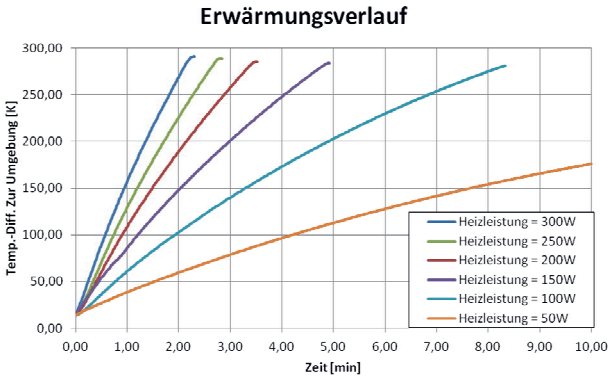


Abb. 3: Aufheizkurve der Hochtemperaturfassung ohne DUT [11]

Daher kommt eine Heizpatrone mit einer Leistung von 160 W bei einer Betriebsspannung von 24 V zum Einsatz. Diese bildet zusammen mit einem Aluminiumblock und einem integrierten Temperatursensor das Heizelement der Fassung. Die Regelung erfolgt digital. Die Genauigkeit der Temperierung liegt im Bereich von +/- 5 K bei 300 °C.

Kontaktierung des DUT

Die elektrische Isolation der Pins des DUT und die thermische Isolation des DUT zur Umgebung und zu externen Bauteilen müssen gewährleistet sein. Als Gehäuse für die zu untersuchenden integrierten Schaltungen wurden einfach verfügbare keramische DIL-Gehäuse mit maximal 48 Pins und einem Reihenabstand von 600 mil = 15,24 mm ausgewählt.
Diese bieten auch genügend Platz für größere Schaltkreise und haben eine ausreichend große Fläche im Gehäuseboden zur thermischen Kopplung zum Heizelement der Fassung. Der Abstand der Pins ist ausreichend groß, um eine sichere Kontaktierung sowie eine gute Isolation und eine stabile Führung der Kontaktierungsvorrichtung zu realisieren (Abb. 4).

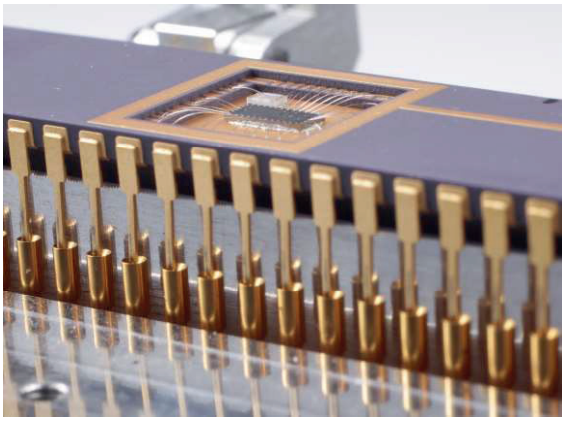


Abb. 4: Federkontakte zur Kontaktierung der Pins des DIL-Gehäuses in der Hochtemperaturfassung [8]

Die beschriebenen Ansätze und Gesichtspunkte können auf Vorrichtungen für weitere präzisere Messungen direkt übertragen werden. Falls sich diese Notwendigkeit ergibt ist allerdings ein anderes Design des Heizelementes zu verwenden, um Verfälschungen der Messergebnisse zu vermeiden und die Toleranzen der Temperierung zu verringern. Alternativ kann die Heizung während der Messung abgeschaltet werden.
Außerdem stellen die hohen Temperaturen hohe Ansprüche an die Aufbau- und Verbindungstechnik. So überschreitet die Maximaltemperatur den Schmelzpunkt herkömmlicher Lote. Herkömmliche kostengünstige Leiterplattenmaterialien sind bereits ab 125 °C nicht mehr einsetzbar, Hochtemperaturmaterialien sind bis 230 °C einsetzbar [7].
Die Kontaktierung der Pins erfolgt mit Federkontakten (Abb. 5), die folgende Eigenschaften aufweisen müssen:

- Abstand von 100 mil (2,54 mm), Breite < 2,4 mm
- Kopf: Innenkegel zur Ausrichtung der Pins
- Länge = Fassungshöhe (Realisierung der thermischen Isolation durch Abstand)
- Temperaturbeständigkeit

Die Maximaltemperatur der verfügbaren Federkontakte beträgt 200 °C [8]. Daher müssen diese nicht nur elektrisch, sondern auch thermisch vom Heizelement isoliert gehaltert werden. Die zwei Federkontaktreihen werden mit hitzefestem Isolierschlauch überzogen und in

passenden Aussparungen mit je zwei Aluminium-Klemmblöcken mechanisch fixiert. Diese Klemmblöcke sind von den beheizten Teilen der Fassung separiert und dienen gleichzeitig der Kühlung durch Konvektion.
Die Kontaktfläche der Innenkegel zum Pin des DIL Keramikgehäuses ist klein genug, so dass hier nur wenig thermische Energie direkt auf die Federkontakte übertragen wird.

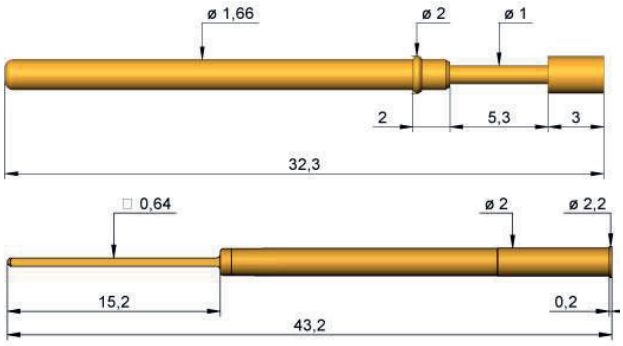


Abb. 5: Federkontakte und Hülsen zur Kontaktierung der Pins [8]

Das DUT wird mit Hilfe von zwei Aluminiumbügeln in der Fassung mit großflächigem Kontakt zum Heizelement fixiert. Die am DUT vorhandenen Temperaturen können mit einem integrierten Temperatursensor (Pt1000) oder mit einem Infrarot-Thermometer kontrolliert werden. Der Temperaturabfall vom Heizelement bis zum DUT lag bei 300 °C unter fünf Kelvin.

Applikation Drucksensor

Am Beispiel dient der am IMMS entwickelte Hochtemperatur Auswerteschaltkreis in SOI-Technologie für Drucksensoren, welcher Sensorsignale, die von einer resistiven Messbrücke erzeugt werden, weiterverarbeitet [10].
Anhand dieses Beispiels soll gezeigt werden, wie die entsprechende Hochtemperaturelektronik evaluiert und getestet wird. In der konkreten Anwendung sollen die Temperaturgänge der resistiven Messbrücke und des Messbrücken-Signalverstärkers linearisiert werden.

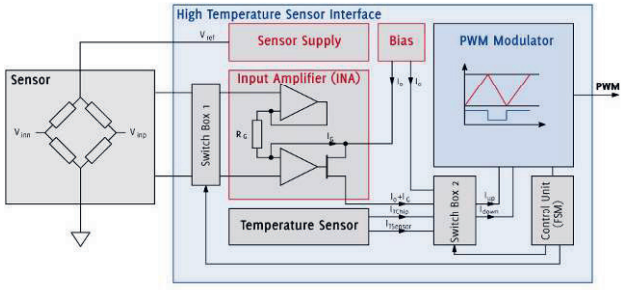


Abb. 6: Prinzipschaltbild des Inferfaces zur Verarbeitung der Sensorsignale [10]

Die Ausgangsspannung der Messbrücke, die im Bereich weniger Millivolt liegt, wird von der Signalverarbeitungsschaltung, die im applikationsspezifischen Schaltkreis (ASIC) D2040A integriert ist, aufgenommen, verstärkt und in einen Strom umgewandelt (Abb. 6).
Weiterhin werden zusätzliche Signale aufgenommen, die zur Kompensation von Nullpunkt- und Verstärkungsfehlern dienen. Diese aufbereiteten Signale werden pulsmoduliert an eine digitale Verarbeitungseinheit übertragen, die nicht mehr den hohen Temperaturen ausgesetzt ist.
Zur Bestimmung der Temperatur des ASICs wird zusätzlich ein Pt1000-Sensor im IC Gehäuse eingesetzt. Auch dieser Messwert wird an die Auswerteelektronik übertragen. So kann die Temperaturdifferenz zwischen Heizpatrone der Hochtemperaturfassung und des DUTs bestimmt werden und entsprechend ausgeregelt werden.

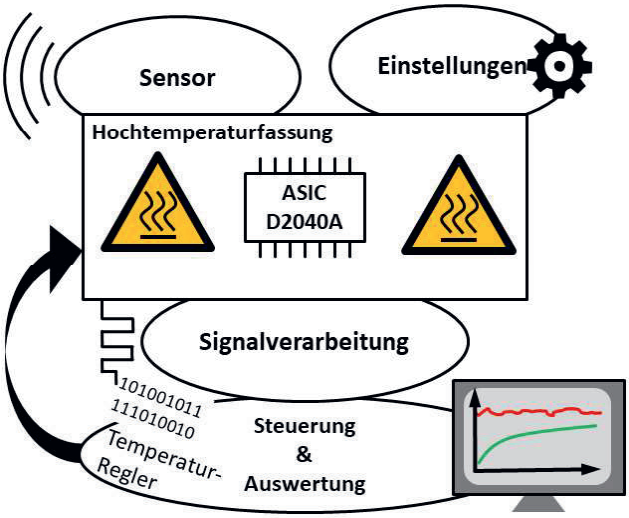


Abb. 7: Systemkonzept der modularen Hochtemperatur Testplattform [11]

Die Auswertung der pulsmodulierten Daten erfolgt mit Hilfe eines FPGA, welcher die temperaturkompensierten Messwerte ermittelt und ausgibt. Ursprünglich war ein analoges 20 mA Stromschleifen-Interface zur Datenübertragung und Stromversorgung des Sensor-ASICs D2040A geplant. Auf schaltungstechnische Details des Drucksensors soll an dieser Stelle nicht weiter eingegangen werden [10].
Der Deckel des verwendeten DIL-Gehäuses ist abnehmbar und die Hochtemperaturfassung hat an dieser Stelle eine Aussparung. So kann die IC-Temperatur auch mit einem Infrarotthermometer bzw. einer Thermo-Kamera bestimmt werden. Außerdem können die Temperaturdifferenzen im DUT ortsaufgelöst kontrolliert und gemessen werden.

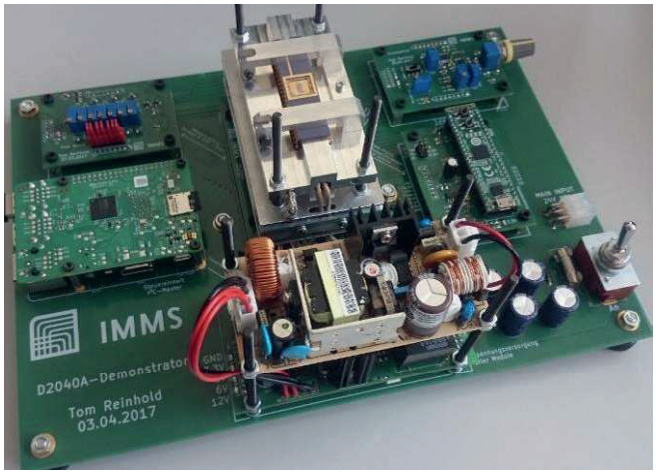


Abb. 8: Finaler Demonstrator der Hochtemperatur Testplattform mit Inferface- und Messmodulen für den ASIC D2040A

Modulare Testplattform als Demonstrator

Die entwickelte Hochtemperaturfassung wird zur Testplattform als Demonstratoraufbau mit modularem Konzept erweitert (Abb. 7). Auf einer Grundplatine werden die benötigten Module angeordnet und kontaktiert [11]. Diese können anwendungsspezifisch ausgeführt werden.

Die Stromversorgung erfolgt über ein Schaltnetzteil, woraus die 24 V für die Heizung und verschiedene Spannungen zwischen 3,3 und 12 V für die Steuerungselektronik bereitgestellt werden. Durch dieses externe Schaltnetzteil erfolgt gleichzeitig die Schutzisolation vom Stromnetz für den gesamten Demonstrator.

Die Temperaturregelung der Fassung erfolgt mit Hilfe eines Einplatinen-Computers vom Typ Raspberry Pi® mit Linux-basiertem Betriebssystem. Dieser kann zur Ansteuerung der Messhardware verschiedene serielle Bussysteme (I2C, SPI) und Interfaceleitungen (GPIOs) als konfigurierbare Steuerleitungen bedienen. Die Programmierung erfolgt in C++ oder Python™.

Das Systemkonzept ist dafür ausgelegt, dass die gewonnenen Messdaten und der Temperaturverlauf auf einem externen Monitor grafisch dargestellt werden können (Abb. 7).

Die DUT-spezifische Messhardware wird in Form verschiedener Module auf die Grundplatine (Motherboard) aufgesetzt (Abb. 8). In diesen wird die Signalauswertung im FPGA realisiert.

Um verschiedene Eingangssignale und Fehlerfälle des ASICs und der Messbrücke (Kurzschlüsse bzw. offene Verbindungen) emulieren zu können, wird der Sensor als resistive Brücke aus einzelnen Widerständen nachgebildet. Einige Teilwiderstände sind einstellbar ausgeführt, um die Messbrücke gezielt verstimmen zu können. Mit Hilfe eines Oszilloskops können die Signalverläufe überprüft und weitergehend analysiert werden [11].

Der gesamte Demonstrator nimmt etwa die Fläche eines DIN A4 Blattes ein und kann dadurch einfach transportiert werden.

Ergebnisse und Ausblick

Es wurde eine modulare Hochtemperatur-Testplattform entwickelt, in der Bauelemente im keramischen DIL-Gehäuse bis 300 °C charakterisiert werden können. In der Hochtemperaturfassung werden die Bauelemente kontrolliert temperiert und sicher elektrisch kontaktiert. Die Gehäuse werden mit Deckeln verschlossen, um die Schaltungsfunktion nicht mit Licht zu beeinträchtigen. Durch eine Aussparung im Deckel der Fassung können die Testobjekte auch optisch überwacht werden.

Die notwendigen zusätzlichen Module zur Verarbeitung der Messwerte für das jeweilige DUT werden auf einer Grundplatine als kompakter Aufbau zusammengefasst.

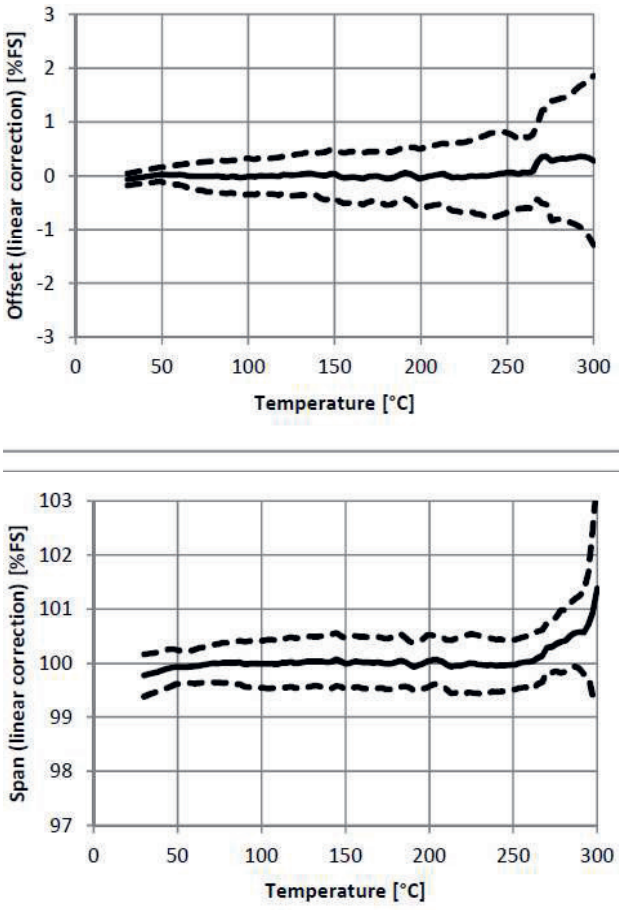


Abb. 9: Abweichungen von Nullpunkt und Vollausschlag der Messergebnisse des ASICs D2040A über der Temperatur nach der Kalibrierung [10]

Mit der vorgestellten innovativen Hochtemperatur-Testplattform konnte die Funktionsfähigkeit des Drucksensor-ASICs D2040A [10] im Temperaturbereich bis 300 °C nachgewiesen werden. Hierzu wurde relative Änderung der ausgegebenen Messwerte im Nullpunkt und bei Vollaussteuerung nach einer notwendigen Kalibrierung für verschiedene Schaltkreise aufgetragen. Die Daten der Brückenelektronik (Nullpunkt und Verstärkung) blieben über der Temperatur stabil.
Im Diagramm sind die Mittelwerte und maximalen Abweichungen dargestellt (Abb. 9).
Die gemessenen Daten können auf einem externen Monitor in Echtzeit dargestellt werden. Die grafische Aufbereitung und Anzeige der Daten erfolgt mit der Software im Steuerungs-Rechner.
Das Aufheizen des DUT auf 300 °C erfolgt in einer Zeit von sechs Minuten. Der Abkühlvorgang verläuft exponentiell durch freie Konvektion mit einer Zeitkonstante von etwa zehn Minuten [11].
Der modulare Aufbau erlaubt eine einfache Anpassung der Testplattform an andere Testobjekte bzw. Testschaltungen und ist als transportable Lösung ausgeführt. Die Visualisierung der Messergebnisse kann durch den verwendeten Einplatinen-Computer Raspberry Pi® weiter ausgebaut werden. Durch den Austausch der entsprechenden Module kann die Testplattform als Demonstrator für verschiedene DUT-ASICs eingesetzt werden.
Auf Grund des modularen Konzepts der Messplattform kann die Hochtemperaturfassung außerdem für andere Gehäusevarianten ausgelegt werden. Weiterhin könnten Hybrid-Schaltungen auf hochtemperaturfesten Keramiksubstraten bei entsprechender Fixierung ebenfalls für Hochtemperaturmessungen temperiert werden. Damit sind die Messmöglichkeiten auf einfache Weise erweiterbar.

Mit der neuen modularen Hochtemperatur-Testplattform wurde der Messbrücken-ASIC D2040A untersucht [12], der bereits im Projekt HoTSens vom IMMS erarbeitet worden war. Dieses Projekt wurde vom Bundesministerium für Bildung und Forschung im Programm „IKT 2020 – Forschung für Innovationen“ unter den Kennzeichen 16ES0008 gefördert.

Literaturnachweis

[1] J. Agostino, Ch. Schweizer: “Hochtemperatur Elektronik,” Auto & Elektronik 3/2004, S 42-43.
[2] P. Rolfes, Hochtemperaturschaltungstechnik, Aachen: Shaker Verlag, 1999. ISBN-13: 9783826563072
[3] St. Richter, W. Göttlich, D. Nuernbergk, V. Nakov,. S. Bormann: „Hitzefest - Design von Mixed-Signal-Hochtemperatur-ICs“, Design & Elektronik, Sonderdruck, März 2002,. S. 56ff
[4] M. Reinhard, U. Liebold, A. Richter, I. Gryl: „Hochtemperatur-Wafertest bis 300°C“, 28. GMM/GI/ITG Workshop, TuZ 2016, 06.03.2016 - 08.03.2016, Siegen, Germany
[5] R. Klieber; R. Lerch: „Evaluation of materials for high temperature IC packaging“, 15th International Workshop on Thermal Investigations of ICs and Systems, 2009, page 117-120.
[6] Reliability System Test SOLUTIONS Inc., „RS Solutions,“ http://rssolutionsinc.com/high-temperature-sockets/. (31 Januar 2017])
[7] Multi-CB Leiterplatten GmbH: „Hoch-TG-Leiterplatten (HTG)“, „https://www.multi-circuit-boards.eu/produkte/leiterplatten/hoch-tg.html“ (18.01.2018)
[8] INGUN Prüfmittelbau GmbH, „ingun,“ http://www.ingun.com (24.04.2019)
[9] Janis Research Company, LLC, „Janis - Products - CryogenicAccessories“, https://www.janis.com/Products/AccessoriesandAncillaryEquiment/CryogenicAccessories.aspx. (20. 04. 2019).
[10] G. Gläser, D. Kirsten, A. Richter, M. Reinhard, G. Kropp, D. Nuernbergk: “High-Precision Mixed-Signal Sensor Interface for a Wide Temperature - Range [0°-300°C]”, High Temperature Electronics Network (HiTEN 2017), S. 36-41, Cambridge, UK, July 10-12, 2017
[11] T. Reinhold: “Konzeption und Realisierung eines Demonstrator-Aufbaus für einen integrierten Hochtemperaturschaltkreis”, BA TU Ilmenau 2017
[12] B. Bieske. T. Reinhold, M. Reinhard,: „Modulare Hochtemperatur-Testplattform bis 300 °C“, 30. GMM/GI/ITG Workshop Test und Zuverlässigkeit, TuZ 2018, 04.03.2018 - 06.03.2018, Freiburg (Breisgau), Germany

DOI 10.5162/sensoren2019/2.4.4

Impedanz-basierter NO_x-Sensor für Abgasanwendungen

Gunter Hagen[1], Andreas Müller[1], Dirk Bleicker[2], Frank Noack[3], Julia Lattus[1], Jaroslaw Kita[1], Ralf Moos[1]
[1]Bayreuth Engine Research Center (BERC), Lehrstuhl für Funktionsmaterialien, Universität Bayreuth, D-95440 Bayreuth, Germany
[2]CARIT Automotive GmbH&Co. KG, Gildenstr. 18, D-48157 Münster, Germany
[3]CPK Automotive GmbH & Co. KG, Gildenstr. 4c, D-48157 Münster, Germany

Zusammenfassung

Die Messung der Stickoxid-Konzentration in automobilen Abgasen erfordert robuste aber auch kostengünstige Lösungen. Als Alternative zu kommerziellen amperometrischen Sensoren auf Zirkon-Oxid-Basis wird vorgeschlagen, die Impedanz eines Funktionsmaterials zu messen. Die einfach herzustellenden planaren Sensoren zeigen bei einer Betriebstemperatur von 650 °C eine schnelle und ausreichend hohe Reaktion auf NO. Allerdings ist die NO-Empfindlichkeit von der Grundgas-Atmosphäre abhängig. Im Labor wurden deshalb mehrere Kennlinien bei variablem Lambda-Wert unter gleichzeitiger Anpassung des Sauerstoff-, Kohlendioxid- und Feuchte-Gehalts aufgenommen. Legt man dieses Kennfeld (NO_x-Konzentration als Funktion der Sensorantwort unter gleichzeitiger Berücksichtigung des Lambda-Wertes) bei Messungen im realen Diesel-Abgas zu Grunde, so ergibt sich eine gute Übereinstimmung mit Messwerten eines kommerziellen Sensors.

Keywords: Impedanz, NO_x-Sensor, Diesel-Abgas, Lambda

Motivation

Abgastaugliche NO_x-Sensoren zur Überwachung und / oder Regelung von Diesel-betriebenen Fahrzeugen sind notwendig, um einen effizienten Betrieb des jeweiligen Abgasnachbehandlungssystems zu gewährleisten und die Gefahren für Mensch und Umwelt durch zu hohe Stickoxid-Emissionen zu verhindern. Kommerzielle, amperometrische NO_x-Sensoren auf ZrO_2-Basis sind aufwändig in der Herstellung und es müssen sehr kleine Messsignale verarbeitet werden. Im vorliegenden Beitrag wird alternativ ein Impedanz-basierter NO_x-Sensor vorgestellt, der diese Nachteile überwinden kann.

Experimentelles

Grundlage dieses neuartigen Sensors ist die Änderung der elektrischen Impedanz einer Funktionsschicht unter NO_x-Beaufschlagung. Das gassensitive Funktionsmaterial besteht dabei aus $KMnO_4$ auf einem Al_2O_3-Träger (imprägniert und kalziniert) und liegt als siebgedruckte Dickschicht auf einer interdigitalen Elektrodenstruktur vor [1]. Das Sensorelement ist auf einem Al_2O_3-Substrat aufgebaut und beinhaltet auf der Rückseite eine Dickschicht-Heizerstruktur in Vierleiter-Technik, mit deren Hilfe die gewünschte Betriebstemperatur (hier ca. 650 °C) erreicht und gehalten wird. Ein abgastaugliches Gehäuse mit entsprechender Schutzkappe erlaubt den Einbau in Labor- oder Abgasanlagen. Messwerte werden bei 100 kHz als Betrag $|\underline{Z}|$ und Phase φ der Impedanz aufgenommen und dann in einen Widerstandswert R unter Annahme eines $R||C$-Ersatzschaltbildes umgerechnet.

Ergebnisse im synthetischen Laborabgas

Zunächst wurden Kennlinien des Sensors im Labor in Abhängigkeit der NO-Konzentration (Variation von 0 bis 1500 ppm) und von Lambda aufgenommen (Abbildung 1). Das Grundgas bestand dabei jeweils aus O_2, CO_2 und H_2O in N_2 und wurde für die Lambda-Einstellung entsprechend angepasst. Die Empfindlichkeit des Sensors (dargestellt in Werten der relativen Änderung des gemessenen Widerstandswertes R mit NO bezüglich des Grundwiderstands R_0 im Grundgas ohne NO-Zudosierung) sinkt mit sinkender Sauerstoffkonzentration im Abgas.

Dieses Sensorverhalten (verschiedene Kennlinien in Abhängigkeit von Lambda) wurde in den späteren Realabgasmessungen zur Berechnung der NO_x-Konzentration aus dem Messsignal herangezogen. Lambda-Werte wurden jeweils mit einer handelsüblichen Breitband-O_2-Sonde (LSU) und entsprechender Elektronik (ETAS) bestimmt.

DOI 10.5162/sensoren2019/2.4.4

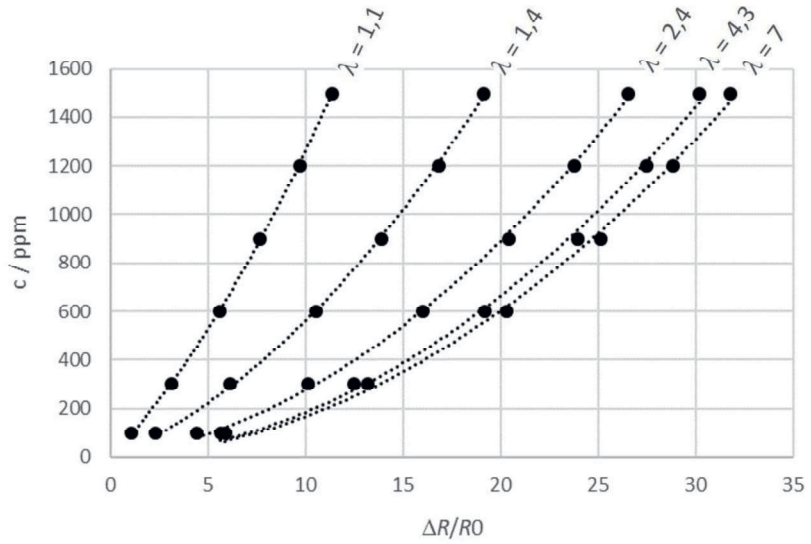


Abb. 1: Kennlinien des Impedanz-basierten NO_x-Sensors in Abhängigkeit von Lambda, gemessen in synthetischem Abgas (Labor-Anlage). Das Grundgas wurde in den Werten für O_2, CO_2 und H_2O entsprechend angepasst, um verschiedene Lambda-Werte (Messwerte eines kommerziellen O_2-Sensors) einzustellen. Aufgetragen ist die relative Änderung des gemessenen Widerstands.

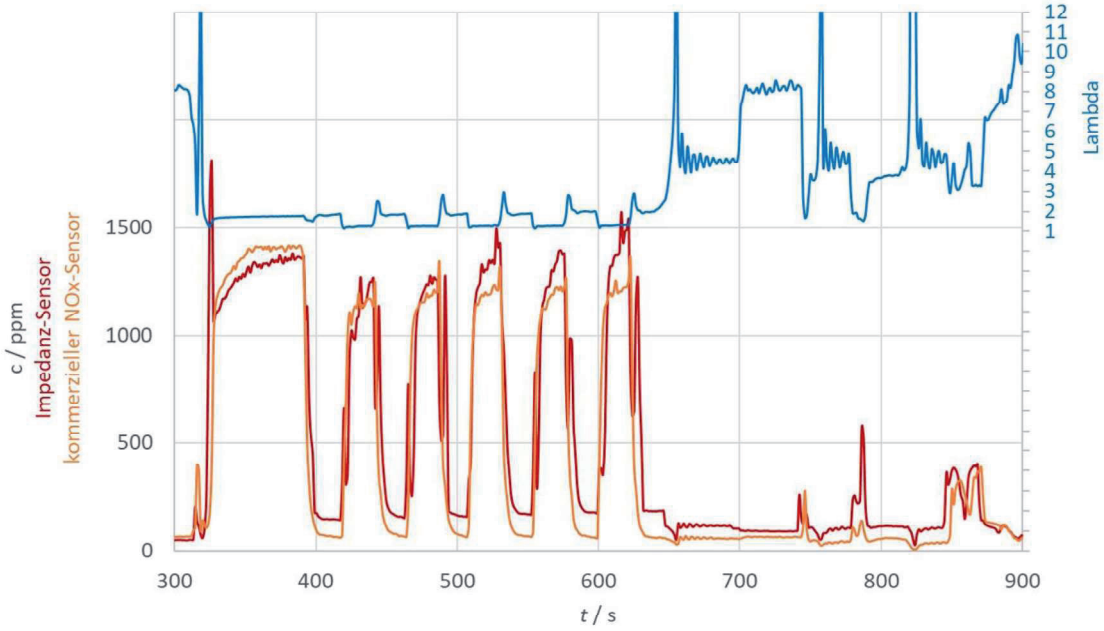


Abb. 2: Realabgas-Messung im Hauptstrom eines Diesel-Vierzylinders bei verschiedenen Betriebsbedingungen / Abgasatmosphären: Zeitlicher Verlauf der Signale des Impedanz-Sensors (zu einer Konzentration umgerechnete Widerstandsänderung) und kommerzieller NO_x-Sensor im Vergleich.

Ergebnisse im realen Abgas

Um die Tauglichkeit des Sensorelements im realen Einsatzfall nachzuweisen, wurden Messungen im Dieselabgas eines 2,1 l – Vierzylindermotors (Hauptstrom) im direkten Vergleich mit einem kommerziellen NO_x-Sensor durchgeführt (Abbildung 2). Verschiedene Abgasbedingungen wurden dabei durch Änderungen im Motorbetrieb bzgl. Drehzahl und Last herbeigeführt. Die zeitlichen Verläufe beider Sensorsignale zeigen schnelle Reaktionen und eine stabile Grundlinie. Die gemessenen Konzentrationen stimmen im Wesentlichen gut überein. Leichte Abweichungen ergeben sich bei der zeitlichen Anpassung der Datensätze und damit auch bei der Berechnung der Konzentration aus dem Impedanz-Signal unter Berücksichtigung von Lambda.

Zusammenfassung

Insgesamt ist das Ergebnis aber als sehr vielversprechend zu beurteilen und eröffnet damit Möglichkeiten zur weiteren Entwicklung kostengünstiger und abgastauglicher Hochtemperatur-NO_x-Sensoren.

Literaturnachweis

[1] A. Groß, M. Kremling, I. Marr, D. Kubinski, J. Visser, H. Tuller, R. Moos, Dosimeter-Type NOx Sensing Properties of KMnO4 and Its Electrical Conductivity during Temperature Programmed Desorption, *Sensors* 13(4), 4428-4449 (2013); doi: 10.3390/s130404428.

DOI 10.5162/sensoren2019/3.1.1

Aktuelle Entwicklungen von MEMS IR-Sensorkomponenten

Martin Schädel[1], Julia Baldauf[1], Li Long[1], Jan Freitag[1], Nicole Thronicke[1], Neckermann Kristin[1], Sabine Nieland[1], André Grün[1], Indira Käpplinger[1], Thomas Ortlepp[1]
[1]CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Deutschland

Zusammenfassung

Dieser Beitrag gibt einen kurzen Überblick zu der Thematik MEMS-basierter IR-Gassensoren und Sensorkomponenten und berichtet ausführich über den Stand aktueller Entwicklungen des CiS Forschungsinstitutes in diesem Themenfeld. Es werden verschiedene, auf thermischer Emission basierende IR-Strahler vorgestellt, die sich z.B. durch herausragenden Minaturisierungsgrad oder Wafer-Level-Packaging hervorheben. Als Alternative werden auch Ergebnisse von IR-Strahlern präsentiert, die durch Drucktechniken auf dünnen Membranträgern erzeugt werden.
Auf Detektionsseite kommen Thermopiles zum Einsatz, die ebenfalls auf Waferebene hermetisch gehaust werden können. Der Siliziumdeckel kann neben der mechanischen Schutzfunktion dabei gleichzeitig als Träger für optische Filterschichten oder siliziumintegrierte Infrarot-Optiken dienen. Im Beitrag wird ein Einblick in die eingesetzten Technologien gegeben und es werden die bisher erreichten Bauteileigenschaften vorgestellt.

Keywords: Gassensorik, optische Messverfahren, Infrarotemitter, Thermopiles, Packaging

Motivation

Infrarotsensoren gewinnen zunehmend an Bedeutung, inbesondere in Form von Gassensoren im Bereich der Medizintechnik und Umweltsensorik [1,2]. Ermöglicht wird dies vor allem durch die zunehmende Miniaturisierung, die neben dem Energiebedarf auch die Kosten pro Baugruppe weiter reduziert. Hinzu kommt ein wachsendes Bewustsein der Gesellschaft für gesundheitliche und umweltliche Einflussfaktoren. Verschiedene am CiS Forschungsinstitut für Mikrosensorik durchgeführte Forschungsprojekte zielen darauf ab Technologien für die Fertigung von Infrarotkomponenten zu entwickeln, die diesen Ansprüchen gerecht werden.

Miniaturisierte Infrarotemitter

Infrarotemitter basierend auf einer leitfähigen strukturierten Schicht, welche auf eine Membran aufgebracht oder eingebettet ist, werden vielfach in der Literatur beschrieben [3, 4,5,6]. Diese Infrarotemitter können durch eine an Stegen aufgehängte Membranstruktur [3,4] oder durch eine flächige Membran mit Öffnungen [5] oder ohne Öffnungen [6] ausgeführt werden. Zumeist wird für den Widerstandsheizer eine Platinschicht verwendet [3,4,7].
Durch die Verwendung von Platin als Heizelement kann der Infrarotemitter nicht vollständig in einem CMOS-kompatiblen Prozess gefertigt werden, wodurch die Kosten gegenüber einem vollständig in CMOS-kompatiblen Dünnschichtprozessen gefertigter Infrarotemitter erhöht sind.
Die miniaturisierten Infrarotemitter des CiS Forschungsinstitutes für Mikrosensorik bestehen aus einer dielektrischen Membran in welche eine leitfähige Schicht aus einer leitfähigen Keramik eingebettet ist, welche als Widerstandsheizer dient und ein Aufheizen der Membran um mehrere hundert Grad ermöglicht. Abb. 1 zeigt beispielhaft eine thermische Simulation eines miniaturisierten Infrarotemitters mit einem derartigen Membranstapel unter Betriebsbedingungen. Die erreichbare Temperatur und der Widerstand des Bauteiles können durch eine Strukturierung der leitfähigen Schicht verändert werden.
Die Fertigung der Emitter erfolgte durch die Abscheidung der dielektrischen Schichten und der leitfähigen Schicht auf einen Siliziumwafer und der anschließenden rückseitigen Strukturierung des Siliziums mit Hilfe eines Inductive Coupled Plasmaätzverfahrens. Durch die Ausnutzung dieses Ätzverfahrens können

nahezu beliebige Membrangeometrien realisiert werden (vgl. Abb. 2).

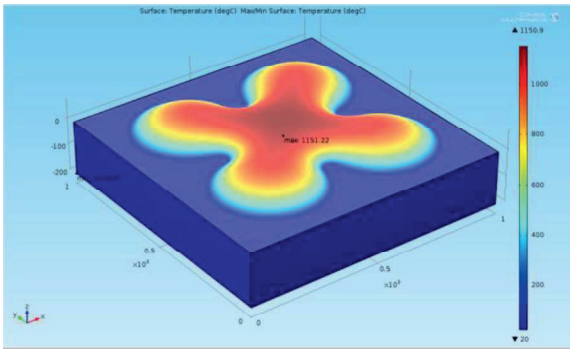


Abb. 1: Thermische Simulation eines miniaturisierten Infrarotemitters mit 1 mm Kantenlänge.

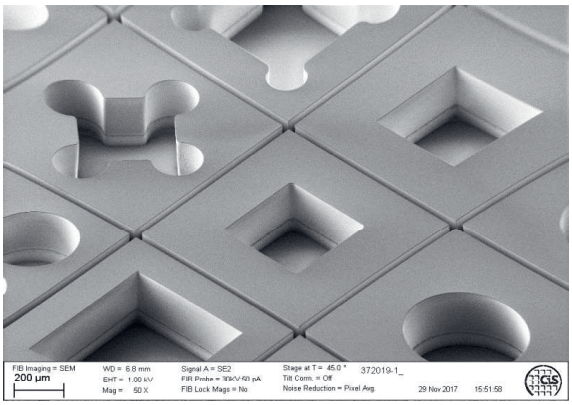


Abb. 2: Rasterelektronenmikroskopaufnahme der Rückseiten der freigestellten Membranen verschiedener Versuchsgeometrien.

Druckbare Materialsysteme für Infrarotemitter

Alternativ zu den Infrarotemittern basierend auf Dünnschichttechnologien untersucht das CiS auch Dickschichttechnologien für die Fertigung von Widerstandsheizerstrukturen für Infrarotemitter.

Bei den untersuchten Technologien handelt es sich vornehmlich um den Siebdruck von leitfähigen Pastensystemen (Silberpasten und Platinpasten). Aufgrund der CMOS-Inkompabilität kann der Druck hier esrt nach der Mikrostrukturierung des Siliziumträgers erfolgen.

Zunächst wird eine Passivierung in einem Dünnschichtprozess auf Siliziumwafern abgeschieden, anschließend werden Siliziummembranen mittels KOH-Ätzen von der Rückseite erzeugt.

Diese Pasten werden mittels Siebdruck auf die Membranen appliziert und mit speziell an die Bauteile angepassten Temperaturprofilen gesintert. Eine derartige gedruckte Widerstandsheizerstruktur bestehend aus einer Platinpaste ist in Abb. 3 gezeigt.

Unter Betrieb erhitzen die gedruckten Widerstandsheizerstrukturen die Membran auf mehrere hundert Grad (vgl. Abb. 4) und senden in einem breiten Wellenlängenbereich infrarote Strahlung aus (vgl. *Abb. 5*). Die in Abb. 5 zu erkennenden Interferenzeffekte sind von der Dicke der verwendeten Siliziummembranen abhängig.

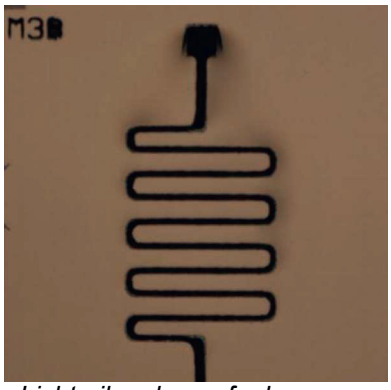


Abb. 3: Lichtmikroskopaufnahme einer siebgedruckten Platinpaste als Widerstandsheizerstruktur.

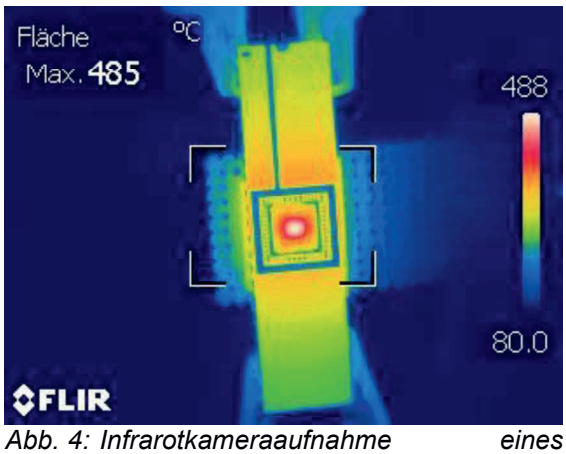


Abb. 4: Infrarotkameraaufnahme eines gedruckten Strahlers im Betrieb.

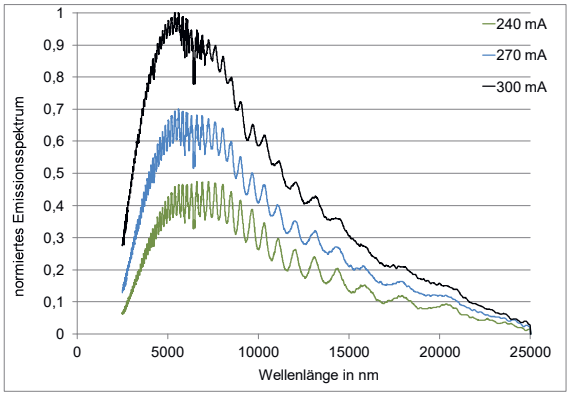


Abb. 5: Normierte Emissionsspektren eines gedruckten Strahlers mit Betrieb bei verschiedenen Stromstärken.

Thermopiles

Thermopiles werden in vielen Anwendungen als Detektoren verwendet. Diese Anwendungen umfassen unter anderem die nicht-dispersive Infrarotgasanalyse [8,9] und die thermoelektrische Energieerzeugung [10]. Thermopilesensoren bestehen aus einer Reihenschaltung von Thermoelementen. Diese Thermoelemente wiederum werden dadurch realisiert, dass zwei thermoelektrisch verschiedene Materialien in elektrischen Kontakt gebracht werden. Ein Ende jedes Thermoelementes befindet sich auf dem Siliziumrahmen, welcher zur Kühlung dieses Endes führt, das andere Ende liegt auf einer Membran, welche zu einer thermischen Entkopplung dieses Endes führt. Bei Bestrahlung mit Licht erwärmt sich dieses Ende des Thermoelementes. Durch die Temperaturdifferenz zwischen den beiden Enden des Thermoelementes wird eine Spannung erzeugt. Die Spannung des gesamten Thermopiles ist von der Spannung der einzelnen Thermoelemente und somit von der Temperaturdifferenz, der Anzahl der Thermoelemente und der verwendeten Materialien abhängig.

Thermopiles werden am CiS Forschungsinstitut basierend auf verschiedenartig dotiertem polykristallinen Silizium erzeugt.

Die Verwendung des Materiales Silizium hat einige Vorteile, wie einen großen Seebeckkoeffizienten und vergleichsweise geringe Wärmeleitfähigkeit des polykristallinen Siliziums. Ein weiterer Vorteil ist die breite Verteilung und das gute Verständnis der Polysilizium-Dünnschichttechnologie, durch die es möglich ist große Thermowiderstände zu realisieren, in massentauglichen CMOS-Prozessen zu fertigen und Thermopiles kostengünstiger zu erzeugen als dies mit einer Technologie basierend auf Bismut-Telur-Legierungen oder Bismut-Antimon-Legierungen der Fall wäre. Darüber hinaus sind auf polykristallinem Silizium basierende Thermopiles hochtemperaturstabiler als Bismut-Antimon-basierende und Bismut-Telur-basierende Thermopiles und können ein kürzeres Ansprechverhalten erreichen.

Als Substratmaterial für die Dünnschichtabscheidungen dienen einkristalline Siliziumwafer mit einer <100> Kristallorientierung. Auf diese wird zunächst ein dielektrischer Schichtstapel aufgebracht, anschließend wird eine erste polykristalline Siliziumschicht aufgebracht, dotiert und strukturiert. Daran anschließend wird eine weitere dielektrische Schicht aufgebracht und es erfolgt eine Strukturierung, so dass ein Kontakt zur nachfolgend abgeschiedenen dotierten polykristallinen Siliziumschicht ermöglicht wird. Diese zweite polykristalline Schicht wird mit einer weiteren dielektrischen Schicht vor Umwelteinflüssen geschützt. Die Freistellung der Membran erfolgt durch eine Kombination aus KOH-Ätzprozess mit einem anschließenden Inductive Coupled Plasmaätzprozess.

Die wichtigsten Schritte des Prozessablaufes der Fertigung eines Thermopiles sind schematisch in Abb. 6 dargestellt und ein Beispiel für einen gefertigten Thermopile ist in Abb. 7 zu sehen. Die entstehenden Thermopiles weisen eine sehr dünne Membran einer Dicke von weniger als 2 µm auf.

Um diese Membranen zu schützen kann eine Verkapselung auf Modulebene oder Waferebene erfolgen.

Darüber hinaus werden in laufenden Forschungsvorhaben durch eine verringerte Membrangröße und eine integrierte Optik (vgl. Abb. 8) Thermopilesensoren mit einer hohen Messfrequenz basierend auf CMOS kompatiblen Prozessen entwickelt.

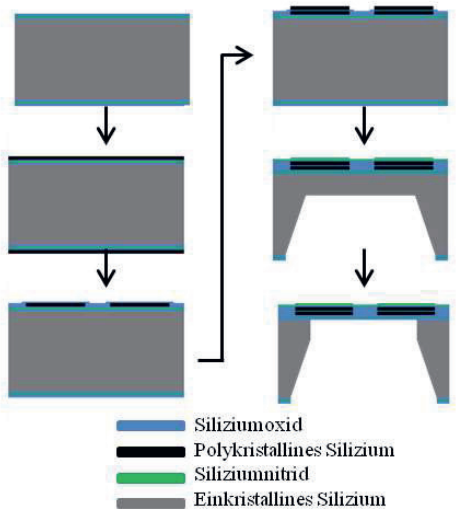


Abb. 6: Schematische Darstellung des Prozessablaufes der Fertigung eines Thermopiles.

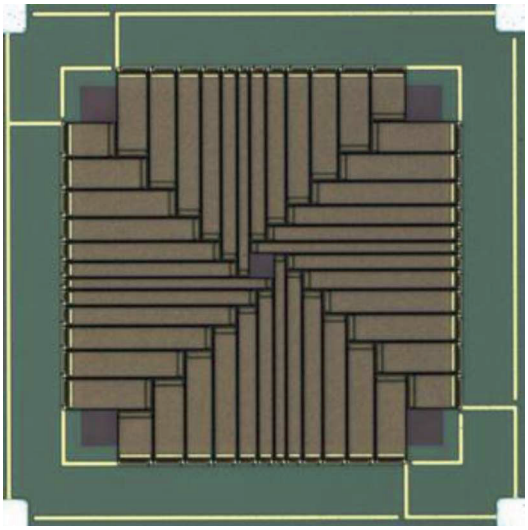


Abb. 7: Lichtmikroskopaufnahme eines Thermopiles.

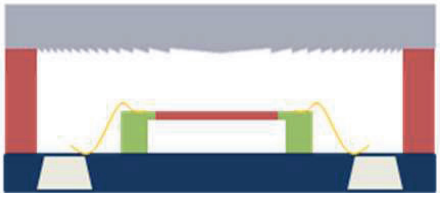


Abb. 8: Schematische Darstellung eines Thermopiles mit integrierter Optik

Waferlevelpackaging von Infrarotkomponenten

Da für einige Anwendungen eine hermetische Verkapselung der Thermopiles vorteilhaft oder zwingend notwendig ist, wurden auch Untersuchungen hinsichtlich des Packaging von Thermopilesensoren durchgeführt.
Gegenüber zu kommerziell erhältlichen Thermopiles auf Basis einer Bismut-Antimon-Legierung bieten die Siliziumbasierten Thermopiles dabei den entscheidenden Vorteil auch Temperaturen von 450°C beim Packaging mit Hilfe von Glaslot zu ohne signifikante Änderungen überstehen.
Für die Verkapselung wurde das Glaslot sowohl auf den Substratwafer (Träger), als auch auf den Kappenwafer gedruckt, was ein Waferlevelpackaging der in Abb. 9 gezeigten Komponenten (Substratwafer, Thermopile und Kappenwafer) zu einem dichten verkapselten Thermopile ermöglichte.
Da Infrarotemitter im Allgemeinen höheren Temperaturen standhalten können als Thermopiles, kann mit der für Thermopiles entwickelten Technologie problemlos auch ein Packaging der Infrarotemitter durchgeführt werden.

Sowohl Thermopilesensoren, als auch Infrarotemitter wurden bereits erfolgreich im Waferverbund verkapselt.



Abb. 9: Lichtmikroskopaufnahmen der für die Verkapselung notwendigen Komponenten und eines auf Waferebene gekapselten Thermopiles.

Photonische Kristalle

Um die optischen Eigenschaften eines Infrarotemitters einzustellen, können photonische Kristalle verwendet werden.
Als photonischer Kristall wird ein Festkörper mit einem periodischen Verlauf des Brechungsindexes bezeichnet. Photonische Kristalle können durch Strukturieren von bulk-Materialien oder durch die Erzeugung eines Metall-Dielektrikum-Metall-Schichtstapels realisiert werden. Die Größe, Form, Gitteranordnung der Strukturelemente und Materialparameter und Dicken der Schichten nehmen einen Einfluss auf die spektralen Eigenschaften des photonischen Kristalles.

Auch mit Hilfe von strukturiertem Silizium können photonische Kristalle erzeugt werden. Am CiS werden diese photonischen Kristallstrukturen zunächst als Lochstrukturen in Silizium realisiert. Dazu wurden umfangreiche Simulationen mit einem kommerziell erhältlichen Programm (VirtualLab von LightTrans International UG (www.lighttrans.com) Built 7.0.035.) durchgeführt.
Da nach dem Kirchhoffschen Gesetz im thermischen Gleichgewicht die Emission eines Materials seiner Absorption entspricht, kann für die Bestimmung der möglichen Emission einer

Struktur eine Simulation der Absorption herangezogen werden.

Als Strukturen wurden kreisförmige Löcher verschiedener Durchmesser in quadratischen Gittern mit verschiedener Periode (Abstand Kreismitte zu Kreismitte) in verschiedenartig dotiertem Silizium mit unterschiedlichen Dicken simuliert. Außerdem wurden mit Hilfe der Simulationen Toleranzanalysen hinsichtlich des Einflusses von Strukturfehlern durch den Ätzprozess untersucht.

Abb. 10 zeigt Simulationen von 12 µm dicken Siliziumschichten mit 10^{20} Phosphoratomen pro cm^3, einer Lochstruktur (Durchmesser 10 µm) und verschiedener Periode.

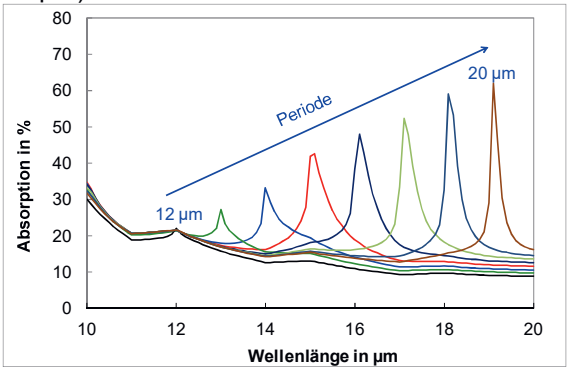


Abb. 10: Simulierte Absorption verschiedener photonischer Kristalle aus Silizium.

Danksagung

Die Forschungs- und Entwicklungsarbeiten für die vorgestellten Infrarotkomponenten wurden im Rahmen mehrerer Projekte gefördert

-Miniaturisierter InfraRot-Emitter-Chip (MIREC), InnoKom MF160085

-Evaluierung druckbarer Materialsysteme für IR-Emitter, InnoKom VF160028

-Schneller Infrarotsensor mit integrierter Optik, InnoKom 49MF180032

-Entwicklung einer CMOS-kompatiblen Thermopile-Infrarotsensoren in Stapeltechnologie für Gasdetektoren, InnoKom MF150013

-Sipamig System-in-Package für einen miniaturisierten Gassensor bis 200°C, ZIM ZF4000401WM5

-Photonische Kristalle für die InfraRot Gasanalyse (PhIR), InnoKom 49MF180072

gefördert durch das Bundesministerium für Wirtschaft und Energie.

Literaturnachweis

[1] IVAM-Befragung 2017 – Europas Mikrotechnikbranche stellt sich auf Wachstum ein https://www.ivam.de/research/economic_data/europe_s_microtechnology_industry_signals_growth (abgerufen am 18. Jan. 2019)

[2] B. Roussel, E. Mounier; Yole Dévelopement: Gas Sensor Technology and Market, www.yole.fr, (Feb.2016)

[3] J. Hildenbrand, J. Korvink, J. Wöllenstein, C. Peter, A. Kürzinger, F. Naumann, M. Ebert, and F. Lamprecht, IEEE Sensors Journal 10, 353-362 (2010); doi: 10.1109/JSEN.2009.2034624

[4] J. Hildenbrand, C. Peter, F. Lamprecht, A. Kürzinger, F. Naumann, M. Ebert, R. Wehrspohn, J. G. Korvink, J. Wöllenstein, Microsystem Technologies 16, 745-754 (2010); doi: 10.1007/s00542-010-1049-1

[5] P. Zhou, R. Chen, N. Wang, H. San, and X. Chen, Micromachines 7, 166-172 (2016); doi: 10.3390/mi7090166

[6] H. San, X. Chen, M. Cheng, F. Li, Proceedings of SPIE 6836, 68360N1-68360N-8, (2007); doi: 10.1117/12.755923

[7] J. Spannhake, O. Schulz, A. Helwig, A. Krenkow, G. Müller, and T. Doll, Sensors 6, 405-419 (2006); doi: 10.3390/s6040405

[8] S. Udina, M. Carmona, G. Carles, J. Santander, L. Fonseca, S. Marco, Sensors and Actuators B: Chemical 134, 551–558 (2008); doi: 10.1016/j.snb.2008.05.043

[9] A. Roncaglia, F. Mancarella, G.C. Cardinali, Sensors and Actuators B: Chemical 125, 214-223, (2007); doi: 10.1016/j.snb.2007.02.018

[10] V. Leonov, T. Torfs, P. Fiorini, and C. Van Hoof, IEEE Sensors Journal 7, 650-657 (2007); doi: 10.1109/JSEN.2007.894917

DOI 10.5162/sensoren2019/3.1.2

High sensitivity 80x64 Thermopile infrared array sensors

Schieferdecker, J.; Schnorr, M.; Forg, B.; Simon, M.; Leneke, W.; Herrmann, F.; Schmidt, C.;
Heimann Sensor GmbH, Maria-Reiche-Str. 1, D-01109 Dresden, Germany, Tel. +49351-888885-0;
info@heimannsensor.com;

Abstract

Heimann Sensor introduced in 2016 the first monolithic 80x64 thermopile array sensor. Since the market shows an increased demand of low cost infrared sensing solutions with improved performance Heimann presents an update of the 80x64d in this year.

Key words: Thermopile array, 80x64 pixels, infrared detection, true shutterless, medium resolution

Introduction

Since the introduction of the first fully monolithic thermopile array Heimann Sensor consequently broadens its array portfolio [1].

Already in 2016 Heimann introduced the first digital 80x64 arrays [2]. In recent years the first array types were completed to a full family of digital thermopile arrays from 8x8, 16x4, 16x16, 32x32 and 80x64.

Since thermopile solutions are very insensitive against environment changes (such as ambient temperature) it is not necessary to use offset correcting measures, such as shutters and / or thermal stabilizers such as TECs. Also the sensitivity has a very low temperature coefficient (TC), which can be considered in calibration without any further effort, The competitive extended object temperature measurement range allows not only usage in the “classic” infrared imaging applications such as hand held thermal imagers, security cameras remote temperature measurement, it in fact enables also usage in industrial process control for very high temperatures and hotspot detection. The competitive price of the thermopile arrays also allows entrance to the high volume markets such as HVAC, person detection, energy management and automotive.

The low power consumption of these systems renders also battery powered applications possible. The arrays are further offered with a wide range of optics to adapt to various applications and additional measures to increase sensitivity and fill factor.

Functional description of the digital thermopile array sensors

The 80x64 sensor has several internal amplifiers (OPA) for signal conditioning and Analog to Digital Converters (ADC) with a resolution of 16 bit for digitizing.

In addition two proportional to ambient temperature circuits (Vtemp) are placed on the chip to measure the ambient temperature. The architecture of the 80x64 array chip was described in [2].

Readout of the 80x64 sensor

Thanks to the internal 16 bit ADC and an integrated EEPROM inside the TO package with optics, the complete array can be provided as fully calibrated array with a six pin package. The calibration data and the pixel data stream can be read via the integrated serial peripheral interface (SPI).

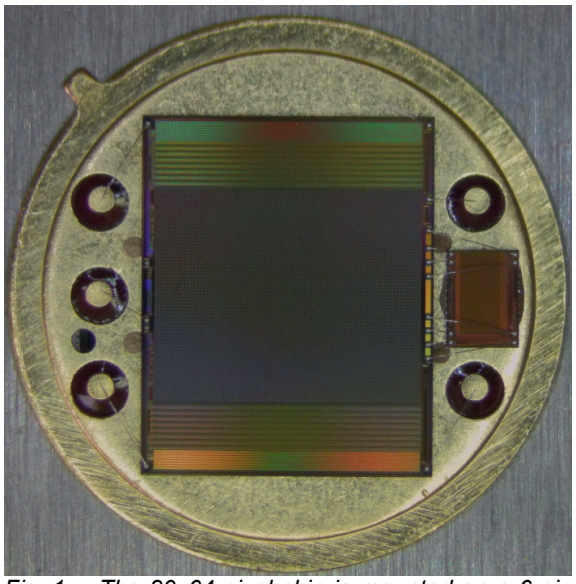


Fig. 1. The 80x64 pixel chip is mounted on a 6 pin TO8 header together with an EEPROM chip. The optics is mounted inside the TO cap.

Since the signal from the pixel decreases with smaller pixel sizes, there are different options to maintain a good signal to noise ratio (SNR):

- The use of a special chip design with thermally isolated membrane from the Si heat sink by thin beams
- Reducing the heat conductivity of the gas around the pixel and
- Enhancing the fill factor by additional absorber structures (so called umbrellas!)

Fig. 2 a) shows a thermopile array pixel with a thin reticulated membrane 4, which is suspended on two beams 3 between the Si frame (substrate) 2. The beams 3 are carrying the thermoelectric sensitive materials n- and p-type Poly-Si as thermocouples. They are isolated by slots 7 from the membrane. The Si frame is also acting as a heat sink. The hot junctions 9 are located on the suspended membrane and the cold junctions 8 are at the heat sink. The membrane is covered by an IR absorbing layer 4.

In Fig. 2 b) such array pixel is depicted as cross sectional view. For enhanced signal the fill factor can be increased by adding an umbrella type radiation collector. In Fig. 2 c), the much bigger absorbing area of the umbrella compared to the area of the absorber on the membrane only is clearly shown.

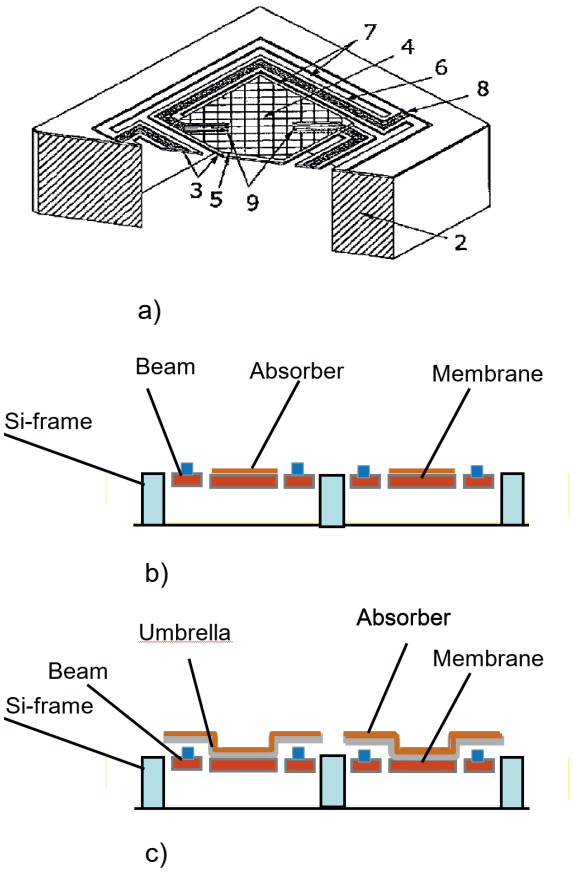


Fig. 2. Thermopile array pixel with reticulated membrane, suspended on thin beams, a) in a 3D view, b) as cross sectional view with a normal absorber on the membrane and c) with additional umbrella type radiation collector covered with additional absorber layer.

The realization of thermopile pixels with absorber layer can be seen in Fig. 3:

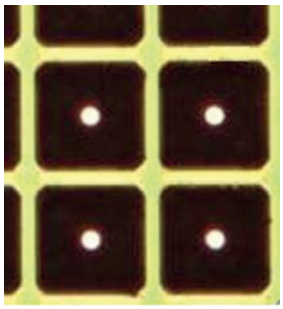
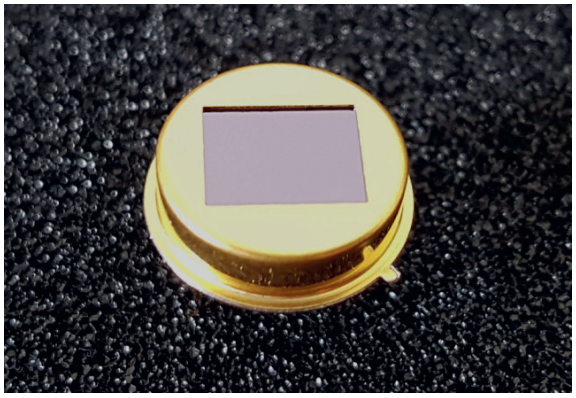


Fig 3.: Thermopile array pixels covered by umbrella and absorber.

The array chip is mounted onto a TO-8 header inside a metal can, which is hermetically sealed to maintain a Xe atmosphere. Xe is used due to its about 4.5 x lower heat conductivity compared to air or N2.

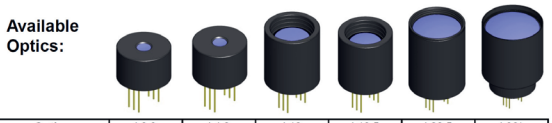
A further signal improvement can be reached by sealing the TO-8 can under vacuum atmosphere). For that purpose we soldered the

lens or entrance filter to the metal can for reaching stable hermeticity.



Optical design of the 80x64 sensor

There are currently 6 different optics designed and available for this sensor. The field of view (FoV), the F-number and optics dimensions are listed in Fig. 4.



Optic	L3.9	L4.8	L10	L10.5	L22.5	L33*
FoV [°]	115 x 90	90 x 72	41 x 33	39 x 31	18 x 14	12 x 9
Length of cap [mm]	12.6	14.4	25.7	24.1	36.5	46
Diameter of cap [mm]	20	20	23	23	28	37
F-number	0.8	0.8	0.7	0.95	1.0	1.05

Fig. 4. The sensor equipped with different optics.

Unlike the smaller arrays (8x8, 16x16 and most 32x32), where a single lens optics is sufficient, for 80x64 arrays two lens optics are necessary for good optical performance.
While the 80x64 thermopile array is in mass production, we are planning for an 120x84 array. Thanks to smaller pitch of 60 µm, the chip size is similar to the 80x64, so that also the 120x84 fits in same TO-8 housing as the 80x64.

Measurement results and settings

The frame rate can be chosen up to 22 Hz with the full 16 bit resolution and can achieve >100 Hz for a lower ADC resolution.

Maximum possible frame rate is reached with an SPI-clock of 20 MHz for the full readout of the sensor

The NETD of a vacuum packaged 80x64d array can reach 19 mK@1Hz@25°C with the latest Pixel design and the best optics.

NETD values reach 150….300 mK for the standard version depending on optics.

Special optics allow to measure temperatures up to 1000 °C and above.

In the following Figures we show thermal images taken with our application set using the highest resolution thermopile arrays with vacuum sealing:

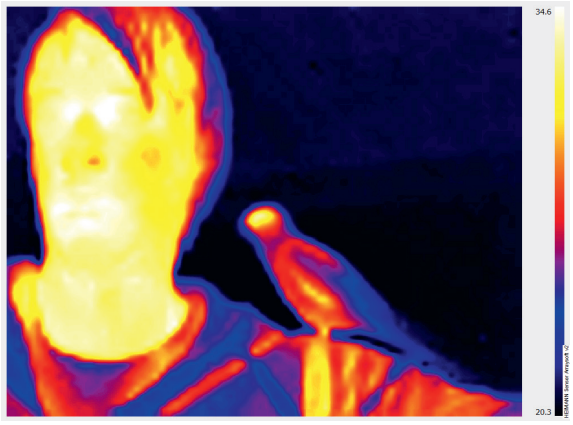


Fig. 5. Thermal image with 80x64d array showing a bird sitting on a human´s shoulder.

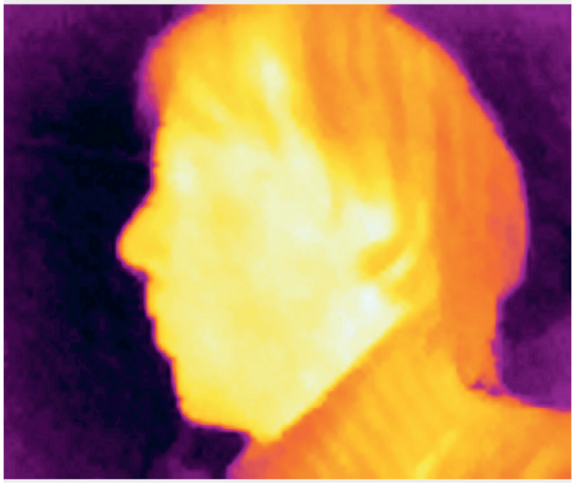
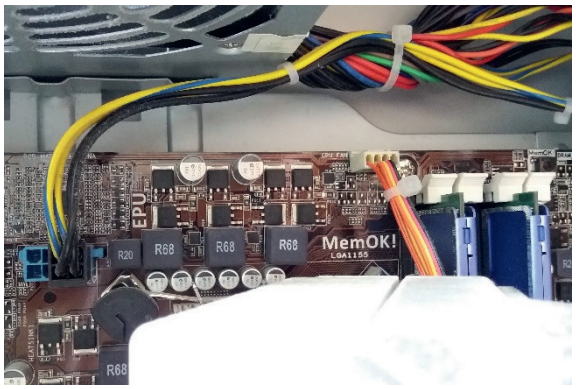
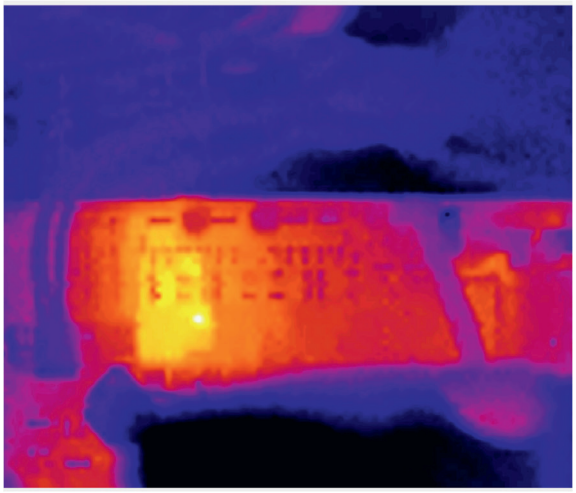


Fig. 6. Another thermal image showing a person from the side.



7 a)



7 b)



7 c)

Fig. 7. Thermal images of a CPU showing hot spots, saved with different color scalings: a) visual image, b) false color thermal image and c) black and white thermal image.

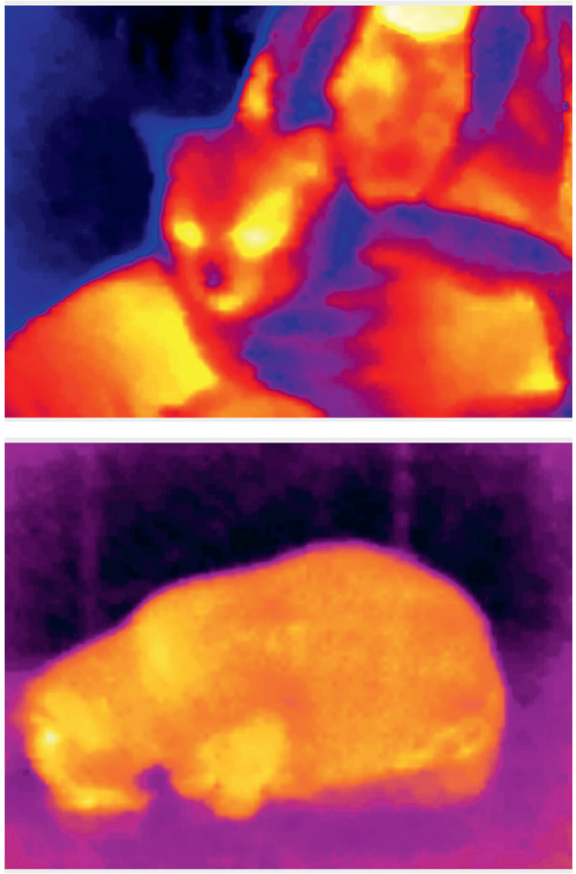


Fig. 8. Thermal images of a cat sitting on a person and eating on the floor.

Application set

For customer´s application support we developed a further improved application set allowing thermal imaging direct to any connected PC. The new software allows to show simultaneously the thermal videos of up to 8 IR modules in parallel on a PC or larger screen.



References

[1] W. Leneke, J. Schieferdecker, M. Schulze, M. Simon, K. Storck, B. Forg, F. Völklein: “Thermopile Linear Array Sensors and Modules with Short Time Constant and High Sensitivity”; *Sensor Conference 2007*, Proc. Vol.1, Nürnberg, Germany, pp. 329-333

[2] M. Schnorr, B. Forg, Dr. M. Simon, Dr. W. Leneke, F. Herrmann, C. Schmidt, Dr. J. Schieferdecker: “New 80x64 Thermopile Arrays”; *Sensor & Test 2016*, Proceedings, Nürnberg, Germany, pp..

[3] B. Forg, W. Leneke, J. Schieferdecker, M. Schulze, M. Simon, K. Storck: “Thermopile Sensor Arrays with Internal Amplifiers and Digital Out”; *Sensor & Test 2008*, Proceedings, Nürnberg, Germany, pp. 249-252.

DOI 10.5162/sensoren2019/3.1.3

Infrared-based sensor system for contactless detection of wetness and ice on surfaces

Lakshan Tharmakularajah[1], Jakob Döring[1], Leonard Friedrich[1], Karl-Ludwig Krieger[1]
[1]Institute of Electrodynamics and Microelectronics (ITEM.ae), University of Bremen, Germany

{lakshan, doering, friedrich, krieger}@item.uni-bremen.de

Abstract

In order to differentiate between a wet and a dry road surface, the water film height is measured by using an infrared-based sensor system. With the help of different wavelengths it is also possible to distinguish between ice and water. In this article, a sensor system for determination of the physical state of water on different surfaces with infrared LEDs and one photodiode is presented. This shall serve as a basis for the calculation of the road condition. Here, emitted and reflected signals of five different infrared emitters are acquired with an infrared receiver. In commercial sensors this method (with one emitter) is used to determine a distance. At a constant distance, different intensities for the reflected signal are measured depending on the color and texture of the surface. Similar sensor systems work with a broadband light, usually a halogen lamp, and several photodiodes. This type is more cost-intensive and not robust enough for the desired application. The transmission spectrum of water shows that light is absorbed little at a wavelength of 900 nm and strongly at a wavelength of about 1400 nm – 1500 nm. Therefore, the intensity depends not only on the material surface, but also on the wavelength of the transmitting and receiving diode. In order to distinguish between ice and water, the wavelengths 1450 nm and 1550 nm are compared, which is shown in this article by the measurements. In ice, the infrared light is more strongly absorbed at 1550 nm and less absorbed at 1450 nm.

Keywords: optical measurement method, infrared light sensor, ice and water detection, contactless sensor, road condition sensor

I. Introduction

Common causes of personal injury in traffic are poor road conditions [1]. Between 2014 and 2017, slippery roads caused around 45900 accidents in Germany due to rain, snow and ice. That is around 3.5 % to 4.1 % of all accidents involving personal injury in Germany and the trend is rising. Drivers regularly lose control of their vehicle due to slippery roads. As a result, the demand for technologies that reliably detect critical road conditions of this kind is growing. Environmental detection and knowledge of road conditions are becoming increasingly important as automation and autonomous driving increase in the automotive sector. This is necessary for autonomous driving in demanding weather conditions. The environmental sensors already installed in vehicles are not sufficient, as they can only indirectly determine road conditions. For example, the rain-light sensor can only be used on a wet road surface when it is raining. Afterwards, the road is still wet but the sensor cannot be used to determine the road condition. There is an approach of indirect road condition monitoring with the use of capacitive sensors in the wheel arch liner [2]. To monitor the road surface, directly visible image sensors, infrared image sensors and millimeter-wave radar sensors are used [3, 4]. These sensors are complex and partly not suitable for vehicles. Therefore, the motivation is to determine the water film height with a novel sensor.

The technology for the determination of ice and water using infrared radiation is already known [4]. However, expensive spectrometers are used for the analyses. In [5], three photodiodes were used as a receiver and a broadband 20 W halogen lamp was used as the light source for the measurements. This method (with one emitter and receiver) is used to determine a distance with commercially available sensors [6]. In this approach a broadband photodiode and five selected LEDs with defined wavelengths are used as a light source.

II. Infrared-based sensor system

The aim of the sensor system presented here is to detect wetness or water on a surface and to record the height of the water film. Furthermore, only one broadband photodiode is used. For the

selection of suitable LEDs, the wavelengths in the transmission spectrum of water and the absorption spectrum of ice are considered.
A theoretical approach implies that due to the presence of a water film on a surface, the light intensity of the wavelength with a high transmittance decreases less than the intensity with a low transmittance [4]. In [7] the comparison of wavelengths 1.4 µm and 1.5 µm for ice detection is described. From the absorption spectrum of ice, it can be concluded that the wavelengths up to 1.4 µm are absorbed less than the wavelength at 1.5 µm. The transmittance spectrum (blue solid line) and the absorption spectrum (red dashed line) are shown in Fig. 1. By comparing the transmittance of water and the absorbance of ice, it can be shown that in particular the wavelengths at 1.4 µm and 1.5 µm can give information about the state of aggregation of water.

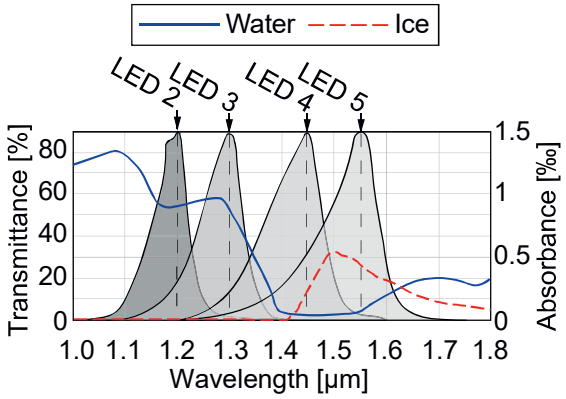


Fig. 1: Transmission spectrum of water (blue solid line) [7] and absorption spectrum of ice (red dashed line) [8] with the selected four LEDs.

To build the sensor system, five different LEDs are selected as the light source and one photodiode is used as the receiver. LED 1 lies outside the shown transmission spectrum in Fig. 1, this has an emission spectrum of 0.92 µm to 0.97 µm and the highest intensity is at 0.945 µm. From the emission spectra of LED 2 to 5 it can be seen that the spectra overlap.

Tab. 1: Comparison of the main features of the selected LEDs

	Designation	*Wavelength* [µm]	*Radiant power* [mW]
LED 1	L9338	0.945	15.0
LED 2	L13072	1.200	5.0
LED 3	L12771	1.300	3.1
LED 4	L10660	1.450	2.8
LED 5	L12509	1.550	3.8

The LEDs are arranged in a circle with the same distance to the photodiode, so that they have a similar influence. To extract the signals of the LEDs, in the signal of the photodiode, the LEDs are pulsed with different frequencies. This sensor system is shown in Fig. 2. The selected photodiode (G12180-030) is used as receiver. The photodiode consists of the semiconductor material indium gallium arsenide also referred to as InGaAs photodiode. This has a round, light-sensitive surface with a diameter of 0.3 mm, which reacts in a wavelength range of 0.9 µm to 1.7 µm.



Fig. 2: Infrared-based sensor system with five LEDs and one photodiode.

III. Measurement setup

In this section, the setup of the measurement environment is presented. For the measurements, a test bench was set up which sprays a water film with a defined height in 1 mm steps onto a surface. The test bench shown in Fig. 3 can be controlled in horizontal and vertical direction with a resolution of 1 mm. In order to reproducibly generate water film heights with an accuracy of 1 mm, a strain gage balance was implemented under the surface. The strain gage balance measures the weight, while a pump applies a fine spray to the surface. According to equation (1), the given density ρ of the water allows the weight m to be calculated into the volume

$$V = \frac{m}{\rho}. \tag{1}$$

By the calculated volume V and a given surface of a tile $a \cdot b$ the water film height

$$WFH = \frac{V}{a^2} \tag{2}$$

can be determined. In this case $a = b$. Therefore, the area of the surface is a^2. The white and black tiles measuring 20 cm x 20 cm are used for the tests. Boundaries were placed on the sides of the surfaces to prevent sprayed water from overflowing.

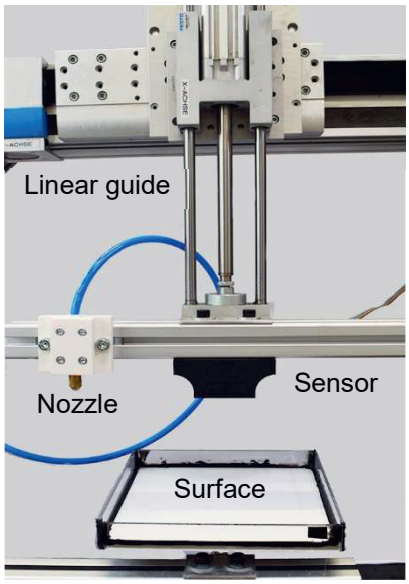


Fig. 3: Test bench for defined water film heights.

IV. Experimental study

In this section the performed experiments and results are presented. In order to be able to make tangible statements on the individual changed parameters of the distance and the surface, a protection against extraneous light is used.

First of all, two differently colored surfaces, a white and a black tile, were examined. At first, the measurements were carried out on a dry surface to investigate the influence of the different surface properties. In this case, the surfaces in a dry state consequently had a water film height of 0 mm were investigated. The intensities of the individual LEDs were measured from 5 cm to 10 cm in 1 cm steps.

Subsequently, wet measurements were carried out. A water film of 1 mm to 10 mm was measured in 1 mm steps at a distance of 5 cm on the white surface. Finally, on the white surface ice with a thickness of 1 mm was measured with the sensor system and the results were compared with those of water.

A. Distance and intensity analysis

The measurements were performed at a distance, between the sensor and the surface, of 5 cm to 15 cm in 1 cm steps, which are shown in Fig. 4 and Fig. 6. For all steps 5 measurements with 10s each were performed. The measurement data were acquired with a sampling rate of 500 kHz and then averaged. The light emitted by the LEDs is reflected by the two different surfaces and converted into a current by the photodiode. A voltage is generated by a transimpedance converter, which is digitally acquired. At a distance of 5 cm to 15 cm from the surface, different intensities of each LED can be observed.

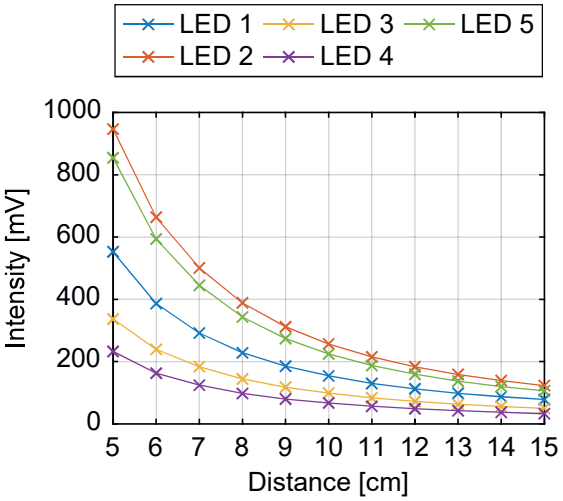


Fig. 4: Intensity at different distances from 5 cm to 15 cm in 1 cm steps with the white surface.

On the one hand this can be attributed to the different spectral sensitivity of the photodiode and on the other hand to the different radiant powers of the LEDs listed in Tab. 1. LED 1 has the highest radiant power, but it has the shortest wavelength. This results in increased reflections at the surface molecules. Due to the coarseness of the tile, the light beams of LED 1 are not reflected uniformly but scattered. This means that fewer light rays penetrate into the sensitive area of the photodiode and therefore the intensity is lower. In addition, the spectral transmission of the protective glass and the sensitivity of the photodiode contribute to LED 2 having the highest intensity value and LED 4 the lowest intensity value.

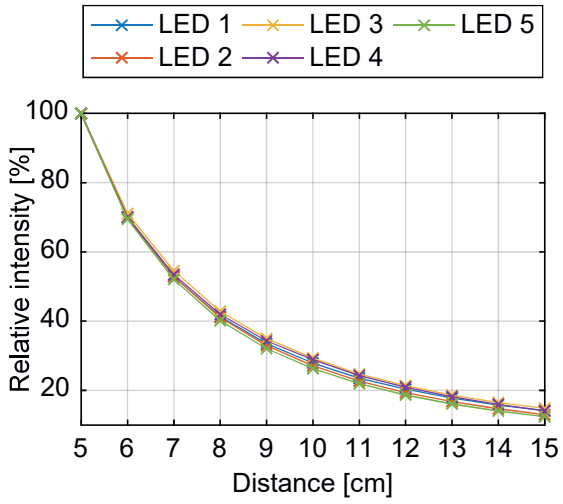


Fig. 5: Relative intensity decrease at distances of 5 cm to 15 cm with white surface.

By observing the course of the intensities I with increasing distance r, it is possible to recognize that these follow the inverse-square law of distance. The inverse-square law describes the decrease of a physical quantity with increasing distance to the source or transmitter [9].

$$I \sim \frac{1}{r^2} \tag{3}$$

This behavior is more clearly recognizable by the representation of the relative intensity to the smallest distance on a white surface. Fig. 5 shows that the intensities of all LEDs decreases quadratically with increasing distance. The intensity is reduced to a quarter of the initial intensity when the distance is doubled from 5 cm to 10 cm. Analogous to the measurements with the white surface, the measurements and results with the black surface are considered in the following. Fig. 6 shows the measured intensities at different distances with the black surface. Here it can be seen that the order of the intensity of the individual LEDs is almost identical for white and black tiles. The light of LED 2 is reflected most strongly and the light of LED 4 most weakly from the black surface back to the photodiode. LED 1 and LED 3 have an almost identical behavior.

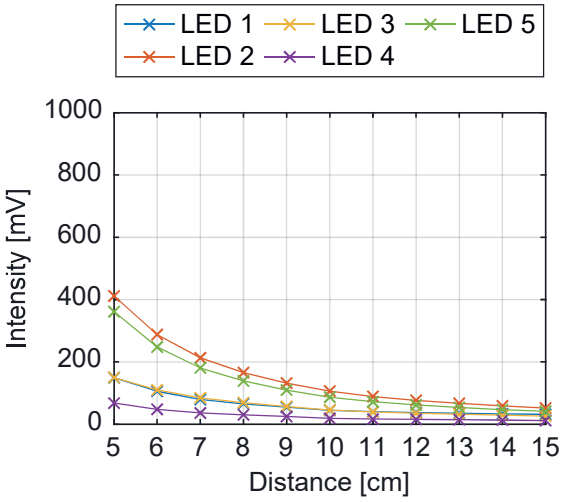


Fig. 6: Intensity at different distances from 5 cm to 15 cm in 1 cm steps with the black surface.

The direct comparison of the intensity values between the white and black surface at a distance of 5 cm shows a difference. This difference is due to the increased reflectivity of the white surface and the increased absorption of the black surface. A quantification of the values at a distance of 5 cm for the white and black surface and the resulting decrease in intensity are given in Tab. 2. The intensity of LED 2, 3 and 5 are absorbed almost uniformly by the surface of the black tile. LED 1 and 4 show a decrease in intensity of about 70 %. The wavelength-dependent refractive index and the absorption behavior of the black tile could be decisive for this.

Tab. 2: Intensity decrease between the white and black tiles

	White tile [mV]	*Black tile* [mV]	*intensity decrease* [%]
LED 1	552.51	149.72	72.90
LED 2	946.67	411.48	56.53
LED 3	336.78	149.80	55.52
LED 4	232.59	67.82	70.84
LED 5	854.55	361.33	57.72

B. Investigation of water film heights

The measurements are carried out at a fixed distance of 5 cm between sensor and surface. A water film layer of 1 mm to 10 mm was applied to the substrate in 1 mm steps. The measured intensity values of the individual LEDs for water film height from 0 mm to 10 mm with a sensor distance of 5 cm are shown in Fig. 7. The water film height of 0 mm describes the dry condition of the surface. The measurement with a water film height of 1 mm shows that the intensity of the individual LEDs decrease. This decrease in intensity can be explained by the light beam path described in [10]. The light beam is not only reflected on the surface of the ground, but first partially reflected on the water surface. The transmitted light beam is partly absorbed by water. The strength of the absorption is different for each wavelength of the LEDs. The light beam is then reflected at the surface of the substrate, absorbed again by the water, refracted at the transition between water and air and received by the photodiode.

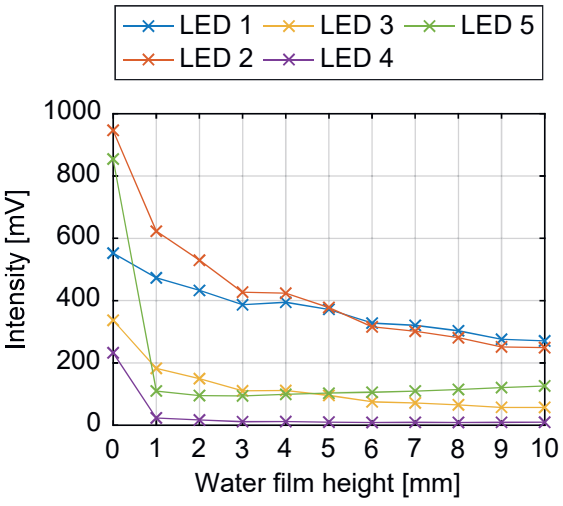


Fig. 7: Measurements of water film heights from 1 mm to 10 mm on a white surface.

In general, the light with the shorter wavelengths 0.945 µm, 1.2 µm and 1.3 µm decreases less than that of the longer wavelengths 1.45 µm and 1.55 µm. This is due to the fact that light with the wavelengths 0.945 µm, 1.2 µm and 1.3 µm is more effectively transmitted by water and light with the wavelengths 1.45 µm and 1.55 µm is more strongly absorbed by water. By analyzing the progression with increasing water film height, it can be observed that the measurement at the water film height of 4 mm shows a slight measurement deviation. The intensity of LED 5 with a wavelength of 1.55 µm increases from a water film height of 1 mm. This increase in intensity suggests that the reflective component on the surface of the water is higher than the transmission and absorption component. As the distance between sensor and water surface decreases, the intensity values of LED 5 increase. For detailed analyses of the intensity change, the transmittance is formed. The transmittance T is calculated for each wavelength λ from the ratio of the intensities with water I present to the intensity with a dry surface I_0. The following equation (4) shows the calculation of transmittance [11]

$$T(\lambda) = \frac{I}{I_0}. \qquad (4)$$

Fig. 8 shows the transmittance for the respective wavelengths with increasing water film heights from 1 mm to 10 mm on the white surface. Since only discrete values can be recorded with the five LEDs, a diagram with discrete values is formed instead of an infrared spectrum.

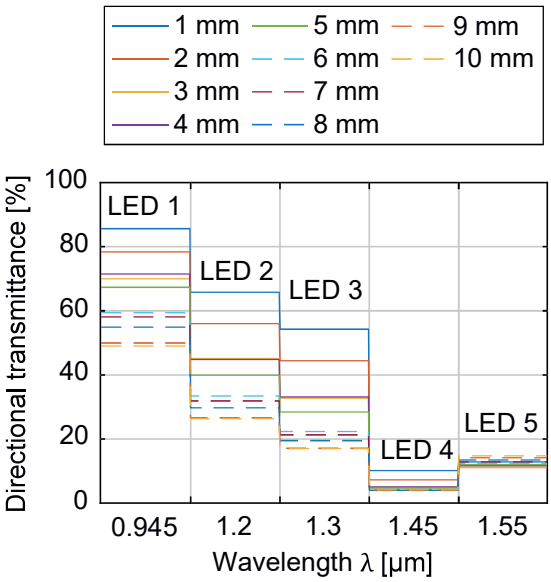


Fig. 8: Measured directional transmittance for increasing water film heights from 1 mm to 10 mm with white surface.

The following statements can be made with the help of the directional transmittance shown in Fig. 8. The discrete values of the transmittance at a water film height of 1 mm correspond approximately to those of the theory in Fig. 1. The transmittance is lowest at a wavelength of 1.45 µm. Subsequently, the transmittance increases for the wavelengths 1.55 µm, 1.3 µm and 1.2 µm and the highest is at 0.945 µm.
A further statement refers to the resolution of the respective LEDs with increasing water film height. This is shown by the steps of the diagram with increasing water film height. The difference between the transmittances of two water film heights is approximately the same for the wavelengths 0.945 µm, 1.2 µm and 1.3 µm. By considering the transmittance, those wavelengths are suitable for quantitative statement about the water film height of 1 mm to 8 mm. At a wavelength of 1.45 µm, the transmittance remains unchanged from a water film height of 3 mm. The increase in intensity of the LED 5 with the wavelength of 1.55 µm can be explained by the fact that the distance between sensor and surface of the water decreases with increasing water film height. Accordingly, the wavelengths 1.45 µm and 1.55 µm are suitable for a qualitative measurement of water film heights.

C. *Investigation of ice and water*

The aim of this measurement was to verify the detection of ice with the help of the theoretical approach. The water film height and the thickness of the ice layer on the white surface were 1 mm. As the red solid line shows in Fig. 1, the wavelength of 1.5 µm is strongly absorbed by ice. Thus, the classification of ice should be possible by comparing the different wavelengths.

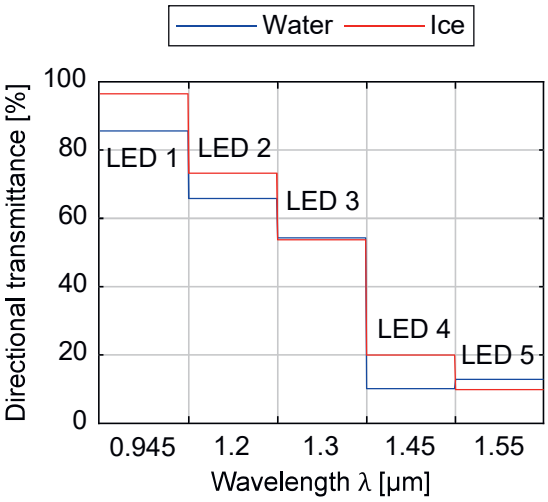


Fig. 9: Measured directional transmittance of water and ice.

Fig. 9 shows that the intensities of LED 1 with a wavelength at 0.945 µm, LED 2 with a wavelength at 1.2 µm and LED 4 with a wavelength at 1.45 µm are absorbed more by water than by ice. The intensity of LED 3 with the wavelength at 1.3 µm shows, for water and ice, almost the same transmittance. The theoretical approach to the detection of ice includes the comparison of the last two wavelengths. The wavelength at 1.45 µm is more absorbed by water than a wavelength at 1.55 µm. Ice, on the other hand, transmits the wavelength at 1.45 µm more efficiently and absorbs at 1.55 µm more strongly. Other combinations such as LED 2 and LED 5 or LED 1 and LED 5 can also be used to distinguish between water and ice. Because here the transmittance of LED 1 and LED 2 increases like the transmittance of LED 4 when ice is present. Consequently, a distinction can be made between water and ice by comparing these wavelengths.

V. Conclusion

This work presented an infrared-based sensor system with one photodiode as the receiver and five different LEDs as the transmitter for contactless detection of wetness and ice on a surface. In the first part of this work, a theoretical approach was used to determine the appropriate emitters and receiving diodes. Five LEDs were selected as transmitter diodes. A photodiode in the wavelength range from 0.9 µm to 1.7 µm was selected as the receiver.

Subsequently, a measuring environment was created which reproducibly sprayed a defined water film onto a surface. This has a strain gage balance that measures the weight of the sprayed water to determine the water film height.

In the next step, the measurements were carried out. First, the influence of the brightness of the surface was investigated. A white and a black tile were used as the surface. The measurements showed that the white surface reflects much more light back than the black surface. The intensity of the monochromatic lights decreases from 55 % to 72 % depending on the wavelength compared to the black surface. With the help of the relative intensity the inverse-square law of distance could be shown. The intensity decreases by 75 % when doubling the distance and is consequently only a quarter of the original intensity.

The measurements with water on the surface showed that the sensor system has a similar course as the transmission spectrum of water. Here it was evident that a part of the light of LED 5 with 1.55 µm wavelength was absorbed by water and a part was reflected back from the water surface. Due to the intensity increased with the water film height, it is possible to measure the water film height with this sensor system.

The investigations carried out with ice showed that ice can be detected qualitatively by comparing different wavelengths. The light with a wavelength of 0.945 µm, 1.2 µm and 1.45 µm is more strongly absorbed by water than by ice. For light with a wavelength of 1.55 µm it is exactly the opposite. This sensor system is able to measure water film height and ice.

VI. References

[1] Statistisches Bundesamt (Destatis), "Fachserie 8 Reihe 7 Verkehr Vekehrsunfälle," Wiesbaden, 2018.

[2] J. Döring, L. Tharmakularajah, J. Happel and K.-L. Krieger, “A novel approach for road surface wetness detection with planar capacitive sensors”, J. Sens. Sens. Syst. 8, pp. 57-66 (2019); doi: 10.5194/jsss-8-57-2019

[3] R. Kurata, H. Watanabe, M. Tohno, T. Ishii and H. Oouchi, “Evaluation of the detection characteristics of road sensors under poorvisibility conditions”, IEEE Intelligent Vehicles Symposium, 2004, pp. 538-543 (2004); doi: 10.1109/IVS.2004.1336441

[4] F. Holzwarth and U. Eichhorn, “Non-contact sensors for road conditions, Sensors and Actuators”, Sensors and Actuators A: Physical, 37-38, 121-127 (1993); doi: 10.1016/0924-4247(93)80023-A

[5] P. Jonsson, “Remote sensor for winter road surface status detection”, SENSORS, 2011 IEEE, pp. 1285-1288 (2011); doi: 10.1109/ICSENS.2011.6127089

[6] O. Schrey, W. Brockherde and B. Hosticka, "Distance sensor and method for distance detection". Germany Patent US 20060214121 A1, 09 05 2005.

[7] F. Klempau, “Untersuchungen zum Aufbau eines Reibwertvorhersagesystems im fahrenden Fahrzeug”, Darmstadt: Technische Universität Darmstadt, 2003.

[8] S. G. Warren, „Optical constants of ice from the ultraviolet to the microwave”, Appl. Opt., 23, 1206-1225 (1984); doi: 10.1364/AO.23.001206

[9] U. Harten, „Physik“, Berlin Heidelberg: Springer-Verlag, (2014); doi: 10.1007/978-3-642-53854-4

[10] L. Tharmakularajah, J. Döring and K.-L. Krieger, "Infrared-based determination of the type and condition of the road surface" in 19. Internationales Stuttgarter Symposium Automobil- und Motorentechnik, Stuttgart, 2019.

[11] H. Günzler and H.-U. Gremlich, „IR-Spektroskopie“, Weinheim: WILEY-VCH Verlag GmbH & Co. KGaA, (2003); doi: 10.1002/9783527662852

DOI 10.5162/sensoren2019/3.1.4

RoHS-konforme, thermoelektrisch gekühlte Detektoren für das langwellige Infrarot auf Basis antimonidischer Übergitter

Raphael Müller, Jasmin Niemasz, Volker Daumer, Robert Rehm
Fraunhofer Institut für Angewandte Festkörperphysik (IAF)
Tullastraße 72, 79108 Freiburg, Deutschland

Zusammenfassung

In diesem Beitrag präsentieren wir antimonidische Übergitter, im Speziellen InAs/GaSb-Übergitter, als ein zu Quecksilbercadmiumtellurid alternatives Materialsystem für die schnelle und sensitive Photodetektion im langwelligen Infrarot bei Betriebstemperaturen die mit thermoelektrischer Kühlung erreichbar sind. Die Detektorentwicklung für InAs/GaSb-Übergitter, welche unter anderem durch die mangelnde RoHS-Konformität von Detektoren aus Quecksilbercadmiumtellurid motiviert wird, wird für das Detektionskonzept eines lateral betriebenen Photoleiters vorgestellt. Die in diesem Beitrag diskutierten Detektoren zeigen eine Grenzwellenlänge von 10,5 µm bei einer Betriebstemperatur von 210 K, erreichbar mithilfe eines 3-stufigen Peltierkühlers. Die durchschnittliche Detektivität der undotierten Übergitterdetektoren ist vergleichbar mit derer von dotieroptimierten Detektoren auf Basis von Quecksilbercadmiumtellurid. Die experimentelle Untersuchung der Übergitterdotierung stellt eine Erhöhung der Detektivität und eine zunehmende Konkurrenzfähigkeit mit Quecksilbercadmiumtellurid für Detektoranwendungen im langwelligen Infrarot unter thermoelektrischer Kühlung in Aussicht.

Keywords: Infrarotdetektor, Peltierkühler, RoHS, InAs/GaSb-Übergitter, MCT

Thermoelektrisch gekühlte Infrarot-Photodetektoren

Für die Identifizierung von Substanzen, die Bestimmung ihrer Konzentration in Gasen und Flüssigkeiten sowie die Messung zeitlich veränderlicher Signale werden in Forschung und Industrie schnelle und sensitive Infrarotdetektoren (IR-Detektoren) benötigt und standardmäßig eingesetzt.
Die Kombination aus schneller Signaldetektion und hohem Signal-Rausch-Verhältnis, welches die aussagekräftige Detektion selbst schwacher Signale ermöglicht, erfordert den Einsatz von IR-Photodetektoren. Diese basieren auf Halbleitermaterialien und werden für verbesserte Leistungsfähigkeit zumeist gekühlt betrieben. Die thermoelektrische Kühlung, welche auf dem Einsatz mehrstufiger Peltierkühler beruht, ermöglicht Betriebstemperaturen zwischen Raumtemperatur und etwa 200 K und ist für viele Anwendungsszenarien das einzige praktikable und kosteneffiziente Kühlverfahren für IR-Detektoren.
In Abb. 1 ist ein IR-Detektor mit und ohne Gehäuse auf einem 4-stufigen Peltierkühler exemplarisch dargestellt. Oberhalb des Peltierkühlers befindet sich eine Mikrolinse, welche die einfallende Strahlung auf das darunterliegende Detektorelement fokussiert und dadurch dessen Signal-Rausch-Verhältnis steigert.



Abb. 1: IR-Detektor mit Immersionslinse auf 4-stufigem Peltierkühler [1]

Thermoelektrisch gekühlte Detektoren auf Basis von Quecksilbercadmiumtellurid (MCT, von engl.: *mercury cadmium telluride*) sind in verschiedenen Größen, mit unterschiedlichen Grenzwellenlängen und für den durch thermoelektrische Kühlung erschließbaren Temperaturbereich kommerziell erhältlich. Die Nutzung von Quecksilber und Cadmium für die Herstellung von IR-Detektoren wird jedoch zunehmend durch die RoHS–Richtlinie »2011/65/EU« der europäischen Union bedroht, welche deren Nutzung beschränkt. Derzeit existieren für bestimmte Anwendungen befristete Ausnahmegenehmigungen, für die mittel- und langfristige Detektorentwicklung werden Alternativen jedoch intensiv beforscht.

Flexible Materialsysteme für die Infrarotdetektion: Quecksilbercadmiumtellurid und InAs/GaSb-Übergitter

Die zentrale Eigenschaft für die Eignung eines Halbleiters für die Herstellung von IR-Photodetektoren ist dessen Bandlücke E_g, welche maßgeblich die Grenzwellenlänge des Detektors bestimmt.
Die Bandlücke von MCT, respektive die Grenzwellenlänge von MCT-Detektoren, kann mithilfe verschiedener Epitaxieverfahren durch das Verhältnis von Cadmium und Quecksilber eingestellt werden. Dabei sind Grenzwellenlängen von 2-30 µm möglich, wobei die Einstellbarkeit der Bandlücke sowie flächig homogenes Wachstum für die höheren Grenzwellenlängen in diesem Bereich jedoch technologisch zunehmend komplizierter wird.
Das Legieren von III/V-Halbleitern wie In, Ga, As, Sb ermöglicht ebenfalls eine flexible Bandlückeneinstellbarkeit, jedoch in einem schmaleren Spektralbereich und mit geringerer maximal erreichbarer Grenzwellenlänge.
Eine größere Flexibilität besteht durch den Einsatz sogenannter Kompositionsübergitter, einer periodischen Abfolge alternierender Schichten verschiedener Materialien, die im vorgestellten Fall der InAs/GaSb Übergitter aus Indiumarsenid und Galliumantimonid bestehen. Diese Schichten können durch Molekularstrahlepitaxie mit Schichtdicken im Bereich weniger Monolagen (ML) und abrupten Schichtübergängen hergestellt werden. In diesem künstlichen Halbleitermaterial werden die energetische Lage von Leitungs- und Valenzbandkante durch die Quantenmechanik definiert. Durch die Superposition der Wellenfunktionen der Ladungsträger in den Übergitterteilschichten entstehen sogenannte Minibänder. Die energetische Lage und der Abstand dieser Minibänder definiert die Größe der effektiven Halbleiterbandlücke. Durch Veränderung der Teilschichtdicken von InAs und GaSb in einer Übergitterperiode können die energetische Lage der Minibänder und die Größe der Bandlücke zielgenau eingestellt werden.
In Abb. 2 ist die Bandanordnung von InAs und GaSb sowie das resultierende Bandstrukturdiagramm eines InAs/GaSb-Übergitters dargestellt. In der Bandstruktur des Übergitters sind das energetisch tiefstliegende Elektronenminiband (Übergitterleitungsband) und das energetisch höchstliegende Lochminiband (Übergittervalenzband) exemplarisch dargestellt. Eine Verbreiterung der InAs-Teilschichtdicke führt beispielweise zur Absenkung des Übergitterleitungsbandes und zur Verringerung der Übergitterbandlücke.

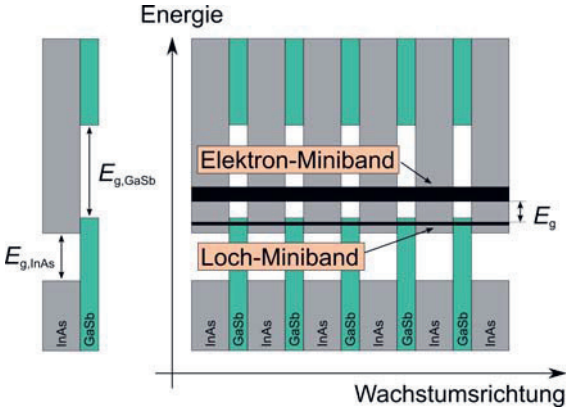


Abb. 2: Bandanordnung von InAs und GaSb sowie Bandstrukturdiagramm eines InAs/GaSb-Übergitters

Trotz der theoretisch aufwändigen Beschreibung lassen sich InAs/GaSb-Übergitter für viele entwicklungs- und anwendungsspezifische Aspekte als anisotropes Halbleitermaterialsystem mit flexibel einstellbarer Bandlücke betrachten mit dem Detektorgrenzwellenlängen im Infraroten für verschiedene Anwendungen im Bereich von 3-20 µm maßgeschneidert definiert werden können.
Im Vergleich zu MCT zeichnen sich Übergitter durch höhere Prozessausbeuten und geringere Kosten aus. Aus diesen wirtschaftlichen Gründen und aufgrund einer theoretisch postuliert, höheren Leistungsfähigkeit als MCT werden InAs/GaSb-Übergitter seit einigen Jahren als Alternative zu MCT intensiv beforscht und zunehmend konkurrenzfähig. Der Fokus der Übergitterentwicklung liegt auf bildgebenden Anwendungen bei denen das Detektormaterial kryogen gekühlt betrieben wird. Auch das Fraunhofer IAF erforscht und fertigt seit langem IR-Detektoren auf Basis von InAs/GaSb-Übergittern, was zur Entwicklung von bispektralen Kameras für das mittelwellige Infrarot [2] und zur ersten europäischen InAs/GaSb-Übergitterkamera für das langwellige Infrarot führte [3].

Entwicklung und Herstellung photoleitender Detektoren

Die hier vorgestellte Entwicklung photoleitender Detektoren erfolgte in Kooperation mit VIGO System, im Folgenden kurz »VIGO«, für MIRPHAB. MIRPHAB ist eine europäische Pilotlinie für die Entwicklung und Kommerzialisierung von sogenannten »Advanced MIRPHAB Modules«, bestehend aus Lichtquelle, Lichtleiter und Detektor. Diese Module werden für den Spektralbereich von 3-12 µm entwickelt und finden Anwendung in der chemischen Analyse und der Spektroskopie (siehe auch: www.mirphab.eu).

Der Schichtentwurf und das Wachstum der Übergitterschichten erfolgen am Fraunhofer IAF, ebenso wie die photolithographische Prozessierung. Die im Folgenden knapp vorgestellte Schichtstruktur, das Detektorkonzept, sowie die Prozessierung wurden bereits in Ref. [4] und Ref. [5] diskutiert.
In Abb. 3 sind die Schichtstruktur sowie ein prozessiertes Detektorelement schematisch dargestellt. Für die in Abb. 3 oben gezeigte Schichtstruktur wird auf 3“-GaAs zunächst ein geeigneter Wachstumspuffer abgeschieden, gefolgt vom photosensitiven Übergitterabsorber und einer Kontaktschicht. Für die Übergitterkomposition der Absorberschicht, welche die Grenzwellenlänge des Detektors definiert, wurden 14 ML InAs und 7 ML GaSb gewählt.
Im prozessierten Detektorelement, in Abb. 3 unten, sind neben den strukturierten Schichten von Absorber und Kontakt eine Passivierungsschicht sowie die Metallisierung der Kontakte gezeigt. Um für ein durch das Substrat beleuchteten Detektor einen Doppeldurchgang der einfallenden Strahlung zu ermöglichen, wird zudem eine als Spiegel wirkende Metallisierung im mittigen Bereich des photosensitiven Absorbers zwischen den metallischen Kontakten aufgebracht.

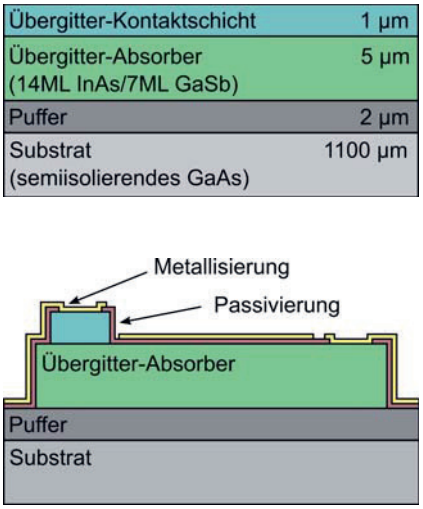


Abb. 3: Schichtstruktur für die Herstellung von Übergitterdetektoren (oben) und Schema eines prozessierten Übergitterdetektors (unten)

Für die Vereinzelung der Detektorstrukturen und die weitere Integration werden VIGO Waferchips zur Verfügung gestellt. Einer dieser Waferchips, welcher 15x15 Detektorelemente umfasst, ist in Abb. 4 dargestellt. Nach der Detektorvereinzelung werden Mikrolinsen aus dem GaAs-Substrat abgeformt, die Detektoren mit Verstärkerelektronik versehen und in einem dedizierten Messaufbau charakterisiert.

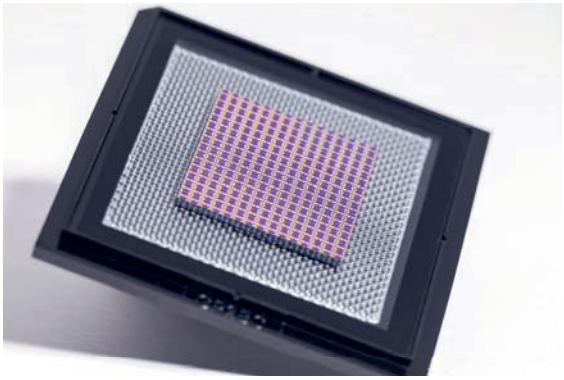


Abb. 4: Waferchip mit 15x15 InAs/GaSb-Übergitterphotoleitern (22x22mm²)

Spektrale Detektivität:
Zentrale Vergleichsgröße für die Bewertung von Infrarotdetektoren

Die zentrale Vergleichsgröße für die Bewertung der Leistungsfähigkeit von Infrarotdetektoren ist die spektrale Detektivität $D^*(\lambda)$, ein Maß für das Signal-Rausch-Verhältnis des Detektors. Sie verknüpft das durch die eingestrahlte Leistung normierte Photosignal, die Responsivität R, mit dem Rauschstrom I_n, normiert mit der optischen Detektorfläche A und der Bandbreite Δf:

$$D^*(\lambda) = \frac{R(\lambda)}{I_n(f)} \sqrt{A\,\Delta f} \ .$$

Die Bestimmung der Detektivität erfordert folglich neben der Kenntnis der Detektorgeometrie die Bestimmung der spektralen Responsivität sowie des frequenzabhängigen Rauschverhaltens.
Für den Fall des hier untersuchten Detektorkonzepts eines photoleitenden InAs/GaSb-Übergitterdetektors, welcher unter thermoelektrischer Kühlung betrieben wird, ist die Responsivität proportional zur Quanteneffizienz und dem Photoleitungsgewinn, das heißt der Lebensdauer der Ladungsträger sowie deren Beweglichkeit. Der frequenzabhängige Rauschstrom wird als Kombination aus thermischem Rauschen, Generations-Rekombinationsrauschen sowie 1/f-Rauschen beschrieben, welche vom Widerstand des Detektors, dem Photoleitungsgewinn sowie unter anderem der Detektorprozessierung abhängig sind.
Eine Verbesserung der Detektivität kann aus einer Erhöhung der materialspezifischen, dotierabhängigen Ladungsträgerlebensdauer oder Beweglichkeit, sowie durch die Erhöhung des Detektorwiderstands resultieren.
Nachfolgend werden nun zunächst Charakterisierungsergebnisse undotierter Übergitterphotoleiter diskutiert und dann Ergebnisse einer Dotierstudie vorgestellt.

Spektrale Detektivität undotierter InAs/GaSb-Übergitterphotoleiter

Für die elektrooptische Bewertung der vorgestellten Detektoren ist neben der potenziellen Leistungsfähigkeit auch die Homogenität über den Wafer von großer Bedeutung für eine mögliche Kommerzialisierbarkeit.

Für eine erste Bewertung der Materialhomogenität wurden 10 zufällig ausgewählte Detektoren mit einer Detektorfläche von 100x100 µm^2 aufgebaut. Auf den detektivitätssteigernden Einfluss einer GaAs-Immersionslinse wurde hier verzichtet, die Kompatibilität wurde bereits in Ref. [5] gezeigt. Nach dem Aufbau der Detektoren mit planer Substratoberfläche erfolgte die Charakterisierung der Responsivität und des Rauschverhaltens bei einer Betriebstemperatur von 210 K, welche mit einem 3-stufigen Peltierkühler erreichbar ist.

Die Charakterisierung der einzelnen Detektoren wurde bereits in Ref. [4] diskutiert. Dabei konnte ein sehr einheitliches Verhalten festgestellt werden, sowohl für die Responsivität, wobei die Streuung der Maximalwerte im Bereich der messplatzbedingten Unsicherheit lag, als auch für das Rauschverhalten, welches sowohl qualitativ als auch quantitativ sehr einheitlich war. Demzufolge ergab sich ebenfalls ein einheitliches Verhalten bezüglich der Detektivität der untersuchten Übergitterphotoleiter. Dies motiviert für die weitere Diskussion den Mittelwert der Detektivität dieser 10 Detektoren als repräsentativen Wert zu nutzen.

In Abb. 5 ist die gemittelte spektrale Detektivität der 10 InAs/GaSb-Übergitterphotoleiter vorgestellt. Außerdem ist die Detektivität für kommerziell erhältliche MCT-Photoleiter von VIGO mit unterschiedlicher Grenzwellenlänge gezeigt, welche den direkten Vergleich ermöglichen. Die Messdaten aller Detektoren, von MCT und Übergitter, wurden am gleichen Messaufbau bei gleicher Betriebstemperatur charakterisiert und sind für das Rauschen des Detektors bei 20 kHz angegeben. Die gemittelte Detektivität der Übergitterphotoleiter mit einer Grenzwellenlänge von 10,5 µm liegt bei 3,3x10^8 cmHz$^{0.5}$W^{-1} im Maximum und 1,6x10^8 cmHz$^{0.5}$W^{-1} bei der typischen Grenzwellenlänge von 10,6 µm. Der Vergleich mit der Kurve entsprechender Grenzwellenlänge bei MCT (blau) offenbart geringere Werte für die Übergitterdetektoren, der Maximalwert für MCT ist dabei jedoch weniger als zweimal so hoch.

Trotz der etwas höheren Detektivität von dotieroptimiertem MCT im Vergleich zu undotierten Übergittern, zeigt dies die Relevanz des Übergittermaterialsystems für den Einsatz unter Betriebstemperaturen die mit Peltierkühlern erreichbar sind.

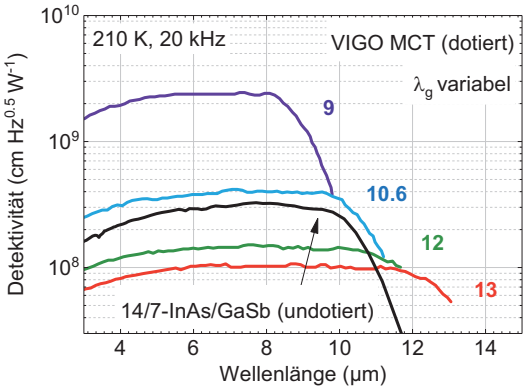


Abb. 5: Gemittelte spektrale Detektivität von 10 undotierten InAs/GaSb-Übergitterphotoleitern im Vergleich zu dotiertem kommerziellem MCT mit Grenzwellenlänge λ_g von 9 bis 13 µm

Entwicklung und Bewertung dotierter InAs/GaSb-Übergitterphotoleiter

Die Responsivität und der Dunkelwiderstand sind wichtige, die Detektivität des Photoleiters beeinflussende Eigenschaften, welche maßgeblich durch die Dotierung der photosensitiven Absorberschicht beeinflusst werden. Dies motiviert die Untersuchung der Übergitterdotierung im Rahmen der Detektoroptimierung. Wir zeigen im Folgenden erste Ergebnisse einer Dotierstudie für InAs/GaSb-Übergitterphotoleiter mit gleicher Übergitterkomposition von 14 ML InAs und 7 ML GaSb bei variierter Dotierung der Absorberschicht, ebenfalls basierend auf GaAs-Substrat. Die Dotierung der Absorberschicht erfolgt durch den Einbau von Beryllium in die InAs-Teilschicht, was zunächst zur Kompensation der n-artigen Hintergrunddotierung und mit steigender Dotierkonzentration zunehmend zu p-leitenden Eigenschaften der Schicht führt.

Die vorgestellten Messdaten wurden für jeweils einen Detektor der jeweiligen Dotierniveaus in einem Tauschgaskryostaten. Die untersuchten Detektorstrukturen haben eine optisch sensitive Fläche von 1x1 mm^2, keine Frontseitenmetallisierung und wurden unter Frontseitenbeleuchtung charakterisiert. Für alle Detektorstrukuren werden der Dunkelwiderstand und der Photostrom gemessen, letzterer mithilfe einer Lockin-Messtechnik bei frequenzmodulierter Beleuchtung mit einem Schwarzkörperstrahler der Temperatur 1100°C.

In Abb. 6 sind die Widerstände für einen undotierten Übergitterdetektor sowie drei Übergitterdetektoren mit zunehmender Be-Dotierung für eine Betriebstemperatur von 210 K gezeigt. Bei der angegebenen Dotierung handelt es sich um die nominelle Dotierung der InAs-Teilschicht, bezogen auf die Gesamtdicke einer Übergitterperiode. Die Messdaten zeigen

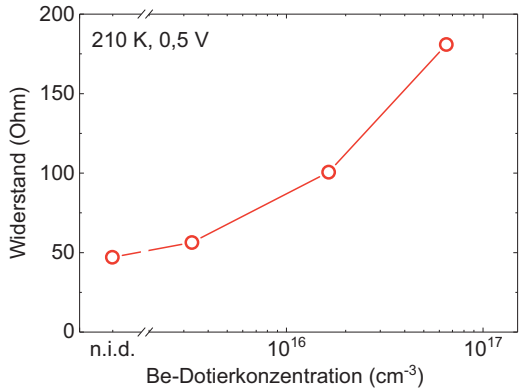


Abb. 6: Dotierabhängigkeit des Widerstands photoleitender InAs/GaSb-Übergitter-detektoren bei 210 K und 0,5 V

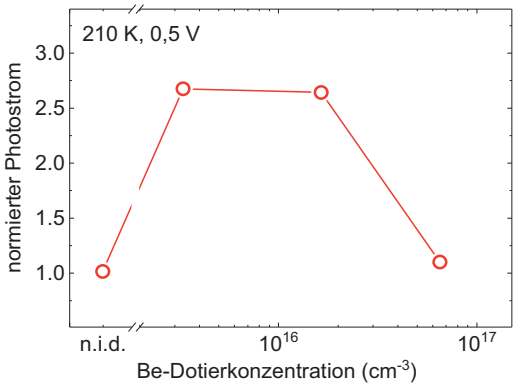


Abb. 7: Dotierabhängigkeit des Photostroms photoleitender InAs/GaSb-Übergitter-detektoren bei 210 K und 0,5 V

einen eindeutigen Trend eines mit zunehmender Dotierung steigenden Widerstands mit 46 Ohm für einen undotierten Übergitterdetektor bis zu 181 Ohm für den mit 6,7x10^{16} cm^{-3} höchstdotierten Übergitterdetektor dieser Untersuchung. Dies führt zu einer Reduktion des thermischen Rauschens des Detektors.

In Abb. 7 ist der Photostrom für die gleichen Detektorstrukturen bei gleicher Betriebstemperatur und Betriebsspannung dargestellt. Die vertikale Achse ist auf den Photostrom der undotierten Probe normiert. Für die untersuchten Proben zeigt sich ein mit zunehmender Dotierung zunächst steigender Photostrom, der ein um einen Faktor 2,7 erhöhtes Photostromniveau ermöglicht. Bei hoher Dotierung fällt dieses jedoch wieder nahezu auf den Wert des undotierten Übergitters ab. Die Verringerung des Photostroms für stark erhöhte Dotierpegel ist vermutlich auf eine dotierbedingte Zunahme von Rekombinations- und Streuzentren zurückzuführen, welche die Lebensdauer und die Beweglichkeit der Ladungsträger einschränken, was Responsivität und Detektivität wieder reduziert. Der generelle Trend des Photostroms deutet die Existenz eines von der Absorberdotierung abhängigen Maximalwertes an, der im Rahmen dieser Untersuchung jedoch nicht weiter quantifizierbar ist.

Zusammenfassend ist festzuhalten, dass sowohl der Detektorwiderstand als auch Photostrom durch Absorberdotierung eine Erhöhung der Detektivität in Aussicht stellen. Für die vorgestellte Betriebstemperatur von 210 K, welche mithilfe eines dreistufigen Peltierkühlers erreichbar ist, wird diese bei der gegebenen Übergitterkomposition für eine Dotierkonzentration um 10^{16} cm^{-3} erwartet, bei welcher der Detektorwiderstand steigend und das Photostromniveau verglichen mit dem undotierten Fall noch stark erhöht ist.

Eine quantitative Aussage über die Erhöhung der spektralen Detektivität für dotierte Übergitterphotoleiter erfordert die Untersuchung einer größeren Anzahl von Detektoren sowie die ausstehende Charakterisierung der Detektoren im finalen Aufbauzustand unter substratseitiger Beleuchtung im Messaufbau bei VIGO. Eine vollständige Analyse wird an anderer Stelle veröffentlicht.

Fazit

In diesem Beitrag wurden InAs/GaSb-Übergitterphotodetektoren als mögliche Alternative zu Quecksilbercadmiumtellurid für mit Peltierkühlern erreichbare Betriebstemperaturen vorgestellt. Die Entwicklung eines photoleitenden Übergitterdetektors mit einer Übergitterkomposition aus 14 Monolagen InAs und 7 Monolagen GaSb im photosensitiven Absorber wurde vom Schichtentwurf bis zur Charakterisierung skizziert. Es wurde gezeigt, dass die spektrale Detektivität undotierter photoleitender Detektoren bei 210 K bereits vergleichbar ist mit der Detektivität kommerziell erhältlicher MCT-Detektoren und dass durch gezielte Dotierung sowohl Widerstand als auch Photostrom erhöht werden können. Dies resultiert in einem geringeren thermischen Rauschen und einer erhöhten Responsivität, wovon beides eine Detektivitätserhöhung in Aussicht stellt. Eine Quantifizierung dieses Effekts für den finalen Aufbauzustand mit oder ohne Mikrolinse ist noch ausstehend. Nichtsdestotrotz sind InAs/GaSb-Übergitter für thermoelektrisch gekühlte IR-Detektoranwendungen bereits im vorgestellten Entwicklungsstadium als mögliche Alternative zu Quecksilbercadmiumtellurid anzusehen, dessen Nutzung in absehbarer Zeit von der RoHS-Richtlinie beschränkt wird.

Ausblick

Die für die Entwicklung von photoleitenden InAs/GaSb-Übergitterdetektoren eingesetzten Prozesse im Bereich des Schichtwachstums und der Übergitterprozessierung am Fraunhofer IAF ermöglichen, wie in Ref. [4] gezeigt, die Herstellung einheitlich leistungsfähiger IR-Detektoren. Dies stellt hohe Prozessausbeuten und eine kosteneffiziente Nutzung der Übergittertechnologie in Aussicht.

Darüber hinaus besteht die begründete Annahme, dass die Optimierung photoleitender InAs/GaSb-Übergitterdetektoren durch Be-Dotierung, welche hier lediglich für die Schichtkomposition aus 14 Monolagen InAs und 7 Monolagen GaSb für 210 K gezeigt wurde, auch für andere Übergitterkompositionen und andere durch thermoelektrische Kühlung erreichbare Betriebstemperaturen möglich ist.

Im Kontext unterschiedlicher Übergitterkompositionen, und der damit verbundenen unterschiedlichen Detektorgrenzwellenlängen, sei an dieser Stelle nochmals auf Abb. 5 verwiesen. Sie wurde zuvor für die Einordnung der Leistungsfähigkeit undotierter Übergitterdetektoren genutzt. Der Maximalwert der Detektivität der MCT-Detektoren unterschiedlicher Grenzwellenlänge nimmt für steigende Grenzwellenlängen ab. Dieser generelle Trend ist nicht materialspezifisch, sondern auf eine allgemeine Eigenschaft von Photodetektoren zurückzuführen. Für größere Grenzwellenlängen wird eine geringe Bandlücke benötigt, was jedoch die thermische Band-Band-Anregung im Halbleiter verstärkt. Diese Anregung führt zu hohen Ladungsträgerdichten, hohen Dunkelströmen, zunehmendem Rauschen und einer entsprechend verringerten spektralen Detektivität. Das gleiche qualitative Verhalten, die Reduktion der spektralen Detektivität für größere Grenzwellenlängen, ist daher auch bei InAs/GaSb-Übergitterphotodetektoren zu erwarten. Der quantitative Verlauf ist jedoch außerdem abhängig von fundamentalen Halbleitereigenschaften, die für Übergitterdetektoren bei hohen Betriebstemperaturen bislang noch wenig erforscht sind.

Literaturnachweis

[1] www.vigo.com.pl/products/pci-4te/
[2] R. Rehm, M. Walther, J. Schmitz, J. Fleißner, J. Ziegler, W. Cabanski, R. Breiter, Proc. of SPIE 6294, 629404 (2006); doi:10.1117/12.680495
[3] R. Rehm, V. Daumer, T. Hugger, N. Kohn, W. Luppold, R. Müller, J. Niemasz, J. Schmidt, F. Rutz, T.Stadelmann, M. Wauro, A. Wörl, Proc. of SPIE 9819, 98190X-1 (2016); doi: 10.1117/12.2223887
[4] R. Müller, J. Niemasz, L. Kirste, V. Daumer, A. Janaszek, J. Jureńczyk, R. Rehm, *Proc. of SPIE* 10914, 1091416 (2019); doi: 10.1117/12.2506624
[5] R. Müller, V. Gramich, M. Wauro, J. Niemasz, L. Kirste, V. Daumer, A. Janaszek, J. Jureńczyk, R. Rehm, Infrared Physics & Technology 96, 141-144 (2019); doi: 10.1016/j.infrared.2018.10.019

DOI 10.5162/sensoren2019/3.2.1

Droplet-on-cantilever approach for determining the mass of magnetic particles

Wilson Ombati Nyang'au[1,2], Andi Setiono[1,3], Prabowo Puranto[1,3,4], Maik Bertke[1], Hutomo Suryo Wasisto[1], Thilo Viereck[5], Harald Bosse[4], Erwin Peiner[1]

[1] Institute of Semiconductor Technology (IHT) and Laboratory for Emerging Nanometrology (LENA), Technische Universität Braunschweig, Braunschweig, Germany
[2] Department of Metrology, Kenya Bureau of Standards, Nairobi, Kenya
[3] Research Center for Physics, Indonesian Institute of Sciences (LIPI), Banten, Indonesia
[4] Precision Engineering Division, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany
[5] Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik (EMG), Technische Universität Braunschweig, Braunschweig, Germany

E-mail: wilombat@tu-braunschweig.de, a.setiono@tu-braunschweig.de, m.bertke@tu-braunschweig.de, h.wasisto@tu-braunschweig.de, t.viereck@tu-braunschweig.de, harald.bosse@ptb.de, e.peiner@tu-braunschweig.de

Abstract

In this study, we present a technique that involves depositing a droplet laden with microparticles (µPs) onto a defined sensing area on a microcantilever sensor. This process entailed the optimization of dispensing parameters (pressure *p* and pulse time *t*) to the generation of a small water droplet volume (V_d = 53.0 ± 2.72 pL). Subsequently, the evaporation trends of the particle carrier fluid, i.e., the water droplets on the sensor were monitored and a total evaporation time of t_{ev} = 40 ± 2 s was determined. By using monodispersed particles in water, i.e., magnetic polystyrene matrix particles (MPS) and poly methyl methacrylate (PMMA) µPs, and adsorbing them on a dynamic cantilever sensor, the mass and number of these µPs were measured and determined comparatively using both resonant frequency measurements and scanning electron microscopy (SEM), respectively. As a result, we have observed and reported monolayer particles assembled on the sensor with a particles count of at least 203.

Key words: Microcantilever, droplet, mass, magnetic particles, resonant frequency

1. Introduction

Over the years, there has been a growing dynamic demand in determining the mass of liquidborne particles. Much of such interests and applications have increasingly been exploited in fields such as medicine and biological sciences, optics, printing and coating, and electronics manufacture. In medicine, for instance, magnetic particles have been used for cell sorting and separation, and as contrast reagents in magnetic resonance imaging (MRI). Moreover, droplet dispensing has been utilized in drug discovery and delivery as well as hyperthermia treatment [1–3]. The mass of such particles plays a critical role in these applications. However, most available equipment for determining their masses, e.g. mass spectrometers and microchannel resonators [4], utilize complicated technology and are quite expensive. Thus, a low-cost and direct approach involving dipping of a cantilever [5] into a magnetic particles solution to adsorb such particles has also recently been employed. This approach, however, yields a low quality factor (Q ~180) [5] due to randomized adsorption of the particles on the whole cantilever surface. Contrarily, our previous study has utilized a liquid-dispensing mechanism to sub-sample and load water droplets (~3.44 µL) laden with and without particles, i.e., ~1.96 µm poly methyl methacrylate (PMMA) and ~4.97 µm polystyrene (PS) from Sigma-Aldrich Inc., onto a silicon-based cantilever sensor (see Figure 1) to achieve a Q ~1224 [6]. As a self-validating process, particle-laden droplets were also deposited on surface-treated silicon substrates and upon solvent evaporation, particle count estimations were performed using scanning electron microscopy (SEM). The particle counting process was, however, curtailed due to the observed clustering of non-

uniform multilayered particles at the edges of the droplet after solvent evaporation [6].
The current study has therefore been intended to optimize dispensing parameters so as to generate and deposit a small droplet (< 3.44 µL) on a microcantilever using a liquid dispenser (1500XL, Nordson EFD Inc., USA); to monitor and determine the total evaporation time (t_{ev}) and trends thereof for water droplets of known volume (V_d); and also to determine the mass or number of monodispersed microparticles (µPs). The µPs used were mainly magnetic polystyrene matrix microparticles (MPS, Micromod GmbH), and PMMA microparticles (Sigma-Aldrich Inc.). The two test µPs have comparable diameters (~2 µm) and densities; but, have contrasting magnetic properties, i.e., MPS are magnetic whereas PMMA particles are non-magnetic. Magnetic polymer particles (e.g., MPS) are synthesized by embedding magnetic nanoparticles (pure metal such as Fe, Co and Ni or their oxides) into a non-magnetic spherical (polymer) matrix [7]. The ~2 µm MPS, for instance, have a polymer core (PS) and magnetic shell consisting of iron oxide nanoparticles. Hence, MPS exhibit a superparamagnetic behavior. However, similar to silicon, PMMA has magnetic susceptibility of less than 1 and is therefore diamagnetic in nature. The essence of using these two µPs was to compare the results and assess the performance of our measurement approach.

2. Substrate preparation and sensor fabrication

Initially, silicon substrates were diced from *n*-type (100) silicon wafers (Siegert Wafer GmbH, Germany) with resistivity and thickness of 1 - 10 Ωcm and 275 ± 15 µm, respectively, into ~15×15 mm^2 and ~30×30 mm^2 pieces. The latter were used for fabrication of cantilever sensors whereas the former, upon surface treatment with oxygen plasma for 30 s, were used as particle adsorbing medium. All the diced samples were cleaned by (a) boiling them for 5 min in an oxidant solution containing a mixture of sulfuric acid (H_2SO_4, 96%) and hydrogen peroxide (H_2O_2, 30%) in a volume ratio of 1:1, (b) immersing in a water bath for 5 min, (c) thoroughly rinsing with deionized water and (d) blow-drying with nitrogen [6] prior to surface treatment or fabrication processes. All the cantilever sensors were fabricated utilizing *n*-type (100) silicon (as base material) and bulk-micromachining technology, discussed elsewhere [8, 9]. The fabrication process mainly comprises thermal oxidation, photolithography, dopant diffusions (boron and phosphorus), etching processes (i.e., HF dip and cryogenic etching), contact holes formation, and metallization. Upon fabrication and prior to use, the cantilever sensors were thoroughly cleaned by immersing or soaking them in acetone, rinsing with deionized water and freely drying them at ambient conditions. Thereafter, we characterized the fabricated cantilever(s) by determining for instance their resonant frequency f_0 and geometry (e.g. thickness h_c, length L and width w).

3. Experimental design and procedure

3.1 Optimization of dispensing parameters and measurement of droplet volume

Large droplets on microstructures are bound to spill-off and sometimes move towards the clamped-end of the cantilever [6]. Furthermore, since the target droplet deposit area on the cantilever is limited and small, it is therefore necessary to optimize the dispensing parameters such as output air pressure (p) and pulse duration (t) and dispensing tip size so as to generate small droplets. The knowledge of the droplet size is thus helpful in precisely depositing and localizing the droplet on a defined region of the cantilever. In our case, the target droplet deposit point on the cantilever was ~ 400 µm from the free-end. For optimizing p and t, and determining the droplet volume V_d (Figure 3), deionized water was used as the test liquid. Due to the fragility of the cantilever sensors, we initially dispensed single water droplets on treated bulk Si substrates so as to estimate optically the droplet volume (V_d = 2.78 ± 0.28 nL). This was done by capturing droplet images, and analyzing them using the low-bond axisymmetric drop shape (LB-ADSA) method [10]. For comparability and better precision, V_d was also determined from the cantilever-based resonant frequency shift Δf measurements.
The water droplet was deposited on the cantilever using a liquid dispenser fitted with: a barrel (filled with deionized water), a dispensing needle (with internal diameter ≈ 0.10 mm) and air pressure supply. The desired output air pressure and pulse time were set and adjusted accordingly, and initially the test liquid was freely dispensed on a test paper

instead of a test sample. At this stage, the dispensing tip would be aligned perpendicularly on top of the test surface (see Figure 1). The positioning and alignment of the tip on the sample surface were accomplished using the 3D coordinate micropositioning system. To minimize the errors due to misalignment and possibility of breaking the fragile cantilever, a USB digital microscope (Mz-902, Oowl Tech Ltd.) was further integrated in the alignment process in order to provide a clear view.

Before and after depositing the water droplet on the cantilever, the resonant frequency f_0 was measured (fundamental in-plane vibration mode). The actuation and vibration sensing mechanisms of the cantilever consisted of a piezoelectric shaker and a U-shaped piezoresistive Wheatstone bridge (WB), respectively. A Lock-in amplifier (MFLI, Zurich Instruments Ltd.) was used to supply the WB voltage (1 V_{dc}) and the excitation signal (9.9 V_{pp}). The MFLI was connected and controlled (by software) from a computer, and connections to and from our electronics were done through coaxial cables and SMA connectors.

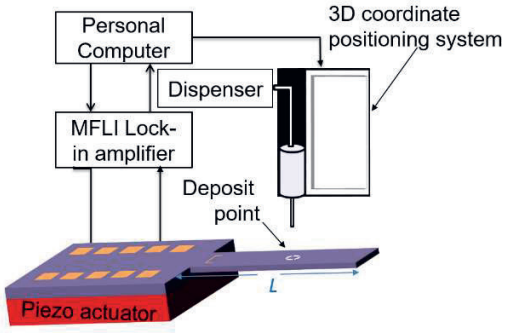


Figure 1: (a) Droplet deposition system and resonant frequency measurement set-up and (b) dispensed droplet on a cantilever.

During droplet dispensing, the tip was moved in close proximity of the cantilever sensing area and an output pressure p was applied for a specified pulse duration t. Consequently, the droplet mass m_d would be determined at a point x from the fixed-end of the cantilever of length L. With the measured resonant frequency shift Δf, the effective mass of the cantilever m_{eff} and the normalized mode-shape function $\phi(x_{md})$, the drop mass can be calculated using Equations (1a) and (1b) [11]. The parameter λ (1b) represents a modal constant, which is about 1.875 for the first mode.

By varying dispensing time and output pressure, various droplet sizes were obtained. It was clearly observed (Figure 2) that for the same dispensing pressure (60 mbar), the size of the water droplet increased with the pulse duration (t). By multiplying m_d with the water density (ρ_w = 1 gcm^{-3}), we obtain the drop volume. Obviously, a small droplet size V_d = 53.0 ± 2.7 pL was measured for t = 10 ms with p = 60 mbar. Similarly, V_d = 76.8 ± 7.7 pL corresponds to t = 50 ms for p = 60 mbar. Consequently, we used low output air pressure (60 mbar) and the shortest pulse duration (10 ms) to generate and realize all the subsequent droplet volumes. The quality factor before ($Q_b \approx 2100 \pm 210$) and after (Q_a = 181 ± 43) depositing the water droplet shows a significant reduction due to damping of the resonator. By assuming a spherical-shaped droplet of $V_d \approx$ 53 pL, a diameter ~46.6 μm was computed; which is much smaller than the maximum width of the cantilever sensor ($w \approx$ 170 μm).

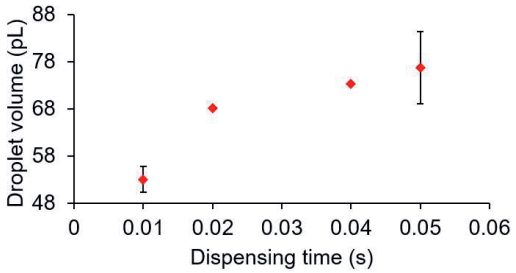


Figure 2: Volume of water droplets dispensed with a needle tip (diameter 0.10 mm) on a cantilever under constant air pressure (60 mbar) at ambient conditions (T ≈ 22.2 ± 0.5 °C; rH ≈ 30 ± 3%)

$$m_d = 2m_{eff}\frac{\Delta f}{f_0}\phi^2(x_{m_d}) \quad (1a)$$

$$\phi(x_{m_d}) = \frac{1}{2}\left[\cos\left(\lambda\frac{x}{L}\right) - \cosh\left(\lambda\frac{x}{L}\right) - \frac{\cos(\lambda)+\cosh(\lambda)}{\sin(\lambda)+\sinh(\lambda)}\left(\sin\left(\lambda\frac{x}{L}\right) - \sinh\left(\lambda\frac{x}{L}\right)\right)\right] \quad (1b)$$

4. Evaporation dynamics of small droplets on microcantilever surface

Water droplets on a cantilever sensor are bound to evaporate at ambient conditions. In our study, the loss of water droplet mass from the sensing surface was monitored by measuring the shift in resonant frequency over time. This was necessary to determine the droplet evaporation time t_{ev}, and the minimum particle adsorption or waiting time prior to resonant frequency measurement. Therefore, in determining t_{ev}, the loss of droplet mass was initially monitored over time under ambient conditions through an offline resonant frequency measurement system (using a MFLI lock-in-amplifier). It should be noted (Figure 3) that as the water droplet was deposited on the cantilever sensor, an immediate shift in resonant frequency was observed, i.e., from f_0 ~ 205.113 kHz to f_r ~ 203.843 kHz), which corresponds to the added mass by the droplet. The total time taken for the cantilever to restore from f_r to f_0 is equivalent to the drop evaporation time, t_{ev} = 40 ± 2 s.

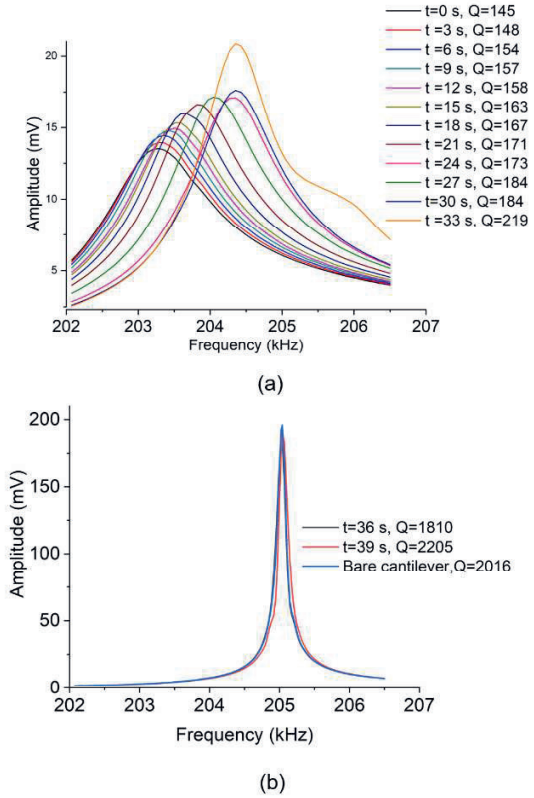


Figure 3: Offline monitoring and measurement of frequency shifts with time for an evaporating water droplet on a cantilever sensor at ambient conditions.

It has clearly been observed (Figure 3) that once the water droplet is loaded onto the cantilever, the frequency (f_0) shifts for ~5 s to a minimum whereupon the sensor momentarily stabilizes after droplet-deposit impact before uniformly and exponentially losing the water into the ambient atmosphere. Afterwards, it was however observed that the remaining water on the sensor surface evaporates rapidly and instantaneously (within ~3 s); which then culminates to a rapid shift of frequency back to f_0. This evidently shows an instantaneous change (increase) in the evaporation rate. Contrarily, Arcamone *et al.* [12] observed a decrease in the evaporation rate with time during the evaporation process.

A spherical droplet cap is characterized by four different parameters: the drop height (h_d), the contact radius (r_c), the radius of the sphere forming the spherical cap (R_S), and the contact angle (θ, in radians). By geometry, the relationships between the two radii (r_c and R_s), θ, and the volume of the spherical cap (V_c) at any instant in time are [13]:

$$r_c = R_s \sin\theta \qquad (2a)$$

$$V_c = \pi R_s^3 \frac{\beta}{3\sin^3\theta} \qquad (2b)$$

where β = 2 - 3cos(θ) + cos^3(θ). The height of the spherical cap above the solid surface is given by: $h_d = R_s[1 - \cos(\theta)]$ or $h_d = r_c\tan(\theta/2)$. During droplet evaporation, the volume of the droplet on the solid surface at any time t can be computed using the following equation [13]:

$$V_c^{2/3} = V_{c,i}^{2/3} - \frac{2}{3} Kf(\theta)\cdot t \qquad (3)$$

where, $V_{c,i}$ denotes the initial droplet volume (at the time t = 0 s) and 2/3$Kf(\theta)t$ represents the volume decrease with time. The function $f(\theta)$ = -cos{θ/[2ln(1-cosθ)]} is for the contact angle and for constant contact-angle evaporation mode: $f(\theta) = f(\theta_0)$, where θ_0 is the contact angle prior to evaporation; and $K \approx 12.37467091 D\Delta c/(\rho_L \beta^{1/3})$, where the symbols D, Δc and ρ_L denote the diffusivity of the vapor molecules in the gas (air) (cm^{2}s^{-1}), the difference between the saturation concentration of water vapor at the sphere surface c_s (gcm^{-3}) and the ambient vapor concentration at infinite distance c_∞ (gcm^{-3}) to the droplet, and the liquid density (gcm^{-3}), respectively. For a water droplet, $\rho_l = \rho_w \approx$ 1 gcm^{-3}, $D \approx$ 0.282 cm^{2}s^{-1} and $c_S = p_v \times M_w/(RT)$ [13] is ~1.7081×10^{-4} gcm^{-3} given a molecular mass M_w of 18.01528 gmol^{-1}, the gas constant

R = 8.314 J(mol·K)$^{-1}$ and the vapor pressure p_v in air of 22.2 mmHg (≅ 29.6 mbar) at a temperature $T \approx 24°C$ [14]. Based on Equation (4), we can therefore estimate the total time taken for the liquid droplet to completely vanish from the surface of a cantilever sensor ($\theta \sim 40°$) to be about 38.8 s, which agrees very well with experimental droplet evaporation time ($t_{ev} \approx 40$ s). The quality factor was also observed to rapidly increase from about 200 (see Figure 3a) to more than 1800 (Figure 3b) towards the last phase of evaporation.

$$t_{ev} = \frac{3/2\, V_{c,i}^{2/3}}{Kf(\theta)} \qquad (4)$$

5. Particles adsorption and measurements

5.1 Adsorption of particles

A particle-laden droplet (~53 pL) was sampled and dispensed on Si (100) substrates, and cantilever sensors (at a point ~400 µm from the free end). In this study, the approximate particle concentrations used were ~1% (w/v) and ~1.25% (w/v). Before drawing a test particle solution, the aqueous µPs suspensions were thoroughly mixed and agitated in an ultrasonic bath to homogenize the particle solution. Consequently, we assumed a uniform mass concentration in the sampled droplet(s). Upon dispensing the droplet containing µPs, a time lapse equivalent to t_{ev} was allowed for the water to completely evaporate and particles to adsorb on the sensing surface (Figure 4).

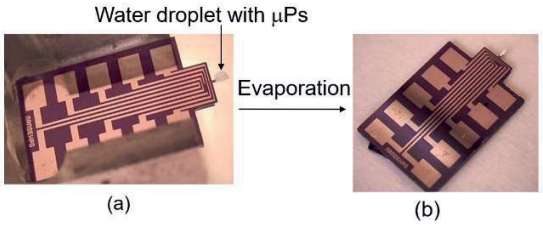


Figure 4: (a) Particle-laden droplet on a cantilever. Solvent evaporation leads to (b) particle adsorption on sensor.

5.2 Resonant frequency measurements

The wet mass and dry mass of the droplet can be measured immediately after depositing the droplet and evaporating the droplet, respectively. In both cases, a shift in resonant frequency is observed. Firstly, in case of a wet drop, the resonant frequency of the unloaded cantilever (f_0) shifts to f_{r1} as a droplet lands on a sensor (i.e., cantilever + water + µPs). Secondly, after a time lapse of t_{ev} the water is expected to completely vanish or evaporate from the sensor thereby shifting the resonant frequency (f_{r1}) to f_{r2} (i.e., resonant frequency of cantilever – water + µPs). The mass m_p of the adsorbed µPs can therefore most conveniently be computed (Equation 1) from the frequency shift of f_{r1} to f_{r2}, i.e., $\Delta f = f_{r2} - f_{r1}$. Similarly, the mass of the particle carrier fluid (i.e., water) can at the same time be determined from the frequency changes of f_{r1} (wet mass) to f_{r2} (dry mass); which nearly corresponds to the previously discussed evaporation trends. In this study, however, since evaporation time and droplet size have previously been determined, only f_{r2} and f_0 were therefore of essence. Subsequently, the mass of the adsorbate (Figure 5a) would then be determined from the frequency shift (Figure 5b) using Equation (1). For a mass at point $x \approx 600$ µm and $x \approx L = 1000$ µm (see Figure 1), the value of $\phi(x_{md})$ is about -0.46 and -1, respectively.

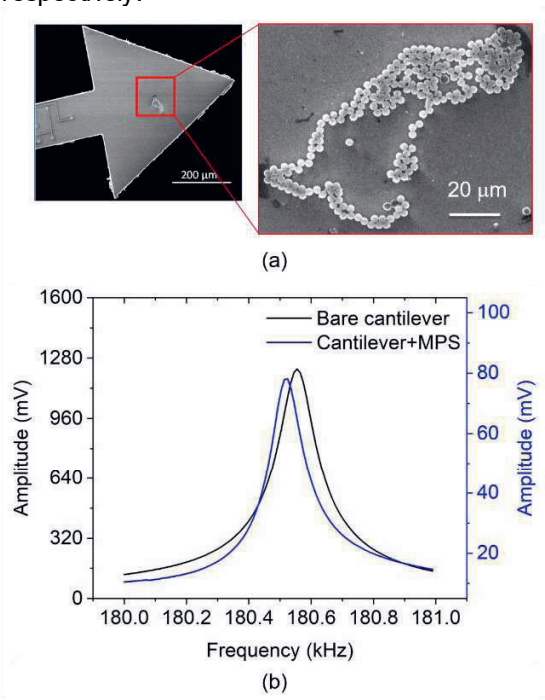


Figure 5: (a) SEM image of a triangular cantilever with MPS µPs, and a completely self-assembled monolayer of µPs (inset). (b) Resonant frequency measurement of a bare cantilever with and without MPS µPs.

Furthermore, given the knowledge of the particle sizes, i.e. 1.83 ± 0.03 µm (MPS) and 1.90 ± 0.05 µm (PMMA), together with their respective densities of 1.1 gcm^{-3} (MPS) and

1.18 gcm^{-3} (PMMA), the corresponding number of adsorbed particles can be estimated from the mass ratio of the adsorbate to the respective single particle. The mass of the latter is calculated from the product of particle volume (assuming spherical-shape) and particle density.

5.3 Microscopic particle analysis

Particles dispersed in a water droplet can assemble themselves on a solid surface as the solvent evaporates. Depending on solid-surface properties and particle concentration, particles may form monolayers or multilayers or a monolayer segment (at the center) combined with a cluster-ring of multilayers [6]. In the present study, µPs were deposited and adsorbed on a hydrophilic Si surface. The latter is highly adhesive and allows convective flow of fluid carrier thereby leading to the self-assembly of particles upon solvent evaporation. Initially, a particle count estimate from SEM micrographs for single droplets deposited on Si substrates was found to be 104470 ± 420. However, by tuning the particle concentration and droplet size, a small particle assembly was realized on the cantilever sensors.

Firstly, the particles were adsorbed on the cantilever and the resonant frequency thereof was measured (cf. Figure 5b). Secondly, the same sample was mounted on the SEM and imaging of particles was done. Subsequently, a particles count would be done manually or with ImageJ [15]. Lastly, prior to the next measurement series, the cantilever would be cleaned as earlier prescribed. Typically, a monolayer of particles was adsorbed on the sensor (Figure 6a) with a particles count of 203 and in some instances, a monolayer plus cluster(s) of non-uniform multilayers would be observed (Figure 6b). In an instance like the latter, only particles on the top layer can be counted with certainty to give minimum number of the adsorbate (e.g. > 300 µPs, see Table 1).

In Table 1, we present a summary of particle counts (from SEM imaging) and compare the same with the number of particles calculated from resonant frequency measurements. For the same adsorbate, the number of particles due to both point mass ($\Phi(x_{md}) > -1$) and distributed mass ($\Phi(x_{md}) \approx -1$) was calculated based on Equation (1). From both mass distributions, disparities in the number of particles were observed. Nonetheless, the number of µPs due to distributed mass was

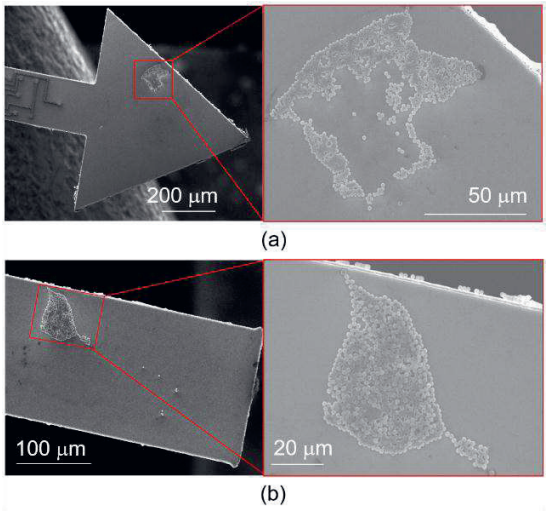


Figure 6: SEM images of adsorbed µPs on cantilevers with triangular (a) and rectangular (b) free-ends. Particle distribution profiles consisting of monolayer and mono/multilayers are in inset, respectively, in (a) and (b).

Table 1: A comparison of number of µPs from resonant frequency shifts and SEM imaging

µPs type/ concentration (%w/V)	Cantilever sensor free-end configuration and target point of µPs deposition	Resonant frequency measurements				SEM µPs count
		f_0 [Hz]	Δf [Hz]	Calculated number of µPs		
				with point mass	with distributed mass	
MPS (~1)	Triangular (x ~ 600 µm)	181452	-15.2	1375	292	203
	Triangular (x ~ 600 µm)	180530	-26.2	2379	506	229
	Rectangular (x ~ 716 µm)	205468	-36.4	1007	377	> 300
PMMA (~1)	Triangular (x ~ 600 µm)	204142	-117.7	5149	1095	664
	Rectangular (x ~ 600 µm)	205468	-94.2	4094	871	504
MPS (~1.25)	Rectangular (x ~ 685 µm)	204142	-378.3	10551	3440	1009

fairly comparable with particles count from the SEM. Generally, the observed disparities in particle counts (from the two approaches) would be attributed to some drift and inconsistencies in positioning of the dispensing tip, and therefore the droplet, on the sensor and to the formation of multilayers. The former is a consequence of misalignment and may have led to spilling or adsorbing of particles along the edges of the beam (cf. inset in Figure 6b), while the latter introduces uncertainties in particle count estimations. This study therefore intends to pick up from here to resolve the aforementioned challenges, and calibrate the cantilever sensors for better comparative particle measurement and count estimations.

6. Conclusion

In this work, the droplet dispensing parameters (p and t) were optimized and a small water droplet volume of $V_d = 53.0 \pm 2.7$ pL was realized. Furthermore, a droplet evaporation time of $t_{ev} \approx 40$ s was determined and found to compare well with the theoretical estimation. Particle adsorption on hydrophilic Si substrates and piezoresistive microcantilevers has also been presented, followed by a determination of their mass and number based on both resonant frequency shift measurements and SEM image analysis for particle counting.

Acknowledgments

W. O. Nyang'au is grateful to the German Federal Ministry for Economic Cooperation and Development (BMZ) for a doctoral scholarship within the Braunschweig International Graduate School of Metrology (B-IGSM). A. Setiono and P. Puranto acknowledge support from the Ministry of Research, Technology and Higher Education of the Republic of Indonesia under no. 343/RISET-Pro/FGS/VII/2016 (World Bank Loan No. 8245-ID). M. Bertke is grateful for the support received from Niedersächsisches Vorab, Germany, through the "Quantum- and Nanometrology (QUANOMET)" initiative within the project NP 2-2. The authors are further grateful to Angelika Schmidt, Juliane Breitfelder, Aileen Michalski, and Karl-Heinz Lachmund for their assistance during preparation of research tools and valuable technical assistances.

7. References

[1] T. Lindemann and R. Zengerle, "Droplet Dispensing," in *Encyclopedia of Microfluidics and Nanofluidics*, D. Li, Ed., New York, NY: Springer New York, 2015, pp. 641–652.

[2] T.-J. Yoon *et al.,* "Multifunctional nanoparticles possessing a "magnetic motor effect" for drug or gene delivery," (eng), *Angewandte Chemie (International ed. in English)*, vol. 44, no. 7, pp. 1068–1071, 2005.

[3] W.-L. Chou, P.-Y. Lee, C.-L. Yang, W.-Y. Huang, and Y.-S. Lin, "Recent Advances in Applications of Droplet Microfluidics," *Micromachines*, vol. 6, no. 9, pp. 1249–1271, 2015.

[4] T. P. Burg *et al.,* "Weighing of biomolecules, single cells and single nanoparticles in fluid," (eng), *Nature*, vol. 446, no. 7139, pp. 1066–1069, 2007.

[5] M. Park, H. Choi, S. J. Kim, C. S. Kim, and T. Kouh, "Resonance Frequency Shift of Microcantilever via Surface Adsorption of Magnetic Nanoparticles," *JKMS*, vol. 28, no. 2, pp. 49–52, 2018.

[6] W. O. Nyang'au *et al.,* "Adsorption and detection of microparticles using silicon microcantilevers," (to be published in the Journal of Physics: Conference Series, 29th Micromechanics and Microsystems Europe workshop).

[7] C. Ruffert, "Magnetic Bead-Magic Bullet," (eng), *Micromachines*, vol. 7, no. 2, 2016.

[8] H. S. Wasisto, S. Merzsch, E. Uhde, A. Waag, and E. Peiner, "Handheld personal airborne nanoparticle detector based on microelectromechanical silicon resonant cantilever," *Microelectronic Engineering*, vol. 145, pp. 96–103, 2015.

[9] M. Bertke *et al.,* "Analysis of asymmetric resonance response of thermally excited silicon micro-cantilevers for mass-sensitive nanoparticle detection," *J. Micromech. Microeng.*, vol. 27, no. 6, p. 64001, 2017.

[10] A. F. Stalder *et al.,* "Low-bond axisymmetric drop shape analysis for surface tension and contact angle measurements of sessile drops," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 364, no. 1-3, pp. 72–81, 2010.

[11] S. Schmid, L. G. Villanueva, and M. L. Roukes, *Fundamentals of*

Nanomechanical Resonators. Cham: Springer International Publishing, 2016.
[12] J. Arcamone, E. Dujardin, G. Rius, F. Pérez-Murano, and T. Ondarçuhu, "Evaporation of femtoliter sessile droplets monitored with nanomechanical mass sensors," (eng), *The journal of physical chemistry. B*, vol. 111, no. 45, pp. 13020–13027, 2007.
[13] H. Y. Erbil, G. McHale, and M. I. Newton, "Drop Evaporation on Solid Surfaces: Constant Contact Angle Mode," *Langmuir*, vol. 18, no. 7, pp. 2636–2641, 2002.
[14] E. L. Cussler, *Diffusion: Mass transfer in fluid systems,* 3rd ed. Cambridge, New York: Cambridge University Press, 2009.
[15] C. T. Rueden *et al.,* "ImageJ2: ImageJ for the next generation of scientific image data," (eng), *BMC bioinformatics*, vol. 18, no. 1, p. 529, 2017.

DOI 10.5162/sensoren2019/3.2.2

Ein neuer Ansatz für eine präzise automatische Temperaturkalibrierung von Microhotplates

Robin Diener[1], Caroline Schultealbert[1], Tobias Baur[1], Andreas Schütze[1], Tilman Sauerwald[1]
[1]Lehrstuhl für Messtechnik, Universität des Saarlandes, Campus A5 1, 66123 Saarbrücken, Deutschland
E-Mail: c.schultealbert@lmt.uni-saarland.de

Zusammenfassung

In dieser Arbeit wird eine Methode für die Temperaturkalibrierung von Microhotplates am Beispiel von MEMS-Gassensoren vorgestellt. Als Bezugspunkte dienen Fest-Flüssig-Phasenübergänge von Referenzmaterialien. Die latente Wärme des Phasenübergangs führt zu einem Anstieg der Wärmekapazität, was sich bei einer definierten Heizrampe in einem lokalen Maximum der Heizleistung widerspiegelt. Die Methode wird mit einem AS-MLV-P2 Gassensor (ams Sensor Germany GmbH) mit integriertem Membranheizer getestet. Als Referenzmaterial für die Entwicklung der Methode wird Hexatriacontan verwendet, dessen Schmelzpunkt bei 76 °C als erster und die Raumtemperatur (22 °C) als zweiter Kalibrierpunktpunkt dient. Durch den hohen Siedepunkt von 497 °C kann die Reproduzierbarkeit des Verfahrens getestet werden. Der Schmelzpunkt kann als Onset-Punkt des Schmelzvorgangs mit einer Wiederholgenauigkeit von ±0,15 K ermittelt werden. Bei variierenden Versuchsbedingungen (Heizrate) liegt die Wiederholgenauigkeit weiterhin deutlich unter 1 K. Weitere Versuche werden mit den Referenzmaterialien Pyrazin und Urotropin durchgeführt, da diese durch einen relativ niedrigen Siedepunkt von (115 °C und 263 °C) und ihre gute Löslichkeit in Wasser und Alkoholen auch innerhalb von Fertigungsprozessen reversibel appliziert und entfernt werden können. Bei beiden Materialien konnten die Phasenübergangstemperaturen aus dem Heizleistungsverlauf präzise bestimmt werden.

Keywords: Temperaturkalibrierung, MEMS Sensoren, Gassensorik, Microhotplate, Phasenübergang

Motivation

Die präzise Messung und Steuerung der Temperatur von Microhotplates ist oft fehlerbehaftet. Für eine Messung der Temperatur nach dem Stefan-Boltzmann-Gesetz sind die Emissionskoeffizienten und die Transparenz der Membranen häufig nicht tabelliert [1] und auch experimentell sehr schwer zu bestimmen. Bei einer Kalibrierung über den Temperaturkoeffizienten des Heizerwiderstands wird das gesamte Bauteil (z.B. in einer Klimakammer) erhitzt. Dabei wirkt sich der Temperaturkoeffizient auch auf die signifikanten Zuleitungswiderstände aus, die außerhalb der geheizten Membran liegen und daher im Regelbetrieb nicht erwärmt werden. Allerdings sind für viele MEMS-Anwendungen exakte Messung und Regelung der Temperatur notwendig. Ein Beispiel sind Halbleitergassensoren, bei denen viele Prozesse (z.B. die thermisch aktivierte Leitfähigkeit und stattfindende Oberflächenreaktionen mit Aktivierungsenergie) stark temperaturabhängig sind [2]. Insbesondere für die Übertragbarkeit von Kalibriermodellen ist eine sensorübergreifende korrekte Temperatureinstellung absolute Notwendigkeit. Derzeit wird hierfür am Lehrstuhl für Messtechnik die meist vorhandene Leistungs-Temperatur-Kennlinie (oft lediglich durch einen Arbeitspunkt im Datenblatt gegeben) in eine gemessene Widerstands-Temperatur-Kennlinie überführt, sodass der Heizerwiderstand als Regelgröße verwendet werden kann [3].

In dieser Arbeit wird deshalb eine Methode für die Temperaturkalibrierung von Microhotplates am Beispiel von MEMS-Gassensoren vorgestellt, die sich sowohl in der Herstellung als auch im Nachgang, beispielsweise beim Anwender, auf jedes Einzelelement anwenden lassen. Die Kalibrierroutine muss daher sowohl einfach und kostengünstig als auch zerstörungsfrei sein. Als Bezugspunkte dienen Fest-Flüssig-Phasenübergänge von Referenzmaterialien, was ein in der Praxis anerkanntes Verfahren zur Temperaturkalibrierung darstellt [4].

Die latente Wärme des Phasenübergangs führt zu einem temporären Anstieg der

Wärmekapazität, was sich bei einer definierten Heizrampe in einem lokalen Maximum der Heizleistung widerspiegelt. Ziel ist es, ein reproduzierbares Merkmal dieses Maximums zu finden, das mit der normalen Betriebselektronik dieser Sensoren vermessen und ausgewertet werden kann. Darüber hinaus werden Materialien gesucht, die einerseits den interessanten Temperaturbereich des Sensors mit ihrer Phasenübergangstemperatur abbilden und andererseits leicht aufzubringen sowie zerstörungsfrei zu entfernen sind, hierfür sind insbesondere sublimierende Materialien von Interesse.

Experimenteller Aufbau und Methoden

Die Methode wird mit einem AS-MLV-P2 Gassensor (ams Sensor Germany GmbH) mit integriertem Membranheizer aus Platin getestet. Der Widerstand der Heizerschicht dient als Führungsgröße der Regelung. Aufgrund des Temperaturkoeffizienten von Platin lässt sich (bei gleichbleibender Temperatur der Zuleitungswiderstände) der Heizerwiderstand eindeutig mit einer Membrantemperatur identifizieren. Der elektronische Aufbau der Heizerregelung ist identisch zu der Elektronik, die am Lehrstuhl für Messtechnik (LMT) für die Vermessung von Gassensoren im temperaturzyklischen Betrieb (engl. temperature cycled operation, TCO) verwendet wird [5]. Sie beinhaltet eine Vierleitermessung, um Zuleitungswiderstände außerhalb des Sensors zu unterdrücken. Um bei der Raumtemperaturmessung Eigenerwärmung zu verhindern, wird der Heizerwiderstand mit einem Stromfluss von unter 1 µA ermittelt. Dieser Widerstand kann daher als zweiter Kalibrierpunkt zur Ermittlung einer Widerstands-Temperatur-Kennlinie verwendet werden.

Für die Methodenentwicklung wird Hexatriacontan als Referenzmaterial verwendet (Schmelzpunkt 74-76 °C [6]), da es aufgrund des geringen Dampfdrucks nach einer Applikation mehrmals reproduzierbar vermessen werden kann. Ein aufgebrachter möglichst kleiner Flake des Materials wird zunächst aufgeschmolzen, sodass sich eine homogene Schicht auf der Hotplate bildet, diese Schicht kann dann mehrmals für die Messung aufgeschmolzen werden.

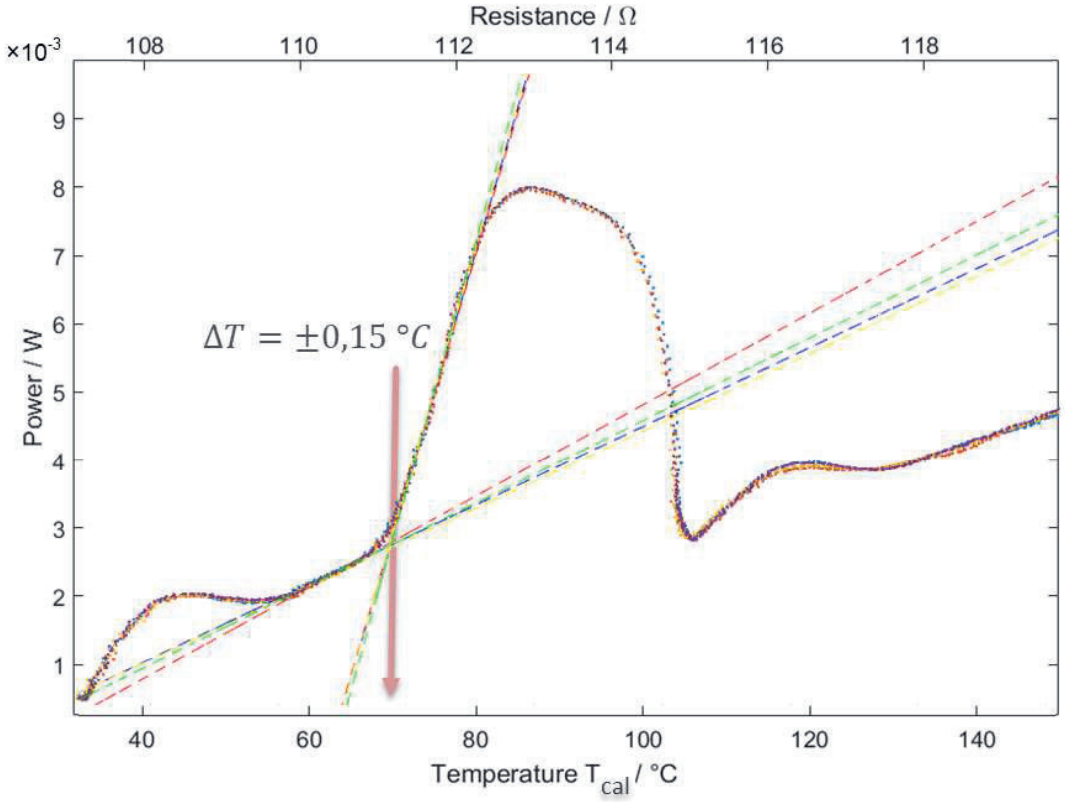


Abb. 1: Mehrmalige Temperaturkalibrierung über die Heizleistung mit derselben Hexatriacontan-Schicht auf der Microhotplate. Die bestimmten Onset-Punkte der Kurven liegen maximal 0,15 °C auseinander.

Als weitere Referenzmaterialien werden Pyrazin (Schmelzpunkt ca. 56 °C, Siedepunkt 115-116 °C [7]) und Urotropin (Sublimationspunkt 263 °C [8]) verwendet. Beide Substanzen können in Wasser oder (in geringerer Konzentration) in Methanol gelöst und dann auf die Hotplate aufgetropft werden.
Für unsere Untersuchungen wurde das Pyrazin als Feststoff (Flake) aufgebracht und das Urotropin als Lösung in Wasser oder Methanol. Das Lösemittel muss bei Raumtemperatur verdunstet werden, bevor der eigentliche Kalibriervorgang startet. Ein Erwärmen der Hotplate, die diese Verdunstung beschleunigt, kann dazu führen, dass die Urotropin-Schicht mit dem entstehenden Gas von der Hotplate abgelöst wird, wodurch die Schicht ohne Phasenübergang von der Hotplate entfernt wird (vermutlich als Partikel im Gasstrom). Für die in Abb. 3 gezeigte Messung wurde eine gesättigte Urotropin-Methanol Lösung verwendet.

Ergebnisse

In Abb. 1 sind mehrere Schmelzkurven derselben Hexatriacontan-Schicht dargestellt. Aufgetragen ist die benötigte Heizleistung über dem Heizerwiderstand (obere x-Achse). Diese kann nach der Temperaturkalibrierung direkt in eine Temperatur überführt werden (untere x-Achse). Angenommen wird hierbei ein linearer Zusammenhang. Die Heizrate entspricht auf Basis der in der Einleitung beschriebenen vorläufigen Temperaturkalibrierung [2] etwa 80 K/s im Temperaturbereich von 30 °C bis 150 °C, geregelt wird auf den entsprechenden Heizerwiderstand.
Die Bestimmung der Schmelztemperatur erfolgt am Onset-Punkt des Heizleistungspeaks, der den Beginn des Phasenübergangs definiert. Bestimmt wird der Onset-Punkt als Schnittpunkt zweier Regressionsgeraden der Heizleistung in Abhängigkeit des Heizerwiderstands. Die erste Gerade wird im linearen Bereich des Aufheizvorgangs nach dem Einschwingen der Regelung und vor dem Auftreten des Schmelzpeaks gefittet. Die zweite Regressionsgerade wird durch den Anstieg des Schmelz-Peaks gelegt. Der Onset-Punkt kann auf diese Art und Weise für vier Messungen mit einer Wiederholgenauigkeit von ±0,15 °C ermittelt werden.

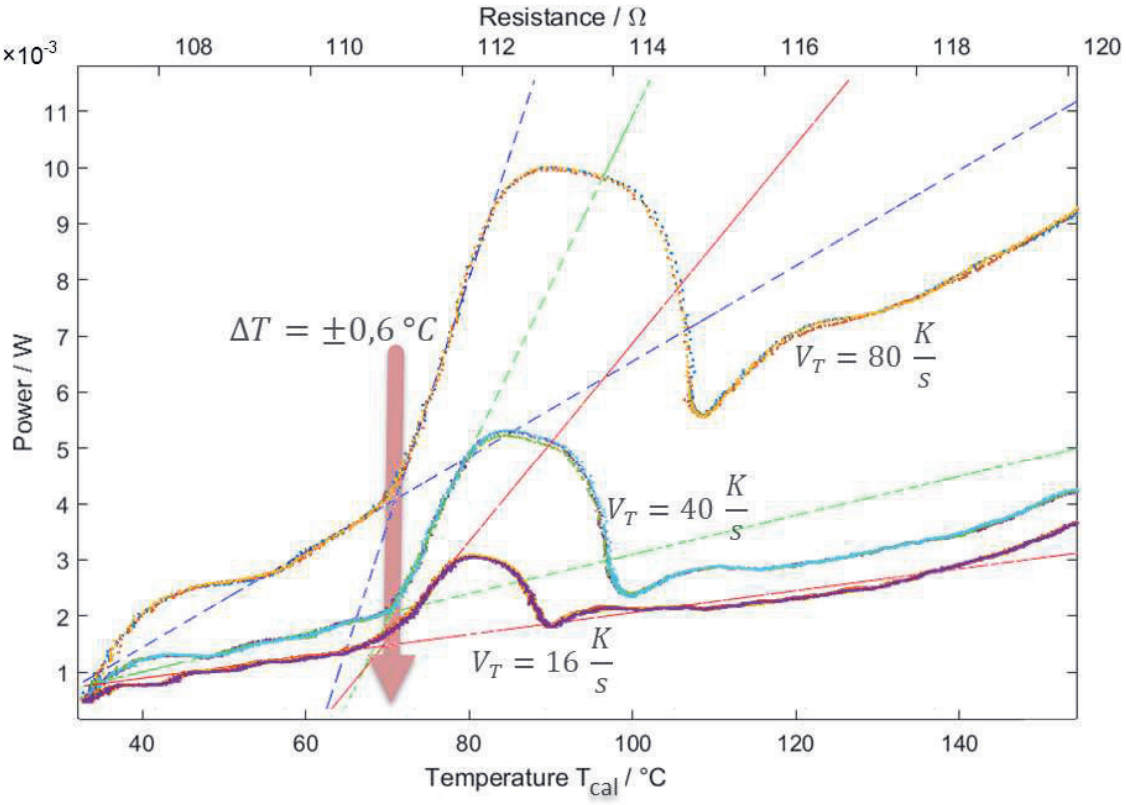


Abb. 2: Auch bei Veränderung der Heizrate, die stellvertretend für andere Einflüsse wie unterschiedliche Mengen Referenzmaterial herangezogen wurde, kann die Bestimmung des Onset-Punktes des Schmelzprozesses auf ±0,6 °C genau erfolgen. Die Verwendung anderer Charakteristika des Schmelzpeaks (z.B. Maximum) sind deutlich weniger erfolgsversprechend, da sie sich durch die Heizratenvariation verschieben.

Abb. 2 zeigt die Vermessung der Hexatriacontan-Schicht bei drei unterschiedlichen Heizraten v_T: 16 K/s, 40 K/s und 80 K/s, Die Auftragung ist analog zu Abb. 1. Durch diese Variation lässt sich abschätzen, wie sich eine Variation der Stoffmenge an aufgebrachtem Phasenwechselmaterial auf die Genauigkeit der Kalibrierung auswirkt. Wir gehen davon aus, dass beide Variationen vergleichbare Fehler hervorrufen, da die Heizleistungen über den gesamten Zeitverlauf der Messung gleich skalieren. Die maximale Abweichung der bestimmten Onset-Punkte beträgt trotz Variation der Heizrate um den Faktor 5 nur ±0,6 °C. Die sehr reproduzierbare Bestimmung der Temperatur bei unterschiedlichen Heizraten lässt auf eine geringe Anfälligkeit für Fehler besonders auch in Bezug auf die aufgebrachte Stoffmenge schließen. In Abb. 2 wird auch deutlich, dass eine Bestimmung anderer Parameter des aufgenommenen Schmelzpeaks, wie beispielsweise das Maximum, keine reproduzierbaren Ergebnisse für die Temperaturkalibrierung liefern, da sie sich in Abhängigkeit von der Heizrate verschieben. Das liegt vermutlich daran, dass der Phasenübergang nicht im gesamten Material homogen erfolgt und dann weitere Parameter wie z.B. Wärmetransport in der Hexatriacontanschicht eine Rolle spielen.

Die Verwendung von Hexatriacontan als Referenzmaterial dient, wie bereits erwähnt, zur Methodenentwicklung, da das Material durch den geringen Dampfdruck vollständig auf dem Heizer verbleibt und die Messungen wiederholbar sind. Eine rückstandsfreie Entfernung des Schmelzmaterials kann bei Temperaturen über 500 °C erfolgen. Eine Beeinflussung der Sensorschicht kann dabei jedoch nicht ausgeschlossen werden. Die Aufbringung von wachsartigen Referenzmaterialien wie Hexatriacontan ist daher nicht für eine zerstörungsfreie Kalibrierung geeignet.

Für eine rückstandsfreie Kalibrierung können Referenzmaterialien mit höherem Dampfdruck bzw. niedrigem Siedepunkt verwendet werden, die, bereits kurz nach dem schmelzen direkt in die Dampfphase übergehen. Ein Beispiel ist Pyrazin, welches beim Erhitzen bereits vollständig verdampft, noch bevor der Siedepunkt erreicht ist.

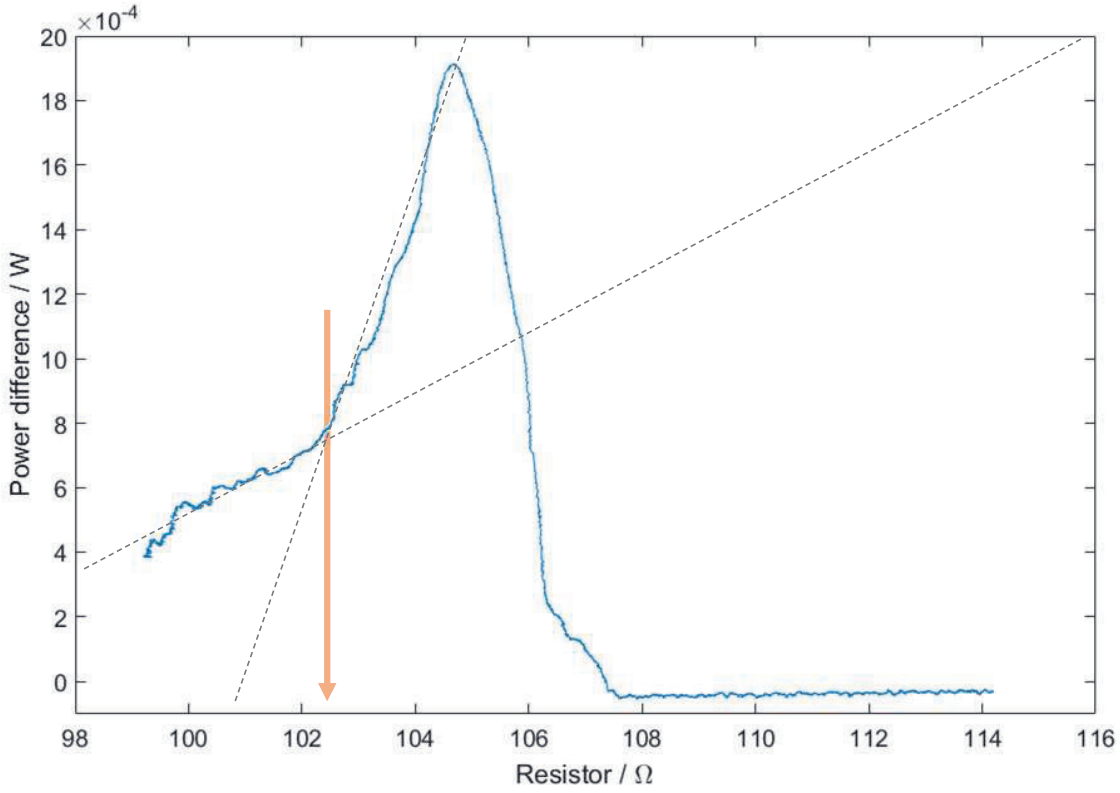


Abb. 3: Kalibrierung der Temperatur über die Bestimmung des Onset-Punkts des Schmelz- und Verdampfungsprozesses von Pyrazin, welches für die automatisierte und fertigungsnahe Kalibrierung besser geeignet ist.

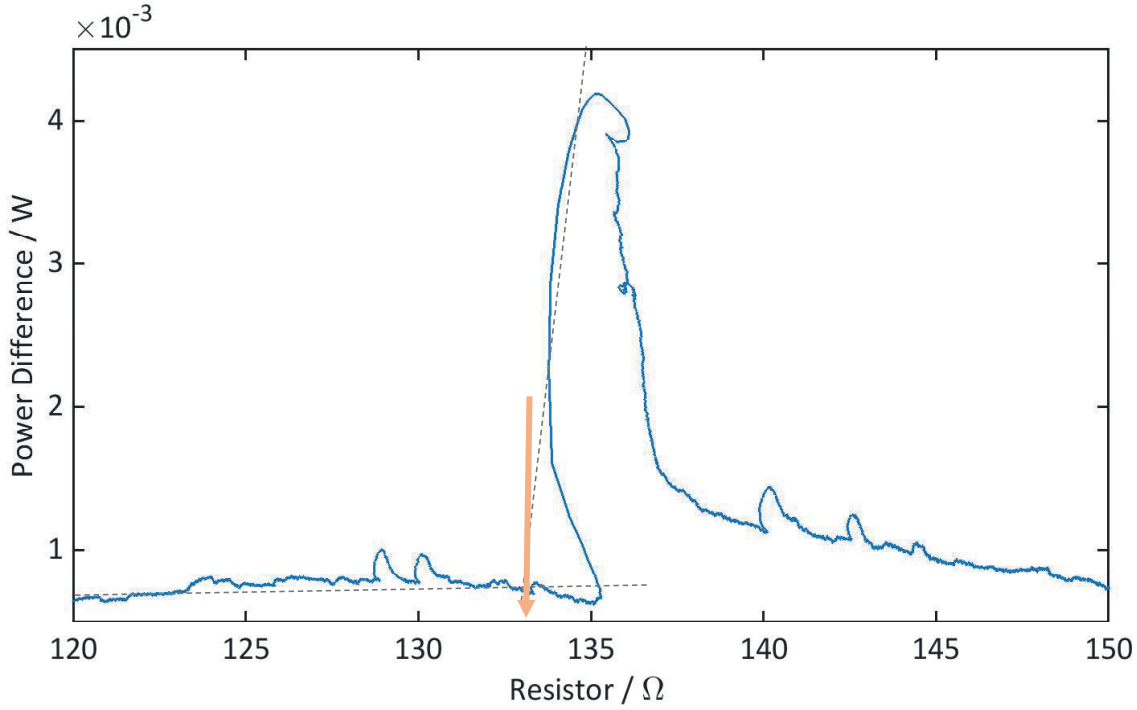


Abb. 4: Kalibrierung der Temperatur über den Sublimationsprozesses von Urotropin, welches aus der Flüssigphase aufgebracht werden kann und dessen Phasenübergang in einem für Gassensoren geeigneteren Temperaturbereich liegt. Die Heizrate und Menge des aufgebrachten Materials sollten noch weiter optimiert werden, um ein reproduzierbares Merkmal, wie den Onset-Punkt bestimmen zu können.

Abb. 3 zeigt eine Heizleistungskurve mit einem aufgebrachten Pyrazin-Flake auf der Schicht, wobei die Basislinie des leeren Heizers abgezogen wurde. Auch hier ist der Onset-Punkt des Schmelzprozesses problemlos zu bestimmen. Nach dem Maximum der Heizleistung liegt die Differenz zwischen dem leeren Heizer und dieser Messung nahezu bei 0 W, sprich es verbleibt kein Material auf dem Heizer, da dieses sich in Form einer erhöhten Wärmekapazität und einer damit erhöhten Heizleistung bemerkbar machen würde (vgl. Kurve vor dem Phasenübergang).
Weitere Versuche mit Pyrazin wurden nicht durchgeführt, da der Temperaturbereich zumindest für Halbleitergassensoren nicht optimal ist. Stattdessen wurden Versuche mit Urotropin durchgeführt, da dieses Material aus der Flüssigphase auf der Microhotplate abgeschieden und anschließend rückstandsfrei sublimiert bzw. zersetzt werden kann. Die Zersetzungsprodukte sind sehr leicht flüchtig, daher ist es naheliegend, dass diese nicht auf der Sensorschicht verbleiben und sie langfristig beeinflussen: Kohlenmonoxid, Kohlendioxid, Stickoxide, Ammoniak und Formaldehyd [8]. Der Sublimationspunkt von Urotropin liegt deutlich besser im Betriebsbereich von Halbleitergassensoren im TCO.

Für die Messung wurde ein Tropfen gesättigte Urotropin-Methanol-Lösung direkt auf der Hotplate platziert. Nach mehreren Stunden an Raumluft ist das Lösungsmittel komplett verdampft und lediglich ein Urotropin-Flake bleibt zurück (ca. 100 nl). Für den Sublimationsprozess wurde die Hotplate bei einer Heizrate von 12 K/s von 50 °C auf 350 °C aufgeheizt. Der Verlauf der Heizleistung (mit abgezogener Basislinie) ist in Abb. 4 dargestellt. Materialmenge, Heizrate und Regelungsparameter sind bei dieser Messung noch nicht optimal abgestimmt, da die Hotplate ein Temperaturmaximum durchläuft bevor sie durch den Sublimationsprozess wieder abgekühlt wird. Für eine eindeutige Bestimmung des Onset-Punktes und damit eine reproduzierbare Temperaturkalibrierung der Hotplate, müssen diese Parameter noch optimiert werden.

Zusammenfassung und Ausblick

Mit den hier gezeigten Messergebnissen konnte gezeigt werden, dass eine Temperaturkalibrierung von Microhotplates mit Genauigkeiten unter 1 K mit Hilfe von Phasenübergängen (Schmelzen, Sublimieren) möglich ist. Die sehr gute Reproduzierbarkeit kann mit schwerflüchtigen Substanzen wie Hexatriacontan gezeigt werden. Für die

Anwendung im Fertigungsprozess sind Substanzen, die aus der Flüssig abgeschieden werden können, wünschenswert, was hier am Beispiel von Pyrazin und Urotropin gezeigt wurde. Ein Nachweis der Reproduzierbarkeit ist hierbei schwieriger. Weitere Untersuchungen zur Optimierung der Prozess-Parameter und Wiederholgenauigkeit der Messungen sind hier noch notwendig. Die vorgestellte Methode ermöglicht die zerstörungsfreie Kalibrierung von MEMS Gassensoren, bei denen die exakte Temperaturkalibrierung jedes einzelnen Sensors aufgrund von Fertigungsschwankungen eine enorme Verbesserung für die Übertragbarkeit von Kennlinien von verschiedenen Sensoren des gleichen Typs bedeutet.

Literaturnachweis

[1] R. Zhang, Y. Gu, D. Zhu, J. Hao, Z. Tang, Thermal measurement and analysis of micro hotplate array using thermography, Sensors and Actuators A: Physical. 2002, 100, 144–152, doi: 10.1016/S0924-4247(02)00061-4

[2] G. Korotcenkov, B. K. Cho, Instability of metal oxide-based conductometric gas sensors and approaches to stability improvement (short survey), Sensors and Actuators B: Chemical. 2011, 156, 527–538, doi:10.1016/j.snb.2011.02.024.

[3] C. Schultealbert, T. Baur, A. Schütze, T. Sauerwald, Facile Quantification and Identification Techniques for Reducing Gases over a Wide Concentration Range Using a MOS Sensor in Temperature-Cycled Operation, Sensors, 2018, 18(3), 744, doi:10.3390/s18030744

[4] S. M. Sarge, E. Gmelin, G. W. H. Höhne, H. K. Cammenga, W. Hemminger, and W. Eysel, The caloric calibration of scanning calorimeters, Thermochimica Acta, 1994, 247, 2, 129–168, doi:10.1016/0040-6031(94)80118-5

[5] T. Baur, C. Schultealbert, A. Schütze, T. Sauerwald, Devicce for the detection of short trace gas pulses, tm – Technisches Messen, 2018, 85, doi: 10.1515/teme-2017-0137

[6] Thermo Fisher GmbH, Hexatriacontane SDS, 2018. Online verfügbar: https://www.alfa.com/de/msds/?language=DE&subformat=CLP1&sku=31641, aufgerufen am 29.04.2019.

[7] Thermo Fisher GmbH, Pyrazine SDS, 2018. Online verfügbar: https://www.alfa.com/de/msds/?language=DE&subformat=CLP1&sku=16326, aufgerufen am 29.04.2019.

[8] Thermo Fisher GmbH, Hexamethylenetetramine SDS, 2018. Online verfügbar: https://www.alfa.com/de/msds/?language=DE&subformat=CLP1&sku=36462, aufgerufen am 29.04.2019.

DOI 10.5162/sensoren2019/3.2.3

Wafer-Level-3D-Integrationsverfahren für hochsensitive optische Sensoren

Lukas Kalwa[1], Stefan Gläsener[1], Jennifer Ruskowski[1], Martin Figge[1], Holger Kappert[1], Holger Vogt[1,2]
[1]Fraunhofer-Institut für Mikroelektronische Schaltung und Systeme IMS, Finkenstraße 61, 47057 Duisburg, Germany
[2]Universität Duisburg-Essen (UDE), Bismarckstraße 81, 47057 Duisburg, Germany

Zusammenfassung

Dieser Beitrag beschäftigt sich mit der 3D-Integration von hochsensitiven Pixelelementen mit einer separat gefertigten Pixelelektronik. Durch Ausnutzung der dritten Dimension, können optische Sensoren mit der zugehörigen interpretierenden und signalverarbeitenden Schaltungslogik, in einem gemeinsamen Wafer-zu-Wafer Bondprozess gefertigt werden. Als Elemente der optischen Sensorik werden schnell reagierende und maximal lichtsensitive Einzelphoton-Lawinenphotodioden (Single Photon Avalanche Diode, SPAD) verwendet. Diese werden in speziellen Makropixelstrukturen zusammengefasst und durch patentierte Auslese-Algorithmen, integriert auf dem Schaltungswafer, ausgelesen. Durch die integrierte Pixelelektronik kann adaptiv auf sich ändernde Umfeldeinflüsse reagiert werden. Zur Herstellung einer qualitativ hochwertigen mechanischen Verbindung zwischen Sensor- und Schaltungswafer wurden die Fertigungsprozess bezüglich Oberflächenrauheit, Waferbiegung, Topologie und Layout optimiert. Für die elektrische Verbindung wurde eine spezielle Vorderseiten-Silizium-Durchkontaktierung verwendet. Die Zuverlässigkeit der Verbindungen konnte durch Temperaturlastwechsel bestätigt werden. Für die SPAD-Sensorelemente wurden die wichtigsten Kenngrößen (Durchbruchspannung, Dunkelzählrate und Quanteneffizienz) im Waferverbund vermessen.

Keywords: SPAD, Optische Sensoren, 3D-Integrationsverfahren, Wafer-to-Wafer Bonding, Backside Illumination (BSI)

1. Einleitung

Forschungs- und Entwicklungsaktivitäten im Bereich der Mikro- und Nano-Elektronik führten in den letzten Jahrzehnten zu immer kleineren Strukturgrößen, was zu einer kontinuierlich wachsenden Integrationsdichte von Sensoren, Speichern und verarbeitenden Schaltungen führte. Auf Grund der physikalischen Limitierung durch die Atomabmessungen stagniert jedoch das enorme technologische Wachstum bei der Beschränkung auf eine Waferebene. Es gibt bereits heute vielfältige Ansätze eine Integration von verschiedenen mikro-/ nanoelektrischen Bauelementen in drei Raumdimensionen zu realisieren und den Trend von kompakten und leistungsstarken Bauelementen fortzusetzen („More than Moore“). Mit der 3D-Integration erschließen sich neben der Möglichkeit einer heterogenen Integration, eine höhere Integrationsdichte, eine Kostenreduktion und kürzere Verbindungswege.

Am Fraunhofer IMS hat sich für den Verbindungaufbau von Sensor und Schaltung das Chip-zu-Wafer-Verfahren etabliert. Hierbei werden über einen Flip-Chip-Prozess die zuvor aus dem Sensorwafer vereinzelten Sensor-Chips auf den Schaltungs-Wafer ausgerichtet und fixiert. Über galvanische Strukturen auf der Chip- und auf der Waferseite wird mit dem Solid-Liquid-Interdiffusion-Verfahren (SLID) ein hochtemperaturbeständiger Kontakt bei niedriger Prozesstemperatur erzeugt. Der Sensor und die Schaltungseinheit müssen sich nicht auf gleich großen Wafern befinden, was eine heterogene 3D-Integration von unterschiedlichen Materialien ermöglicht. Zudem besteht die Möglichkeit eine Vorsortierung durchzuführen, so dass ausschließlich fehlerfrei qualifizierte Chips (KGDs, known-good-dies) aufgebaut werden. Dieser Einzelchipaufbau erfordert jedoch lange Prozesszeiten, insbesondere für kleine Chip-Größen. Eine deutliche Kostenreduktion und Durchsatzsteigerung, im Vergleich zum Chip-zu-Wafer-Verfahren, kann nur durch ein Wafer-zu-Wafer Fügeverfahren erzielt werden.

Dazu wurde in dieser Arbeit ein CMOS-kompatibler Bondprozess für Sensor- und Schaltungs-Wafer aus der hauseigenen 8-Zoll CMOS-Linie entwickelt.

Zur Qualitätsprüfung der entwickelten 3D-Integration und der hergestellten Waferverbindung, wurden die elektrische

Kontaktierung charakterisiert und eine Zuverlässigkeitsprüfung durchgeführt. Um die Kompatibilität des Verfahrens zur Herstellung eines Sensorsystem aus SPAD und ROIC (ReadOut-Integrated Circuit) zu prüfen, wurden abschließend die kritschen Kenngrößen Durchbruchspannung, Dunkelzählrate, Quanteneffizienz der SPADs sowie der Kettenwiderstand der Durchkontaktierung charakterisiert.

2. Prozessablauf

Zur Integration der beiden Wafer wurde ein direkter Bondprozess („Direct-Wafer-Bonding“) mit anschließender Nachbearbeitung, bestehend aus 6 Einzelschritten, entwickelt (siehe Abbildung 1):

1) Separate Fertigung der Sensor- und Schaltungs-Wafer in der CMOS-Linie
2) Direktes Oxid-zu-Oxid-Bonden mit optimierten Oberflächeneigenschaften
3) Ganzflächiges Dünnschleifen mit anschließendem Plasmaätzen
4) Ätzung der Öffnungen zur Durchkontaktierung
5) Elektrische Kontaktierung durch Abscheidung eines Materialschichtstapels
6) Auffüllung der Öffnungen mit Aluminium und Strukturierung der Kontaktierungspads

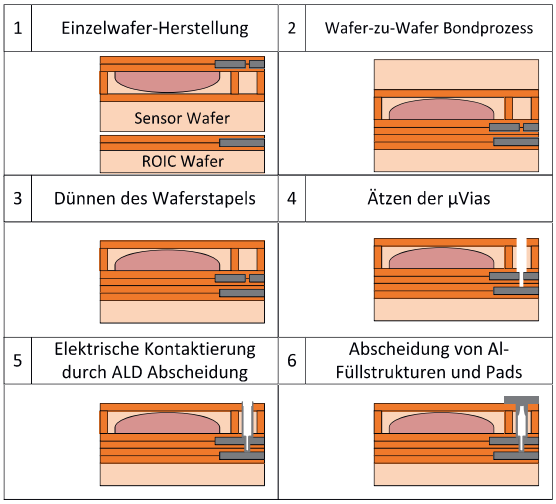


Abb. 1: Schematischer Ablauf des direkten Wafer-zu-Wafer-Bondprozesses mit nachfolgender elektrischer Durchkontaktierung

3. Optimierung des Bondprozesses

Als Wafer-zu-Wafer Fügeverfahren wurde in dieser Arbeit das hydrophile Siliziumdirektbonden verwendet. In diesem Verfahren basiert die Herstellung der mechanischen Verbindung auf der Ausbildung von kovalenten Siloxan (Si-O-Si)-Bindungen im Bereich des Bondinterfaces. Dazu werden zunächst die Waferoberflächen in einem Sauerstoffplasma hydrophilisiert und aktiviert um anschließend in einer H_2O-Atmosphäre Silanol (Si-OH)-Gruppen zu erzeugen. Die Wafer werden aufeinander ausgerichtet und mit einer Kraft von mehreren Kilonewton in Kontakt gebracht um die Bildung von Wasserstoffbrücken zu ermöglichen. Abschließend wird das Waferpaar auf eine Temperatur von über 250 °C erhitzt, wodurch Wassermoleküle entlang der Grenzflächen als auch in das Oxid diffundieren. Diese Abdiffusion führt zu einer Annäherung der gegenüberliegenden Silanol-Gruppen, die unter Abspaltung eines Wassermoleküls eine kovalente Siloxan-Verbindung bilden [2]. Dieses Bondverfahren stellt hohe Anforderungen an die Oberflächenrauheit und Topologie der Wafer.
Die SPAD- und ROIC-Wafer kommen jeweils aus einem standardisieren CMOS-Prozess und können ohne Prozessanpassung nicht über das Oxid-zu-Oxid-Bondverfahren verbunden werden. Dazu wurde nach dem CMOS-Prozess eine ganzflächige Oxidschicht auf die Kontaktoberflächen der jeweiligen Wafer abgeschieden. Ein Unterschied zwischen den beiden CMOS-Wafern ist das verwendete Substrat. Während für den ROIC standard Bulk-Wafer eingesetzt werden, werden Silicon-on-Insulator (SOI) -Wafer für die Herstellung der SPADs verwendet. Das vergrabene Oxid (buried oxide, BOX) dient für die Nachbearbeitung des Waferstapels als Ätzstopp beim Dünnen des Wafers (siehe Abb.1, Schritt 3) und induziert bereits bei der Herstellung des SPAD-Wafers in der CMOS-Linie eine deutliche Waferdurchbiegung.

Zur Herstellung einer qualitativ hochwertigen Verbindung wurde der Prozessablauf hinsichtlich folgender Oberflächeneigenschaften optimiert:

- Mikrorauheit
- Topologie
- Biegung
- Partikelkontamination.

Um eine hohe Bondfestigkeit zu erhalten sollte die Oberflächenrauheit der zu bondenden Wafer möglichst gering sein [3]. Als Richtwert für das direkte Oxid-Bonden hat sich eine quadratische Rauheit (eng. rms-roughness, Wurzel des Mittelquadrates) von unter 0,5 nm etabliert [4]. Mit einem mehrstufigen chemisch-

mechanischen Polierverfahren nach der Oxidabscheidung konnten folgende Kennwerte (gemessen mit einem Rasterkraftmikroskop, AFM) erreicht werden (siehe Tabelle 1).

Tab. 1: Kennwerte der Oberflächenrauheitsmessung der SPAD- und ROIC-Wafer mittels AFM

Wafer	Mittlere Rauheit R_a [nm]	Quadratische Rauheit R_q [nm]
SPAD	0,22	0,28
ROIC	0,24	0,30

Neben einer erhöhten Oberflächenrauheit können auch ausgeprägte Topologien die Waferkontaktfläche beim Bonden verringern und dadurch die Bondfestigkeit schwächen. Zur Reduktion dieser Topologien, wurde die Passivierungsöffnung erst bei der Durchkontaktierung des Waferstapels realisiert. Zusätzlich wurden kritische Strukturen wie z.B. Metallpads homogen im Layout verteilt, um durch diese Dezentralisierung den Einfluss der Topologie auf den Waferbond zu reduzieren. Auf diese Weise wurden Topologie-Unterschiede größer als 20 nm vermieden.
Selbst kleinste Partikel mit einem Durchmesser von 1 µm können Bonddefekte mit einem Durchmesser von mehreren mm erzeugen [5]. Um Partikelkontaminationen auf den Wafer zu verhindern, wird eine hochfrequente Reinigung (Megasonic-Cleaning) mittels DI-Wasser vorgenommen. Die hochfrequente Reinigung kommt ohne zusätzliche Chemikalien aus und kann einfach in den Prozessablauf integriert werden.
Eine starke Waferbiegung, sowohl des Sensor- als auch des Schaltungswafers, kann zu einer Dejustage des Waferbonds führen und zu einer Delamination des Waferstapels in der Nachbearbeitung. Daher wurde bei der Herstellung des biegungskritischen Sensorwafers (SOI-Wafer) eine Stresskompensationsschicht eingeführt, die eine Reduktion der Waferbiegung um bis zu 75% ermöglicht (siehe Tabelle 2).
Dabei wurden zwei Verfahren untersucht:

1) Es wurde eine Nitridschicht mit Zugspannung auf der Wafervorderseite im Frontend Prozess abgeschieden.
2) Im Backend Prozess wurde Oxid-Schicht mit Druckspannung auf der Waferrückseite abgeschieden.

Tab. 2: Stresskompensationsvarianten und gemessene Waferdurchbiegung des Sensorwafers

Wafer	Kompensations-Varianten	Biegung Wafermitte [µm]
SPAD	Ausgangszustand	-100
	Nitridschicht (Frontend-Prozess)	-57
	Rückseiten-Oxid (Backend-Prozess)	-25

Letzteres Verfahren bietet neben einer deutlichen Reduktion der Waferbiegung den Vorteil, dass der CMOS Frontend-Prozess nicht geändert werden muss und zusätzlich das Oxid in der Nachbearbeitung des Waferstapels beim Dünnschleifen wieder entfernt werden kann. Zudem muss keine zusätzlich Schicht bei der Durchkontaktierungs-Ätzung entfernt werden. Folglich wurde für die weitere Prozessentwicklung die Abscheidung von Oxid auf die Rückseite des Sensorwafers verwendet.
Zur Herstellung der elektrischen Kontaktierung müssen beide Wafer präzise aufeinander ausgerichtet werden. Die Waferbiegung hat hierbei einen erheblichen Einfluss auf die Ausrichtungsgenauigkeit [6]. Mit dem entwickelten Bondprozess kann eine Platziergenauigkeit von unter 7 µm erreicht werden.
Während des ganzflächigen Dünnschleifens und der anschließenden Plasmaätzung wird zusätzlicher Stress in den Wafer induziert. Um diesem Stress standzuhalten, muss das Bondinterface über eine hohe Bondfestigkeit verfügen. Schertests an unstrukturierten Wafern, die mit dem optimierten Oxid-Bondprozess in Verbund gebracht wurden, zeigen Scherfestigkeiten von über 4 kg/mm^2.

4. Micro-Vias

Bei der Darstellung des Bondprozesses in Kapitel 2 wurde bereits erwähnt, dass die elektrische Kontaktierung der beiden Wafer (SPAD- und ROIC-Wafer) durch eine Vorderseiten-Silizium-Durchkontaktierungen (TSV: Through Silicon Vias) realisiert wird. Eine Besonderheit des Prozessablaufs ist, dass die Ätzung der Strukturen erst in der Nachbearbeitung des Waferstapels erfolgt. Diese Durchkontaktierungen – im Folgenden als µVias bezeichnet - sind geätzte Strukturen mit einigen Mikrometern im Durchmesser, die auf den jeweiligen Metall-Pads von SPAD- und ROIC-Wafer stoppen. In Abbildung 2 ist eine schematische Zeichnung der µVia-Struktur darge-

stellt. Abbildung 3 zeigt die Realisierung dieser Struktur auf einem gebondeten Wafer im REM-Querschnitt. Zur elektrischen Kontaktierung wird im Anschluss ein leitfähiger mittels ALD (Atomic Layer Deposition)optimierter Materialstapel in die µVias bzw. auf den Waferstapel abgeschieden

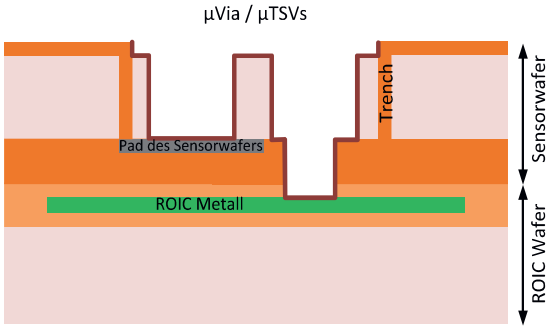


Abb. 2: Schematische Darstellung eines µVias mit zwei Einzelätzungen bis auf die Metalle von SPAD – und ROIC-Wafer, sowie der elektrischen Verbindung über eine leitfähige Schicht.

Wichtig bei der Füllung der µVias ist die Erzeugung einer durchgehenden elektrisch leitenden Schicht. Ein reiner Sputterprozess würde die 6x2,4 µm² großen Strukturen nicht konform füllen.
Daher wurde zur Optimierung der Kontaktierung im Anschluss an den Sputterprozess eine leitfähige homogene ALD-Schicht abgeschieden.
Durch die Füllung der beiden Einzelätzungen im µVias entsteht eine elektrische Brückenverbindung zwischen den Wafern.

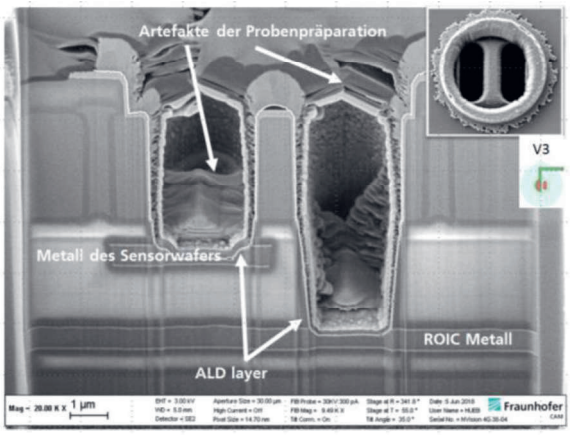


Abb. 3: REM-Querschnitt eines µVias (Version V3). Inserts zeigen die REM sowie die schematische Aufsicht.

Die Qualität des elektrischen Kontakts ist dabei maßgebliche von Größe, Form und Anzahl der µVias abhängig und kann an beliebige Anforderungen angepasst werden. In diesem Beitrag werden vier untersuchte Versionen von µVias vorgestellt (siehe Abbildung 4). Prinzipiell sind beliebig viele weitere Formgebungen möglich.

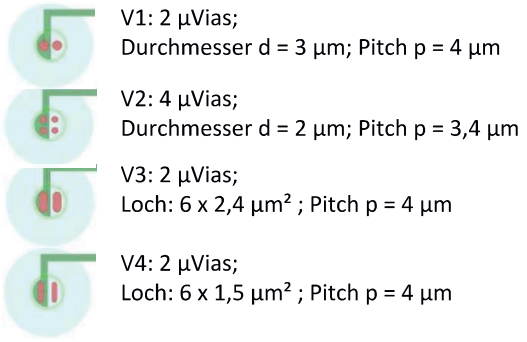
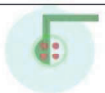


Abb. 4: Schematische Aufsicht der vier vorgestellten µVia Varianten (links), sowie die zugehörige Beschreibung (rechts).

5. Charakterisierung der elektrische Verbindung

Die Hauptmerkmale bei der Charakterisierung der elektrischen Verbindung von zwei Wafern ist der Widerstand der µVias, sowie dessen Zuverlässigkeit. Zur Untersuchung dieser Parameter wurden zunächst 100 µVias in Reihe geschaltet und deren Kettenwiderstand direkt nach der Wafer-Prozessierung gemessen.
Zur Zuverlässigkeitsuntersuchung (ZUV) wurden jeweils Viertelwafer (40 Chips) mit bis zu 500 Temperaturzyklen gestresst und dabei Komplettausfälle sowie die prozentuale Änderung des µVia-Widerstands fortlaufend analysiert. Die Kettenwiderstände für die vier in Abbildung 4 vorgestellten µVia-Varianten sind in Tabelle 3 aufgeführt.

Tab. 3: µVia Varianten mit ihren zugehörigen Widerständen / Via

µVia Version	Layout	Widerstand Ω/Via
V1		43
V2		71
V3		14
V4		25

Die Widerstandswerte pro µVia liegen dabei zwischen 14 und 71 Ω/Via und stehen wie erwartet in Anhängigkeit zur Größe, Form und Anzahl der µVias.
Die Version 3, bei der die Verbindung über zwei Langlöcher realisiert ist, zeigt dabei den geringsten Widerstand (siehe Querschnitt und Aufsicht in Abbildung 3). Das Ergebnis entspricht den Erwartungen, da bei dieser Variante der größte Füllfaktor in Bezug zur Padfläche besteht.
Die an die AECQ100 grade 1 und JESD22 A104E Normen angelehnte Zuverlässigkeitsuntersuchung (500 Temperaturzyklen von -65 bis 150 °C, 2 Zyklen pro Stunde) der Verbindung unter Temperatureinwirkung ist exemplarisch für die µVia-Variante-Version 3 in Abbildung 4 dargestellt.

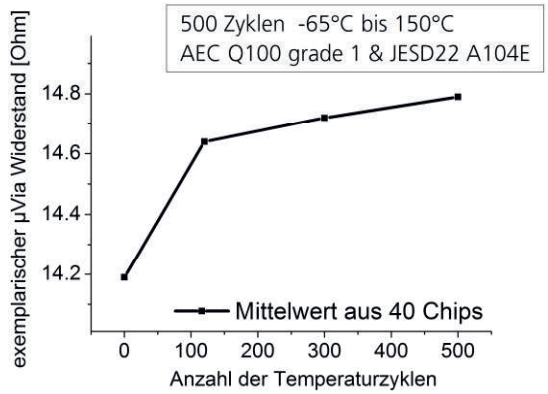


Abb. 4: Messergebnisse der Zuverlässigkeitsuntersuchung einer µVia Kette: Widerstand eines µVias über die Anzahl der Temperaturzyklen gemäß dem AECQ100 Grade 1 Standard.

Defekte wurde durch die Belastung nicht festgestellt und die gemittelte Widerstandsänderung ist mit kleiner 5% ein positives Ergebnis im Hinblick auf eine Prozessqualifizierung.

6. Funktionalitätsprüfung der SPADs

Neben der Charakterisierung der elektrischen Verbindung zwischen den Wafern, ist für einen hochempfindlichen Sensor das Pixel – in diesem Fall die maximal lichtsensitive Einzelphoton-Lawinenphotodiode (SPAD) – ein wesentlicher Bestandteil.
Wie oben gezeigt ist der vorgestellte Prozess an Hand der Ergebnisse der elektrischen Verbindung bereits als erfolgreich und robust einzustufen. Da dieser jedoch hauptsächlich für die 3D-Integration von SPADs und Ausleseelektronik verwendet wird, ist es essentiell auch die grundlegenden Eigenschaften der SPADs im Waferverbund auf Plausibilität zu prüfen um Beschädigungen der SPADs im Verlauf des Prozesses auszuschließen. Als Referenz für die Plausibilitätsprüfung wurden SPAD-Daten aus einem leicht abgewandelten Prozess verwendet werden [7].
Abbildung 5 zeigt exemplarisch den Aufbau eines solchen SPADs in Aufsicht und im Querschnitt. Dabei ist das SPAD von einer Grabenisolation umgeben. Das isolierte Substrat wird vollständig durch eine niedrige Dotierung zu einer N-Wanne umgeformt. Ein sogenannter Guard Ring (GR) verhindert Lawinendurchbrüche an der Oberfläche zwischen Anode (p+) und Kathode (n+), sodass die Lawinendurchbrüche nur zwischen dem p+ und einer speziellen SPAD-Implantation „n-SPAD“ stattfindet. Dieser Bereich ist der aktive Bereich der SPAD.

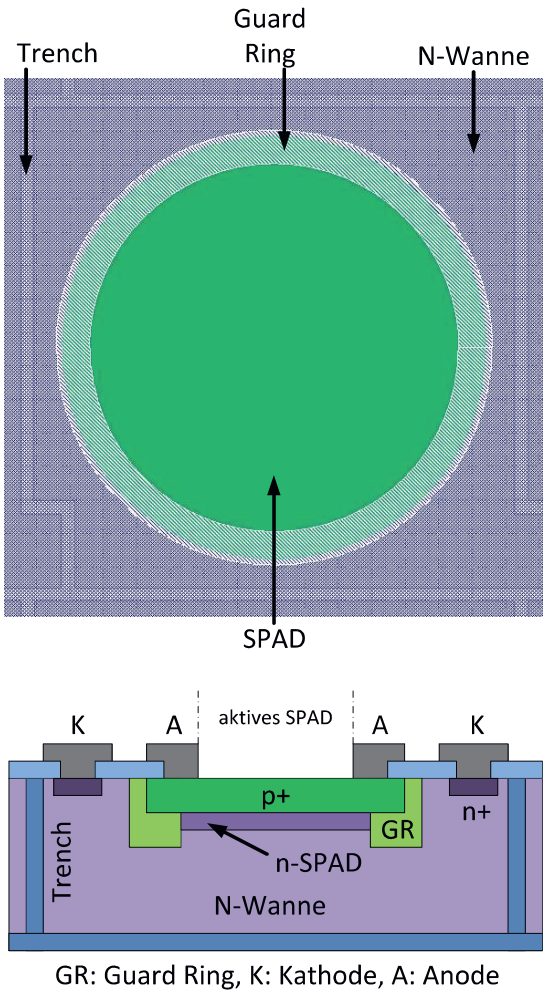


Abb. 5: Schematischer Aufbau eines SPADs: Aufsicht (oben), Querschnitt (unten).

Abbildung 6 zeigt die Strom-Spannungskennlinie des SPADs. Die Messung wurde dabei zunächst im Dunkeln und anschließend bei Licht durchgeführt. Dabei konnten die folgenden Messergebnisse gesammelt werden:

1) Die SPADs brechen jeweils bei einer Spannung von ca. 24 V durch, was durch einen signifikanten Anstieg des Stromes gekennzeichnet ist. Der Arbeitspunkt der SPAD sollte deshalb oberhalb von 24 V liegen. Die Durchbruchspannung der etablierten SPAD-Technologie liegt bei ca. 26 V [7].
2) Der Einfluss des Lichts ist deutlich am Strom unterhalb des Diodendurchbruchs zu sehen, der bei Beleuchtung um mehr als 4 Größenordnungen ansteigt. Dies demonstriert bereits die generelle Lichtempfindlichkeit des Bauelements. Eine quantitative Aussage über die Lichtempfindlichkeit ist nur bei definierter Beleuchtung möglich.

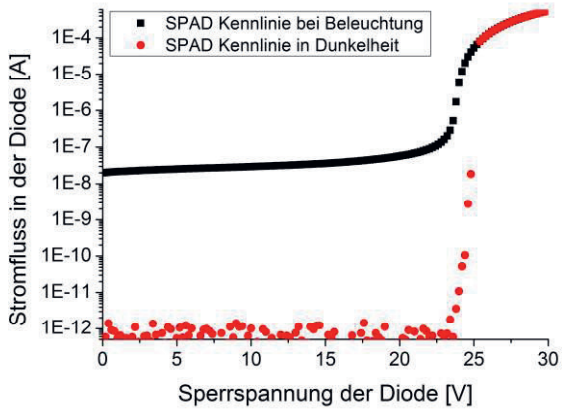


Abb. 6: Messergebnisse der I-V-Charakterisierung eines SPADs bei Raumtemperatur in Dunkelheit und bei Beleuchtung.

Eine weitere wichtige Größe einer SPAD ist die Dunkelzählrate. Diese wird im Folgenden quantitativ untersucht. Dazu wird das Bauelement im Geiger-Modus mit einer passiven externen Schaltung (Widerstand von 200 kΩ) betrieben. Bei der Messung befindet sich die Diode in Dunkelheit. Die Betriebsspannung liegt dabei jenseits der Durchbruchspannung und setzt sich aus der Durchbruchsspannung und einem Volt Überspannung zusammen. Die im Geiger-Modus entstehenden Spannungs-Pulse über der Diode werden gezählt und ergeben die Dunkelzählrate. Diese liegt für die verwendeten Betriebsbedingungen bei den untersuchten SPADs bei Raumtemperatur im Mittel bei 310 Hz und ist damit weltweit gesehen ein SPAD mit der niedrigsten Dunkelzählrate [8].
Die optische Sensitivität der Diode kann über die Quanteneffizienz (Anzahl generierter Elektronen-Loch-Paare / Anzahl der einfallenden Photonen) angegeben werden. Abbildung 7 zeigt die Quanteneffizienz für die Wellenlängen $\lambda = 440$ nm, $\lambda = 530$ nm, $\lambda = 625$ nm und $\lambda = 850$ nm. Die spektrale Abhängigkeit der Quanteneffizienz mit einem Maximum im roten Wellenlängenbereich entspricht dabei den typischen Charakterisierungskennwerten einer solchen Diode [7].

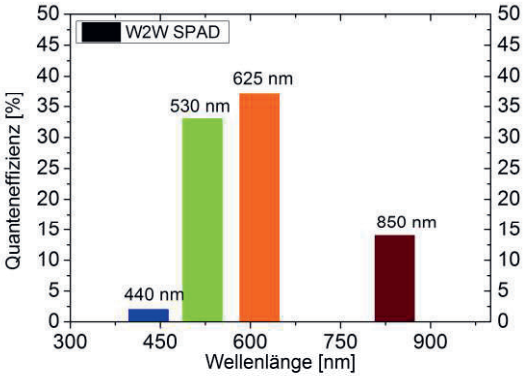


Abb. 7: Ergebnisse der Quanteneffizienz-Messung für die Wellenlängen $\lambda = 440$ nm, $\lambda = 530$ nm, $\lambda = 625$ nm und $\lambda = 850$ nm.

Die aufgenommenen Messwerte befinden sich im Größenbereich der Vergleichskennwerte und zeigen, dass die SPADs nach der Prozessierung kontaktiert und funktionsfähig sind.

7. Zusammenfassung & Ausblick

Im Rahmen der Entwicklung eines alternativen Aufbau- und Verbindungsprozesses wurde ein Wafer-zu-Wafer-Prozess für die Verbindung und elektrische Kontaktierung von Sensoren und Schaltungen entwickelt.
Mit den vorgestellten Prozessen konnten die Oberflächeneigenschaften der SPAD- und ROIC-Wafer in Hinblick auf einen qualitativ hochwertigen Wafer-zu-Wafer-Bond bereits in der hauseigenen CMOS-Linie optimiert werden.
Das anschließende Oxid-zu-Oxide-Bonden kann bei CMOS-kompatiblen Temperaturen durchgeführt und auf unter 7µm präzise genau ausgerichtet werden.
Zur elektrischen Kontaktierung des Waferstapels wurde ein spezielles µVia-Verfahren zur Vorderseiten-Silizium-Durchkontaktierungen mit anschließender Abscheidung eines ALD-optimierten Schichtstapels entwickelt. Mit diesem Verfahren können zuverlässige Kontaktierungen zwischen Sensor und Schaltung erzeugt werden, die über 500 Temperaturzyklen

von -65 bis 150 °C ohne Ausfall mit einer Widerstandsänderung kleiner 5% standhalten.
Zusätzlich wurde die Funktionalität der SPADs im Anschluss an den Wafer-zu-Wafer-Prozess untersucht. Sowohl die Quanteneffizienz als auch die Dunkelzählrate und der Arbeitspunkt der SPADs entsprechen den Vergleichswerten der bisherig etablierten Technologie und zeigen, dass der entwickelte Wafer-zu-Wafer-Prozess neben einer verlässlichen Kontaktierung zu keiner Funktionalitätsbeeinträchtigung der SPADs führt.
Der Prozess kann als kostengünstigere und durchsatzstärkere Alternative, zum am IMS etablierten, Chip-zu-Wafer-Verfahren eingesetzt werden.

Aktuell werden am Fraunhofer IMS weitere Entwicklungen zur Optimierung der SPAD-Performance im Wafer-zu-Wafer-Prozess vorgenommen. Hierzu zählen Verbesserungen des Füllfaktors und eine Erhöhung der Empfindlichkeit im nahinfraroten Bereich.

Mit dem Chip-zu-Wafer und Wafer-zu-Wafer-Verfahren stehen zwei leistungsfähige Prozesse zur Realisierung von 3D-Integrationen am Fraunhofer IMS zur Verfügung

Literaturnachweis

[1] M. Beer, O.M. Schrey, J.F. Haase, J.Ruskowski, W. Brockherde, et al., ”SPAD-based flash LiDAR sensor with high ambient light rejection for automotive applications”, *Proc. of SPIE* 10540 (2018); doi: 10.1117/12.2286879

[2] Q-Y.Tong, U. Gösele, “Wafer Bonding and Layer Splitting for Microsystems”, *Advanced Materials* 11 (17), 1409-1425 (1999); doi 10.1002/(SICI)1521-4095(199912)11:17<1409::AID-ADMA1409>3.0.CO;2-W

[3] C.Gui, M.Elwenspoek, N.Tas, J.Gardeniers, “The effect of surface roughness on direct wafer bonding”, *Journal of Applied Physics* 85, 7448 (1999);doi: https://doi.org/10.1063/1.369377

[4] L.Nie, T.Shi, Z.Tang, G.Liao, “Pressure Aided Low Temperature Direct Bonding of Silicon Wafers with High Surface Roughness”, *1st IEE International Conference on Nano/Micro Engineered and Molecular Systems* (2006); doi: 10.1109/NEMS.2006.334753

[5] U. Gösele, Q-Y.Tong, “SEMICONDUCTOR WAFER BONDING”, *Annual Review of Materials Science* 28, 215-241 (1998); doi: https://doi.org/10.1146/annurev.matsci.28.1.215

[6] S.E. Steen, D. LaTulipe, A.W. Topol, D.J. Frank, K. Belote, D. Posillico, “Overlay as the key to drive wafer scale 3D integration”, *Microelectronic Engineering* 84(5-8), 1412-1415 (2007); doi: https://doi.org/10.1016/j.mee.2007.01.231

[7] D.Bronzi, F.Villa, S. Bellisai, B. Markovic, S. Tisa, A. Tosi, F. Zappa, S. Weyers, D. Durini, W. Brockherde, U. Paschen, “Low-noise and large-area CMOS SPADs with Timing Response free from Slow Tails”, *Proceedings of the European Solid-State Device Research Conference (ESSDERC)*; doi: 10.1109/ESSDERC.2012.6343375

[8] D. Bronzi, F. Villa, S. Tisa, A. Tosi, F. Zappa, D. Durini, S. Weyers , W. Brockherde, “100 000 Frames/s 64 × 32 Single-Photon Detector Array for 2-D Imaging and 3-D Ranging”, *IEEE Journal of Selected Topics in Quantum Electronics* 20, 354–363 (2014); doi: 10.1109/JSTQE.2014.2341562

DOI 10.5162/sensoren2019/3.2.4

Programmable reference parameters for resonance locking in electro-thermal piezoresistive cantilever sensors

Andi Setiono[1,2], Michael Fahrbach[1], Jiushuai Xu[1], Wilson Ombati Nyang'au[1,3], Maik Bertke[1], Hutomo Suryo Wasisto[1], Erwin Peiner[1]
[1]Institute of Semiconductor Technology (IHT) and Laboratory for Emerging Nanometrology (LENA), Technische Universität Braunschweig, Hans-Sommer-Straße 66, D-38106 Braunschweig, Germany
[2]Research Centre for Physics, Indonesia Institute of Sciences (LIPI), Kawasan Puspitek Serpong, 15314 Tangerang Selatan, Indonesia
[3]Department of Metrology, Kenya Bureau of Standards (KEBS), Nairobi 00200, Kenya
a.setiono@tu-braunschweig.de, m.fahrbach@tu-braunschweig.de, jiushuai.xu@tu-braunschweig.de, wilombat@tu-braunschweig.de, m.bertke@tu-braunschweig.de, h.wasisto@tu-braunschweig.de, e.peiner@tu-braunschweig.de

Abstract

Resonant-frequency tracking is essential when measuring analyte targets using resonance-based sensors in real-time. We developed a self–sensing and –actuating silicon cantilever resonator that can be functionalized as an environmental sensor for analytes such as particles, humidity, and gases. However, direct thermal parasitic causes the sensor to generate a non-ideal output signal which hinders implementing a real-time phase-locked loop (PLL) based setup. To alleviate these problems, software de-embedding solutions to create an optimized signal, especially a monotonic phase response, are investigated. The de-embedding approach is based on subtracting a selectable reference, which mimics the parasitic behavior, from the measured data. All reference calculations are performed real-time on a microcontroller, which allows for on-line PLL based frequency tracking. Experiment results show a good response in resonant-frequency locking and tracking under exposure to cigarette smoke.

Keywords: De-embedding, programmable reference, electro-thermal piezoresistive cantilever sensor, monotonic phase response, phase-locked loop.

Introduction

Environmental health has been a concern of all parties globally in the past decades, which further has led to the development of environmental monitoring technology. Sensor technology is an important part of environmental monitoring, i.e., to perform one or more measurements that are necessary to assess the status of an environment. Measurements of particles, toxic gases, and volatile organic compounds (VOCs) are a part of the monitoring systems to protect persons against environmental deterioration. Nowadays, environmental monitoring emerges with the development of microsensors and microactuators in the field of micro-electro-mechanical system (MEMS) technologies. There are many environmental sensors available already, such as particle-, and gas-sensors. Most frequently, they are based on optical scattering and dispersion [1, 2]. Compared to optical-based sensor, MEMS resonance-based sensors have advantages in size, energy consumption, and high-volume fabrication. Electro-thermal piezoresistive cantilever sensors (EtPCS) are MEMS-based sensor-actuator systems that can reliably measure environmental conditions at high resolution [3–5]. Electro-thermal based actuation system has advantages in low fabrication complexity and high responsivity [6]. Furthermore, utilizing silicon piezoresistors as a sensing part results in high signal-to-noise ratio, high miniaturization capability, and compatibility with available electronics. Nevertheless, parasitic factors inside of the EtPCS are a drawback that has to be eliminated. Thermal parasitics can generate a Fano resonance (asymmetric shape) and a reversing phase response at the sensor output. A reversing phase response is an inhibitive factor in PLL-based real-time measurements. Therefore, the aim of this work is to optimize the phase response of EtPCS in order to work properly in a real-time measurement setup, i.e., by de-embedding the signal of the sensor using an approximation of the parasitic signal,

which is subtracted from the measured signal. This approximated parasitic is also called reference.

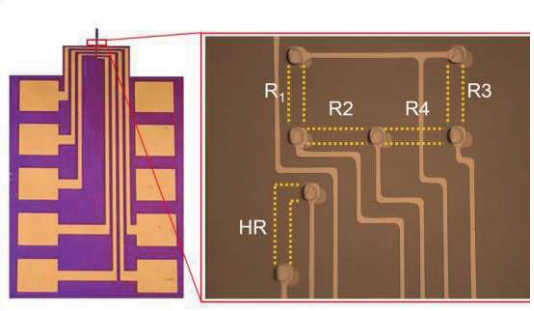


Fig. 1: Photograph of a MEMS electro-thermal cantilever sensor and optical-micrograph magnification of the cantilever top surface showing the heating resistor (HR) and the U-shape Wheatstone bridge (WB).

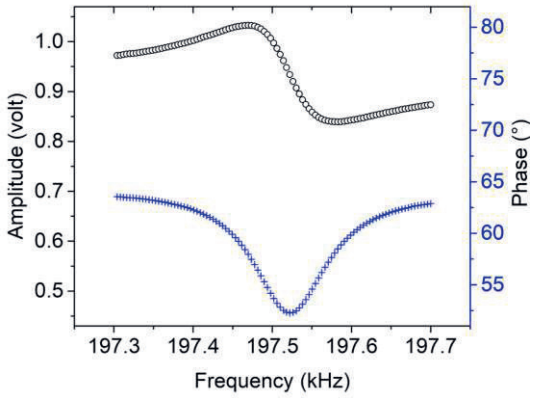


Fig. 2: Measured amplitude (black circle) and phase (blue plus signs) response of a piezoresistive cantilever excited by on-chip electro-thermal actuation.

Output characteristics of EtPCS

We developed an electrothermal piezoresistive, i.e., a self-actuating and self-sensing, resonant cantilever (Fig. 1) for detecting various analyte targets. The main electrical components are a *p*-doped piezoresistive Wheatstone bridge (for sensing) and a *p*-doped heating resistor (for actuation).
The heating resistor (HR) creates a mechanical actuation by converting thermal power through Joule heating [7]. This thermal power is generated by supplying the HR with an actuation signal consisting of a sine-wave (V_{AC} sin(ωt)) superimposed onto a constant voltage (V_{DC}). The dynamic amplitude V_{AC} leads to two frequencies being excited: ω and 2ω. The constant component V_{DC} creates constant stress on the cantilever beam. If V_{DC} is much higher than V_{AC}, the cantilever beam is driven into vibration at ω, as the frequency component of 2ω is negligible. Here, the main component of the output signal has the same frequency as the excitation signal. The generated excitation force is then detected by the WB and converted into an electrical signal.
However, direct thermal coupling exposed from the HR to the WB, creates a parasitic effect on the sensor output that induces an asymmetric resonance line-shape (Fano resonance) and phase deformation as shown in Fig. 2. The reversing phase response presents an ambiguity and is detrimental for real-time operation using a phase-locked loop (PLL) [8]. Therefore, there is a need for optimizing the amplitude and the phase response. This can be done by creating an approximation of the parasitic, also called reference, and subtracting it from the measured values. This way the ideal signal of the sensor is de-embedded from the measured signal.

Reference signal subtraction

Signal subtraction is intended to eliminate the parasitic parameters inside the EtPCS. In [8], Bertke et al. illustrated how a Fano resonance results from a superposition of a constant amplitude signal with Lorentzian-line-shape signal. Therefore, subtracting a constant-amplitude signal from the output of the EtPCS is expected to reveal a Lorentzian amplitude shape and optimize the sensor phase response.
In previous work [9, 10], a system based on reference-signal subtraction for eliminating the parasitic effect and optimizing the EtPCS has been realized. The reference signal was created using passive electronic components to build an RCL (Resistor, Capacitor, Inductor) filter. Through this approach, symmetrical amplitude shapes and monotonic phase responses were demonstrated. However, as passive references require much effort to be tuned, a programmable reference is more desirable to be used. Thus, the current study seeks to implement a programmable reference as schematically shown in Fig. 3. The proposed instrument comprises three main parts, i.e., a cantilever sensor (EtPCS), an electronic measurement system, and a display and acquisition system. The electronic measurement system has bi-directional capability for communicating and commanding between the sensor and the display and acquisition system. It generates a sinusoidal signal through a wave generator component, which subsequently actuates the cantilever beam. It also transmits the measured data and

de-embedding results to the acquisition system.

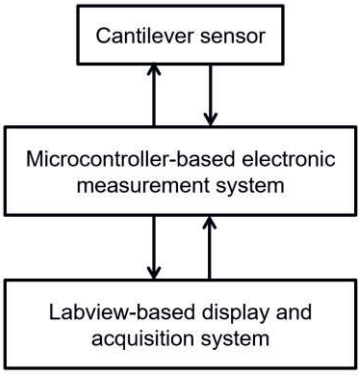


Fig. 3: Block diagram of the de-embedding microcontroller system that has bi-directional communication for actuating and sensing of the cantilever sensor and further display and acquisition using a LabVIEW-based system.

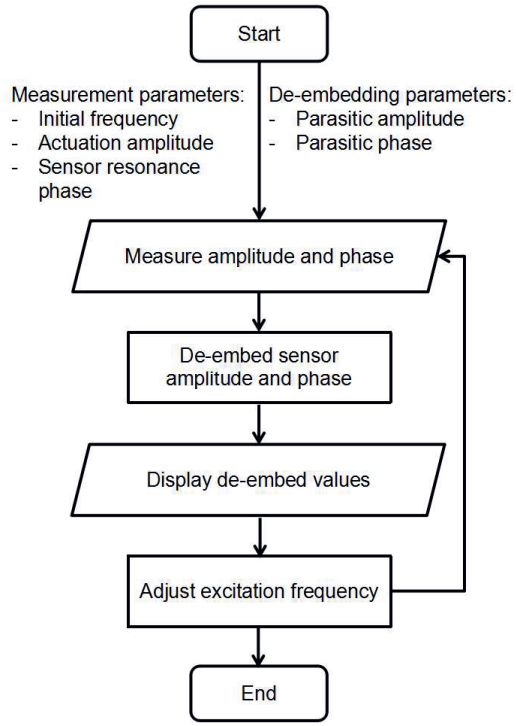


Fig. 4: Flowchart of de-embedding sensor signals using reference parameters which are transmitted to a microcontroller.

According to the flowchart delineated in Fig. 4, a microcontroller acts as a central processor for reading, processing and transmitting data. It measures the cantilever output signal amplitude, and simultaneously calculates the phase difference between the actuation and sensing signals. Afterwards, these sensor responses are plotted on a LabVIEW-based display system. On the other hand, reference parameters (i.e., amplitude and phase) are constant values, which are sent to the microcontroller by the LabVIEW system. Reference parameters are determined by considering the baselines of both amplitude and phase responses in the range of the working frequency (near resonance) [9, 11], as shown in Figs. 5a and 5b. These values are then used to de-embed the signal of the sensor from the measured signal. The corresponding formula is described in more detail in [11]. The resulting cantilever response shows optimized parameters (i.e., a nearly symmetrical amplitude shape and a monotonic phase response), as depicted in Fig. 5c. The resultant monotonic phase response can now be used in a PLL-based setup. All measurement and calculation results (i.e., amplitude and phase) are latterly captured on the display and acquisition system. In addition, the occurrence of the double peak might be avoided by implementing a lock-in amplifier principle that calculate the phase by using a DC component as a result of signal mixing.

PLL-based frequency locking

A monotonic phase response is needed for stable operation of a PLL-based system. With a monotonic phase response, the PLL can lock the resonance phase (φ_0) at a certain value of frequency, which is subsequently interpreted as resonant frequency (f_r). A block diagram of PLL-based frequency locking is shown in Fig. 6, which consists of a phase detector, an I-controller, and a feedback system. These three main elements are represented by the microcontroller in software. The phase detector measures the phase between the actuation signal and the output of the sensor, which is achieved by measuring the difference in time between the zero crossings of both signals. This time difference is proportional to the phase. The measured phase value is then used to calculate the de-embedded phase of the sensor. Afterwards, the phase is used by the I-controller for tracking the resonance phase. The I-controller is formulated by:

$$f_r(t) = f_0 + K_i \int_0^t \Delta\varphi_{set}(\tau) - \Delta\varphi_{meas}(\tau) d\tau \qquad (1)$$

where f_0, K_i, φ_{set}, φ_{meas}, t, τ are initial frequency, integral coefficient, set-point phase, measured phase, time, and variable of integration, respectively.

The flowchart shown in Fig. 7 illustrates how the PLL operates to lock the resonant frequency. In the beginning, a desired phase is

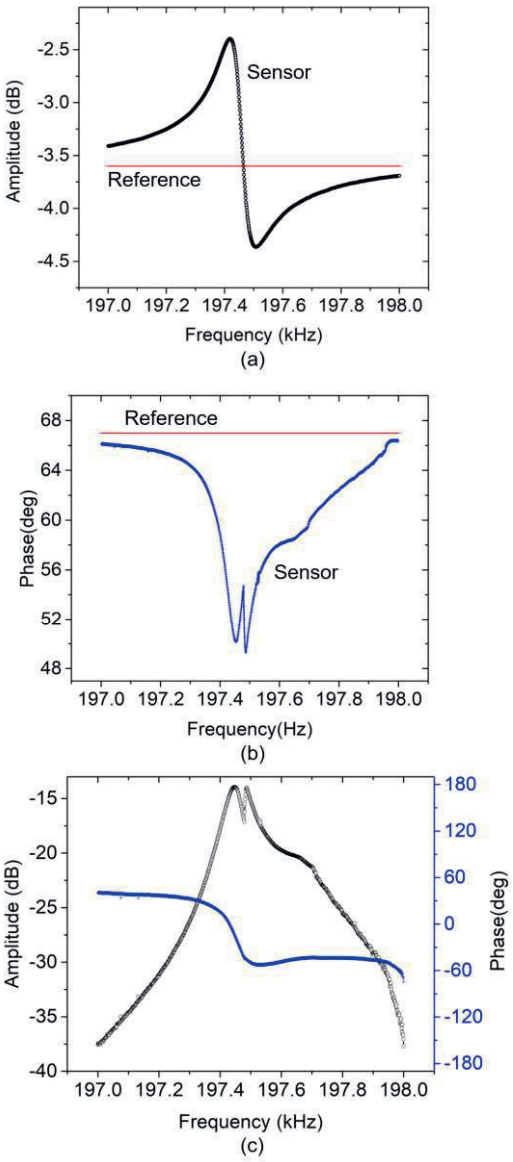


Fig. 5: Measured sensor signals. (a) Asymmetric amplitude response. (b) and reversed phase response. The occurrence of the double peak is expected due to losing time in phase calculation during a short resonance transition. (c) Amplitude (black circle line) and phase (blue plus line) responses with programmable references in place. De-embedding results in symmetrical amplitude shape and nearly monotonic phase response.

set as a set-point, which is related to the resonance phase of the sensor. Hereinafter, the microcontroller measures the phase between the sensor output and actuation signal and uses it in the de-embedding process. The resulting phase is then employed

by the I-controller for adjusting the excitation frequency as described by Eq. (1).
Figure 8 shows that the de-embedded phase response can be used to track the resonant frequency. Increasing the integral coefficient (K_i) of the I-controller leads to a decrease of response time, but on the other hand, an overshoot is generated. Therefore, an optimum value of K_i is needed for obtaining the locked condition as fast and precise as possible.
Furthermore, the EtPCS is investigated under slight exposure to cigarette smoke in a chamber as illustrated in Fig 9a. Frequency tracking uses the de-embedded sensor signal using the PLL-configuration shows good results as depicted in Fig 9b. In section A, the baseline of the resonant frequency was measured, showing a decreasing trend. The suspected reason is an increase in temperature inside the chamber. The temperature coefficient of Young's modulus and the thermal expansion coefficient are the intrinsic parameters of silicon that contribute most to its fast and reversible temperature response. As a confirmation, stability tests were conducted by using an MFLI Lock-in Amplifier (Zurich Instruments) under temperature and humidity conditions of 23.15 ± 0.17° and 38.5 ± 0.7%. The result shows a stable resonant frequency of 197405.05 ± 0.01 Hz and 197405.16 ± 0.06 Hz during 5 minutes and 1 hour of measurement duration, respectively.

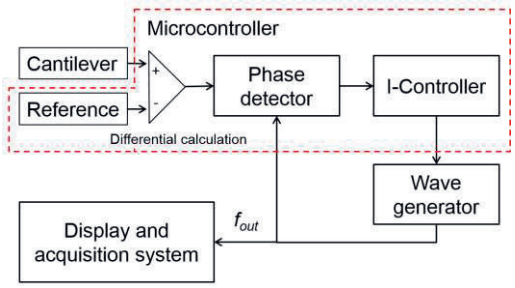


Fig. 6: A block diagram of the programmable reference subtraction in a PLL configuration.

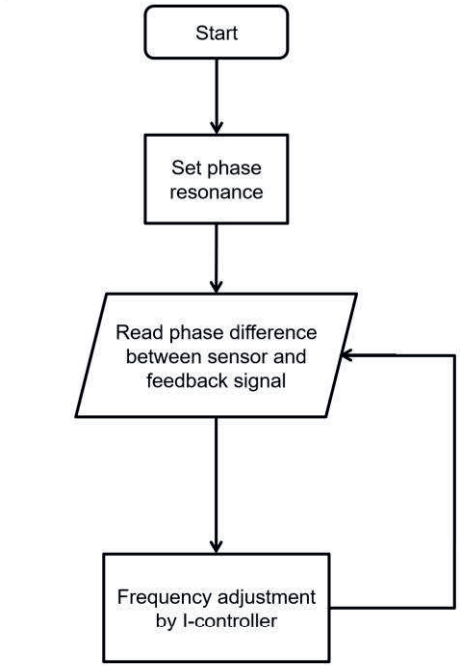


Fig. 7: Flowchart of the digital I-controller for managing a frequency actuation based on a phase error.

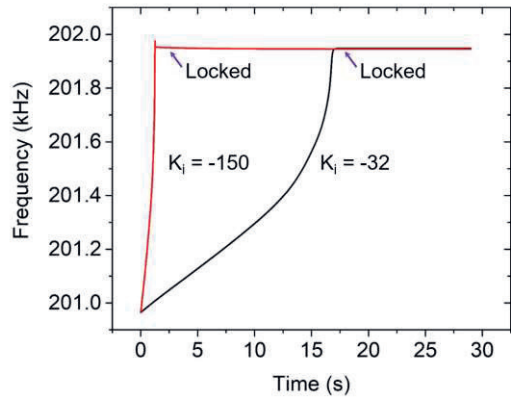


Fig. 8: Time-dependent resonance locking under a PLL set-up and an I-controller with K_i = -150 Hz/°/s (red) and K_i = -32 Hz/°/s (black).

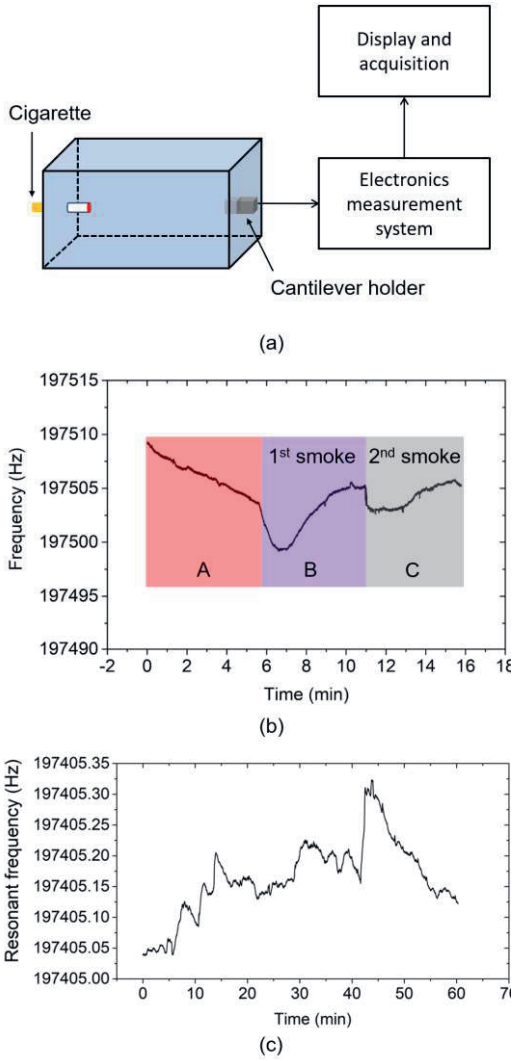


Fig. 9: (a). Measurement setup for real-time measurements of an EtPCS exposed to cigarette smoke. Time-dependent resonant frequency using an electro-thermal piezoresistive cantilever sensor (b) under temperature changing and smoke exposing and (c) stability tests under relatively constant of temperature and humidity.

Furthermore, in the beginning of sections B and C respectively, a burning cigarette was introduced into the chamber. The frequency drops in the beginning of each smoke exposure with the response times of 41.8 s and 2.8 s for the 1st smoke and the 2nd smoke, respectively. The ensuing responses show an act of recovering that is caused by transient loss of deposited cigarette smoke constituents from the resonator by evaporation [12]. The sensor shows a good recovery under exposure to cigarette smoke with recovery times of 126 s and 129.5 s for the 1st smoke and the 2nd smoke, respectively. However, a wet sensor cleaning by immersing in a special solution (e.g., acetone, isopropanol) [12–14] is still required to obtain optimal initial conditions. This method has even shown possibility to completely regenerate a cantilever sensor that has been fully loaded by nanoparticle agglomerates leading to instability during its operation [15].

Conclusions

The presented programmable de-embedding scheme works out properly with the reversing phase responses caused by crosstalk between heating resistor and piezoresistor. It provides flexibility in determining the reference parameters that latterly yields a monotonic phase response and demonstrate good frequency tracking under cigarette smoke exposure. The microcontroller has response time of 10 ms during measurements. In addition, the resolution is limited to ca. 0.1 Hz by the wave-generator. However, this programmable scheme requires a software-PLL to be able to track the resonant frequency in real-time. Consequently, the main controller needs to have sufficient speed to calculate and execute the program more accurately measure the phase and de-embed the sensor signal. The controller used in this study can perform up to 64 million instructions per second. Further development and testing, however, are still needed, especially for calculating phase difference and providing a reference setup as quantitative comparison.

Acknowledgments

A. Setiono would like to thank the Ministry of Re-search, Technology and Higher Education of the Republic of Indonesia (RISTEKDIKTI) for the PhD scholarship of RISET-Pro under no. 343/RISET-Pro/FGS/VIII/2016 (World Bank Loan No. 8245-ID) which is revised by no. 37/RISET-Pro/FGS/III/2019 and Indonesian-German Center for Nano and Quantum Technologies (IG-Nano) for the support. This work received funding from the European Union's Horizon 2020 research and innovation programme under no. 17IND05 MicroProbes. We are also grateful to Angelika Schmidt, Juliane Breitfelder, Aileen Michalski, Karl-Heinz Lachmund, and Ratna Indrawijaya for their assistance during preparation of research tools as well as many fruitful discussions.

References

[1] W. Shao, H. Zhang, and H. Zhou, "Fine Particle Sensor Based on Multi-Angle Light Scattering and Data Fusion," (eng), *Sensors (Basel, Switzerland)*, vol. 17, no. 5, 2017.

[2] Y. Wu, B. Yao, C. Yu, and Y. Rao, "Optical Graphene Gas Sensors Based on Microfibers: A Review," (eng), *Sensors (Basel, Switzerland)*, vol. 18, no. 4, 2018.

[3] B. Sviličić, E. Mastropaolo, R. Zhang, and R. Cheung, "Tunable MEMS cantilever

resonators electrothermally actuated and piezoelectrically sensed," *Microelectronic Engineering*, vol. 145, pp. 38–42, 2015.
[4] H. S. Wasisto, S. Merzsch, E. Uhde, A. Waag, and E. Peiner, "Handheld personal airborne nanoparticle detector based on microelectromechanical silicon resonant cantilever," *Microelectronic Engineering*, vol. 145, pp. 96–103, 2015.
[5] M. Bertke *et al.,* "Strategy toward Miniaturized, Self-out-Readable Resonant Cantilever and Integrated Electrostatic Microchannel Separator for Highly Sensitive Airborne Nanoparticle Detection," (eng), *Sensors (Basel, Switzerland)*, vol. 19, no. 4, 2019.
[6] HuiChao Shi, Shangchun Fan, Weiwei Xing, and Jinhao Sun, "2012 8th IEEE International Symposium on Instrumentation and Control Technology (ISICT): 11-13 July 2012, London, United Kingdom,"
[7] O. Brand, I. Dufour, S. M. Heinrich, and F. Josse, *Resonant MEMS: Fundamentals, Implementation and Application / edited by Oliver Brand, Isabelle Dufour, Stephen M. Heinrich, and Fabien Josse*. Weinheim, Germany: Wiley-VCH, 2015.
[8] M. Bertke *et al.,* "Analysis of asymmetric resonance response of thermally excited silicon micro-cantilevers for mass-sensitive nanoparticle detection," *J. Micromech. Microeng.*, vol. 27, no. 6, p. 64001, 2017.
[9] A. Setiono *et al.,* "Sensoren und messsysteme: Beiträge der 19. ITG/GMA-Fachtagung 26-27 Juni 2018 in Nürnberg,"
[10] A. Setiono *et al.,* "Real-Time Frequency Tracking of an Electro-Thermal Piezoresistive Cantilever Resonator with ZnO Nanorods for Chemical Sensing," *Chemosensors*, vol. 7, no. 1, p. 2, 2019.
[11] A. Setiono *et al.,* "Phase optimization of thermally actuated piezoresistive resonant MEMS cantilever sensors," *J. Sens. Sens. Syst.*, vol. 8, no. 1, pp. 37–48, 2019.
[12] J. Toledo *et al.,* "Piezoelectric MEMS Resonators for Cigarette Particle Detection," (eng), *Micromachines*, vol. 10, no. 2, 2019.
[13] H. S. Wasisto *et al.,* "Effect of Photoresist Coating on the Reusable Resonant Cantilever Sensors for Assessing Exposure to Airborne Nanoparticles," *Procedia Engineering*, vol. 47, pp. 302–305, 2012.
[14] H. S. Wasisto *et al.,* "Evaluation of photoresist-based nanoparticle removal method for recycling silicon cantilever mass sensors," *Sensors and Actuators A: Physical*, vol. 202, pp. 90–99, 2013.
[15] H. S. Wasisto, S. Merzsch, E. Uhde, A. Waag, and E. Peiner, "Partially integrated cantilever-based airborne nanoparticle detector for continuous carbon aerosol mass concentration monitoring," *J. Sens. Sens. Syst.*, vol. 4, no. 1, pp. 111–123, 2015.

DOI 10.5162/sensoren2019/3.3.1

Prozess- und Sicherheitssensorik für das Wasserstoff-Qualitätsmanagement

Jens Zosel[1], M. Eiserbeck[2], A. Hebestreit[2], T. Frosch[3], O. Kiesewetter[4], M. Wienecke[5], F. Altmann[6], Michael Mertig[1,7]
[1]Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg (KSI), Kurt-Schwabe-Straße 4, 04736 Waldheim, Deutschland
[2]Hochschule für Technik, Wirtschaft und Kultur Leipzig (HTWK), Karl-Liebknecht-Str. 132, 04277 Leipzig, Deutschland
[3]Leibniz-Institut für Photonische Technologien e.V. (IPHT), Albert-Einstein-Straße 9, 07745 Jena, Deutschland
[4]UST Umweltsensortechnik GmbH (UST), Dieselstraße 2 und 4, 98716 Geschwenda, Deutschland
[5]Materion GmbH, Alter Holzhafen 15, 23966 Wismar, Deutschland
[6]Fraunhofer-Institut für Mikrostruktur von Werkstoffen und Systemen IMWS, Walter-Hülse-Straße 1, 06120 Halle (Saale), Deutschland
[7]Professur für Physikalische Chemie, Mess- und Sensortechnik, Technische Universität Dresden, 01062 Dresden, Deutschland

Zusammenfassung

Zentrales Ziel des Konsortiums HYPOS ist die umfassende Nutzung von Strom aus Wind, Sonne und Wasser zur wirtschaftlichen Erzeugung von Wasserstoff mittels Elektrolyse in großtechnischem Maßstab und dessen Nutzung als Energieträger für mobile Anwendungen, chemischer Grundstoff sowie zur Wärme- und Elektroenergieerzeugung. Diese neu entstehende Wasserstoff-Infrastruktur benötigt eine leistungs- und zukunftsfähige, innovative Wasserstoffsensorik, die auf verschiedenen, den jeweiligen Anforderungen angepassten Messprinzipien beruht und damit weitaus flexibler und kostengünstiger als bisher an die Erfordernisse der H_2-Wertschöpfungskette hinsichtlich Qualitätsüberwachung und Gewährleistung maximaler Anwendungssicherheit beim Einsatz von H_2 anpassbar ist. Diese Zielstellung wird im Verbundvorhaben „Prozess- und Sicherheitssensorik für das Wasserstoff-Qualitätsmanagement (HyProS)“ verfolgt, in dem 12 Partner an der Entwicklung von fünf verschiedenen Sensorprinzipien zur Messung von Wasserstoff arbeiten. Diese Prinzipien betreffen die impedimetrische Messung von Spillover-Wasserstoff, die Nutzung der Ramanstreuung in optischen Fasern, ein Multisensor-Messsystem aus Metalloxid-Halbleiter-Gas-, Wärmeleitfähigkeits- und Drucksensor, sowie die Wasserstoffmessung mit Festelektrolyt- bzw. palladiumgestützen Sensoren für die Leckageüberwachung in einem breiten Konzentrationsbereich. Im Beitrag werden die Entwicklungen dargestellt.

Keywords: Wasserstoff-Gassensor, Spillover-H_2, Raman-Sensor, Festelektrolyt-Detektor, Palladium-MEMS-Sensor

Einleitung

Wasserstoff (H_2) bildet bereits seit langer Zeit einen wichtigen Ausgangsstoff für die chemische Industrie [1]. Heute wird sein Einsatz in zunehmendem Maß in immer mehr Anwendungen als sauberer Energieträger aus erneuerbaren Quellen in Erwägung gezogen. Beispiele dafür sind die Entwicklung von Brennstoffzellenfahrzeugen aller Leistungsklassen [2] und der Aufbau von Hochleistungs-Elektrolysesystemen für die Umsetzung überschüssiger Elektroenergie aus Wind-, Wasser- und Solarkraftwerken (*power-to-gas*) [3, 4]. Diese Applikationen führen zu einer Dezentralisierung der H_2-Versorgungskette von der Herstellung über den Transport und die Zwischenspeicherung bis hin zu den Endnutzern [5].

Wegen des weiten Konzentrationsbereichs, in dem H_2 entflammbare bzw. explosive Gemische mit Luft bildet, bestehen Risiken, sobald Leckagen an Tanks, Transportleitungen oder Anlagen zur Herstellung bzw. Verarbeitung von H_2 auftreten. Diese Risiken lassen sich durch eine angepasste Auslegung dieser Infrastruktur und den Einsatz hochsensitiver und hochselektiver Leckagesensoren mit unteren Messgrenzen < 10 Vol.-ppm weitgehend beherrschen [6]. Obwohl für diese Applikationen bereits heute viele verschiedene Sensoren verfügbar sind [7], bleiben dennoch mit diesen Sensoren immer noch wichtige Forderungen der Anwenderseite

unerfüllt, die beispielsweise die Selektivität, die Zuverlässigkeit und die Kosten betreffen [8]. Deshalb werden für die H_2-Sicherheits- und -qualitätsüberwachung im Verbundvorhaben HyProS hochminiaturisierte palladiumbasierte Sensoren (PBS) entwickelt, die im explosionsgeschützten Bereich einsetzbar sind [9]. Diese werden zur Erweiterung des Messbereichs hin zu Spurenkonzentrationen ab 0,5 Vol.-ppm mit coulometrisch arbeitenden Festelektrolytsensoren (FES) kombiniert [10], die zur Selektivitätsverbesserung eine chromatographische Vortrennung des H_2-haltigen Messgases aufweisen.

Weiterhin werden im Vorhaben katalytisch-impedimetrische H_2-Sensoren (IMS) [11] sowie faseroptische Raman-Sensoren (FRS) [12] entwickelt, die für die Überwachung der Wasserstoffkonzentration (< 20 Vol.-%) in Erdgasnetzen vorgesehen sind. Diese beiden Entwicklungsrichtungen werden durch ein diversitär-redundantes H_2-Multisensorsystem (MWD), komplettiert, das aus einem selektivem Metalloxid(MOX)-Halbleitergassensor, einem miniaturisierten Wärmeleitfähigkeitsdetektor und einem H_2-stabilen Drucksensor besteht [13], die mit einer intelligenten Signalverarbeitung verknüpft werden. Mit dem MWD soll in H_2-Druckspeichern bis 100 bar die Qualität des eingelagerten H_2 überwacht werden. Die Einsatzorte der neu zu entwickelnden Sensoren in der kompletten Versorgungskette sind in Abbildung 1 dargestellt.

Themenübergreifend wurden im Vorhaben weiterhin Querschnittsaufgaben vorgesehen, in denen Fragestellungen zur Materialauswahl für die Sensoren, zu deren Zuverlässigkeit und zu Alterungsvorgängen unter rauen Umgebungsbedingungen bearbeitet werden. Weiterhin werden Simulationsrechnungen zur Optimierung von Sensorlayouts, zu Bauteildeformationen und zu den Einsatzbedingungen der Sensoren im Projektkonsortium durchgeführt.

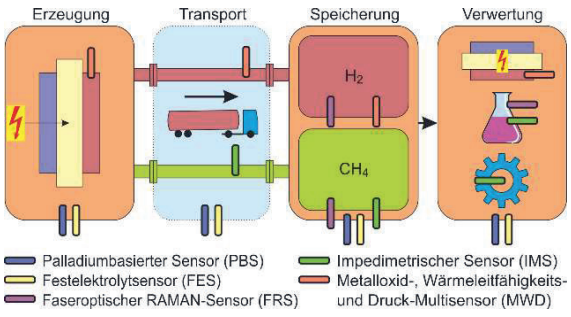


Abb. 1: Schematische Darstellung der Applikation von Sensoren und Mess-systemen, die im Verbundvorhaben HyProS entwickelt werden.

Beschreibung der Sensoren und Messprinzipien

A) IMS

Das Prinzip des impedimetrischen H_2-Sensors wird an der HTWK gemeinsam mit den Partnern Leibniz-Institut für Katalyse e.V., Rostock, Sciospec GmbH, Wurzen, und Union Instruments GmbH, Karlsruhe, entwickelt.

Das in Abbildung 2 dargestellte Prinzipschema des IMS zeigt, dass die in der Gasphase vor dem Sensor vorhandenen Wasserstoffmoleküle an einer mit Platin-Keimen versehenen Oberfläche aus einem platinhaltigen H_2-Akzeptor adsorbieren und dort in einzelne angeregte H-Atome dissoziieren. Diese Atome diffundieren an der Sensoroberfläche in einem zweiten Schritt auf einen für die H_2-Dissoziation nicht katalytisch aktiven Bereich (Spillover) und binden dort unter Abgabe eines Elektrons an oberflächliche Sauerstoffatome unter Bildung eines OH-Moleküls. Durch diese OH-Moleküle wird die Oberflächenleitfähigkeit des H_2-Akzeptors modifiziert, was durch Messungen der komplexen Wechselstromimpedanz an Interdigitalelektroden auf der Oberfläche des Akzeptors nachweisbar ist.

Da H_2 das einzige Molekül ist, welches auf diese Weise schnell und reversibel mit der speziell gestalteten Oberfläche in Wechselwirkung tritt, ist das Sensorsignal hochselektiv und weist auch in komplexen Gasgemischen mit vielen verschiedenen Einzelkomponenten wenig Querempfindlichkeiten auf.

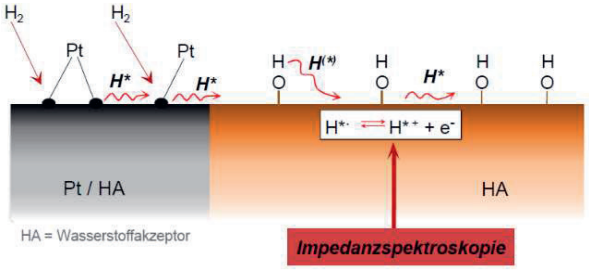


Abb. 2: Funktionsschema des impedimetrischen Sensors zur Messung des Spillover-Wasserstoffs.

B) FRS

Das in Abbildung 3 aufgeführte Schema der selektiven Gasdetektion mittels Raman-Streuung wird im Teilvorhaben des IPHT Jena in einem faseroptischen Sensor umgesetzt. Hierzu wird ein spezifisch angepasster Laserstrahl in eine speziell entwickelte Hohlkernfaser eingekoppelt, durch die das Messgas strömt. Auf dem Weg des Lichts durch die Faser kommt es in Abhängigkeit von der H_2-Konzentration zu zahlreichen inelastischen Streuvorgängen an H_2-Molekülen.

Dadurch treten im Streulichtspektrum neben der Laser-Wellenlänge ν_i Seitenbänder bei den Frequenzen $\nu_i \pm \nu_{vib}$ auf, die durch den Raman-Effekt erster Ordnung verursacht werden [12].
Die Intensität dieser Seitenbänder hängt von der Konzentration der korrespondierenden H_2-Moleküle ab. Da die Schwingungsfrequenz ν_{vib} von H_2 selektiv für dieses Molekül ist und sich von Frequenzen anderer Gasmoleküle deutlich unterscheidet, kann dieser Analyt mittels Spektrometer und hochempfindlichem Detektor mit hoher Selektivität bis zu Konzentrationen im oberen ppm-Bereich erfasst werden.

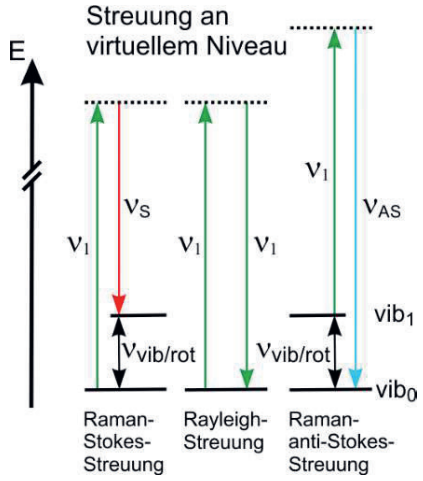


Abb. 3: Prinzipschema der selektiven H_2-Messung mittels Raman-Streuung.

C) MWD

Das in Abbildung 4 gezeigte Konzept des diversitär-redundanten H_2-Multisensorsystems wird von der UST und der CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt, entwickelt und erprobt. Es verknüpft drei Einzelsensoren, die jeweils ganz unterschiedliche Informationen aus dem Messgas liefern.
Während mit dem Wärmeleitfähigkeitsdetektor vor allem die Konzentration des vorhandenen Wasserstoffs charakterisiert werden kann, liefert der parallel ausgelesene Multielektroden-MOX-Sensor Informationen zu möglichen Verunreinigungen, die beispielsweise in Kavernenspeichern auftreten können. Der Drucksensor dient einerseits zum Abgleich der Signale der beiden anderen Sensoren und wird andererseits auch für die Druckmessung selbst an der Messstelle benötigt.

D) PBS

Das Prinzip des palladiumbasierten H_2-Sensors wird von der Materion GmbH in Wismar entwickelt und gemeinsam mit der Analytical Control Instruments GmbH (ACI), Berlin, in einen industrietauglichen Sensor für den Einsatz in explosionsgeschützten Bereichen umgesetzt.

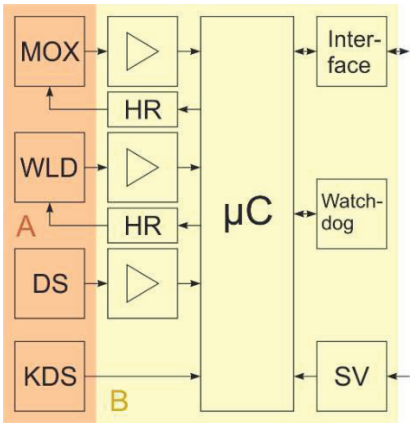


Abb. 4: Funktionsschema des diversitär-redundanten H_2-Multisensorsystems für die Überwachung von Großspeichern und Elektrolyseuren; MOX = Metalloxid-Gassensor, WLD = Wärmeleitfähigkeitsdetektor, DS = Drucksensor, KDS = Kalibrierdatenspeicher, HR = Heizungsregelung, µC = Mikrocontroller mit EEPROM, SV = Stromversorgung, A = austauschbarer intelligenter Sensorkopf, B = Datenaufnahme und –verarbeitung.

Es ist bereits seit längerem bekannt, dass verschiedene Materialien ihre Eigenschaften wie Farbe, Leitfähigkeit oder mechanische Eigenschaften ändern, sobald sich H_2 in ihr Gefüge einlagert. Insbesondere Palladium (Pd) ermöglicht die Lösung höherer H_2-Mengen und ändert dabei seine optischen, elektrischen und mechanischen Eigenschaften. Bei dem im Vorhaben zu entwickelnden Sensoren wird dabei der Effekt genutzt, dass die konzentrationsabhängige H_2-Einlagerung eine Volumenvergrößerung von Pd sowie von Pd-haltigen Legierungen und Kompositen hervorruft. Die damit verbundene Volumenausdehnung kann hochminiaturisiert mit dem in Abbildung 5 gezeigten mikroelektromechanischen System (MEMS) mittels einer Dehnungsmessung ausgelesen werden.

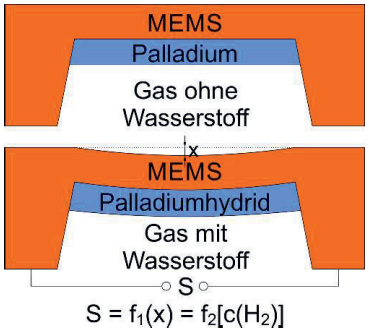


Abb. 5: Prinzipschema der selektiven H_2-Messung mittels Dehnungsmessung an H_2 einlagernden palladiumhaltigen Schichten.

Wie in Abbildung 5 dargestellt, wird das Signal S eines PBS als Funktion der Durchbiegung x ermittelt, die ihrerseits von der Menge des eingelagerten Wasserstoffs abhängt. Da diese vom H_2-Partialdruck in der Gasphase abhängt, korreliert das Signal S mit der H_2-Konzentration in der Gasphase. Durch die Aufbringung sehr dünner Pd-haltiger Schichten wird eine kurze Ansprechzeit erreicht.

E) FES

Das Konzept des coulometrischen Festelektrolytsensors mit chromatographischer Vortrennung wird gemeinsam von KSI und ACI entwickelt, um ein Gerätesystem für die verteilte H_2-Leckageüberwachung zu schaffen. Die Messung von H_2 mit einem zunächst nichtselektiven coulometrischen Festelektrolytsensor auf der Basis von stabilisiertem Zirconiumdioxid erfordert eine Vortrennung des Messgasgemisches, vor allem dann, wenn hohe Sauerstoffgehalte wie an Luft vorliegen. Diese Trennung kann beispielsweise in chromatographischen Säulen erfolgen, wie in Abbildung 6 dargestellt ist. Der Festelektrolytsensor wird dann als hochempfindlicher coulometrischer Detektor eingesetzt, indem durch konstante Polarisation der Messelektrode gegenüber einer Pt-Luft-Referenzelektrode ein gleichbleibender Sauerstoffpartialdruck im Bereich 0,01 - 0,1 Pa eingestellt wird. Ein einlaufender Peak eines oxidier- bzw. reduzierbaren Gases in einem sauerstoffarmen Trägergas wird dann durch einen Faraday-Strom mit positivem bzw. negativem Vorzeichen angezeigt. Die Vorteile dieses Detektors bestehen in seiner hohen Ansprechgeschwindigkeit, seinem breiten Messbereich über etwa 6 Zehnerpotenzen und seiner hohen Langzeitstabilität, die kalibrierfreie Messungen von vielen Monaten ermöglicht.

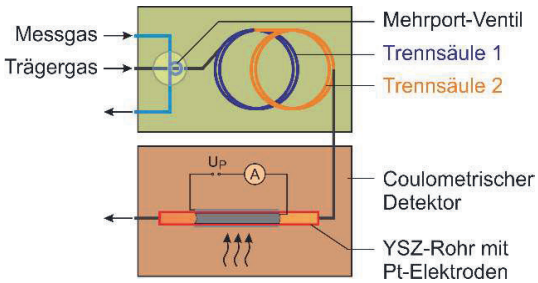


Abb. 6: Schema des Festelektrolytsensors mit chromatographischer Vortrennung zur selektiven Messung von H_2-Spurenkonzentrationen im ppb-Bereich.

Resultate der Sensorerprobungen

A) IMS

Der in Abbildung 7 dargestellte Versuchsstand zur Erprobung von IMS wurde an der HTWK aufgebaut.

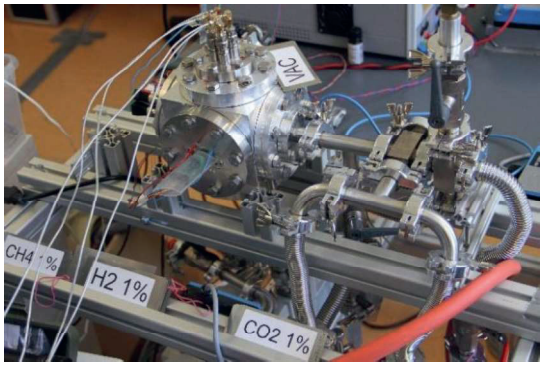


Abb. 7: Versuchsaufbau zur Charakterisierung impedimetrischer H_2-Sensoren.

Das bislang für die Sensorerprobung eingesetzte hochflexible Labor-Impedanzmessgerät wird demnächst durch sensorspezifische miniaturisierte Impedanzmesssysteme ersetzt, die vom Projektpartner Sciospec neu entwickelt wurden. Der Prüfstand ermöglicht den flexiblen Einsatz verschiedener Gasmischungen zur Nachbildung der Einsatzbedingungen in Erdgasnetzen.

Das in Abbildung 8 dargestellte Beispielergebnis belegt, dass die gewählten Sensorbeschichtungen den Spillover-Effekt zeigen und bereits geringe H_2-Konzentrationen von 1 Vol.-% zu deutlichen Änderungen des Betrages und der Phase der komplexen Wechselstromimpedanz führen. Mit zunehmender H_2-Konzentration ändert sich der Betrag der komplexen Impedanz vor allem bei niedrigen Frequenzen signifikant. Die Sensitivität der Phasenverschiebung auf H_2 ist bei Frequenzen um 1 kHz besonders groß.

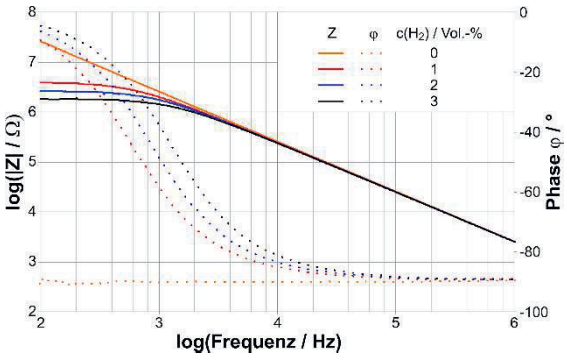


Abb. 8: Impedanzspektren an einem IMS bei verschiedenen H_2-Konzentrationen.

B) MWS

Der in Abbildung 9 gegebene Signalverlauf eines miniaturisierten WLD-Sensors belegt dessen hohe Sensitivität und Ansprechgeschwindigkeit. In Kombination mit den Signalen eines Metalloxid- und Drucksensors lassen sich neben der Wärmeleitfähigkeit noch weitere für die Beurteilung der Wasserstoffqualität wichtige Informationen aus dem Messgas gewinnen.

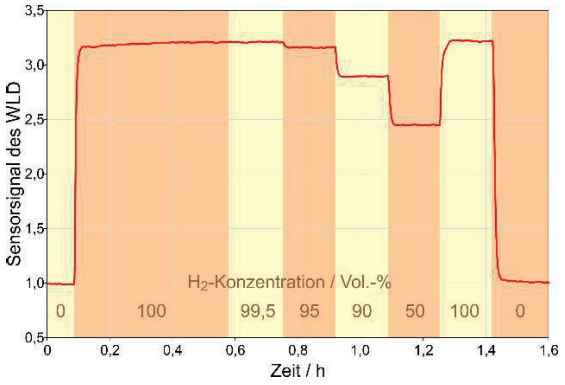


Abb. 9: Signalverlauf eines Wärmeleitfähigkeitsdetektors (WLD) des MWD gemäß Abbildung 4 bei verschiedenen Konzentrationen c(H_2).

Weitere Resultate zu diesem Sensorprinzip werden im Beitrag von O. Kiesewetter *et al.* „Diversitär-redundanter H_2-Sensor für H_2-Konzentrationen bis 100 % und einen Umgebungsdruck bis 100 bar für Elektrolyseuranwendungen“ vorgestellt.

C) PBS

In Abbildung 10 ist eine Signalkurve eines PBS in Gasen mit verschiedenen H_2-Konzentrationen dargestellt. Das Ergebnis belegt relativ kurze Ansprechzeiten unter 30 s, die weitgehend unabhängig von der Richtung des Konzentrationswechsels sind. Weiterhin zeigt das untersuchte Versuchsmuster nach einer Warmlaufphase von einem Beladungszyklus eine ausgezeichnete Wiederholgenauigkeit des Spannungssignals zwischen den einzelnen Zyklen, das logarithmisch von der H_2-Konzentration abhängt. Bei konstanter Konzentration treten nach Abschluss des Einstellvorgangs keine Signaldriften auf. Weitere Resultate zu diesem Sensorprinzip werden von M. Wienecke *et al.* im Beitrag „Sensitivität und Selektivität optischer und neuartiger MEMS-Wasserstoffsensoren“ vorgestellt.

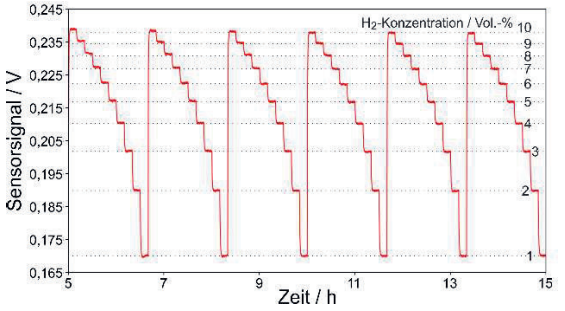


Abb. 10: Signalverlauf eines palladiumbasierten Sensors bei wechselnden H_2-Konzentrationen.

D) FES

Die in Abbildung 11 dargestellten Chromatogramme wurden an einem FES gemäß Abbildung 6 aufgenommen. Der Detektorstromverlauf zeigt, dass oxidierbare Gaskomponenten wie Wasserstoff und Methan bei ihrer Umsetzung im Detektor in positiven Peaks resultieren, die einen Sauerstofftransfer durch den Festelektrolyt in das Messgas hinein anzeigen. Im Gegensatz dazu rufen reduzierbare Gase wie Sauerstoff negative Strompeaks hervor, da dann Sauerstoff aus dem Messgas entfernt werden muss, um dessen Partialdruck im Detektor konstant zu halten. Ein Vorteil dieses Detektors besteht in seiner langzeitstabilen Funktion auf der Basis des Faraday-Gesetzes, die bei vollständigem Umsatz der zu messenden Gaskomponenten im Detektor gegeben ist [15]. Weitere Resultate zu diesem Sensorprinzip werden im Beitrag von P. Sood *et al.* „*Development and characterization of a highly selective hydrogen sensor system*“ vorgestellt.

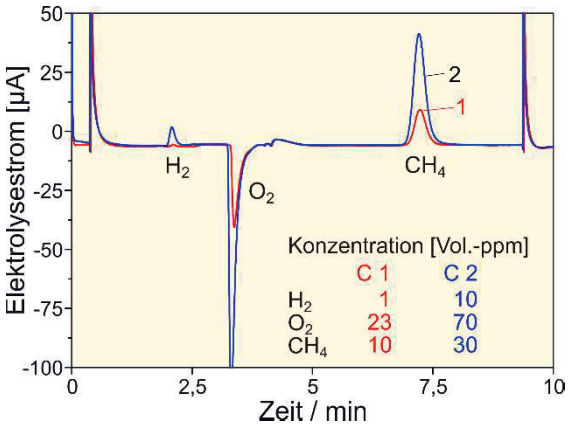


Abb. 11: Chromatogramme eines Festelektrolytsensors gemäß Abbildung 6, Einspritzvolumen = 1 ml.

E) Materialauswahl und Zuverlässigkeit

Neben den Arbeiten an den Sensorprinzipien und entsprechenden Versuchsmustern werden am IMWS Halle und an der TU Bergakademie Freiberg Untersuchungen zur Materialauswahl für die Sensoren und zur Zuverlässigkeit von Strukturelementen durchgeführt. Mit solchen Untersuchungen lassen sich auch Alterungsvorgänge charakterisieren, die vor allem unter den im Feldbetrieb teilweise auftretenden rauen Umgebungsbedingungen erwartet werden. Ein Beispiel für diese Untersuchungen ist in Abbildung 12 illustriert. Es zeigt einen Materialverbund aus einem gasdicht gesinterten stabilisierten Zirconiumdioxid (YSZ), einer porösen YSZ-Schicht und einer Platin-Netzelektrode nach einer Präparation mittels Laserablation, die am IMWS durchgeführt wurde.

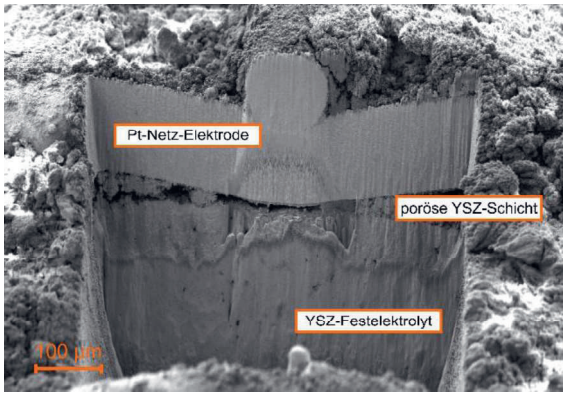


Abb. 12: REM-Aufnahme eines Materialverbundes aus dicht gesintertem YSZ, poröser YSZ-Schicht und Platin-Netzelektrode, Querschliff hergestellt durch Laserablation.

Schlussfolgerungen

Im Verbundvorhaben „Prozess- und Sicherheitssensorik für das Wasserstoff-Qualitätsmanagement (HyProS)“ werden neuartige Wasserstoffsensoren für eine zukünftige H_2-Infrastruktur zur Herstellung, Verteilung und Applikation von H_2 aus erneuerbaren Energiequellen entwickelt. Die von den 12 Projektpartnern verfolgten Projektziele sind auf fünf verschiedene Sensorprinzipien gerichtet, die auf der Nutzung des Spillover-Effektes, des H_2-Nachweises mittels Raman-Streuung, der Kombination verschiedener Einzel-Gassensoren zu einem neuartigen Multisensor, der Wechselwirkung zwischen H_2 und Palladium sowie der coulometrischen Umsetzung von H_2 beruhen. Für alle untersuchten Sensorprinzipien konnte im Vorhaben bereits der Funktionsnachweis erbracht und erste praxistaugliche Versuchsmuster geschaffen werden, die intensiv im Labor erprobt werden. Erste Feldtests werden z.Z. mit den am weitesten fortgeschrittenen Sensoren geplant und vorbereitet. Neben der eigentlichen Sensorentwicklung werden im Projekt Querschnittsaufgaben zur Materialauswahl und -charakterisierung für alle Sensorentwickler bearbeitet.

Danksagung

Das diesem Beitrag zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 03ZZ0724 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren, die für die Förderung danken.

Literaturnachweis

[1] R. Ramachandran, R.K. Menon, *International Journal of Hydrogen Energy* 23, 593-598 (1998); doi: 10.1016/S0360-3199(97)00112-2
[2] T. Sinigaglia, F. Lewiski, M. Eduardo, S. Martins, J. Cezar, M. Siluk, *International Journal of Hydrogen Energy* 42, 24597-24611(2017), doi: 10.1016/j.ijhydene.2017.08.063
[3] A. Maroufmashat, M. Fowler, *Energies* 10, 1089-1111 (2017); doi: 10.3390/en10081089
[4] D. Ferrero, M. Gamba, A. Lanzini, M. Santarelli, *Energy Procedia* 101, 50-57 (2016); doi: 10.1016/j.egypro.2016.11.007
[5] J. M. Ogden, *Annual. Review of Energy and the Environment.* 24, 227–279 (1999); doi: 10.1146/annurev.energy.24.1.227
[6] ISO 26142, ‚Hydrogen detector apparatus-stationary applications', 2010.
[7] T. Hübert, L. Boon-Brett, G. Black, U. Banach, *Sensors and Actuators B: Chemical* 157, 329-352 (2011); doi: 10.1016/j.snb.2011.04.070
[8] W. J. Buttner, M. B. Post, R. Burgess, C. Rivkin, *International Journal of Hydrogen Energy* 36, 2462-2470 (2011); doi: 10.1016/j.ijhydene.2010.04.176
[9] P. Fedtke, M. Wienecke, M.-C. Bunescu M. Pietrzak, K. Deistung, E. Borchardt, *Sensors and Actuators B: Chemical* 100, 151-157 (2004), doi: 10.1016/j.snb.2003.12.062
[10] M. Schelter, J. Zosel, W. Oelßner, U. Guth, M. Mertig, *Sensors and Actuators B* 187, 209–214 (2013); doi: 10.1016/j.snb.2012.10.111
[11] F. Rößner, U. Simon, M. Franke, U. Roland, Verfahren und Sensor zur Bestimmung der Wasserstoffkonzentration in fluiden Medien, *Schutzrecht*, DE 10011164B4 (2004).
[12] S. Hanf, R. Keiner, D. Yan, J. Popp, T. Frosch, *Analytical Chemistry* 86, 5278–5285 (2014); doi: 10.1021/ac404162w
[13] A. Ewert, O. Kiesewetter, S. Kittelmann, V. Melchert, Verfahren zur Erfassung brennbarer Gase, insbesondere zur Erfassung von Wasserstoff, *Schutzrecht*, EP 000001621882B1 (2009).
[14] R. Burch, F. A. Lewis, *Platinum Metals Rev* 15, 21-25 (1971)
[15] U. Guth, J. Zosel, W. Oelßner, M. Schelter, Gasanalysator, Schutzrecht, DE 102012004756 A1 (2013).

DOI 10.5162/sensoren2019/3.3.2

Sensitivität und Selektivität optischer und neuartiger MEMS-Wasserstoffsensoren

Marion Wienecke[1], Laura Godenrath[2], Bernd Zacharias[2]
[1]Hochschule Wismar, Institut für Oberflächen- und Dünnschichttechnik, PF1210, 23952 Wismar, Deutschland
[2] Materion GmbH, Alter Holzhafen 15, 23966 Wismar, Deutschland

Zusammenfassung

In diesem Beitrag werden die Potenziale von optisch schaltenden Pd-Dünnfilmsensoren (POS) untersucht sowie neuartige mikro-elektro-mechanischen Systeme (MEMS), die die Volumenänderung bei der Hydrierung des Pd ausnutzen. Beide Systeme können Wasserstoffkonzentrationen von 0,1 bis 100 vol % sicher und reproduzierbar nachweisen. Das physikalische Schalten von Pd bei der Einlagerung von Wasserstoff ist frei von Querempfindlichkeiten. Die Ansprechzeiten betragen unter 30 s (optischer Sensor) bzw. 5 s (MEMS-Sensor). Mit Pd-basierten MEMS-Sensoren können minimale Nachweisgrenzen von 100 ppm erreicht werden. Es werden mögliche Einsatzgebiete dieser Sensoren diskutiert.

Keywords: Palladium, optischer Wasserstoffsensor, MEMS-Wasserstoffsensor, Ansprechzeit, Querempfindlichkeit

1. Motivation

Seit mehr als 100 Jahren, seit Wasserstoff technisch genutzt wird, besteht die Notwendigkeit Wasserstoffkonzentrationen nachzuweisen und verlässlich zu messen. Für viele Industriezweige haben sich inzwischen Wasserstoffsensoren etabliert, so z.B. in der chemischen Industrie, in der Metallurgie, im Bergbau, bei der Überwachung von Kernkraftwerken. Die physikalisch-chemischen Grundlagen, auf denen diese Sensoren beruhen, sind seit vielen Jahrzehnten gut untersucht [1]. Eine Vielzahl von marktverfügbaren H_2-Sensoren beruht auf einigen wenigen Messprinzipien, darunter hauptsächlich elektro-chemische Sensoren, katalytische Pellistoren, Metall-Oxid Sensoren, Wärmeleitsensoren, Metall-Oxid-Halbleiter-Sensoren. Diese Sensoren dienen überwiegend zur Kontrolle und zum Explosionsschutz von stationären Anlagen und erfüllen entsprechende Standards. Nach ISO 26142 sind dies folgende Anforderungen [2]:

- Messbereich bis 4 vol% in Luft, Überlebensfähigkeit bis 100 vol%,
- Nachweisgrenze 100 ppm,
- Ansprechzeit (T90): 30 s,
- Erholungszeit (T10): 60 s,
- Umgebungstemperatur: -20 bis +50°C,
- Umgebungsdruck: 80-110 kPa,
- Feuchtigkeit: 20-80% relative Feuchte,
- Messgenauigkeit: $\pm$25 oder 50% abhängig von der H_2-Konzentration.

Spezifische Anforderungen ergeben sich für Anwendungen zur Prozesskontrolle sowie insbesondere für neue Anwendungsfelder im Bereich der erneuerbaren Energien [2], z.B. zur Überwachung von Brennstoffzellen, im Automobilbereich, bei der Einspeisung von regenerativ erzeugtem Wasserstoff in Erdgasleitungen, sowie auch in der Medizintechnik zur Messung von H_2-Anteilen im Atemgas, z.B. zum Nachweis von Laktoseintoleranz [3]. Diese neuen Anforderungen beziehen sich vor allem auf einen erweiterten Messbereich, für die Überwachung von mobilen Brennstoffzellen z.B. bis Wasserstoffkonzentrationen von 100 vol%, schnellere Ansprechzeiten, im Automobilbereich z.B. kleiner als 1 s und im Medizintechnikbereich deutlich geringere Nachweisgrenzen von weit unter 100 ppm. Ein Nachteil der o.g. marktverfügbaren H_2-Sensoren, insbesondere für die erwähnten neuen Einsatzgebiete, besteht zudem darin, dass diese alle, außer WLD-Sensoren, mehr oder weniger starke Querempfindlichkeiten zu anderen brennbaren Gasen aufweisen [4]. Seit jüngerer Zeit sind H_2-Sensoren auf dem Markt, die auf dem physikalischen Schalten von Pd oder Pd-Legierungen beruhen und von daher frei von Querempfindlichkeiten sind. Dazu gehören optische, resistive und MOS-Sensoren mit Pd-basierten Dünnfilmen [5]. Weitere neue Herausforderungen beziehen sich auf die Akzeptanz in den genannten Einsatzfeldern, da

etwa in der regenerativen Energiegewinnung, für mobile Geräte und im Automobilbereich und erst recht in der Medizintechnik die Messsysteme dezentral, nahe am Endverbrauer zum Einsatz kommen. In diesem Zusammenhang bestehen neue Anforderungen hinsichtlich der Sicherheit, des Energieverbrauchs, der Miniaturisierung und des Preises. Auch in dieser Hinsicht weisen Pd-basierte Dünnfilmsensoren Potenziale auf, da sie mittels Siliziumtechnologie in einer Massenfertigung hergestellt werden können.
In diesem Beitrag werden die Potenziale von optisch schaltenden Pd-Dünnfilmsensoren (POS) untersucht sowie neuartige mikro-elektro-mechanischen Systeme (MEMS), die die Volumenänderung bei der Hydrierung des Pd ausnutzen.

2. Material und Methoden

2.1. physikalische Eigenschaften des Pd-H Systems

Die schaltbaren physikalischen Eigenschaften beruhen auf dem Effekt der atomaren Einlagerung von Wasserstoff auf Zwischengitterplätzen im Metallgitter. In Abhängigkeit von seiner Konzentration wird der Wasserstoff beim Auftreffen auf Metall zunächst adsorbiert, dissoziiert und durch Diffusion im Metallgitter gelöst. Die α-Phase, eine feste Lösung, entsteht. Bei Überschreitung einer bestimmten metallspezifischen Wasserstoffkonzentration wird die Hydridphase (ά-Phase oder oft als β-Phase bezeichnet) gebildet. Dieser Prozess läuft schon bei Raumtemperatur ab und ist reversibel. Im Palladium besteht die α-Phase bei Raumtemperaturen bis zu einer Wasserstoffkonzentration von 1,68 at%. Im Konzentrationsbereich von 1,68 bis 37,6 at% ist das Stoffsystem zweiphasig und besteht aus der α- und der β-Phase. Ab 37,6 at% ist nur noch die β-Phase zu finden [6]. Die Löslichkeit von Wasserstoff in Palladium sinkt mit steigender Temperatur und steigt mit steigendem H_2-Partialdruck. Die α-Phase ist metallisch leitfähig und opak. Der Widerstand steigt mit zunehmender Konzentration von H im Metallgitter aufgrund von durch die Einlagerung bedingter Elektronenstreuung. Die β-Phase ist halbleitend und transparent [7].
Während der Wasserstoffeinlagerung und insbesondere beim Phasenwechsel kommt es zu einer Volumenänderung. Die Gitterkonstante für reines Pd beträgt 3,887 Å, bis zum Wasserstoffgehalt von ca. 2 at% steigt sie auf 3,895 Å und für die β-Phase mit ca. 37,6 at% Wasserstoff beträgt die Gitterkonstante 0,402 Å [8]. Damit ist eine Volumenausdehnung von 3,5% verbunden [6].

Diese Volumenausdehnung folgt der Formel

$$\Delta V/V=\Delta v/\omega\, c_H \quad (1)$$

$$\Delta a/a=1/3\ \Delta v/\omega\, c_H \quad (2)$$

$$\Delta a/a= \alpha_H\, c_H \quad (3)$$

- ΔV - Volumenwechsel des Metalls bei Lösung von n Wasserstoffatomen
- V - Ausgangsvolumen des Metalls
- Δv - charakteristische Volumenwechsel pro Wasserstoffatom, Δv steht in direkter Beziehung zum molaren Volumen ($V_M= \Delta v \cdot L$, L- Avogadro Konstante
- ω - Atomvolumen eines Metallatoms
- c_H - Verhältnis n/N, Verhältnis der Anzahl der Wasserstoffatome zur Anzahl der Metallatome (bei 50 at% beträgt $c_H=1$)
- a - Gitterkonstante,
- α_H - Ausdehnungskoeffizient bezüglich H-Einlagerung, beträgt für Pd 0,063

Diese Volumenänderung des Pd ist die Ursache für mangelnde Haftfestigkeit von Pd-Dünnfilmen. Strategien, diese zu verbessern bestehen darin, das Pd zu legieren oder haftvermittelnde Schichten anzuwenden, wir verweisen auf Fedtke et al. und dort zitierte Autoren [9].
Die Wasserstoffkonzentration in der α-Phase ist eine streng Partialdruckabhängige Funktion. Die Einlagerung von Wasserstoff im Metall erfolgt proportional der Wurzel aus dem Wasserstoffpartialdruck (Sievertsches Gesetz).

$$c_H \approx p(H_2)^{1/2} / K \quad (4)$$

- $p(H_2)$ - Wasserstoffpartialdruck
- K - temperaturabhängige Sieverts-konstante, für Pd = 0,12 bei 300 K

2.2. Sensitivität und Selektivität Pd-basierte Wasserstoffsensoren

Pd-basierte Wasserstoffsensoren sind heute marktverfügbar und werden mit verschiedenen Messprinzipien angeboten.
Der Nachweis der Wasserstoffkonzentration erfolgt optisch meist über die Bestimmung der Änderung der Transmission oder Reflexion des Pd-dünnfilmes mittels Transmissions- bzw. Reflexionsmessungen. Erstmalig wurde dies 1993 von Oh et al. berichtet [10]. Mehr sensitive aber aufwendigere Verfahren zur Bestimmung der Änderung der optischen Eigenschaften sind *Surface plasmon resonance* und interferometrische Messungen, Fiber-Bragg-Gitter sowie Messungen der Evaneszenz an Lichtleitern [5].

Resistive Pd-basierte Sensoren nutzen den um bis zu 80% steigenden Widerstand der Dünnfilme beim Übergang von der α- zur β-Phase als Maß für die Wasserstoffkonzentration. Diese Sensoren können ebenso wie die optische Sensoren Wasserstoffkonzentrationen bis 100% messen. Weiterhin kommen MOS-Sensoren mit Pd-Metall als Gate-Elektrode zum Einsatz und die Kapazität der MOSFET als Folge des steigenden Widerstands am Gate ist das Maß für die Wasserstoffkonzentration. Diese Sensoren können kleine Wasserstoffkonzentrationen nachweisen, sie beruhen darauf, dass die metallische α-Phase ihren Widerstand mit steigender Wasserstoffeinlagerung ändert. Eine Übersicht über die Prinzipien und Messbereiche von marktverfügbaren Pd-basierten Wasserstoffsensoren zeigt Tabelle 1.

Tab. 1: Übersicht marktverfügbarer Pd-basierter Wasserstoffsensoren [5]

Hersteller	Messprinzip	Messbereich
H2scan	resistiv	0,5-100vol%
H2scan	kapazitiv	15 ppm-1vol%
Makel Engineering	resistiv	0,1-100vol%
Applied Nanotech	Resonanz	100ppm-5vol%

Betrachtet man in referierten Fachzeitschriften Angaben zur Sensititvität und Ansprechzeit von Pd-basierten Sensorsystemen so zeigt sich, dass Ansprechszeiten von unter 3 s nicht erreicht werden. Boon-Brett berichtet in einem Review Artikel Ansprechzeiten (T_{90}) von >30 s für optische H_2-Sensoren [2]. Liu etal. [11] zeigen für optische, auf Transmission beruhende Sensoren für 4 vol% Wasserstoff eine Ansprechszeit von wenigen Sekunden. Die Ansprechzeit ist invers proportional zur Wasserstoffkonzentration, je kleiner die Konzentration, desto langsamer der Sensor. Für kleine Konzentrationen zwischen 10 und 100 ppm berichtet Lim [12] 2 min bis 30 s Ansprechzeit. Untersucht wurden hier mittels Mikrosystemtechnik hergestellte freitragende nanoskalige Pd/C-Nanodrähte.
Bezüglich der unteren Nachweisgrenze wird übereinstimmend berichtet, dass nanoskalige Schichtdicken und Kristallitgrößen sowie nanoskalig strukturierte Trägermaterialien zu höherer Sensitivität führen: Su et al. [13] berichten über resistive Messungen an CNT getragene nanoskaligen Pd-Strukturen und erreichen 200 ppm als untere Nachweisgrenze. Über Graphene getragene Strukturen berichten Panday et al. [14] und zeigen in resistiven Messungen minimale Nachweiskonzentration ebenfalls von 200 ppm. Die Ansprechzeiten in beiden Berichten liegen bei mehreren Minuten. Eine systematische Untersuchung von nanoskaligen Pd-Strukturen zeigen Noh et al. [15]. Es werden 2-D Palladium Dünnfilme, 1-D-(Nanowire)-Sensorstrukturen und Nanopartikel (0-D) in resistiven Messungen untersucht. Es wird gezeigt, dass je nanoskaliger die Strukturen desto kürzer wird die Ansprechzeit. Es werden minimale Nachweisgrenzen von 100 ppm und Ansprechzeiten zwischen 10 und 30 s erreicht.
Untersuchungen der Selektivität von Pd-basierten Sensoren in internationalen Fachzeitschriften zeigen übereinstimmend, dass Pd-basierte H_2-Sensorstrukturen unabhängig vom Messprinzip keine Querempfindlichkeiten zu anderen für die o.g. Anwendungen relevanten Gasen aufweisen. Vergleichende Untersuchungen von marktverfügbaren Wasserstoffsensoren sind z.B. von Palmisano et al. [4] durchgeführt worden. Pd-Dünnfilmsensoren zeigen demnach keine Querempfindlichkeit zu 1 vol% CH_4, CO_2 und 50 ppm CO. Panday et al. [14] zeigen für resistive Messungen an Graphene getragenen Pd-Nano-Dünnfilmen, dass gegen CO, Ethanol und Toluol keine Querempfindlichkeiten bestehen.

2.3. optisch schaltende Wasserstoffsensoren

In dieser Untersuchung haben wir PE-Folien beidseitig (siehe Abbildung 1) mit wenigen Nanometer dicken Pd/Au-Legierungen besputtert (Magnetron-Sputtern mit Patchwork-target). Die Au-Anteile wurden variiert von 18 bis 64 at%. Die Schichtdicken betrugen zwischen 10 und 15 nm. Gemessen wurde die Transmission von mono-chromatischem Licht mit einer Wellenlänge von 650 nm. Hinsichtlich der verwendeten Messtechnik verweisen wir auf Bramann et al. [16].

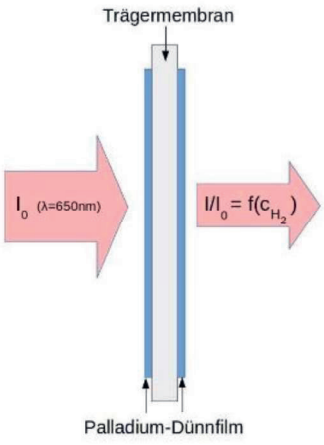


Abb. 1: Messprinzip optischer H_2-Sensor, Intensität des transmittierten Lichtes mit 650 nm Wellenlänge ist Funktion der H_2-Konzentration

2.4. MEMS basierte Wasserstoffsensoren

Mikro-elektro-mechanische Sensorstrukturen zum Nachweis von Wasserstoff nutzen physikalisch die Volumenänderung des Pd oder von Pd-Legierungen bei der Einlagerung von Wasserstoff. Die Messtechnik zum Nachweis der Volumenänderung kann unterschiedlich sein. Baselt et al. [17] berichten über einen Mikrocantilever, der ähnlich wie ein Bimetallstreifen von einer Seite mit Pd-beschichtet, sich unter H_2-Einfluß verformt. Diese Verformung wird kapazitiv gemessen und erlaubt den Nachweis von 1000 ppm bis 100% Wasserstoff. Es wird für 1 vol% Wasserstoff eine Ansprechzeit von 90 s angegeben.
Gurusamy et al. [18] beschreiben ebenfalls einen Microcantilever. Es können 50 ppm H_2 durch optische Auslesung der Verformung des Cantilevers sicher nachgewiesen werden.
Lee et al. [19] beschreiben Design-Vorschläge für sogenannte Nanogap-basierte Sensoren. Sie beruhen auf resistiven Messungen an Pd Nano-Strukturen, die bei Volumenausdehnung perkolieren und dadurch leitfähiger werden. Für eine einfache on-off Anordnung wurde gezeigt, dass 100 ppm H_2 sicher nachgewiesen werden können.
In dieser Arbeit haben wir marktverfügbare MEMS-Drucksensoren mit Pd/Au Dünnfilmen beschichtet (100 nm, Magnetron-Sputtern mit Patchworktarget). Eine Prinzipskizze zeigt Abbildung 2. Die MEMS-Drucksensoren beruhen auf piezoresistiven Eigenschaften von Silizium. Widerstandsstrukturen in Form einer Wheatstone-Brücke werden verstimmt, wenn die Si Membran sich verformt. MEMS-Drucksensoren werden in Massenfertigung mittels Siliziumtechnologie hergestellt. Das Ausgangssignal des MEMS-Sensors wird in den nachfolgenden Abbildungen als Maß für die Wasserstoffkonzentration herangezogen.

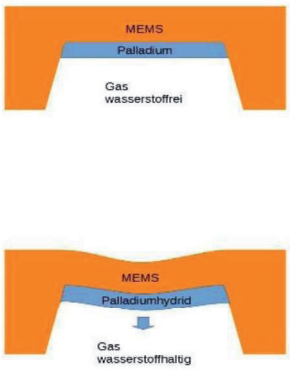


Abb. 2: Sensorprinzip eines Pd modifizierten MEMS-Sensors

3. Ergebnisse

3.1. Sensitivität und Selektivität optischer Wasserstoffsensoren

Abbildung 3 zeigt Messungen der optischen Transmission an einer oben beschriebenen Struktur (Abbildung 1) im Konzentrationsbereich 1 bis 100 vol% Wasserstoff in Luft. Dargestellt ist ein Messzyklus für Pd/Au mit 18 at% Au. Es zeigte sich, dass Filme mit höherem Au-Anteil weniger reproduzierbare Messergebnisse und insbesondere eine starke Drift der Basislinie durch Alterungserscheinungen zeigen.

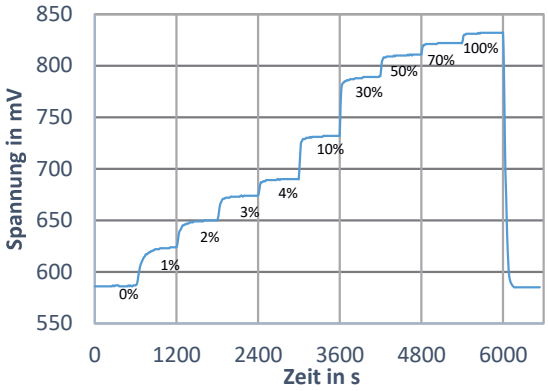


Abb. 3: Messung Transmission von Licht für H2-Konzentrationen 1 bis 100 vol% in Luft, Pd/Au Dünnfilm mit 18 at% Au

In Abbildung 4 ist die Sensorkennlinie in logarithmischer Form dargestellt. Die Transmission als Maß für das Verhältnis der H/Pd Atome im Kristallgitter folgt der Wurzel des H_2-Partialdruckes (Gleichung 4), sodass in lorarithmischer Darstellung ein linearer Verlauf entsteht.

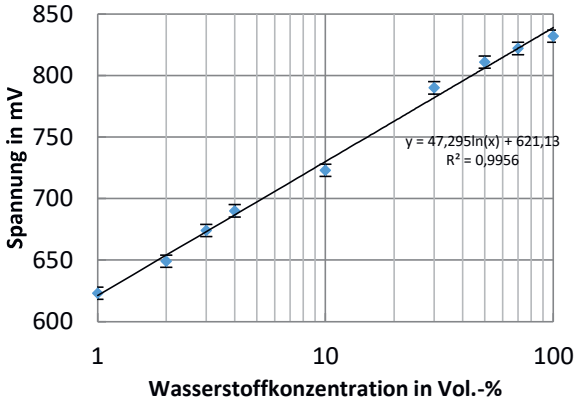


Abb. 4: Kennlinie optischer H2-Sensor aus Messung in Abbildung 3

Abbildungen 5 zeigt im Detail den Zeitverlauf des Signals beim Sprung von 3 auf 4 vol% Wasserstoff bei der optischen Messung. Es zeigt sich eine T_{90} Ansprechzeit von 20 s.
Messungen der optischen Transmission unter Wasserstoffeinfluss wurden auch für kleinere Konzentrationen zwischen 1000 und 5000 ppm durchgeführt. Konzentrationsverläufe lassen sich sicher und reproduzierbar messen bis etwa 1000 ppm als untere Nachweisgrenze. Abbildung 6 zeigt im Detail die Abklingzeit auf 10% des Ausgangswertes (T_{10}) beim Sprung von 5000 auf 1000 ppm Wasserstoff in Luft. Die T_{10}-Zeit beträgt ebenfalls etwa 20 s.

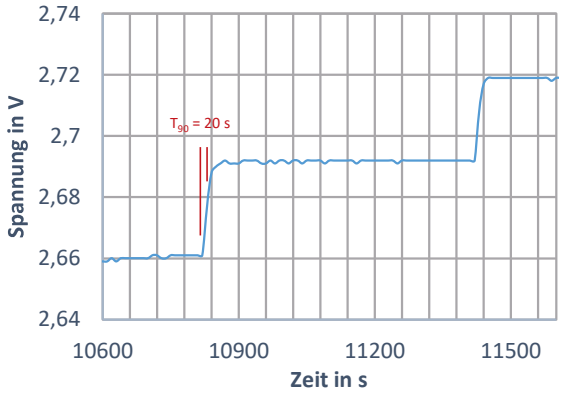


Abb. 5: Messung Transmission von Licht beim Wechsel von 2 auf 3 vol% H_2 Konzentration in Luft

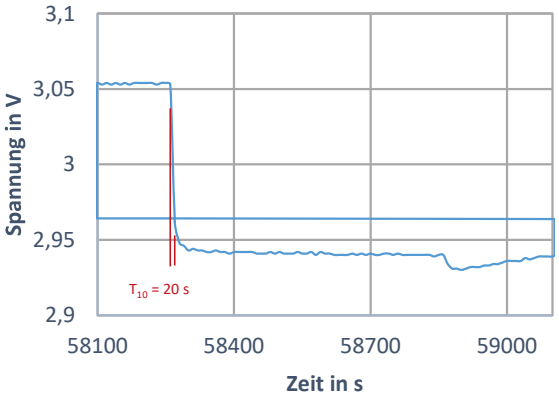


Abb. 6: Messung Transmission von Licht beim Wechsel von 5000 auf 1000 ppm H_2 Konzentration in Luft

Erste Untersuchungen zur Querempfindlichkeit wurden mit der beschriebenen Sensorstruktur durchgeführt, indem der Sensor CH_4-Luft-Gemischen anstelle von Wasserstoff-Luft ausgesetzt wurde. Eine entsprechende Messung mit CH_4 Konzentrationen stufenweise von 0 auf 4 vol % zeigt Abbildung 7. Der optische Pd-Dünnfilmsensor (blaue Kurve) zeigt darauf keine Messignale.

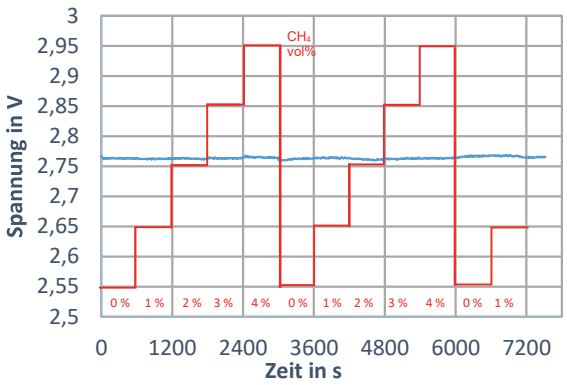


Abb. 7: Messung Transmission von Licht (blaue Kurve) beim Wechsel von CH_4 Konzentrationen zwischen 0 und 4 vol% in Luft

3.2. MEMS basierte Wasserstoffsensoren

In Abbildung 8 sind mehrere Messzyklen mit wechselnden H_2-Konzentrationen zwischen 1 und 10 vol % in Luft dargestellt, gemessen mit der Pd-modifizierten MEMS-Sensorstruktur (siehe Abbildung 2). Jeder Zyklus besteht aus 10 Stufen mit 10 min Haltezeit. Es zeigt sich nach kurzem Einschwingen der MEMS Struktur ein reproduzierbarer Signalverlauf. Abbildung 9 zeigt eine Detailansicht aus weiteren zyklischen Messungen an der gleiche MEMS-Sensor-Struktur beim Sprung von 1 auf 10 vol% Wasserstoff in Luft. Die T_{90} Zeit beträgt 5 s.

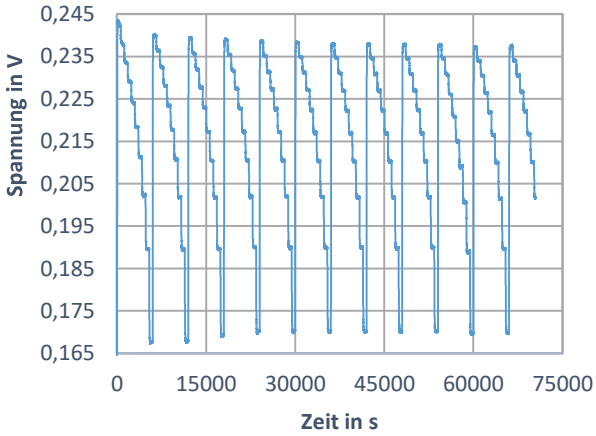


Abb. 8: Messsignal Pd-beschichteter MEMS Sensor, zyklische Messungen von 1 – 10 vol% und stufenweise Absenkung in 1% Schritten von 10 auf 1vol% Wasserstoff in Luft

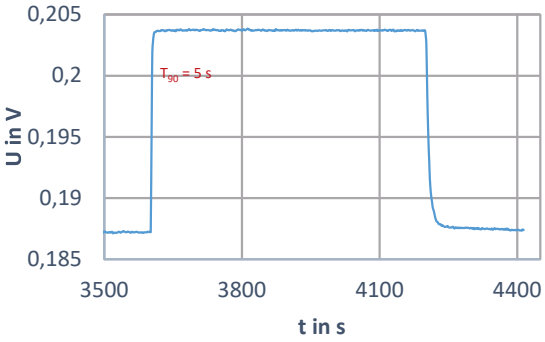


Abb. 9: Pd-modifizierter MEMS-Sensor beim Wechsel von 1 auf 10 vol % Wasserstoff in Luft, Messung bei Raumtemperatur

In Abbildung 10 ist ein Messzyklus über 22 h dargestellt mit Wechsel der H_2-Konzentration von 0 auf 10 vol% und zurück auf 0 vol% in Luft nach jeweils 10 min. Die Messungen erfolgten thermostatiert bei 60°C. Die Ansprechzeit T_{90} beträgt jeweils wenige Sekunden. Es zeigt sich ein stabiler Signalverlauf mit konstanter Signalhöhe und eine Drift der Nulllinie. Es wurde ein Referenz MEMS Sensor gleichzeitig aufgezeichnet, der keine Pd-Beschichtung aufweist. Demnach ist die Signaldrift durch die Umgebung bedingt. Infrage kommt die Umgebungstemperatur, z.B. Tag-Nacht Rhythmus, möglicherweise ist die Thermostatierung nicht konstant. Die Abweichungen betragen weniger als 10%.

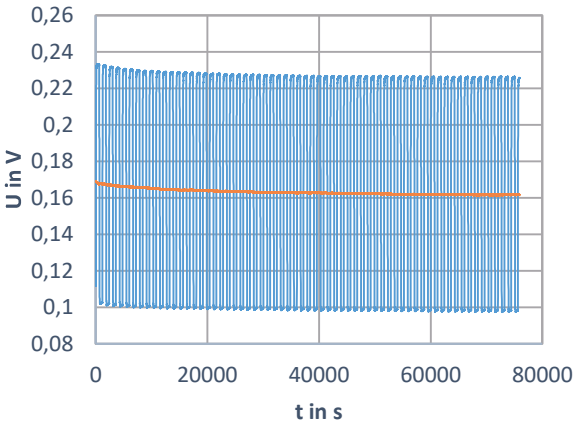


Abb. 10: Messsingnal Pd- beschichteter MEMS Sensor (blaue Kurve), zyklischer Wechsel zwischen 0 und 10 vol% H_2 Konzentration in Luft, thermostatiert bei 60 °C, zum Vergleich unbeschichteter MEMS-Referenzsensor (rote Kurve)

Untersucht wurden auch Konzentrationen unter 1000 ppm bis 100 ppm. Abbildung 11 zeigt mehrere Zyklen mit Wechsel von 400 ppm auf 1000ppm und in 200 ppm Stufen zurück auf 400 ppm, gemessen mit dem gleichen Pd-modifizierten MEMS-Sensor.
400 ppm lassen sich sicher nachweisen. Jedoch beträgt die T_{90} Zeit hier beim Sprung von 400 auf 1000 ppm etwa 3 min. In weiteren Messungen konnten auch 100 ppm sicher nachgewiesen werden. Durch weitere Messungen mit zyklischen Wechseln der Wasserstoffkonzentration zwischen 0 und 100 ppm haben wir T_{90} und T_{10} Zeiten von mehr als 5 min festgestellt.

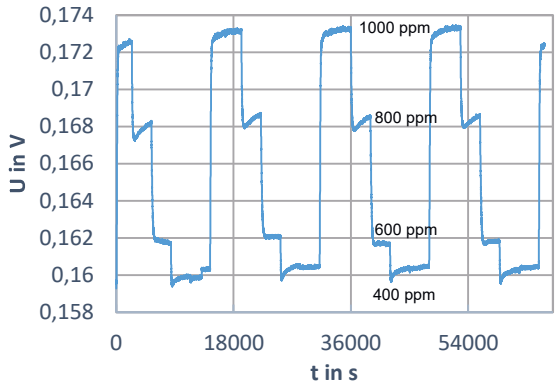


Abb. 11: Pd-beschichteter MEMS-Sensor beim Wechsel 1000 bis 400 ppm H_2 in Luft

4. Diskussion

Die dargestellten Messergebnisse zeigen, dass die sensorischen Eigenschaften von Pd durch physikalische Eigenschaften des Pd-H-Systems sehr gut erklärbar sind. Die Aufnahme von Wasserstoff durch Pd ist reversibel, sie folgt dem äußeren Partialdruck im Messgas (Sievertsches Gesetz). Der zunächst auf der Oberfläche adsorbierte Wasserstoff diffundiert in das Metallgitter, Triebkraft ist der Konzentrationsgradient, es stellt sich eine dem Phasengleichgewicht entsprechende Gleichgewichtskonzentration von H-Atomen im Metallgitter ein. Die Transparenz der Pd-Dünnfilme entsteht mit Auftreten der β-Phase ab der Löslichkeitsgrenze der α-Phase. Der optische Sensor kann von daher sehr kleine Wasserstoff-Partialdrücke nicht nachweisen, da die α-Phase metallisch opak ist. Der Übergang von α- zur β-Phase kann durch innere Spannungen im ultra-dünnen Pd/Au-Film in die eine oder andere Richtung gegenüber der Gleichgewichtslöslichkeitsgrenze, die aus dem

Pd-H-Phasendiagramm [6] [8] bekannt ist, verschoben sein. Wie auch immer, wir können zeigen, dass mittels optischer Messungen in Transmission 1000 ppm Wasserstoff sicher nachgewiesen werden können. Die T_{90} Zeit beträgt für den optischen Sensor unter 30 s.
Die Volumenänderung des Metallgitters kann durch MEMS-Systeme ebenfalls für sensorische Zwecke zum Nachweis H_2 genutzt werden. Für Konzentrationen zwischen 1 vol% und 100 vol% beträgt die T_{90} Zeit wenige Sekunden. Die Messungen sind reversibel und zeigen keine Drift. Eine Volumenänderung tritt auch bereits in der α-Phase auf, sodass mit der hier vorgestellten MEMS-Sensor-Struktur minimale Wasserstoffkonzentrationen von 100 ppm nachgewiesen werden können. Die T_{90} Zeit für kleine Konzentrationen unter 1000 ppm beträgt jedoch mehrere Minuten. Ungewöhnlich ist die in Abbildung 11 sichtbare Drift der Messsignale bei der Messung kleiner Konzentrationen nach Einstellung des Messwertes. Diese Erscheinung ist scheinbar reversibel, da sie sich bei der zyklischen Messung der Wasserstoffbeladungen exakt wiederholt. Dies kann an den elasto-machanischen Eigenschaften der MEMS-Struktur liegen oder durch innere Spannungen im Metallfilm (Rekristallisation, Relaxation [6]) verursacht sein. Dies ist Gegenstand weiterer Untersuchungen bezüglich elasto-mechanisch angepasster MEMS-Pd-Dünnfilm-Strukturen zum Nachweis kleiner H_2-Konzentrationen.

5. Schlussfolgerungen

Die hier dargestellten Pd-basierten optischen und MEMS Sensoren können sicher Wasserstoffkonzentrationen von 0,1 bis 100 vol% messen und weisen Potenziale für neue Anwendungsfelder auf, da sie auf dem physikalischen Schalten von Pd bei Raumtemperatur beruhen. Sie haben keine Querempfindlichkeiten, sie messen H_2-Konzentrationen in O_2-freier Umgebung und müssen nicht beheizt werden. Optische Sensoren sind zudem intrinsich sicher, da sie keine elektrischen Komponenten im zu überwachenden Volumen führen (Lichtleiterkabel). Pd-basierte MEMS Systeme haben im Konzentrationsbereich 1 -100 vol% kurze Ansprechzeiten von wenigen Sekunden. Wir sehen von daher Anwendungspotenziale für die Einspeisung von regenrativ erzeugtem Wasserstoff in das Erdgasnetz. Pd-basierte optische und MEMS Sensoren erfüllen die ISO 26142 zur Kontrolle und zum Explosions-schutz von stationären Anlagen.
Weitere Einsatzgebiete sehen wir für die Kontrolle von Prozessen, bei denen in reduzierender Atmosphäre von bis zu 100 vol% Wasserstoff gearbeitet werden muß, etwa bei der Herstellung von Reinstmetallen oder in der chemischen Verfahrenstechnik.
Potenziale weisen die hier vorgestellten Pd-MEMS-Sensoren für die Atemgasmesstechnik auf, wo verläßlich kleine H_2-Konzentrationen unter 100 ppm gemessen werden müssen. Hier kommt es besonders auf die Querempfind-lichkeitsfreiheit insbesondere gegenüber Ethanol und anderen Kohlenwasserstoffen an (Desinfektionsmittel in klinischer Umgebung). Die beschriebenen relativ langen Ansprech-zeiten wären in diesem Anwendungsfeld akzeptabel, da keine Konzentrationen im Explosivbereich auftreten können. Sichere Messwerte könnten nach einigen Minuten ausgelesen werden.

Dank

Das Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 03ZZ0724 gefördert.

Literaturnachweis

[1] T. Hübert, L. Boon-Brett, V. Palmisano, M. A. Bader, Developments in gas sensor technology for hydrogen safety, *International Journal of hydrogen energy* 39, 20474–20483 (2014); doi: 10.1016/j.ijhydene.2014.05.042.

[2] L. Boon-Brett, J. Bousek, G. Black, P. Moretto, P. Castello, T. Hübert, U. Banach, Identifying performance gaps in hydrogen safety sensor technology for automotive and stationary applications, *International Journal of hydrogen energy* 35, 373–384 (2010); doi: 10.1016/j.ijhydene.2009.10.064.

[3] Simon C Fleming, Evaluation of a hand-held hydrogen monitor in the diagnosis of intestianl lactase deficiency, *Ann Clin Biochem* 27, 499–500 (1990).

[4] V. Palmisano, L. Boon-Brett, C. Bonato, F. Harskamp, W. J. Buttner, M. B. Post, R. Burgess, C. Rivkin, Evaluation of selectivity of commercial hydrogen sensors, *International Journal of hydrogen energy* 39, 20491–20496 (2014); doi: 10.1016/j.ijhydene.2014.03.251.

[5] Prabhu Soundarrajan and Frank Schweighardt: Hydrogen Sensing and Detection, In: Gupta, Ram B., ed., *Hydrogen fuel: Production, transport, and storage*. CRC; Taylor & Francis, Boca Raton, Fla., London; 2009.

[6] Astrid Pundt, *Nanoskalige Metall-Wasserstoff-Systeme*. Universitätsverlag Göttingen; 2005.
[7] A. Mandelis, J. A. Garcia, *Sensors and Actuators B: Chemical* 49, 258–267 (1998).
[8] F.D. Manchaster, A.San-Martin, and J.M. Pitre, The H-Pd (Hydrogen-Palladium) System, *Journal of Phase Equilibria* 15, 62–83 (1994).
[9] P. Fedtke, M. Wienecke, M.-C. Bunescu, M. Pietrzak, K. Deistung, E. Borchardt, Hydrogen sensor based on optical and electrical switching, *Sensors and Actuators B: Chemical* 100, 151–157 (2004); doi: 10.1016/j.snb.2003.12.062.
[10] Y. Oh, J. Hamagami, Y. Watanabe, M. Takata, A novel palladium thin film hydrogen-detector, *J. Ceram. Soc. Jpn.* 101, 605 (1993).
[11] Y. Liu, Y. Li, Enhanced sensitivity of transmission based optical fiber hydrogen sensor with multi-layer Pd–Y alloy thin film, *Sensors and Actuators B: Chemical* 227, 178–184 (2016); doi: 10.1016/j.snb.2015.11.112.
[12] Y. Lim, Y. Lee, J.-I. Heo, H. Shin, Highly sensitive hydrogen gas sensor based on a suspended palladium/carbon nanowire fabricated via batch microfabrication processes, *Sensors and Actuators B: Chemical* 210, 218–224 (2015); doi: 10.1016/j.snb.2014.12.109.
[13] P.-G. Su, Y.-S. Chuang, Flexible H2 sensors fabricated by layer-by-layer self-assembly thin film of multi-walled carbon nanotubes and modified in situ with Pd nanoparticles, *Sensors and Actuators B: Chemical* 145, 521–526 (2010); doi: 10.1016/j.snb.2009.12.068.
[14] P. A. Pandey, N. R. Wilson, J. A. Covington, Pd-doped reduced graphene oxide sensing films for H2 detection, *Sensors and Actuators B: Chemical* 183, 478–487 (2013); doi: 10.1016/j.snb.2013.03.089.
[15] J.-S. Noh, J. M. Lee, W. Lee, Low-dimensional palladium nanostructures for fast and reliable hydrogen gas detection, *Sensors (Basel, Switzerland)* 11, 825–851 (2011); doi: 10.3390/s110100825.
[16] G. Bramann, B. Zacharias, M. Wienecke, New optical hydrogen sensor with fast response time based on multilayer palladium-nickel-PTFE thin film for explosion-proof detection of high H 2 concentrations of 1-100%. In: *Third European Workshop on Optical Fibre Sensors*, 66191L. SPIE; 2007.
[17] D. R. Baselt, B. Fruhberger, E. Klaassen, S. Cemalovic, C. L. Britton, S. V. Patel, T. E. Mlsna, d. McCorkle, B. Warmack, Design and performance of a microcantilever-based hydrogen sensor, *Sensors and Actuators B: Chemical* 88, 120–131 (2003); doi: 10.1016/S0925-4005(02)00315-5.
[18] J. T. Gurusamy, G. Putrino, R. d. Jeffery, K.K.M.B. D. Silva, M. Martyniuk, A. Keating, L. Faraone, MEMS based hydrogen sensing with parts-per-billion resolution, *Sensors and Actuators B: Chemical* 281, 335–342 (2019); doi: 10.1016/j.snb.2018.07.118.
[19] Junmin Lee, Wooyoung Shim, Jin-Seo Noh, and Wooyoung Lee, Design Rules for Nanogap-Based Hydrogen Gas Sensors, *ChemPhysChem* 13, 1395–1403 (2012).

DOI 10.5162/sensoren2019/3.3.4

Development and characterization of a highly selective hydrogen sensor system

Pramit Sood[1], Jens Zosel[1], Michael Mertig[1,2], Wolfram Oelßner[1], Andreas Klockow[1], Olaf Herrmann[3], Michael Woratz[3]
[1]Kurt-Schwabe Institut für Mess- und Sensortechnik e.V. Meinsberg, Kurt-Schwabe-Straße 4, 04736, Waldheim, Deutschland
[2] Technische Universität Dresden, Professur für Physikalische Chemie, Mess- und Sensortechnik, 01062 Dresden, Deutschland
[3]ACI Analytical Control Instruments GmbH, Volmerstr.9, 12489 Berlin, Deutschland

Abstract

Distributed use of hydrogen (H_2) has grown with its increasing application in the mobility and energy sector. The need for highly sensitive, selective, long-term stable and miniaturized H_2 sensor systems is essential for the safety and quality management of a future H_2 supply chain. The present contribution reports on the development and characterization of a H_2 sensor system based on a solid-electrolyte flow-through cell (8 mol.-% yttria-stabilized zirconia, YSZ) operated in coulometric mode and combined with gas chromatographic (GC) pre-separation. Miniaturization of system components and results on parametric characterization are provided. A valve injection system for gas sampling was developed as a stainless steel 3D-printed valve manifold with miniaturized solenoid valves. The GC column oven (heating/cooling module) was accomplished using resistive heating and forced convective cooling of an cylindrical carrier. Both design concepts are reported together with results of their performance. A commercial solid electrolyte coulometric detector (SEC) was characterized under different operating conditions, i.e. at different detector temperatures, oxygen partial pressures, injected H_2 concentrations and carrier gas flowrates, to gain insights into the essential sensor parameters like H_2 peak attributes, accuracy and long-term stability.

Keywords: Safety monitoring, hydrogen sensor system, yttria-stabilized zirconia (YSZ), solid-electrolyte coulometry, gas chromatography

1. Introduction

Hydrogen has been widely used in chemical industry as well as petrochemical refineries for a century [1]. An increasing recognition of hydrogen as a clean energy source has extended it's application to various other fields. On the one hand, it is increasingly used in the e-mobility sector after transformation in fuel cells [2]. And, on the other hand, its importance is growing rapidly in a mixed energy infrastructure, where hydrogen is being produced from renewable energy sources (e.g. power-to-gas technologies) [3, 4]. These applications have resulted in the decentralization of the hydrogen supply chain [5] in the context of production (electrolyzer facilities), transport (integrated with natural gas distribution pipelines) and various end users.
Because of its wide flammability/explosion range in air and small molecular size, hydrogen poses potential danger due to leakage. A widely used risk management approach in safety monitoring applications is to reliably and selectively detect hydrogen at trace concentrations (<10 vol.-ppm) at start of a leakage [6].

Above-mentioned requirements dictate the need for more sensitive, selective, and long-term stable sensor systems, which can be miniaturized for flexible field installations. In [7, 8], different commercial sensor technologies and test methods are described and evaluated for hydrogen detection in safety applications. In [9-11] the usage of a solid-electrolyte-based coulometric gas sensor as a GC detector is described, which is capable to detect trace concentrations of oxidizable gas components (H_2, CH_4) in the presence of air. The detector measures the titration current, which is related to the gaseous analyte concentration. The analytes are oxidized or reduced at the triple phase boundary (TPB) of the SEC electrode (Pt/YSZ/gas phase), which is polarized versus a Pt-air reference electrode to establish a constant low oxygen partial pressure in the carrier gas [10]. Eq. 1 describes the relation between the amount of analyte *x* and the detector signal *I* based on Faraday law:

$$x = \int_{t_S}^{t_E} \frac{(I - I_{BL})}{z \cdot F} \cdot dt \qquad (1)$$

where t_S and t_E are the starting and ending times of the peak, respectively. F denotes the Faraday constant, I_{BL} stands for the baseline current and z stands for the number of electrons transferred in the reaction.
The detector was characterized and optimized in terms of material properties and instrumentation related noise sources to enable selective measurement of H_2 concentrations down to 0.2 vol.-ppm [10, 12]. The present contribution describes the design and test of some essential components of a miniaturized chromatographic measuring system with a SEC, to develop a miniaturized field device. It also reports the results from operational characterization of a commercially available SEC as test platform for future improvements and design optimizations.

2. Experimental

2.1 Detector characterization

Fig. 1 shows a scheme and a picture of the existing measurement system. A commercial GC (8610C, MG #1, SRI Instruments Europe GmbH, Bad Honnef, Germany) was modified by adding a SEC with temperature control (F) and purpose-built potentiostatic circuit (E) as described in [10]. The gas sample injection system is enabled through a 10-port rotational valve (A) and 1mL sample loop. The sample is separated in a temperature programmed air bath oven (B) containing a combination of molecular sieve 13X and silica gel stainless steel (SS) packed columns (length: 2m; OD/ID: 1/8”/2.1 mm) and analyzed using a commercial SEC (D) (O2-DF-28.0 from ZIROX Sensoren & Elektronik GmbH, Greifswald, Germany). The separation process is described in detail elsewhere [13]. Nitrogen (N_2) of quality 5.0 was used as carrier gas. The sample loop was continuously flushed with sample gas mixtures, prepared with calibrated mass flow controllers (G) from Brooks Instruments (Hatfield, USA) using test gases with 1000 and 102.7 vol.-ppm H_2 in synthetic air.
For the characterization of the essential sensor parameters like peak behavior, sensitivity, accuracy and reproducibility, the measurement system was inspected at different detector temperatures (600 – 750 °C), oxygen partial pressures/polarization voltages (-300 mV to -450 mV), injected H_2 concentrations (0.2 – 180 vol.-ppm) and carrier gas flowrates (~6 – 16,5 mL/min). All chromatograms were recorded at a sample loop volume of 1 mL, sample loop temperature of 60 ± 1°C and ambient pressure conditions.

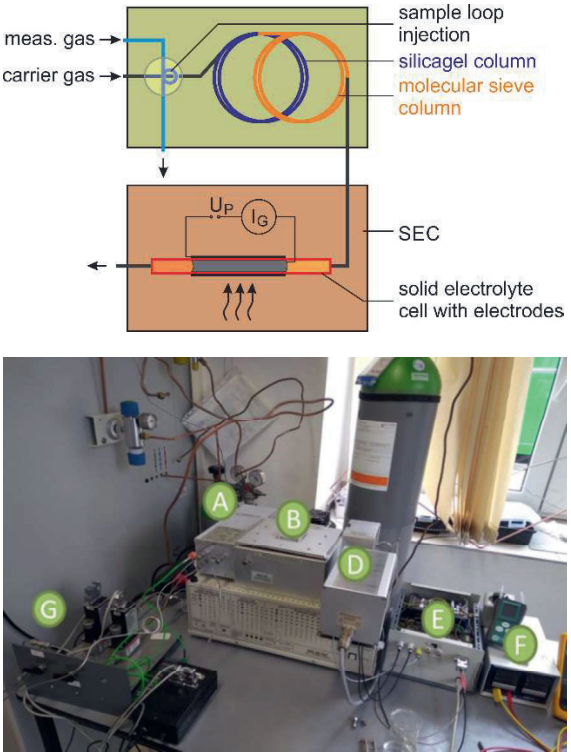


Fig. 1: Scheme (above) and picture (below) of the present measurement system, letter labelling is described in the text.

2.2 Component miniaturization

For the sample loop injection, a miniaturized SS valve manifold was developed, which is equipped with 8 miniaturized solenoid valves (S0705DX-CO, 3/2 way solenoid valves, SMC GmbH, Germany) mounted on the manifold for flow actuation. A schematic of the concept is shown in Fig. 2. It also shows a laboratory prototype which was manufactured by 3D-printing (AM Metals GmbH, Freiberg, Germany). The manifold was tested for gas tightness of the external fittings and internal channels, using helium leak test system (ASM 310, Pfeiffer Vacuum, Asslar, Germany). To characterize the gas dispersion in the manifold, it was connected to carrier (N_2) and sample gas flows (0.1 vol.-% H_2 in N_2) as well as to a commercial SEC detector, operated by a potentiostat (Interface 1000, Gamry Instruments, Warminster, USA) and a sample loop with the volume 1 mL. The column ports were short-circuited. Subsequently the coulometric current peaks were measured at this arrangement. Potentiostatic scans with multiple load and injection steps were performed at different detector temperatures (600 – 650 °C), carrier gas flowrates (10 – 50 mL/min) and injected H_2 concentrations (10 – 300 vol.-ppm).

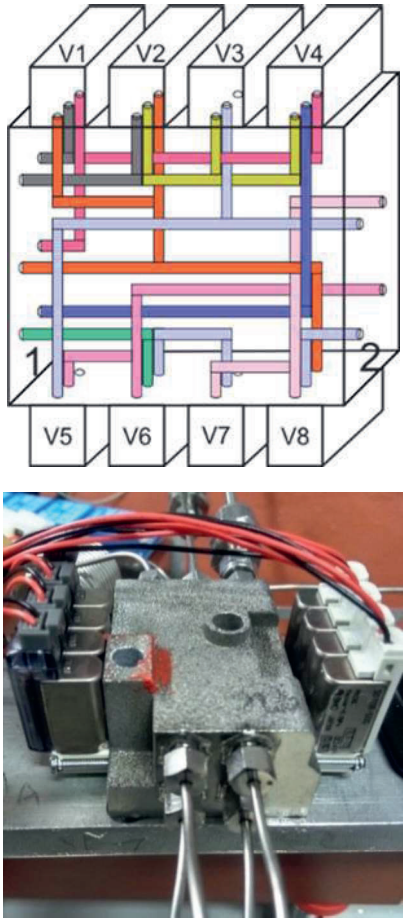


Fig. 2: Flow circuitry (above) of the miniaturized 3D-printed valve manifold (45x35x25 mm³), where 1) shows connections to sample loop, measuring gas, carrier gas and detector and 2) shows connections to columns, below: prototype for laboratory tests.

Fig. 3 shows the design concept and the corresponding prototype of a miniaturized column heating/cooling module (HCM). The GC columns were wound upon an aluminum (Al) cylinder as carrier together with an isolated heating wire (ThermoExpert GmbH, Stapelfeld, Germany) and a Pt-200 temperature sensor (Heraeus Sensor Technology GmbH, Kleinostheim, Germany). A silicon free heat conductive paste (Techtronix GmbH, Cologne, Germany) was applied to ensure uniform thermal contact. The Al carrier is equipped with an embedded heat sink structure and cooled by an axial fan. The columns were thermally isolated with glass wool (not shown). The performance of the prototype was evaluated by heating/cooling cycles with constant or increasing heating power. The influence of a swing-type flap at the outlet of the cylindrical Al-carrier was tested as well. The cooling cycle was recorded always at maximum power of the axial fan.

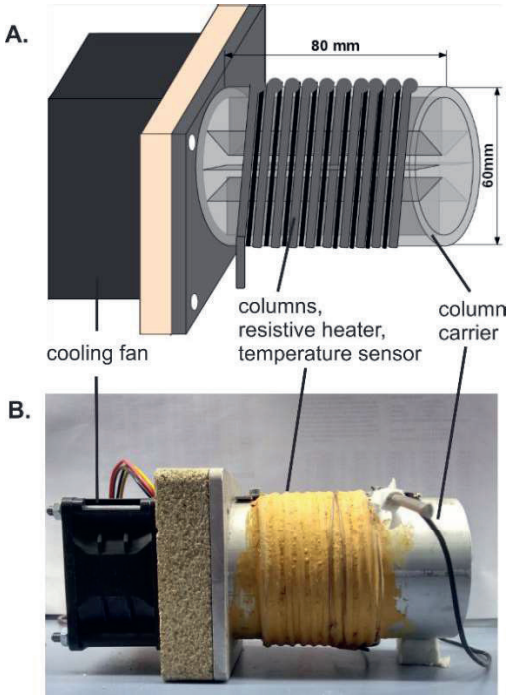


Fig. 3: Scheme (A) and picture (B) of a miniaturized heating/cooling module for the columns.

3 Results

3.1 Effect of detector temperature and polarization voltage

Fig. 4 shows hydrogen peaks measured at different SEC temperatures. The peaks were measured at constant polarization voltage and injected H_2 concentration. The average carrier gas flow rate was adjusted to 10.5 mL/min at standard conditions (= standard cubic centimeters per minute, sccm).

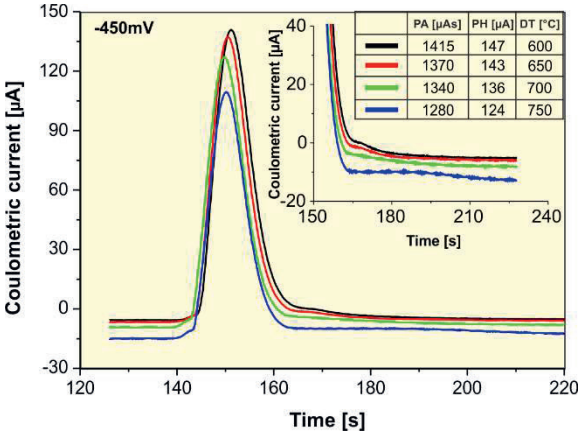


Fig. 4: Chromatogram segment showing H_2 peaks and dependence of their peak areas (PA) and peak heights (PH) on detector temperatures (DT), at polarization voltage = -450 mV, c(H_2) = 180 vol.-ppm in synthetic air.

The results show a decrease of the H_2-related peak area and peak heights with increasing detector temperatures. A closer look at the coulometric currents in the time segment between 160 and 200 s in the inset of Fig. 4 indicates an apparent peak tailing, which prolongs with increasing SEC temperature. This peak tailing indicates a H_2 storage near the triple phase boundary (TPB) of the detector at higher temperatures. It is assumed that this temporary H_2 storage (approx. 10% if the peaks at 750 and 600°C are compared) occurs at the electrode/electrolyte interface. As illustrated in Fig. 5 this interface consists of a porous YSZ layer sintered onto the YSZ tube to increase the TPB at the surface of the Pt wire mesh electrode. The stored H_2 subsequently elutes from this layer and results the tailing. This phenomenon could be correlated to the findings reported in [14] where an increase of peak heights and areas was seen with increasing temperatures in cyclic voltammetry (CV) investigations on an identically constructed detector at comparable flow rates of measurement gas. The peak tailing duration at 750 °C prevents the complete peak integration since the following oxygen peak limits the integration time. At 600 °C the amount of charge found in the peak corresponds to the injected amount of H_2 according to Eq. 1.

To investigate the peak dependency on oxygen partial pressure (pO_2), different polarization voltages from -300 mV (pO_2 = 25 mPa) to -450 mV (pO_2 = 0.27 mPa) were applied. As shown in Fig. 6, the described peak tailing can be observed at all applied polarization voltages, indicating that the H_2 storage is occurring in a broader range of $p(O_2)$. No other influence of the polarization voltage is visible.

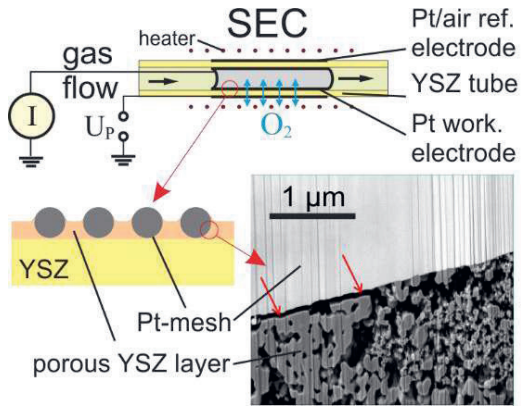


Fig. 5: Scheme of the solid electrolyte cell in the SEC and SEM image of the interface between the Pt mesh electrode and the porous YSZ layer situated between electrode and YSZ electrolyte to enhance TPB.

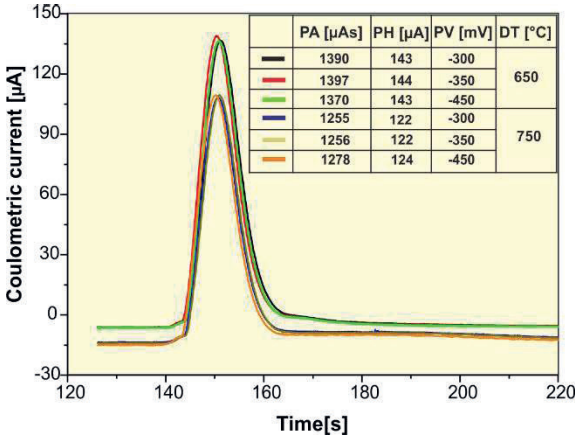


Fig. 6: Chromatogram segment showing H_2 peaks and dependencies of peak area (PA) and peak height (PH) on polarization voltage (PV) at two detector temperatures (DT), $c(H_2)$ = 180 vol.-ppm in synthetic air.

3.2 Peak behavior at different hydrogen concentrations

The SEC was tested for chromatographic peak behavior and detection uncertainties at different injected H_2 concentrations in the range 2 – 180 vol.-ppm. The detector response was recorded for varying detector temperatures between 600 – 750 °C at constant polarization voltage -450 mV. The resulting chromatogram segments given in Fig. 7 indicate that the peak tailing, described in Section 3.1, occurs at all measured concentrations. In addition, front-end and back-end shoulders are observed at concentrations < 50 vol.-ppm. This effect is mainly caused by the gas chromatographic columns. It vanishes at trace concentrations <10 vol.-ppm.

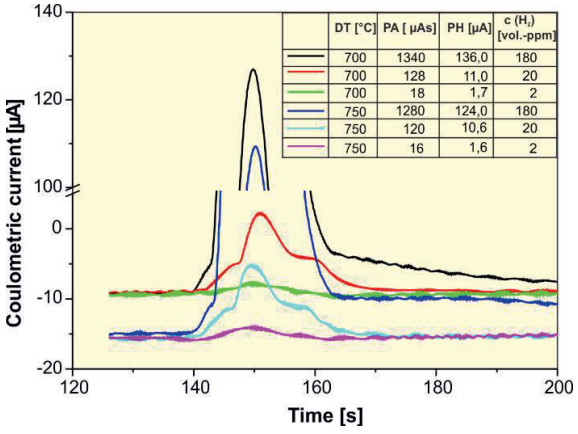


Fig. 7: Chromatographic H_2 peaks with their peak areas (PA) and peak heights (PH) at varying injected $c(H_2)$ = 2 – 180 vol.-ppm, polarization voltage = -450 mV, detector temperature (DT) = 700 and 750 °C.

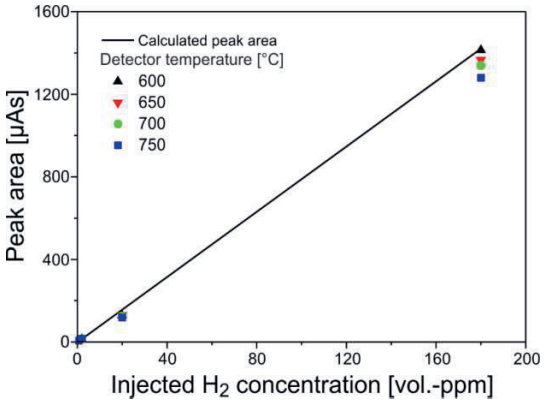


Fig. 8: Comparison between peak areas calculated by Faraday law (line) and peak areas derived from chromatograms at different temperatures and H_2 concentrations (dots).

The peak areas measured in different experiments were summarized with respect to the analyte concentration as shown in Fig. 8. The results prove that the measured peak areas follow Faraday law closely within the accessible precision at 600 °C. Due to the above mentioned storage effect a deviation of around 4% was found at 750 °C.

3.3 Influence of carrier gas flow rate

The hydrogen peak parameters were evaluated at varying carrier gas flow rates. Fig. 9 shows these results at two selected detector temperatures. The results in Fig. 9 reveal that with increasing flow rate, the peak height (and correspondingly the signal-to-noise ratio) increases while peak width and retention time decrease as expected. The peak area does not change.

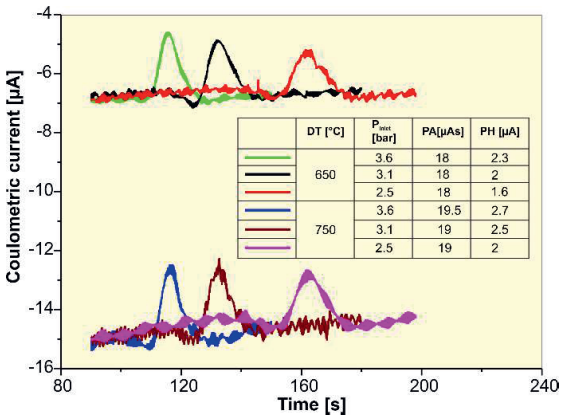


Fig. 9 Chromatogram segments showing H_2 peaks and dependency of H_2 peak areas (PA) and peak heights (PH) on carrier gas inlet pressures p_{inlet}= 2-3.6 bar (flowrates = 6 – 16 mL/min), c(H_2) = 2 vol.-ppm; polarization voltage = -450 mV; detector temperatures (DT) = 650 and 750 °C.

The neglectable deviations between the measured peak areas indicate that the detector oxidizes the incoming hydrogen completely at all investigated flow rates. This result is supported by the findings at a similar SEC given in [10]. These insights will be used for optimizing parameters like carrier gas flow rate, detector geometry, peak retention time, signal-to-noise ratio and completeness of H_2 conversion. They would also play a key role in further diminishment of the lower limit of detection (LLOD) and size of the measuring system.

3.4 Characterization of miniaturized components

3.4.1 Valve injection system

The newly developed injection manifold was characterized with respect to overall gas tightness and sample dispersion after injection. The leakage rates amount roughly to 10^{-8} mbar·L/s. The average values of open circuit potentials (OCP) at tested carrier gas flow rates and detector temperatures indicate oxygen concentrations between 0.3 – 0.7 vol.-ppm. Fig. 10 shows experiments performed with the manifold with different flow rates of the carrier gas. The illustrated sections of different potentiostatic scans contain one injection event. The stable baseline current proves an appropriate gas tightness of the prototype. As expected, the H_2 peak height increases with flow rate while the peak area remains constant. Furthermore, the H_2 peaks show an acceptable tailing, which indicates remaining possibilities to reduce dead volumes in the manifold and the fittings. Another evidence for these dead volumes are negative oxygen peaks, occurring after the first several actuations from inject to load events. These negative peaks decrease rapidly and vanish after 4-6 injections due to the ongoing flushing of these dead volumes.

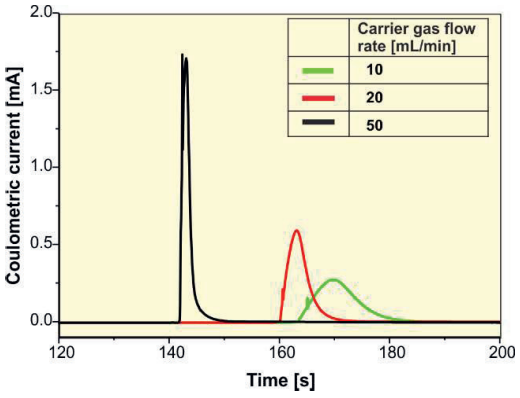


Fig. 10: Sections of potentiostatic scans at different flow rates, showing H_2 peaks at one injection step, polarization voltage = -300 mV, detector temperature = 600 °C, c(H_2) = 300 vol.-ppm.

3.4.2 Heating/cooling module (HCM)

The temperature courses illustrated in Fig. 11 were obtained by heating and cooling the prototype shown in Fig. 3. The measured temperature gradients during heating (10°K/min) and cooling (~55°K/min) already meet the desired values for final operation. Furthermore, the temperature gradient can be stabilized if the heating power increases linearly with temperature. As expected, the heating efficiency can be increased significantly by installing a swing-type flap at the outlet of the cylindrical Al-carrier (acting as a heat trap), which is closed during heating and opened by airflow during cooling.

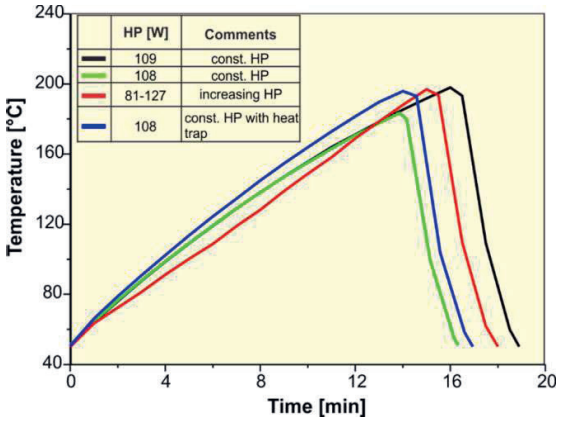


Fig 11: Heating/cooling ramps measured at the prototype given in Fig. 3 at different operating conditions as listed in the inset table, HP = heating power.

4 Conclusions

In this work, the response of a commercial solid electrolyte coulometric detector (SEC) was investigated at different operating conditions like temperature, oxygen partial pressure, injected H_2 concentration and carrier gas flow. The observed results enable method and component optimization for improvement of selectivity, long-term stability and lower limit of detection.

At detector temperatures above 700 °C a pronounced H_2 peak tailing was observed, which indicates temporary storage and slow release of H_2. It is assumed that this storage occurs in a porous ceramic layer between the electrode and the electrolyte. This result is supported by dynamic measurements with cyclic voltammetry at a similar detector.

This tailing can be prevented almost completely by diminishing the volume of the porous layer or by decreasing the detector temperature. The peak areas measured at different H_2 concentrations and detector temperatures demonstrated a close agreement with Faraday law at 600 °C and a peak area decrease of about 4 % at 750 °C. The separation reproducibility could be quantified with the standard deviation of retention time = 1.5 s. The repeatability of the whole system results in a standard deviation of the measured peak areas of < 2 % for the H_2 concentration 180 vol.-ppm. This outstanding detector stability enables its calibration free operation over long periods. The constancy of peak areas in the investigated carrier gas flow range proves the completeness of H_2-conversion during the passage through the electrolysis cell. The results offer a test reference for the various construction tradeoffs between electrode geometry (length and cross section area) and chosen flow rates.

The limit of detection of the system was quantified to range at around 200 vol-ppb H_2. A further diminishment of this parameter is expected with a newly developed miniaturized detector.

A miniaturized sample injection system was developed and manufactured as a 3D-printed manifold made of stainless steel. It is equipped with side-mounted miniaturized solenoid valves. The manifold tests show reproducible and stable H_2 peaks and drift-free baseline current. Main advantages of this, more than 40 % miniaturized, injection system are an increased flexibility enabling intelligent injection procedures and significant cost reduction.

A cylindrical carrier was developed and tested for resistive heating and forced convection cooling of separation columns. It provides more than 50 % size reduction and meets the requirements of the chromatographic method.

5 Acknowledgements

The presented work was carried out within a research project funded by the Federal Ministry of Education and Research under the support code 03ZZ0724A. The authors are responsible for the content and acknowledge the funding. Furthermore, the authors would like to thank CS Chromatographie Service GmbH, Langerwehe, and Zirox- Sensoren & Elektronik GmbH, Greifswald, for professional and device related support. The material related investigations and analysis done by F. Altmann and his team at Fraunhofer Institute for Microstructure of Materials and Systems IMWS in Halle is gratefully acknowledged.

6 References

[1] R. Ramachandran, R.K. Menon, *International Journal of Hydrogen Energy* 23, 593-598 (1998); doi: 10.1016/S0360-3199(97)00112-2

[2] T. Sinigaglia, F. Lewiski, M. Eduardo, S. Martins, J. Cezar, M. Siluk, *International Journal of Hydrogen Energy* 42, 24597-24611(2017); doi: 10.1016/j.ijhydene.2017.08.063

[3] A. Maroufmashat ,M. Fowler, *Energies* 10, 1089 (2017); doi: 10.3390/en10081089

[4] D. Ferrero, M. Gamba, A. Lanzini, M. Santarelli, *Energy Procedia* 101, 50-57(2016); doi: 10.1016/j.egypro.2016.11.007
[5] J. M. Ogden, *Annual. Review of Energy and the Environment.* 24, 227–279 (1999); doi: 10.1146/annurev.energy.24.1.227
[6] ISO 26142, Hydrogen detector apparatus-stationary applications, 2010.
[7] W. J. Buttner, M. B. Post, R. Burgess, C. Rivkin, *International Journal of Hydrogen Energy* 36, Issue 3, 2462-2470 (2011); doi: 10.1016/j.ijhydene.2010.04.176
[8] R. O. Cebolla, E. Weidner, W. Buttner, C. Bonato, K. Hartmann, K. Schmidt, *International Journal of Hydrogen Energy* 43, 21149-21160 (2018) doi: 10.1016/j.ijhydene.2018.09.107
[9] J. Zosel, M. Schelter, V. Vashook, U. Guth, *Tagungsband (2012)*, 324–331, doi: 10.5162/sensoren2012/3.2.4
[10] M. Schelter, J. Zosel, W. Oelßner, U. Guth, M. Mertig, *Sensors and Actuators B* 187, 209–214 (2013); doi: 10.1016/j.snb.2012.10.111
[11] K. Teske, P. Popp, J. Baumbach, *Journal of Chromatography* 360, 417–420(1986); doi: 10.1016/S0021-9673(00)91690-X
[12] M. Schelter, J. Zosel, V. Vashook, U. Guth, M. Mertig, *Solid State Ionics* 288, 266-270 (2016); doi:10.1016/j.ssi.2016.01.020.
[13] https://srigc.com/cn/downloads/29/MG1man.pdf
[14] A. Ruchets, N. Donker, D. Schönauer-Kamin, R. Moos, J. Zosel, U. Guth, M. Mertig, *Sensors and Actuators B: Chemical* 290, 53-58 (2019); doi: 10.1016/j.snb.2019.03.063

DOI 10.5162/sensoren2019/3.4.2

Combination of laser topography measurements with 3D magnetic field mapping for permanent magnets, magnet assemblies and PM rotors

Dr. Koen Vervaeke, Magcam NV, Romeinse straat 18, B-3001 Leuven, Belgium

Abstract

As a further development of Magcam's magnetic field camera scanners, we report on a combined stand-alone scanner for very fast measurements of the magnetic field distribution of flat magnets, magnet assemblies and permanent magnet rotors, together with a measurement of the surface topography using a scanned laser triangulation sensor. This allows to correlate features in the measured magnetic field distribution with surface topography features.
The integrated laser distance sensor allows to accurately measure the height, lateral dimensions and full surface map of magnets and magnet assemblies under test. These dimensional measurements complement the high accuracy magnetic field maps measured by the standard integrated 3D magnetic field camera, used for very fast and accurate magnetic field mapping of permanent magnets and magnet assemblies. Both maps can be overlaid, resulting in a correlation between geometric and magnetic data.
Applications of such correlation include magnet height variations as well as lateral positioning errors in magnetic assemblies, surface shape effects, rotor eccentricity in PM rotors etc.
The scanner is controlled by Magcam's powerful MagScope software, which automatically steers the motorized slides, performs the measurements and analyzes the resulting data using powerful algorithms.
The new Combi Scanner is an allround magnetic and dimensional measurement and analysis solution for inspection of permanent magnets, magnet assemblies and PM rotors for use in R&D labs, QC divisions and production lines and for various applications, such as automotive sensor systems, electric motors, medical devices, consumer electronics and magnet production.

1 Introduction

Fast and accurate quality inspection of permanent magnets is increasingly important in development, production and quality control of ever more accurate and reliable sensors as widely used in automotive applications, more efficient electric drives for e.g. electric vehicles, new medical devices, consumer electronics and many other magnet applications. Furthermore, in many high-end applications a 100% magnet quality control is desired, but not feasible with classical magnetic measurement equipment.
In the present paper, first the main components of the Magcam Combi Scanner are described. Subsequently examples of typical measurement results are presented.

1.1 Magcam 3D Magnetic Field Camera

Magcam developed the magnetic field camera technology for inspection of permanent magnets, based on a high resolution and high speed quantitative 2D mapping of the magnetic field distribution of a magnet, using a patented sensor chip with an integrated 2D array of microscopic Hall sensors [1-7]. A further development of this technology was a three-axis magnetic field camera ('MiniCube3D', see **Figure 1**), which makes the three components of the magnetic field distribution available in a 2D area of 12.7mm x 12.7mm, with 0.1mm pixel resolution, equaling 128x128 = 16384 measurement points [8-10]. A full resolution frame is recorded in less than one second. Small magnets with sizes up to about 10mmx10mm can be measured by simply placing the magnet onto the



magnetic field camera surface.

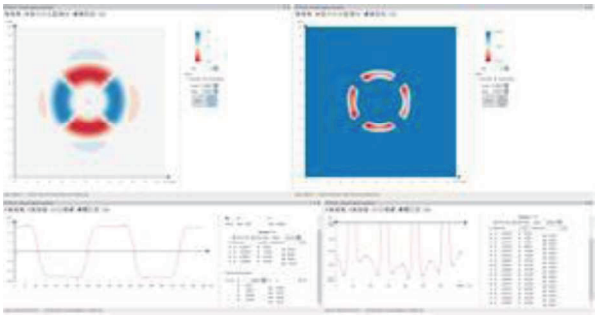


Figure 1: Magcam MiniCube3D magnetic field camera (top) and magnetic field map of a small ring magnet with analysis (bottom)

1.2 Magcam Combi Scanner

For larger magnets, magnet assemblies and permanent magnet rotors, Magcam offers its Portal and Rotor Scanners [1,5-7,10], as well as the newest Combi Scanner, combining the Portal and Rotor scanner functionalities into a single Combi Scanner machine.



Figure 2: Magcam Combi Scanner

1.2.1 Portal Scan mode

To map the 3D magnetic field distribution in a large area and/or a volume, the MiniCube3D magnetic field camera is mounted onto a XYZ scan stage to sequentially measure multiple small-scale magnetic field maps which are subsequently stitched together to obtain a large area image. This configuration is shown in **Figure 3**.

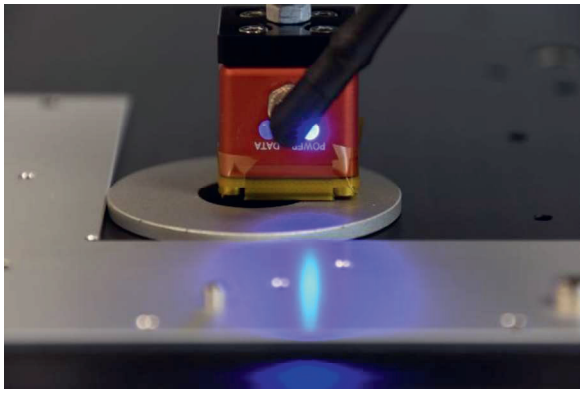


Figure 3: Measuring a flat magnet in Portal Scan mode

1.2.2 Rotor Scan mode

For measuring permanent magnet rotors, the MiniCube3D magnetic field camera is rotated over 90° such that it faces sideways, and a motorized rotary stage is mounted on the sample plate. The rotor under test is mounted vertically in a high accuracy rotary chuck that is mounted on the motorized rotation table, as shown in **Figure 4**.

Figure 4: Measuring a permanent magnet rotor in Rotor Scan mode



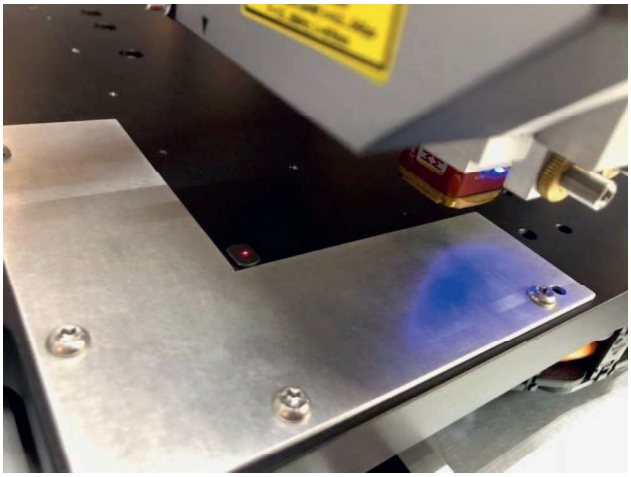
The magnetic field camera is brought in close vicinity to the rotor surface, which is fully automatic thanks to an accurate calibration of the mechanics. During one rotation of the rotor, the magnetic field camera measures a band of 12.7mm axial length (this is the length of the MiniCube sensor array). By stitching together multiple bands, the full rotor surface is measured with high axial resolution (0.1mm) and arbitrary angular resolution (typically ranging from 0.1° to 1°).

1.2.3 Laser distance sensor

A recently developed add-on to the Magcam Portal Scanner and Combi Scanner is a high accuracy Laser Distance Sensor, which allows to measure the height, lateral dimensions and full surface topography map of magnets and magnet assemblies with micrometer accuracy. The laser distance sensor setup is shown in **Figure 5**. These dimensional measurements complement the high accuracy magnetic field maps measured by the standard integrated 3D magnetic field camera, used for very fast and accurate magnetic field maps. Both maps can be overlaid, resulting in a correlation between geometric and magnetic data.

The Laser Distance Sensor can be retrofitted on existing PLC-controlled Magcam Portal Scanner setups and takes advantage of the high accuracy mechanics of the scanner,

converting the scanner in a highly performant hybrid dimensional and magnetic measurement system for in-



spection of permanent magnets and magnet assemblies.
Figure 5: Laser distance sensor for measuring magnet height and sample topography

2 Measurement results

2.1 Flat magnet

In Portal Scan mode, the Combi Scanner is used to measure flat magnets and magnet assemblies. As an example, the 3D magnetic field distribution of a 2-pole axially magnetized rotary encoder magnet (shown in **Figure 5**) is measured at a height of 0.5mm above the magnet surface. This measurement height is accurately determined by the laser distance sensor first measuring the magnet surface level, combined with accurately calibrated mechanics. The result is shown in , presenting the Bx, By, Bz component distributions in the measurement plane (top color graphs), as well as the full B field, Bxy and azimuthal angle distributions (bottom graphs), which are calculated from the former.

Figure 6, presenting the Bx, By, Bz component distributions in the measurement plane (top color graphs), as well as the full B field, Bxy and azimuthal angle distributions (bottom graphs), which are calculated from the former.

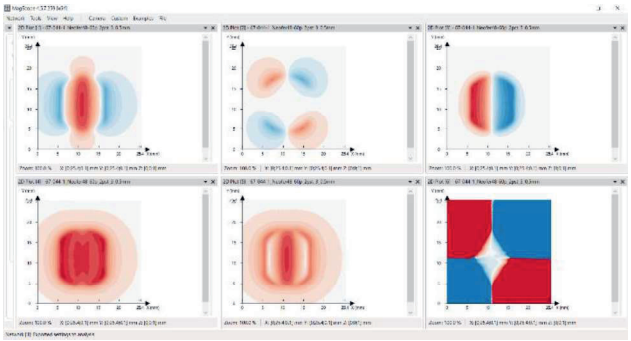
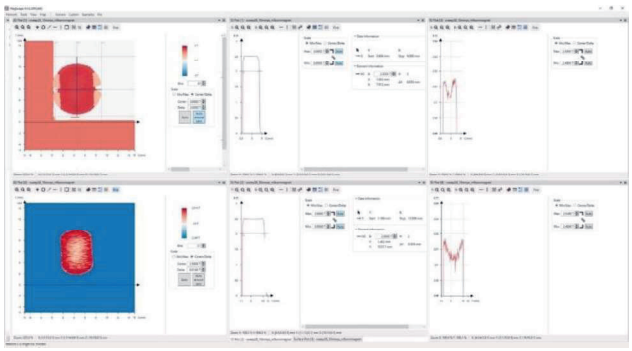


Figure 6: Measured magnetic field distribution of 2-pole rotary encoder magnet. Top: Bx, By, Bz; Bottom: B, Bxy, azimuth angle distributions.

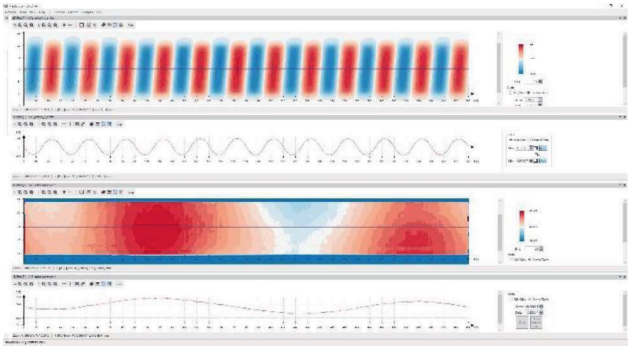
2.2 Laser topography scan on flat magnet

The laser distance sensor allows to measure a full topography map of the magnet under test. The result on the



magnet of **Figure 5** is shown in **Figure 7**. The measurement reveals that the magnet is lower in the center and higher towards the edges.
This measured topography can be accurately analyzed using the MagScope software and can be compared to the magnetic field maps of **Figure 6**.
Figure 7: Magnet surface topography measured by laser sensor scanning

2.3 Permanent magnet rotor



In Rotor Scan mode the 3D magnetic field distribution of permanent magnet rotors is measured as shown in **Figure 4**, resulting in the magnetic field distribution presented in **Figure 8** (top color and line plots, only the radial component is shown).
Figure 8: Radial magnetic field distribution of a permanent magnet rotor (top) and laser topography map of the rotor mantle (bottom)

The magnetic field distribution is shown in cylindrical coordinates and can be analyzed in multiple ways, including automatic finding of zero-crossings and local extremes, pole angle count and analysis and FFT analysis.

2.4 Laser topography scan on PM rotor

In Rotor Scan mode the laser distance sensor can be used to measure a full topography map of the rotor under test, as shown in **Figure 9**. The result of the scan is shown in **Figure 8** (bottom color and line plots), showing the measured radius of the rotor as a function of axial and angular positions. The plot shows a second harmonic, indicating a slightly elliptic shape of the rotor. This kind of observations can subsequently be combined with the measured magnetic field distribution in order to analyze

the source of magnetic features, perform mechanical quality control etc.

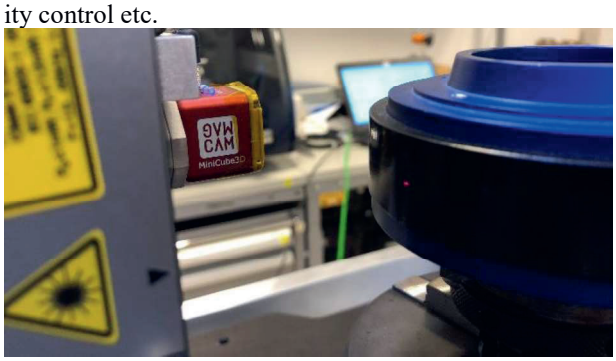


Figure 9: Laser topography measurement configuration of the Combi Scanner in Rotor Mode

3 Data analysis

All the measurement data presented in the preceding section can be thoroughly analyzed using Magcam's MagScope measurement and analysis software. Examples of data analysis are available elsewhere [1]. Features include:

- Magnetization vector size and angle deviation analysis
- Interpolation of magnetic field maps for micrometer-resolution analysis
- 2D color plots of magnetic field maps
- 2D and 1D (cross-section) region selection in Cartesian and polar coordinates for cut-out and analysis
- Line plots of cross sections in Cartesian or polar coordinates
- Automatic multipole segment detection and measurement of pole sizes/angles
- Statistical analysis on images and cross sections
- Accurate distance/radius/angle measurements on image features (e.g. pole segment sizes)
- Advanced image processing
- Noise reduction and cancellation
- Fourier analysis of cylindrical data (e.g. rotor data)
- Crack detection algorithms
- Integrated Python Scripting Module for e.g. automated pass/fail analysis, integration in production lines and data logging

4 Conclusion

We reported on a combined stand-alone scanner for very fast measurements of the magnetic field distribution of both flat magnets or magnet assemblies and permanent magnet rotors. This new 'Combi' scanner is an integration of Magcam's existing Portal and Rotor scanners into a high precision, PLC controlled 4-axis scanner.

We demonstrated the main features of the Combi scanner, including the Portal Scan mode, Rotor Scan mode and the integrated laser distance sensor for accurate dimensional measurements of samples and a design safety housing.

This laser distance sensor allows to accurately measure the height, lateral dimensions and full surface map of magnets and magnet assemblies under test, as well as for permanent magnet rotors, complementing the high accuracy magnetic field maps measured by the standard integrated 3D magnetic field camera.

Both the magnetic field maps and topography maps can be analyzed using the versatile MagScope software, resulting in a thorough and fast analysis of the magnet, assembly or rotor. Many quantitative magnet properties are extracted from the measurement data, making the Magcam system suitable for both R&D and production of sensor magnets, PM rotors, consumer electronics, medical devices, magnet production and others.

5 Literature

[1] www.magcam.com

[2] K. Vervaeke, Conference Proceedings of 15th International Conference on Sensors and Measurement Technology SENSOR 2011, 7.-9.6.2011, Nürnberg, P5.1

[3] K. Vervaeke, Conference Proceedings of 1st International Electric Drives Production Conference (E|DPC 2011), 28-29 September 2011, Nürnberg, ISBN 978-1-4577-1371-2, pp. 172

[4] K. Vervaeke, Conference Proceedings of 16. GMA/ITG Fachtagung Sensoren Und Messsysteme 2012, 22. und 23. Mai 2012, Nürnberg, 'Magnetic field camera for fast - high resolution inline magnet inspection', 2.3.1

[5] K. Vervaeke, Conference Proceedings of 2nd International Electric Drives Production Conference (E|DPC 2012), 15-18 October 2012, Nürnberg, ISBN 978-1-4673-3006-0, pp. 186

[6] K. Vervaeke, Conference Proceedings of 16th International Conference on Sensors and Measurement Technology, 14-16 May 2013, Nürnberg (SENSOR 2013), DOI 10.5162/sensor2013/A7.1

[7] K. Vervaeke, Conference Proceedings of 3rd International Electric Drives Production Conference (E|DPC 2013), 29-30 October 2013, Nürnberg, ISBN 978-1-4799-1102-8, pp. 38

[8] K Vervaeke, Conference Proceedings of 17. ITG / GMA-Fachtagung „Sensoren und Messsysteme", 3.-4. Juni 2014, Nürnberg, P5.2

[9] K. Vervaeke, "3-axis magnetic field camera for ultrafast and high-resolution inspection of permanent magnets" Conference Proceedings of 4th International Electric Drives Production Conference (E|DPC 2014), 30 September - 1 October 2014, Nürnberg, ISBN 978-1-4799-5010-2, pp. 164

[10] K. Vervaeke, "6D magnetic field distribution measurements of permanent magnets with magnetic field camera scanner", Proceedings of 'SENSOR 2015' Conference, 19-21 May 2015, Nürnberg

[11] K. Vervaeke, "Combined 3D magnetic field camera scanner for both flat permanent magnets and PM rotors", Conference Proceedings of 19. GMA/ITG-Fachtagung Sensoren und Messsysteme, 26-27 June 2018, Nürnberg

DOI 10.5162/sensoren2019/3.4.3

Development of a rotating-coil scanner for superconducting accelerator magnets

Piotr Rogacki[1,2], Lucio Fiscarelli[1], Stephan Russenschuck[1], Kay Hameyer[2]
[1]CERN, Switzerland
[2]Institute of Electrical Machines (IEM), RWTH Aachen University, Aachen, Germany
p.rogacki@cern.ch

Abstract

The superconducting magnets for High-Luminosity Large Hadron Collider (HL-LHC) upgrade project at CERN require magnetic measurements with high precision in field angle and multipole field errors. The paper describes the design, development, and metrological characterization of a new, rotating-coil magnetometer and a transport system to scan the magnetic field along the axis of the accelerator magnet.

Keywords: Magnets, Accelerators, PCB, Magnetometer, 3D Printing

Introduction

The High-Luminosity upgrade of the Large Hadron Collider (HL-LHC) [1] requires Nb_3Sn superconducting magnets with strict requirements for their field quality. Therefore, the magnets must be measured with highest possible precision of up to 1-ppm levels, both locally and averaged over their entire length of up to ten meters. To acquire all required quantities now, the superconducting quadrupole magnets have to be measured using both rotating-coil magnetometers [5] and stretched-wire systems [7]. It is therefore advantageous to develop a new magnetometer for the upcoming measurement campaign that allows us to measure all required quantities simultaneously.

Harmonic fields and rotating-coil magnetometers

A 2D magnetic field in a domain that is free of current and magnetized material (such as the bore of accelerator magnets) can be described as a complex valued harmonic field expansion [9]:

$$B_y + \mathrm{i}B_x = \sum_{n=1}^{\infty} (B_n + \mathrm{i}A_n) \left(\frac{x + \mathrm{i}y}{r_0} \right)^{n-1} , \quad (1)$$

where r_0 is the reference radius, and B_n and A_n are called the normal and skew harmonic coefficients, or in short, field multipoles. In practice, the n-th coefficient corresponds to a field generated by a magnet with n[1]pole pairs, for example, for a quadrupole magnet $n = 2$.

[1] This convention, established for the accelerator magnets, differs from the usual electrical machines naming convention in using n instead of p for the pole pair number.

In the case of accelerator magnets, the field along the magnet length can be integrated and shown to satisfy the two-dimensional Laplace equation. The same applies to the field integrated along the length of an induction coil, with the constraint that the ends of the coil are in the domain that is free of axial field components. Otherwise, the coil ends will pick up the field in the axial (z) direction, in which case the measurement results do not correspond to the 2D mathematical model but must be treated with the theory of pseudo-multipoles [2].

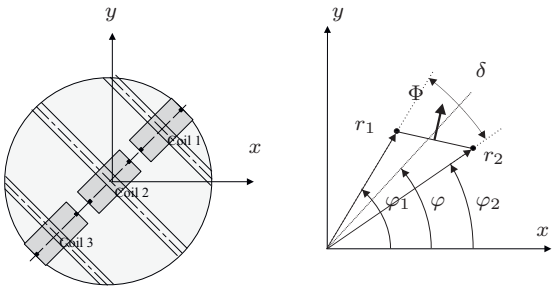


Fig. 1: Left: Radial (sagital) coil array. Right: Naming convention for angles and radii of a single wire loop.

The rotating-coil measurement principle is especially suited to characterize the fields described as in Eq. 1. This is due to the fast and robust analysis of the output signals.
According to Faraday's law:

$$U = -\frac{\mathrm{d}\Phi_B}{\mathrm{d}t} , \quad (2)$$

using the principal coil arrangement in the shaft and the geometric relations shown in Fig. 1, we

can calculate the magnetic flux intercepted by the induction coil.
Knowing that [9]

$$\int_{z_1}^{z_2} \boldsymbol{B}(\boldsymbol{z})\ \mathrm{d}\boldsymbol{z} = \int_{z_1}^{z_2} (B_y\ \mathrm{d}x - B_x\ \mathrm{d}y) + \mathrm{i}\int_{z_1}^{z_2} (B_y\ \mathrm{d}y + B_x\ \mathrm{d}x)\,, \quad (3)$$

the flux intercepted by the induction coil can be expressed as

$$\begin{aligned}\Phi &= N\ell\,\mathrm{Re}\left\{\int_{z_1}^{z_2} \boldsymbol{B}(\boldsymbol{z})\ \mathrm{d}\boldsymbol{z}\right\}\\ &= N\ell\,\mathrm{Re}\left\{\int_{z_1}^{z_2} \sum_{n=1}^{\infty}(B_n + \mathrm{i}A_n)\left(\frac{\boldsymbol{z}}{r_0}\right)^{n-1}\mathrm{d}\boldsymbol{z}\right\}\\ &= \mathrm{Re}\left\{\sum_{n=1}^{\infty}\frac{N\ell}{n\,r_0^{n-1}}(B_n + \mathrm{i}A_n)\,(\boldsymbol{z_2^n} - \boldsymbol{z_1^n})\right\}\\ &= \mathrm{Re}\left\{\sum_{n=1}^{\infty}\boldsymbol{C_n}\,\boldsymbol{S_n}\,e^{\mathrm{i}n\varphi}\right\}\,, \end{aligned} \quad (4)$$

where $\boldsymbol{C_n} := B_n + \mathrm{i}A_n$ (at the reference radius r_0) and the $\boldsymbol{S_n}$ are complex valued sensitivity factors given by

$$\begin{aligned}\boldsymbol{S_n}(r_0) &= S_n^{\mathrm{rad}} - \mathrm{i}\,S_n^{\mathrm{tan}}\\ &= \frac{N\ell}{n\,r_0^{n-1}}(\boldsymbol{z_2^n} - \boldsymbol{z_1^n})\\ &= \frac{N\ell}{n\,r_0^{n-1}}(r_2^n\,e^{\mathrm{i}n(\varphi_2-\varphi)} - r_1^n\,e^{\mathrm{i}n(\varphi_1-\varphi)})\,. \end{aligned} \quad (5)$$

Radial[2] coils, as sketched in Fig. 1 (left), intercept the azimuthal component of the magnetic flux density. For $\varphi_1 = \varphi_2 = \varphi$, inner radius r_1 and outer radius r_2 the S_n^{tan} are zero and we obtain:

$$\boldsymbol{S_n} = S_n^{\mathrm{rad}} = \frac{N\ell}{n\,r_0^{n-1}}\left(r_2^n - r_1^n\right)\,. \quad (6)$$

The S_n^{rad} are calculated using the geometry of the coils and their arrangement in the probe. In the most straightforward case the field multipoles can then be calculated using

$$B_n + \mathrm{i}\,A_n = \frac{\boldsymbol{\psi_n}}{\boldsymbol{K_n}}\,, \quad (7)$$

where $\boldsymbol{K_n} = \boldsymbol{S_n} r_0^{n-1}$ are the sensitivity factors defined to be independent of the reference radius and $\boldsymbol{\psi_n}$ are complex coefficients of a Fourier transform of the acquired integrated voltage.
The harmonic content of the field is used as one of the quality measures. Generally, a field of an acceptable quality is defined to have at most a few hundred ppm of multipole coefficients.
In realistic measurement, affected by mechanical imperfections and signal noise, some additional steps are required to achieve the specified accuracy. First of all, for the integration of voltage it is necessary to keep exact track of the timing. This is complicated due to the instabilities in shaft rotation caused by the motor and bearings. They introduce vibrations, which can influence the measured field harmonics. The established solution to that problem is the use of digital integrators triggered by an angular encoder coupled with the shaft. The voltage between two triggers is integrated, yielding a measure of the flux linked between the two positions. Mathematically this corresponds to a re-parametrization of the signal with respect to the rotation angle instead of time.
Multiple coils in the probe can be connected together in an appropriate configuration to provide compensation of the main field components [10]; in our case the quadrupole and dipole field components. This allows to amplify the compensated signal, thus increasing the SNR for the higher-order field harmonics. Moreover, compensation may reduce the impact of mechanical vibrations and displacements, if all connected coils move together.
Finally, to correct the misalignment of the measurement coils with the magnetic axis, a so called feed-down correction is applied [9]. In the case of a centered quadrupole measurement, the dipole component should vanish, except if the coil is decentered with respect to the magnetic axis. This misalignment can be calculated and corrected by translating the reference frame of the results to where the dipole harmonic is minimized. This yields the centered field harmonic coefficients as well as the position of the magnetic field center in the local sensor coordinates.

Requirements

The requirements for the measurement system are based on the design of the HL-LHC magnets and the accuracy specifications laid out in table 1.

Tab. 1: The target accuracy for LHC insertion region quadrupole measurements [1].

Parameter	Unit	Accuracy
Main field	[ppm]	100
Main field direction	[mrad]	0.1
Harmonics	[ppm]	1
Magnetic center	[mm]	0.15

The measurement system must be able to measure different quantities with highest achievable accuracy. All those quantities must be provided

[2] Or better sagital.

both locally and integrated along the entire length of the magnet (up to 10 m). They are important as they allow to intercept manufacturing errors at an earliest state in the production.
The two types of systems commonly used for magnetic measurements of long accelerator magnets are rotating-coil magnetometers and stretched-wire systems. A summary of the currently achievable measurement accuracy for those systems is given in table 2.

Tab. 2: Existing systems accuracy [3, 6, 7, 8]

Parameter	Accuracy	Units	Type	System
Main field	100 1000	[ppm]	Integral Local	Stretched wire Rotating coils
Magnetic center	0.05	[mm]	Integral Local	Stretched wire Rotating coils
Main field direction	0.1	[mrad]	Integral Local	Stretched wire Rotating coils
Harmonics	1	[ppm]	Integral Local	Rotating coils Rotating coils

The accuracy for the integral field-center and field-direction measurements has been obtained with a Single-Stretched Wire system as used for the LHC main quadrupoles [7]. These magnets posed similar challenges as the ones to be measured in the incoming HL-LHC upgrade project. The expected accuracy for the field harmonics measurement are based on the probes developed by Fermi National Accelerator Laboratory (FNAL) and Lawrence Berkeley National Laboratory (LBNL) for the LHC Accelerator Research Program (LARP) as described in [6]. These probes are designed for measuring a high-gradient quadrupole magnet and are based on a similar principle, with the use of induction coils produced in PCB technology. The accuracy of the other parameters retrieved by rotating-coil systems are based on the performance of the legacy QIMM and DIMM systems, used for LHC dipole and quadrupole measurements at ambient temperatures [8].
It can be observed that even though for most parameters a system exists that allows a measurement with the required accuracy, it is necessary to use at least two or three complementary systems to measure and derive all required data. Given the considerable time and effort for each measurement, a new system had to be developed, able to measure all aforementioned quantities simultaneously.
Additionally, as the new magnets will have multiple different lengths and apertures, the system needs to be easily adaptable. The size and radius of the induction coils have an influence on the amplitude of the acquired signal and thus on the signal-to-noise ratio (SNR). This is especially important for magnets such as the quadrupoles, sextupoles and higher orders.

The prototype magnetometer

Taking into account all the requirements as listed above, the necessity for the local and integral measurements of long magnets, and the required accuracy of the measurements, the best approach was to design a new rotating-coil scanner, able to provide all values locally, and after scanning the entire magnet bore, providing the integral values as well.
By comparing the specification with the designs of existing rotating-coil scanners, like the QIMM [8] or the *ferret* [6], we derived the main design concepts:

- Construction adaptable to magnet apertures ranging from 90 mm to 150 mm and above,
- 3D printing for complicated structural parts,
- built in encoder and level meter,
- PCB based induction coils,
- open end for mounting retro-reflectors for the laser tracker,
- and transport in the magnet bore with extension rods.

The placing of the encoder and the level meter close to the induction coils is necessary to precisely track their angular position, that in turn enables the accurate measurement of the angle of the magnetic field. In order to relate the magnetic-axis measurement (given in the probe coordinates) to the magnet reference, one end of the scanner must be open to leave free line-of-sight for the laser tracker. During the rotation, the tracker follows the retro-reflectors mounted on the shaft. By fitting a circle to the measured points, it is possible to find the mechanical rotation axis and indirectly the axis of the magnetic field. The longitudinal positioning is realized by the use of extension rods that transmit the rotation from a motor placed outside the magnet.

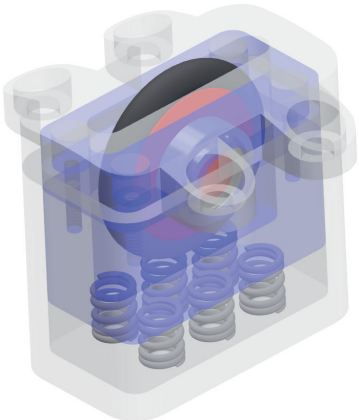


Fig. 2: Rendering of the stabilizing wheel.

The base mechanical structure of the probe was

fabricated using 3D printing technology due to its complicated shape and because the prototype version was designed to accommodate multiple solutions for testing. The two main features to be tested are the positioning of the level meter, as well as supporting the probe in the magnet aperture. In the first tested variant, the level meter was mounted to the base of the scanner, which was held in position during the measurement using the stabilizing wheels as shown in Fig. 2.
The main goal of the wheels is to facilitate longitudinal displacements, while retaining the angular position of the base during the coil rotation.

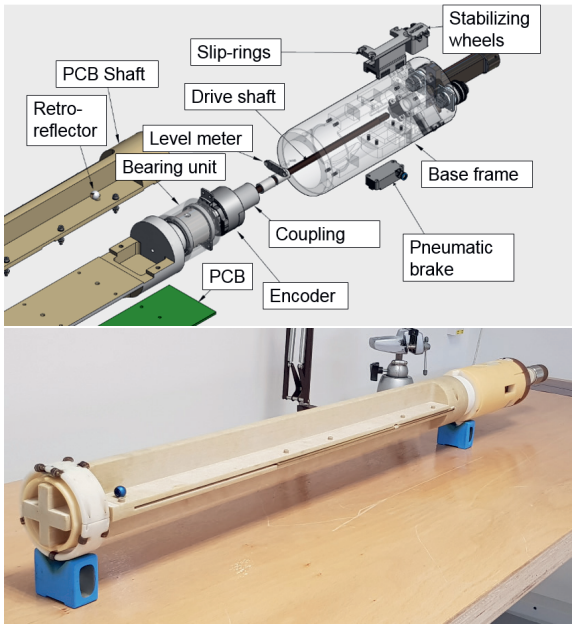


Fig. 3: First prototype measurement system. Top: exploded view of the probe base structure. Bottom: The picture of an assembled system.

Considering the high positioning accuracy required for the PCB coils, together with the size limitations for the 3D printer, we decided that the support shaft for the PCB board should be machined from a glass fiber composite. The support shaft has been designed to allow mounting differently sized PCBs without major effort, in particular without disassembling the base structure. Moreover, the part is relatively easy to machine and modify thanks to parametric design. Therefore, it goes in line with the adaptability requirement; a single probe can be used for measuring magnets of different bore radii. The final design of the first prototype is shown in Fig. 3.
The fabricated PCBs (see Fig. 4) are equipped with two sets of five induction coils. To balance the cost and complexity of manufacturing we decided to use 24 active layers in the board. The number of turns per layer has been established by taking into account the clearances of 125 µm between the tracks, and the copper track widths of 125 µm, to minimize the chances for short circuits or track breaks. Generally, we aimed at maximizing the number of turns, without reducing the average measurement radius of the coil below 35 mm, which is derived from the measurement specification. Finally, we were able to fit 11 turns per layer, which resulted in a total of 264 coil-winding turns.

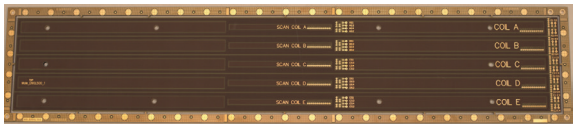


Fig. 4: The PCB board containing an array of 10 coils.

The 500 mm-long coils have been designed to cover the entire length of the magnet end region and provide a high SNR. The smaller, nested coils, that span 100 mm can be used in case a higher longitudinal resolution is required. In the case of induction-coil arrays manufactured in PCB technology, adding the nested coils comes at a very little price both in design time and production cost.
To study the uncertainty in the measurements due to manufacturing tolerances in the PCB production and layer stacking, the ”witness tracks” were inspected under a microscope; see Fig. 5.

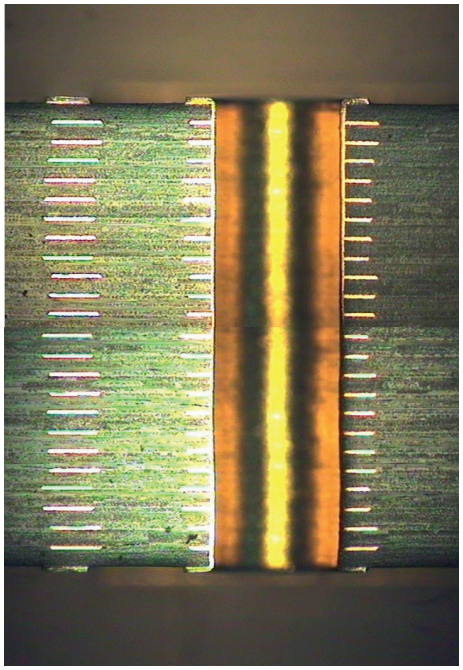


Fig. 5: The PCB stack cross-section under microscope.

By measuring the positions of the tracks, we established the accuracy of the layer positioning to be within ±30 µm. Using the CERN field computation program ROXIE [9], we calculated the sensitivity factors of the coil with random errors

of layer placement in this range of tolerance. The results of over 500 iterations show that for the coefficients of lower orders, the standard deviation in the sensitivity factors is negligible, while for the coefficients of higher orders (15-th and above) the standard deviation can be close to the percent range. This, in turn, would result in a relative error on the order of ppm in the measurement, as the higher order components in the field rarely exceed 100 ppm.

Coil calibration

The first step in the assembly process, was to measure the spanned surface of the PCB coils [4]. Even if one can rely on the optical measurement of the track position, the calibration step is necessary for verifying the correctness of the coil production and detecting the inter-turn shorts that cannot be detected by resistance measurements. The measured surface is used for computing the transfer function between the magnetic flux intercepted by the induction coil and the acquired voltage signal. The surface of the coil can be found by flipping it upside down in the uniform magnetic field of a reference magnet, while integrating the output voltage. The integration yields

$$-\int_0^t V_c \,\mathrm{d}t = \Phi - (-\Phi) = 2A\bar{B}\,, \qquad (8)$$

where $\bar{B}$ is the average magnetic flux density, and A is the surface spanned by the coil. The average surface measured for the three produced PCBs is 1.8729 m^2. These results have been corrected for the resistance of the coil (2.7 kΩ), as the input stage of the integrator has an impedance of 400 kΩ. The measured values show very high homogeneity between the coils on one PCB as well as between the boards; the differences are in the order of 10^{-4} m^2.

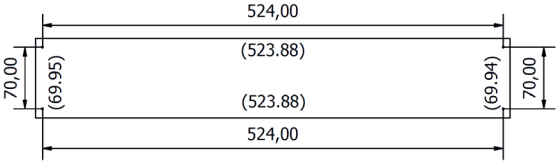


Fig. 6: Comparison of measured and design values for coil reference points. Measured values are in brackets.

We then compared the measurement results with the nominal surface from the PCB design and found a relative difference of approximately 0.1%. The reason for this discrepancy is the shrinking of the PCB board during the production. A precise geometric measurement of the reference points on the board (see Fig. 6) provided values to calculate the appropriate shrinking factor, assuming homogeneous shrinking of the whole board. The computed area, corrected by the shrinking factor equals 1.8727 m^2, which differs from the measured surfaces on the order of 10^{-4} m^2. These values are very close to the standard deviation of the calibration results that is on the order of 10^{-5}.

More PCB samples are needed to create proper statistics and identify the important factors for refining the design and fabrication process. The ultimate goal of these efforts is to be able to assume, with high certainty, that the final coil area corresponds to the design within the 10^{-5} relative tolerance. This would limit the role of the surface calibration to a mere electrical verification of the PCB coils. This is especially important for large PCBs that do not fit into the available reference magnets.

Measurement results

The first validation measurement campaign has been conducted in two reference quadrupoles used in the Magnetic Measurements (MM) section at CERN. Fig. 7 shows a measurement setup in one of the short magnets.

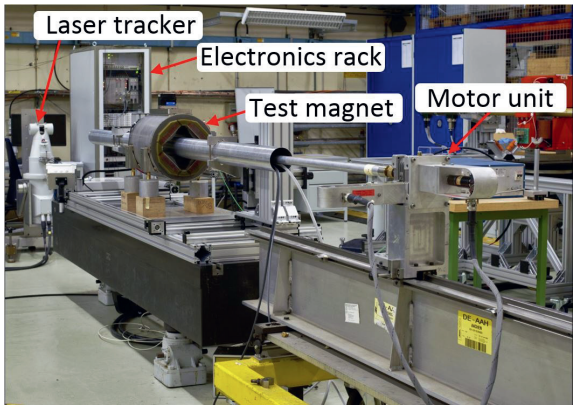


Fig. 7: Measurement setup with all system components. The probe is inside the tube that goes through the magnet.

Both magnets have been scanned along their entire length, using the laser tracker for longitudinal positioning as well as for tracking the rotation axis. In each position, the coils have been rotated 20 times in both directions to compensate for angular positioning errors and provide data for uncertainty analysis. The precision of local measurements shown in Table 3 and Fig. 8 is consistent among the different positions in both magnets.

Tab. 3: The precision of main field parameters measurement of a single position.

Parameter	Unit	Repeatability
Main field	[ppm]	60
Main field direction	[mrad]	0.1
Magnetic center	[mm]	0.1

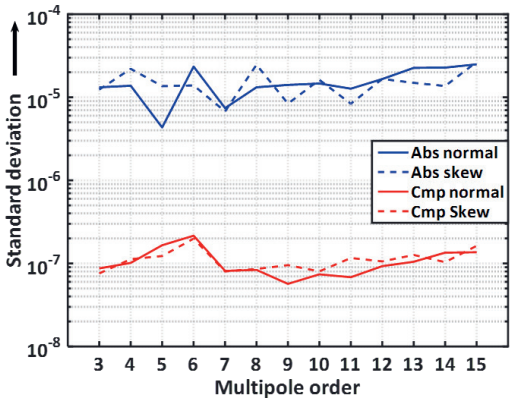


Fig. 8: The precision of field harmonics measurement of a single position.

The plot also shows the advantages of the compensation scheme for the the main field components. If the field harmonics are calculated using the signal from a single coil, the accuracy is at least 100 times worse. The quality of the coil production and positioning can be assessed by comparing the field strength measured by a single coil and the coils connected in a compensation scheme. The resulting proportional factor is commonly called the compensation (or bucking) ratio. As seen from the scale differences in Fig. 9 the achieved bucking ratio for dipole and quadrupole components are on the order of 1000.

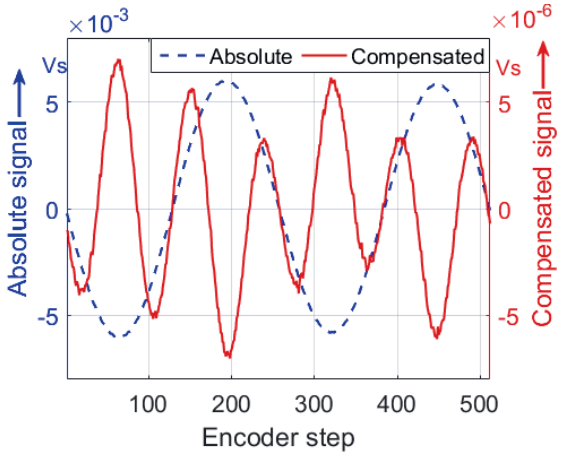


Fig. 9: The acquired integrated voltage between steps of the encoder over one rotation.

The integral values are computed from the local measurement results at all positions along the magnet. The local values have to be multiplied by the induction-coil length and then added together. In this case, two additional factors play an important role: the knowledge of the exact length of the coil and the precise longitudinal positioning, especially in the end regions of the magnet. The measurements have been conducted with both long and short coils several times to evaluate the repeatability of the system. The results are shown in Fig. 10 and Fig. 11.

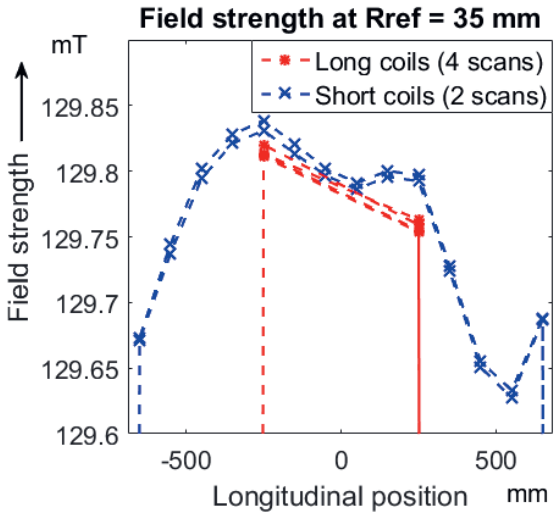


Fig. 10: Results of repeated scans with both long and short coils. External positions are removed from the plot to emphasize the details in the straight field region.

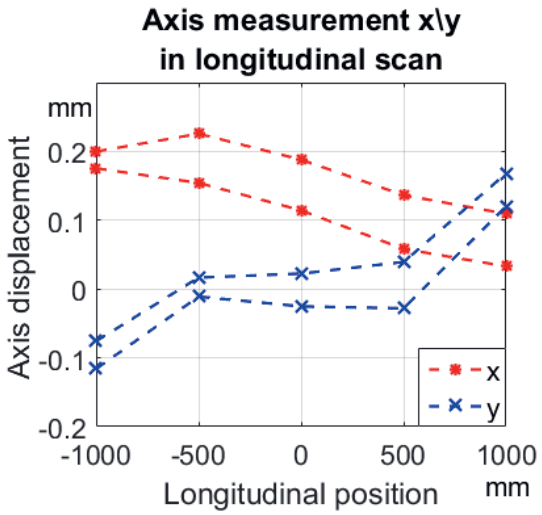


Fig. 11: Axis measurement results of two scans with long coils. The plot is centered at the average axis measured by a stretched wire system.

During the scan we discovered that the system mechanics is not precise enough, which results in movements of the base support during the rotation. Due to the compensation, the effect on the field strength and field harmonics measurements is negligible, but it has a noticeable influence on the axis and, especially, on the angle measurement. Since the integral values depend on the longitudinal positioning, the movement of the probe reduces the accuracy in their measurement as well.

The scan results have been compared to a stretched wire measurement that is the reference for integral values. The integrated gradient is in a good agreement with wire measurement; the

accuracy is at the level of 300 ppm, which is a promising result. The axis is within 0.15 mm from the wire measurement. This difference between the results of two systems needs to be further investigated. Given the high repeatability of local measurements, it should still be possible to improve these values. In order to characterize the limits of the system performance, these results have been obtained in controlled and favorable environmental conditions. The reference quadrupole has a relatively high and pure field of approximately 130 mT at the reference radius of 35 mm.

After the first test campaign, we substituted the stabilizing wheels with a pneumatic brake and placed the level meter directly on the encoder. In this way, and by leveling the probe before the measurement, we can rely on the more linear range of the tilt sensor that moves with the encoder. This modification resulted in more stable and precise angle measurements, even without changing the mechanical structure. The accuracy of the field angle measurement has been established to be better than 0.1 mrad.

Conclusions

A new rotating coil magnetometer has been developed and characterized. The results of the measurements already indicate that its performance is at least as good as other existing systems, while offering the functionality of those systems combined. A prototype of a long insertion-region quadrupole for HL-LHC has been measured and the results have been used as feedback to the magnet development.

The adaptable design of the system and the use of 3D printing allowed us to easily test multiple configurations of the mechanical structure and induction coil setup. Some of the encountered mechanical issues have been addressed. While the measurement of the field strength and field harmonics is only slightly affected by mechanical imperfections (due to the compensation scheme), the measurements of the angle and magnetic axis require much higher mechanical precision.

Ultimately, we aim to substitute the measurements with the stretched-wire system with the integrated measurements of the rotating coil scanner, substantially reducing the time and effort necessary for the measurements.

Acknowledgements

The authors would like to thank the Magnetic Measurements section laboratory and workshop staff for their help in assembling and testing the system. We would also like to thank D. Giloteaux, O. Dunkel and J. G. Perez for helpful suggestions and discussions. This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research (grant no. 05E15CHA)

References

[1] G. Apollinari et al. *CERN Yellow Reports: Monographs, Vol 4 (2017): High-Luminosity Large Hadron Collider (HL-LHC) Technical Design Report V. 0.1*. 2017. DOI: 10.23731/CYRM-2017-004.

[2] Pasquale Arpaia, Gianni Caiafa, and Stephan Russenschuck. “A Rotating-Coil Magnetometer for Scanning Transversal Field Harmonics in Accelerator Magnets”. In: *Scientific reports* 9.1 (2019), p. 1491. DOI: 10.1038/s41598-018-37371-3.

[3] L. Bottura, M. Buzio, S. Pauletta, and N. Smirnov. “Measurement of magnetic axis in accelerator magnets: critical comparison of methods and instruments”. In: *Proceedings of the 23rd IEEE Instrumentation and Measurement Technology Conference, 2006, IMTC 2006*. IEEE Operations Center, 2006, pp. 765–770. DOI: 10.1109/IMTC.2006.328154.

[4] M. Buzio. “Fabrication and calibration of search coils”. In: *CAS - CERN Accelerator School: Specialised course on Magnets*. Ed. by Daniel Brandt. CERN, 2009, pp. 387–421. DOI: 10.5170/CERN-2010-004.

[5] W. G. Davies. “The theory of the measurement of magnetic multipole fields with rotating coil magnetometers”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 311.3 (1992), pp. 399–436. DOI: 10.1016/0168-9002(92)90637-J.

[6] J. DiMarco et al. “Application of PCB and FDM Technologies to Magnetic Measurement Probe System Development”. In: *IEEE Transactions on Applied Superconductivity* 23.3 (2013), p. 9000505. DOI: 10.1109/TASC.2012.2236596.

[7] J. DiMarco et al. “Field alignment of quadrupole magnets for the LHC interaction regions”. In: *IEEE Transactions on Appiled Superconductivity* 10.1 (2000), pp. 127–130. DOI: 10.1109/77.828192.

[8] J. G. Perez, J. Billan, M. Buzio, P. Galbraith, D. GiloteauX, and V. Remondino. “Performance of the Room Temperature Systems for Magnetic Field Measurements of the LHC Superconducting Magnets”. In: *IEEE Transactions on Applied Superconductivity* 16.2 (2006), pp. 269–272. DOI: 10.1109/TASC.2006.871221.

[9] Stephan Russenschuck. *Field computation for accelerator magnets: Analytical and numerical methods for electromagnetic design and optimization*. Weinheim and Chichester, 2010. DOI: 10.1002/9783527635467.

[10] P. Schmüser. “Magnetic measurements of superconducting magnets and analysis of systematic errors”. In: *CAS - CERN Accelerator School*. Ed. by Stuart Turner. CERN, 1992, pp. 240–273. DOI: 10.5170/CERN-1992-005.240.

DOI 10.5162/sensoren2019/3.4.4

Tunnelmagnetoresistive Sensoren für Weg-, Winkel- und Positionsmessung in der elektrischen Antriebstechnik

Dr. Rolf Slatter
Sensitec GmbH, Lahnau, Deutschland

Zusammenfassung

Längen- und Winkelmesssysteme auf magnetischer Basis erleben in den letzten Jahren ein rasantes Wachstum. Allen voran finden Sensoren auf Basis des magnetoresistiven (MR) Effekts zunehmend Anwendung in der Antriebstechnik, sowohl im industriellen als auch im automotiven Bereich. MR-Sensoren werden nicht nur bei der Regelung von mechatronischen, sondern auch bei pneumatischen und hydraulischen Antrieben eingesetzt. Es können inkrementelle als auch absolute Messungen in einer Vielzahl von unterschiedlichen Konfigurationen durchgeführt werden. Neue Sensoren auf Basis des tunnelmagnetoresistiven (TMR) Effektes eröffnen neue Möglichkeiten aufgrund ihrer sehr geringen Abmessungen, ihrem exzellenten Temperaturverhalten und dem niedrigen Energiebedarf. Damit bilden sie eine wichtige Ergänzung für die bisher eingesetzten AMR- und GMR-Sensoren in antriebstechnischen Anwendungen.

Keywords: Magnetische Sensoren, Winkelmessung, Wegmessung, Positionsmessung, magnetoresistive Sensoren

1. Einleitung

Der Bedarf an mechatronischen Antrieben steigt in mehreren Industriebereichen. Nicht nur bei der Industrieautomatisierung, sondern auch in der Fahrzeugtechnik und der Luftfahrt nimmt die Anzahl der eingesetzten Antriebe stetig zu. „Dezentralisierte Antriebstechnik“, „elektrifizierte Antriebstrang“ oder „more electric aircraft“ sind nur einige der Begriffe, die diesen Trend in den verschiedenen technologischen Bereichen beschreiben. Diese Entwicklung stellt neue Anforderungen an die eingesetzte Sensortechnologie für die Messung von linearen und rotatorischen Bewegungen. Abb. 1 zeigt die Schematik eines elektrischen Antriebes, typischerweise bestehend aus einem Umrichter, Servomotor, Drehzahl- und Positionssensoren (Encoder) sowie mechanischer Maschinenelemente. Aktuell entstehen neue Bedürfnisse, die die technischen und wirtschaftlichen Anforderungen, die sowohl an Antriebe als auch an Sensoren gestellt werden, neu definieren. Die hohe Leistung und Flexibilität von magnetoresistiven (MR) Sensoren spielt eine zunehmend wichtige Rolle bei der Unterstützung von Maschinen- und Antriebsentwicklern, um diesen neuen Bedürfnissen gerecht zu werden. Diese kompakten, präzisen und robusten Sensoren bieten Eigenschaften, die von optischen, kapazitativen oder induktiven Sensoren nicht geboten werden.

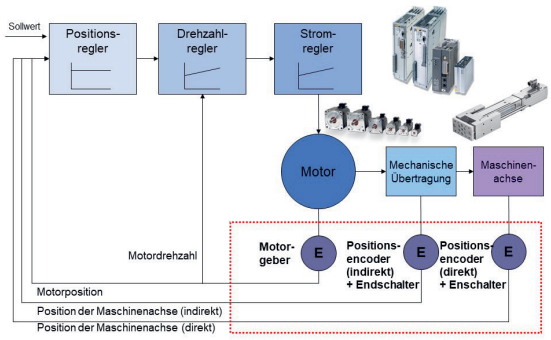


Abb. 1: Schematik eines elekromechanischen Antriebs (Bildquellen: LTi Motion GmbH, Festo AG)

2. Entwicklungstrends in der elektrischen Antriebstechnik

Der Markt für Antriebstechnik wird durch mehrere aktuelle Trends bestimmt, die neue Anforderungen sowohl an Antriebe als auch Sensoren stellen:

a) Höhere Genauigkeit
Aus der zunehmenden Bedeutung von Mikrosystemtechnik sowie Nanotechnologie ergeben sich neue Ziele bezüglich der Genauigkeit von Robotern, Werkzeugmaschinen und anderen Formen der Industrieautomatisierung. Es steigen nicht nur die Anforderungen hinsichtlich Positioniergenauigkeit, sondern auch die An-

forderungen an eine höhere Auflösung zwecks höherer Bahngenauigkeit. Hinzu kommt, dass die gewünschte Genauigkeit über die Lebensdauer der Maschine unter unterschiedlichen Betriebsbedingungen konstant bleiben soll. Dies führt wiederum zum verstärkten Einsatz von Direktmesssystemen.

b) Höhere Leistungsdichte
Maschinen erhalten zunehmend kompaktere Bauformen, um den verfügbaren Bauraum optimal zu nutzen und auch um die Energieeffizienz zu steigern. Dadurch steigt die Leistungsdichte von Antrieben stetig an, was sich in höheren Abtriebsdrehmomenten sowie höheren Geschwindigkeiten ohne einen proportionalen Anstieg der physikalischen Größe niederschlägt.

c) Dezentralisierte Antriebstechnik
Die verbesserte Rechenkapazität von Mikro-Controllern führt zu einer Verlagerung von Mess- und Regelungsaufgaben bis auf die Ebene des einzelnen Antriebs herunter. Dezentralisierte Antriebe, bestehend aus Servomotoren mit integrierter Elektronik, bedürfen weniger Verkabelung, geringeren Montageaufwand und entlasten höher geordnete Steuerungen. Damit lassen sich die Kosten von Maschinen, die über viele Antriebe verfügen, drastisch senken.

d) Zunehmender Einsatz von geschlossenen Regelkreisen
Elektromechanische Antriebe werden zunehmend in Anwendungen eingesetzt, die früher ausschließlich durch hydraulische oder pneumatische Antriebe besetzt waren. Um diesem Trend entgegenzuwirken, setzen die Hersteller der letztgenannten immer häufiger geschlossenen Regelkreise ein, um die Genauigkeit und Dynamik ihrer Antriebe zu steigern.

e) Immer anspruchsvollere Umgebungsbedingungen
Immer mehr Anwendungen finden unter sehr niedrigen oder sehr hohen Umgebungstemperaturen statt. Es gibt bekannte Beispiele in der Luft- und Raumfahrt, aber auch im industriellen Bereich steigt die Zahl der Anwendungen, die unter extremen Temperatur- und Umgebungsbedingungen funktionieren müssen.

f) Energieeffizienz
Der Anzahl batteriebetriebenen Anwendungen steigt stetig. Sensoren mit sehr niedrigen Energiebedarf ermöglichen eine längere Batterie-Lebensdauer bzw. den Einsatz kleinerer Batterien in mobilen Anwendungen.

g) Schnelleres „Time-to-market“
Die Lebenszyklen von nahezu allen Maschinentypen werden verkürzt. Das bedeutet, dass auch die Entwicklungszeiten kürzer werden, um eine schnelle Amortisation des Entwicklungsaufwandes zu gewährleisten. Dies hat zur Folge, dass Maschinen- und Antriebskonstrukteure nach besonders verlässlichen Sensorlösungen mit einem hohen Integrationsgrad suchen.

Diese Trends haben eine unmittelbare Auswirkung auf die Sensoren, die für Längen- und Winkelmessaufgaben eingesetzt werden.

3. Neue Anforderungen an die Sensorik

Der Konstrukteur muss ein Messsystem auswählen, das folgende Eigenschaften vereint:

- Hohe Auflösung für hohe Regelgüte
- Hohe Genauigkeit für präzise Positionierung
- Hohe Bandbreite, um den Einsatz auch bei sehr hohen Geschwindigkeiten und Drehzahlen zu ermöglichen
- Einfache, günstige Montage und Einrichtung
- Einsetzbar unter schwierigen Umgebungsbedingungen (hohe oder niedrige Temperaturen, Verschmutzung, hohe mechanische Belastung usw.)
- Unempfindlichkeit gegen Schock und Vibration
- Geringes Bauvolumen
- Hohe Energieeffizienz
- Hohe Zuverlässigkeit

Diese komplexe Anforderungsliste kann durch Messsysteme, die auf magnetoresistiven Sensoren basieren, vollständig erfüllt werden. Lineare und Dreh-Geber auf MR-Basis bieten mehrere Vorteile gegenüber anderen Messprinzipien. Optische Geber verfügen über eine hohe Genauigkeit, bieten aber auf Grund begrenzter Linearität eine niedrigere Auflösung. Die höchste zulässige Einsatztemperatur für optische Geber liegt oft nur bei 85 °C und nur selten bei 100 °C. MR-Sensoren können bei deutlich höheren Temperaturen eingesetzt werden und sind bis 150 °C qualifiziert, sowohl für Industrie- als auch Automobilanwendungen. Optische Geber bieten zudem ein begrenztes Miniaturisierungspotenzial auf Grund des komplexen Aufbaus. Ein entscheidender Vorteil von magnetischen Gebern ist die deutlich höhere Robustheit und Widerstandsfähigkeit gegenüber Verschmutzung. Diese Vorteile bieten die Grundlage für den stetig steigenden Marktanteil magnetischer Messsysteme.

4. Grundlagen der tunnelmagnetoresistiven Sensortechnologie

Der magnetische Tunnelwiderstand ist ein magnetoresistiver Effekt, der in magnetischen Tunnelkontakten auftritt. In seiner einfachsten Form handelt es sich um ein Bauelement bestehend aus zwei Ferromagneten (Free Layer und Pinned Layer, dargestellt in Abb. 2), die durch einen dünnen Isolator getrennt sind. Ist die isolierende Schicht oder Tunnelbarriere (Barrier Layer, dargestellt in Abb. 2) extrem dünn (einige wenige Nanometer, ähnlich groß wie der Durchmesser der menschlichen DNA), so können Elektronen zwischen den beiden Ferromagneten „tunneln“. Dieser Vorgang ist mit Hilfe der klassischen Physik nicht erklärbar und ist daher ein rein quantenmechanisches Phänomen [1].

Unter dem Einfluss eines äußeren Magnetfeldes kann die Richtung der Magnetisierung der beiden magnetischen Schichten unabhängig voneinander gesteuert werden. Wenn die Magnetisierungen gleich ausgerichtet sind, ist die Wahrscheinlichkeit, dass Elektronen durch die Isolatorschicht hindurch tunneln größer als bei gegensätzlicher Ausrichtung. Damit kann der elektrische Widerstand zwischen unterschiedlichen Widerstandszuständen ($\Delta R/R$) variieren.

Obwohl der TMR-Effekt seit den 1970er Jahren bekannt war, wurde es erst durch verbesserte Prozesstechnik ab Mitte der 1990er Jahre möglich, reproduzierbare Ergebnisse zu erzielen. Seit wenigen Jahren hat die Entwicklung nun auch in die Industrie Einzug gehalten, so dass inzwischen Festplatten mit TMR- statt GMR- (Giant MR-Effekt) Schreib-Leseköpfen angeboten werden. Trotzdem ist auf dem Gebiet der Sensorik bisher vornehmlich mit AMR (Anisotroper MR-Effekt) und GMR gearbeitet worden (Abb. 2). Die TMR-Sensoren werden auch in Dünnschichttechnologie hergestellt, wobei wesentliche Teile der Produktionskette für GMR-Sensoren übernommen werden können.

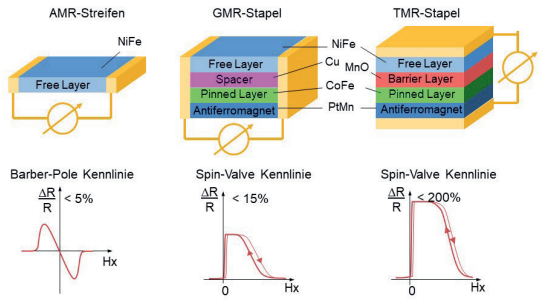


Abb. 2: Vergleich der xMR-Technologien

Der Widerstand des TMR-Sensors ist definiert durch die Fläche der Tunnelverbindung und die Schichtdicke bzw. den Flächenwiderstand der Barriere. Ein typischer Flächenwiderstand einer Al_2O_3 Barriere liegt bei 10 MΩ/µm². Gestaltet man eine Tunnelbarriere von 100 µm² Fläche, so ergibt sich ein Widerstand von 100 kΩ aus einem einzigen Tunnelelement.

Im Gegensatz zu AMR- und GMR-Sensoren gilt: je kleiner das Sensorelement, desto größer der Widerstand. Kleinere Tunnelelemente führen zu einem höheren Widerstand. Dies ist ein elementarer Unterschied zu herkömmlichen AMR- und GMR-Sensoren. Er ermöglicht auf gleichem Platz wesentlich hochohmigere Sensoren zu platzieren als es mit AMR- oder GMR-Sensoren möglich ist. Dadurch kann mit TMR der Leistungsverbrauch um einen Faktor 100 bis 1000 reduziert werden [2, 3, 4].

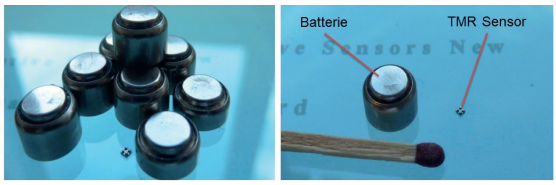


Abb. 3: Bedeutung des Energieverbrauchs in der Sensorik: Deutlich reduzierter Energiebedarf durch TMR-Sensor (rechts) im Vergleich mit AMR-Sensor (links) in einer beispielhaften Anwendung

5. Anwendungsbeispiele

5.1 Winkelsensorik

Ein weiterer ganz wesentlicher Unterschied zwischen AMR-Sensoren und GMR- bzw. TMR-Sensoren ist das Verhalten im Drehfeld. Dreht bei einem AMR-Sensor die Magnetisierung um 90° von parallel zur Stromrichtung auf senkrecht zur Stromrichtung, ist bereits der komplette MR-Hub erschöpft und bei 180° ist der Ausgangszustand wiedergegeben. Die Periodizität ist demnach nur 180°. Im Gegensatz dazu ist bei GMR- und TMR-Systemen der Ausgangszustand erst nach einer vollen Umdrehung wieder erreicht und somit eine Periodizität von 360° gegeben [5, 6].

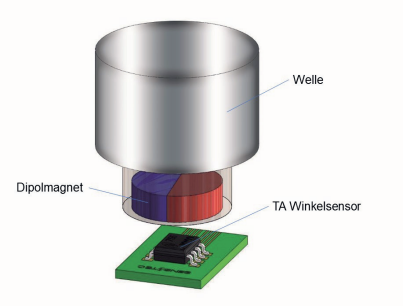


Abb. 4: Typische Messanordnung

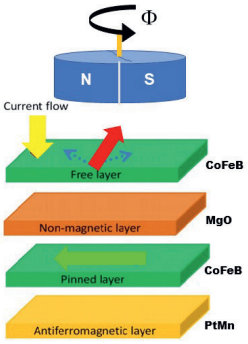


Abb. 5: Funktionsprinzip eines TMR-Winkelsensors

TMR Sensoren sind damit sehr gut geeignet für Anwendungen in der Winkelsensorik, z. B. in Drehgebern am Wellenende eines Antriebs. Eine typische Messanordnung sowie das Funktionsprinzip sind in Abb.4 bzw. Abb. 5 dargestellt. Üblicherweise wird ein Dipolmagnet an der drehenden Welle angebracht. Eine Drehbewegung der Antriebswelle resultiert in einer Änderung in der Magnetisierungsrichtung des Free Layers, was wiederum zu einer Widerstandsänderung führt.
Um höchste Anforderungen an Signalqualität zu erfüllen, werden die MR-Streifen im PerfectWave® Design ausgeführt. Die Streifen haben eine gekrümmte Form (Abb. 6), die zur Oberwellenfilterung bei der Abbildung der Magnetfeldrichtung in ein elektrisches Signal genutzt wird. Diese Filterung wird durch die spezielle Geometrie und Anordnung der TMR-Strukturen realisiert und verursacht keine zusätzlichen Signallaufzeiten. PerfectWave® wirkt sich besonders bei kleineren Magnetfeldern in verbesserter Linearität, höheren Genauigkeiten und dadurch besserer Regelgüte aus. Durch die Oberwellenfilterung werden mehrere Oberwellen unterdrückt, und das analoge Sensorausgangssignal gleicht dadurch nahezu einem perfekten Sinus.

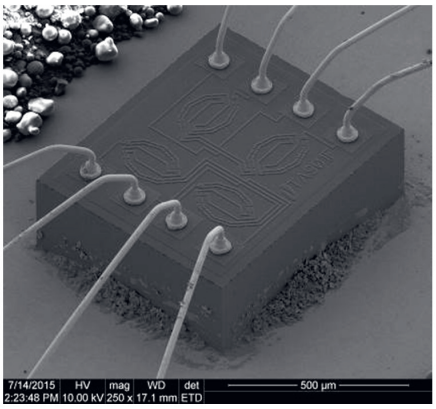


Abb. 6: REM-Aufnahme eines TMR-Winkelsensor-Chips

Durch dieses Verfahren können absolute Winkelgenauigkeiten von besser als 0,4° bei Raumtemperatur und spezifizierte Einbautoleranzen über einen breiten Feldstärkenbereich erreicht werden (Abb. 7) [7].
Bei dem quantitativen Vergleich von Leistungseigenschaften von gängigen AMR- und TMR-Winkelsensoren, dargestellt in Tabelle 1, erkennt man, dass der AMR Sensor weiterhin eine höhere Genauigkeit und niedrigere Hysterese aufweist. Der TMR-Sensor kann jedoch eine absolute Winkelmessung über 360° ermöglichen und weist eine exzellente Temperaturstabilität aus, was in vielen Anwendungen einen Temperaturabgleich überflüssig macht. Weiterhin bedeutet die hohe Signalamplitude, dass oft auf eine Verstärkung des Analog-Sensorsignals verzichtet werden kann, was wiederum zu Kostenersparnissen führen kann.

Tab. 1: Vergleich AMR- und TMR-Winkelsensoren

	AMR	TMR
Periodizität	180°	360°
Genauigkeit	0.1°	0.4°
Signal Amplitude	12 mV/V	80 mV/V
Amplitude TK	-0,35%/K	-0,15%/K
Offset	2 mV/V	3 mV/V
Relativer Offset TK	~ 1,6% @100 K	~ 1% @100 K
Widerstand TK	0,42%/K	-0,1%/K
Hysterese	vernachlässigbar	0,1 %
Sensorwiderstand	Meanderfläche	Barrieredicke

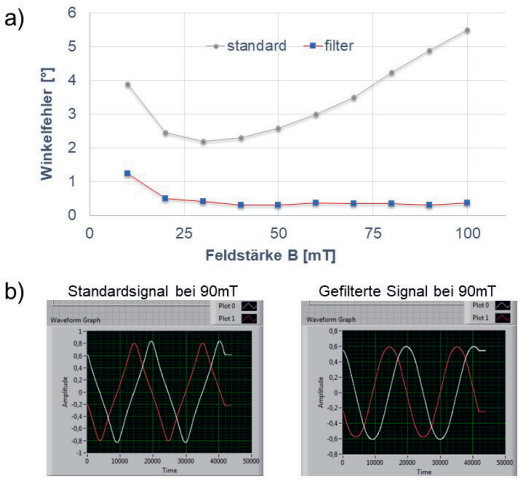


Abb. 7: Oberwellenfilterung

TMR-Sensoren sind nicht nur für die Winkelmessung am Wellenende anwendbar, sondern auch am Wellenumfang. Um letzteres zu ermöglichen, wird das Nonius-Prinzip angewandt, d. h. es werden zwei oder drei Spuren mit unterschiedlichen Polteilungen benutzt und der Winkel wird durch die unterschiedlichen Phasenlagen der Signale der jeweiligen Sensoren (ein Sensor pro Spur) berechnet. In Zusammenarbeit mit NTN-SNR hat Sensitec eine neue Lösung entwickelt mit einer Auflösung bis 17 bit/Umdrehung (131.072 Puls/Umdrehung) sowie einer sehr hohen absoluten Winkelgenauigkeit von ±0,1° (Abb. 8).

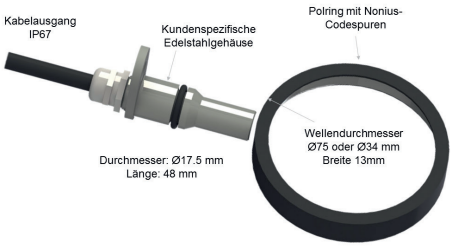


Abb. 8: Absolut-Winkelmesssystem auf TMR-Basis (Bildquelle: NTN-SNR Bearings GmbH)

5.2 Längensensorik

Das Nonius-Prinzip kann auch bei Absolut-Längenmesssystemen angewandt werden (Abb. 9). Auch hier verfügt der magnetische Maßstab über zwei oder drei Spuren mit unterschiedlichen Polteilungen. Alternativ verfügt der Maßstab über eine inkrementelle Code-Spur in Verbindung mit einer Spur mit einer Pseudo-Random-Codierung.

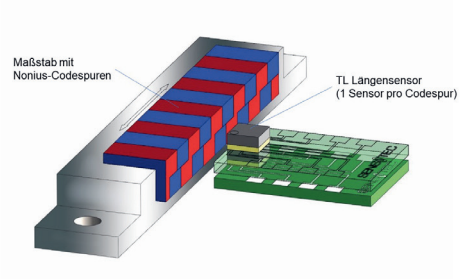


Abb. 9: Absolute Wegmessung mittels Nonius-Prinzip

Für solche Anwendungen kommen Längensensoren auf TMR-Basis zum Einsatz. Diese Sensoren sind an der Polteilung des Maßstabes angepasst und verfügen weiterhin über eine Oberwellenfilterung. Sie weisen eine extrem konstante Signalamplitude und absolute Genauigkeit über einer breiten Arbeitsabstandsbereich auf (Abb. 9 bzw. Abb.10).

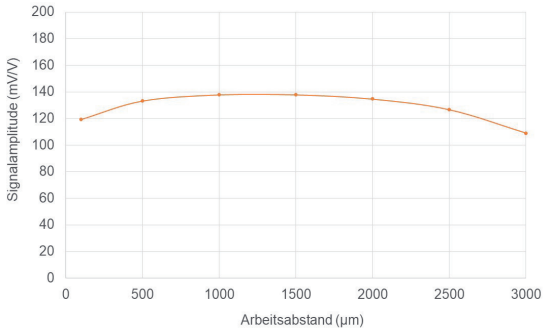


Abb. 10: Signalamplitude vs. Arbeitsabstand

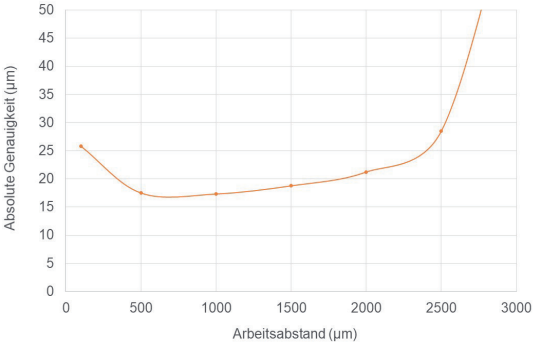


Abb. 11: Absolute Genauigkeit vs. Arbeitsabstand

Innerhalb des BMBF-geförderten Forschungsprojekts KaLiPso - Kabellose Linearführungssysteme mit integrierter Positionsmessung - ist es gelungen, durch die Zusammenarbeit zwischen NTN-SNR sowie dem Institut für Fertigungstechnik und Werkzeugmaschinen an der Universität Hannover als auch Sensitec, ein kabelloses Absolut-Längenmesssystem auf TMR-Basis in einer Linearführung zu integrieren (Abb 12). Wie in Abb. 13 gezeigt ist das Längenmesssystem in der Linearführung integriert und wird mittels eines berührungslosen induktiven Verfahrens mit Energie versorgt. Die Primärspule ist in die Führungsschiene integriert und die Sekundärspule befindet sich im Schlitten [8].

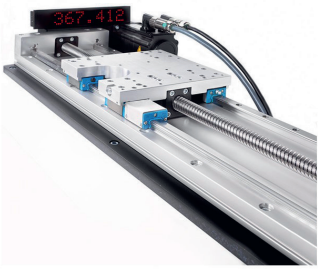


Abb. 12: Linearachse (Bildquelle: NTN-SNR Bearings GmbH)

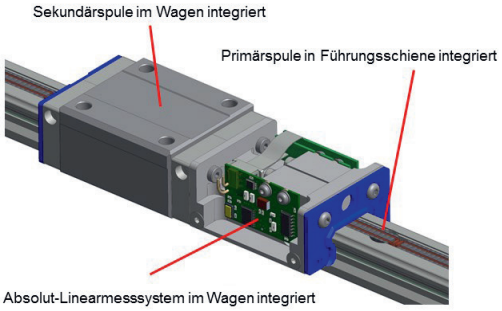


Abb. 13: Integriertes Absolut-Linearmesssystem auf TMR-Basis (Bildquelle: NTN-SNR Bearings GmbH)

Es gibt gleichzeitig einen Bedarf an Längenmesssystemen für kurze Hübe mit sehr kompakten Abmessungen für die Integration in z. B. Miniatur-Linearführungen oder sehr kleine Wälzlager (Abb. 14). Im kürzlich abgeschlossenen BMBF-geförderten Projekt „Aquila“ wurden lithographisch hergestellte Maßverkörperungen entwickelt, um absolute Positioniergenauigkeiten von < 1 µm zu erreichen [9].

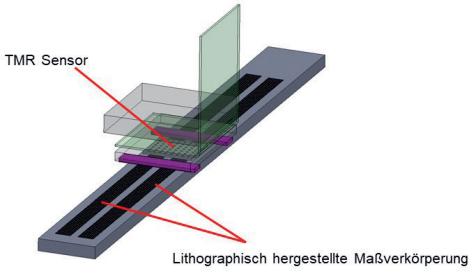


Abb. 14: Grundkonzept des miniaturisierten Wegmesssystems

Der magnetische Maßstab verfügt über zwei hartmagnetische Code-Spuren (eine inkrementelle, eine absolute), die mittels lithographischer Verfahren auf einem Silizium-Substrat strukturiert werden (Abb. 15).

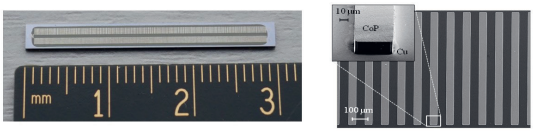


Abb. 15: Lithographisch hergestellte magnetische Maßstäbe

Die Polteilung beträgt nur 20 µm, eine Größenordnung feiner als bei bisherigen magnetischen Maßstäben, bei denen die kleinsten Polteilungen üblicherweise zwischen 200 und 500 µm lagen. Ein speziell entwickelter TMR-Längensensor liest beide Code-Spuren gleichzeitig, um eine Absolut-Längenmessung zu realisieren (Abb. 16, 17).

Die extrem kleine Polteilung bedeutet, dass der Luftspalt zwischen Sensor und Maßstab so klein ist, dass konventionelle Wirebonds zwischen Sensorchip und Leiterplatte nicht anwendbar sind. Daher mussten neue Aufbau- und Verbindungstechniken untersucht werden, um die empfindliche Fläche des Sensorchips so nah wie möglich an die Maßverkörperung zu bringen.

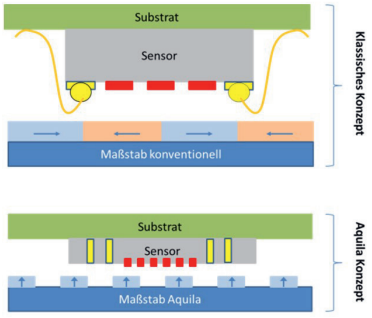


Abb. 16: TMR Längensensor mit „Through-Silicon-Vias“ (TSV)

Through-silicon vias (TSV) sind vertikale elektrische Verbindungen, die durch den Sensorchip geführt werden und eine Kontaktierung auf der Rückseite des Chips erlauben. Die TSVs ermöglichen zudem eine deutliche Reduzierung der Chipfläche und vermindern damit auch die Gesamtabmessungen des Messsystems [10]. Wie aus Abb. 17 zu entnehmen ist, verfügt der Maßstab über zwei Kodespuren – eine Spur für hochauflösende inkrementelle Weginformation, die zweite Spur mit einer speziellen Kodierung für Absolut-Weginformation (PBRS – Pseudo-Random-Kodierung). Der Sensorchip hat zwei empfindliche Bereiche, die an die jeweiligen Kodespuren angepasst sind.

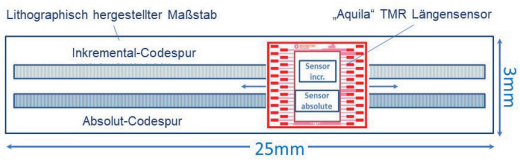


Abb. 17: Prinzip der absoluten Wegmessung

Mit einem Laserinterferometer als Referenzpositionsgeber (Messunsicherheit <120 nm/m) wurde die Genauigkeit des AQUILA-Systems gemessen. Abb. 18 zeigt oben die Auswertung der Absolutmessung: Das PRBS-Signal ist über den gesamten Maßstab eindeutig. Damit wird die “true-power-on” Ermittlung der Absolutposition nachgewiesen. Direkt nach dem Einschalten kann fehlerfrei die Absolutposition an jeder Position auf dem Maßstab bestimmt werden, ohne auf gespeicherte Daten zurück-

greifen zu müssen oder eine Referenzfahrt durchzuführen. Zur Bewertung der Systemgenauigkeit wurde die inkrementelle Spur gegen das Referenzinterferometer verglichen (Abb. 18, unten). Über den gesamten Bewegungshub beträgt die Abweichung maximal ±700 nm. Dabei wurde auf die Rohdaten lediglich eine globale Offset- und Amplitudenkorrektur angewandt. Da durch die PRBS-Spur die absolute Position bekannt ist, kann das System zusätzlich kalibriert werden. Mit Hilfe einer Fehlertabelle lässt sich die Genauigkeit somit noch weiter verbessern [10].

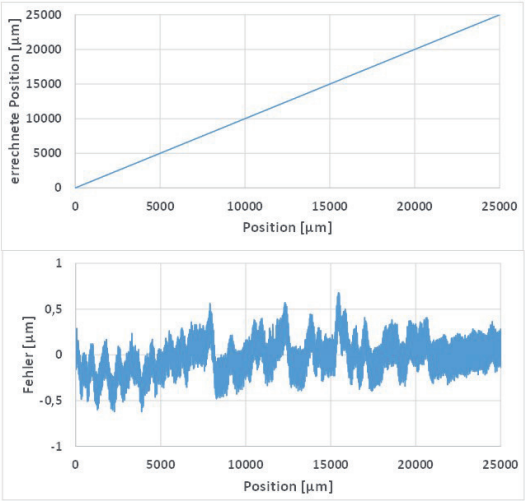


Abb. 18: Messergebnisse für miniaturisiertes Wegmesssystem

Aufgrund dieser hervorragenden Ergebnisse wurden die Maßstäbe und TMR-Längensensoren unmittelbar in industrielle Anwendungen übertragen. Abb. 19 zeigt eine Fokussierachse für einen Mikroskop-Tisch der Projektpartner Märzhäuser. Mit diesem neuen Ansatz hat Sensitec mit den weiteren Projektpartnern die erreichbare Auflösung und Genauigkeit von magnetischen Längenmesssystemen in einen bisher unerreichten Bereich gebracht und bietet damit eine robuste Alternativ zu hochauflösenden optischen Längenmesssystemen.

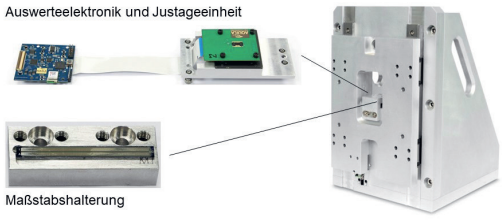


Abb. 18: Anwendungsbeispiel – Fokussierantrieb für Mikroskop (Bildquelle: Märzhäuser GmbH & Co. KG)

5.3 Feldsensorik

Aufgrund der sehr kleinen Abmessungen und hohen Empfindlichkeit liegt es nahe, auch Feldsensoren auf TMR-Basis zu entwickeln. Abb. 20 zeigt eine typische Ausführung, wo der Feldsensor als Chip-Scale-Package umgesetzt wurde.

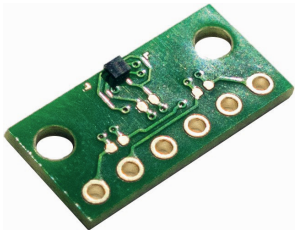


Abb. 20: Feldsensor auf TMR-Basis (Flip-Chip auf Leiterplatte)

Eine typische Anwendung ist die als Schaltsensor. Die Anordnung in Abb. 21 kann benutzt werden für Endlagen-Schalter in Linearzylindern auf elektrischer, pneumatischer oder hydraulischer Basis. Der Dipolmagnet wird auf der Schubstange angebracht und der TMR Feldsensor üblicherweise in das Zylindergehäuse integriert. Die hohe Empfindlichkeit des TMR-Sensors führt dazu, dass entweder der Abstand zwischen Magnet und Sensor vergrößert werden kann oder kleinere Magnete eingesetzt werden können.

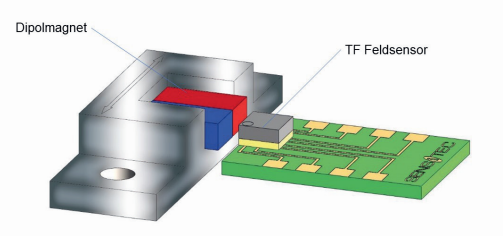


Abb. 21: Typische Messanordnung für TMR-Schaltsensor

Eine weitere Anwendungsmöglichkeit für TMR-Feldsensoren findet sich in Wegmesssystemen (Abb. 22). In Lineardirektantrieben ist es vorteilhaft die Statormagnete als Maßverkörperung zu nutzen, d .h. es bedarf keines separaten Maßstabs, wie es bei optischen Messsystemen erforderlich wäre. Um die Auswirkungen von geometrischen Fehlern der Magnetsegmente bzw. Teilungsfehlern zu unterdrücken, wird üblicherweise ein Array an Sensoren ausgeführt. Diese Anordnung ist besonders attraktiv in Handling-Applikationen, wo ausreichende Genauigkeit geboten wird trotz eines sehr einfachen und kompakten Aufbaus des Messsystems.

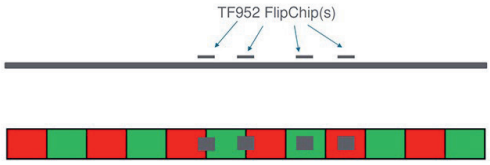


Abb. 22: Wegmesssystem mit TMR-Feldsensoren

6. Zusammenfassung

Steigende Anforderungen an Sensoren in der Antriebstechnik führen zu einem wachsenden Anteil an magnetischen Sensoren, insbesondere jener, die auf dem magnetoresistiven Effekt basieren. MR Sensoren sind ein gelungenes Beispiel für die Mikro-Nano-Integration. Aktuelle Neuentwicklungen bei Sensoren auf Basis des Tunnelmagnetoresistiven-Effekts öffnen neue Möglichkeiten für Antriebs- und Maschinenentwickler. Alle wichtigen Applikationen, um Bewegungen in der elektrischen Antriebstechnik zu erfassen, können mit TMR-Sensoren abgedeckt werden. Damit bilden TMR-Sensoren eine wichtige Ergänzung für die etablierten AMR- und GMR-Technologien.
Bei TMR-Sensoren verfügt eine neue Winkelsensorfamilie über eine höhere Winkelgenauigkeit und höhere Dynamik als bisherige TMR-Sensoren und auch Hall-Effekt Sensoren. Neue Fertigungstechnologien ermöglichen miniaturisierte Längenmesssysteme für komplett neue Anwendungen. Extrem kleine Feldsensoren sind sehr gut geeignet als Endlagen-Schalter und ermöglichen deutlich kompaktere Konstruktionen.
TMR-Sensoren sind nicht nur für Winkel- und Längenmessaufgaben geeignet, sondern auch für Anwendungen im Bereich der mobilen Navigation, des Verkehrsmanagements, der Strommessung und Zustandsüberwachung [11, 12, 13].

7. Literatur

[1] Duret, C. et al, “TMR : a new frontier for magnetic sensing”, Proc. of 11th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar, 2011
[2] Paul,J., Lehndorff, R. & Duret, C.; „Ressourceneffizienz am Beispiel magnetoresistiver Sensoren“, Mikrosystemtechnik Kongress 2011, Darmstadt
[3] Slatter, R., Doms, M., “Energy efficient magnetoresistive sensors for low-power and wireless applications”, Proc. of Smart Systems Integration Conference, Vienna, 2014
[4] Slatter, R., Buss, R.; „Energieeffiziente, miniaturisierte magnetische Sensoren für energieautare Anwendungen“, 8. GMM-Workshop Energieautarke Sensorik - EAS 2016, Renningen, 2016
[5] Doms, M. et al; “AMR vs. GMR vs. TMR – Eigenschaften, Unterschiede, Anwendungen“, Proc. of 5. Mikrosystemkongress, Aachen, Germany, 2013
[6] Slatter, R.; „Neuartige magnetoresistive Sensoren für Winkel- und Längenmessaufgaben im Automobil“, Sensoren im Automobil IV, Expert Verlag, Renningen, 2011
[7] Slatter, R., von Manteuffel, G.; „Neue Entwicklungen bei magnetoresistiven Sensoren für Weg-, Winkel-, Strom- und Feldmessung in und um das Automobil“ in Automobil-Sensorik, Tille, T. (Hrsg.), Springer Vieweg, Wiesbaden, 2016
[8] Lindemann, R.; “KaLiPso - Wireless position measuring in linear guides using integrated TMR Sensors”, Proc. of 14th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar, 2017
[9] Slatter, R. & Paul, J.; “A new approach to miniaturize high resolution linear position measurement systems”, Proc. of 14th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar, 2017
[10] Paul,J. *et al*; „High Resolution Magnetic Position Sensors for Precise Control of Production Tools“, Proc. of 15th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar 2019
[11] Helwig, N.; “Modular Sensor Systems for real time Process Control and Smart Condition Monitoring using XMR sensor technology“, Proc. of 14th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar, 2017
[12] Slatter, R., Holland, L. & Abele, E.; „Magnetoresistive sensors for the condition monitoring of high-frequency spindles”, Proc. of 7th HPC – CIRP Conference on High Performance Cutting, Chemnitz, 2016
[13] Traute, J., Casper, F. & Paul, J.; „Robuste xMR-Sensoren für die Zustandsüberwachung“, 20. GMA/ITG Fachtagung Sensoren & Messsysteme, Nürnberg, 2019

DOI 10.5162/sensoren2019/4.1.1

Messung mechanischer Beanspruchungen in laufenden Schleifprozessen

Andreas Tausendfreund [1], Gabriela Alexe [1], Dirk Stöbener [1,2], Andreas Fischer [1,2]
[1]Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft (BIMAQ),
Linzer Straße 13, 28359 Bremen, Deutschland
[2]MAPEX Center for Materials and Processes,
Postfach 330 440,28334 Bremen, Deutschland

Zusammenfassung

Die Kenntnis der während eines Fertigungsprozesses (z.B. Schleifen) auftretenden Beanspruchungen und der im Material verbleibenden Veränderungen wird im Konzept der Prozesssignaturen genutzt, um den Fertigungsprozess zu optimieren und mit anderen zu vergleichen (z.B. Laserbearbeitung). Voraussetzung für die Erstellung einer Prozesssignatur ist, dass die Beanspruchungen während des laufenden Prozesses charakterisiert werden können. Aufgrund der rauen Prozessbedingungen gibt es beim Schleifen bisher keine prozessbegleitende Technik zur Messung der Beanspruchungen in Form von Verschiebungen und Dehnungen in der bearbeiteten Randzone. Aus diesem Grund wird die Eignung der Speckle-Fotografie für prozessbegleitende Messungen von Materialbeanspruchungen in einem laufenden Schleifprozess ohne Kühlschmierstoff nachgewiesen und ein Konzept für eine Messung unter dem Einsatz von Kühlschmierstoffen vorgestellt.

Keywords: In-Prozess-Messung, Speckle-Fotografie, Schleifprozess, Verformungs- und Dehnungsmessung

Einleitung

In der Metallverarbeitung geht es darum, aus einem Grundmaterial ein Produkt mit funktionalen Eigenschaften zu erzeugen. Hierfür sind in der Regel verschiedene Fertigungsprozesse erforderlich, um die Form, die Oberflächenrauheit bzw. die Randzoneneigenschaft des Werkstücks gezielt einzustellen. Eine Prozesskette wie z.B. das Fräsen mit anschließenden Schleifen und Laserhärten lässt sich unter Umständen kostengünstiger gestalten, indem nach dem Fräsprozess ein Schleifprozess mit kombiniertem Schleifverfestigen erfolgt. Aktuelle Forschungsanstrengungen im Rahmen des Sonderforschungsbereichs SFB/TRR 136 zielen deshalb darauf ab, Fertigungsschritte prozessunabhängig über die sogenannte Prozesssignatur zu beschreiben. Die Grundvoraussetzung dafür ist jedoch eine In-Prozess-Charakterisierung der mechanischen Beanspruchungen des Werkstoffs in Form von Verformungen und Dehnungen.

Prädestiniert für die berührungslose Messung von In-Plane-Verschiebungsfeldern sind die seit den 1970er Jahren erforschten Speckle-Korrelations-Verfahren [1, 2]. Bei diesen Verfahren wird die zu messende Oberfläche mit einem Laser beleuchtet und die Auswertebereiche in den mit der Kamera aufgenommenen Specklebildern – jeweils vor und während des Auftretens der Verformung - miteinander korreliert. Aus den korrelierten Bildausschnitten lassen sich lokale Ortsverschiebungen bestimmen, deren theoretische Auflösungen lediglich durch die Heisenbergsche Unschärferelation beschränkt sind [3]. Für Verschiebungsmessungen können mit aktueller Kamera- und Lasertechnologie Unsicherheiten von weniger als 7 nm erreicht werden [3]. Als berührungsloses optisches Messverfahren ist die Speckle-Fotografie daher prädestiniert für schnelle und präzise In-Prozess-Messungen. Die Speckle-Fotografie konnte kürzlich für Verformungsmessungen im Walz- [4] und Fräsprozess [5] eingesetzt werden. Aufgrund der rauen Umgebungsbedingungen mit starkem Funkenflug oder dem Einsatz von Kühlschmierstoffen ist die Speckle-Fotografie beim Schleifprozess unter realen Bedingungen jedoch noch nicht zur Anwendung gekommen.

Der vorliegende Beitrag beschäftigt sich deshalb mit speckle-fotografischen In-Prozess-Dehnungsmessungen beim Schleifen. Die sich hierbei ergebenden wissenschaftlichen Fragestellungen sind, ob trotz der anspruchsvollen Prozessbedingungen Speckle-Fotografie-Messungen möglich sind, wie eine robuste Bildauswertung erzielt werden kann und welche Messabweichungen sich durch den Einsatz von Kühlschmierstoffen ergeben.

Verschiebungs- und Dehnungsmessungen mittels Speckle-Fotografie

Die Speckle-Fotografie ist eine bekannte Technik zur Charakterisierung von Verschiebungsfeldern in der Ebene einer zu messenden Oberfläche [1, 2]. Eine Kamera erfasst hierbei das an der Oberfläche diffus gestreute kohärente Laserlicht. Die von der Oberfläche modulierten Phasenelemente der Lichtwelle erzeugen ein Bild der Topographie in der Bildebene (auf dem Kamerachip), welches aufgrund von konstruktiver und destruktiver Interferenz mit einem Pünktchen- bzw. dem Specklemuster überlagert ist (Abbildung 1). Die Größe der Speckle kann durch die Öffnung der Blende in einem weiten Bereich variiert werden. Ein herausragendes Merkmal der subjektiven Speckle ist, dass sie – entsprechend dem Abbildungssystem – spezifischen Oberflächenpunkten zugeordnet werden können. Daher lassen sich Verschiebungen einzelner Speckle als Verschiebungen lokaler Oberflächenpunkte in der Objektebene interpretieren. Zur Berechnung des globalen Verschiebungsfelds der gemessenen Oberfläche, rastert ein kleines Auswertefenster über das Messfeld, hierbei wird in jedem Punkt jeweils die Kreuzkorrelation zwischen zwei aufeinanderfolgenden Specklebildern gebildet. Die Größe des Auswertefensters lässt sich automatisch berechnen und sollte mindestens drei Speckle auf der Diagonalen enthalten. Ein optimaler Speckledurchmesser beträgt nach Zhou und Goodson ca. 3-5 Pixel [6]. Durch die Bestimmung der Verschiebung der Korrelationsmaxima können elastische und plastische Verformungen mit einer räumlichen Auflösung von etwa 20 nm und zeitlichen Auflösung im Mikrosekundenbereich rekonstruiert werden [5].

Nach der Berechnung des Punktverschiebungsfeldes $\mathbf{P}(u,v)$ (wobei u und v die Verschiebungen in x- und y-Richtung sind) lässt sich das Gradientenfeld $\mathbf{F}_{2D}$ der Verformung nach Kajberg [7] bestimmen:

$$\mathbf{F}_{2D} = \begin{pmatrix} 1+\frac{\partial u}{\partial x} & 1+\frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & 1+\frac{\partial v}{\partial y} \end{pmatrix}. \quad (1)$$

Die Polardarstellung des $\mathbf{F}_{2D}$ Felds führt zu

$$\mathbf{F}_{2D} = \mathbf{R}_{2D}\mathbf{U}_{2D}\ , \quad (2)$$

wobei $\mathbf{R}_{2D}$ der Rotationstensor ($\mathbf{R}^{T}_{2D}\mathbf{R}_{2D} = \mathbf{I}$) und $\mathbf{U}_{2D}$ der gesuchte positiv definite und symmetrische Dehnungstensor ($\mathbf{U}^{T}_{2D} = \mathbf{U}_{2D}$) ist. Zunächst kann $\mathbf{F}^{T}_{2D}\mathbf{F}_{2D} = (\mathbf{R}_{2D}\mathbf{U}_{2D})^{T}\mathbf{R}_{2D}\mathbf{U}_{2D} = \mathbf{U}^{T}_{2D}\mathbf{U}_{2D}$ geschrieben werden als

$$\mathbf{U}^{2}_{2D} = \mathbf{F}^{T}_{2D}\mathbf{F}_{2D}. \quad (3)$$

Um dann die Quadratwurzel zu lösen, wird folgende spektrale Zerlegung verwendet:

$$\mathbf{U}_{2D} = \lambda_1\mathbf{n}_1\mathbf{n}^{T}_1 + \lambda_2\mathbf{n}_2\mathbf{n}^{T}_2. \quad (5)$$

Hierbei sind λ_1 und λ_2 die Quadratwurzeln der Eigenwerte der symmetrischen Matrix $\mathbf{U}^{2}_{2D}$ und $\mathbf{n}_1$ und $\mathbf{n}_2$ die dazugehörigen Eigenvektoren. Die logarithmische (Hencky-) ln-Plane-Dehnung $\boldsymbol{\varepsilon}_{2D}$ ergibt sich schließlich zu [8]

$$\boldsymbol{\varepsilon}_{2D} = \ln(\mathbf{U}_{2D}). \quad (6)$$

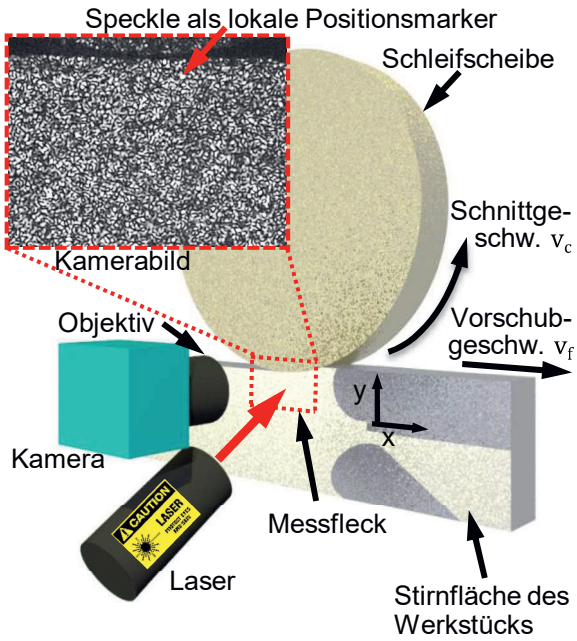


Abb. 1: Mess- und Einsatzprinzip der Speckle-Fotografie beim Flachschleifen.

Das Grundprinzip für die Dehnungsmessungen im Flachschleifprozess ist in Abbildung 1 dargestellt. Das Messfeld befindet sich auf der Stirnfläche – der unbearbeiteten Seite des Werkstücks – direkt an der Grenze zur bearbeiteten Oberfläche. Das Objektiv und die CMOS-Kamera (complementary metal-oxide-semiconductor) mit globalen Shutter sind relativ zum Werkstücksystem fixiert, so dass auftretende Specklebewegungen direkt mit Verschiebungen von Oberflächenpunkten korreliert werden können.

Out-of-Plane-Verformungen der Oberfläche im Schleifprozess führen zu einer Dekorrelation des Specklemusters, wodurch der gewünschte Messeffekt zerstört wird. Der Einsatz einer Hochgeschwindigkeitskamera (Optronis: CP70) mit einer schnellen Bild-zu-Bild-Auswertung bei einer Bildrate von 167 fps reduziert diese Dekorrelation, da die Out-of-Plane-Verschiebung in der kurzen Zeit von 1/167 s so gering ist, dass fast keine Speckledekorrelation auftritt.

Die gemessenen Verschiebungen in der Objektebene sind in dieser kurzen Zeit entsprechend klein. Mit einer geeigneten zweidimensionalen Sub-Pixel-Interpolation können jedoch Verschiebungen von wenigen Hundertstel Pixeln (hier 1 Pixel ≙ 1,94 µm) bestimmt werden [9].

Experimenteller Aufbau

Die Versuche wurden auf einer *Blohm Profimat 412 HSG* Profilschleifmaschine durchgeführt. Die Umsetzung des Messprinzips ist im unteren, vergrößerten Bereich von Abbildung 2 zu sehen.

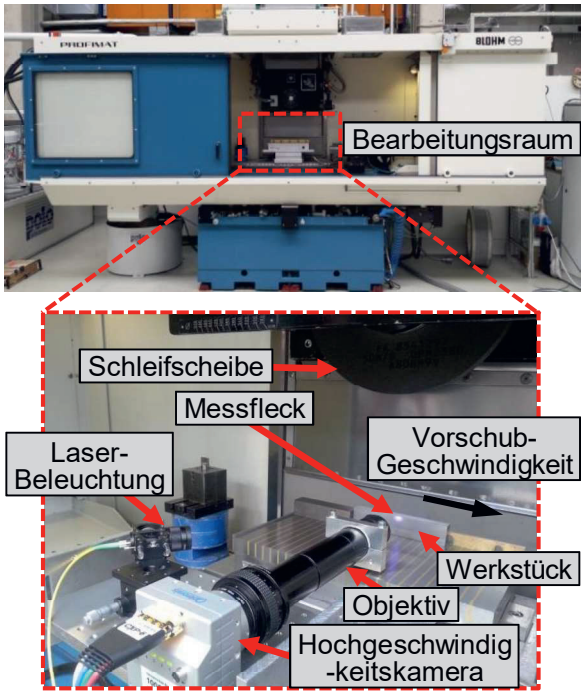


Abb. 2: Blohm Profimat 412 HSG Schleifmaschine und integrierter Messaufbau

Das Werkstück ist als Quader mit einer Länge von 165 mm, 18 mm Breite und 41,5 mm Höhe ausgeführt. Um die Peripheriegeräte (Kamera, Laser (λ = 405 nm, I = 80 mW) und optische Komponenten) fest mit dem Messobjekt zu verbinden, werden die Adapter und Halterungen gemeinsam mit dem Werkstück magnetisch gespannt. Die dem Specklemesssystem zugewandte Werkstückseite wird vor den Experimenten kugelgestrahlt, um eine raue Oberfläche zu erzeugen, was zu einer ausreichenden Phasenmischung der Laserwellenfront und zur Bildung eines voll entwickelten und homogenen Specklemusters führt. Die Messung erfolgt an einer Stelle im mittleren Abschnitt des Werkstücks. Der Messaufbau ist mit der Linearbewegung des Werkstücks in x-Achsrichtung verbunden. Daher bezieht sich die Vorschubgeschwindigkeit von v_f = 2 m/min auf die relative Vorschubgeschwindigkeit der Schleifscheibe gegenüber dem Werkstück. Der Vorschub des Werkstücks ist gegenläufig zur Schnittgeschwindigkeit v_c = 20 m/s der Schleifscheibe. Somit bewegt sich die Vorschubgeschwindigkeit v_f der Schleifscheibe in die gleiche Richtung wie v_c (siehe Abbildung 1).

Ein sich verändernder Flüssigkeitsfilm auf der Oberfläche führt zu einer ständigen Specklekorrelation, wodurch der eigentliche Messeffekt verschwindet. Um die In-Prozess-Fähigkeit der Speckle-Fotografie beim Flachschleifen zu demonstrieren, wird im ersten Ansatz deshalb ein trockener Schleifprozess betrachtet und auf den Einsatz von Kühlschmierstoff verzichtet.

Das Fehlen der Kühlschmierstoffe führt jedoch zu einer starken Entwicklung von Funkenflug (siehe Abbildung 3). Glühende Späne können die optischen Komponenten beschädigen, diese müssen deshalb abgedeckt oder in einem entsprechend großen Abstand vom Punkt des Werkzeugeingriffs positioniert werden. Der große Arbeitsabstand von ca. 90 mm wird mit einem Hochleistungs-Vergrößerungs-objektiv (*Rodenstock Apo-Rodagon N*) erreicht.

Der Einsatz verschiedener Tuben ermöglicht eine dreifache Vergrößerung und führt so zu einem Messfeld von ca. 6 mm × 8 mm. Um die Optik zu stabilisieren und eine feste Position der Kamera zum Werkstück beim Schleifen zu gewährleisten, wurde eine spezielle Montagevorrichtung entwickelt. Eine Fixierung am vorderen Ende des Objektivs verhindert die Übertragung von Vibrationen von der Maschine auf die Kamera (siehe Abbildung 2). Dadurch werden Schwingungen sowohl für die Messeinheiten als auch für das Werkstück ausgeglichen und es ist möglich, auch während des Schleifvorgangs scharfe Bilder des Objekts aufzunehmen.

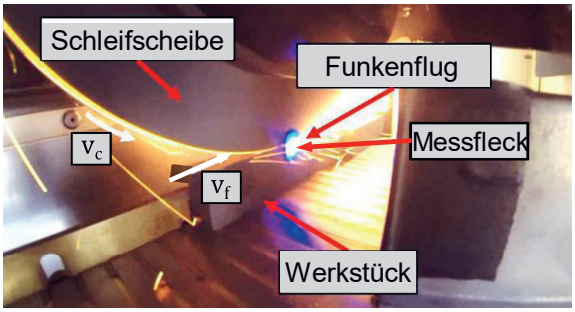


Abb. 3: Flachschleifprozess mit Funkenflug

Nach dem Schleifprozess konnte am Rand der bearbeiteten Oberfläche die Bildung eines Grats festgestellt werden. Die Gratbildung entspricht einer unterschiedlichen Neigung der Messfläche, wodurch das Licht aus dem oberen Teil des Werkstückbereichs nicht in die Kamera reflektiert wird. Dies führt zu einem

Kontrastverlust des Specklemusters an der Oberkante des Werkstücks und damit zu einer hohen Messunsicherheit in diesem Bereich. Daher lässt sich das Verschiebungsfeld der oberflächennahen Kontaktzone nur bis zu einer Tiefe von ca. 800 µm unterhalb der bearbeiteten Oberfläche berechnen. Unter Berücksichtigung der Schnitttiefe von a_e = 200 µm wird der Bereich zwischen y = 0 µm und y = -1000 µm unter der Oberfläche nicht bewertet (siehe Abbildung 4).

In den herkömmlichen Auswerteverfahren der Speckle-Fotografie wird der starke Partikel- bzw. Funkenflug des trockenen Schleifprozesses als Verschiebung von Oberflächenpunkten fehlinterpretiert. Die Partikelbewegung kann durch den Einsatz optimierter Algorithmen allerdings nahezu vollständig herausgerechnet werden. Zu diesem Zweck wird die Größe der Speckle innerhalb des Auswertefensters mit der durchschnittlichen Specklegröße im globalen Specklebild verglichen. In der Regel sind störende Partikel durch Funkenbildung deutlich größer als der eingestellte Speckledurchmesser und somit sofort erkennbar. In diesem Fall ist die lokale Verformung für die Messposition auf null zu setzten. Der Fehler, der beim Summieren der Verschiebungen auftritt, ist aufgrund der hohen Taktfrequenz der Kamera vernachlässigbar.

Geht man von einem homogenen Werkstück aus, so kann durch die zeitliche Abfolge der Einzelbilder, der Messbereich in Bewegungsrichtung der Schleifscheibe (x-Richtung) von 8 mm nahezu beliebig erweitert werden. In den gezeigten Messungen ist der Messbereich durch den Speicher der Kamera jedoch auf 32 mm beschränkt.

Ergebnisse der In-Prozess-Messung für den Flachschleifprozess ohne Kühlschmierstoff

Abbildung 4 zeigt die mittels Speckle-Fotografie an der Stirnfläche des Werkstücks gemessenen Verformungsfelder in x- und y-Richtung. Die tangentiale Kraftkomponente verschiebt das gesamte Werkstück um 10 µm in Richtung der Werkzeugbewegung v_c. Die maximale x-Verschiebung liegt ca. 2,5 mm nach der ersten Interaktion zwischen Werkzeug und Werkstück und beträgt 22 µm in Vorschubrichtung. Gegen Ende der Kontaktfläche zwischen Schleifscheibe und Werkstück verringert sich die x-Verschiebung auf ca. 15 µm.

Für die y-Komponente (Bild unten) ist eine Verschiebung nach unten, in negative y-Richtung, zu beobachten. Zehn Millimeter hinter der Schleifscheibe hat sich die Verschiebung in negative Richtung auf 25 µm erhöht.

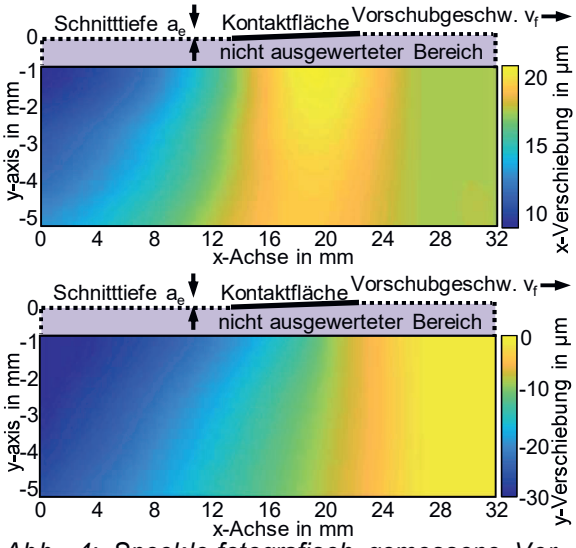


Abb. 4: Speckle-fotografisch gemessene Verschiebungsfelder, jeweils in x- und y-Richtung

Die gemessenen Verschiebungen decken sich im Vergleich (bis auf 20 % Abweichung) mit einem simulierten Schleifprozess bei gleichem Energieeintrag [10]. In der Simulation wurde der trockene Schleifprozess allerdings dadurch angenähert, dass sich eine konstante Wärmequelle über das Werkstück bewegt.

Aus den x- und y-Verschiebungen wurde gemäß Gln. (1)-(6) die Dehnung in der Ebene der Stirnfläche des Werkstücks berechnet. Im oberen Teil von Abbildung 5 ist die Dehnung in x-Richtung aufgetragen. Das Minimum der x-Dehnung liegt ungefähr 2-3 mm vor der Schleifscheibe und hat einen Wert von $-1{,}2 \times 10^{-3}$ mm/mm. Direkt hinter der Kontaktzone der Schleifscheibe ist das Maximum der x-Dehnung zu finden, es besitzt einen Spitzenwert von $1{,}5 \times 10^{-3}$ mm/mm.

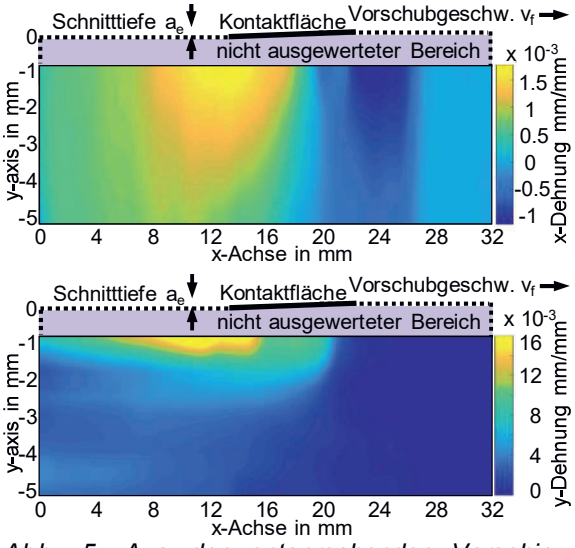


Abb. 5: Aus den entsprechenden Verschiebungsfeldern berechnete Dehnungen in x- und y-Richtung

Im unteren Teil von Abbildung 5 ist das Dehnungsfeld für die y-Richtung gezeigt. Das Maximum der y-Dehnung liegt am Ende der Kontaktfläche und steigt bis auf Werte von 16 x 10^{-3} mm/mm an.
Auch die gemessenen Dehnungen stehen mit einer Abweichung von weniger als 20 % in guter Übereinstimmung zu den Simulationen der Literatur [10].
Die Speckle-Fotografie ist also hinreichend robust, um auch unter den anspruchsvollen Bedingungen des trockenen Flachschleifens mit starkem Funkenflug In-Prozess-Messungen zu ermöglichen.

Speckle-Fotografie-Messungen unter Einsatz von Kühlschmierstoff

Insbesondere der Einsatz unter realitätsnahen Bedingungen, wie bei der Verwendung von Kühlschmierstoffen, erfordert neue Mess- und Auswertekonzepte der Speckle-Fotografie. In der Regel werden in der Fertigungstechnik bearbeitete Oberflächen mit Kühl- oder Schmiermitteln kontaminiert. Auf mikroskopischer Ebene verändern sich die Schichtdicken dieser Kontaminationsfilme permanent, so dass selbst bei transparenten Fluiden keine statischen, stabilen Specklemuster zu beobachten sind. Kann die Variation der Schichtdicke verhindert werden, so lässt sich eine Dekorrelation der Speckle – zumindest bei transparenten, homogenen Fluiden – verhindern. Eine Auswertung mit Verformungs- und Dehnungsanalyse ist somit möglich.
Diesbezüglich zeigen erste Vorstudien, dass Messungen auch durch begrenzte Flüssigkeitsfilme möglich sind. In einem Laboraufbau (siehe Abbildung 6) kann die Probe variabel in eine Petrischale mit Kühlschmierstoff getaucht und die über den x/y-Versteller eingestellte Verschiebung von unten speckle-fotografisch gemessen werden. Selbst ein Schmierstofffilm von h = 1,5 mm Höhe führt beim Vergleich von eingestellter und gemessener Verschiebung lediglich zu Messabweichungen von unter 5%. Voraussetzung ist allerdings eine begrenzende, transparente Fläche, durch die eine konstante Schichtdicke garantiert ist.
Die praktische Umsetzung dieses Laboraufbaus kann relativ einfach erfolgen, indem die Stirnfläche des Werkstücks durch eine Glasplatte abgedeckt und der Bearbeitungsprozess durch die Glasscheibe hindurch mit dem Messsystem beobachtet wird.
Der Spalt zwischen Glasplatte und Werkstück kann nach Abbildung 6 bis zu 1,5 mm betragen und sollte von unten kontinuierlich mit Kühlschmierstoff gespült werden. Partikel und Fremdkörper lassen sich so permanent mit dem Kühlschmierstoff noch oben aus dem Spalt herausdrücken, wobei durch die über das Werkstück heraustretende Flüssigkeit gleichzeitig die Kühlung der Oberfläche erfolgt.
Prinzipiell ist die im trockenen Schleifprozess erprobte Speckle-Fotografie somit auch geeignet, um Fertigungsverfahren unter realitätsnahen Bedingungen zu untersuchen und somit einen Beitrag zum Verständnis der Prozesse, bzw. zum Aufstellen einer Prozesssignatur zu liefern.

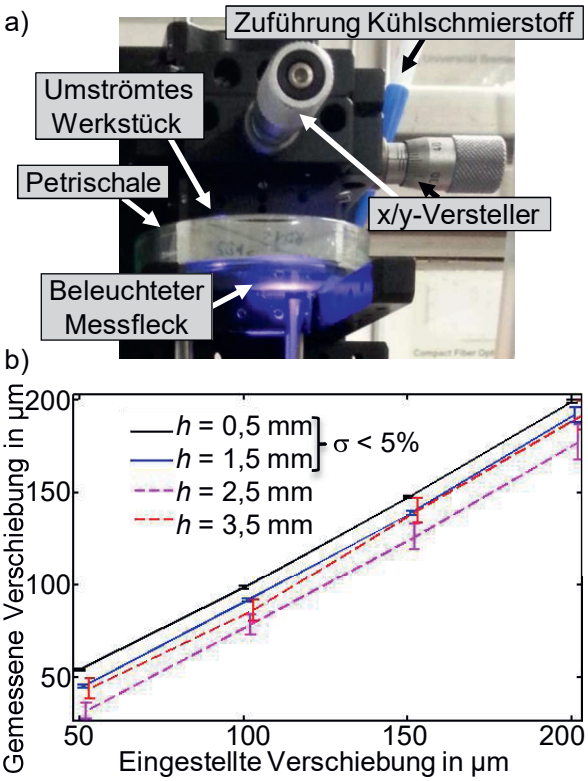


Abb. 6: Speckle-Fotografie-Messungen an einem in Öl getauchten Werkstück (a). Gemessene Verschiebung in Abhängigkeit vom eingestellten Wert und der Höhe des Flüssigkeitsfilms h (b).

Zusammenfassung und Ausblick

In dem vorliegenden Beitrag werden während eines Schleifprozesses prozessbegleitend Verschiebungs- und Dehnungsfelder optisch gemessen. In dem zunächst betrachteten trockenen Schleifprozess ohne Kühlschmierstoff entsteht starker Funken- bzw. Partikelflug, welcher sich mit Schnittgeschwindigkeit durch das Sichtfeld des Messsystems bewegt. Trotz dieser anspruchsvollen Bedingungen mit störendem Funkenflug sind, mit entsprechend angepassten Auswerteverfahren der Speckle-Fotografie, In-Prozess-Dehnungsmessungen möglich.
Die Dehnungsanalysen zeigen, dass der Energieeintrag sehr lokal erfolgt. Da die Ergebnisse nur um 20 % von Simulationen unter Annahme

einer konstant bewegten Wärmequelle [10] differieren, sind die eingebrachten Verformungen und Dehnungen im trockenen Schleifprozess offensichtlich in erster Linie thermischer Natur.
Labormessreihen durch einen begrenzten Kühlschmierstofffilm verdeutlichen den Einfluss der Schichtdicke des Fluids auf die Speckle-Fotografie-Messung. Selbst bei begrenzten Schichtdicken von 1,5 mm steigt die Messabweichung auf lediglich 1,5 %. Die Ergebnisse geben somit Anlass, den Einsatz der Speckle-Fotografie auch auf Fertigungsprozesse unter realitätsnahen Bedingungen und der Verwendung von Kühlschmierstoffen zu erweitern.
Daher wird aktuell ein Messaufbau für das Flachschleifen implementiert, bei dem der Spalt zwischen Werkstück und Messsystem bzw. einer begrenzenden Glasscheibe vollständig mit Kühlschmierstoff gefüllt ist.
Nach der erfolgreichen Messung der Beanspruchung und der im Werkstück verbleibenden Modifikation unter dem Einsatz von Kühlschmierstoffen, lässt sich die Prozesssignatur für dieses Fertigungsverfahren aufstellen.
Weiterhin kann das Messsystem zukünftig an die Kinematiken weiterer Fertigungsprozesse angepasst bzw. entsprechend erweitert werden, um einen Vergleich von verschiedener Fertigungsverfahren und eine Prozessübergreifende Beschreibung mittels Prozesssignatur zu realisieren.

Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft (DFG) für die Förderung des Transregionalen Sonderforschungsbereichs SFB/TRR 136 "Prozesssignaturen", insbesondere des Teilprojekts C06. Sie danken auch Ewald Kohls (Teilprojekt F01) und seinen Kollegen aus dem Leibniz-Institut für Werkstofforientierte Technologien (IWT) für die freundliche Unterstützung bei der Durchführung der In-Prozess-Messungen.

Literaturnachweis

[1] Archold, E.; Burch, J.M.; Ennos, A.E.: Recording of in-plane surface displacements by double-exposure speckle photography. Opt. Acta 1970, 17, 12, S. 883-898.
[2] Stetson, K.A.: A Review of Speckle Photography and Interferometry. Opt. Eng. 1975, 14, 5, S. 482-489.
[3] Fischer, A.: Fundamental uncertainty limit for speckle displacement measurements. Applied Optics 2017, 56, S.7013-7019.
[4] Tausendfreund, A.; Stöbener, D.; Dumstorff, G.; Sarma, M.; Heinzel, C.; Lang, W.; Goch, G.: Systems for locally resolved measurements of physical loads in manufacturing processes. CIRP Annals - Manufacturing Technology2015, 64, 1, S. 495-498.
[5] Tausendfreund, A.; Stöbener, D.; Fischer, A.: Precise in-process strain measurements for the investigation of surface modification mechanisms. Journal of Manufacturing and Materials Processing 2018, 2(1), 9 (11 S.).
[6] Zhou, P.; Goodson, K.E.: Subpixel displacement and deformation gradient measurement using digital image/speckle correlation (DISC). Optical Engineering 2001, 40, 8, S. 1613-1620.
[7] Kajberg, J.; Lindkvist, G.: Characterisation of materials subjected to large strains by inverse modelling based on in-plane displacement fields. International Journal of Solids and Structures 2004, 41, S. 3439-3459.
[8] Eman, J.; Sundin, K.G.; Oldenburg, M.: Spatially resolved observations of strain fields at necking and fracture of anisotropic hardened steel sheet material. International Journal of Solids and Structures 2009, 46, S. 2750-2756.
[9] Tausendfreund, A.; Stöbener, D.; Goch, G.: Measuring technique for the detection of process force related component deformations in metalworking manufacturing processes. In: Zoch, H. W.; Lübben, T. (Hrsg.): 5th International Conference on Distortion Engineering (IDE), Bremen 2015, S. 355-364.
[10] Tausendfreund, A.; Borchers, F.; Kohls, E.; Kuschel, S.; Stöbener, D.; Heinzel, C.; Fischer, A.: Investigations on material loads during grinding by speckle photography. Journal of Manufacturing and Materials Processing 2018, 2(4), 71 (12 S.).

DOI 10.5162/sensoren2019/4.1.2

An Optical Sensor System for Characterization of Ceramics based on Laser Speckle Photometry

Lili Chen[1], Ulana Cikalova[1], Beatrice Bendjus[1], Stefan Muench[1], Mike Roellig[1]
[1]Fraunhofer Institute for Ceramic Technologies and Systems IKTS, Maria-Reiche-Straße 2, 01109 Dresden, Germany

Abstract

Advanced ceramic components are widely used in the industrial applications, e.g. substrates of power electronic modules, and electrolytes of new generation batteries. As a brittle material, the crack propagation is very sensitive to the stress distribution of ceramics. Existing defects under extreme stress condition will lead to the functional failures of the whole component. In order to avoid the crack initiation and propagation, and guarantee a good quality state of ceramic products, stress condition and defects should be monitored and detected.
In this paper, an optical non-destructive testing method – Laser Speckle Photometry (LSP) will be demonstrated to determine mechanical stresses in ceramic substrates and detect defects in ceramic products for the industrial applications. The LSP method is based on the analysis of a time resolved speckle dynamics which is activated by an external excitation. The speckle signal can be correlated to the external introduced signal and further be related to our targeting parameters – mechanical stress and existing defects in the ceramics. The preliminary results of LSP has shown its great potential for non-destructive characterization of ceramics in terms of stress monitoring and defect detection.

Keywords: Optical sensor system, Laser speckle photometry, Non-destructive testing, Strain/stress characterization, defects detection in ceramics

Introduction

Due to the special structural and functional properties, the advanced ceramic technology plays an important role in many fields of the industry. For example, direct copper bonding (DCB) substrates are widely used in electronic applications because of its high thermal conductivity performance, and Na-ß-Alumina can be utilized to produce high-temperature, low-cost batteries. During the manufacturing process or under the service loading, high thermal or mechanical stress would be occurred. Extreme stress condition might lead to the crack initiations. The propagation of cracks in ceramic components will result in security risks and great financial losses for manufacturers and customers. In order to ensure a continuously high quality of ceramic products in the industrial applications, it is important to know exactly the stress distribution and be able to detect the existing defects on ceramics. Currently, there is no technique that allows direct, non-contact and fast determination of mechanical stresses in ceramics. Techniques such as X-ray diffraction method, the digital speckle pattern interferometry, and ultrasonic scanning microscopy are time-consuming, very expensive and sometimes destructive. For the defect detection of ceramics are common used methods like X-ray transmission and backscattering techniques, ultrasound reflectance, with the aim of detecting defects in near surface zone as well as inner zone. These procedures are usually complicated, and the sample are individually scanned by laboratory equipment. It is difficult to integrate the devices into production process. However, there is a high economic demand for the integration of inspection setups into production lines in order to locally verify important quality parameters and to detect defects as well. Component failures in the application will result in security risks and great financial losses for manufacturers and customers.
In this case, the authors propose a novel non-destructive and non-contact optical technique – Laser Speckle Photometry (LSP) for the characterizations of ceramics including stress monitoring and defect detection in a fast and cost effective way.

Laser Speckle Photometry

A speckle pattern is generated when an optical rough surface illuminated by a coherent light source, e.g. laser [1]. The scattered waves from various points of the illuminated surface

interfere in the observation plane, generating the speckle pattern, i.e. a spatial structure with randomly distributed intensity minima and maxima. Different luminosities can be recorded by a CCD and CMOS chip respectively. The theory of speckle or dynamic speckle belongs to the category of statistical optics of which a detailed account of phenomena is given by Goodman [2] including the mathematical description of speckle properties, for instance the central limit theorem of probability theory, random processes. The LSP method is based on the analysis of speckle patterns generated by a laser source. Actually the recorded speckle patterns are not only 2-D graphics, but also the fingerprint containing the 3-D information of sample surface. The basic experimental setup of LSP contains three parts – an illumination source, an external excitation source and a detector, as shown in Fig. 1. A laser diode as the illumination source generate a static speckle pattern on sample surface. Speckles can be activated by an external thermal or mechanical excitation to yield a time resolved dynamic speckle pattern, which can be also called speckle movement. Thanks to the development of hardware and mathematical tools for registration and processing of non-stationary optical fields, the time resolved speckle patterns can be recorded by a digital camera and be evaluated with different algorithms for different specific applications.

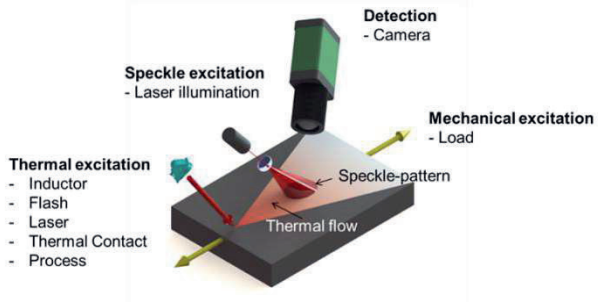


Fig. 1: Schematic structure of Laser Speckle Photometry.

Generally, the LSP technique can be used for characterization of a material's defects [3]. It provides a solution of fast in-line defect inspection in industrial field. Much work has been done by LSP method for metallic materials including the characterization of material properties, for example, damage, hardness and porosity [3-5]. Since ceramic materials have total different thermal and mechanical properties compared with metals, characterization of ceramics based on LSP show its special research values in both academic and industrial fields.

Stress characterization of ceramics by LSP

In the research of stress characterization, the strain field was introduced by mechanical tension with a 3-point-bending device. Rough surfaces of ceramic samples respond to this stimulation with changing the position of surface reflectors respectively scattering points and converting the variations into changes of speckle pattern recorded by the camera system in form of video sequences.

In order to connect the speckle signal to mechanical signal, a calibration function of speckle and stress should be established. A concept of the calibration was developed which describes the correlation between both variables, and the details were presented in [6]. Briefly, in one hand, the process of time-resolved dynamic speckle is led by surface strain changes introduced by the bending test. In the other hand, the behavior between mechanical stress and strain can be described by the Hooke's law in the elastic region and for small stretches. Therefore, strain is a transfer parameter connecting the speckle and mechanical signals. Since the temperature of sample was controlled in a constant during the measurement, the surface change caused by the thermal properties (e.g. thermal expansion) was not taken into consideration any more.

As mentioned above the LSP signal is dependent on the surface strain of sample, in that case, the mechanical strain can be denoted by the speckle signal $S_{LSP}(t)$, the calibration function is given by:

$$\sigma(t) = P_1 * S_{LSP}(t) + P_2 \qquad (1)$$

where P_1 and P_2 should be determined by the experimental calibration test. Because the initial state of sample was defined to be stressless, the parameter P_2 was set to zero. P_1 could be estimated by linear fitting of mechanical stress computed by finite element method (FEM) simulation from the calibration measurement and the LSP signal $S_{LSP}(t)$. After the determination of parameter P_1, the real stress condition of ceramic sample can be calculated directly from LSP signal based on the calibration function (1) [6].

Defect detection of ceramics by LSP

As it is mentioned above, change of surface strain can activate dynamic speckle patterns. And the strain change could also be produced by the change of temperature, i.e. thermal expansion. When ceramic sample is illuminated by an external thermal source, e.g. a high power laser, the thermal energy will be absorbed by the sample, and it will propagate on the sample surface, which results in the local temperature differences. The change of temperature causes

strain change on sample surface activating time resolved dynamic speckle patterns, which is shown in Fig. 2.

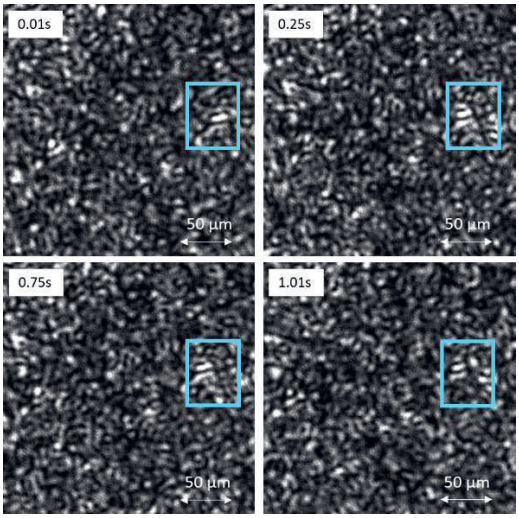


Fig. 2: Time-resolved speckle patterns.

When thermal waves propagate on an un-defective sample surface, a homogeneous spatial gradient of temperature will be formed. If surface defects exist, the local magnitude of the thermal propagation will be amplified, i.e. the presence of defects change distribution of thermal energy around the defective area [7]. To be exact, the mentioned local change of heat distribution results in the thermal energy accumulating around the defective area, as presented in Fig. 3.

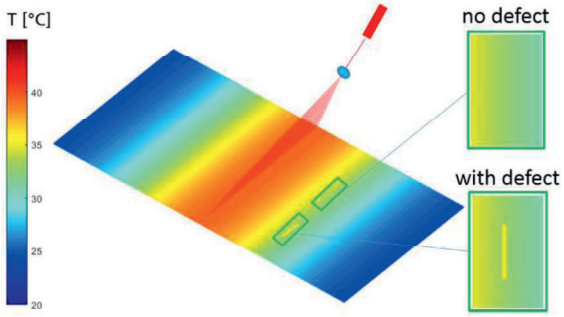


Fig. 3: Simulation of thermal propagation on a defective surface.

The thermal accumulation distorts the normal spatial gradient and leads to a higher level of speckle movement at the border of defects compared with the surroundings. As a result, a contrast of dynamic speckle movement between defective area and background could be obtained. On the basis of special algorithms, dynamic speckle signal can be evaluated and analyzed to extract the contrast, which provides the possibility for defect detection by the technique of LSP.

Unlike other speckle based inspection techniques, for example, speckle interferometry and shearography, the LSP works without reference beams. This allows a simple but robust construction, and last but not least, measurement and evaluation of obtained data could be performed in real time. Therefore, it can be easily integrated into the manufacture process to implement the in-line inspection.

Algorithms

There are two basic evaluation methods of LSP, static and quasi-static algorithms, to evaluate speckle signals in the different situations. The static algorithm focuses on the image scale, which uses the gray value related parameters combined with statistical properties of speckle recordings. Parameters, for instance, mean value of intensity, speckle contrast, speckle size and texture properties calculated from the grey-level co-occurrence matrix (GLCM) [8, 9] can be utilized to present speckle signals in LSP technique. In the quasi-static evaluation algorithms, the differences of gray value distribution of a single pixel in time sequence rather than image scale are interested. In our research of characterization of ceramics including stress characterization and defect detection, quasi-static algorithms were used to calculate the speckle parameters for final results.

For the purpose of stress characterization, an algorithm based on correlation method was proposed. In the field of image processing, the correlation can be explained by another word, i.e. similarity. A high correlation means a great similarity of two different signals. In [10], strain information of a brass plate in tensile test was calculated by one dimensional speckle correlation. And microscopic deformations of the uniform tensile test specimen were measured by digital speckle correlation in [11]. Taking into account the characteristics of the speckle pattern generated by a laser source on the surface of ceramic materials, a parameter based on correlation techniques was selected to present speckle signals in the correlation function. The parameter is 2-D correlation coefficient whose mathematical expression is shown in equation (2).

$$r(t) = \frac{\sum_i \sum_j (I_{ij}(t_0) - \overline{I_0})(I_{ij}(t) - \overline{I_t})}{\sqrt{(\sum_i \sum_j (I_{ij}(t_0) - \overline{I_0})^2)(\sum_i \sum_j (I_{ij}(t) - \overline{I_t})^2)}} \quad (2)$$

The coefficient $r(t)$ was renamed by correlation coefficient of speckles (CCoS), which presents the similarity between two recorded speckle images. In equation (2), $I_{ij}(t_0)$ is gray value of pixel (i, j) in speckle image at the stressless state, which is the first frame of speckle recording. And $\overline{I_0}$ means the average gray value of all pixels in the first frame. Analogously,

$I_{ij}(t)$ and $\overline{I_t}$ present the intensity of pixel (i,j) in speckle images at different stress states and the average intensity of corresponding frames. In one speckle recording, the value of CCoS demonstrates the change of time-resolved speckle patterns in a specific frame compared with the first frame, which reveals the changes of stress in tested samples.

In the research of defect detection, the algorithm was based on a time-resolved analysis shown in Fig. 4. The intensity of each pixel was processed in the time domain. During the movement of speckle patterns, the time-varied intensity can be related to the spatial gradient of temperature with a special correlation function. It has been proven in the research of Kazak [12] that the value of the correlation function was proportional to the local temperature of a certain region of interest. As distinguished from the CCoS mentioned above, the correlation function used in the defect detection calculated the intensity differences of pixels in the speckle recordings. And it is given by:

$$C(\tau) = \langle I(1+\tau,x,y) - I(1,x,y)\rangle^2 \qquad (3)$$

where τ is the time shift or the number of frame intervals, $I(n,x,y)$ is the intensity of pixel whose location is determined by the coordinates x and y in the n-th frame. This correlation function is modified from the so-called semivariogram, which is a geostatistical tool for studying the relationship between collected data as a function of distance and direction [13].

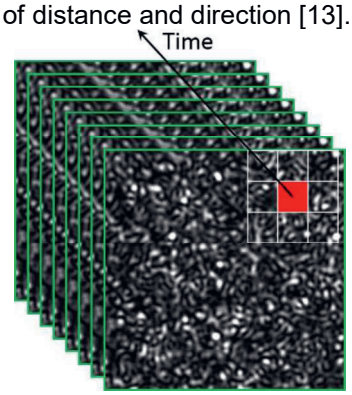


Fig. 4: Schematic of the time-resolved image analysis.

The relationship between the value of correlation function and the change of local temperature is shown in Fig. 5. From the diagram, it can be concluded that before saturation point, the numerical value of correlation function exhibits similar behavior to that of the local temperature, which is entirely consistent with the conclusion of Kazak mentioned above.

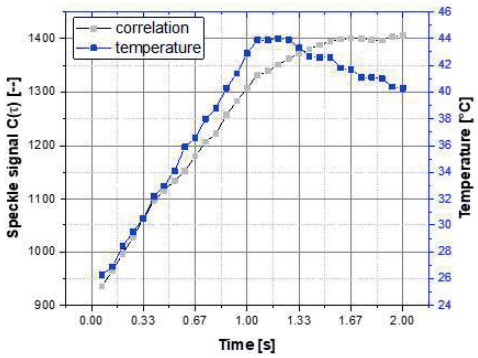


Fig. 5: relationship between correlation function and local temperature.

As it is known, an image can be demonstrated by a two-dimensional matrix in the digital processing. Similarly, the video file recording the speckle signals can be treated as a three-dimensional matrix. The time-varied intensity of each pixel is a dimensionless signal. Since the thermal energy would be accumulated around defective area, the dynamic speckle movement is stronger than that of background. Therefore, frequency analysis could be utilized to extract the contrast of dynamic speckle signals between defective area and background. After the calculation of correlation function for each pixel, the Short Time Fourier Transformation (STFT) or Wavelet Transformation (WT) is applied to convert the dimensionless signal into time-frequency domain [14, 15]. In this case, the amplitudes of STFT or WT for all pixels at a specific time-frequency point can be selected to build the resulting image.

Experiments and results

Stress characterization of ceramics

The experimental setup of calibration measurements combined the optical setup for LSP recordings and the mechanical configuration for stress and strain introduction. The setup is shown in Fig. 6. The speckle patterns were generated on sample surface by a single mode laser diode with a wavelength of 650 nm. The diameter of total illumination area was approximately 8 mm. Dynamic speckle signals were recorded by a CMOS digital camera, whose full resolution of the chip was 2048*1088 pixels. A macro lens with 8 times magnification was mounted on the camera. The beam splitter provided the possibility to make both camera and laser diode be perpendicular to sample surface. In this case, during the bending test there is no out-of-plan motion at the measuring position. It ensured that the dynamic speckle patterns were generated only by the change of surface strain, which reduced the noise coming from the setup in final results.

Samples were fixed in a 3-point bending device. The distance between the outer loading pins was 40 mm. The samples under test were a thin Al_2O_3 plate (45*10*0.62 mm^3) and a thin low temperature co-fired ceramic (LTCC) plate (45*10*0.67 mm^3). The measurements occurred at the center of sample surfaces. During the bending measurement, surface strain was introduced into sample via the displacement of loading pins of the device, which was under controlled. The speckle recordings were evaluated by means of the software MATLAB®. Meanwhile the mechanical stress at the measuring position was estimated by FEM simulations.

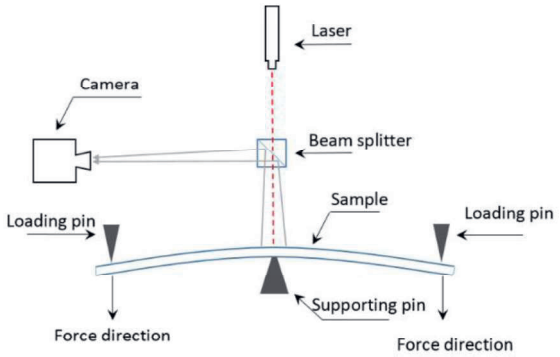


Fig. 6: Experimental setup of LSP for stress characterization.

The preliminary results of calibration measurements are shown in Fig. 7 (a-c). The data were obtained from a 200*200 μm^2 area on sample surface. The stress was estimated from FEM simulation, and the speckle parameter CCoS was calculated based on a quasi-static algorithm, i.e. equation (2) mentioned above. The measurements were repeated 4 times at same position for each sample. Both variables were plotted against the time of measurement in the diagrams.

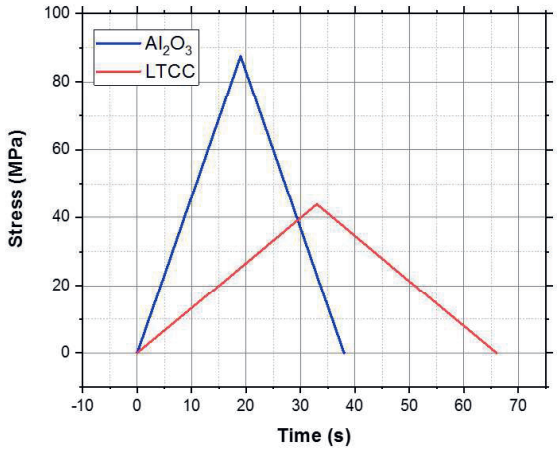


Fig. 7(a): Stress signals calculated from FEM.

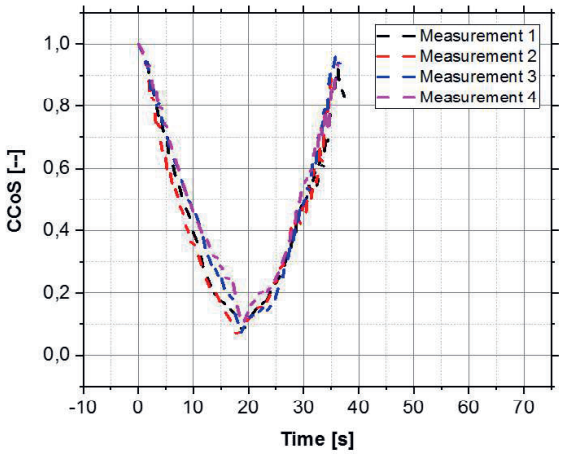


Fig. 7(b): Resulting speckle signal of Al_2O_3 sample calculated by equation (2).

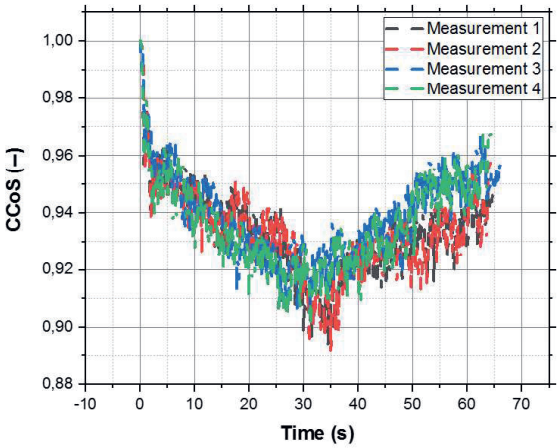


Fig. 7(c): Resulting speckle signal of LTCC sample calculated by equation (2).

From the diagrams, it can be seen that the speckle parameter CCoS has a linear relationship with the mechanical stress introduced by the bending device for both Al_2O_3 and LTCC samples in the elastic region. The speckle signals changed direction while the loading condition of bending device turned to unloading. Since the CCoS assesses the similarities of speckle patterns at different stress states, if the speckle patterns change too much, it will lose the correlation to the initial state. In this case, there is a measuring limitation ranging from 0 to circa 90 MPa of Al_2O_3 samples at this moment. For LTCC samples, stress differences up to 45 MPa are detectable. However, the measuring range of LTCC is mainly dependent on the strength of the material rather than the correlation range of speckle signal compared with Al_2O_3 samples. This phenomenon is caused by different scattering models of these ceramics. In terms of LTCC, it is opaque to the laser at 650 nm and speckle patterns generated on its surface is determined by surface scattering. On the

contrary, Al_2O_3 is partial transparent to the laser source, and therefore, it has volume scatterings. As we all know, speckle pattern is a result of interference of coherent sources, and the interference is controlled by the differences of optical paths. In surface scattering, surface height fluctuation modulates the optical paths, while in volume scattering it is dominated by the average distribution of path lengths [16], which is caused by the scattering on grain boundaries in the volume. In that case, under the same surface strain condition, the speckle patterns generated by surface scattering changes less than that of volume scattering. From the observations of our experiments, the speckle patterns on the surface of LTCC didn't change so much while the patterns on Al_2O_3 changed rapidly and then lost the similarity compared to the speckle pattern at initial stress state. In additional, the noise caused by the camera is around 2% of CCoS, i.e. 0.02 in numeric value, and it causes the fluctuations in the Fig. 7(c).

Defect detection of ceramics

The research measurements for defect detection were carried out based on a thermal excitation method. The experimental setup is shown in Fig. 8. Static speckle patterns were generated on sample surface by a laser diode with the wavelength of 650 nm. A camera system including a CMOS chip and an objective worked as a detector to record speckle signal. The camera was focused on sample surface so that the near field subjective speckle patterns were captured. Near field speckle have an interesting feature that their statistical properties are closely related to the structure of the scattering object [17], which means it can be connected to the structure properties of sample surface to provide information of defects for the experiments. The angle between laser diode and camera was approximately 30°. A high power laser with a wavelength of 808 nm combined with a reflective collimator was used to provide a thermal pulse to excite the sample, which produced dynamic speckle signals. And the pyrometer was a temperature detector. The samples under detection were a TiO_2 tube and a Na-ß-Alumina tube. The defects e.g. holes and pores were introduced into sample surface during the manufacturing process. In our measurements, the samples were excited by a 10 W laser pulse, and the pulse length was around 0.5 s.

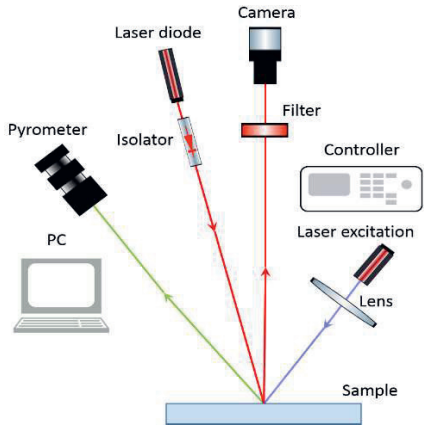


Fig. 8: Experimental setup of LSP for defect detection.

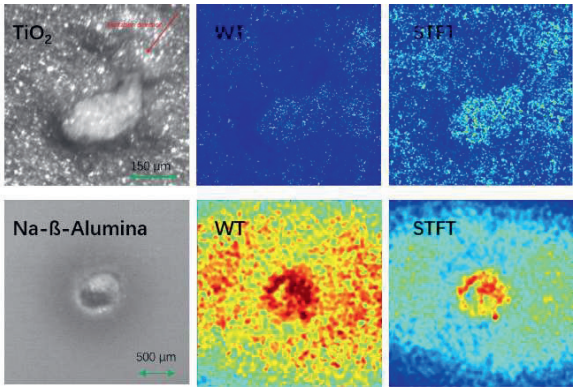


Fig. 9: Resulting images of defect detection. Top: results of TiO_2. Bottom: results of Na-ß-Alumina

The preliminary results of current research are shown in Fig. 9. The middle and right parts in Fig. 9 are the results of detection calculated from WT and STFT respectively. And the left parts are the corresponding optical images showing the defects. From the resulting images, it can be observed that the defected are detected by the presence of features with enhanced brightness. And the features show the position and approximate dimension of which correspond well with the defects observed directly in the optical images.

In general, the ideal detectability of LSP is mainly determined by the sensor size of the digital camera, sin the sensor size is far greater than the wavelength of laser source. In reality, at current research the LSP method is constrained by special thermal properties of ceramics, for example, absorption coefficient, thermal diffusivity and coefficient of thermal expansion. These parameters influence the speckle recordings. Extremely, if a material has a small absorption coefficient and a small thermal expansion coefficient, the dynamic speckle movement is difficult to be activated by an external thermal source. In that case, the

defects cannot be detected due to the lack of speckle information. Meanwhile, the type of defects is also an important factor affecting the detectability. Normally, the LSP technique has a relative higher detectability to sharp defects, because sharp edges give prominence to the effect of thermal accumulation mentioned above. In summary, the detectability of LSP in ceramics is approximately 300 µm for crack in length and 200 µm for pore and hole in diameter after repeating the measurements with a variety of defective samples in the current research.

Conclusion

This paper presents an optical sensor system based on the LSP technique for the characterization of ceramics including stress measurement and defect detection. The experimental setups and corresponding evaluation algorithms are proposed. The preliminary results show the possibility for the measurement of stress condition on the surface of ceramic sheet by using specific speckle parameter CCoS. Moreover, LSP has a simple but robust design, reducing cost compared with other conventional optical non-destructive testing methods. And there is a great potential for the LSP technique to develop its detectability in ceramic materials.

Acknowledgments

The authors would like to thank Federal Ministry for Economic Affairs and Energy for the support within “Zentrales Innovationsprogramm Mittelstand – ZIM – Kooperationsnetzwerke” for the project “SubStress”, and Fraunhofer-Gesellschaft for providing the opportunity to participate in the research of Bundesministerium für Wirtschaft und Energie, IGF Forschung project “Entwicklung eines optischen Inline-Verfahrens zur zerstörungsfreien Prüfung keramischer Hochleistungsbauteile”. Meanwhile, the authors would like to show the appreciations to all the colleagues in the institutions for the assistance in the investigations.

Reference

[1] Joseph W. Goodman, [Laser Speckle and Related Phenomena], Springer, New York, chapter 2, 9-75, (1975).
[2] Joseph W. Goodman, [Statistical Optics], Wiley, New York (1985).
[3] U. Cikalova, B. Bendjus, and J. Schreiber. "Laser-Speckle-photometry–A method for non-contact evaluation of material damage, hardness and porosity." *Materials Testing 54.2 (2012): 80-84*; doi: 10.3139/120.111308
[4] U. Cikalova, B. Bendjus, and J. Schreiber, “Material Characterization by Laser Speckle Photometry”, *Speckle 2012: V International Conference on Speckle Metrology (2012)*; doi: 10.1117/12.978246
[5] U. Cikalova, B. Bendjus, and J. Schreiber, "Laser speckle photometry: contactless nondestructive testing technique." *Speckle 2012: V International Conference on Speckle Metrology*. Vol. 8413. International Society for Optics and Photonics, 2012; doi: 10.1117/12.978247
[6] S. Muench, et al. "A laser speckle photometry based non-destructive method for measuring stress conditions in direct-copper-bonded ceramics for power electronic application." *2017 18th International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Microsystems (EuroSimE)*. IEEE, 2017; doi: 10.1109/EuroSimE.2017.7926273
[7] D. Y. Tzou, [Macro-to Microscale Heat Transfer: The Lagging Behavior], Second ed. *Wiley*, 282-288 (2014).
[8] Hector J. Rabal, and Roberto A. Braga Jr, eds. [Dynamic laser speckle and applications]. *CRC press*, 2008.
[9] Robert M. Haralick, "Statistical and structural approaches to texture." *Proceedings of the IEEE 67.5* (1979): 786-804; doi: 10.1109/PROC.1979.11328
[10] Ichirou Yamaguchi. "A laser-speckle strain gauge." *Journal of Physics E: Scientific Instruments* 14.11 (1981): 1270; doi: 10.1088/0022-3735/14/11/012
[11] Zili Sun, Jed S. Lyons, and Stephen R. McNeill. "Measuring microscopic deformations with digital image correlation." *Optics and Lasers in Engineering* 27.4 (1997): 409-428; doi: 10.1016/S0143-8166(96)00041-3
[12] N. S. Kazak, et al, "Speckle-photometry method of measuring the coefficient of thermal diffusion", *Proc. International Conference Optical Techniques and Nano-Tools for Material and Life Sciences*, 1, 61 – 72 (2010).
[13] A. Mazzella, A. Mazzella, "The importance of the model choice for experimental semivariogram modeling and its consequence in evaluation process", *Journal of Engineering, 2013*, 1 – 10 (2013); doi: 10.1155/2013/960105
[14] M. A. Namuq, M. Reich, and S. Bernstein, (2013). "Continuous wavelet transformation: A novel approach for better detection of mud pulses." *Journal of Petroleum Science and Engineering*, 110, 232–242. doi: 10.1016/j.petrol.2013.08.052
[15] R. X. Gao, and R. Yan, (2006). "Non-stationary signal processing for bearing health monitoring" *International Journal of Manufacturing Research*, 1(1), 18. doi: 10.1504/ijmr.2006.010701
[16] Joseph W, Goodman, [Speckle phenomena in optics: theory and applications]. *Roberts and Company Publishers*, 2007.
[17] M. Giglio, et al (2001). "Near-field intensity correlations of scattered light." *Applied Optics*, 40(24), 4036. doi: 10.1364/ao.40.004036

DOI 10.5162/sensoren2019/4.1.3

IMPACTSCHÄDEN SCHNELL UND EINFACH FINDEN

M. Sc. J. Rittmann[1], Dr.-Ing. M. Rahammer[1], Dr.-Ing. N. Holtmann[2], und Prof. Dr. rer. nat. habil. M. Kreutzbruck[1]
[1]IKT, Institut für Kunststofftechnik der Universität Stuttgart, Pfaffenwaldring 32, 70569 Stuttgart
[2]edevis GmbH, Handwerkstraße 55, 70565 Stuttgart
Tel.: +49 711 685-62887, E-Mail: johannes.rittmann@ikt.uni-stuttgart.de

Zusammenfassung

Impactschäden in Faserkunststoffverbunden wie CFK und GFK stellen hohe Herausforderungen an die zerstörungsfreie Prüftechnik (ZfP). Die anisotrope Werkstoffstruktur erschwert die Ergebnisinterpretation herkömmlicher Prüfverfahren erheblich. Die Resonante Frequenzsweep Thermografie (RFST) basiert auf dem Prinzip der lokalen Defektresonanz in Verbindung mit der bekannten Ultraschall-angeregten Thermografie. Mit ihr ist es möglich, in kürzester Zeit auf einfache Art und Weise die Existenz von signifikanten Impactschäden festzustellen. RFST nutzt eine Frequenzsweep-Anregung im unteren und mittleren Kilohertzbereich, um Defekt-, Anregungs- und Bauteilresonanzen zu aktivieren. Resonanzen verstärken die akustische Aktivierung der Defekte um ein Vielfaches und führen zu einer signifikanten Steigerung des thermischen Signals. Dieses basiert auf Rissuferreibung sowie visko-elastischer Erwärmung und wird an der Oberfläche von einer Thermografiekamera detektiert. Die Defekterkennungsschwelle hängt dabei von der Anregungsleistung sowie dem Abstand der Defekte zur Schallquelle ab. Zu diesem Zweck wurde ein Prototyp in Form eines Dreibeins mit integrierter IR-Kamera und Ultraschallanregung entwickelt. Dieser zeichnet sich durch einfache Handhabung und flexible Einsatzmöglichkeiten aus. Mithilfe von Augmented Reality wird dem Prüfer die Ergebnisinterpretation und Defektmarkierung erleichtert, indem das ausgewertete Prüfergebnis ortsgetreu auf das Bauteil projiziert wird. Vorgestellt werden erste Ergebnisse an einer Serie von Impactschäden in CFK mit variierenden Einschlagenergien und Delaminationsflächen. In diesen Bauteilen lassen sich für eine Prüfdauer von einer Minute Impactschäden mit projizierter Schadensfläche von mindestens ein Zoll Durchmesser sicher detektieren.

Keywords: Impactschäden, Faserkunststoffverbunde, Ultraschall Thermografie, Resonanz

Einleitung

Impactschäden sind kurzzeitige, punktförmige Belastungen senkrecht zur Bauteilebene. In dieser Lastrichtung sind Faserkunststoffverbunden (FKV) am anfälligsten, da dies keine primäre Faserorientierung ist. Oft genannte Ursachen sind fallengelassene Werkzeug („Tool Drop“), Hagel oder Vogelschlag („Bird Strike“). Eines der häufigsten Beispiele aus der Luftfahrt ist jedoch ein Zusammenstoß mit anderen Fahrzeugen während der Abfertigung („Ramp Crash“) [1]. Je nach Energie des Einschlags/Aufpralls kommt es zu Matrixbrüchen, Delamination oder gar Faserbrüchen und einem Durchdringen des Impaktors. Bereits relativ geringe Energien erzeugen Delaminationen, obwohl rein äußerlich kein Schaden entstanden zu sein scheint („Barely Visible Impact Damage“, BVID). Das Resultat der Spannungsverteilung bei Impactbelastung ist eine mit der Tiefe anwachsende Schadensfläche. Ein solches Schadensbild ist in Abb. 1 aus [2] abgebildet. Ähnlich einer Drei-Punkt-Biegung sind bei einem Impact Druck- bzw. Zugspannungen an den beiden Oberflächen maximal, während in der Mitte Schubspannungen maximal werden. Die Druckfestigkeit des Verbunds ist vergleichsweise hoch, weshalb es hier kaum zu Schäden kommt. Die Schubspannungen führen jedoch zu Matrixbrüchen (i), welche sich an Grenzflächen zu Delaminationen (ii) akkumulieren. Dieser Effekt wird nach unten hin durch die zunehmenden Zugspannungen zusätzlich begünstigt, weshalb das Schadenausmaß zunimmt. Aufgrund der Rotationssymmetrie der Last treten in den tiefsten Lagen noch zusätzlich Matrixrisse (iii) auf, welche parallel zur Faser verlaufen.

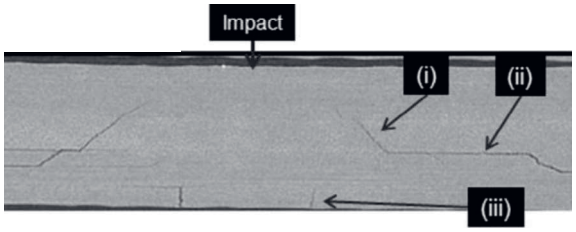


Abb. 1: Computertomografie-Schnittbild eines Impactschadens in CFK [2]

Von den aktuell bekannten Prüfverfahren eignen sich sowohl Ultraschallprüfung [3] als auch Computerthomografie [4] sehr gut für die Detektion von Impactschäden, aber nur bedingt für einen schnellen und einfachen mobilen Einsatz. Die normativ verankerten Verfahren sind entweder sehr aufwendig durchzuführen oder sehr kompliziert zu bewerten. Neben diesen beiden Standardverfahren etabliert sich die ZfP mittels Thermografie immer weiter. Hauptanwendungsgebiet der optisch angeregten Thermografie (OTT) für FKV ist die schnelle ZfP an großen, dünnen Flächen, wie sie z. B. in vielen Luftfahrtanwendungen oder bei Windkraftanlagen vorkommen. Defekte, die sich dabei sehr gut finden lassen, sind Delaminationen (z. B. aufgrund von Impact) [5], welche das Abfließen von Wärme verhindern.
Wenig bekannt ist hingegen die Ultraschall-angeregte Thermografie (UTT), welche im Gegensatz zu den meisten ZfP-Verfahren ein defekt-selektives Verfahren ist. Dies bedeutet, dass Defekte ein sehr viel stärkeres Signal erzeugen, als intakte Strukturen. Diese Selektivität hat mehrere Gründe. Erstens tritt Rissuferreibung überhaupt nur an Rissen oder rissähnlichen Defekten auf. Zweitens findet man Spannungsspitzen vornehmlich in der Nähe von geschädigten Bereichen, und drittens ist die Biegesteifigkeit durch einen Defekt lokal geschwächt und die Schwingung damit erhöht. Seit Mignogna et al. [6] ist es üblich, Ultraschallschweißsonotroden aus der Kunststoffschweißtechnik für die Anregung zu verwenden. Schweißsonotroden arbeiten bei einer fixen Resonanzfrequenz (meistens 20 oder 40 kHz), um im Anpresspunkt so viel Energie wie möglich einzukoppeln. Sehr viel weniger Energie ist jedoch notwendig, wenn man die Resonanzfrequenz des Bauteils [7] oder des Defektes [8, 9] zur Anregung nutzt.
Auf Basis dieser Entdeckungen entstand am IKT ein komplett neuer Prüfansatz basierend auf der Ultraschall-angeregten Thermografie [10], welcher bereits in einen ersten Prototypen überführt werden konnte.

Das RFST-Prüfsystem

Kernbausteine des Verfahrens sind die Anregungsquelle und die Thermografiekamera. Zur Anregung werden drei eigens für dieses Verfahren entwickelte Vakuum-Piezo-Shaker verwendet, welche sowohl mit Überdruck als auch Unterdruck betrieben werden können. Das erzeugte Vakuum dient dabei einerseits der Fixierung des Systems auf der zu prüfenden Struktur, als auch der Erzeugung eines Anpressdrucks für die Ankopplung der Shaker. Versorgt werden diese von einem Verstärker, ebenfalls isi-sys GmbH, mit einer Wechselspannung von 60 V Amplitude. Das System aus Verstärker und Shakern hat seine Eigenresonanz bzw. seine höchste Effizienz im Bereich von 15–25 kHz. Aus Gründen des Lärmschutzes wird ein Bereich von 18–25 kHz gewählt.
Aus Kosten- und Gewichtsgründen fällt die Wahl der Thermografiekamera auf ein Mikrobolometersystem. Solche Systeme sind am Markt bereits für unter 15 k€ erhältlich und liefern NETD von bis zu kleiner 30 mK sowie eine Bildauflösung von 640 x 512 Pixel. Im vorgestellten Prototypen (Abb. 2) wird eine FLIR A655sc eingesetzt, die zudem Thermogramme mit einer Wiederholungsrate von 50 Hz aufnehmen kann.
Weiterhin verfügt das System über einen integrierten Miniaturprojektor. Dieser dient dazu, das automatisch ausgewertete Prüfergebnis auf die Struktur zurück zu projizieren. Diese Augmented Reality Anwendung ermöglicht dem Prüfer eine schnelle und einfache Interpretation der Ergebnisse und erlaubt es den Defektbereich exakt zu kennzeichnen. Hierzu sind Kamerasichtfeld und Projektionsfläche optisch aufeinander abgestimmt.

Einfluss der Anregungsfrequenz

Die Defektresonanzfrequenzen von Impact-induzierten Schäden liegen oftmals über 100 kHz. Diese Frequenzen sind mit üblichen Verstärkern und Piezo-Shakern nicht effektiv erzeugbar. So konnten für die hier verwendeten Impactschäden z. B. LDR-Frequenzen bei 114 kHz und. 128 kHz identifiziert werden.
Wie bereits erwähnt lassen sich aber noch weitere optimale Frequenzen nutzen. Die elektrische Anregung z. B. hat ihr Optimum im Bereich um die 20,5 kHz, wie die blaue Messkurve in Abb. 3 zeigt. Für sehr viel höhere Frequenzen ist der genutzte Verstärker nicht ausgelegt. Der Piezo-Shaker selbst hat ebenfalls natürliche Eigenresonanzen. Bei diesen Frequenzen kann er die elektrische Energie besonders effizient in mechanische

Schwingung umsetzen. Diese sind 21,8, 23,7 und 25,4 kHz und ebenfalls in Abb. 3 dargestellt. Ein Vergleich der beiden Spektren zeigt, dass das kombinierte Optimum im Bereich 20–22 kHz liegt.

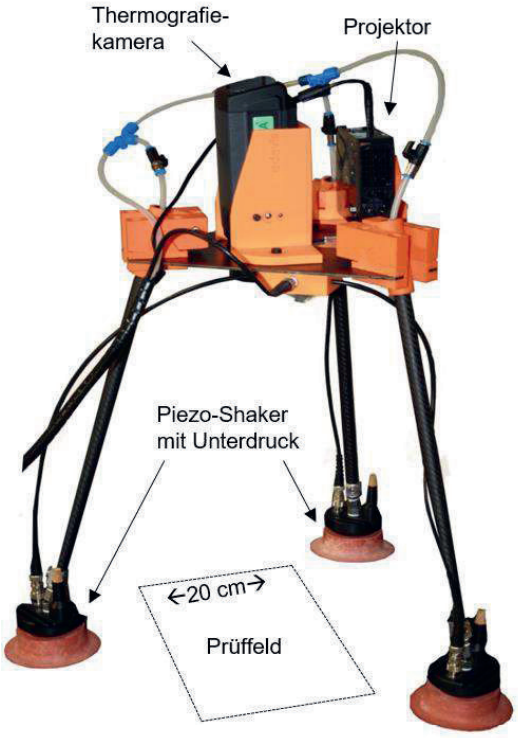


Abb. 2: Prototyp für ein mobiles RFST System

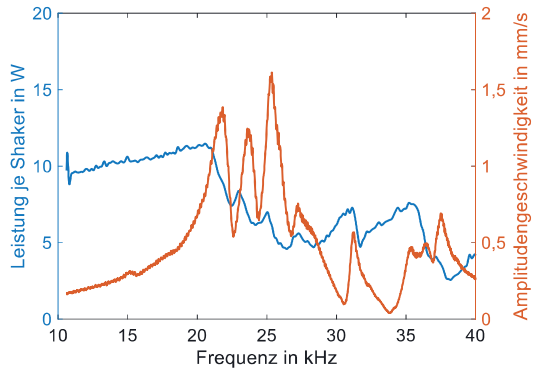


Abb. 3: Elektrische Leistung des Verstärkers (blau) und Amplituden-geschwindigkeit eines Shakers (rot) in Abhängigkeit der Frequenz

Dieser Umstand ist auch in den folgenden beiden Prüfergebnissen (Abb. 4) dargestellt. Diese zeigen jeweils ein nach der RFST-Methode ausgewertetes Amplitudenbild von einer Platte mit vier Impactschäden im Sichtbereich der Kamera. Schon eine subjektive Betrachtung zeigt, dass das schmalere, optimierte Frequenzband eine effektivere Anregung der Defekte und damit eine bessere Detektierbarkeit bewirkt.

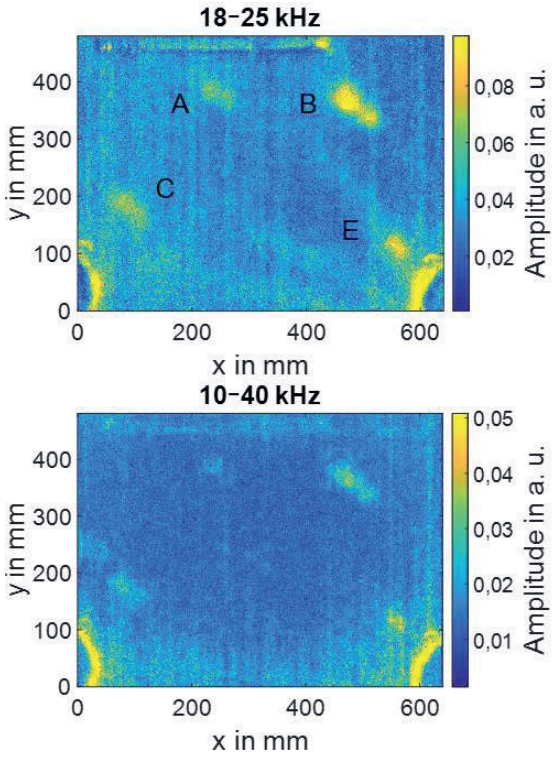


Abb. 4: Prüfergebnisse für verschiedene Anregungsbandbreiten

Einfluss der Anregungsleistung

Die Defekterkennungsrate skaliert direkt mit der Wärme, welche am Defekt umgesetzt wird. Diese erhöht sich mit steigender Ultraschallamplitude. Es wurde daher schon immer versucht, die elektrische Leistung am Piezo-Shaker zu erhöhen. Eine Testreihe mit variabler Anregungsleistung ist in Abb. 5. dargestellt und verdeutlicht diesen Zusammenhang.

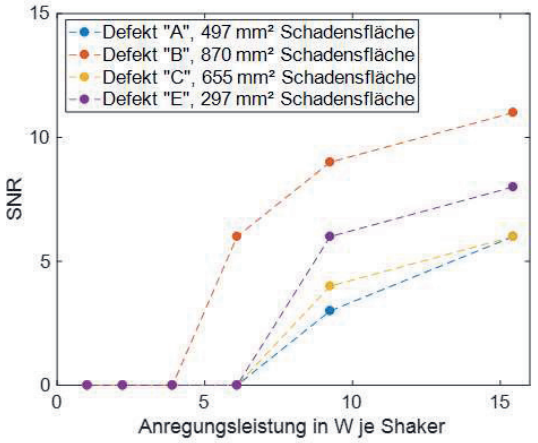


Abb. 5: Defekterkennung in Abhängigkeit der Anregungsleistung

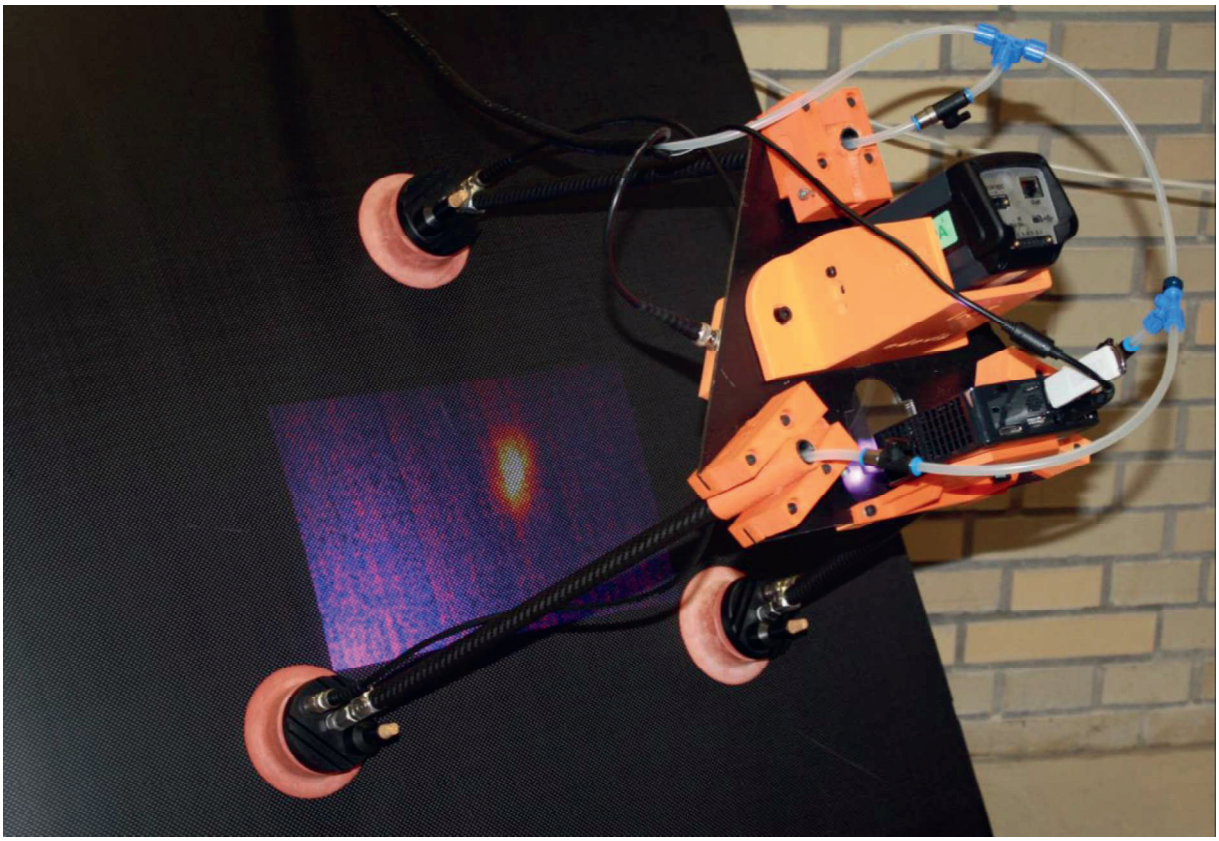


Abb. 6: Mobiles RFST-Prüfsystem im Einsatz

Ab einer Leistung von 8 W je Shaker lassen sich alle vier Defekte detektieren. Anzumerken ist hierbei, dass die Defektgröße nur teilweise Einfluss auf die Erkennungsrate hat. Ein hohes SNR kann auch für vergleichsweise kleine Defekte auftreten, wenn diese sich sehr nah an der Anregungsquelle befinden (vgl. Defekt „E“ in Abb. 5).

Gesamtergebnis

In Abb. 6 ist das Prüfsystem im Einsatz abgebildet, welches eine ca. DIN A4 große Fläche prüfen kann. Mit wahlweise Überdruck (Venturi-Prinzip) oder Unterdruck lässt es sich am zu prüfenden Bauteil fixieren und leitet dabei gleichzeitig über die Shaker in den Saugfüßen den Schall ein. Die durchströmende Luft dient zusätzlich der aktiven Kühlung der Piezo-Shaker, deren Langlebigkeit so verbessert wird. Parallel zur Beschallung nimmt die Thermografiekamera eine Thermogrammsequenz auf, welche unmittelbar online mittels Fast Fourier Transformation verarbeitet wird, um sofort ein Prüfergebnis zu erhalten. Dieses wird anschließend ortsgetreu auf die Bauteiloberfläche projiziert und ermöglicht dem Prüfer so eine schnelle und einfache Interpretation der Ergebnisse. Im Beispiel Abb. 6 ist z. B. ein Defekt vorhanden und wird in einer eindeutigen Falschfarbendarstellung abgebildet.

Insgesamt acht ähnliche Bauteile, zwei geviertelte, 3 mm dicke CFK-Dächer eines Automobils, wurden mit je 5 Impactschäden versehen. Beispielhaft sind die Front- und Rückseite eines Viertel-Daches in Abb. 7 dargestellt. Die Impactenergien und daraus resultierenden Schadensgrößen reichen von 4,5 J und 15 mm² bis 22,4 J und 1704 mm².



Abb.7: Viertel-Dach aus CFK mit Impactschäden. Oben: Frontseite mit Markierungen. Unten: Rückseite mit prominenten Delaminationen.

Die Prüfung dieser Bauteile für zu den in Abb. 8 dargestellten Ergebnissen. Auch hier ist die Detektionsqualität mit dem SNR-Wert angegeben, welcher aus den jeweiligen Amplitudenbildern (siehe [10]) extrahiert wurde.

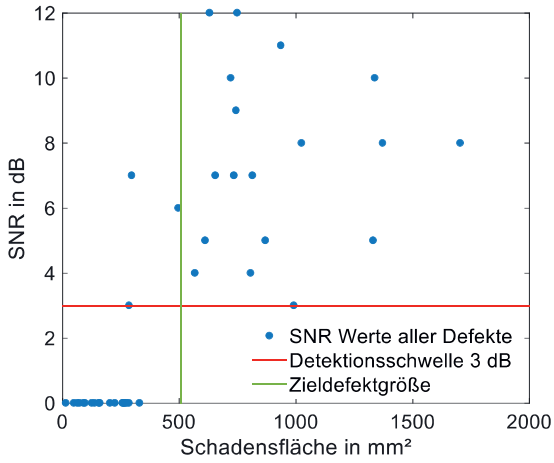


Abb.8: SNR-Werte aller untersuchten Impactschäden

Die Detektionsschwelle, ab der ein Defekt als erkannt deklariert wird (hier: 3 dB), muss für jedes Prüfverfahren und Prüfsystem neu definiert werden. Sie muss ebenso für jeden Anwendungsfall neu identifiziert werden und hängt neben dem Bauteil und Werkstoff auch von der Art der Auswertung (automatisch, manuell) ab. Für die hier durchgeführte, manuelle Bewertung der Prüfergebnisse ist ein Schwellwert von 3 dB gerechtfertigt, wie Abb. 9 zeigt. Der Defekt oben rechts im Bild ist mit einem SNR von 3 dB für das menschliche Auge noch eindeutig erkennbar. Ein zweiter Defekt, der ebenfalls im Prüfbereich liegt, ist hingegen nicht mehr erkennbar.

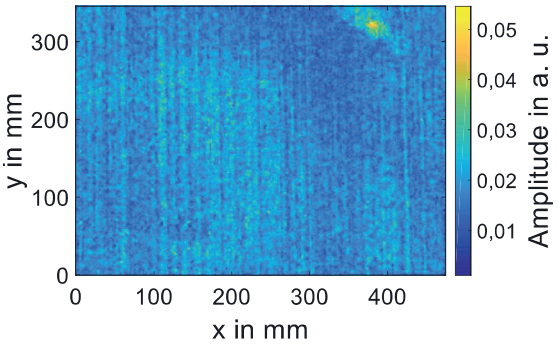


Abb. 9: Prüfergebnis einer CFK-Platte mit zwei Impactschäden (0 dB und 3 dB)

Wie Abb. 8 zeigt, lassen sich für die hier verwendeten Bauteile mit Impactschäden alle Defekte deren projizierte Schadensfläche größer als ein Quadratzoll beträgt eindeutig detektieren. Aus Zeit- und Kostengründen wurde auf eine ausführliche POD-Studie (Probability of Detection) verzichtet, die Reproduzierbarkeit des Verfahrens wurde aber bereits in früheren Veröffentlichungen nachgewiesen [11].

Fazit

Im Rahmen eines gemeinsamen Projektes mit der edevis GmbH, gefördert durch das BMWi im Rahmen des Zentralen Innovationsprogramm Mittelstand, konnte ein völlig neues zerstörungsfreies Prüfverfahren entwickelt werden. Dies wurde gezielt für den Zweck der schnellen und einfach interpretierbaren Prüfung auf Impactschäden entworfen. Es erlaubt eine mobile Prüfung von Impact-gefährdeten FKV-Bauteilen mit einer Prüfdauer von einer Minute. Das Ergebnis wird ortsgetreu auf die Oberfläche zurückprojiziert und erlaubt es so dem Prüfer Defekte direkt zuzuordnen. Aufgrund der inhärenten Dunkelfeldmethode sind die Prüfergebnisse auch von ungeschultem Personal leicht zu interpretieren, was die Flexibilität des Systems nochmals erhöht.
Anhand eines exemplarischen CFK-Bauteils konnte gezeigt werden, dass relevante Schäden sicher detektiert werden können. Grundlagenuntersuchungen deuten weiterhin an, dass die Effizienz des Verfahrens bezogen auf Prüfzeit und Detektionsrate durch eine

Optimierung der Anregungshardware noch deutlich gesteigert werden können.

Literaturnachweis

[1] KANKI, B.G., C.L. BRASIL. Analysis of ramp damage incidents and implications for future composite aircraft structure. *NASA Technical Report*, 2009.

[2] BULL, D.J., S.M. SPEARING, I. SINCLAIR, L. HELFEN. Three-dimensional assessment of low velocity impact damage in particle toughened composite laminates using micro-focus X-ray computed tomography and synchrotron radiation laminography. In: *Composites Part A: Applied Science and Manufacturing*, 2013, 52, 62–69. Doi:10.1016/j.compositesa.2013.05.003

[3] KRAUTKRÄMER, J., H. KRAUTKRÄMER. *Werkstoffprüfung mit Ultraschall*: Springer-Verlag, 2013. ISBN 3662134268.

[4] YANG, P., R. ELHAJJAR. Porosity content evaluation in carbon-fiber/epoxy composites using X-ray computed tomography. In: *Polymer-Plastics Technology and Engineering*, 2014, 53(3), 217–222.

[5] MAIERHOFER, C., P. MYRACH, M. REISCHEL, H. STEINFURTH, M. RÖLLIG, M. KUNERT. Characterizing damage in CFRP structures using flash thermography in reflection and transmission configurations. In: *Composites Part B: Engineering*, 2014, 57, 35–46. Doi:10.1016/j.compositesb.2013.09.036

[6] MIGNOGNA, R.B., R.E. GREEN, J.C. DUKE, E.G. HENNEKE, K.L. REIFSNIDER. Thermographic investigation of high-power ultrasonic heating in materials. In: *Ultrasonics*, 1981, 19(4), 159IN1161-160IN2163.

[7] RENSHAW, J., S.D. HOLLAND, R.B. THOMPSON. Measurement of crack opening stresses and crack closure stress profiles from heat generation in vibrating cracks. In: *Applied Physics Letters*, 2008, 93(8), 81914.

[8] SOLODOV, I., D. DERUSOVA, M. RAHAMMER. Thermosonic Chladni figures for defect-selective imaging. In: *Ultrasonics*, 2015, 60, 1–5.

[9] SOLODOV, I., M. RAHAMMER, D. DERUSOVA, G. BUSSE. Highly-efficient and noncontact vibro-thermography via local defect resonance. In: *Quantitative InfraRed Thermography Journal*, 2015, 12(1), 98–111.

[10] RAHAMMER, M., M. KREUTZBRUCK. Fourier-transform vibrothermography with frequency sweep excitation utilizing local defect resonances. In: *NDT & E International,* 2017, **86**, 83–88. Doi:10.1016/j.ndteint.2016.11.012

[11] RAHAMMER, M., M. KREUTZBRUCK. Validierung der Resonanten Frequenzsweep-Thermografie mittels einer POD-Analyse. In: *Materials Testing*, 2018, 60(5), 483–488. Doi:10.3139/120.111179

DOI 10.5162/sensoren2019/4.1.4

Ultrasonic characterisation of steel fibre orientation in ultra-high performance concrete

Benjamin Ankay [1], *Kevin Metje* [2], *Tom Mönig* [1,2], *Chuanzeng Zhang* [1], *Torsten Leutbecher* [2]

[1] *Chair of Structural Mechanics, University of Siegen, 57068 Siegen, Germany*
[2] *Chair of Structural Concrete, University of Siegen, 57068 Siegen, Germany*

Summary

Ultrasonic technology has proven to be a reliable non-destructive method for testing engineering applications. Especially for complex composites, like fibre-reinforced plastics or fibre-reinforced concrete, the pulse velocity method can be used to analyse the fibre-induced anisotropy. In the present study, ultrasonic testing was applied to investigate the fibre orientation in ultra-high performance fibre-reinforced concrete (UHPFRC). UHPFRC is a composite consisting of concrete and steel fibres, which distinguishes itself from conventional concrete by higher compressive and tensile strength as well as a particularly high durability. Steel fibres influence the tensile behaviour after the formation of micro-cracks decisively and increase the ductility of UHPFRC members. The orientation and distribution of fibres significantly influence the bearing behaviour and are thus important parameters when designing UHPFRC members. In ultrasound experiments, the main orientation of the fibres in UHPFRC specimens was determined by a method based on analysing the wave velocities of polarised shear waves. Subsequently, the specimens were cut perpendicular and parallel to the predominant fibre alignment predicted by ultrasonic testing and the actual orientation of the fibres in the cross-sections was evaluated by an optoanalytical method. Comparison of test data suggests that the proposed ultrasonic testing method is a suitable tool for predicting the predominant fibre alignment in UHPFRC members.

Keywords: ultrasonic method, concrete, steel fibres, fibre orientation, UHPC

1. Introduction

1.1. Ultrasonic testing in concrete

Since many decades non-destructive testing using ultrasound technology has proven to be a flexible, robust, meaningful, and comparatively cost-effective method for a wide range of engineering applications. Today different standard methods for ultrasonic testing of structural components are available. Especially in the fields of aircraft construction, mechanical engineering and railway transport, many standardised methods have been developed and established. The application in these fields is related to the ultrasonic non-destructive testing of adhesive bonds, joint connections, damage detection and generally the quality control of metallic and plastic materials.

Ultrasonic testing of concrete in civil engineering applications is still relatively new and has not been widely applied compared to other engineering fields mentioned above. However, over the years several ultrasonic methods for non-destructive concrete testing have been developed and have been included in guidelines and standards worldwide (e.g. ASTM C 597, EN 12504-4).

Basically, the major problem in ultrasonic testing of concrete is related to the highly heterogeneous structure of the material [1]. The complex meso- and microstructure of concrete, such as different aggregate sizes or voids inside the composite, leads to multiple scattering, reflections and absorption of the emitted ultrasonic waves. These factors influence the evaluation of the received ultrasonic signals considerably. Therefore, ultrasound experiments for concrete are usually carried out in the low-frequency range (50-250 kHz) in order to achieve a smaller disturbance due to the coarse aggregate in the material and thus to obtain a high signal to noise ratio. However, it must be considered that a reduction in the frequency always results in a worse spatial resolution and thus smaller defects can no longer be detected or resolved. With progress in the development of ultrasonic sensors nowadays efficient low-frequency broadband transducers allow testing of concrete structures reliably also in the so-called pulse-echo mode [2]. In the pulse-echo mode the transducer

acts simultaneously as a transmitter and receiver of ultrasonic waves. Here, the emitted ultrasonic pulse is reflected at a material boundary and the resulting echo is received at the transmission point. This is an improvement and a simplification for practical applications compared to the through-transmission mode [3], where the component must be accessible from both sides, which is often difficult or impossible.

As an established standard method, the pulse velocity measurement is used [1], where the required travel time of the ultrasonic pulse through the specimen is determined. From the data obtained, the pulse velocity of the tested concrete can be calculated, which can be used to assess the quality and strength of the concrete. Furthermore, damages or defects such as cracks, pores, and cavities as well as large-scale inhomogeneities can be detected [1]. Similarly, pulse velocity measurements on fibre reinforced concrete (FRC) have shown the influence of the distribution, orientation, and volume fraction of steel fibres on the acoustic characteristics [4-7]. In these studies, mostly longitudinal ultrasonic waves were used to detect the change in the measured wave velocity in concrete due to different fibre volume fraction and fibre orientation.

Basically, it can be assumed that the addition of fibres influences the elastic properties (e.g. the modulus of elasticity and Poisson's ratio) of the concrete depending on the specific FRC mix design and fibre orientation. In particular, a pronounced uniaxial or two-dimensional fibre orientation leads, similar to fibre-reinforced plastics (CFRP, GFRP), to a transverse anisotropy.

1.2. Ultrasonic testing in UHPFRC

Ultra-high performance concrete (UHPC) [8] is a heterogeneous composite made of cement, water, quartzitic or basaltic aggregate as well as reactive and non-reactive admixtures and additives. The compressive strength of UHPC typically ranges between 150 and 200 N/mm². This high strength and a very dense microstructure are obtained by reducing the water-cement ratio to typically about 0.25 and by grading the fine raw materials which provides an increased packing density. Superplasticiser is used to ensure the workability of the fresh concrete. As it is typical for concrete, the tensile strength of UHPC is low compared to its compressive strength. It typically ranges between about 7 and 11 N/mm² [8].

With increasing compressive strength the concrete becomes more brittle. To counteract this tendency and to increase the deformation capacity after appearance of cracks, steel fibres are added to the concrete mixture. So-called ultra-high performance fibre-reinforced concrete (UHPFRC) typically contains fibre volume fractions between 0.5 to 2.5 % using smooth straight steel fibres with lengths of typically 9 to 25 mm and diameters of typically 0.10 to 0.25 mm.

Beside increasing the UHPFRC's ductility, steel fibres enable the concrete to transfer tensile stresses through cracks and thus improve the tensile behaviour [9, 10] and significantly increase the bearing capacity of UHPFRC members subjected to bending, shear and/or torsion [8]. To benefit from these effects in structural design, knowledge of the specific fibre distribution and orientation inside a structural member is required [11, 12].

Ultrasonic testing in UHPC has been performed only occasionally so far. It has already been found that UHPC is particularly suitable for high-resolution ultrasonic testing due to its dense microstructure and comparably high homogeneity [4]. Furthermore, as shown in [4], the dynamic modulus determined by ultrasonic measurement corresponds well with the modulus of elasticity obtained from uniaxial compressive tests. For fine-grained UHPC, high testing frequencies in the range of 1 MHz can still be used avoiding the problem mentioned above [4], since the maximum aggregate size, which is in the range of less than 1 mm, is still below the associated wavelength, which leads to a higher resolution of the material structure. For UHPFRC the same frequency range may be used, since the fibres have no major influences on the signal to noise ratio.

A fibre-induced anisotropy has already been observed for UHPFRC specimens under static compressive loading [13]. There, various approaches considering the transverse anisotropy have been applied, which predict a variation in the elastic parameters between 5 to 10 % depending on the loading direction. In [11] direct tensile tests were conducted on specimens with different fibre orientations. In this study, the modulus of elasticity also varied in the range of 5 to 10 %.

The magnitude of variation suggest that the anisotropy of the elastic parameters can also be detected by ultrasonic measurements using elastic waves. However, the challenge is that due to the low fibre volume fraction only a small variation in the acoustic parameters, such as the wave velocity, is to be expected.

2. Analytical and experimental investigations

2.1. Analytical consideration and experimental procedure

Since an absolute measurement of the acoustic parameters and its subsequent comparison is often limited by the achievable measurement accuracy, a relative ultrasonic measurement was performed in this study to capture the expected small velocity variations. Here, a pulse-echo method using polarised shear waves is applied to determine the relative pulse velocity change in a specimen while rotating the ultrasound transducer.
The use of transverse waves for ultrasonic inspection is a well-established method for determining the material anisotropy caused by the structure itself or by directional damage effects. Similar methods based on the principle of shear wave birefringence/splitting have recently been used in the field of fibre-reinforced plastics [14, 15] and have also been applied to anisotropic materials such as wood [16]. Generally, the anisotropy level is much larger in this kind of materials in comparison with fibre-reinforced concrete. However, the presented studies show that even small changes in the material behaviour are detectable with this method. To classify the direction-dependent changes in the wave velocity present in UHPFRC, an analytical examination based on a layer model [17] is performed (Fig. 1).

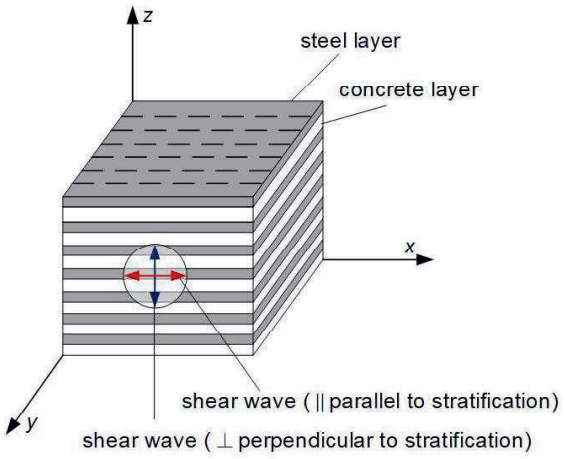


Fig. 1: Shear wave propagation in fibre-reinforced concrete idealised by a layer model

According to the layer model for a shear wave with propagation direction and particle motion parallel to the stratification (but perpendicular to eachother) the corresponding wave velocity is given by Eq. (1).

$$v_{\parallel} = \sqrt{\frac{\left(n_A \rho_A v_A^{\,2} + n_B \rho_B v_B^{\,2}\right)}{\left(n_A \rho_A + n_B \rho_B\right)}} \tag{1}$$

where

ρ : mass density of the material
v : velocity of the shear wave
n : volume fraction of the material

The subscript indices denote the two materials A and B under consideration.
If the particle motion is perpendicular to the stratification (normal to the layer) the corresponding wave velocity is given by Eq. (2).

$$v_{\perp} = \sqrt{\frac{1}{\left(\frac{n_A}{\rho_A v_A^{\,2}} + \frac{n_B}{\rho_B v_B^{\,2}}\right)\left(n_A \rho_A + n_B \rho_B\right)}} \tag{2}$$

As a specific example, the material A is taken as steel with a density $\rho_A = 7850$ kg/m^3 and a shear wave velocity $v_A = 3200$ m/s, while the material B is selected as UHPC with a density $\rho_B = 2400$ kg/m and a shear wave velocity $v_B = 2900$ m/s. Regarding realistic fibre volume fractions in UHPFRC the volume fraction of both materials is set to $n_A = 0.025$ and $n_B = 0.975$, which means that the volume fraction of steel in the total volume is assumed to be 2.5 %. This results in the following values for the corresponding shear wave velocities determined by Eqs. (3) and (4).

$$v_{\parallel} \approx 2924 \text{ m/s} \tag{3}$$

$$v_{\perp} \approx 2848 \text{ m/s} \tag{4}$$

Therefore, the shear wave velocity with polarisation parallel to the layer results in an absolute value that is about 2.6 % higher than the corresponding velocity perpendicular to the layer. This gives a good estimation on the velocity changes expected in ultrasonic experiments on UHPFRC specimens.
In order to evaluate the wave velocity change of polarised shear waves by ultrasonic experiments, the UHPFRC specimens should ideally provide a nearly uniaxial fibre orientation.

2.2. Specimen preparation

To provoke a predominant fibre alignment in the specimen, thin panels were cast in different arrangement (3 x upright/1 x reclined) as well as with varying casting method. Previous studies demonstrated the influence of the casting method on the flow of the fresh concrete and thus on the distribution and orientation of the fibres inside the specimen [11, 18, 19].

For the present study, a fine-grained UHPFRC with a maximum aggregate size of 0.5 mm was used in order to ensure a microstructure as homogeneous as possible. The fibre volume fraction was 2.5 %. The steel fibres had a length of 10 mm and a diameter of 0.175 mm. Tab. 1 shows the mix design.

Tab. 1: UHPFRC mix design

Cement 52,5 R	765 kg/m³
Silica fume	160 kg/m³
Quartz flour 0/0.125 mm	195 kg/m³
Quartz sand 0.125/0.5 mm	950 kg/m³
Superplasticiser (30 % solids)	26 kg/m³
Steel fibres 10/0.175 mm	196 kg/m³
Water	195 kg/m³
Water-cement ratio	0.28*

* Taking into account the water content of the superplasticiser

The panels were about 450 mm high, 1200 mm long and 40 mm thick. The concrete was poured from one end. The mould of the three vertically cast panels was inclined by an angle of 63° at the end where the concrete was poured into the mould in order to ensure a steady flow velocity (Fig. 2a). The mould of the horizontally cast panel was inclined by an angle of about 4° with regard to the horizontal (Fig. 2b). Using a funnel, the casting took about 7 minutes for each panel. The mean cube compressive strength (100 mm cubes) was 106 MPa, 136 MPa and 165 MPa measured 2, 7 and 28 days after casting, respectively.

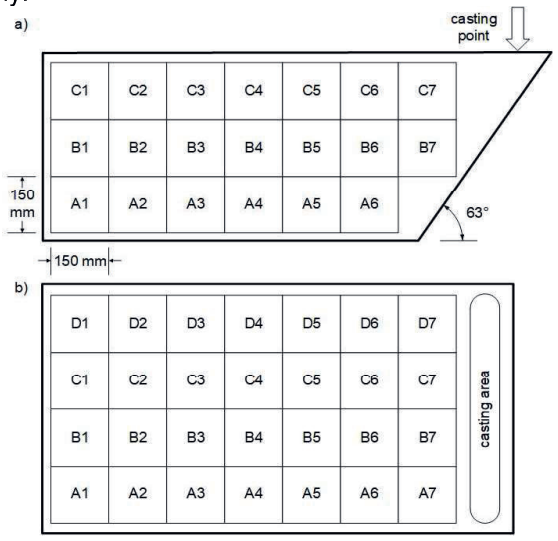


Fig. 2: a) Front view of the vertically casted panels, b) top view of the horizontally casted panel

After hardening, the panels were sawn into square specimens with a side length of 150 mm for ultrasonic testing. The specimens were specified according to Fig. 2.

2.3. Ultrasonic shear wave testing

For the experiments, a low-frequency shear wave transducer with a centre-frequency of 500 kHz was used in a pulse-echo mode. A crystallised honey was used as coupling medium between the transducer and the UHPFRC specimens. The wave velocities of the generated shear waves were determined via a specially designed measuring system (Ritec-RAM 5000-SNAP), which evaluates analogously the phase spectrum of the received waves.
This method can be explained by assuming two signals u_0 and u_1 with different wave velocities according to Eqs. (5) and (6).

$$u_0 = \mathrm{A} \cdot \sin(2\pi f \cdot t_0) = \mathrm{A} \cdot \sin(\varphi_0) \tag{5}$$

$$u_1 = \mathrm{B} \cdot \sin(2\pi f \cdot t_1) = \mathrm{B} \cdot \sin(\varphi_1) \tag{6}$$

where

f: excitation frequency
φ: phase angle
t: time of flight

By using the path length D, the relation between the phase angle φ and the wave velocity v_0 is given for signal u_0 by Eq. (7).

$$t_0 = \frac{\varphi_0}{2\pi f} \Rightarrow v_0 = \frac{2D}{t_0} \tag{7}$$

In the same way, the wave velocity can be calculated for signal u_1.
Now, by measuring the corresponding phase difference $\Delta\varphi = \varphi_0 - \varphi_1$ between both signals the relative change in the wave velocity $\Delta v / v_0$ can be calculated by Eqs. (8) and (9) respectively.

$$\Delta t = t_0 - t_1 = \frac{\Delta\varphi}{2\pi f} \tag{8}$$

$$\frac{\Delta v}{v_0} = \frac{\Delta t}{t_0} \tag{9}$$

Using this method, the shear wave velocities v_s were calculated for different rotation angles. Therefore, the transducer was rotated from a stipulated starting position ($\theta = 0°$) in 15° increments and the associated relative changes in the wave velocity with respect to the starting position were recorded (Fig. 3).



Fig. 3: Experimental setup: Shear wave transducer with a polar disk

Finally, this measurement procedure results in a diagram as given exemplarily in Fig. 4. Here, the percentage change in the shear wave velocity $\Delta v_s / v_{s,0}$ is given as a function of the rotation angle θ. Based on this diagram, the maximum wave velocity change and the corresponding rotation angle θ_{max} were determined for each specimen. The latter quantity describes the expected predominant fibre alignment at the measuring point. Subsequently, the wave velocity change at the position perpendicular to the maximum wave velocity change was determined ($\theta_{90°} = \theta_{max} \pm 90°$) and correlated to the maximum value. So the increase of the shear wave velocity starting from the position according to the predominant fibre alignment to the position perpendicular thereto can be described.

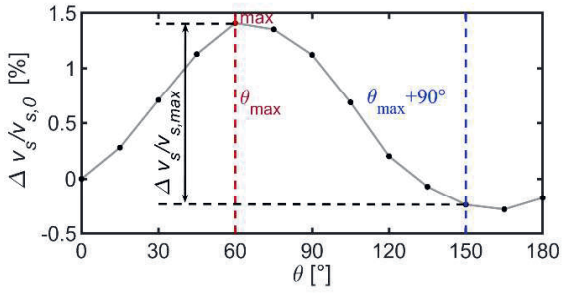


Fig. 4: Percentage change of the shear wave velocity v_s as a function of the rotation angle θ

2.4. Optoanalytical method

After ultrasonic testing, specimens with particularly high differences in the relative wave velocities were selected for subsequent analysis of fibre orientation. The objective was to validate the results of the non-destructive ultrasonic measurement by means of a proven test method for evaluating the fibre distribution and orientation. As presented in [20-24], the optoanalytical method has proven to be such a suitable tool.

The square specimens were sawn parallel and perpendicular to the rotation angle θ_{max}, which was predicted by ultrasonic testing to represent the predominant fibre alignment. By sawing, the steel fibres are melted at the cut surface and show deformed shapes. Thus, the cut sections were polished to recover the real shape of the fibres. After polishing, the cross-sections were photographed (Fig. 5) and analysed with the software FiDiOr [22].

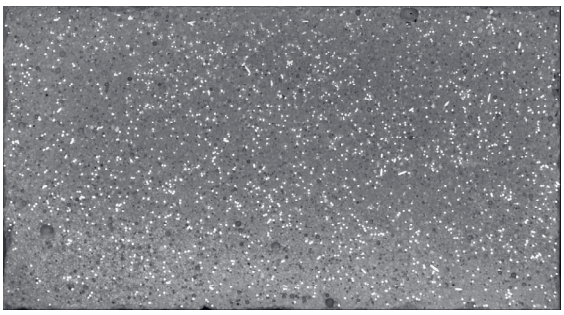


Fig. 5: Detail of a photography of a polished cross-section orientated perpendicular to the predominant fibre alignment

FiDiOr uses the contrast between darker concrete and lighter steel to determine the elliptical shapes of the cut fibres via grey value analysis. This provides number, individual coordinates, long and short principal axis and cross-sectional area of the ellipses (= fibres). With these data the inclination of the ellipse's principal axis against the axis of the cross-sectional plane as well as their out-of-plane inclination α can be calculated. Fig. 6 shows the situation and the notations for an inclined fibre in a cut section.

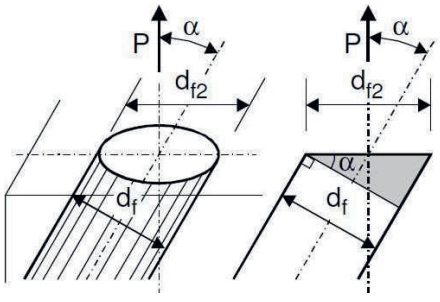


Fig. 6: Inclined fibre in a cut section [23]

For an individual fibre the coefficient of fibre orientation $\eta_{f,i}$ can be calculated by Eq. (10).

$$\eta_{f,i} = \cos\alpha = \frac{d_f}{d_{f2,i}} \tag{10}$$

where

d_f : length of the short principal axis of the elliptical cross-section (= fibre diameter)

$d_{f2,i}$: length of the long principal axis of the elliptical cross-section of fibre *i*

After evaluating the individual fibres of a cross-section, the mean coefficient of fibre orientation η_S in this cross-section can be calculated by Eq. (11) [20-24].

$$\eta_S = \frac{1}{N_f} \cdot \sum_{i=1}^{N_f} \eta_{f,i} \tag{11}$$

where

N_f : number of fibres in the cross-section

However, in order to represent the fibre orientation of a specific volume, the mean coefficient of fibre orientation η_V is to be calculated by Eq. (12) [24].

$$\eta_V = \frac{N_f}{\sum_{i=1}^{N_f} \frac{1}{\eta_{f,i}}} \tag{12}$$

Based on η_V, the fibre volume fraction of a specimen can be predicted by Eq. (13) [20, 22, 24].

$$\rho_f = \frac{N_f \cdot \pi \cdot d_f^2}{4 \cdot A_c \cdot \eta_V} \tag{13}$$

where

A_c : cross-section of the concrete

3. Results and evaluation

The results of the ultrasonic and photographic analyses are summarised in Tab. 2.

Tab. 2: Mean coefficient of fibre orientation of the volume, predicted fibre volume fraction, number of fibres and percentage wave velocity change of the selected specimens sawn from panel no. 1, 2, 3, and 4

No.	$\eta_{V,max}$ [-]	$\eta_{V,90°}$ [-]	ρ_{max} [%]	$\rho_{90°}$ [%]	N_{max} [-]	$N_{90°}$ [-]	$\Delta v_{s,max} / v_{s,0}$ [%]
1_A1	0.738	0.466	2.604	2.388	4289	2484	1.948
1_A6	0.709	0.475	2.502	2.534	3712	2536	1.633
1_B2	0.752	0.433	2.159	1.864	3734	1835	1.416
1_B3	0.736	0.395	2.374	1.917	4088	1772	1.658
1_C1	0.802	0.363	2.069	1.840	3281	1208	1.647
1_C4	0.699	0.513	1.969	1.906	2685	1817	1.244
1_C7	0.764	0.513	2.074	1.737	3063	1620	1.562
2_A1	0.746	0.497	2.363	2.006	3179	1784	0.709
2_A4	0.815	0.380	2.267	2.116	3321	1447	0.697
2_A6	0.796	0.393	2.262	1.873	3125	1287	0.849
2_B1	0.746	0.439	2.236	2.017	2847	1513	0.678
2_C4	0.776	0.412	2.382	2.056	3182	1466	0.862
3_A1	0.717	0.474	2.774	2.801	4376	2991	1.253
3_A3	0.712	0.454	2.573	2.476	4009	2491	1.603
3_A6	0.601	0.620	2.217	2.370	2920	3203	1.443
3_B1	0.764	0.439	2.290	2.228	3493	1976	1.585
3_B7	0.709	0.407	2.347	2.137	3528	1821	1.167
3_C6	0.681	0.517	2.198	1.867	2956	1794	1.144
4_A1	0.743	0.492	2.430	2.277	4000	2540	1.102
4_A2	0.754	0.436	2.611	2.112	4529	2109	0.992
4_A4	0.752	0.418	2.268	1.944	4066	1950	0.905
4_A6	0.748	0.425	2.325	2.197	4041	2170	0.938
4_B1	0.787	0.375	2.342	1.835	4183	1579	1.039
4_C1	0.788	0.341	2.317	1.963	3528	1119	1.423
Avg.	0.743	0.445	2.331	2.103	3589	1938	1.229
CV	6.14 %	13.90 %	8.01 %	12.54 %	14.96 %	27.91 %	29.29 %

The experimental data are subdivided into four groups corresponding to the four panels from which the test specimens were taken from. Panel no. 1, 3, and 4 were cast vertically. Panel no. 2 was cast horizontally.

As already mentioned, the optoanalytical evaluation was performed in two cross-sections perpendicular to each other. For a clear distinction, the cross-section perpendicular to the expected predominant fibre alignment is denoted by the index “max” while the section parallel to the expected predominant fibre alignment is denoted with the index “90°”.

The coefficient of fibre orientation η_V, the fibre volume fraction ρ and the number of fibres N obtained by optoanalytical evaluation are listed in Tab. 2 for each specimen. Furthermore, the maximum percentage wave velocity change $\Delta v_{s,max} / v_{s,0}$ measured by the ultrasonic method is indicated. The average values (Avg.) and the associated coefficients of variation (CV) are presented at the end of the table.

Looking at the fibre volume fraction it is noticeable that the values for individual specimens deviate by up to 20 % from the actual fibre volume fraction of the mixture (2.5 %). This may be due to sedimentation of the fibres in the fresh concrete.

The wave velocity changes of individual panels deviate in some cases by up to 30 %. While panels no. 1 and no. 3 have high velocity change values of 1.24 to 1.95 % and 1.14 to 1.60 %, panels no. 2 and no. 4 have much lower values in the range of 0.71 to 0.86 % and 0.91 to 1.44 % respectively. Since visual inspection of the cross-sections showed a significantly increased proportion of air voids for panels no. 2 and no. 4, these air voids might be responsible for a major part of the deviations mentioned before. The air voids can be explained as a result of poor ventilation during casting due to insufficient compaction. In principle, air voids will influence the ultrasonic measurement, depending on their form and distribution in the concrete. Further investigation is required in order to finally explain the observed deviations in velocity changes.

With a value of $\eta_{V,max} = 0.74$ the cross-sections perpendicular to the rotation angle θ_{max} on average show a high coefficient of fibre orientation while the average coefficient of fibre orientation is significantly smaller for the cross-sections parallel to θ_{max} ($\eta_{V,90°} = 0.44$). Thus, it can be basically stated that ultrasonic measurement is able to identify directions with high coefficient of fibre orientation corresponding to the predominant fibre alignment.

4. Conclusions and outlook

In this study, a novel ultrasonic method for determining the fibre orientation in UHPFRC was applied. The evaluation of the results obtained with ultrasonic and optoanalytical method confirms that the method is generally suitable for qualitatively determining the predominant fibre alignment in UHPFRC. However, a quantitative assessment for a specific coefficient of orientation for individual specimens depends on different factors, which may be investigated and evaluated in further studies. The main influencing factors are inhomogeneities, such as air voids, aggregate, etc. as well as heterogeneous fibre distribution in the concrete. Furthermore, the coupling condition between the transducer and the specimen’s surface is influencing the ultrasonic measurement.

The present study may be considered as a meaningful starting point for future work in this research field. For example, new ultrasonic methods using air-coupled transducers or surface acoustic wave (SAW) methods can be adopted to overcome some drawbacks of the present method and to automate the measuring procedure.

References

[1] V.M. Malhotra, J.C. Nicholas, Handbook on nondestructive testing of concrete. *Boca Raton, Fla: CRC Press*. (2003); doi: 10.1201/9781420040050

[2] M. Krause, M. Bärmann, R. Frielinghaus, F. Kretzschmar, O. Kroggel, K.J. Langenberg, C. Maierhofer, W. Müller, J. Neisecke, M. Schickert, V. Schmitz, H. Wiggenhauser, F. Wollbold, Comparison of pulse-echo methods for testing concrete. *NDT & E International*. 30(4), 195-204 (1997); doi: 10.1002/best.200900679

[3] D. Breysse (ed.), Non-destructive assessment of concrete structures: reliability and limits of single and combined techniques. Springer Netherlands. (2012); doi: 10.1007/978-94-007-2736-6

[4] G. Washer, P. Fuchs, A. Rezai, H. Ghasemi, Ultrasonic measurement of the elastic properties of ultra-high performance concrete (UHPC). In P.J. Shull, A.L. Gyekenyesi, A.A. Mufti (eds.), *Nondestructive Evaluation and Health Monitoring of Aerospace Materials, Composites, and Civil Infrastructure IV*. *SPIE* (2005); doi: 10.1117/12.601815

[5] M. Benaicha, O. Jalbaud, A. Hafidi Alaoui, Y. Burtschell, Correlation between the mechanical behavior and the ultrasonic velocity of fibre-reinforced concrete. *Construction and Building Materials*. 101, 702-709 (2015); doi: 10.1016/j.conbuildmat.2015.10.047

[6] E. Reufi, J. Marku, T. Bier, Ultrasonic pulse velocity investigation of polypropylene and steel fibre reinforced concrete. *Zenodo*. (2016); doi: 10.5281/zenodo.1112125

[7] B.T. Gebretsadik, Ultrasonic pulse velocity investigation of steel fibre reinforced self-compacted concrete. *UNLV Theses, Dissertations, Professional Papers, and Capstones*. (2013).
[8] E. Fehling, M. Schmidt, J. Walraven, T. Leutbecher, S. Fröhlich, Ultra-High Performance Concrete UHPC: Fundamentals - Design - Examples. Berlin: Ernst & Sohn. (2014); doi: 10.1002/9783433604076
[9] T. Leutbecher, E. Fehling, Rissbildung und Zugtragverhalten von mit Fasern verstärktem Stahlbeton am Beispiel ultrahochfesten Betons - Teil 1: Rissmechanische Zusammenhänge. *Beton- und Stahlbetonbau*. 104(6), 357-367 (2009); doi: 10.1002/best.200900677
[10] T. Leutbecher, E. Fehling, Rissbildung und Zugtragverhalten von mit Fasern verstärktem Stahlbeton am Beispiel ultrahochfesten Betons - Teil 2: Experimentelle Untersuchungen und Anwendungsbeispiele. *Beton- und Stahlbetonbau*. 104(7), 406-415 (2009); doi: 10.1002/best.200900679
[11] L.F. Maya Duque, B.A. Graybeal, Fiber orientation distribution and tensile mechanical response in UHPFRC. *Materials and Structures*. 50:55, (2017); doi: 10.1617/s11527-016-0914-5
[12] K. Wille, N.V. Tue, G.J. Parra-Montesinos, Fiber distribution and orientation in UHP-FRC beams and their effect on backward analysis. *Materials and Structures*. 47(11), 1825-1838 (2014); doi: 10.1617/s11527-013-0153-y
[13] M. Curbach, K. Speck, Versuchstechnische Ermittlung und mathematische Beschreibung der mehraxialen Festigkeit von Ultra-Hochfestem Beton – Zweiaxiale Druckfestigkeit. Arbeitsbericht zum Forschungsvorhaben CU 37/6-1 der Deutschen Forschungsgemeinschaft (DFG), TU Dresden, (2007)
[14] I. Solodov, K. Pfleiderer, D. Döring, G. Busse, Nondestructive evaluation of anisotropy in composite materials via acoustic birefringence. *Research in Nondestructive Evaluation*. 19(3), 129-143 (2008); doi: 10.1080/09349840801954223
[15] S.T. Rakotonarivo, C. Payan, J. Moysan, C. Hochard, Local damage evaluation of a laminate composite plate using ultrasonic birefringence of shear wave. *Composites Part B: Engineering*. 142, 287-292 (2018); doi: 10.1016/j.compositesb.2018.01.006
[16] J.P. Zhou, J.X. Liu, W.Y. Gao, Z.W. Cui, W.G. Lv, K.X. Wang, Effect of anisotropy on shear wave velocity in wood. *Advanced Materials Research*. 535–537, 1923-1926 (2012); doi: 10.4028/www.scientific.net/amr.535-537.1923
[17] J.E. White, F.A. Angona, Elastic wave velocities in laminated media. *The Journal of the Acoustical Society of America*. 27(2), 310-317 (1955); doi: 10.1121/1.1907520
[18] B. Zhou, Y. Uchida, Influence of flowability, casting time and formwork geometry on fiber orientation and mechanical properties of UHP-FRC. *Cement and Concrete Research*. 95, 164-177 (2017); doi: 10.1016/j.cemconres.2017.02.017
[19] R. Lovichová, J. Fornůsek, L. Soukupová, J. Valentin, Ultrasound gel as suitable tool for simulation of the fiber orientation in the fiber reinforced concrete. *EAN 2016 - 54th International Conference on Experimental Stress Analysis*, Srní, CZ. (2016).
[20] H.K. Hilsdorf, W. Brameshuber, R. Kottas, Weiterentwicklung und Optimierung der Materialeigenschaften faserbewehrten Betons und Spritzfaserbetons als Stabilisierungselemente der Felssicherung. Forschungsbericht, Universität Karlsruhe, (1985).
[21] K. Schönlin, Ermittlung der Orientierung, Menge und Verteilung der Fasern in faserbewehrtem Beton. *Beton- und Stahlbetonbau*. 83(6), 168-171 (1988); doi: 10.1002/best.198800280
[22] N.V. Tue, S. Henze, M. Küchler, G. Schenck, K. Wille, Ein optoanalytisches Verfahren zur Bestimmung der Faserverteilung und -orientierung in stahlfaserverstärktem UHFB. *Beton- und Stahlbetonbau*. 102(10), 674-680 (2007); doi: 10.1002/best.200700568
[23] I. Marković, High-Performance Hybrid-Fibre Concrete: Development and Utilisation. Delft University Press, (2006); http://resolver.tudelft.nl/uuid:44ed51cd-fc27-4353-ab81-cf8348f52443
[24] P. Hadl, S. Henner, H. Kim, N.V. Tue, Experimental investigations on the scattering in the post cracking tensile behaviour of UHFRC. In E. Fehling, B. Middendorf, J. Thiemicke (eds.), *Proceedings of HiPerMat 2016*, Kassel, D. (2016); doi: 10.19211/KUP9783737600958

DOI 10.5162/sensoren2019/4.2.1

Konzeption und Evaluierung eines Laboraufbaus zur photoakustischen Detektion von CH4 mit modulierter Breitband MEMS-IR-Strahlungsquelle

B. Vatter[1,2], A. Magi[1], F. Liebold[1,3]

[1]Micro-Hybrid Electronic GmbH, Forschung und Entwicklung, 07629 Hermsdorf
[2]Westsächsische Hochschule Zwickau, Fakultät für Physik und Informatik, 08056 Zwickau
[3]Duale Hochschule Gera-Eisenach, Fachbereich Technik, 07546 Gera

benedikt.vatter.f4u@fh-zwickau.de, a.magi@micro-hybrid.de, falk.liebold@dhge.de

Zusammenfassung

In dieser Arbeit wird eine photoakustische Messzelle zur Bestimmung von Methan vorgestellt. Die Grundlage stellt eine kommerzielle NDIR-Gasmesszelle dar, die um ein MEMS-Mikrofon erweitert wurde. Zur Erzeugung des photoakustischen Signals wurde ein MEMS-Breitband-IR-Strahler bei einer Frequenz von 9 Hz elektrisch moduliert. Mittels Lock-In-Verfahren wurde das Signal in Abhängigkeit der Konzentration von Methan detektiert. Das zu erwartende Signal konnte über eine rechnerische Modellgleichung abgeschätzt und so mit den experimentellen Werten verglichen werden. Es wird gezeigt, dass diese Abschätzung mit den experimentellen Daten gut übereinstimmt und somit prinzipiell für die Realisierung von photoakustischen Messzellen verwendet werden kann. Der Sensor wurde mit zwei Strahlertypen erprobt. Es wurde ein Detektionslimit von 108 ppmV bei einer Auflösung von 43 ppmV ermittelt.

Keywords: Gassensorik, photoakustische Spektroskopie, Methan, Absorptionsspektroskopie, MEMS-Membranstrahler, NDIR, IR-Strahler, Hot-plate Emitter

Einleitung

Die Bestimmung der Konzentration IR-aktiver Gase wie z.B. CO_2 und CH_4 spielt in der Umwelt- und Prozessüberwachung eine tragende Rolle. Als Messprinzip bietet sich neben der weit verbreiteten NDIR-Spektroskopie die Photoakustik an. Es wurde bereits mehrfach gezeigt [1][2][3], dass die Photoakustik für die Bestimmung geringster Konzentrationen bis in den unteren pptV-Bereich geeignet ist. Jedoch sind diese Methoden oft sehr komplex und nur für den Laboraufbau geeignet. Für die einfache Integration in bereits bestehende Gassensoren und zur Minimierung mechanischer Bauteile, wurde in dieser Arbeit ein Messzellenkonzept aus elektrisch moduliertem MEMS-Breitbandstrahler und einem MEMS-Mikrofon untersucht. Das zu erwartende photoakustische Signal für Methan wurde im Vorfeld durch eine geeignete Modellgleichung abgeschätzt.

Grundlagen

Das photoakustische Prinzip wurde erstmals 1880 von A. G. Bell erwähnt. Es beschreibt die Erzeugung einer Schallwelle durch einen modulierten Energieeintrag, in Form von Strahlung, in ein System [4]. Die periodisch eingebrachte Energie Q aus einer Strahlungsquelle führt nach Gl. (1) zu einer Änderung der Temperatur ΔT und nach der idealen Gasgleichung Gl. (2) auch zu einer Druckänderung Δp [5]. Aus der optischen Leistung der Strahlungsquelle ergibt sich die in das System eingebrachte Wärmeenergie pro Zeitintervall.

$$Q = C_V \, \Delta T \tag{1}$$

$$\Delta p = \frac{n \, R \, \Delta T}{V} \tag{2}$$

Diese Druckänderung ist abhängig von den Eigenschaften der Strahlungsquelle, dem Absorptionsverhalten des zu untersuchenden Materials, der Geometrie der Messzelle und der Konzentration des Stoffes. Damit lässt sich für kleine Konzentrationen c die Druckänderung nach Gl. (3) berechnen.

$$\Delta p = \frac{R \, P_{opt} \, \alpha \, l \, c \, dc \, m \, \delta \, T_{opt}}{c_V \, V_{Zelle} \, f} \tag{3}$$

Die Druckänderung ergibt sich aus der idealen Gaskonstante R, der optischen Leistung der

Strahlungsquelle P_{opt}, dem Absorptionskoeffizienten α, der Länge l und dem Volumen V_{Zelle} der Messzelle. Weiterhin geht der duty cycle dc (Verhältnis zwischen An- und Auszeit der Strahlungsquelle), der Intensitäts-Modulationsgrad m des Strahlers (Verhältnis der Intensität im An-Zustand zur max. möglichen Intensität im unmodulierten Betrieb), sowie ein Korrekturfaktor δ, der die Überlagerung von Emissionspektrum der Strahlungsquelle und Absorptionsspektrum der Probe berücksichtigt mit ein. Auch der Transmissionskoeffizient T_{opt} der optischen Komponenten im Strahlengang, die spezifischen Wärmekapazität c_v sowie die Frequenz f der modulierten Strahlungsquelle werden in der verwendeten Modellgleichung Gl. (3) berücksichtigt. Die daraus resultierenden Ergebnisse wurden zur Abschätzung im Vorfeld der eigentlichen Messung verwendet. Vorgänge bei der Absorption und Verlustmechanismen, wie sie in [6][7] diskutiert werden, sollen hier nicht weiter betrachtet werden.

Vorstellung der Komponenten

Im Vorfeld wurden die Eigenschaften der Strahlungsquelle, der Messzelle und der Probe untersucht. Als Strahlungsquelle kommt ein MEMS-Membranstrahler JSIR-350 der Firma Micro-Hybrid Electronic GmbH (MHE) zum Einsatz. Das Emissionsspektrum bei 0,8 W elektrischer Leistung ist in Abb. 1 dargestellt.

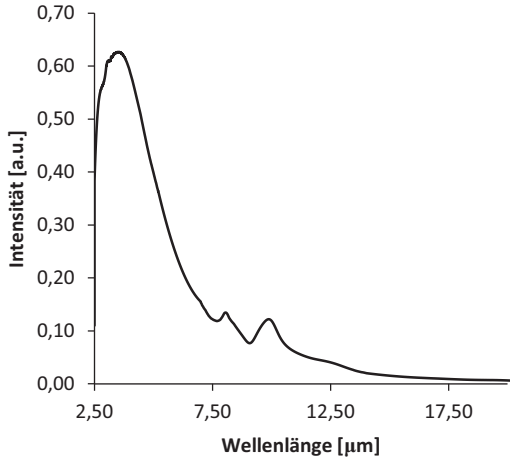


Abb. 1: Normiertes Emissionsspektrum des MEMS-Breitbandstrahlers JSIR-350 zwischen 2,5 und 20 µm.

Der Strahler wird sowohl offen als auch mit Krypton (Kr) verschlossen eingesetzt. Der mit Kr verschlossene Strahler kann, aufgrund der schlechteren Wärmeleitung gegenüber Luft, bei gleicher elektrischer Leistung höhere Temperaturen und damit auch höhere optische Leistungen erzielen. Da es sich um einen thermischen Dünnschicht- bzw. Membran-Strahler handelt, kann dieser mit einer Frequenz im Hz-Bereich betrieben werden. Mit Modulationsfrequenzen größer 20 Hz nimmt der Modulationsgrad m merklich ab. Zur Detektion wurde daher ein MEMS-Mikrofon ICS-40300 der Firma Invensense verwendet, welches bis 6 Hz sensitiv ist.

Als Messkammer dient eine NDIR-Gasmesszelle der Firma MHE (Abb. 2). Der Absorptionspfad ist 10 cm lang und besitzt einen Radius von 4 mm. Damit ergibt sich ein Zellvolumen von rund 5 ml. Zu- und Abfluss der Zelle können über Magnetventile verschlossen werden, um Störgeräusche aus der Leitung zu minimieren.

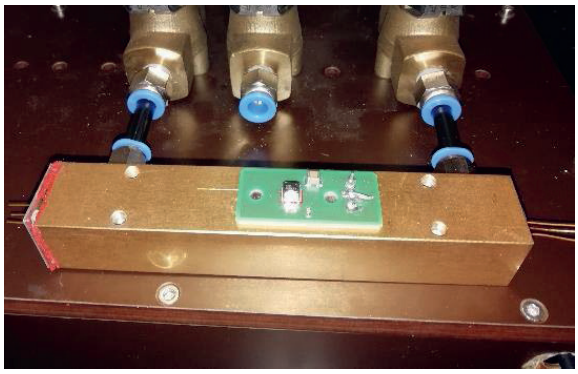


Abb. 2: NDIR-Gasmesszelle mit zusätzlichem MEMS-Mikrofon. Zu- und Abfluss über Magnetventil regelbar.

Die Mikrofonspannung wurde mittels LockIn-Verstärker (LIA, 5210 Dual Phase Analog Lock-in Amplifier Fa. Signal Recovery) um Faktor 10^4 verstärkt und mit einem erstellten LabVIEW-Programm ausgelesen. Das Ansteuersignal des Strahlers diente als Referenzsignal. Die Integrationszeit des LIA betrug 3 s.

Zur Bestimmung des Optimums aus Modulationsfrequenz der Strahlungsquelle und Empfindlichkeit des Mikrofons, wurde die Messzelle zunächst mit CO_2 befüllt und der Frequenzgang bis 15 Hz aufgenommen. Wie in Abb. 3 zu erkennen ist, tritt ein Maximum des Sensorsignals bei 9 Hz auf. Die indirekte Proportionalität der Anregungsfrequenz führt zu einer Minimierung des Schallwechseldrucks für Frequenzen größer 9 Hz (siehe Gl. 3). Bei höheren Frequenzen führt auch der sinkende Modulationsgrad m des Strahlers selbst zu einer Abnahme des Ausgangssignals.

Messung

Als Probe wird Methan verwendet. CH_4 weist eine gute Überlagerung der Absorptionsbanden

mit dem Maximum des Emissionsspektrums der Strahler auf. Da kein optischer Filter verwendet wurde, tragen alle Absorptionsbanden, die mit der Strahlungsquelle angeregt werden können zur Schallgenerierung bei. Für gasspezifische Anwendungen können zur Minimierung von Quereinflüssen auf spezielle Absorptionsbanden angepasste Filtergläser verwendet werden.

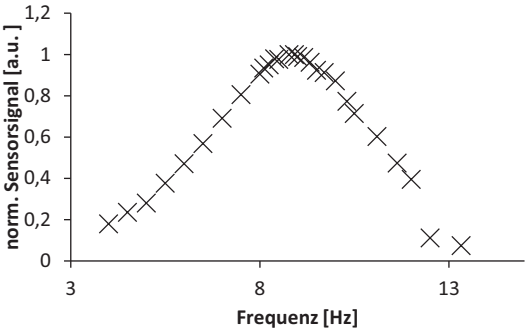


Abb. 3: Frequenzgang des photoakustischen Sensors. Maximales Signal bei 9 Hz.

Zur Charakterisierung des Messsystems wurde eine Messreihe von 0-1073 ppmV CH_4 in N_2 verdünnt aufgenommen. Dabei wurden die Messungen mit einem offenen und mit einem Kr gefüllten Strahler untersucht. Die Zelle wurde vor jeder Messung mit N_2 gespült und danach 4 min mit dem CH_4-N_2-Gemisch durchspült. Nachdem die Messzelle vollständig mit dem Messgas geflutet war, wurden die Magnetventile verschlossen und das photoakustische Signal detektiert. Das zeitliche Verhalten des Signals zeigt Abb. 4.

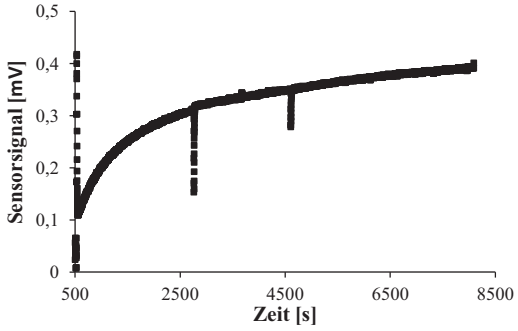


Abb. 4: Zeitlicher Verlauf des photoakustischen Signals für 4000 ppmV CH_4.

Zu Beginn erkennt man einen linearen Anstieg des Messsignals, welcher nach etwa 1000 s abflacht. Der Verlauf lässt sich durch die Erwärmung der Messzelle erklären. Dem System wurde stetig Energie zugeführt, sodass es sich nicht im thermischen Gleichgewicht befindet. Über ein entsprechendes Kühlsystem könnte eine Temperaturstabilisierung erfolgen, dies wurde jedoch nicht weiter untersucht, da der Anstieg des Messsignals unabhängig von der CH_4-Konzentration war und nur von der Temperaturzunahme der Messzelle abhing. Der Störeinfluss konnte deshalb über eine rechnerische Temperaturkompensation berücksichtigt werden. Es erfolgte eine Mittelung über 150 Messwerte je Messung (siehe Abb. 5).

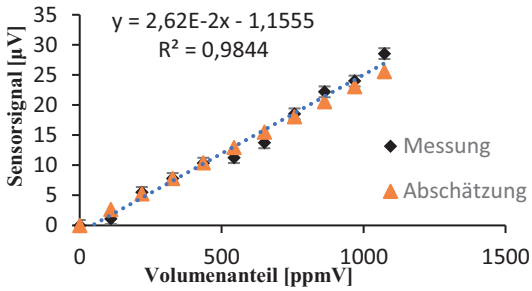


Abb. 5: Photoakustisches Signal (offsetkorrigiert) für 0-1000 ppmV CH_4 in N_2 bei Verwendung des offenen Strahlers.

In Abb. 5 wird die Mikrofonspannung für Volumenanteile von 0 bis 1073 ppmV CH_4, in N_2 verdünnt, dargestellt (offener Strahler). Der nach Gl. 3 zu erwartende lineare Anstieg des photoakustischen Signals mit steigendem Volumenanteil ist zu erkennen. Ebenso stimmt die Abschätzung in guter Näherung mit den experimentellen Messwerten überein. Das Detektionslimit wurde zu 108 ppmV (3σ) bestimmt.

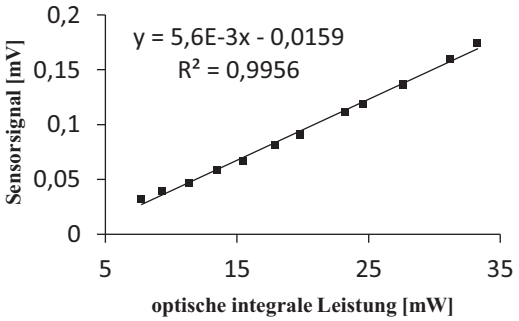


Abb. 6: Korrelation von Mikrofonspannung und optischer Leistung des offenen Strahlers für 4000 ppmV CH_4.

Abb. 6 stellt den Zusammenhang von optischer Leistung der Strahlungsquelle und dem Anstieg des photoakustischen Signals dar. Dazu wurde 4000 ppmV CH_4 in die Messkammer eingeleitet und das Signal bei 0,4 bis 1 W elektrischer Leistung aufgenommen. Die optische Leistung

am Ende der Zelle wurde über ein breitbandiges optisches Leistungsmessgerät vom Typ PM16-401 (Fa. Thorlabs) detektiert. Der nach Gl. 3 zu erwartende lineare Zusammenhang von der optischen Strahlerleistung und dem photoakustischem Signal konnte so experimentell bestätigt werden werden.

Als zweite Strahlungsquelle wurde ein, mit Kr geschlossener, MEMS-Membranstrahler verwendet. Auf dem Reflektor des Strahlers wurde ein Saphirfilterglas aufgebracht, welches den Strahler hermetisch verschließt. Der Strahler konnte so direkt in die Messzelle d.h. ohne Eintrittsfenster eingesetzt werden. Auch hier wurde eine Kalibriergerade von 0 bis 1073 ppmV CH_4 aufgenommen. Das Ergebnis ist in Abb. 7 dargestellt.

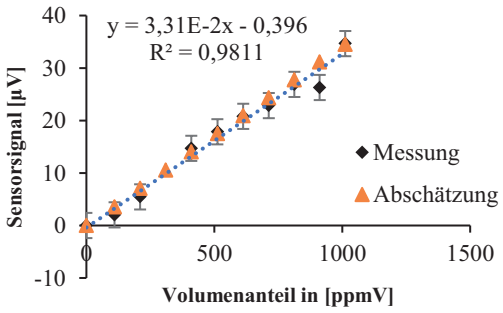


Abb. 7: Photoakustisches Signal (offsetkorrigiert) für 0-1000 ppmV CH_4 in N_2 bei Verwendung des geschlossenen mit Kr gefüllten Strahlers.

Die erhöhte optische Leistung spiegelt sich auch in einem größeren Anstieg der Kalibriergerade wieder. Da der Strahler nicht mehr durch ein zusätzliches Eintrittsfenster vom Gasvolumen entkoppelt ist, steigt auch das Hintergrundgeräusch, hervorgerufen durch den Strahler selbst, an. Aufgrund des hohen akustischen Grundpegels sind Messungen unter 200 ppmV (3σ) so nicht möglich. Das Auflösungsvermögen beträgt deshalb nur noch 74 ppmV (1σ), im Gegensatz zu 36 ppmV (1σ) bei Einsatz des offenen Strahlers mit Eintrittsfenster.

Zusammenfassung

Es wurde ein photoakustisches Gasmesssystem bestehend aus einem elektrisch moduliertem MEMS-Breitbandstrahler und einem MEMS-Mikrofon untersucht. Messungen haben gezeigt, dass die Detektion von Methan bis 108 ppmV möglich ist. In Kombination mit dem bereits bestehenden NDIR-Gasmesssystem ergeben sich so Möglichkeiten im Bereich Sensor-Fusion, um weitere Informationen über die Gasphase zu erhalten. Da beide Verfahren im selben Sensor gleichzeitig eingesetzt werden können, kann so ein redundantes System entwickelt werden. Dies bietet die Möglichkeit, die Ausfallsicherheit des Gassensors zu erhöhen, bzw. Fehlfunktionen des Sensors zu detektieren. Dies ist insbesondere für Anwendungen zur funktionalen Sicherheit (SIL) ein denkbarer Ansatz. Aufgrund der niedrigen Modulationsfrequenz wirken sich Störgeräusche aus der Umgebung auf die Messung aus. Der Sensor sollte daher zukünftig von seiner Umgebung möglichst akustisch entkoppelt sein. Da die Zelle während der Messung verschlossen sein muss, kommt es zu einer Erwärmung des Gassensors. Idealerweise sollte dieser daher temperaturreguliert werden, oder wie in dieser Arbeit gezeigt, über eine Temperaturkompensation das Messsignal korrigiert werden. Über optische Filter kann die Selektivität für verschiedene Gase erhöht und Quereinflüsse minimiert werden.

Danksagung

Die Autoren danken dem Bundesministerium für Bildung und Forschung für die Förderung des Verbundprojektes µPAS2 (Förderkennzeichen 03ET1317B).

Literaturnachweis

[1] T. Rück, R. Bierl, *Sensors and Actuators A 263 (2017), 501-509,* doi:10.1016/j.sna.2017.06.036
[2] T. Tomberg, M. Vainio,T. Hieta, L. Halonen, *Scientific Reports, Bd. 8, S. 1-7*, doi:10.1038/s41598-018-20087-9
[3] P. Patimisco, G. Scamarcio, Frank K. Tittel, V. Spagnolo, *Sensors 2014 (14), S. 6165-6206,* doi: 10.3390/s140406165
[4] A. Bell, *American Journal of Science, 118, S. 305-324,* doi: 10.2475/ajs.s3-20.118.305
[5] J. Huber, A. Ambs, J. Wöllenstein, *Procedia Engineering 120 (2015), S. 283-288,* doi: 10.1016/j.proeng.2015.08.616
[6] J. Huber, K. Schmitt, J. Wöllenstein, *J. Sens. Sens. Syst., 5 (2016), S. 293–299,* doi:10.5194/jsss-5-293-2016

DOI 10.5162/sensoren2019/4.2.2

Miniaturisiere photoakustische Detektoren für den Nachweis fluorhaltiger Kältemittel

Christian Weber[1,2,], Mahmoud El-Safoury[1,*], André Eberhardt[1], Dr. Katrin Schmitt[1,2], Prof. Dr. Jürgen Wöllenstein[1,2]*

[1] Fraunhofer Institut für physikalische Messtechnik IPM, 79110 Freiburg i. Br.
[2] Institut für Mikrosystemtechnik – IMTEK, Universität Freiburg 79110 Freiburg i. Br.
** Beide Autoren haben im gleichen Umfang zu dieser Arbeit beigetragen.*
christian.weber@ipm.fraunhofer.de; mahmoud.el-safoury@ipm.fraunhofer.de; andre.eberhardt@ipm.fraunhofer.de; katrin.schmitt@imtek.uni-freiburg.de; juergen.woellenstein@ipm.fraunhofer.de

Kurzzusammenfassung

Durch die Umstellung auf klimafreundliche Kältemittel im KFZ-Bereich, kam es in den letzten Jahren zu einer wachsenden Nachfrage nach Messtechnik für diese Kältemittel. Wir präsentieren im Folgenden unsere Arbeiten an photoakustischen Detektoren für die flourhaltigen Kältemittel R123a und R1234yf. Durch Simulationen konnten die Detektordimensionen optimal für beide Kältemittel ausgelegt und somit zuverlässige und empfindliche Detektorelemente geschaffen werden. Das in ersten Versuchen bestimmte Signal-zu-Rausch-Verhältnis reicht aus, um das Kältemittel R1234yf mit einer Auflösung 0,5 vol.% über den gesamten Konzentrationsbereich zuverlässig bestimmen zu können

Keywords: Photoakustische Detektoren, Photometer, Kältemittel, R1234yf, R134a

Einleitung

Kältemittel aus mobilen und stationären Kälteanlagen sind maßgeblich an den Emissionen von fluorierten Treibhausgasen weltweit beteiligt. In KFZ Klimaanlagen betrifft dies vor allem das weit verbreitete R134a (1,1,1,2-Tetrafluorethan). Es weist ein Global Warming Potential (GWP) von 1430 verglichen mit CO_2 auf [1] und ist aufgrund seiner Klimaschädlichkeit seit dem Jahr 2017 für Neuwagen innerhalb der EU verboten [2]. Als alternatives Kältemittel kommt seitdem das ebenfalls fluorierte R1234yf (2,3,3,3-Tetrafluorpropen) zum Einsatz, welches mit einem GWP von 4 deutlich weniger klimaschädlich ist [1]. Die thermodynamischen Eigenschaften beider Kältemittel ähneln sich sehr stark, weshalb die gleichen Kältekreislauf-Komponenten zum Einsatz kommen können. Da jedoch die Öle und Dichtungen der Kälteanlagen auf das jeweilige Kältemittel hin optimiert sind, gilt es im Servicefall das Kältemittel eindeutig zu identifizieren, um eine Falschbefüllung zu verhindern. Daneben sollen auch Verunreinigungen der Kältemittel erkannt werden. Das gilt besonders, wenn ein Teil des wesentlich günstigeren R134a dem R1234yf zugemischt wird um mit dieser Mischung den Kältekreislauf zu befüllen.
In diesem Artikel beschreiben wir den Aufbau und die Auslegung photoakustischer NDIR-Detektoren zum optischen Nachweis und der eindeutigen Trennung beider Kältemittel. Die Detektoren sind dazu geeignet in einem *Zwei-Kammer-System* eingesetzt zu werden, welches nur die Zielgase zur optischen Filterung verwendet und somit keine weiteren Filter benötigt. Somit dienen die Detektoren als kostengünstige Grundlage für ein sensitives Messsystem zur Konzentrationsbestimmung dieser Kältemittel.

Grundlagen und Messprinzip

Die entwickelten Detektoren beruhen auf dem potoakustischen Effekt zur Intensitätsmessung von Infrarotstrahlung in gasspezifischen Wellenlängenbereichen. Der Effekt wurde 1880 von Alexander Greyham Bell erstmalig publiziert [3]. Er beschreibt den Zusammenhang zwischen der Absorption von Strahlung in abgeschlossenen Gasvolumina und des darauffolgenden Druckanstiegs, der als Schallwelle wahrgenommen werden kann. Stimmt die Photonenenergie der Infrarotstrahlung mit der Energie eines möglichen Schwingungs-Rotations-Überganges der Gasmoleküle überein, so kann das Photon absorbiert und dabei der Zustandsübergang angeregt werden [4]. Durch nichtstrahlende Relaxation des Zustandes wird die absorbierte Energie in Wärme umgewandelt, die im angeschlossenen

Gasvolumen zu einem Druckanstieg führt, der sich im dynamischen Fall als Schallwelle im Gas zeigt. Über ein Mikrofon lässt sich die Amplitude der Schallwelle sehr sensitiv nachweisen, weshalb sich der Effekt messtechnisch nutzen lässt.
Der photakustische Effekt eignet sich direkt zur Bestimmung von Gaskonzentrationen, da die absorbierte Leistung über das Lambert-Beer'sche Gesetz direkt proportional zur Konzentration des Zielgases ist.
Da die Energien der Schwingungs-Rotations-Übergänge für jedes Molekül spezifisch sind, lässt sich über die Auswahl des infraroten Spektrums, eine Selektivität des Sensors auf das Zielgas erreichen. Für Breitbandquellen, wie thermische Emitter, wird dies in der Regel über optische Filter realisiert.
Der hier beschriebene Detektor arbeitet nach dem 2-Kammer-Prinzip, das heißt, er detektiert die Gaskonzentration indirekt als Detektor eines Absorptionsspektrometers (vergl. Abbildung 1). Dazu wird das Zielgas hermetisch im Detektor eingeschlossen, womit er über die fixierte Konzentration eine definierte Sensitivität auf Infrarotstrahlung aufweist. Die Besonderheit gegenüber anderen Infrarotdetektoren ist dabei die mit dem Absorptionsspektrum des eingeschlossenen Gases verknüpfte, spektral spezifische Detektorantwort.

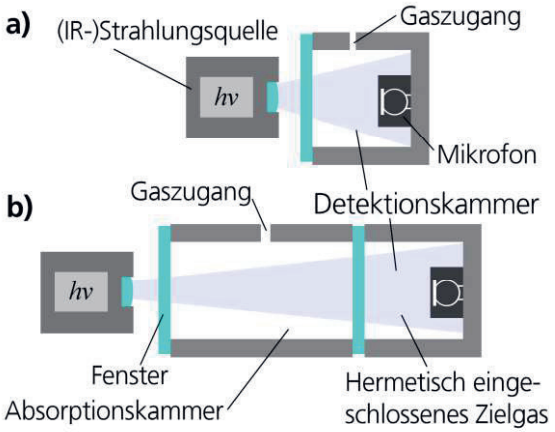


Abb. 1: Vergleich des Ein-Kammer-Prinzips (a) mit dem Zwei-Kammer-Prinzip (b) bei photoakustischen Messsystemen. Bei (a) tritt das zu messende Gas direkt in die Detektionskammer ein, während bei (b) das Zielgas in dieser Kammer hermetisch eingeschlossen ist. Das Gas wird somit als Strahlungsdetektor verwendet, mit dem die Reststrahlung, selektiv nach Durchtritt durch die Absorptionskammer, gemessen wird.

Der Detektor ist über den photoakustischen Effekt nur für die Spektralanteile empfindlich, in denen das Zielgas absorbiert.

Absorptionsspektren der zu messenden Kältemittel

Die Auslegung der photoakustischen Detektoren wurde mittels numerischer Verfahren simuliert. Die Grundlage dieser Simulationen bilden hochauflösende Infrarotspektren beider Kältemittel mit quantitativer Aussagekraft. Da diese in der Literatur nur mit unzureichender Genauigkeit zu finden waren, wurden sie mittels FTIR-Spektroskopie bestimmt. Dazu wurden Konzentrationen von 0,5 bis 100 vol.% der Kältemittel in einer 10 cm Küvette eingestellt und die Absorptionsspektren vermessen. Anschließend wurden daraus die spektralen Absorptionskoeffizienten von R1234yf und R134a errechnet. In Abbildung 2 sind die spektralen Absorptionskoeffizienten beider Kältemittel, über der IR-Wellenlänge auftragen, dargestellt.

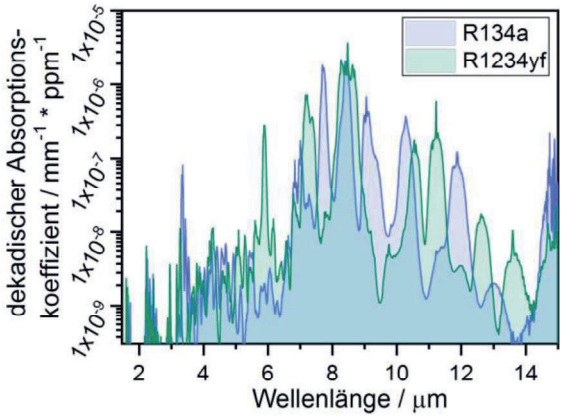


Abb. 1: Spektrale Absorptionskoeffizienten von R134a und R1234yf über der Wellenlänge der anregenden Infrarotstrahlung aufgetragen.

Die Bestimmung der Koeffizienten zeigt, dass beide Kältemittel im gesamten Bereich von 3 bis 15 µm Infrarotstrahlung absorbieren. Sowohl bei R1234yf als auch bei R134a liegt die Hauptabsorptionsbande bei ca. 8,5 µm. Daneben zeigen sich mehrere starke Banden zwischen 7 und 12 µm. Die Auffächerung in einzelne Absorptionslinien ist, besonders im Langwelligen, sehr gering ausgeprägt. Mit Ausnahme des Q-Zweigs, zeigen sich die starken Banden solide und detailarm.

Simulationen zur Auslegung der Detektoren

Die optimalen der Detektordimensionen konnte für beide Kältemittel über numerische Simulationen der Infrarotabsorption bestimmt

werden. Um die Komplexität des photoakustischen Effektes auszuklammern, wurde sich auf die spektrale Leistungsdichte, welche im Detektor absorbiert wird, beschränkt. Quantitative Aussagen zur absoluten Druckamplitude im Detektor selbst lassen sich so nur eingeschränkt über Vergleiche des Simulationsmodells mit Messungen an bereits vorhandenen, mit anderen Gasen gefüllten, Detektoren machen.
Hauptziel der Simulationen war es, eine möglichst hohe Sensitivität auf das jeweilige Ziel-Kältemittel zu erreichen, während nur geringe Querempfindlichkeiten auf das andere Kältemittel und andere Umgebungsgase vorhanden sind. Dabei muss der Detektor das Zielgas über den gesamten Bereich von 0 bis 100 % detektieren können.
Da die optimale Detektorgeometrie auch vom Rest der Messstrecke abhängt, wird er innerhalb eines sogenannten *Drei-Kammer-Systems* simuliert. Dabei wird das Zwei-Kammer-Prinzip um eine weitere Kammer erweitert, die sich zwischen Strahlungsquelle und Absorptionskammer (vergl. Abb. 1) befindet, ergänzt.
Diese *Filterkammer* kann Gase (sog. Filtergase) enthalten, gegenüber denen Querempfindlichkeiten verringert werden sollen.

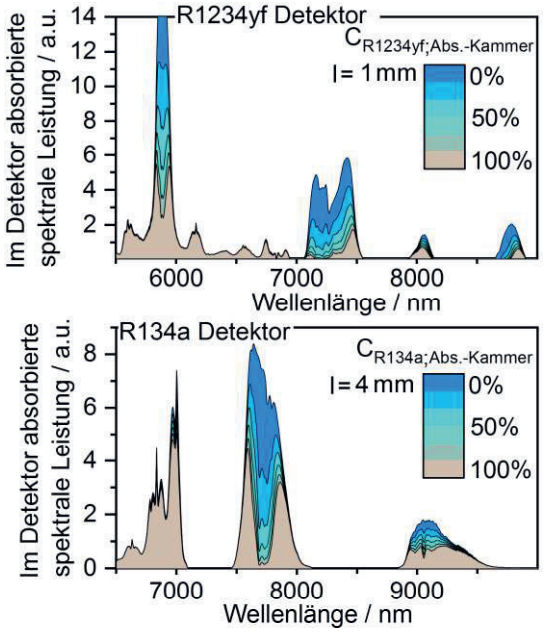


Abb. 3: Berechnete spektrale absorbierte Leistung im Detektor mit R1234yf und R134a unter verschiedenen Konzentrationen des Zielgases in der Absorptionskammer (Länge: 1 mm bzw. 4 mm). In beiden Fällen war der Detektor rund 3 mm lang und hatte einen Durchmesser von 8 mm.

Dies funktioniert, in dem in der Filterkammer alle Spektralbereiche entfernt werden, die für Querempfindlichkeiten auf die Filtergase verantwortlich sind. Folglich werden die Bereiche, in denen kein spektraler Überlapp zwischen Filter- und Zielgas vorhanden ist, am stärksten gewichtet.
Als Strahlungsquelle wurde in den Simulationen ein thermischer Strahler mit 2x2 mm strahlender Fläche, einem Emissionsgrad von 0,9 und einer Temperatur von 700°C angenommen.
Die Simulationen zeigen, dass bereits 3 mm Absorptionslänge in der Detektorkammer ausreichen, um die einfallende Strahlung im Bereich der Hauptbanden vollständig zu absorbieren. Als Durchmesser für die Absorptionskammer wurden 8 mm gewählt, um das Mikrofon später gut integrieren zu können. Die optimale Absorptionskammerlänge liegt im Fall von R1234yf bei 1mm und bei R134a bei 4 mm. Zur Verringerung der Querempfindlichkeit auf das jeweils andere Kältemittel wurden je 50 mm lange Filterzellen verwendet. Sie wurde im Fall des R1234yf Detektors mit 100% R134a simuliert und im Fall des R134a Detektors mit 10% R1234yf.

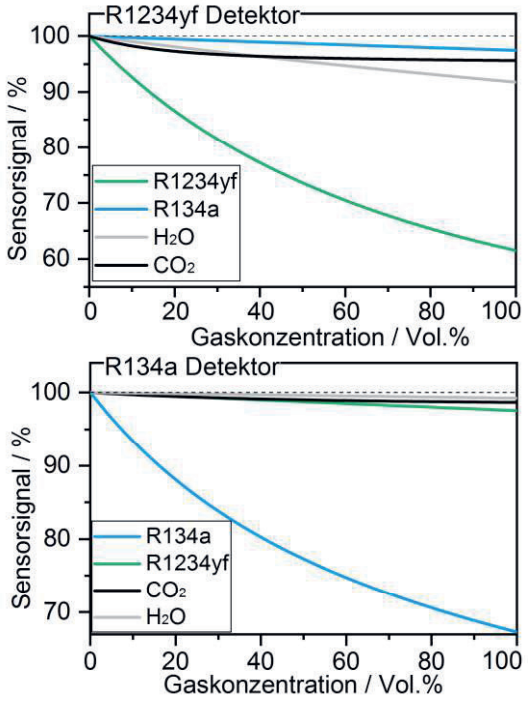


Abb. 4: Simulierte Gasreaktion der Detektoren auf die beide Kältemittel, Luftfeuchte und CO_2. Die Dimensionen und die Filterzellenfüllung sind analog zu denen in Abb. 3.

In Abbildung 3 sind die simulierten spektralen, im Detektor absorbierten Leistungen bei Variation der Gaskonzentration in der Absorptionskammer dargestellt. Durch die Filterzelle trägt die Hauptbande bei 8,5 µm in beiden Fällen kaum zum Detektorsignal bei. Beim R1234yf Detektor wird die Änderung des Gassignals hauptsächlich durch die Banden bei 5,8 und 7,3 µm bestimmt. Zwischen 5 und 7 µm liegt ein Spektralbereich, der zwar zum Gesamtsignal beiträgt, für die Gasreaktion aber unerheblich ist.
Die Änderung des Detektorsignals des R134a Detektors wird fast ausschließlich durch die Bande bei 7,7µm bestimmt. Auch hier tragen Banden < 7 µm stark zum Gesamtsignal bei, reagieren aber kaum auf das Gas in der Absorptionskammer.
Um die Empfindlichkeiten und Quer-Empfindlichkeiten zu vergleichen, wurde in weiteren Simulationen die Reaktion der Detektoren auf beide Kältemittel und die Umgebungsgase H_2O und CO_2 betrachtet. Ausgewertet wurde die integrale, im Detektor absorbierte Leistung. Die Simulationsergebnisse sind in Abb. 4 gegen die jeweiligen Gaskonzentrationen aufgetragen dargestellt.
Die simulierte Empfindlichkeit des R1234yf Detektors auf R1234a ist mit 0,35% pro vol.% ca. 17-mal stärker als die auf R134a. Bei hohen Konzentrationen sinkt sie bis auf 0,07% pro vol.% ab, bleibt dabei aber noch über der Empfindlichkeit auf R134a, die konstant bei 0,012% pro vol.% liegt. Die Empfindlichkeit des Detektors auf CO_2 ist mit bis zu 0,1 % pro vol% bei niedrigen Konzentrationen verhältnismäßig hoch. Im Falle einer Kalibrierung mit Umgebungsluft sollte der Filterzelle daher CO_2 zugegeben werden.
Insgesamt zeigen die Simulationen, dass es mit beiden Detektoren in diesen Dimensionen möglich ist, die Kältemittel zuverlässig zu trennen und auch quantitativ zu messen.

Aufbau des Detektors

Ausgehend von den Maßen die in den Simulationen bestimmt wurden, konnten die Detektoren auf Basis eines TO-5 Gehäuses hergestellt werden. Die elektrische Kontaktierung wurde über sechs Drähte realisiert, die durch Glasdurchführungen im Sockel eingelassen sind (Sonderanfertigung der Fa. SIMEK; Ilmenau).
Zusätzlich befinden sich im Sockel noch zwei Rohre, über die eine Befüllung des Detektors mit den Kältemitteln möglich wird. Im Detektor dient eine Platine mit 0,8 mm Stärke als

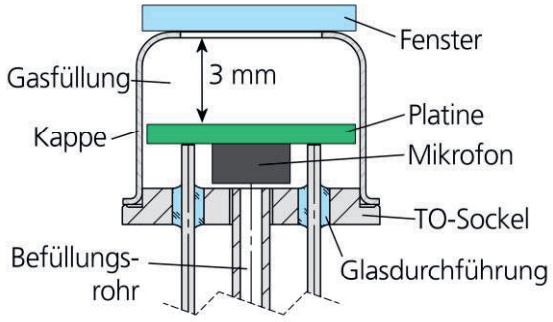


Abb. 5: Querschnitt durch den Detektoraufbau basierend auf einem TO-5 Gehäuse. Das Mikrofon wird über eine Platine auf die Drähte aufgelötet. Es handelt sich um ein Bottom-Port Mikrofon (ICS-40720 der Firma TDK)

elektrische Verbindung zwischen einem SMD MEMS-Mikrofon (ICS-40720 von TDK) und den Drähten der Glasdurchführung. Der Schallzugang zum Mikrofon findet seitlich über eine Fräsnut in der Platine statt (Siehe Abbildung 6).
Die gesamte photoakustische Detektorkammer wurde über das Auflöten einer TO-Kappe mit einem ebenfalls aufgelöteten Fenster aus unterseitig metallisiertem Calciumfluorid realisiert.

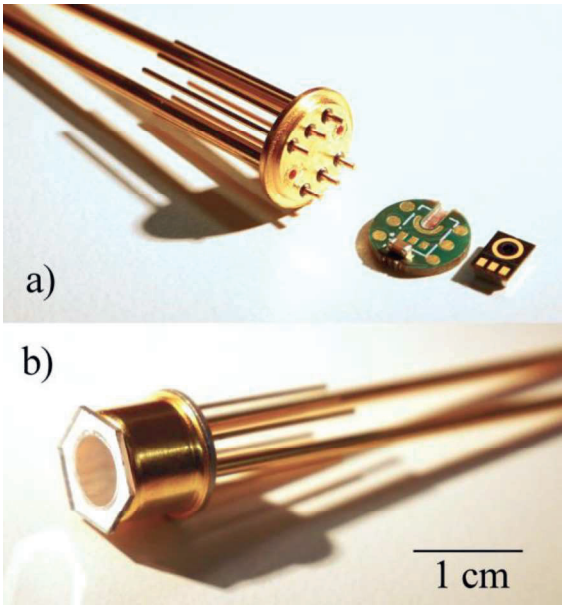


Abb. 6: Fotos des TO-Sockels mit Platine und Mikrofon vor dem Zusammenbau (a) und des aufgebauten Detektors, inklusive aufgelötetem CaF_2-Fenster (b). Der Schallzugang zum Mikrofon findet über die Fräsnut in der Platine statt.

In Abbildung 5 ist der Querschnitt durch den gesamten Detektoraufbau dargestellt.
Alle Lötverbindungen von Fenster und Elektronik konnten in einem Reflow-Löt-Schritt hergestellt werden. In Abbildung 6 ist ein Foto des fertig verlöteten Detektorgehäuses zu sehen.

Befüllung des Detektors mit den Kältemitteln

Vor der Füllung des Detektors mit den Kältemitteln wurde er bei 110°C für 24h mit reichen Stickstoff gespült. Dies dient in erster Linie dazu, ad- und absorbierte Atmosphärengase, insbesondere Wasser, aus dem Detektor zu entfernen. Ein zusätzlicher Anteil infrarotaktiver Fremdgase würde die Querempfindlichkeiten auf diese signifikant erhöhen.
Um den photoakustischen Effekt während der Gasfüllung abschätzen zu können, wurde der Detektor in einen Versuchsaufbau eingebaut (Abbildung 7). In diesem steht er einem thermischen Strahler gegenüber (JSIR-350-4 Fa. MicroHybrid). Der Strahler wird mit 250 Hz von einer eigens angefertigten Elektronik sinusförmig moduliert. Gleichzeitig wird das Mikrofonsignal des Detektors über die Elektronik aufgenommen und nach einer Lock-In-Filterung ausgegeben. In Abbildung 8 ist das gefilterte Detektorsignal während der Befüllung mit R1234yf gegen die Dauer der Befüllung aufgetragen.

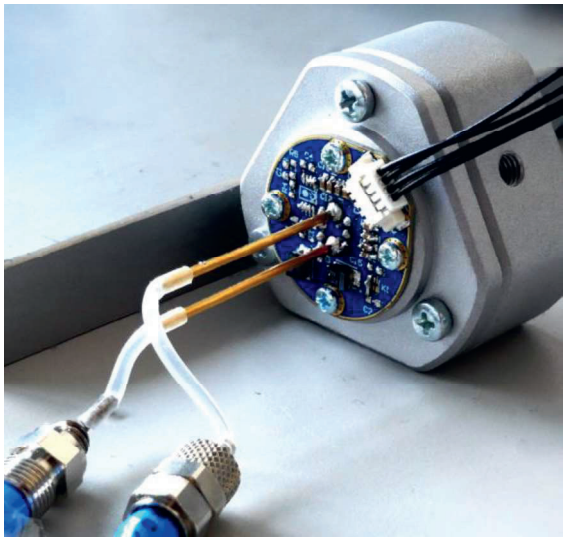


Abb. 7: Versuchsaufbau während der Befüllung des Mikrofons mit R1234yf. Der Detektor ist innerhalb des Aluminium-Gehäuses. Dem Detektor direkt gegenüber ist ein thermischer Strahler angeordnet.

Zum Zeitpunkt t=0 wird der Gasfluss im Detektor von reinem Stickstoff auf R1234yf umgeschaltet. Etwa 100 s nach dem Umschalten erreicht das Kältemittel den Detektor. Das Signal steigt durch den photoakustischen Effekt während der Befüllung stetig an und erreicht ab etwa 1200 s einen stationären Wert, ab dem davon ausgegangen werden kann, dass der Detektor komplett mit R1234yf gefüllt ist.
Vor der Füllung mit R1234yf zeigt der Detektor bereits ein Signal, das in etwa 8% des Signals mit Kältemittel entspricht. Dieses Nullsignal kann durch die Strahlungsabsorption an der Detektorstruktur und der damit einhergehenden Photoakustik erklärt werden.

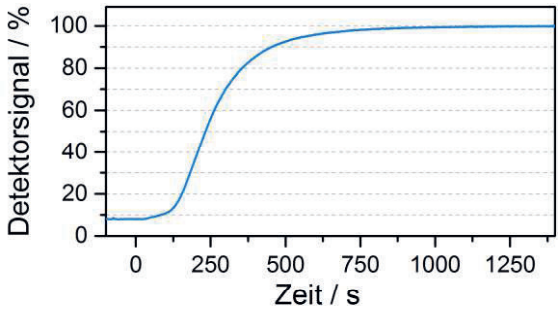


Abb. 8: Detektorsignal bei der Befüllung gegen die Zeit aufgetragen. Zum Zeitpunkt t=0 wird der Gasfluss von Stickstoff auf R1234yf umgeschaltet.

Nach der Befüllung des Detektors mit R1234yf, wurden die Befüllungsrohre mithilfe eines Schraubstockes abgequetscht und anschließend mit Lötzinn verschlossen, um die hermetische Dichtigkeit der Gasfüllung zu gewährleisten.

Bestimmung von Stabilität und Rauschen des Detektorsignals

Als Maß für die Signalqualität wurde das Detektorsignal über 64 h mit Umgebungsluft zwischen thermischen Emitter und R1234yf-Detektor aufgezeichnet. Der Emitter wurde mit 250 Hz sinusförmig moduliert und der Abstand zum Detektor betrug bei dieser Messung 5 cm. Sie fand unter Laborbedingungen statt, wobei die Umgebungstemperatur eine maximale Variation von 4,9 °C (Mittelwert 20,9°C) aufwies. Die Mittelungszeit des Lock-In-Filters wurde auf 1 s eingestellt und die Daten alle 0,5 s ausgelesen.
Abbildung 9 zeigt die aufgenommenen Daten der Messung im Zeitbereich und die daraus errechnete Allan-Abweichung. In den ersten 30 Minuten erwärmt sich die Messkonfiguration

durch das Betreiben des thermischen Emitters um 11,4 °C auf knapp 31°C. Dabei sinkt das Detektorsignal um ca. 1,5 %, verglichen mit dem Anfangswert, ab. Im weiteren Verlauf, driftet das Temperatursignal um ca. 0,7%, was größtenteils auf die Schwankungen der Umgebungstemperatur zurückzuführen ist. Ein durchgängiges Abnehmen des Signals, was auf eine Undichtigkeit des Detektors hindeuten würde, ist nicht zu beobachteten. Somit kann ein signifikantes Gasleck der Detektorkammer ausgeschlossen werden.
Die aus den Daten extrahierte Allan Abweichung [5] zeigt, dass bei der Datenrate von 0,5 s die dreifache Standardabweichung und

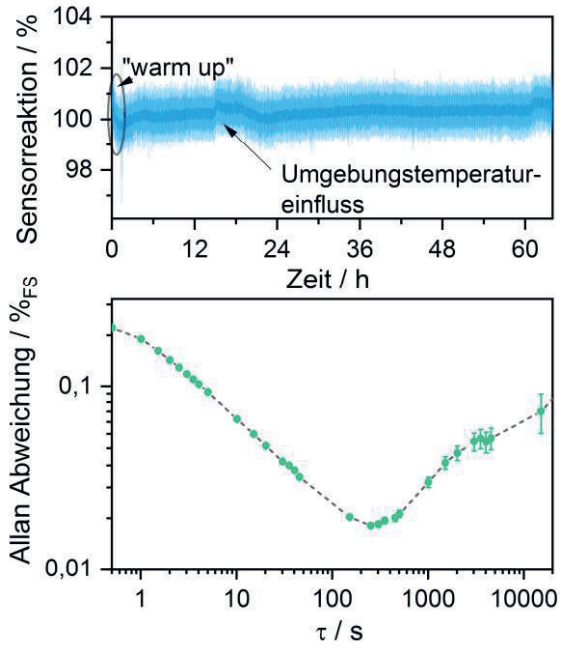


Abb. 9: Ergebnisse der Signalqualitätsmessung am R1234yf Detektor. Oben: Über 64h aufgenommenes Detektorsignal mit Umgebungsluft zwischen thermischem Emitter und Detektor. Unten: Daraus abgeleitete dreifache Allan Abweichung.

damit die Auflösungsgrenze bei ca.0,66 % des „Fullscale"-Signals liegt. Unter der Annahme der simulierten Empfindlichkeit, reicht eine Mittelungszeit von acht bis zehn Sekunden um die geforderte Sensorauflösung von 0,5 vol.% R1234yf zuverlässig zu erreichen.
Die auf die Temperatur zurückzuführende Detektordrift muss im späteren Messsystem über einen Temperatursensor kompensiert werden, oder sie kann über eine Temperaturstabilisierung der Messeinheit vermieden werden.

Zusammenfassung

Ziel dieser Arbeit war die Auslegung photoakustischer Detektoren für die Kältemittel R134a und R1234yf. Anhand von Simulationen wurden die wichtigsten Detektorabmessungen festgelegt und die Sensorperformance abgeschätzt. Durch ein speziell angefertigtes TO-Gehäuse konnte ein Detektor realisiert werden, der mit dem Kältemittel R1234yfgefüllt wurde. Erste Messungen zeigen ein vielversprechendes Signal-zu-Rausch-Verhältnis für die Reinheitsbestimmung der Kältemittel.

Ausblick

In Zukunft soll neben dem Detektor für R1234yf auch ein Detektor für R134a befüllt und vermessen werden. Anschließend erfolgt die Integration beider Detektoren in ein Messsystem, das in der Lage sein soll die Reinheit beider Kältemittel zuverlässig zu bestimmen.

Danksagung

Die durchgeführten Arbeiten wurden vom Bundesministerium für Bildung und Forschung im Rahmen des KMU-Innovativ Verbundprojektes CONTECT-R finanziert.

Literatur

[1] B. Minor, M. Spatz, HFO-1234yf low GWP Refrigerant Update. In: InternationalRefrigeration and Air Conditioning Conference at Purdue, West Lafayette, IN,USA, 2008, Paper No. 2349.
[2] Richtlinie 2006/40/EG des europäischen Parlaments und des Rates, 17.06.2006 über Emissionen aus Klimaanlagen in Kraftfahrzeugen und zur Änderung der Richtlinie 70/156/EWG des Rates
[3] A. G. Bell, "On the Production and Reproduction of Sound by Light," Am. J. Sci., vol. XX, no. 118, pp. 305– 324, 1880.
[4] Z. Bozóki, A. Pogány, G. Szabó: "Photoacoustic In-struments for Practical Applications: Present, Poten-tials, and Future Challenges", Applied Spectroscopy vol. 46; 01/2011
[5] P. Werle, R. Mücke, and F. Slemr, "The limits of signal averagingin atmospheric trace gas monitoring by tunable diodelaser absorption spectroscopy," Appl. Phys. B 57, 131–139 (1993).

DOI 10.5162/sensoren2019/4.2.3

Filterrotationsspektrometer für den Nachweis von Ethen im ppb- Bereich

André Eberhardt[1], Marie-Luise Bauersfeld[1], Katrin Schmitt[1,2], Jürgen Wöllenstein[1,2]
[1] Fraunhofer-Institut für Physikalische Messtechnik IPM, Heidenhofstr. 8, 79110 Freiburg
[2] Professur für Gassensoren, Institut für Mikrosystemtechnik (IMTEK), Technische Fakultät der Universität Freiburg, Georges-Köhler-Allee 102, 79110 Freiburg

Zusammenfassung

Ethen ist eines der wichtigsten Phytohormone im Reifungsprozess klimakterischer Früchte. Die Menge des emittierten Ethens liefert Informationen über den aktuellen Reifezustand dieser Früchte. Deshalb kann es in der Lieferkette für die Umsetzung fortschrittlicher Prozesse, wie dem First-Expired-First-Out-Prozess (FEFO), eingesetzt werden. Diese Arbeit beschreibt den aktuellen Stand der Entwicklung eines kostengünstigen Filterrotationsspektrometers für den Nachweis von Ethen in der klimakterischen Reifephase von Früchten. Das Spektrometer basiert im Wesentlichen auf Komponenten, die auch in nichtdispersiven Infrarotsensoren (NDIR) zum Einsatz kommen. Durch eine modellbasierte Auswertung der Spektren konnte die Stabilität eines ersten Laborprototyps gegenüber klassischen NDIR-Sensoren deutlich verbessert werden. In Verbindung mit einer optimierten Langwegzelle konnte die Nachweisgrenze bis auf 150 ppb reduziert werden.

Keywords: Ethen, Gassensorik, Optik, Spektrometer, NDIR, Fruchtlager

Einleitung

Die anhaltend hohe Verschwendung von leicht verderblichen Lebensmitteln, wie Obst und Gemüse, ist aktuell ein großes ökologisches und ökonomisches Problem. Im Jahr 2011 gingen bis zu 50% der Ernte auf dem Weg zum Konsumenten verloren [1]. Durch die Einführung moderner Transport-, Lager- und Verteilungsprozesse können diese Verluste reduziert werden [2]. Für klimakterische Früchte sind Ethenemissionen ein guter Indikator für den aktuellen Reifezustand. Je nach Fruchtart und Reifephase können die nachzuweisenden Konzentrationen bis in den unteren ppb-Bereich reichen [3]. Deshalb werden für die Überwachung der Ethenkonzentration auf dem Transport oder in Fruchtlagern hochempfindliche, langzeitstabile und kostengünstige Messsysteme benötigt. Es gibt eine Vielzahl unterschiedlicher Sensoren, wie Gaschromatografen, Laserspektrometer, Metalloxidgassensoren oder elektrochemische Zellen, die sich grundsätzlich für den Nachweis von Ethen im ppb-Bereich eignen [4-7]. Jedoch sind diese Sensoren entweder nicht selektiv und langzeitstabil genug oder sie sind für den flächendeckenden, kommerziellen Einsatz zu kostenintensiv.

Als Ausgangspunkt für diese Arbeit wurden nichtdispersive optische infrarot Sensoren (NDIR) ausgewählt. Sie zeichnen sich durch einen einfachen und kostengünstigen Aufbau aus. Die erreichbare Nachweisgrenze dieser Sensoren liegt im niedrigen zweistelligen ppm-Bereich; zu wenig für diese Anwendung [8, 9]. Die meisten NDIR-Sensoren verwenden mehrere Spektralkanäle für die Bestimmung der Gaskonzentration. Dabei befindet sich der Referenzkanal in einem Spektralbereich in dem möglichst kein anderes relevantes Gas absorbiert. Die restlichen Kanäle decken Teile oder ganze Absorptionsbanden der nachzuweisenden Gase und eventueller Querempfindlichkeitsgase ab. Der Referenzkanal befindet sich in den meisten Fällen im Bereich um 3,9 µm, weil hier ein atmosphärisches Fenster vorliegt. Für den Nachweis von Ethen wird der Wellenlängenbereich um 10,5 µm verwendet, weil sich hier die stärksten Absorptionslinien befinden und nur wenige Querempfindlichkeitsgase wie Ammoniak (NH_3), Kohlenstoffdioxid (CO_2) oder Wasser (H_2O) vorkommen. Der große Spektrale Abstand zwischen dem Referenz- und Messkanal stellt sehr große Anforderungen an den NDIR-Sensor. Die Ursache ist die thermische Strahlungsquelle, die üblicherweise in diesen Systemen eingesetzt wird. Das Spektrum dieser Quellen zeichnet sich durch eine starke Temperatur- und Wellenlängenabhängigkeit aus. Somit führen geringe Temperaturschwankungen zu großen Abweichungen der bestimmten Gaskonzentration.

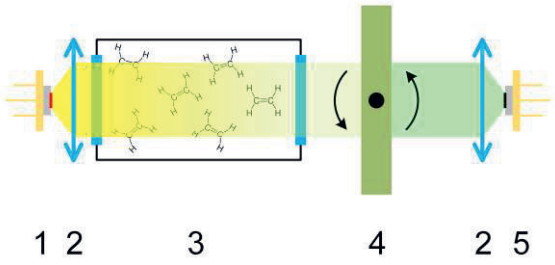


Abb. 1: Schematischer Aufbau des entwickelten Filterrotationsspektrometers: (1) Strahlungsquelle, (2) optische Komponenten, (3) Messzelle, (4) rotierender Interferenzfilter, (5) Strahlungsdetektor.

Um mehrere Spektralkanäle realisieren zu können, wird die Strahlung auf mehrere Detektorelemente mit unterschiedlichen Interferenzfiltern aufgeteilt. Dabei reduziert sich die verfügbare Energie pro Spektralkanal sehr stark.
Durch den Einsatz eines dispersiven Elements in Verbindung mit einem einzelnen Detektor können diese Nachteile überwunden werden. Diese Arbeit stellt eine Möglichkeit für ein solches Element vor, ohne den einfachen Aufbau der NDIR-Sensoren grundlegend zu ändern. Das Drehen eines Interferenzfilters im Strahlengang verschiebt dessen Zentralwellenlänge zu kleineren Werten, wodurch der Filter über Teile der Ethenabsorptionsbande bei 10,5 µm abgestimmt werden kann. Das auf diese Weise aufgenommene Spektrum erlaubt den Einsatz modellbasierter Auswertungsverfahren, wodurch die Robustheit gegenüber Systemdrift und Querempfindlichkeitsgasen zunimmt. In Verbindung mit einer für thermische Quellen optimierten optischen Langwegzelle können somit Nachweisgrenzen im ppb-Bereich erreicht werden. Diese Arbeit beschreibt einen ersten Prototyp, der die Bewertung der Leistungsfähigkeit dieses Konzepts erlaubt.

Messprinzip

Das Schema eines Filterrotationsspektrometers ist in Abb. 1 dargestellt. Es orientiert sich an klassischen NDIR-Sensoren und besteht aus einer breitbandigen thermischen Strahlungsquelle E_{BB} (1), optischen Komponenten im Strahlengang mit einem zusammengefassten Transmissionsspektrum T_{Opt} (2), einer Gasmesszelle mit definiertem optischen Weg T_{Gas} (3), einem rotierenden Interferenzfilter T_F (4) und einem Strahlungsdetektor mit einer spektralen Empfindlichkeit R_D (5).

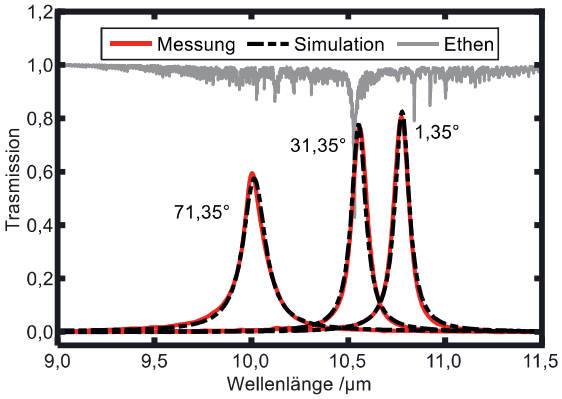


Abb. 2: Vergleich der simulierten und gemessenen Transmissionsspektren des eingesetzten Interferenzfilters im Vergleich zum Spektrum von Ethen (400 cm, 50 ppm).

Das Messsignal S hängt von der Gaskonzentration c und dem Rotationswinkel θ des Interferenzfilters ab. Es wird durch die Überlagerung der Teilspektren der Komponenten des Systems charakterisiert:

$$S(c,\theta) \propto \int_0^\infty E_{BB}(\lambda,T)\cdot T_{Gas}(\lambda,c)\cdot T_F(\lambda,\theta) \cdot T_{Opt}(\lambda)\cdot R_D(\lambda)\,d\lambda. \quad (1)$$

Die Schlüsselkomponente ist der Interferenzfilter. Da der Gangunterschied der im Filter interferierenden Teilwellen mit zunehmenden Rotationswinkeln kleiner wird, kann die Zentralwellenlänge λ_c des Filters stets nur zu kleineren Wellenlängen verschoben werden. Für dieses Ethenspektrometer wurde ein Filter mit einem Lorentz-Profil ausgewählt:

$$T_F(\lambda,\theta) = T_{max}(\theta)\cdot\frac{w(\theta)}{4\left(\lambda-\lambda_c(\theta)\right)^2+w(\theta)^2} \quad (2)$$

Aufgrund des komplexen Aufbaus von Interferenzfiltern sind die Peaktransmission T_{max} und die Halbwertsbreite w auch vom Rotationswinkel abhängig. Abb. 2 zeigt die Rotationswinkelabhängigkeit des verwendeten Filters mit einer Zentralwellenlänge von 10,78 µm und einer Halbwertsbreite von 90 nm für drei ausgewählte Winkel (1,35 °, 31,35 ° und 71,35 °) im Vergleich zu einem simulierten Spektrum der ν_7-Bande von Ethen für einen optischen Weg von 4 m und eine Konzentration von 50 ppm unter Normalbedingungen im Bereich zwischen 9,0 µm und 11,5 µm. Es wird deutlich, dass sich der Filter über die Hauptabsorptionslinien bei 10,53 µm abstimmen lässt. Die Rotationswinkelabhängigkeit der drei Parameter des Lorentz-Profils aus Gl. 2 wurde mithilfe von Spektralmessungen im Bereich zwischen -80 ° und +80 ° bestimmt. Die Ergebnisse sind als

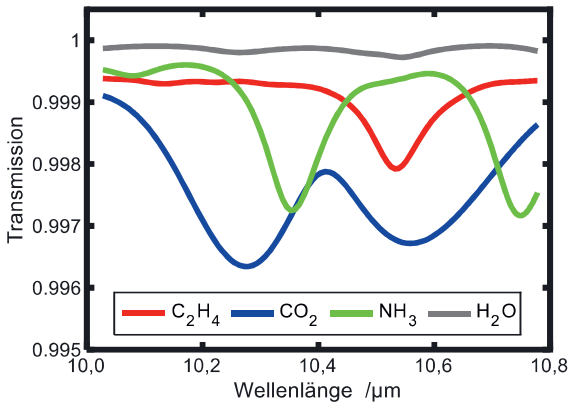


Abb. 3: Simulierte Transmissionsspektren für Ethen (1 ppm) und Ammoniak (1 ppm) in einer Lageratmosphäre für Bananen (14 °C, 90%-r.F., 3% CO_2) [10].

Simulation in Abb. 2 im Vergleich zu den Messwerten dargestellt. Die integrale geometrische Abweichung zwischen Messung und Simulation beträgt maximal 1,3%.
Mithilfe des rotationswinkelabhängigen Interferenzfilterspektrums wurden Transmissionsspektren:

$$T_{Gas} = {S_{Gas}(c,\theta)}\Big/{S_{Gas}(c=0,\theta)}, \qquad (3)$$

für 1 ppm Ethen (rot) im Vergleich zum wichtigsten Querempfindlichkeitsgas Ammoniak (1 ppm, grün) in einer Lageratmosphäre für Bananen (14 °C, 90%-r.F., 3%-CO_2) berechnet und in Abb. 3 dargestellt. Als Grundlage für die Simulationen wurde ein idealer schwarzer Strahler mit einer Temperatur von 800 °C und ein optischer Weg von 4 m unter Normaldruck verwendet. Diese Werte orientieren sich an den Parametern des vorgestellten Spektrometers. Für Ethen zeigt sich der charakteristische Absorptionspeak bei 10,53 µm und der Einfluss der schwächeren Absorptionslinien der ν_7-Bande, die eine Signalabnahme über den gesamten Abstimmbereich bewirken. Aufgrund der großen spektralen Breite des Interferenzfilters im Vergleich zur Breite der Absorptionslinien sinkt die Stärke des Hauptabsorptionspeaks von ca. 1,5% auf ca. 0,25%. Für das Querempfindlichkeitsgas NH_3 ist ein ähnlich starker Doppelpeak bei 10,34 µm und 10,75 µm erkennbar. Aufgrund der vergleichsweise großen CO_2-Konzentration in der Lageratmosphäre ist der markante Doppelpeak bei 10,26 µm und 10,57 µm stärker als der Ethenpeak. Aus diesem Grund muss CO_2 bei der Spektrenauswertung berücksichtigt werden. Im Gegensatz zu CO_2 sind die Absorptionslinien von H_2O so schwach, dass es

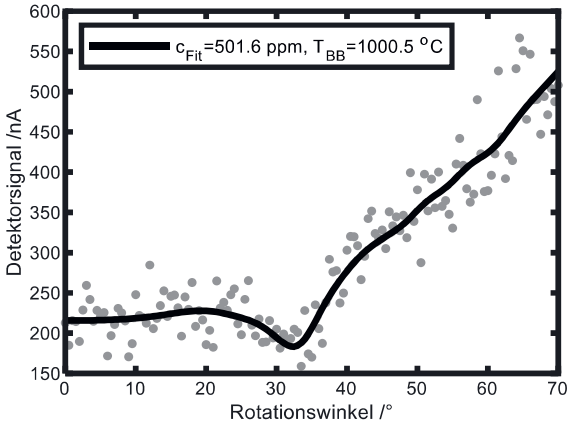


Abb. 4: Typisches Approximationsergebnis für das entwickelte Modell.

je nach untersuchter Fruchtart vernachlässigt werden kann.
Für die Bestimmung des Hintergrundspektrums $S_{Gas}(c=0,\theta)$ muss eine reine Stickstoffatmosphäre vorliegen. Im Feldeinsatz ist das häufige Spülen mit Stickstoff zu aufwendig, weshalb eine Lösung entwickelt wurde, die mit den Rohdaten $S_{Gas}(c,\theta)$ auskommt.
Wir entschieden uns für einen modellbasierten Ansatz, welcher ein möglichst genaues Modell des Spektrometers verwendet, um die Messdaten mit einem Levenberg-Marquardt-Algorithmus zu approximieren. Als Fitvariablen wurden im ersten Schritt die Emittertemperatur und die Ethenkonzentration eingesetzt. Um die Performance dieses Ansatzes zu überprüfen, wurde ein simuliertes Spektrum für Ethen (500 ppm, 400 cm) bei einer Emittertemperatur von 1000 °C simuliert und mit einem gaußförmigen Rauschen mit einer Amplitude von 10% überlagert. Abb. 4 zeigt ein typisches Ergebnis des Lösungsalgorithmus. Über 100 Iterationen wurde für die Temperatur ein Wert von 999,8±4,3°C und für die Konzentration ein Wert von 499,8±9,8 ppm erreicht. Somit beträgt die Unsicherheit für die Konzentration trotz des großen Rauschens nur ca. 2%.
Aufgrund dieser Ergebnisse wurde das Modell um die beiden Querempfindlichkeitsgase NH_3 und CO_2 erweitert. Der Einfluss dieser Gase auf die Ethenkonzentrationsbestimmung wurde in unterschiedlichen Gaszusammensetzungen für eine relative Rauschamplitude von 0,1% über jeweils 1000 Iterationen im Ethenkonzentrationsbereich bis 10 ppm ermittelt. Der Wert der Rauschamplitude ist charakteristisch für das vorgestellte Spektrometer. Der Einfluss auf die bestimmte Ethenkonzentration fällt vernachlässigbar gering aus. Die Abweichung einer Gasatmosphäre ohne Querempfindlichkeitsgase zu einer Atmosphäre mit 10% CO_2 und 50 ppm NH_3 beträgt 35 ppb und

liegt weit unterhalb der avisierten Nachweisgrenze. Über alle untersuchten Gasatmosphären wurde eine Nachweisgrenze (3 σ) von ca. 900 ppb für C_2H_4, 1,17% für CO_2 und 630 ppb für NH_3 simuliert. Die Unterschiede zwischen den einzelnen Gasen entsprechen den Erwartungen aufgrund der Stärke der Absorptionslinien im Abstimmbereich, vgl. Abb. 3.

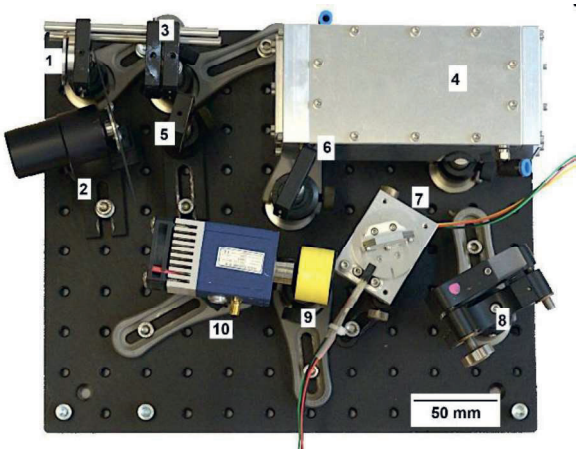


Abb. 5: Fotografie des optischen Teils des Spektrometers: (1) IR-Quelle, (2) opt. Chopper, (3) ZnSe-Linse, (4) optische Langwegzelle, (5) Umlenkspiegel, (6) ZnSe-Linse, (7) Filterrotationseinheit, (8) sphärischer Spiegel, (9) ZnSe-Linsen-Duplett und (10) Strahlungsdetektor (nach [12]).

Messsystem und Datenverarbeitung

Der optische Teil des Spektrometers ist in Abb. 5 dargestellt. Die IR-Strahlung einer thermischen Quelle (1), Hawkeye IR-30, wird mit einem optischen Chopper (2) direkt nach der Quelle mit einer Frequenz von 986 Hz moduliert um Störquellen aus dem Umfeld des Spektrometers auszuschließen. Danach wird die Strahlung mit einer ZnSe-Linse (3, *f = 22,2 mm*) in die Eingangsfeldebene einer speziell für große thermische Quellen optimierten Langwegzelle (4) fokussiert. Die Zelle ermöglicht für eine Eingangsapertur von 5 x 5 mm² einen optischen Weg von 4 m bei einer numerischen Apertur von 0,126 [11]. Die Strahlung, die die Messzelle verlässt wird mithilfe eines Umlenkspiegels (5) und einer weiteren ZnSe-Linse (6) in die zentrale Filterebene der Filterrotationseinheit (7) fokussiert. Die Filterrotationseinheit ist mit einem 1“-Interferenzfilter ausgestattet. Die Apertur der Einheit hat eine Größe von 24 x 4,5 mm² und ermöglichte im Aufbau somit einen maximalen Rotationswinkel von ca. 70 ° bevor es zu Abschattungseffekten kommt. Die transmittierte Strahlung wird mithilfe eines sphärischen Spiegels (8, f = 100 mm) über ein ZnSe-Linsen-Duplett (9, f = 16,7 mm) auf eine MCT-Fotodiode (10, PVM10.6-2TE) mit einer Kantenlänge von 1 mm abgebildet.

Das um 40dB vorverstärkte Signal wird zusammen mit dem Referenzsignal des Choppers an einen Lock-In-Verstärker (Perkin Elmer DSP 7265) mit einer Zeitkonstante von 100 ms weitergleitet. Anschließend wird das Ausgangssignal des Lock-In-Verstärkers mit einer NI PXI-4461 Datenerfassungskarte digitalisiert. Mithilfe einer Lichtschranke, die in die Filterrotationseinheit integriert ist, wird die Datenaufnahme über einen Triggerimpuls für jede Filterseite bei θ = 0 ° gestartet.

Abb. 6 zeigt den Vergleich eines gemessenen und simulierten Filterrotationsspektrums für ±70 ° bei einer Ethenkonzentration von 17 ppm im Stickstoff. Im Bereich zwischen ±50 ° wird eine sehr gute Übereinstimmung erreicht. Die maximale Abweichung zwischen Messung und Simulation beträgt 0,15%.

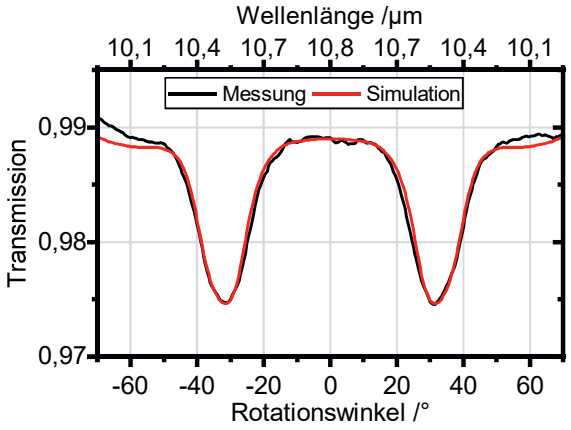


Abb. 6: Vergleich eines gemessenen und simulierten Filterrotations-Transmissionsspektrums für 17 ppm Ethen in Stickstoff.

Experimentelle Ergebnisse

Die Stabilität und Nachweisgrenze des Spektrometers wurden in Langzeitmessungen über einen Zeitraum von 550 h in einer Klimakammer bei einer Temperatur von 25 °C untersucht. Innerhalb dieses Zeitraums wurden zwei Konzentrationszyklen zwischen 20 ppm und 200 ppb Ethen in N_2 vermessen und 350 h mit N_2 gespült.

Die jeweiligen Konzentrationsstufen der Zyklen wurden 45 min konstant gehalten. Anschließend wurde für den gleichen Zeitraum mit N_2 gespült, um mögliche Abweichungen in der Baseline erfassen zu können. Die einzelnen Konzentrationsstufen wurden mithilfe einer Gasmischanlage analog zum System von Kneer et al. [13] aus einer Gasflasche mit

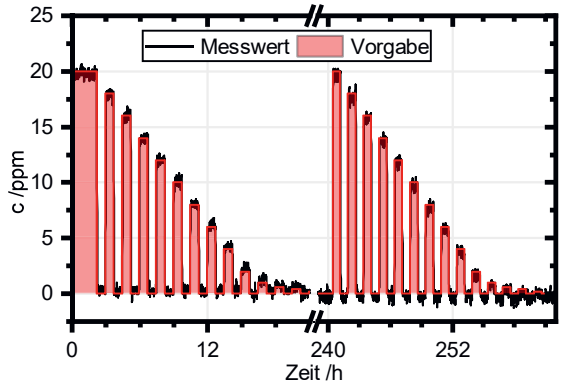


Abb. 7: Transienter Verlauf der bestimmten Ethenkonzentration für zwei Konzentrationszyklen zwischen 20 ppm und 200 ppb Ethen, die im Abstand von 240 h aufgenommen wurden. Die Abweichung zwischen beiden Zyklen beträgt ca. 500 ppb.

100 ppm Ethen bei einem Summenfluss von 0,5 l/min gemischt. Der Aufbau führte zu einer Unsicherheit der bereitgestellten Ethenkonzentration im Bereich zwischen 20 ppb bei 200 ppb Ethen und 400 ppb bei 20 ppm Ethen. Die transienten Signalverläufe der beiden Zyklen sind in Abb. 7 dargestellt. Die jeweils letzten 25 min jeder Konzentrationsstufe wurden gemittelt und in Abb. 8 dargestellt. Aufgrund einer Systemdrift von ca. -55 ppb/d wurden die Daten des zweiten Zyklus für die Darstellung um +550 ppb verschoben. Nach Ausgleich der Drift befindet sich die Abweichung zwischen den beiden Zyklen in der Regel innerhalb der Unsicherheit der Gasmischanlage. Für die Bestimmung der erreichbaren Nachweisgrenze wurde über den Zeitraum, in dem mit Stickstoff gespült wurde, die AllanStandardabweichung [14] berechnet und in Abb. 9 doppelt logarithmisch dargestellt.

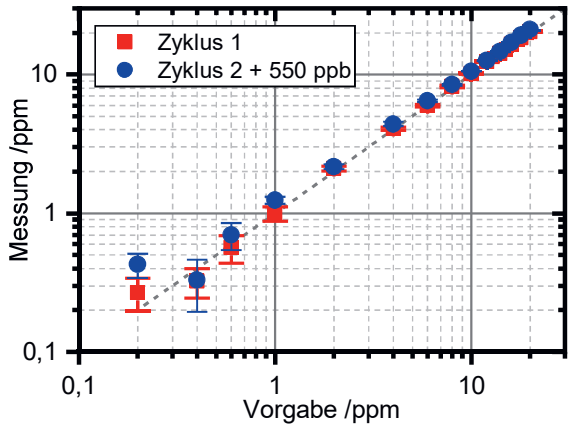


Abb. 8: Auswertung der Konzentrationsstufen des ersten und zweiten Zyklus der Messung aus Abb. 7.

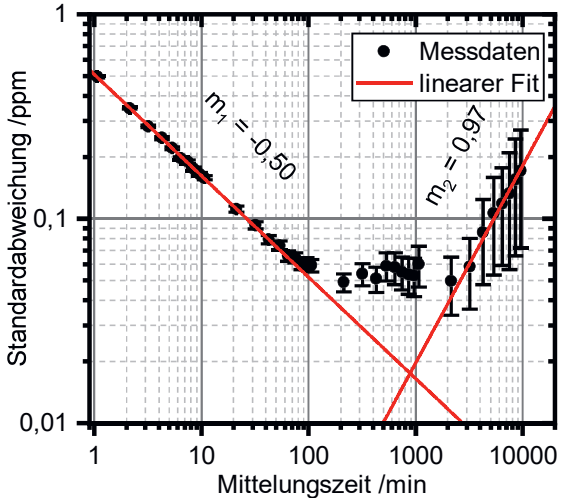


Abb. 9: Darstellung der Allan-Standardabweichung für eine Messzeit von 14,5 Tagen in Stickstoffatmosphäre im Mittelungszeitbereich zwischen 1 min und 10000 min. Die minimale Standardabweichung beträgt bei einer Mittelungszeit von 213,5 min 48,9 ppb Ethen.

Der Effekt der Signalmittelung ist am Anstieg von -0,5 sehr gut erkennbar. Für eine Mittelungszeit von 213 min wurde eine minimale Standardabweichung von 48,9 ppb Ethen erreicht. Ab einer Mittelungszeit von einem Tag zeigt sich der Effekt der Signaldrift, die zu einem Anstieg von +1 führt. Dies stimmt mit den Beobachtungen des transienten Signals überein. Aus diesen Ergebnissen lässt sich eine Nachweisgrenze von ca. 150 ppb bei einer Mittelungszeit von 213 min ableiten.

Zusammenfassung und Diskussion

In dieser Arbeit wurde ein Ansatz vorgestellt, die die Realisierung kostengünstiger Filterrotationsspektrometer für den Nachweis von Ethen mit Komponenten, die auch in NDIR-Sensoren eingesetzt werden, ermöglicht.
Die aufgenommenen Spektren wurden mit einem modellbasierten Ansatz ausgewertet. Simulationen zeigten, dass sich die wichtigsten Querempfindlichkeitsgase CO_2 und NH_3 qualitativ und quantitativ erfassen lassen. Dabei war der Einfluss dieser Gase auf die bestimmte Ethenkonzentration vernachlässigbar gering. Mithilfe eines ersten Prototyps, wurde für eine Mittelungszeit von 213 min eine Nachweisgrenze von 150 ppb erreicht. Im Vergleich zum Stand der Wissenschaft wurde die Nachweisgrenze somit um den Faktor 20 verbessert. Über den gesamten Messzeitraum wurde eine Systemdrift von ca. 50 ppb/d ermittelt, was im Vergleich zum Stand der

Wissenschaft eine theoretische Laufzeit von ca. 600 Tagen ermöglicht bevor die Drift die Nachweisgrenze dieser Sensoren erreichen würde. Durch einen optimierten optischen Aufbau in Verbindung mit einer integrierten Elektronik lassen sich die Stabilität des Systems sowie die Nachweisgrenze bei gleichzeitig geringeren Kosten weiter verbessern.

Danksagung

Diese Arbeit wurde durch die Georg H. Endress Stiftung im Rahmen des InnoSens-Projekts zur Erforschung einer nachhaltigen Lebensmittelproduktion gefördert.

Literaturnachweis

[1] FAO, Global food losses and food waste – Extent, causes and prevention, Rome, 2011
[2] M.L.A.T.M. Hertog et al., “Shelf life modelling for first-expired-first-out warehouse management”, Phil. Trans. R. Soc. A 372, 20130306 (2014); doi: 10.1098/rsta.2013.0306
[3] S. Janssen et al., “Ethylene detection in fruit supply chains”, Phil. Trans. R. Soc. A 372, 20130311 (2014); doi: 10.1098/rsta.2013.0311
[4] A. Giberti et al., “Monitoring of ethylene for agro-alimentary applications and compensation of humidity effects”, Sensors and Actuators B, 103, 272-276 (2004); doi: 10.1016/j.snb.2004.04.058
[5] A. Sklorz et al., “Application of a miniaturised packaged gas chromatography column and a SNO2 gas detector for analysis of low molecular weight hydrocarbons with focus on ethylene detection, Sensors and Actuators B, 180, 43-49 (2013); doi: 10.1016/j.snb.2011.12.110
[6] R. Shekarriz und W.L. Allen, “Nanoporous gold electrocatalysis for ethylene monitoring and control”, European Journal of Horticultural Science, 73, 171-176 (2008)
[7] E. H. Wahl et al., “Ultra-sensitive ethylene post-harvest monitor based on cavity ring-down spectroscopy”, Opt. Express, vol. 14 (4), 1637-1684 (2005); doi: 10.1364/OE.14.001673
[8] J. Fonollosa et al., “Ethylene optical spectrometer for apple ripening monitoring in controlled atmosphere store-houses”, Sens. Actuators B, vol. 136, pp. 546-554, 2009; doi: 10.1016/j.snb.2008.12.015
[9] Hildenbrand et al., “A compact optical multichannel system for ethylene monitoring”, Microsyst. Technol., vol. 14, 637-644 (2008); doi: 10.1007/s00542-007-0475-1
[10] Fruit Control Equipments srl, “Storage Values for several horticultural products”, www.fruitcontrol.it (2019)
[11] A. Eberhardt et al., “Advanced White cell design for investigation of ethylene using a nondispersive infrared photometer”, Proceedings IRS² (2017), 779-782; doi: 10.5162/irs2017/i2.4
[12] A. Eberhardt, “Filterrotationsspektrometer für die Detektion von Ethen”, Shaker Verlag (2018); ISBN: 978-3-8440-6394-3
[13] J. Kneer et al., „Apparatus to characterize gas sensor response under real-world conditions in the lab”, Review of Scientific Instruments 5, vol. 85, 055006 (2014); doi: 10.1063/1.4878717
[14] P. Werle et al., „The limits of signal averaging in atmospheric trace gas monitoring by tunable diode laser absorption spectroscopy“, Applied Physics B, vol. 57 (1993), 131-139; doi: 10.1007/BF00425997

DOI 10.5162/sensoren2019/4.2.4

Lichtabsorption durch freie Ladungsträger in Metalloxiden: Anwendung der Brechungsindex-Modulation in der Gassensorik

Xia Wu[1], Xuyang Zhang[1], Sabrina Amrehn[1], Thorsten Wagner[1]
[1]Department Chemie, Universität Paderborn, Warburger Straße 100, 33098 Paderborn

Zusammenfassung

Einige halbleitende Metalloxide, wie z.B. Wolframtrioxid (WO_3), zeigen gasinduzierte Änderungen ihrer elektronischen Eigenschaften und finden deshalb Anwendung als Wirkschichten in resistiven Halbleitergassensoren. Den mit den elektronischen Änderungen korrelierten, gasinduzierten optischen Eigenschaftsänderungen wurde bisher jedoch wenig Aufmerksamkeit gewidmet. Die verfügbaren Studien zeigen oft nur schwache optische Gasreaktionen in den beobachteten Wellenlängenbereichen, obwohl starke elektrische Signale zu erwarteten wären. Die vorliegende Arbeit soll mit Hilfe von Wolframoxid als Modellmaterial helfen, diese offensichtliche Diskrepanz aufzuklären. Mit einer Reihe von photonischen Kristallen wird die wasserstoffinduzierte Dispersion des Brechungsindexes des WO_3 bzw. einer Wolframbronze durch Analyse der jeweiligen optischen Reaktion rekonstruiert. Es zeigt sich, dass die gasinduzierte Absorption von freien Ladungsträgern der dominierende Mechanismus für die Modulation des Brechungsindexes ist. Diese Modulation ist im nahen Infrarot stark, im sichtbaren Bereich jedoch schwach, was die bisher beobachteten, relativ geringen optischen Reaktionen erklärt. Durch die Verwendung von photonischen Kristallen mit einer optischen Bandlücke im nahen Infrarotbereich erhöht sich die optische Antwort des hier vorgestellten WO_3/H_2-Systems um eine Größenordnung. Des Weiteren wird davon ausgegangen, dass der vorgestellte Mechanismus für optische Gassensoren, die auf anderen Metalloxiden basieren, anwendbar ist. Dies ermöglicht die optische Detektion weiterer Gase mit Hilfe von Metalloxidschichten.

Keywords: Optisches Messverfahren, Gassensor, Metalloxid, chemischer Sensor, Modellbildung

Einleitung

Einige halbleitende Metalloxide zeigen gasinduzierte Änderungen der elektrischen Eigenschaften (Leitfähigkeit, Widerstand) und finden daher Anwendung als aktive Materialien in resistiven Gassensoren.
Theoretisch sollten diese Metalloxide auch optische Gasreaktionen zeigen, da die mikroskopischen elektrischen Eigenschaften der Halbleiter eng mit den makroskopischen optischen Eigenschaften verbunden sind.[1] Der Vorteil von optischen Gaswandlern besteht in der Möglichkeit der Fernauslesung. Dies bedeutet, dass kein physischer Kontakt zum Sensormaterial erforderlich ist. Dadurch kann der Anwendungsbereich von chemischen Gassensoren auf Metalloxidbasis in rauen Umgebungen, wie z.B. unter hohen Temperaturen oder in Umgebungen mit explosiven Gasen, erweitert werden.
Bisherige Studien zeigen vergleichsweise schwache optische Reaktionen von Metalloxiden, wie z.B. Co_3O_4[2], TiO_2[3], SnO_2[3] oder WO_3[4-7], im Vergleich zu ihren elektrischen Antworten auf verschiedene Gasangebote.

Die optische Reaktion basiert in allen Fällen auf dem sogenannten gasochromen Effekt, wobei das bekannteste Beispiel die Änderung der optischen Konstanten von WO_3 unter Wasserstoffatmosphäre ist.[4-11] Für den chemischen Prozess existieren verschiedene Modellvorstellungen, die den Farbeffekten zugrunde liegen. Das weithin akzeptierte Modell im Fall des WO_3 ist das sogenannte Doppelinjektionsmodell. [10,12]

$$\mathrm{WO_3} + x\mathrm{e^-} + x\mathrm{H^+} \rightleftharpoons \mathrm{H_xWO_3}$$

Wasserstoff wird dabei an Edelmetallkatalysatoren (z. B. Platin und Palladium) dissoziiert und die Protonen mit den begleitenden Elektronen in das WO_3 injiziert; bei diesem Prozess bildet sich sogenannte Wolframbronze (H_xWO_3). Der daraus resultierende, weithin akzeptierte physikalische Mechanismus für den beobachteten Umfärbungseffekt im Fall von kristallinem WO_3 ist die Absorption durch freie Ladungsträger; im n-Halbleiter Elektronen.

DOI 10.5162/sensoren2019/4.2.4

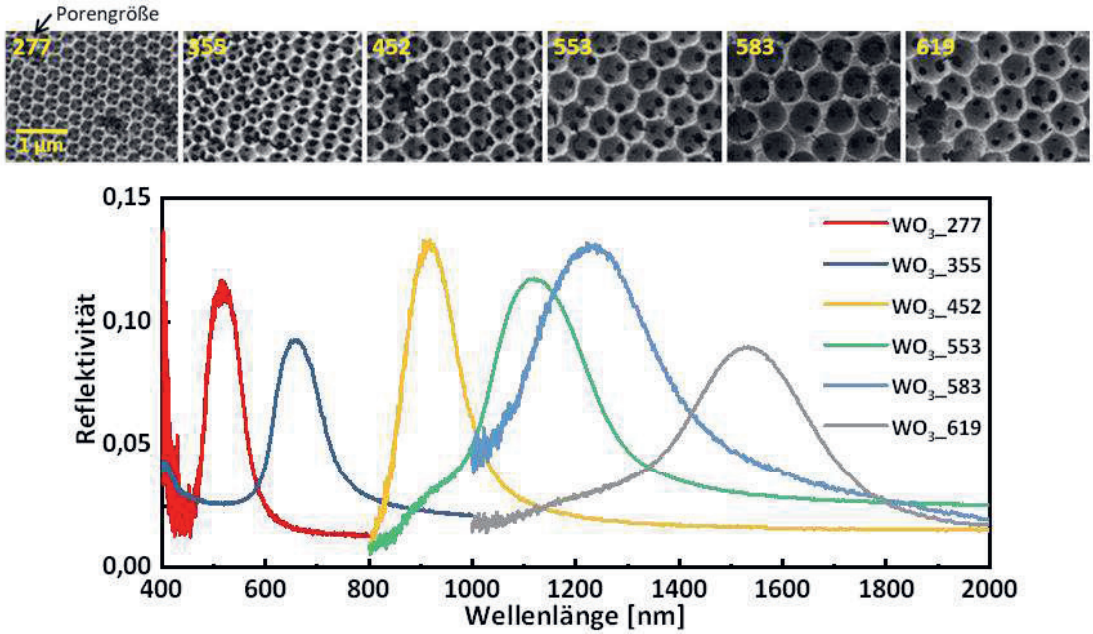


Abb. 1: Reflexionsspektren der verschiedenen WO_3-Inversopale (unten) und elektronenmikroskopische Aufnahmen der Proben (oben); die jeweiligen Porengrößen sind in gelb angegeben.

Der freie Elektronenabsorptionseffekt ist ein Intraband-Übergang im Leitungsband, der die einfallende Photonenenergie auf die freien Elektronen überträgt, die aus dem Gas abgegeben oder durch ein elektrisches Feld injiziert werden.[3,4,8,12] Die Absorption freier Ladungsträger ist das wichtige Bindeglied zum Verständnis der Beziehung zwischen der optischen und elektrischen Gasreaktion von Metalloxiden.

Dies ist wahrscheinlich darauf zurückzuführen, dass die freie Ladungsträgerabsorption der meisten Metalloxide im Infrarotbereich (IR) auftritt und somit für unsere Augen nicht sichtbar ist.[1,3] Im Gegensatz dazu ist der gasochrome Effekt von WO_3 nicht auf den IR-Bereich beschränkt, sondern tritt auch im sichtbaren Bereich auf. Aus diesem Grund betrachten wir hier die Kombination WO_3/H_2 als Modellsystem.

Wir zeigen im Folgenden, dass das tiefere mechanistische Verständnis erlaubt, die optische Antwort des WO_3/H_2-Sensorsystems um eine Größenordnung zu erhöhen. Durch geeignete Wahl der optischen Bandlücke der Inversopale kann die starke Variation des Brechungsindexes von H_xWO_3 bei bestimmten Wellenlängen genutzt werden. Wichtiger ist jedoch, dass das Verständnis der optischen Auswirkung dieses Mechanismus die Verwendung weiterer Metalloxide als optische Sensormaterialien ermöglicht.

Experimentelles

Für die vorgestellte Arbeit wurden Inversopale aus WO_3 mit verschiedenen Bandlücken mit Hilfe einer Abformmethode hergestellt.[13,14]

Mit Hilfe eines löslichen Metalloxid-Präkursors wurden dazu Inversopale an künstlichen Opalen aus Polymethylmethacrylat (PMMA) abgeformt. Die Größe der in der PMMA-Strukturmatrix verwendeten Nanokugeln wurde im Bereich von 298 nm bis 644 nm variiert. Die so hergestellten Inversopale aus WO_3 wurden anschließend mit einem Platinkatalysator beladen. Elektronenmikroskopische Aufnahmen bestätigen die gleichmäßige Porengeometrie und einen hohen Ordnungsgrad (Abb.1).

Die so erhaltenen Proben wurden mit Hilfe eines Spektrometers (FTIR, Vertex 70 mit Hyperion 1000 Mikroskop, Bruker) optisch charakterisiert (Abb.1). Im sichtbaren Bereich kam ein refraktives 4fach Objektiv (NA=0,1) zum Einsatz. Im IR-Bereich wurde ein reflektives 36fach Objektiv verwendet (NA=0,5). Die Gasmessungen wurden mit Hilfe eines selbstentwickelten, optischen Aufbaus durchgeführt (Abb.2). Dazu wurde Licht über eine Faser und einen Strahlteiler auf die Probe und anschließend auf das jeweilige Spektrometer geleitet. Für die Gasmessungen kamen zwei dispersive Spektrometer zum Einsatz (sichtbarer Bereich: flame, Ocean Optics; IR-Bereich: AvaSpec-USB2, Avantes). Die so erhaltenen Spektren wurden mit Hilfe

des Fano-Resonanzmodels ausgewertet (Details siehe [13]).



Abb. 2: Mobiler Aufbau für Gasmessungen mit Metalloxid-Inversopalen. Zentrales Element ist die Gasmesszelle mit eigebauter Heizkartusche und Fenster für optische Messungen.

Die Proben wurden auf eine Temperatur von 140°C aufgeheizt. Die Testgasmischungen bestanden aus H_2 in trockenem Stickstoff (N_2), mit einem Konzentrationsbereich von 0,15 vol% bis 1 vol% H_2. Nach dem jeweiligen Gasangebot wurden die Proben in synthetischer Luft regeneriert. Anschließend wurde die Messzelle vor dem nächsten Gasangebot mit trockenem Stickstoff gespült.

Ergebnisse und Diskussion

Die Abbildung (Abb. 3) zeigt exemplarisch eine Gasmessung mit dem zuvor beschriebenen Aufbau. Es ist zu erkennen, dass die Beaufschlagung mit Wasserstoff zu einer deutlichen Verschiebung des Reflexions-maximums des WO_3 Inversopals führt. Die Reflexionsbande vor H_2-Exposition liegt bei 646 nm und verschiebt sich in Richtung kürzerer Wellenlängen (Blauverschiebung) unter Wasserstoffeinfluss.
Für die weitere Analyse wurden die zuvor beschriebenen Inversopale mit unterschiedlichen Bandlücken bzw. Lagen der Reflexionsbanden mit dem gleichen Messprogramm wie in Abb. 3 charakterisiert. Dadurch fungieren die Inversopale als Sonde zur Ermittlung der Dispersion der Wolframbronze (H_xWO_3) in Abhängigkeit der Wellenlänge.

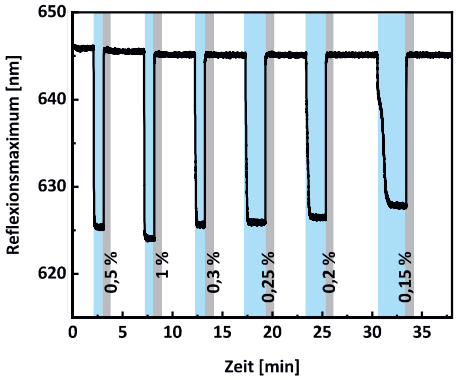


Abb. 3: Verschiebung der Reflexionsbande eines WO_3 Inversopals (Porengröße: 355 nm) während eine Gasmessung mit H_2 bei 140 °C. (blau: H_2 Angebot; grau: Regenration mit synth. Luft; weiß: Spülen mit N_2)

Mit Hilfe der optischen Sensorantwort (R), d.h. mit dem Verhältnis aus Wellenlängen-verschiebung unter Gas ($\Delta\lambda$) zur Lage der Reflexionsbande ohne Gas (λ_0), wurde die Dispersion des Brechungsindexes mit folgendem Zusammenhang rekonstruiert:

$$R = \Delta\lambda/\lambda_0 = \Delta\bar{n}/\bar{n}_0 = (\bar{n}_1 - \bar{n}_0)/\bar{n}_0$$

Der mittlere Brechungsindex des WO_3 Inversopals ($\bar{n}_0$) und des H_xWO_3 Inversopals ($\bar{n}_1$) können mit der allgemeinen Formel für den mittleren Brechungsindex eines photonischen Kristalls berechnet werden [15,16]:

$$\bar{n}^2 = n_s^2 f_s + n_p^2 f_p,$$

mit n_s und $n_p(=1)$ den Brechungsindizes der Festphase und der Gasphase in den Poren, und f_s and $f_p(=1-f_s)$ den Volumenanteilen der beiden Phasen. Für den Brechungsindex des WO_3 (n_{oxide}) wurde konstant 2,03 über den untersuchten Bereich angenommen. [17]
Die so rekonstruierte Dispersion ist in Abb.4 dargestellt. Die Ergebnisse zeigen, dass die stärkste Reaktion im Bereich von 965 nm zu finden ist. Mit Hilfe dieser Dispersionskurve ist es nun möglich, die Theorie der freien Ladungsträgerabsorption als Ursache für die beobachteten Effekte zu prüfen. Ausgehend vom sogenannten Drude-Modell:

$$\epsilon(\omega) = \epsilon_1(\omega) + i\epsilon_2(\omega) = \epsilon_\infty - \frac{{\omega_p}^2}{i\omega\gamma + \omega^2}$$

können folgende Gleichungen abgeleitet werden:

$$\epsilon_1(\omega) = \epsilon_\infty - \frac{{\omega_p}^2}{\omega^2 + \gamma^2}$$
$$\epsilon_2(\omega) = \frac{{\omega_p}^2 \gamma}{\omega^3 + \omega\gamma^2}$$

mit der Frequenz des Lichts ω, der komplexen relative Permittivität $\epsilon(\omega)$, und dem Real- und Imaginärteilen dieser Permittivität $\epsilon_1(\omega)$ und $\epsilon_2(\omega)$. Mit Hilfe dieser Gleichungen kann der Brechungsindex der Wolframbronze $n_{bronze}(\omega)$ wie folgt beschrieben werden:

$$n^2_{bronze}(\omega) = \frac{1}{2}\left(\sqrt{\epsilon_1^2 + \epsilon_2^2} + \epsilon_1\right)$$

Hier sind die freien Parameter für die Kurvenanpassung die Hintergrundpermittivität ϵ_∞bei hohen Frequenzen, die Plasmafrequenz ω_p, und der Dämpfungskoeffizient γ. Zentrales Ergebnis der Kurvenanpassung (Abb. 4) ist, dass der Brechungsindex der Wolframbronze ($n_{bronze}(\omega)$) im Bereich von 500 nm bis 1.5 µm eine starke Dispersion aufweist und die Form dem Drude-Modell gleicht. Da das Drude-Modell die mikroskopische Wechselwirkung zwischen Licht und freien Ladungsträgern in einem Material beschreibt, deutet dies darauf hin, dass die H_2-induzierten freien Ladungsträger zu einer erhöhten Absorption führen und somit der dominante physikalische Effekt bei der beobachteten, optischen Sensorantwort sind.

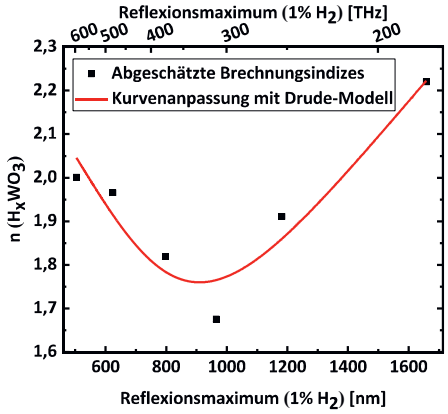


Abb. 4: Abgeschätzte Dispersion des Realteils des Brechungsindexes von H_xWO_3 (schwarz) und angepasstes Drude-Modell (rote Kurve).

Zusammenfassung

Wir haben gezeigt, dass die Absorption durch freie Ladungsträger eine wichtige Verbindung zwischen den mikroskopischen elektronischen und den makroskopischen optischen Eigenschaften bei der Gasreaktion von halbleitenden Metalloxiden darstellt. Mit einem Modellsystem (WO_3/H_2) ist es uns gelungen, die Dispersion des Metalloxids unter Gasangebot mit Hilfe einer Reihe von Inversopalen mit unterschiedlicher photonische Bandlücke zu rekonstruieren. Die so erhaltene Dispersionskurve ließ sich schließlich mit Hilfe des Drude-Modells anpassen. Dies ist ein starker Hinweis auf die Absorption durch freie Ladungsträger als dominierender Mechanismus für die beobachtete Gasreaktion.
Zusätzlich konnte in der Anwendung mit Hilfe der Dispersionskurve die Sensorantwort um eine Größenordnung erhöht werden.
Ausgehend von diesen Ergebnissen gehen wir davon aus, dass die meisten halbleitenden Metalloxide für die optische Gassensorik mit Hilfe der Inversopalstruktur genutzt werden können. Allerdings muss der spektrale Bereich so gewählt, dass dieser nah genug an einem Bereich starker Dispersion der optischen Konstanten des jeweiligen Metalloxids liegt, um eine hinreichend starke optische Sensorantwort zu erhalten. Für die meisten Metalloxide liegt dieser im IR-Bereich und wird bestimmt durch die Variation der freien Ladungsträgerdichte.[3]

Danksagung

Wir danken dem Bundesministerium für Bildung und Forschung (FKZ: 13N12969) für die finanzielle Unterstützung.

Literaturnachweis

[1] Mark Fox, Optical properties of solids, 2. Ed, Oxford University Press, New York, USA 2010
[2] M. Ando, T. Kobayashi, M. Haruta, Catal. Today 1997, **36**, 135
[3] M. –I. Baraton and L. Merhari, Synth. React. Inorg., Met.-Org., Nano-Met. Chem. 2005, **35**, 733
[4] K. Ito and T. Ohgami, Appl. Phys. Lett. 1992, **60**, 938
[5] A. Georg, W. Graf, R. Neumann, V. Wittwer, Solid State Ionics 2000, **127**, 319
[6] D. Nau, A. Seidel, R.B. Orzekowsky, S.-H. Lee, S. Deb, H. Giessen, Opt. Lett. 2010, **35**, 3150
[7] Z. Xie, H. Xu, F. Rong, L. Sun, S. Zhang, Z. Gu, Thin Solid Films 2012, **520**, 4063
[8] P.J. Shaver, Appl. Phys. Lett. 1967, **11**, 255
[9] H. Chen, N. Xu, S. Deng, D.Lu, Z. Li, J. Zhou, J. Chen, Nanotechnology 2007, **18**, 205701
[10] S. H. Lee, H. M. Cheong, P. Liu, D. Smith, C. E. Tracy, A. Mascanrenhas, J. R. Pitts, S. K. Deb, J. Appl. Phys. 2000, **88**, 3076
[11] S. Okazaki, H. Nakagawa, S. Asakura, Y. Tomiuchi, N. Tsuji, H. Murayama, M. Washiya, Sens. Actuators, B. 2003, **93**, 142
[12] C.G. Granqvist, Handbook of Inorganic Electrochromic Materials, Elsevier, Amsterdam, The Netherlands 1995

[13] S. Amrehn, X. Wu, T. Wagner, ACS Sens. 2018, **3**, 191
[14] M. Egen, R. Zentel, Macromol. Chem. Phys. 2004, **205**, 1479
[15] H. Míguez, C. López, F. Meseguer, A. Blanco, L. Vázquez, R. Mayoral, Appl. Phys. Lett. 1997, **71**, 1148
[16] A. Reynolds, F. López-Tejeira, D. Cassagne, F. J. García-Vidal, C. Jouanin, J. Sánchez-Dehesa, Phys. Rev. B. 1999, **60**, 11422
[17] R. B. Goldner, P. Norton, K. Wong, G. Foley, E. L. Goldner, G. Seward, R. Chapman, Appl. Phys. Lett. 1985, **47**, 536

DOI 10.5162/sensoren2019/4.3.1

Erkennung von Landmarken für Roboternavigation mit Luftultraschall und neuronalen Netzen – Feature Engineering und Vergleich

Patrick K. Kroh[1], Ralph Simon[2], Stefan J. Rupitsch[1]
[1]Lehrstuhl für Sensorik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Paul-Gordan-Str. 3/5, 91052 Erlangen, Deutschland
[2]Department of Ecological Science, Vrije Universiteit Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, Niederlande

Zusammenfassung

Ultraschallbasierte Sensoren kommen bei mobilen Robotern und im Automobilbereich häufig zur Abstandsmessung von umliegenden Objekten zum Einsatz. Das akustische Messprinzip unterscheidet sich physikalisch von den Prinzipien anderer geläufiger Sensoren zur Umfelderfassung von Robotern (z. B. Lidar, Radar, Kamera etc.) und folglich können entsprechende Sensorsysteme mit Ultraschallsensoren komplementär ergänzt werden. Es wird dadurch eine Reduktion der Fehleranfälligkeit und somit Erhöhung der Zuverlässigkeit von Roboternavigationssystemen möglich. In unserer Arbeit beschäftigen wir uns daher mit der Identifizierung und Klassifizierung von Ultraschall-reflektierenden Objekten, welche als ortsfeste Referenzpunkte (Landmarken) für die Navigation herangezogen werden können. Hierzu werden künstliche neuronale Netze als Klassifizierer mit unterschiedlichen Merkmalsvektoren (Features) trainiert und untersucht. Es hat sich gezeigt, dass mit den entwickelten Features, auf Basis der Target Strength der Ziele, verschiedene geometrische Formen zu einem hohen Grad und auch bei zusätzlichem Rauschen erfolgreich identifiziert werden können.

Keywords: Landmarkenerkennung, Sonarmessung, Zielerkennung, Roboternavigation, Luftultraschall

Einleitung

In Anwendungsgebieten wie der mobilen Robotik und im Automobilbereich, wird Ultraschallsensorik in Luft häufig eingesetzt, um das Vorhandensein von nahegelegenen Hindernissen zu detektieren und deren Abstand zu bestimmen [1-4]. Im Gegensatz zu anderen Sensorprinzipien zur Umfelderfassung erfolgt diese dabei unabhängig von den optischen und elektromagnetischen Eigenschaften der zu detektierenden Hindernisse. Folglich ist auch eine Erkennung von optisch reflektierenden und transparenten Zielen mit geringem Radarquerschnitt möglich (z. B. Objekte aus transparenten Kunststoffen oder Hindernisse im Dunkeln), welche häufig eine Herausforderung für Systeme wie Kameras, RADAR (Radio Detection And Ranging) und LIDAR (Light Detection And Ranging) darstellen. Im Vergleich zu den genannten Prinzipien weisen ultraschallbasierte Systeme allerdings normalerweise eine geringe Winkelauflösung bei reinem Puls-Echo-Betrieb auf, wodurch eine Identifizierung der Umgebung sowie einzelner reflektierender Objekte anhand deren Geometrie nur bei verhältnismäßig großen Abmessungen möglich ist. Im Rahmen unserer Arbeit entwickeln wir daher ein System zur Identifikation und Ortung von spezifischen schallreflektierenden Objekten (Ziele) direkt anhand ihrer Echosignale. Die Objekte können dabei auch deutlich kleiner sein (charakteristische Abmessungen von ca. 100 mm). Dies kann besonders nützlich sein, wenn ein Roboter in einer Umgebung navigieren muss, die sich ändern kann. So könnten beispielsweise gut zu detektierende Objekte an festen Punkten angebracht werden und als ortsfeste Referenzpunkte dienen (Landmarken) [5,6]. Dies kann z. B. der Fall sein, wenn Roboter in Umgebungen agieren müssen, in welchen sich auch Menschen aufhalten, wie in den Bereichen der Assistenzrobotik und der Mensch-Roboter-Kooperation in der industriellen Fertigung [7,8]. Hierbei ist vor allem noch durch Integration weiterer komplementärer Sensorprinzipien mit Sensordatenfusion eine deutliche Erhöhung der Zuverlässigkeit der Navigation zu erwarten.

In dem vorliegenden Manuskript wird eingegangen auf die grundlegende Methodik zur Identifikation und Klassifizierung von Landmar-

ken. Der Fokus liegt dabei auf den Klassifizierern sowie der Entwicklung von Merkmalsvektoren, mit denen die Klassifizierer effizient arbeiten können. Es werden unter anderem verschiedene akustische Anregungssignale und der Einfluss von Rauschen untersucht. Als Klassifizierer kommen künstliche neuronale Netze (KNNs) zum Einsatz, da sich diese aufgrund ihrer Eigenschaften in bereits ähnlich gelagerten Aufgabenstellungen als vorteilhaft erwiesen haben. Dies wurde beispielsweise in Arbeiten von Eliakim et al., Ayrulu et al. und Dmitrieva et al. gezeigt, welche allerdings andere Features entwickelt sowie verwendet haben, unterschiedliche Signale eingesetzt haben und auch nicht den Einfluss von Rauschen untersucht haben [9-12].

Messaufbau und -ablauf

Echosignale wurden in einem schalltoten Raum für verschiedene Ziele und Positionen aufgenommen sowie anschließend an einem handelsüblichen Desktop-Computer weiterverarbeitet und klassifiziert.

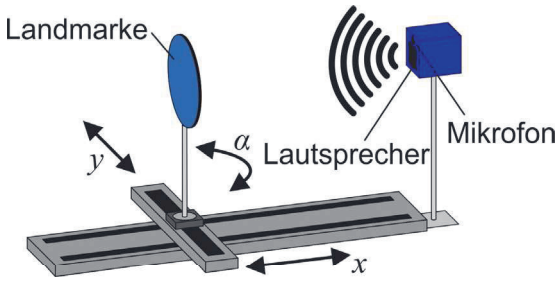


Abb. 1: Versuchsaufbau mit Ultraschalllautsprecher, Landmarke, Mikrofon, translatorischen Verfahreinheiten (x, y) und Rotationseinheit (α).

Hierzu wurden die Ziele mit 1 m Abstand vom Boden auf einer Stange montiert, welche motorisiert in zwei translatorischen Richtungen (*x, y*) verfahren und auch rotiert werden kann. Die Ultraschallsignale wurden von einem breitbandigen kapazitiven Ultraschalllautsprecher (Senscomp 7000, 20 kHz bis > 150 kHz, > 85 dB re 20µPa) emittiert, von den Zielen reflektiert und die Echosignale mit einem ¼"-Messmikrofon (Brüel & Kjær 4939-A-011) mit Vorverstärker (G.R.A.S. 12AK) aufgenommen. Als Signale kamen breitbandige (150 bis 20 kHz) und schmalbandige (52 bis 48 kHz) rechteckförmige Downchirp-Signale mit einer Dauer von 1 ms zum Einsatz [13]. Der Ultraschalllautsprecher und das Mikrofon sind möglichst nahe beieinander montiert (20 mm Mittenabstand) und so angebracht, dass sich die Mitte des Lautsprechers auf Höhe der Mitte der Ziele befindet. Eine Messdatenerfassungskarte (National Instruments NI-USB-6356, 1,25 MS/s, 16 bit) wurde eingesetzt für das Generieren der elektrischen Anregungssignale und das Analog-Digital-Umsetzen der aufgenommenen Echosignale. Der verwendete Lautsprecher erfordert verhältnismäßig hohe Anregungsspannungen im Bereich von 0 bis 400 V. Es wurde daher ein spezieller Verstärker für diesen gebaut und eingesetzt. Die Ziele wurden mit einem 3D-Drucker (Ultimaker 2) aus ABS-Kunststoff hergestellt.

Für jedes Zielobjekt wurden Echosignale für alle Kombinationen aus *x*-Positionen 0,5 m bis 1,8 m (in 0,1 m-Schritten), y-Positionen -0,15 m bis 0,15 m (in 0,05 m-Schritten) und Winkeln *α* von -60° bis 60° (in 15°-Schritten) aufgenommen. Die aufgenommenen Signale wurden weiterverarbeitet, Merkmale aus diesen extrahiert und anschließend für das Training der KNNs herangezogen. Es wurden dabei immer Echosignale für drei benachbarte y-Positionen berücksichtigt. Dies entspricht z. B. den Aufnahmen, die bei Vorbeifahrt eines Roboters an einem Hindernis gemacht würden.

Target Strength und Ziele

Ein wichtiges Merkmal von Sonarzielen ist deren sogenannte Target Strength, welche ein Maß für die reflektierte akustische Intensität in Bezug auf die einfallende akustische Intensität darstellt. Der Begriff der Target Strength kann sich auf die gesamte Intensität beziehen, oder aber auch in Abhängigkeit der Signalfrequenz ausgedrückt werden. Es werden daher nachfolgend auch die Begriffe der gesamten Target Strength und der spektralen Target Strength verwendet.

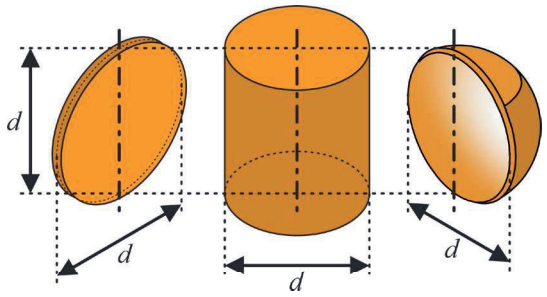


Abb. 2: Zielgeometrien mit charakteristischer Abmessung d; Scheibe, Zylinder und hohle Halbkugel.

Untersucht man die Target Strength verschiedener Objekte, so fällt auf, dass diese auch vom Beschallungswinkel abhängen kann. Es zeigen sich dabei deutliche Unterschiede für flache, konkave sowie konvexe Objekte, falls die Objekte ausreichend groß sind im Vergleich zur Wellenlänge der eingesetzten akustischen Signale. Für die Untersuchungen der vorliegenden Arbeit wurden möglichst einfache, grundlegende Geometrien ausgewählt,

welche die zuvor genannten Eigenschaften aufweisen: kreisförmige Scheiben (flach), Zylinder (konvex) und hohle Halbkugeln (konkav). Für die Trainingsdatensätze der KNNs wurden Ziele in jeweils zwei Größen (Durchmesser *d* jeweils 60 mm und 100 mm) eingesetzt, um eine Generalisierung bezüglich der Form zu ermöglichen.
Bei flachen Objekten zeigt sich ein deutliches Maximum bei senkrechter Einstrahlung und ein starker Abfall mit zunehmendem Winkel. Folglich wird bereits bei leichter Verkippung ein Großteil der Energie nicht mehr zum Ausgangsort zurück reflektiert. Dies ist deutlich zu erkennen im sogenannten akustischen Fingerabdruck, welcher die spektrale Target Strength in Abhängigkeit des Verdrehwinkels eines Objekts angibt (vgl. Abbildung 3).

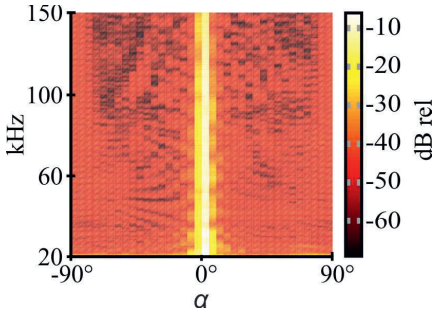


Abb. 3: Akustischer Fingerabdruck einer Scheibe (spektrale Target Strength für verschiedene Drehwinkel α).

Bei konvexen Objekten ist über einen relativ großen Winkelbereich ein Echo detektierbar, allerdings wird auch immer ein Teil der einfallenden Welle in andere Richtungen reflektiert. Es folgt daher, dass die Target Strength gleichmäßiger verteilt ist über den Verdrehwinkel, aber nicht so hoch ist wie bei senkrechtem Einfall bei einem flachen Objekt. Bei einem Zylinder ist die Target Strength optimalerweise unabhängig vom Verdrehwinkel (vgl. Abbildung 4).

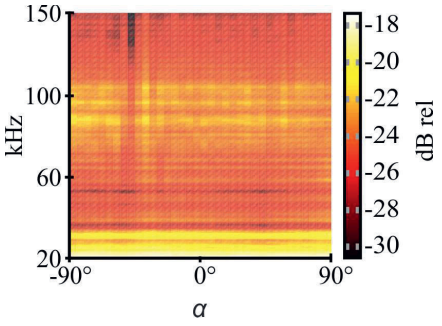


Abb. 4: Akustischer Fingerabdruck eines Zylinders.

Bei konkaven Objekten kann es zur Fokussierung kommen, wodurch eine relativ hohe Target Strength für bestimmte Winkel auftreten kann. Außerdem sind Mehrfachreflexionen innerhalb eines Objektes möglich, die sich überlagern und durch Interferenz zu einer deutlichen Frequenzabhängigkeit der Target Strength führen [14]. Die angesprochenen Zusammenhänge sind deutlich in Abbildung 5 zu sehen.

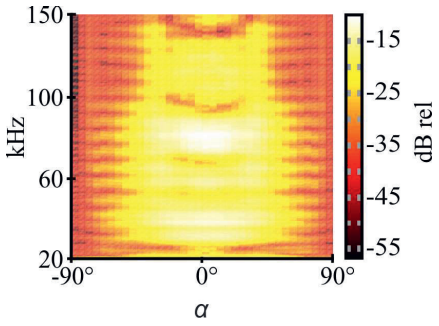


Abb. 5: Akustischer Fingerabdruck einer Halbkugel.

Die akustischen Fingerabdrücke in den Abbildungen wurden mit einem separaten Messverfahren ermittelt, wie in [13] dargestellt.

Künstliche Neuronale Netze

Künstliche neuronale Netze sind Nervensystemen bei realen Lebewesen nachempfunden und werden häufig als Klassifizierer im Bereich der Mustererkennung und der künstlichen Intelligenz eingesetzt. Populäre Beispiele sind dabei die Handschriften- und Spracherkennung. Hierbei zeigen sich besonders die vorteilhaften Eigenschaften künstlicher neuronaler Netze, da diese von gegebenen begrenzten Datensätzen (i) lernen und (ii) generalisieren können sowie (iii) mit verrauschten und verzerrten Bildern und Signalen arbeiten können [15-17]. Aus (i) folgt außerdem, dass eine explizite Modellierung des Problems selbst nicht zwingend notwendig ist. Dies ist besonders vorteilhaft für Aufgabenstellungen, bei welchen bestimmte Merkmale nur aufwändig zu analysieren und modellieren sind, wie bei den zuvor genannten Beispielen.
Die Daten, die am Eingang eines KNNs (Input Layer) anliegen, werden allgemein als Merkmalsvektor oder Featurevektor bezeichnet. Es kann sich dabei um Rohdaten (Bilder, Signalverläufe etc. bei Deep Learning) oder auch um vorverarbeitete Daten handeln, die dann auch als Merkmale oder Features bezeichnet werden. Eine solche Vorverarbeitung dient häufig dazu, eine alternative und ggf. auch kompaktere Darstellung der Daten zu finden, bei welcher die bedeutenden Informationsanteile leichter für die KNNs zu erkennen sind, sodass ein effizienteres Lernen und eine höhere Performanz erzielt werden können. Die Entwicklung

entsprechender Darstellungen unter Berücksichtigung von Wissen über die Problemdomäne wird auch als Feature Engineering bezeichnet und wurde im Rahmen der vorgestellten Arbeit für die Problemstellung der Klassifizierung von reflektierenden Objekten für Ultraschall durchgeführt.
Vorteilhaft beim Einsatz von Engineered Features ist, dass häufig deutlich kleinere Trainingsdatensätze erforderlich sind im Vergleich zu Deep Learning, was besonders bei speziellen technischen Anwendungen wünschenswert ist, da hier häufig nur begrenzte Datensätze vorliegen. Der geringere Bedarf an Trainingsdaten ist darin begründet, dass bei Deep Learning zunächst bedeutungsvolle Features von den Netzen, ausgehend von Rohdaten, selbst gelernt werden müssen (Feature Learning). Diese können bei Engineered Features durch eingebrachtes Domänenwissen bereits vorgegeben sein. Allerdings erfolgt hier das Lernen nur auf den vorab definierten Features, was dazu führen kann, dass bedeutungstragende Informationen aus Rohdaten unberücksichtigt bleiben, da diese explizit erfasst/modelliert werden müssen.
Als Netze kommen bei uns mehrschichtige Perzeptron-Netze zum Einsatz (Abbildung 6), da diese relativ stark abhängig sind von der Qualität der eingesetzten Features und somit anhand der erzielbaren Klassifizierungsperformanz auf die Qualität der entwickelten Features zurückgeschlossen werden kann. Die eingesetzten Netze besitzen vier verdeckte Schichten (Hidden Layers) mit jeweils zehn, fünf, fünf und drei Knoten. Im Vergleich zu einschichtigen Netzwerken ist durch die Tiefe eine Ableitung von abstrakten, komplexen Merkmalen durch die Netzwerke selbst leichter möglich. Außerdem ist das Risiko einer Überanpassung an die Trainingsdatensätze relativ gering, da die Anzahl der Knoten auch niedriger ist als bei anderen Netzwerken für ähnliche Aufgaben [9-12]. Unter Überanpassung ist zu verstehen, dass ein Netz vielmehr die Trainingsdaten auswendig lernt anstatt davon ausgehend zu generalisieren. Dies hat dann zur Folge, dass unbekannte Daten sehr schlecht klassifiziert werden. Dazu kann es kommen, falls die Anzahl der Knoten in einem Netzwerk zu groß ist im Verhältnis zur Menge der vorhandenen Trainingsdaten. Um Überanpassung zu entdecken, wird folglich die Performanz eines Netzes mit Testdatensätzen ermittelt, die nicht Teil der Trainingsdaten sind.
Die KNNs wurden hier mit einem sogenannten Scaled Conjugate Gradient Backpropagation-Algorithmus trainiert, bei welchem noch eine Kreuzvalidierung auf einem separaten Datensatz durchgeführt wird. Die vorhandenen Messdaten wurden zu 20 % auf den Trainingsdatensatz, zu 10 % auf den Validierungsdatensatz und zu 70 % auf den Testdatensatz aufgeteilt. Der vergleichsweise geringe Anteil für den Trainingsdatensatz wurde gewählt, um eine Überanpassung aufgrund von Symmetrien im Aufbau und somit redundanten Echosignalen zu vermeiden.
Für alle Zielgeometrien (Halbkugel, Zylinder, Scheibe) gibt es jeweils eine eigene Klasse für die Klassifikation (Output Layer) sowie eine zusätzliche Klasse für den Fall, dass kein Ziel (kZ) vorliegt. Zur Beurteilung der Performanz von KNNs können verschiedene Maße herangezogen werden. Hier wird die Korrektklassifikationsrate (Accuracy) herangezogen, welche gegen 100 % streben sollte. Die Korrektklassifikationsrate bezeichnet das Verhältnis der Anzahl aller korrekten Klassifizierungen zur Gesamtanzahl an Testbeispielen.
Weitere Maße (z. B. Precision, Recall, F1-Score [18-20]) wurden im Rahmen der durchgeführten Tätigkeiten herangezogen, allerdings führten diese zu den gleichen Ergebnissen wie die Korrektklassifikationsrate und werden daher in diesem Manuskript nicht weiter aufgeführt. In der Veröffentlichung [13] der Autoren wird allerdings auf die Maße in einem erweiterten Kontext eingegangen.

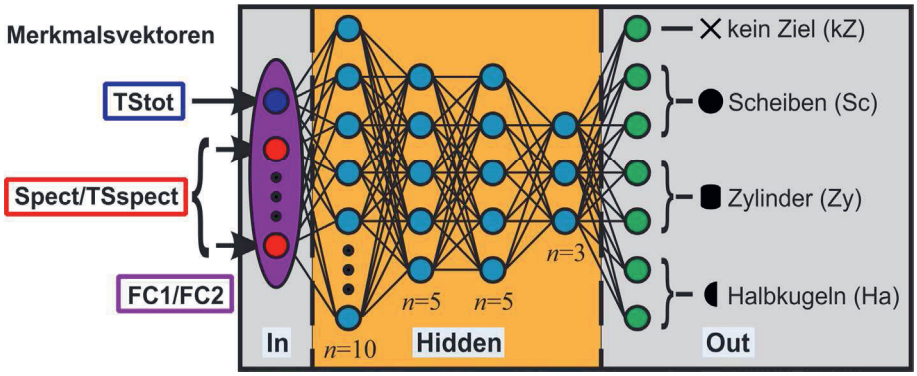


Abb. 6: Künstliches neuronales Netz mit Features (TStot, Spect, TSspect, FC1, FC2), Eingangsschicht, verdeckten Schichten (Hidden) und Ausgangsschicht (Out) mit Zielklassen.

Features

Mit den KNNs wurden verschiedene Featurevektoren untersucht. Diese basieren unter anderem auf Spektrogrammen der Echosignale (*Spect*), Schätzungen der gesamten (*TStot*) sowie der spektralen (*TSspect*) Target Strength und Kombinationen aus den zuvor genannten Merkmalen (*FC1* = TStot und Spect, *FC2* = TStot und TSspect).

Bei der Berechnung der einzelnen Merkmale wurde wie folgt vorgegangen (vgl. auch Abbildung 7). Zunächst erfolgt die Berechnung der Kreuzkorrelationsfunktion des Echosignals mit einer vorhandenen Aufnahme des akustischen Sendesignals, die bei 1 m Mikrofonabstand aufgenommen wurde. Dadurch werden unkorrelierte Signalanteile (z. B. Rauschen und Signale anderer Quellen) herausgefiltert und zudem sind einzelne Reflexionen in der Kreuzkorrelationsfunktion deutlicher erkennbar als im Rohsignal. Im Idealfall besteht nämlich das gefilterte Signal aus einer Überlagerung von um die jeweilige Ausbreitungszeit verschobenen und skalierten Autokorrelationsfunktionen des Sendesignals, die den einzelnen Reflexionen entsprechen und in ihrer Breite abhängig sind von der Bandbreite des Sendesignals. Diese Technik ist im Bereich von Sonar- und Radarsystemen bekannt als sogenannte Pulskompression [21-24].

Aus den pulskomprimierten Signalen wird nun ein Bereich (Region of Interest, ROI) ausgeschnitten, der zentriert ist um den Spitzenwert, der dem Echo des zu klassifizierenden Ziels entspricht. Für die Klasse kZ werden zufällig Spitzenwerte ausgewählt, die sich außerhalb der ROIs von den bekannten Zielen befinden. Für Spect werden ROIs entsprechend in den Spektrogrammen der Rohsignale ausgewählt. Die Dauer der ROIs beträgt 2 ms. Die anschließenden Berechnungen erfolgen dann nur für die ROIs, sodass die KNNs nicht lernen müssen die Signalanteile auszublenden, die nicht im Zusammenhang mit den zu klassifizierenden Zielen stehen.

TStot und TSspect werden aus den pulskomprimierten Echosignalen ermittelt, wobei für TStot nur der Spitzenwert des Pulses relevant ist, der von der Reflexion an einem Ziel herrührt. Die konkreten formelmäßigen Zusammenhänge und deren Herleitung sind gegeben in [13]. Die Herleitung beruht unter anderem auf der Sonargleichung [25] sowie allgemeinen Beziehungen zwischen Signalleistung, Fouriertransformation und Korrelationsfunktionen. Außerdem wird von der Annahme ausgegangen, dass der Signalpegel entlang des Ausbreitungsweges nur aufgrund von geometrischer Aufweitung der Welle abnimmt. Für TSspect wird außerdem anstatt einer einzelnen Fouriertransformation über die ROI in einem pulskomprimierten Signal, eine Kurzzeitfouriertransformation (STFT) mit einer Fensterlänge von 1024 Samples und einem Überlappungsbereich von 50 % durchgeführt. Dies erfolgt in Analogie zu Spracherkennungssystemen, bei welchen sich der Einsatz von STFTs als vorteilhaft gezeigt hat und mittlerweile den Regelfall darstellt [26-30]. Die gleichen Parameter wurden auch für die STFT der Roh-Echosignale für das Feature Spect verwendet. Für jedes Feature werden die Ergebnisse der Berechnungen von drei Positionen in einem 1D-Merkmalsvektor zusammengefasst, um die Ausbreitungsverzögerung (t_t bzw. t_{nt} in Abbildung 7) des Echos ergänzt und dann für die künstlichen neuronalen Netze als Eingangswerte verwendet (wie bereits am Ende des Abschnitts „Messaufbau und -ablauf“ angesprochen).

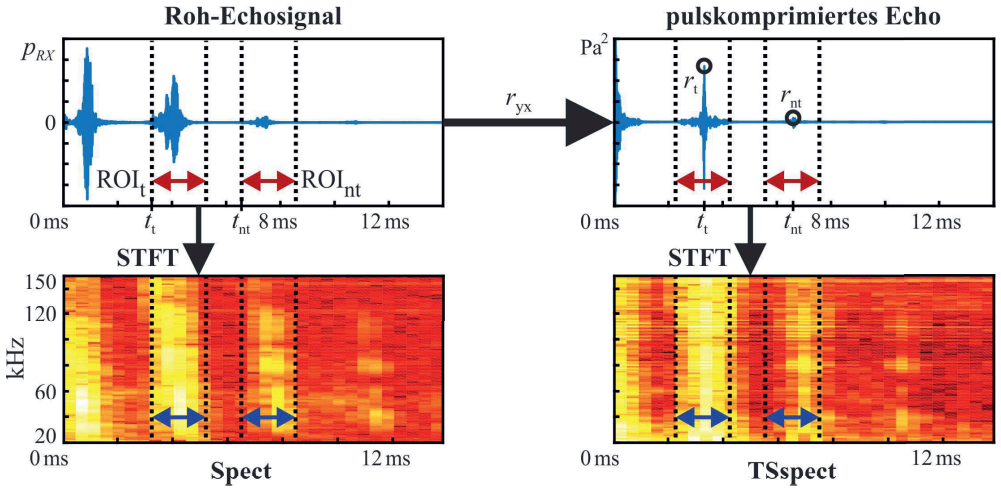


Abb. 7: Vorgehen bei der Berechnung der Features, ausgehend vom Roh-Echosignal mit Kreuzkorrelation (r_{xy}), Kurzzeitfouriertransformation (STFT) und den ROIs für eine Landmarke (ROI_t) sowie kZ (ROI_{nt}). Die Spitzenwerte r_t und r_{nt} werden für die Berechnung von TStot benötigt.

Klassifizierungsergebnisse

Nachfolgend werden Klassifizierungsergebnisse aufgeführt und diskutiert, welche mit KNNs erzielt wurden. Für jeden Punkt auf den Abszissen der Abbildungen 8 bis 12 wurden jeweils 20 KNNs trainiert und deren Performanz statistisch anhand ihrer Korrektklassifikationsraten ausgewertet. Es handelt sich dabei um neue KNNs im Vergleich zur Veröffentlichung [13] der Autoren, wobei die Implementierung noch optimiert worden ist.
Es erfolgt eine statistische Auswertung, da durch die zufällige Initialisierung der KNNs zu Beginn eines Trainingsdurchlaufs verschiedene lokale Optima erreicht werden können. Das KNN mit dem besten Ergebnis würde dann für den praktischen Einsatz herangezogen werden. Daher sind die besten erzielten Ergebnisse auch in den Graphen eingetragen. Die Mittelwerte und Standardabweichungen geben eine Orientierung bezüglich des Trainingsaufwands, der nötig ist, um eine gute Lösung zu erzielen. So wird bei einem großen Abstand zwischen bestem Wert und Mittelwert mit Standardabweichung eine höhere Anzahl an Trainingsdurchläufen nötig sein, um ein KNN mit guter Performanz zu finden.

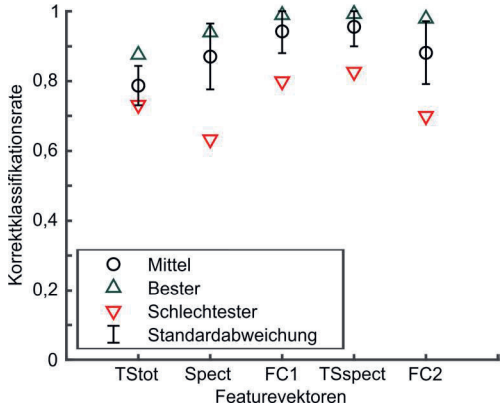


Abb. 8: Korrektklassifikationsrate für breitbandige Downchirp-Anregung, 1,25 MS/s und ohne zusätzliches Rauschen.

Anhand der Abbildungen 8 und 9 ist zu erkennen, dass mit allen Features eine korrekte Klassifizierung der Ziele zu einem hohen Grad (> 80 %) möglich ist. Die schlechtesten Ergebnisse werden mit TStot erzielt, wobei hier der geringste Berechnungsaufwand erforderlich ist, da kein Spektrogramm berechnet werden muss. Für Systeme mit begrenzten Ressourcen kann der Einsatz von TStot also sehr wohl sinnvoll sein. Die besten Ergebnisse sind erkennbar für die Spektrogramm-basierten Features, wobei kaum Unterschiede zwischen den besten Ergebnissen für FC1, TSspect und FC2 zu sehen sind.

Interessant ist die Verbesserung von FC1 gegenüber Spect. Anscheinend können mehr relevante Informationen aus den Rohdaten (Spect) extrahiert werden, wenn TStot direkt mit vorgegeben wird, was einer Kombination aus Feature Engineering und Feature Learning entspricht.

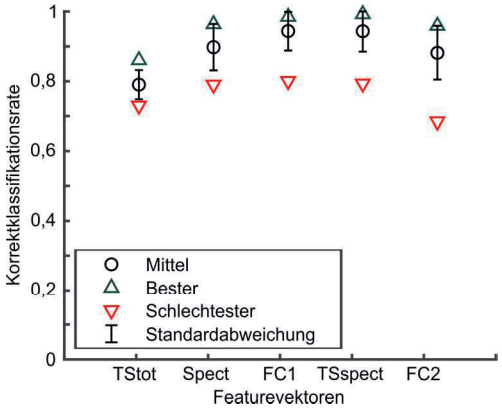


Abb. 9: Korrektklassifikationsrate für schmalbandige Downchirp-Anregung, 1,25 MS/s und ohne zusätzliches Rauschen.

Die Abbildungen 10 bis 12 verdeutlichen den Einfluss von zusätzlichem weißem Rauschen, welches zu den Rohsignalen hinzuaddiert wurde. Eine korrekte Klassifikation von über 90 % der Testdaten ist mit den besten KNNs bei 70 dB re 20 µPa zusätzlichem Rauschen bei einer Abtastrate von 1,25 MS/s noch möglich (Abbildungen 10 und 11). Die Fähigkeit von KNNs mit verrauschten Signalen umzugehen, wird demnach als bestätigt angesehen.

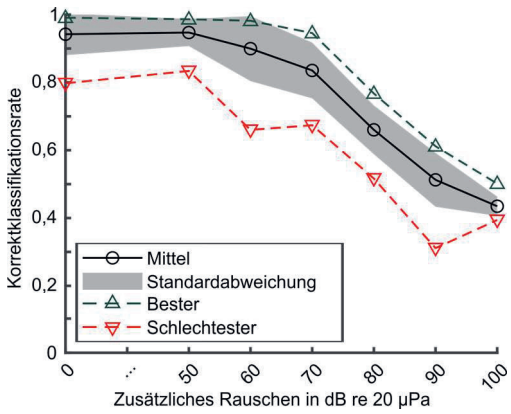


Abb. 10: Korrektklassifikationsrate für FC1 bei breitbandiger Downchirp-Anregung, 1,25 MS/s und mit zusätzlichem Rauschen.

Außerdem ist in Abbildung 12 die erzielbare Korrektklassifikationsrate für schmalbandige Anregungssignale bei einer herabgesetzten Abtastrate von 192 kS/s zu sehen, welche auf Daten basiert, die durch Downsampling der

echten Messdaten bei 1,25 MS/s berechnet wurden. Es ist deutlich zu erkennen, dass auch hier noch über 90 % der Testdaten korrekt klassifiziert werden können bei Rauschschalldrücken von bis zu 60 dB re 20 µPa.

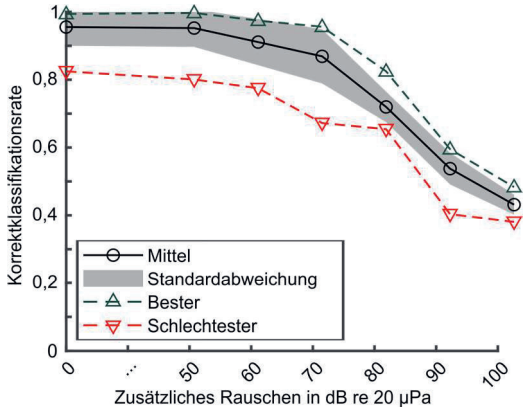


Abb. 11: Korrektklassifikationsrate für TSspect bei breitbandiger Downchirp-Anregung, 1,25 MS/s und mit zusätzlichem Rauschen.

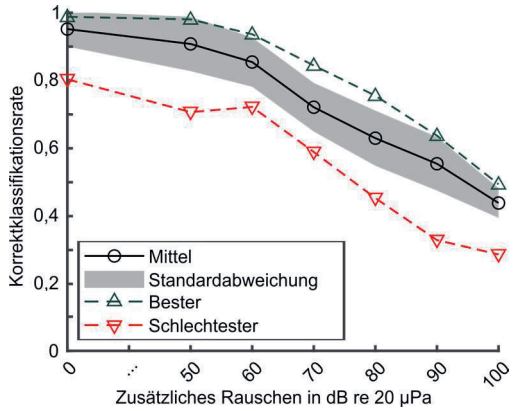


Abb. 12: Korrektklassifikationsrate für FC1 bei schmalbandiger Downchirp-Anregung, 192 kS/s und mit zusätzlichem Rauschen.

Zusammenfassung und Ausblick

Es wurde eine erfolgreiche Klassifizierung von akustischen Landmarken verschiedener Ziele mit den entwickelten Features und KNNs demonstriert. Dabei wurde auch der Einfluss von Anregungssignalen und Rauschen untersucht, wobei sich gezeigt hat, dass sowohl der Einsatz von schmalbandigen wie auch breitbandigen Signalen zu vergleichbaren Ergebnissen führt. Außerdem hat sich die Annahme bestätigt, dass KNNs auch mit verrauschten Signalen die gestellte Aufgabe bewältigen. Die positiven Ergebnisse in Bezug auf Rauschen, Abtastrate und Signalbandbreite motivieren zudem den zukünftigen Einsatz von günstigen MEMS-Mikrofonen mit handelsüblichen Analog-Digital-Umsetzern in Mikrocontrollern und möglicherweise auch schmalbandigen piezoelektrischen Ultraschallwandlern als Lautsprecher [31]. In Zukunft soll außerdem ein eingebettetes Gesamtsystem implementiert werden, das mehrere Landmarken simultan erfassen und orten kann. Ergänzend soll auch der Einfluss von weiteren Reflexionen an unbekannten Objekten untersucht werden, die zusätzlich zu den Landmarken vorhanden sein können.

Literaturnachweis

[1] L. Kleeman, R. Kuc, Sonar Sensing. In Springer Handbook of Robotics; Siciliano, B.; Khatib, O., Eds.; Springer International Publishing: Cham; 491–519 (2016);

[2] J. Steckel, H. Peremans, BatSLAM: Simultaneous localization and mapping using biomimetic sonar. *PloS one*, 8, e54076 (2013); doi: 10.1371/journal.pone.0054076

[3] H.R. Everett, Sensors for mobile robots: Theory and application / H.R. Everett; A. K. Peters: Wellesley, Mass. (1995)

[4] R.J. Przybyla, H.Y. Tang, Guedes, A.; Shelton, S.E.; Horsley, D.A.; Boser, B.E. 3D Ultrasonic Rangefinder on a Chip. *IEEE Journal of Solid-State Circuits*, 50, 320–334 (2015); doi: 10.1109/JSSC.2014.2364975

[5] D. Vanderelst, J. Steckel, A. Boen, H. Peremans, M.W. Holderied, Place recognition using batlike sonar. *eLife*, 5 (2016); doi: 10.7554/eLife.14188

[6] S. Thrun, W. Burgard, D. Fox, Probabilistic robotics; Intelligent robotics and autonomous agents series, MIT Press: Cambridge, Massachusetts and London, England (2006);

[7] P. Rashidi, A. Mihailidis, A Survey on Ambient-Assisted Living Tools for Older Adults. *IEEE Journal of Biomedical and Health Informatics*, 17, 579–590 (2013); doi: 10.1109/JBHI.2012.2234129

[8] T. Dahl, M. Boulos, Robots in Health and Social Care: A Complementary Technology to Home Care and Telehealthcare? *Robotics*, 3, 1–21 (2014); doi: 10.3390/robotics3010001

[9] I. Eliakim, Z. Cohen, G. Kosa, Y. Yovel, A fully autonomous terrestrial bat-like acoustic robot. *PLoS computational biology*, 14, e1006406 (2018); doi: 10.1371/journal.pcbi.1006406

[10] B. Barshan, B. Ayrulu, S.W. Utete, Neural Network-Based Target Differentiation Using Sonar for Robotics Applications. *IEEE Transactions on Robotics and Automation*, 16, 435–442 (2000); doi: 10.1109/70.864239

[11] B. Ayrulu, B. Barshan, Neural networks for improved target differentiation and localization with sonar. *Neural Networks*, 14, 355–373 (2001); doi: 10.1016/S0893-6080(01)00017-X

[12] M. Dmitrieva, M. Valdenegro-Toro, K. Brown, G. Heald, D. Lane, Object classification with convolution neural network based on the time-frequency representation of their echo. *In Proceedings of MLSP2017*; Ueda, N., Ed.; IEEE: Piscataway, NJ; 1–6 (2017); doi: 10.1109/MLSP.2017.8168134

[13] P.K. Kroh, R. Simon, S.J. Rupitsch, Classification of Sonar Targets in Air: A Neural Network Approach. *Sensors* (Basel, Switzerland), 19 (2019); doi: 10.3390/s19051176
[14] R. Simon, M.W. Holderied, O. von Helversen, Size discrimination of hollow hemispheres by echolocation in a nectar feeding bat. *The Journal of experimental biology*, 209, 3599–3609 (2006); doi: 10.1242/jeb.02398
[15] I. Goodfellow, Y. Bengio, A. Courville, Deep learning; Adaptive computation and machine learning, The MIT Press: Cambridge Massachusetts (2016);
[16] C.M. Bishop, Pattern recognition and machine learning; Information science and statistics, Springer: New York (2006);
[17] S.S. Haykin, Neural networks and learning machines, 3rd ed. ed.; Prentice Hall: New York (2009);
[18] G. Dougherty, Pattern Recognition and Classification; Springer New York: New York, NY (2013); doi: 10.1007/978-1-4614-5323-9
[19] C.E. Metz, Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, 8, 283–298 (1978); doi: 10.1016/S0001-2998(78)80014-2
[20] T. Fawcett, An introduction to ROC analysis. *Pattern Recognition Letters*, 27, 861–874 (2006); doi: 10.1016/j.patrec.2005.10.010
[21] M.I. Skolnik, Radar handbook, 3rd ed. ed.; McGraw-Hill: New York (2008);
[22] D.A. Abraham, Signal Processing. *In Applied Underwater Acoustics*; Elsevier; 743–807 (2017); doi: 10.1016/B978-0-12-811240-3.00011-4
[23] J.P. Marage, Y. Mori, Sonar and underwater acoustics; Iste Wiley: London and Hoboken, N.J (2010); doi: 10.1002/9781118600580
[24] D.A. Kiefer, M. Fink, S.J. Rupitsch, Simultaneous Ultrasonic Measurement of Thickness and Speed of Sound in Elastic Plates Using Coded Excitation Signals. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 64, 1744–1757 (2017); doi: 10.1109/TUFFC.2017.2746900
[25] R.P. Hodges, Underwater acoustics: Analysis, design, and performance of sonar; Wiley: Hoboken, NJ (2010); doi: 10.1002/9780470665244
[26] T. Afouras, J.S. Chung, A. Senior, O. Vinyals, A. Zisserman, Deep Audio-visual Speech Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018); doi: 10.1109/TPAMI.2018.2889052
[27] J.H. Connolly, E.A. Edmonds, J.J. Guzy, S.R. Johnson, A. Woodcock, Automatic speech recognition based on spectrogram reading. *International Journal of Man-Machine Studies*, 24, 611–621 (1986); doi: 10.1016/S0020-7373(86)80012-8
[28] S. Ganapathy, Multivariate Autoregressive Spectrogram Modeling for Noisy Speech Recognition. *IEEE Signal Processing Letters*, 24, 1373–1377 (2017); doi: 10.1109/LSP.2017.2724561
[29] J.F. Gemmeke, T. Virtanen, A. Hurmalainen, Exemplar-Based Sparse Representations for Noise Robust Automatic Speech Recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, 19, 2067–2080 (2011); doi: 10.1109/TASL.2011.2112350
[30] V. Zue, L. Lamel, An expert spectrogram reader: A knowledge-based approach to speech recognition. *ICASSP '86. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Institute of Electrical and Electronics Engineers, 1197–1200 (1986); doi: 10.1109/ICASSP.1986.1168798
[31] S.J. Rupitsch, Piezoelectric Sensors and Actuators – Fundamentals and Applications; *Topics in Mining, Metallurgy and Materials Engineering*, Springer Berlin Heidelberg: Berlin, Heidelberg (2019); doi: 10.1007/978-3-662-57534-5

DOI 10.5162/sensoren2019/4.3.2

Vergleich subjektiver und automatisierter Merkmalsextraktion sowie Einsatz maschineller Lernalgorithmen zur mikromagnetischen Materialcharakterisierung

Sargon Youssef[1], Cyril Zimmer[1], Klaus Szielasko[1], Zeeshan Khan Suri[1], Andreas Schütze[2]

[1] Fraunhofer Institut für Zerstörungsfreie Prüfverfahren IZFP
Campus E3 1
66123 Saarbrücken, Deutschland

[2] Lehrstuhl für Messtechnik, Universität des Saarlandes
Gebäude A5 1
66123 Saarbrücken, Deutschland

Zusammenfassung

Im Rahmen dieses Beitrags wurde untersucht, mit welchen Möglichkeiten die datenbasierte Vorhersagequalität verbessert werden kann. Neben der Optimierung des Merkmalsraums wurde die multiple lineare Regression (ML-R) mit weiteren Regressionsalgorithmen (Support Vector Regression, Multilayer Perceptron Regression) verglichen. Diese Regressoren bieten aufgrund ihrer Mächtigkeit und Flexibilität, Nichtlinearitäten in den Daten abzubilden, enormes Potenzial zur Erstellung komplexer Modelle. Beide besitzen jedoch „Stellschrauben“ (sog. Hyperparameter) die signifikant zur Vorhersagequalität beitragen und aufgrund der Vielzahl möglicher Kombinationen kann das Finden einer (quasi-) optimalen Lösung sehr aufwendig sein. Zur Verdeutlichung dieses Aufwandes wurden im Experiment alle Hyperparameter bis auf einen je Regressor vorgegeben. Dieser diente als Stellgröße und sollte aufzeigen, welcher Aufwand schon bei der Variation einzelner Hyperparametern zum Finden eines (quasi-) optimalen Modells erforderlich sind.

Keywords: Zerstörungsfreie Prüfung, Elektromagnetische Prüfverfahren, Materialcharakterisierung, Maschinelles Lernen, Regression

Einleitung

Änderungen mechanischer Kenngrößen wie beispielsweise Härte, Zugfestigkeit oder Streckgrenze, aber auch Last- und Eigenspannungen korrelieren mit Änderungen von Merkmalen der magnetischen Hysteresekurve ferromagnetischer Werkstoffe (z. B. Koerzitivfeldstärke, Remanenz, Permeabilität, etc.) [1]. Diese gegenseitige Abhängigkeit im Werkstoffverhalten erlaubt die Bestimmung mechanischer Kenngrößen auf Basis magnetischer Hystereseeigenschaften, die alternativ teilzerstörend bzw. zerstörend ermittelt werden (z. B. Vickershärte, metallografische Gefügeuntersuchungen, röntgenographische Bestimmung des Eigenspannungszustandes, etc.). Die direkte Bestimmung der magnetischen Hysterese an Bauteilen und Halbzeugen ist häufig in Aufsatztechnik prüfgerecht nicht möglich, teilweise aufgrund der hohen Messzeit, teilweise aufgrund der Notwendigkeit, wohldefiniert geformte Proben zu entnehmen, um die Flussdichte über eine probenumfassende Spule zu messen [1].

3MA-X8 Prüfverfahren

Das am Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP entwickelte 3MA-X8-Verfahren (3MA = Mikromagnetische Multiparameter Mikrostruktur und Spannungs-Analyse) kombiniert in Echtzeit drei Prüfmethoden in Aufsatztechnik mit unterschiedlicher Eindringtiefe und Sensitivität und extrahiert hieraus indirekte Hysteresemerkmale, um auch im Hinblick auf Störeinflüsse zur Zielgröße (z. B. Einsatz verschiedener Werkstoffe, überlagerte Last- und Eigenspannung, etc.) eine zuverlässige Charakterisierung des Werkstoffzustandes zu ermöglichen [2]. Hierbei wird ein Elektromagnet spannungsgesteuert zur Erregung eines magnetischen Wechselfeldes eingesetzt (vgl. Abb. 1).



Abb.1: Unterschiedliche Elektromagnete als 3MA-X8 Prüfkopf

Dieses magnetische Wechselfeld setzt sich zusammen aus einem niederfrequenten Anteil (10 Hz – 200 Hz) und einem höherfrequenten Anteil (500 Hz – 5000 Hz) und wird bei Aufsetzen des Elektromagneten auf einen Werkstoff durch dessen Permeabilität und Leitfähigkeit beeinflusst. Diese Abhängigkeit zeigt sich unter anderem im Realteil Re(Z) und Imaginärteil Im(Z) der Impedanz und im Spulenstrom des Elektromagneten I_C [2].

Subjektive Merkmalsextraktion 3MA-X8

Basierend auf den zur Verfügung stehenden zeitlichen Signale können charakteristische, aber subjektive Merkmale aus den Messkurven extrahiert werden (z. B. Phasenverschiebung des Spulenstroms unten in Abb. 2) [2]. Insgesamt steht ein mehrdimensionaler Merkmalsraum mit insgesamt 21 Merkmalen zur Verfügung.

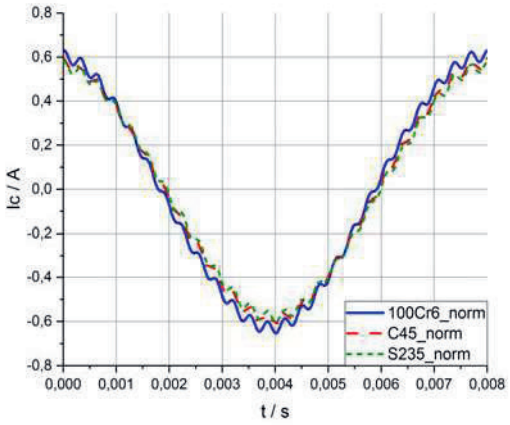


Abb.2: Spulenstrom für drei verschiedene normalisierte Werkstoffe (100Cr6, C45, S235)

Bei dieser Merkmalsextraktion besteht die Möglichkeit, dass geringe, aber signifikante Messeffekte in den Messkurven übersehen werden und der mikromagnetische Merkmalsraum das mechanische Werkstoffverhalten nicht bestmöglich abbildet.

Automatisierte Merkmalsextraktion 3MA-X8

Eine automatisierte Merkmalsextraktion bietet die Möglichkeit, Merkmale auch mit kleinem Messeffekt zur Zielgröße systematisch und unabhängig vom Beobachter zu extrahieren und dadurch die Vorhersagequalität (engl. Performance) von Werkstoffeigenschaften zu verbessern [2].
Bei der automatisierten Merkmalsextraktion werden die zur Verfügung stehenden zeitlichen Signale genauer analysiert. Beispielsweise wird der Spulenstrom mit Hilfe einer Fourier-Transformation in seine Frequenzanteile zerlegt und Betrag und Phase der unterschiedlichen Oberwellen als Merkmale extrahiert [2].

Experiment

In Abhängigkeit der Aufgabenstellung (Klassifizierung bzw. Quantifizierung von Werkstoffen und Werkstoffeigenschaften) werden die extrahierten Merkmale unter Verwendung maschineller Lernalgorithmen zur zerstörungsfreien Materialcharakterisierung eingesetzt. Kleinere Einflüsse auf die magnetische Hysterese (z. B. durch Zugspannungen) lassen sich ebenfalls bestimmen. Bei überlagerten Einflüssen, wie der gleichzeitigen Verwendung verschiedener Werkstoffe verschlechtert sich jedoch die Vorhersagequalität aufgrund von Mehrdeutigkeiten der Daten im Merkmalsraum. Im Vergleich zum ursprünglich 21-dimensionalen Merkmalsraum liefert die automatisierte Merkmalsextraktion mit über 270 Merkmalen einen höherdimensionaleren Merkmalsraum. Zur Beurteilung der Qualität des Merkmalsraum ist jedoch nicht die Anzahl der Merkmale relevant, sondern der Informationsgehalt in und zwischen diesen. Zum Vergleich und zur Bewertung wurde ein Experiment durchgeführt, in der zylinderförmige Proben dreier gehärteter und anschließend angelassener Werkstoffe (100Cr6, C45, S235) in eine Zugmaschine eingespannt und mit elf unterschiedlichen Zugspannungen (von 0 bis 50 % der Streckgrenze R_e des Werkstoffs mit einer Schrittweite von 5 % R_e) belastet wurden (vgl. Abb. 3).

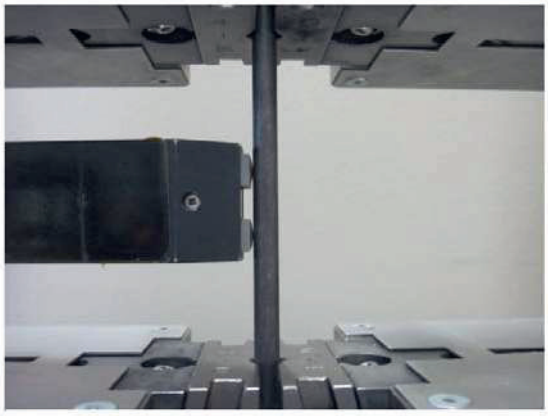


Abb.3: Eingespannte Probe mit aufgesetztem 3MA-X8 Prüfkopf

Im Anschluss wurde auf Basis der beiden Merkmalsräume eine multiple lineare Regressionsanalyse (ML-R) zur Vorhersage des Spannungszustandes durchgeführt.

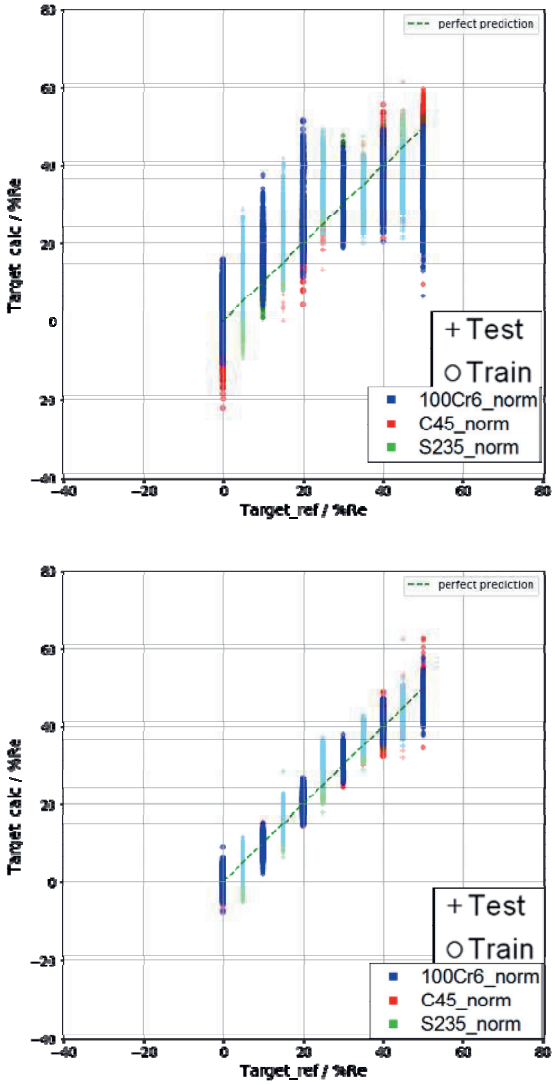


Abb.4: ML-R Abbildung der Trainings- und Testdaten des subjektiv (oben) und automatisiert extrahierten (unten) Merkmalsraum

Im Vergleich beider Merkmalsräume (vgl. Abb. 4) wird eine deutliche Verbesserung der zerstörungsfrei ermittelten Zugspannung sowohl beim Bestimmtheitsmaß R^2_{Train} als auch beim mittleren Fehler $RMSE_{Train}$ (Root Mean Square Error) der Trainingsdaten erreicht. Die Vorhersagequalität der Trainingsdaten verbessert sich von R^2_{Train} = 78,07 % und $RMSE_{Train}$ = 7,99 % R_e auf R^2_{Train} = 98,19 %, $RMSE_{Train}$ = 2,3 % R_e.
Der mittlere Fehler bezogen auf die Testdaten $RMSE_{Test}$, also die Validierung ausschließlich mit Spannungszuständen, die nicht im Training verwendet wurden (5 %, 15 %, 25 %, 35 % und 45 % R_e), zeigt, dass das datenbasierte Modell auf Basis der automatisiert extrahierten Merkmale eine verbesserte Vorhersage des Spannungszustandes von $RMSE_{Test}$ = 7,83 % R_e auf $RMSE_{Test}$ = 4,99 % R_e unabhängig vom Werkstoff erreicht. Der Korrelationskoeffzient der Testdaten verbessert sich von R^2_{Test} = 69,35 % auf R^2_{Test} = 87,54 %.

Einflussfaktoren auf die Vorhersagequalität datenbasierter Modelle

Die Vorhersagequalität datenbasierter Modelle hängt von vielen Einflussfaktoren ab. Die im Rahmen dieses Beitrags untersuchten Einflussfaktoren sind die Wahl des Merkmalsraums (subjektiv / automatisiert), sowie die Wahl des Regressionsmodells.
Im weiteren Verlauf dieses Beitrags wird die gleiche Datenbasis wie bei der ML-R verwendet, die Vorhersagequalität der ML-R dient als Referenz bei der Untersuchung der Einflussfaktoren.

Regressionsmodell

Neben dem Einsatz der multiplen linearen Regressionsanalyse (ML-R) werden im Rahmen dieses Beitrags zwei weitere Regressionsalgorithmen untersucht; die Multi-Layer-Perceptron-Regression (MLP-R) [3] und die Support-Vector-Regression (SV-R) [4]. Im Gegensatz zur ML-R besitzen die SV-R und die MLP-R “Stellschrauben” (sog. Hyperparameter), welche einen erheblichen Einfluss auf die Vorhersagequalität der Zielgröße haben. Es gibt verschiedene Ansätze, geeignete Hyperparameter zu finden. Ein einfacher Weg ist die manuelle Optimierung. Hierbei werden die Hyperparameter durch “Ausprobieren” ermittelt, wobei dies bei komplexeren Modellen mit vielen Freiheitsgraden sehr ineffizient werden kann. Durch Einschränkung von Freiheitsgraden kann dieser Aufwand reduziert werden, wobei

dadurch die Wahrscheinlichkeit, eine (quasi-) optimale Lösungen zu finden, reduziert wird.
Im Rahmen dieses Beitrags soll es nicht primär um das Finden eines optimalen Hyperparametersatzes gehen. Es soll vielmehr aufgezeigt werden, ob und mit welchem Aufwand die Vorhersagequalität der ML-R unter ähnlichen Bedingungen (lineare Separierbarkeit) mit komplexeren Algorithmen erreicht werden kann. Um eine Vergleichbarkeit der Regressionsmodelle zu gewährleisten, werden diese im linearen Modus betrieben, das bedeutet lineare Hyperebene bei SV-R und lineare Aktivierungsfunktionen bei MLP-R.

Support-Vector-Regression

Die Support-Vector-Regression ist ein Regressor aus der Familie der Support-Vector-Machine (SVM) [4]. Ziel der SVM ist es, den Bereich um die Klassengrenzen (engl. Margin) frei von Daten fremder Klassen zu halten. Sind die Klassen linear separierbar, wird die Hyperebene ausgewählt, die den kleinsten Abstand zu den Trainingsdaten maximiert. Sind die Klassen nicht linear separierbar, kann durch nichtlineare Kernelfunktionen der zur Verfügung stehende Merkmalsraum in einem höherdimensionalen Merkmalsraum linearisiert abgebildet und nach der Klassifikation wieder in den ursprünglichen Merkmalsraum projiziert werden (Kernel-Trick).
Die Stellgröße bei der Optimierung der Vorhersagequalität im Rahmen dieses Beitrags ist der Parameter C, welcher ein Maß dafür ist, wie scharf die Klassen voneinander getrennt werden (Breite des Margins).

Subjektiv extrahierter Merkmalsraum:

Um den Einfluss von C zu untersuchen, wurde dieser im Bereich von 10^{-2} bis 10^{4} variiert. Abb. 5 zeigt die Abhängigkeit von R^2 und RMSE für Trainings- und Testdaten bei Variation von C.

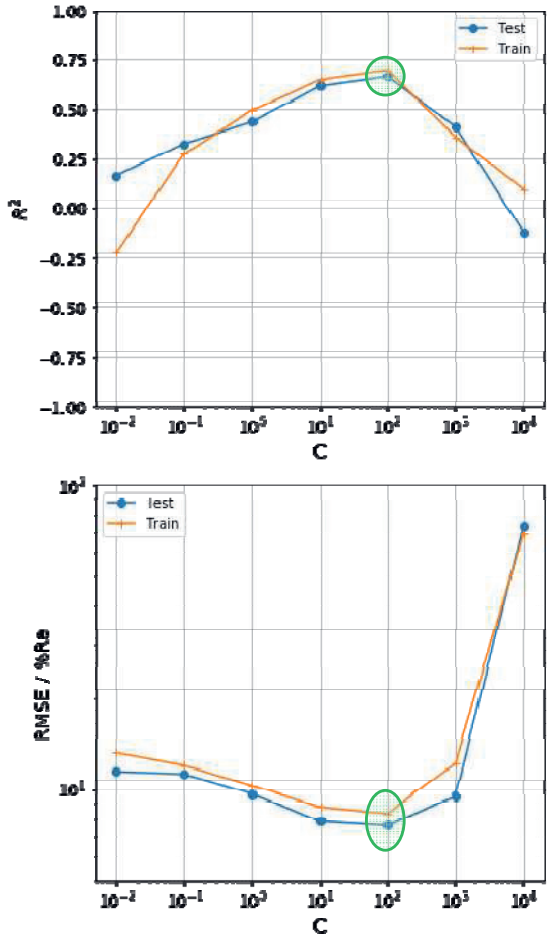


Abb. 5: R² (oben) und RMSE (unten) der Trainings- und Testdaten des subjektiv extrahierten Merkmalsraum für SV-R bei Variation von C

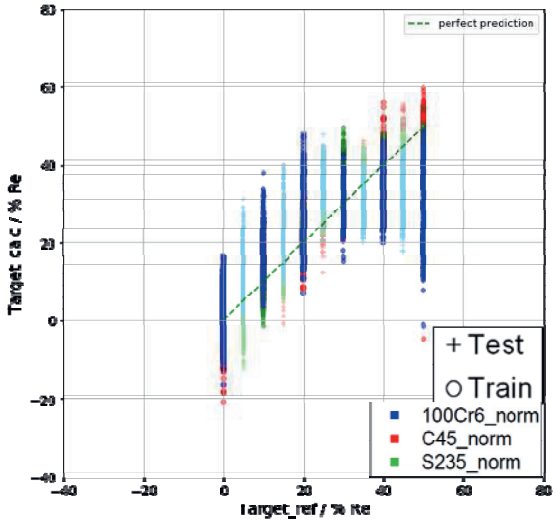


Abb. 6: Abbildung der Trainings- und Testdaten des subjektiv extrahierten Merkmalsraum für SV-R

Das Modell erreicht seine beste Performance von R^2 = 69,85 % und ein RMSE = 8,32 % R_e

bei den Trainingsdaten und R^2 = 66,69 % und ein RMSEV = 7,62 % R_e bei den Testdaten bei C = 100 (vgl. Abb. 5 und Abb. 6).

Automatisiert extrahierter Merkmalsraum:

Um den Einfluss von C zu untersuchen, wurde dieser im Bereich von 10^{-4} bis 10^{1} variiert. Abb. 7 zeigt wiederum die Abhängigkeit von R^2 und RMSE für Trainings- und Testdaten des automatisiert extrahierten Merkmalsraums bei Variation von C. Auch hier hängt die Vorhersagequalität des Modells für die Trainings- und Testdaten signifikant von C ab, siehe auch Abb. 8.

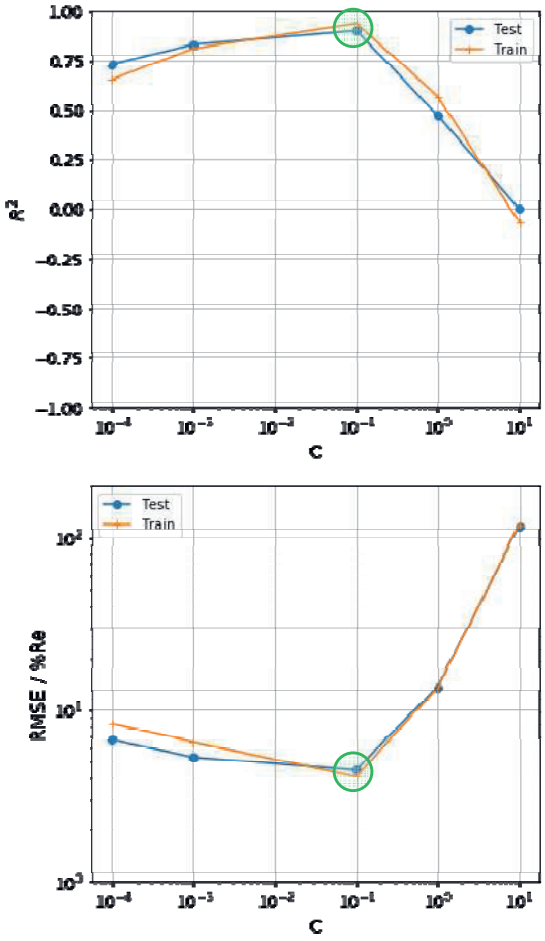


Abb. 7: R^2 (oben) und RMSE (unten) der Trainings- und Testdaten des automatisiert extrahierten Merkmalsraums für SV-R bei Variation von C

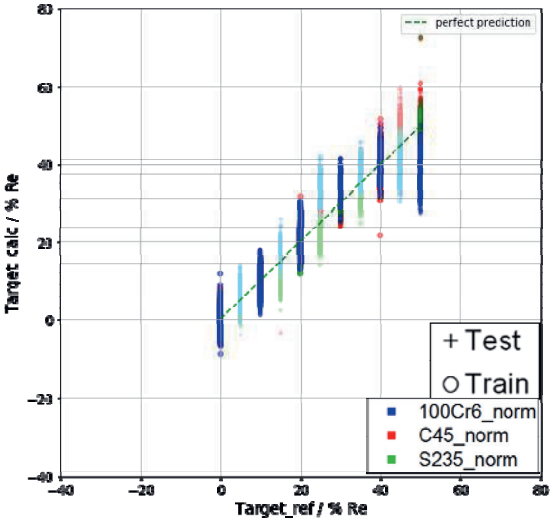


Abb. 8: Abbildung der Trainings- und Testdaten des automatisiert extrahierten Merkmalsraum für SV-R

Das Modell erreicht seine beste Performance mit R^2 = 93,96 % und RMSE = 4,11 % R_e bei den Trainingsdaten und R^2 = 90,33 % und RMSEV = 4,49 % R_e bei den Testdaten bei C = 0,1 (vgl. Abb. 7 und Abb. 8).

Multilayer-Perceptron-Regression

Die Multilayer-Perceptron-Regression ist ein Regressor aus der Familie der künstlichen neuronalen Netze [3]. Es handelt sich hierbei um ein Netz mit mindestens einer verborgenen Schicht (engl. Hidden Layer) von Neuronen. Um den Einfluss der Hyperparameter zu zeigen, werden im Rahmen dieses Beitrags alle Hyperparameter vorgegeben bis auf die Anzahl der Eingangsneuronen im Hidden Layer. Dieser Hyperparameter wird variiert, anschließend wird die Vorhersagequalität der Trainings- und Testdaten bewertet. Die wesentlichen vorgegeben Hyperparameter sind eine lineare Aktivierungsfunktion, ein Hidden Layer, Alpha = 10 und eine Lernrate von 0,01.

Lineare Aktivierungsfunktionen finden häufig Einsatz bei Regressionsproblemen. Aus der linearen Aktivierungsfunktion ergibt sich, dass die Neuronen der ersten verborgenen Schicht Linearkombinationen der Eingangsmerkmale darstellen und der Einsatz weiterer Hidden Layer redundant ist, da der Ausgang jeder verborgenen Schicht letztlich wieder eine Linearkombination der Eingangsmerkmale darstellt.

Subjektiv extrahierter Merkmalsraum:

Um den Einfluss der Anzahl an Neuronen im Hidden Layer zu untersuchen, wurde die Anzahl an Neuronen beginnend mit 1 bis zur

maximal zur Verfügung stehenden Anzahl an Merkmalen im Merkmalsraum erhöht.

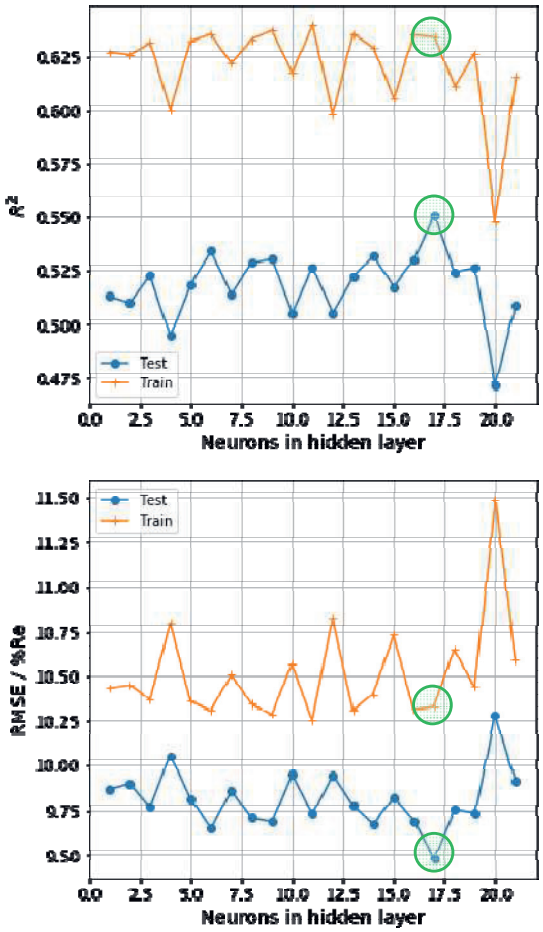


Abb. 9: R^2 (oben) und RMSE (unten) der Trainings- und Testdaten des subjektiv extrahierten Merkmalsraum für MLP-R unter Variation der Anzahl der Neuronen im Hidden Layer

Es zeigt sich, dass die Anzahl der Neuronen im Hidden Layer in dieser Konfiguration im subjektiven Merkmalsraum keinen signifikanten Einfluss hat. Sowohl die Abbildung der Trainings- als auch der Testdaten verbessert sich nur unwesentlich.

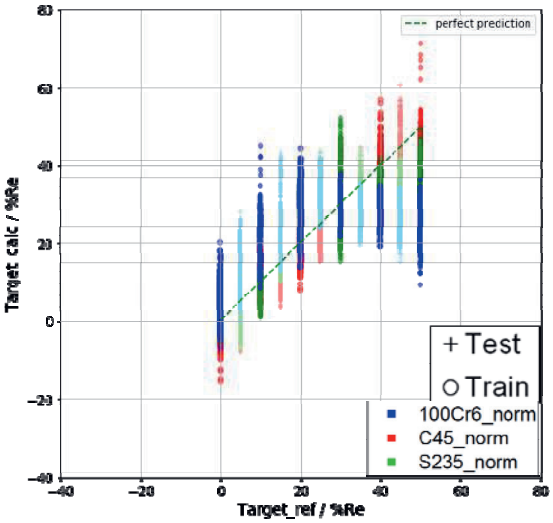


Abb. 10: Abbildung der Trainings- und Testdaten des subjektiv extrahierten Merkmalsraum für MLP-R

Das Modell erreicht seine beste Performance von R^2 = 63,42 % und ein RMSE = 10,33 % R_e bei den Trainingsdaten und R^2 = 55,08 % und ein RMSEV = 9,48 % R_e bei den Testdaten bei 17 Merkmalen (vgl. Abb. 9 und Abb. 10).

Automatisiert extrahierter Merkmalsraum:

Um den Einfluss der Anzahl an Neuronen im Hidden Layer zu untersuchen, stehen im automatisierten Merkmalsraum über 270 Merkmale zur Verfügung. Auch im automatisierten Merkmalsraum wurde die Anzahl der Neuronen beginnend mit 1 bis zur maximal zur Verfügung stehenden Anzahl an Merkmalen im Merkmalsraum erhöht.

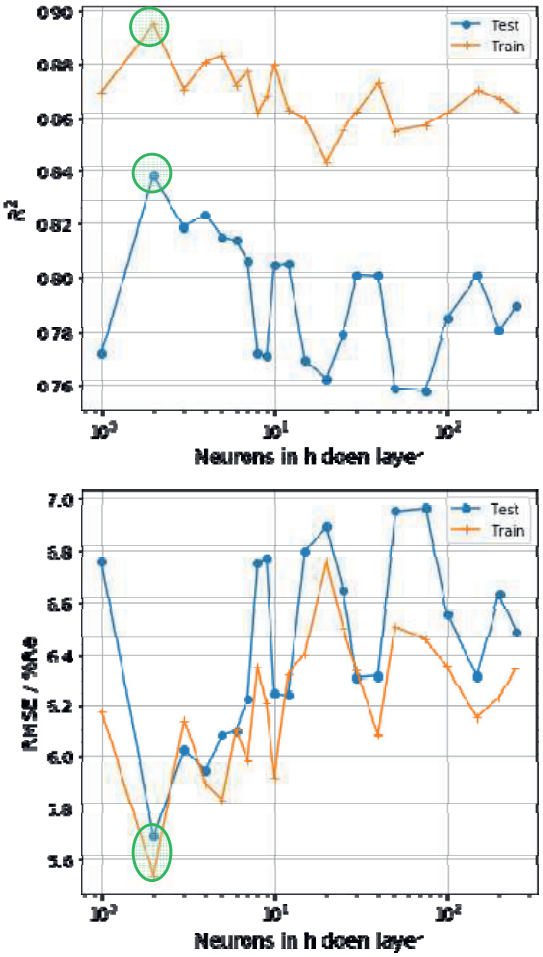


Abb. 11: R² (oben) und RMSE (unten) der Trainings- und Testdaten des automatisiert extrahierten Merkmalsraum für MLP-R unter Variation der Anzahl der Neuronen im Hidden Layer

Es zeigt sich, dass die Anzahl der Neuronen im Hidden Layer in dieser Konfiguration im automatisiert extrahierten Merkmalsraum einen signifikanten Einfluss hat.

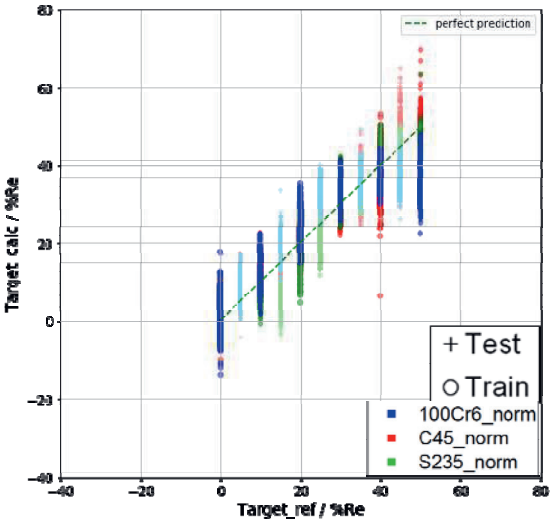


Abb. 12: Abbildung der Trainings- und Testdaten des automatisiert extrahierten Merkmalsraum für MLP-R

Das Modell erreicht seine beste Performance von R^2 = 89,49 % und ein RMSE = 5,54 % R_e bei den Trainingsdaten und R^2 = 83,82 % und ein RMSEV = 5,69 % R_e bei den Testdaten bei lediglich zwei Eingangsneuronen (vgl. Abb. 11 und Abb. 12).

Zusammenfassung und Ausblick

Im Rahmen dieses Beitrags wurde untersucht, mit welchen Möglichkeiten die datenbasierte Vorhersagequalität verbessert werden kann. Neben der Optimierung des Merkmalsraums wurde die multiple lineare Regression (ML-R) mit weiteren Regressionsalgorithmen, nämlich der Support-Vector-Regression (SV-R) und der Multilayer-Perceptron-Regression (MLP-R), verglichen. Diese Regressoren bieten aufgrund ihrer Mächtigkeit und Flexibilität, Nichtlinearitäten in den Daten abzubilden, enormes Potenzial zur Erstellung komplexer Modelle. Beide besitzen jedoch „Stellschrauben“ (sog. Hyperparameter) die signifikant zur Vorhersagequalität beitragen. Aufgrund der Vielzahl möglicher Kombinationen dieser Hyperparameter kann das Finden einer (quasi-) optimalen Lösung sehr aufwendig sein. Zur Verdeutlichung dieses Aufwandes wurden im Experiment alle Hyperparameter bis auf einen (C bei SV-R, Anzahl Neuronen im Hidden Layer bei MLP-R) vorgegeben. Dieser diente als Stellgröße und sollte zeigen, welcher Aufwand schon bei der Variation von einzelnen Hyperparametern zum Finden eines (quasi-) optimalen Modells erforderlich sind. Die ML-R wurde als Referenz verwendet. Um eine Vergleichbarkeit der ML-R mit der SV-R bzw. der MLP-R zu erhalten, wurden diese mit linearer Hyperebene im Merkmalsraum bzw. linearer

Aktivierungsfunktion verwendet. Tab. 1 und Tab. 2 zeigen die modellbasierte Vorhersagequalität der Merkmalsräume sowohl für die Trainings- als auch für die Testdaten.

Tab. 1: R^2 und RMSE für Trainings- und Testdaten des subjektiv extrahierten Merkmalsraum

	ML-R	SV-R	MLP-R
R^2_{Train} in %	78,07	69,85	63,42
$RMSE_{Train}$ in % R_e	7,99	8,32	10,33
R^2_{Test} in %	69,35	66,69	55,08
$RMSE_{Test}$ in % R_e	7,83	7,62	9,48

Tab. 2: R^2 und RMSE für Trainings- und Testdaten des automatisiert extrahierten Merkmalsraum

	ML-R	SV-R	MLP-R
R^2_{Train} in %	98,19	93,96	89,49
$RMSE_{Train}$ in % R_e	2,3	4,11	5,54
R^2_{Test} in %	87,54	90,33	83,82
$RMSE_{Test}$ in % R_e	4,99	4,49	5,69

Unter diesen Bedingungen ist es also nicht gelungen, eine signifikante Verbesserung der Vorhersagequalität in beiden Merkmalsräumen zu erreichen (vgl. Tab. 1 und Tab. 2). Hierbei ist jedoch anzumerken, dass die wahre Stärke dieser Modelle gerade die Behandlung von Nichtlinearitäten in den Daten ist, die Beschränkung auf eine lineare Hyperebene (SV-R) bzw. eine lineare Aktivierungsfunktion (MLP-R) die Mächtigkeit und Flexibilität dieser Modelle somit stark begrenzt. Neben C bzw. der Anzahl der Neuronen im Hidden Layer gibt es jedoch weitere zu optimierende Hyperparameter. Der Optimierungsaufwand hierfür wächst aber exponentiell mit der Dimension des Hyperparameterraums.

Abschließend lässt sich zusammenfassen, dass bei derart hohem Aufwand für die Optimierung von Hyperparametern die Möglichkeit der Optimierung des Merkmalsraums verbunden mit dem Einsatz linearer Modelle gegenübergestellt werden sollte. Beispielsweise kann die multiple lineare Regression weitestgehend ohne A-Priori-Wissen eingesetzt werden, die Berechnung ist sehr schnell und mit einem aussagekräftigen Merkmalsraum werden sehr gute Vorhersagen erreicht.

Im Rahmen dieses Beitrags wurden alle Merkmale der jeweiligen Merkmalsräume ohne Skalierung des Wertebereichs verwendet. Die Skalierung der Daten ist besonders bei der Verwendung von nichtlinearen Aktivierungsfunktionen relevant, da hier stark variierende Wertebereiche in unterschiedlichen Merkmalen das Ergebnis unterschiedlich stark dominieren. Die Wichtigkeit einzelner Merkmale wäre demnach von deren Wertebereich abhängig.

Es ist auch zu beachten, dass für die Zielgröße irrelevante Merkmale die Vorhersagequalität eines Modells verschlechtern können. Eine vorgeschaltete Merkmalsselektion kann die Vorhersagequalität ebenfalls verbessern [5].

Danksagung

Die im Rahmen dieses Beitrags durchgeführten Untersuchungen wurden mit der Python und der freien Softwarebibliothek ScikitLearn durchgeführt [6]. Ein besonderer Dank geht an die gesamte Community, die es uns Anwendern ermöglicht, die zur Verfügung stehende Infrastruktur und Tools einzusetzen.

Literaturnachweis

[1] Kneller, E.: Ferromagnetismus. Springer-Verlag Berlin, Göttingen, Heidelberg. 1962.

[2] Youssef, S.; Zimmer, C.; Szielasko, K.; Schütze, A.: “Bewertung subjektiver und automatisierter Merkmalsextraktion periodischer Zeitsignale am Beispiel des 3MA-X8-Verfahrens“ tm - Technisches Messen, 0.0 (2019): Retrieved 24 Jan. 2019, from doi:10.1515/teme-2018-0074

[3] Robinson, R.: Neural networks offer an alternative to traditional regression Geobyte, February (1991), pp. 14-19

[4] Drucker, H.; Burges, C.; Kaufman, L.; Smola, A.; Vapnik, V. 1997: "Support Vector Regression Machines", in Advances in Neural Information Processing Systems 9, NIPS 1996, 155–161, MIT Press.

[5] Schneider, T; Helwig, N and Schütze, A: Automatic feature extraction and selection for classification of cyclical time series data, tm Technisches Messen (2017), 84(3), 198–206, doi: 10.1515/teme-2016-0072

[6] Pedregosa, F. *et al.*:Scikit-learn: Machine Learning in Python, JMLR 12, pp. 2825-2830, 2011.

DOI 10.5162/sensoren2019/4.3.3

An empirical study of vibration based bearing fault diagnosis methods

Kurt Pichler[1], Ted Ooijevaar[2], Clemens Hesch[1], Christian Kastl[1]
[1]Linz Center of Mechatronics GmbH, Altenberger Straße 69, 4040 Linz, Austria
[2]DecisionS, Flanders Make, Gaston Geenslaan 8, 3001 Leuven, Belgium

Abstract

This paper presents an empirical study in which universally applicable fault diagnosis methods are used to evaluate vibration data of bearings. The data were acquired on two different test beds: a gear box test bed containing various bearings at different health states, and an accelerated life time (ALT) test bed to degrade a bearing and introduce an operational fault. Features are extracted from the raw data of two different accelerometers and used to monitor the actual health state of bearings. For that purpose, feature selection and classifier training is performed in a supervised learning approach. For testing the proposed approach, cross validation is applied to the data. The results of the gearbox test bed show that the classification accuracy data increases with the revolution speed of the bearing. Furthermore, the data of a high-end sensor allow higher classification accuracy than the data of a low-cost sensor. The results of the ALT test bed show that the same features that were identified in the gearbox test start to change significantly when the bearing degrades.

Keywords: Vibration Analysis, Fault Detection, Bearing, Feature Engineering, Machine Learning

Introduction

Manufacturing companies continuously try to increase their productivity, among others by avoiding machine down times. The latter involves considerable costs because of the resulting loss of turnover. Monitoring the condition of, for instance, bearings and gears, plays a vital role in the maintenance program of rotating machines. Early fault detection could allow to move from a time based preventive maintenance program to a condition based predictive maintenance strategy and reduce unexpected machine downtime and cost.

Vibration based condition monitoring is an established approach that has been employed by industries for many years in their maintenance program. However, up to this day, machine operators often still base their maintenance decisions on data from the periodical and manual inspection of single machines, which does not always result in correct conclusions. The common practice is that vibration measurements are periodically recorded using portable vibration sensors and measurement signals are analysed by an expert to interpret the machine's health condition. This approach can, however, lead to serious misinterpretation, where rapidly growing faults could be missed.

A continuous condition monitoring approach enables an early detection of machine faults. In this way, the machine condition is continuously tracked and total failures can be anticipated in advance, hence allowing appropriate maintenance actions. Despite its advantages, continuous monitoring program is however still not well adopted by industry. Firstly, because it often involves a high investment cost. Although recent advancements in sensor, acquisition and processing hardware have demonstrated cost-effective solutions [1,2], the economic benefit of the investment is still not clear and hard to quantify. Secondly, because many of those systems still require an expert to interpret the analysis results. Finally, also because it is not straight forward to select the most appropriate method for a specific application.

A wide range of vibration based bearing fault detection methods have been proposed in literature [3,4,5]. Approaches that utilize time domain features (e.g. crest factor, kurtosis), frequency and cepstral domain features (e.g. envelope analysis, cepstral coefficients) usually assume stationary machine conditions. Other methods such as cyclo-stationary analysis (i.e. second order technique in the frequency domain) and time-frequency domain analysis (e.g. Wigner-Ville distribution, Hilbert-Huang transform and wavelet transform based features) are more appropriate for non-

stationary processes. Some of those methods are purely data driven, whereas others use the physical relation between the bearing geometry, the rotational shaft speed and the bearing specific fault frequencies associated to the impulse behavior introduced by bearing faults.

In this paper, we present a purely data-driven method by extracting a large number of features from vibration data and subsequent selection and classification of those features in a supervised learning approach. Test data were acquired at two different bearing test beds: a gearbox setup that can be equipped with bearings of different degradation status, and a simple rotating shaft with a bearing under certain radial loads for accelerated degradation. The following chapters describe the data analysis method, the experimental setup and the evaluation results.

Description of the method

The fault diagnosis approach adopted here is a purely data-driven one, i.e. it incorporates no physical knowledge about the monitored system. That makes it on one hand more flexible and applicable to many kinds of systems, machines or components. On the other hand, incorporating extra knowledge usually improves the diagnostic ability of a condition monitoring system.

The first step of the proposed procedure is to extract a large number of features (for instance [6-16]) from the raw vibration signal. These features include simple statistical measures like standard deviation or kurtosis as well as more advanced features in time, (time-)frequency, quefrency and wavelet domain.

Subsequently, the dimensionality of the feature space is reduced to avoid the curse of dimensionality. Therefore, the significant features are identified by feature selection procedures as described in [17]. In particular, a standard forward selection filter algorithm selecting one feature per step was applied. As selection criterion in each step we adopted robust distance measures like Dy-Brodley measure [18] or Mahalanobis distance [19]. Feature selection is stopped when the relative gain of the selection criterion falls below 1%. In all of our tests, both distance measures selected the same features until the stopping criterion was reached.

After feature extraction and selection, a classifier in the feature space is trained. For that purpose we use linear or quadratic normal Bayes classifiers [20]. These classifiers assume normally distributed classes and apply Bayes decision rule to those classes. In the linear case, the co-variances of the classes are assumed to be equal. The validity of using normally distributed classifiers was checked by normal probability plots of the feature vectors (see two examples in Fig. 1).

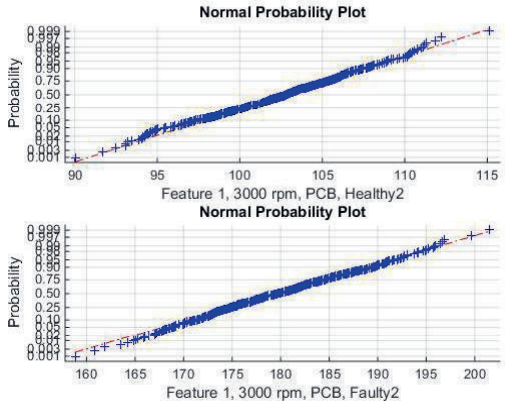


Fig. 1: Normal probability plots of Feature 1, 3000 rpm data, PCB sensor, states Healthy2 and Faulty2

The supervised learning approach implies that the method depends on having a sufficient amount of annotated training data for all states (failure modes) to be monitored. There are also classifiers for one-class classification (also referred to as novelty detection) available [21]. However, those techniques detect only deviation from a nominal state and are thus prone to over detection due to changing operation modes. Furthermore, the feature selection process depends on having an annotated data set as well.

The process of evaluating a new data sample is straightforward: the selected features are extracted from the raw signals, and the classifier is applied to those features. Since many classifiers are able to deliver class membership probabilities, it is generally possible to determine instances lying between two distinct states. However, here we restrict the evaluation to crisp class decisions by detecting the maximum class probability for each observation.

Given annotated training data, the whole process of feature extraction, feature selection and classification can be fully automated. The more useful information the training data contain, the better the resulting feature subset and classifier will be. In this context, information means different states, rotation speed, repeated measurements with different samples of the same bearing type, and so on.

Experimental Setup

Two types of experiments have been performed: (i) an accelerated life time test (ALT) of a ball bearing on a single shaft drive

train set-up and (ii) a test on a more complex gearbox set-up including bearings with various faults. These setups were provided by the scientific partner *FLANDERS MAKE* in Leuven. The ALT allows to create an operational fault in a bearing. This test differentiates from other studies in the fact that they are often limited to artificially induced faults. Moreover, the fault evolution and accumulation can be monitored during the ALT.

The experimental set-up used to perform the accelerated life time test is shown in Fig. 2. The set-up comprises of a single shaft with a test bearing. The shaft is supported with the help of a support bearing on each side. A hydraulic cylinder is used to apply a radial load to the test bearing up to a maximum of 10 kN. The set-up is driven by a motor at a fixed rotation speed of 1500 rpm.

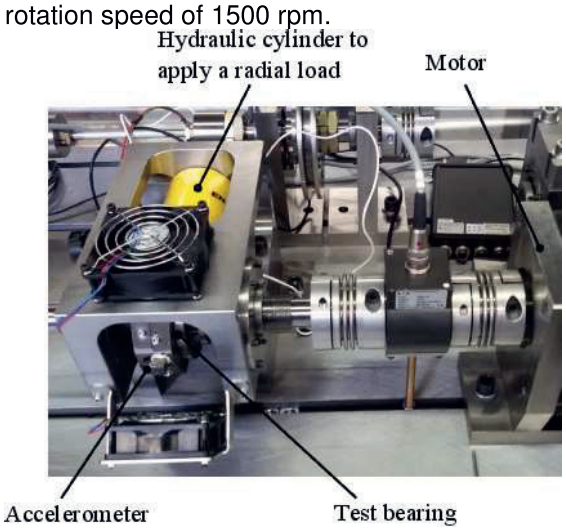


Fig. 2: The drive train set-up used to reduce the life time of a bearing to less than a day, allowing to generate vibration data during the accumulation of an operational bearing fault.

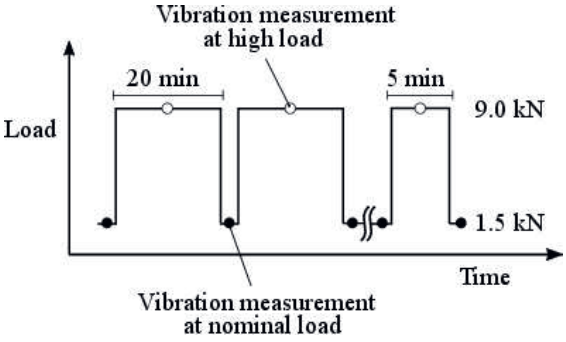


Fig. 3: The load was temporarily increased from 1.5 kN to 9 kN to accelerate the wear of the bearing.

The test procedure is schematically illustrated in Fig. 3. Vibration measurements were performed under a nominal radial load of 1.5 kN (i.e. 10% of the dynamic load rating). The radial load was temporarily increased to 9.0 kN (i.e. 65% of the dynamic load rating) to accelerate the degradation of the bearing. In the beginning the interval was 20 minutes, but this has been reduced as soon as the first indication of an incipient fault was noticed in the measured vibration responses. In total 30 vibration measurements were performed at the nominal 1.5 kN loading condition and 29 vibration measurements at the high 9.0 kN radial load.

The applied radial load, the radial vibrations in the loading direction and the temperature of the bearing housing were measured during the test. The machine vibrations were measured using a piezo film ACH-01-03 accelerometer and digitized at 12.8 kHz by an embedded acquisition platform. In each measurement 20 seconds of data was acquired. The accelerated life time test was stopped when a vibration peak level of +/- 50 g was reached.

The ALT was performed on a FAG 6205 ball bearing. Before the start of the test a small indentation with a diameter of 230 μm was created in the inner race. This indentation is used as a local stress riser and represents a local plastic deformation caused by, for instance, a contamination particle. Subsequently, the ALT was performed for several hours. Although bearings can fail in many different ways, the indentation triggers the bearing to fail in a more repeatable way. The test was stopped when severe rolling contact surface fatigue occurred at the inner race [22]. The start and the end condition of the inner race of the test bearing are shown in Fig. 4.

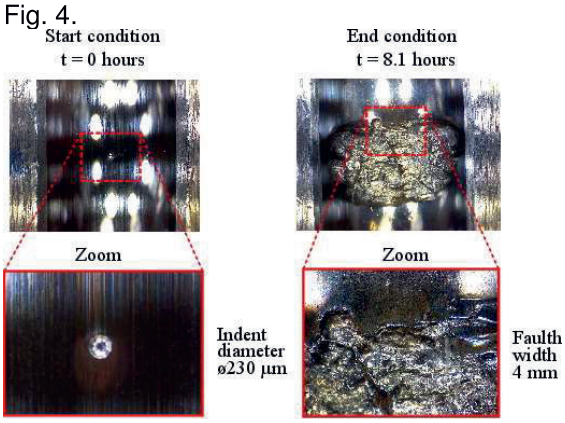


Fig. 4: The indentation at the bearing inner race used as the start condition and the surface fatigue fault at the inner race introduced by the ALT.

The second test performed in this study was an industrially representative gearbox setup. Fig. 5 shows a photograph and a schematic top-view of the gearbox setup. The test setup

consists of (i) an induction electric motor, (ii) a gearbox and (iii) a magnetic brake. The motor is controlled by a variable-frequency-drive (VFD) with either a stationary mode or a transient mode (run-up/run-down). The motor speed can be controlled from 0 to 3000 rpm. The gearbox input shaft is connected to the motor, the gearbox output shaft is coupled to the brake. The torque applied to the brake can be adjusted by the controller from 0 to 50 Nm.
As illustrated in Fig. 5, the gearbox comprises of three-parallel shafts connected through contacting spur gear pairs. Note that the number of gear teeth is indicated in the figure. The input shaft is supported by deep groove ball bearings MB ER-10K, while the other shafts are supported by deep groove ball bearings MB ER-16K. For simulating a healthy or faulty state on the gearbox, the right-side bearing housing that supports the second shaft is equipped either with a healthy or a damaged ball bearing of type FAG 6205-C-TVH.
Two healthy bearings and three faulty bearings with different inner race faults were tested. An indentation fault with a diameter of 490 μm was created. Two other bearings with operational faults were created using the ALT setup as described above. The healthy bearings are referred as 'Healthy1' and 'Healthy2', while the faulty bearings are referred as 'Indent', 'Faulty1', 'Faulty2', in the order of increasing severity and are illustrated in Fig. 6.

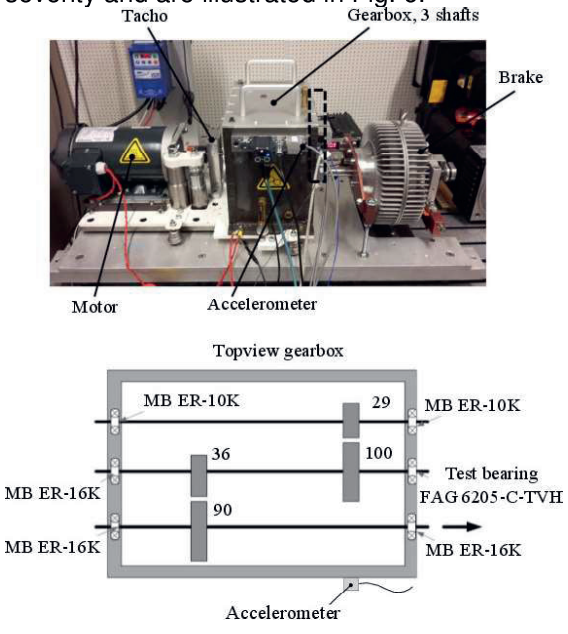


Fig. 5: Gearbox set-up comprising a motor, 3-shaft gearbox and brake to introduce a load.

For each healthy or faulty state, 2 operating conditions were imposed on the gearbox setup, namely two different motor speeds of 1500 and 3000 rpm, respectively. The brake torque was kept constant at 50 Nm. Because of the transmission ratio, the rotational speed of the second shaft is 29/100 lower than that of the motor speed, while the torque applied on the second shaft is 36/90 lower than that of the brake torque. Hence, for the imposed operating conditions, the rotational speeds of the second shaft were 435 and 870 rpm, while the torque applied to the second shaft was 20 Nm. A high-end PCB accelerometer and a low cost MEMS accelerometer were mounted on the gearbox housing as shown in Fig. 5. The vibration signals were sampled at 50 kHz in 20 seconds data batches.

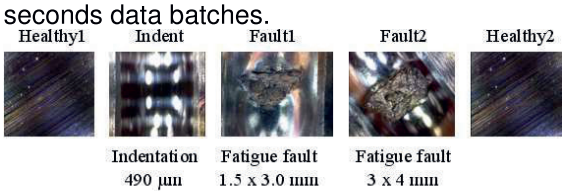


Fig. 6: Five bearing states tested on the gearbox setup comprising two healthy bearings and three faulty bearings with different severities.

Results

In the ALT the features were extracted from the raw sensor signals in an overlapping sliding window approach (window length 0.2 seconds, overlap 0.1 seconds). That yields in 199 observations out of each 20 second data batch. However, for final evaluation, only the mean value of those 199 observations of a data batch is observed. Unlike the gearbox setup, we had no data of different health states available for feature selection and classifier training in the accelerated life time test. Therefore we were restricted to detect significant changes in the feature values. For that purpose, cumulative sum (CUSUM) control charts [23] were applied to the features. Due to the missing feature selection step, we evaluated the top ranked features of the gearbox test. All of those features increased significantly towards the end of the test run for the 9.0 kN as well as for the 1.5 kN load conditions. For example, feature 20 is depicted in Fig. 7. Due to the increasing feature value, the upper threshold of the CUSUM control chart was exceeded after approximately 7.3 hours in the 9.0 kN case and 7.9 hours in the 1.5 kN case, indicating a failure of the bearing. For control purposes, Fig. 8 shows an arbitrarily chosen feature. The feature shows no significant trend, and the CUSUM control charts don't exceed the thresholds.
Just like in the ALT, the features for the gearbox test were extracted from the raw acceleration signals in an overlapping sliding window approach with a window length of 0.2

seconds and an overlap of 0.1 seconds, delivering again 199 observations for each 20 second data batch. After extraction of all features, the feature selection algorithm was applied to the 3000 rpm and 1500 rpm motor speed data independently. Furthermore, we did not use all available states for feature selection. We used only the datasets of the states Healthy1, Healthy2 and Faulty2, but not Indent and Faulty1.

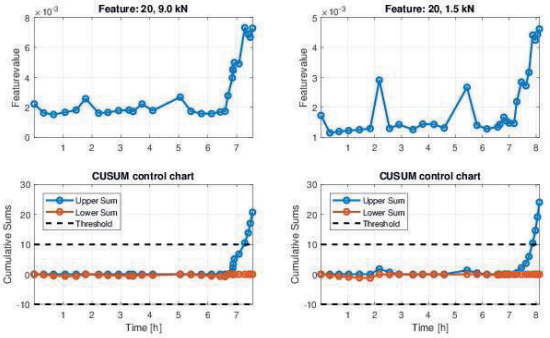


Fig. 7: Feature values and CUSUM control charts for feature 20 in the ALT test for load 9.0 kN (left) and 1.5 kN (right).

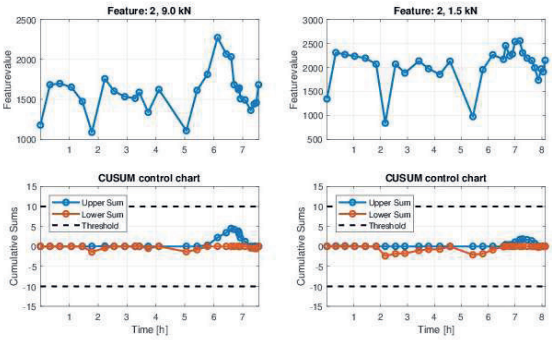


Fig. 8: Feature values and CUSUM control charts for feature 2 in the ALT test for load 9.0 kN (left) and 1.5 kN (right).

We first evaluate the data of the high-end PCB accelerometer. For the 3000 rpm motor speed data, the algorithm selected 3 top ranked features and for the 1500 rpm data it selected 4 top ranked features. However, for a first visual impression we show only the top 2 features for all recorded rotation speeds and all recorded states in a scatter plot in Fig. 9 (3000 rpm) and Fig. 10 (1500 rpm).
The scatterplots indicate already a few possible conclusions:

- Different top features were selected for the different rotation speeds
- The 3000 rpm dataset revealed better separability
- Faulty1 and Faulty2 produced similar feature values
- In the 3000 rpm dataset, the Indent-class lies somewhere in-between the healthy and the faulty states

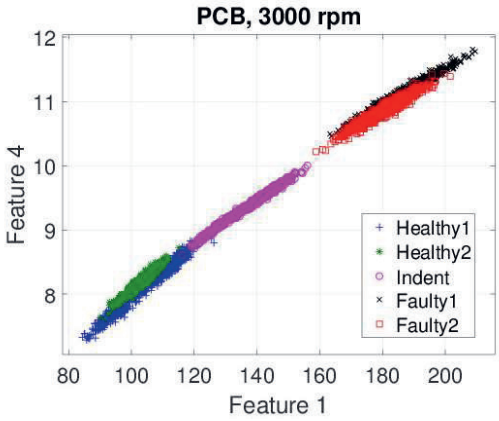


Fig. 9: Scatter plot of top 2 features for the 3000 rpm PCB dataset.

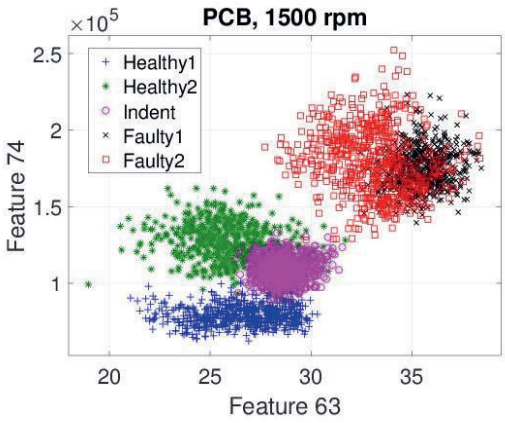


Fig. 10: Scatter plot of top 2 features for the 1500 rpm PCB dataset.

For classification accuracy estimation we trained a quadratic normal Bayes classifier only using observations of the states Healthy2, Indent and Faulty2 (using two 20 second data batches each). After training, validation was performed with another 20 second data batch of all 5 bearing states. However, as target class of the classifier we did not use those five states, but only the simplified states Healthy (Healthy1 and Healthy2), Indent and Faulty (Faulty1 and Faulty2). According to the feature selection step, the 3000 rpm data were validated with the top 3 features and the 1500 rpm data were validated with the top 4 features. The validation yields 99.30% accuracy for the 3000 rpm data (confusion matrix in Tab. 2) and 82.41% accuracy (confusion matrix in Tab. 2) for the 1500 rpm data. That result confirms the first conclusions above: the separability of the states Healthy, Indent and Faulty in the 3000 rpm case is satisfying, while the separability is worse in the 1500 rpm case. Especially the state Healthy1

is misclassified in the 1500 rpm case. Since Healthy1 was not used for classifier training, it is obviously more likely to be misclassified.

Tab. 1: Confusion matrix for 3000 rpm PCB data and top 3 features.

Confusion Matrix		Estimated State		
		Healthy	Ident	Faulty
True State	Healthy1	192	7	0
	Healthy2	199	0	0
	Indent	0	199	0
	ALT2	0	0	199
	ALT3	0	0	199

Tab. 2: Confusion matrix for 1500 rpm PCB data and top 4 features.

Confusion Matrix		Estimated State		
		Healthy	Ident	Faulty
True State	Healthy1	33	166	0
	Healthy2	198	1	0
	Indent	5	194	0
	ALT2	0	0	199
	ALT3	3	0	196

To compare the results of the PCB sensor with the MEMS sensor, we extracted the same features from the raw data of the MEMS sensor and validated them in the same way. For the 3000 rpm case that yields in 93.67% accuracy (confusion matrix in Tab. 3, scatterplot of top 2 features in Fig. 11) and in the 1500 rpm case in 78.99% accuracy (confusion matrix in Tab. 4, scatterplot of top 2 features in Fig. 12). In both cases, the accuracy of the low-cost MEMS sensor data is lower than the accuracy of the high-end PCB sensor data. However, one might argue that the comparison is not fair, since the feature selection was performed with the PCB data. In our experiments, the MEMS sensor did not perform significantly better when the feature selection was performed using the MEMS data.

Tab. 3: Confusion matrix for 3000 rpm MEMS data and top 6 features.

Confusion Matrix		Estimated State		
		Healthy	Ident	Faulty
True State	Healthy1	199	0	0
	Healthy2	138	61	0
	Indent	2	197	0
	ALT2	0	0	199
	ALT3	0	0	199

Tab. 4: Confusion matrix for 1500 rpm MEMS data and top 4 features.

Confusion Matrix		Estimated State		
		Healthy	Ident	Faulty
True State	Healthy1	22	177	0
	Healthy2	199	0	0
	Indent	8	185	6
	ALT2	0	0	199
	ALT3	11	7	181

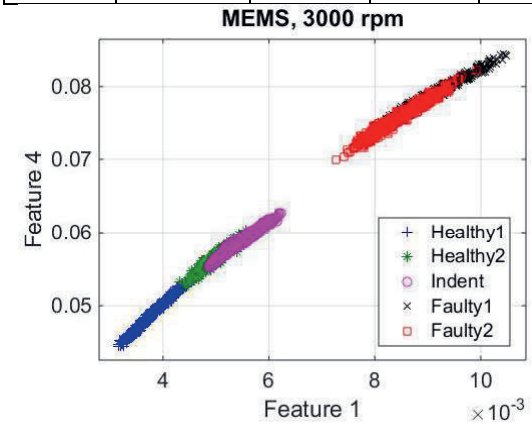


Fig. 11: Scatter plot of top 2 features for the 3000 rpm MEMS dataset.

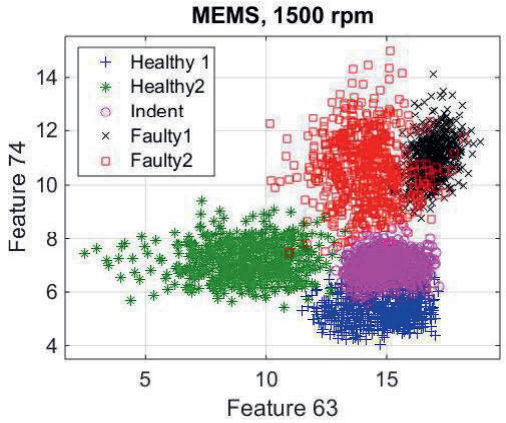


Fig. 12: Scatter plot of top 2 features for the 1500 rpm MEMS dataset.

Conclusions

The results of this empirical study of vibration based fault diagnosis methods for bearings show that the feature based approach can be used for fault monitoring of bearings. Furthermore it can be concluded that the classification accuracy increases with higher rotational speed of the bearing. The high-end PCB sensor allows higher accuracy values than the low-cost MEMS sensor. The ALT test shows that an upcoming fault can be detected before total failure of the bearing occurs.

The proposed method does not require any parameters or physical information about the system/component to be monitored. This

makes it very flexible and useful for manifold industrial applications.

Acknowledgments

This work has been supported by the COMET-K2 Center of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.
This research was partially supported by Flanders Make, the strategic research centre for manufacturing industry.

References

[1] A. Albabar, S. Mekid, A. Starr, R. Pietruszkiewicz, *Suitability of MEMS accelerometers for condition monitoring: An experimental study*, Sensors 8(2), 784-799 (2008)
[2] A.P. Ompusunggu, T. Ooijevaar, B. Kilundu, S. Devos, *Automated bearing fault diagnostics with cost-effective vibration sensor*, In J. Mathew, C. Lim, L. Ma (Eds.), *Asset Intelligence through Integration and Interoperability and Contemporary Vibration Engineering Technologies*, Proceedings of the 12th World Congress on Engineering Asset Management, 469-472 (2018)
[3] P. Henriquez, J. Alonso, M. Ferrer, C. Travieso, Review of automatic fault diagnosis systems using audio and vibration signals, IEEE Transactions on Systems, Man, and Cybernetics: Systems 44(5), 642-652 (2014)
[4] A. Sait, Y. Sharaf-Eldeen, *A Review of Gearbox Condition Monitoring Based on vibration Analysis Techniques Diagnostics and Prognostics*, In T. Prolux (Ed.), *Rotating Machinery, Structural Health Monitoring, Shock and Vibration* 5, Proceedings of the 29th IMAC: A Conference on Structural Dynamics, The Society for Experimental Mechanics, 359-374 (2011)
[5] D. Wang, K. Tsui, Q. Mia, *Prognostics and Health Management: A Review of Vibration Based Bearing and Gear Health Indicators*, IEEE Access 6, 665-676 (2017)
[6] S. Singh, M. Vishwakarma, *A Review of Vibration Analysis Techniques for Rotating Machines*, International Journal of Engineering Research & Technology 3, 757-761 (2015)
[7] M. Jalil, F.A. Butt, A. Malik, *Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals*, The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering, 208-212 (2013)
[8] V. Sharma, A. Parey, *A review of gear fault diagnosis using various condition indicators*, Procedia Engineering 144, 253-263 (2016)
[9] P.D. McFadden, M.M. Toozhy, *Applications of Synchronous Averaging to Vibration Monitoring of Rolling Element Bearings*, Mechanical Systems and Signal Processing 14(6), 891-906 (2000)
[10] H. Heidari Bafroui, A. Ohadi, *Application of wavelet energyband Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions*, Neurocomputing 133, 437-445 (2014)
[11] Y. Lei, Z. He, Y. Zi, Q. Hu, *Fault diagnosis of rotating machinery based on multiple ANFIS combination with GAs*, Mechanical Systems and Signal Processing 21, 2280-2294 (2007)
[12] C. Shen, D. Wang, F. Kong, P.W. Tse, *Fault diagnosis of rotating machinery based on the statistical parameters of wavelet packet paving and a generic support vector regressive classifier*, Measurement 46, 1551-1564 (2013)
[13] R: Bajric, N. Zuber, G.A. Skrimpas, N. Mijatovic, *Feature Extraction Using Discrete Wavelet Transform for Gear Fault Diagnosis of Wind Turbine Gearbox*, Shock and Vibration (2016)
[14] J. Antoni, R.B. Randall, *The spectral kurtosis: application to the vibratory surveillance and diagnostics of rotating machines*, Mechanical Systems and Signal Processing 20, 308-331 (2006)
[15] F.d.A. Boldt, T.W. Rauber, F.M. Varejao, *Feature Extraction and Selection for Automatic Fault Diagnosis of Rotating Machinery* (2013)
[16] H. Konstantin-Hansen, H. Herlufsen, *Envelope and Ceptstrum Analyses for Machinery Fault Identification*, Sound & Vibration, 10-12 (2010)
[17] I. Guyon, A. Elisseeff, *An introduction to variable and feature selection*, Journal of Machine Learning Research 3, 1157-1182 (2003)
[18] J.G: Dy, C.E. Brodley, Feature Selection for Unsupervised Learning, Journal of Machine Learning Research 5 (2004)
[19] G.J. McLachlan, Mahalanobis distance, Resonance 4(6), 20-26 (1999)
[20] D. de Ridder, D.M.J. Tax, B. Lei, G. Xu, M. Feng, Y. Zou, F. can der Heijden, *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB*, John Wiley & Sons, Ltd. (2017)
[21] D. Tax, *One-class classification*, PhD thesis, Delft University of Technology (2001)
[22] J. Halme, P. Anderson, *Rolling contact fatigue and wear fundamentals for rolling bearing diagnostics - state of the art*, Journal of Engineering Tribology, 377-393 (2009)
[23] D.M. Hawkins, D.H. Olwell, *Cumulative Sum Charts and Charting for Quality Improvement*, Springer (1998)

DOI 10.5162/sensoren2019/4.3.4

Enhanced performance of thermographic asset monitoring by an interpolation-based correction algorithm

Aydin Boyaci[1], Stephan Wildermuth[1], Holger Kaul[1], Tomas Kozel[2]
[1]ABB AG, Corporate Research, Wallstadter Str. 59, 68526 Ladenburg, Germany
[2]ABB s.r.o., Vídeňská 117, 619 00 Brno, Czech Republic

Abstract

Thermal condition monitoring of electrical assets requires an accurate temperature sensing to improve the lifetime and to prevent catastrophic overheating failures of the equipment. Low-resolution infrared (IR) cameras allow a permanent installation of temperature supervision in medium voltage switchgears. Besides their benefits of contactless measurement of temperature and capturing a large section of the electrical asset, low-resolution IR cameras lack in accuracy if the hot spot is spread over a small number of pixels in the thermal image. This contribution introduces an interpolation-based temperature correction algorithm to enhance the accuracy considerably of the temperature in the hot spot region. The feasibility of the proposed algorithm is demonstrated for a temperature rise test of a medium voltage switchgear under realistic conditions over the range of hot spot temperatures from 40°C to 100°C.

Keywords: Low-resolution infrared camera, hot spot temperature measurement, medium voltage switchgear, condition monitoring

Introduction

Accurate online temperature measurement becomes increasingly important to perform condition monitoring and diagnostics of electrical assets to avoid catastrophic overheating events as well as to enhance the lifetime of the equipment. Infrared (IR) cameras allow for contactless measurement of temperature, offering a significant advantage compared to contacting methods like RFID- or SAW- sensors [1]. Low-resolution IR cameras (e.g. 32x32 pixels) with wide angle optics are ideally suited to capture a large part of the asset's internal structure with only one sensor, thus offering a commercially attractive solution with minimal impact on the dielectric properties [2]. In such a configuration, however, heated areas like bus bars or spouts only spread over a small number of pixels, in particular if located at the boundaries of the image. During measurements in real medium voltage (MV) switchgear panels our measurements indicate that the temperature of small-sized hot areas may get underestimated by the IR camera compared to a reference measurement with a conventional reference thermocouple. As this would limit the application range of the monitoring system, we have developed an interpolation-based temperature correction algorithm for low-resolution IR cameras which enhances the measurement accuracy of small-sized structures during online thermography.

Temperature monitoring in MV switchgears

Reliable temperature condition monitoring of a MV switchgear comprises different compartments and components (cf. Figure 1). Especially the contact locations are prone to high temperatures. Thus, the highest temperatures can be usually observed in the joints of the bus bar compartment which are the first choice to monitor. In addition to that, the temperature development of the circuit breaker spouts needs to be monitored. During a regular replacement of the breaker, a damage of the mechanical contact spring at the spouts might occur resulting in an increased electrical contact resistance and therefore high temperatures. At the bottom of the switchgear, the cable terminals are a further source of possibly high temperatures to be supervised since the cable connections can be incorrectly fastened in need of a cable rearrangement.

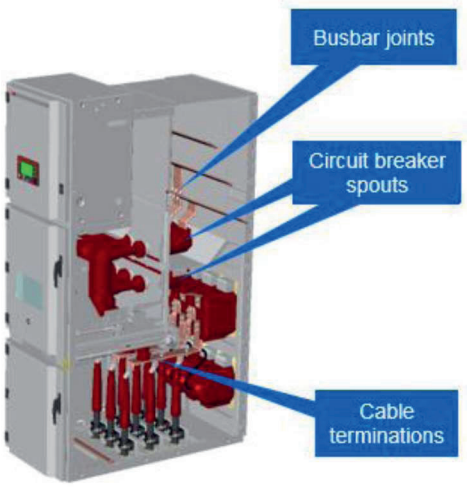


Figure 1: Schematic view of a MV switchgear panel. Temperature monitoring at the following locations: bus bar compartment (top), upper and lower contacts of the circuit breaker (middle), and cable compartment (bottom).

In a previous study [2], the feasibility of low-resolution IR cameras has been proved for the critical locations given in Fig. 1. The main advantage is that a single IR sensor is sufficient to monitor the temperature of the three phases for each location compared to an installation of three point-like sensors, i.e. one for each phase.

Derivation of the interpolation-based temperature correction algorithm

The temperature correction algorithm is derived beforehand from reference measurements with a large variety of hot area topologies. As outlined in Fig. 2a, a simple test rig is built up where the temperature T_{obj} of an object can be controlled individually with a high degree of accuracy. The IR sensor is placed at different lateral and perpendicular positions to the object for capturing the temperature of hot spots of varying size as well as location in the image.

The derivation of the proposed algorithm consists of two main steps. Firstly, a spline interpolation is applied on the raw image of the infrared camera to increase the spatial image resolution which enhances the characterization of the hot spot in temperature and size (cf. Fig. 2b). This way, important features of the hot spot, in particular its size and initial temperature, are determined. Furthermore, the temperature distribution in close proximity to the hot area gets analyzed to extract input parameters for the correction routine and its derivation.

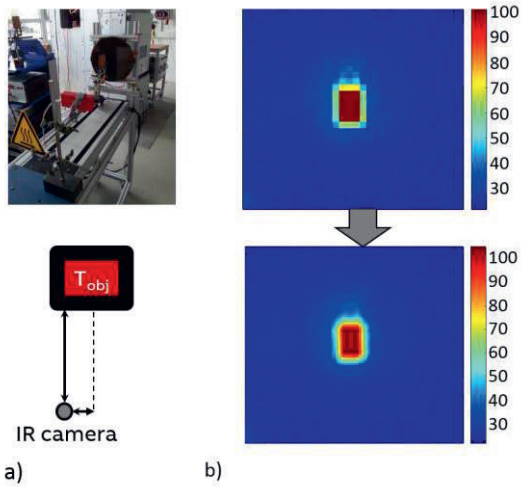


Figure 2: Reference measurements for deriving the temperature correction algorithm of hot spots at different size and location in the image. a) Schematic measurement setup. b) Comparison between a raw image measured by a low-resolution IR camera and the spline interpolation of the raw image.

Secondly, the hot spot size-dependent algorithm is derived from the error between the measured IR temperature and the real object temperature T_{obj} for different hot spot sizes. The curve fitting of the error leads to a global correction function over all hot spot sizes. Thus, this global correction function is dependent on the hot spot size and is valid for hot spot temperatures in the range from 40°C to 100°C.

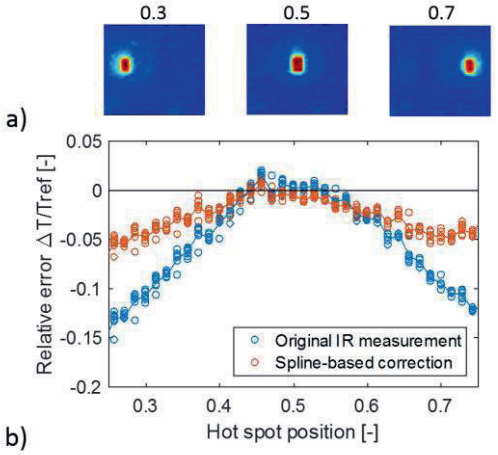


Figure 3: Temperature correction algorithm for a hot spot at different lateral positions in the image: a) Definition of hot spot position. b) Comparison of original IR-measured temperature and spline-based correction.

In Fig. 3., the interpolation-based correction algorithm is applied for a hot spot temperature of 60°C and different lateral positions of the hot spot in the thermal image. The fisheye effect of the IR camera lens leads to a hot spot size reduction and thus to a larger underestimation of the temperature when the object reaches the boundaries of the image. As can be seen from Fig. 3b, the correction algorithm is capable to reduce considerably the temperature error measured by the IR camera, even in case of small-sized hot spots at the image boundaries.

Temperature correction in a MV switchgear

The interpolation-based correction algorithm is applied for a typical temperature rise test in a MV switchgear. The temperature rise test considers the heating up of the entire switchgear by applying the nominal current for several hours until a steady state temperature is reached. Representative for all mentioned monitoring locations, the section of the lower breaker spouts is discussed in detail below.

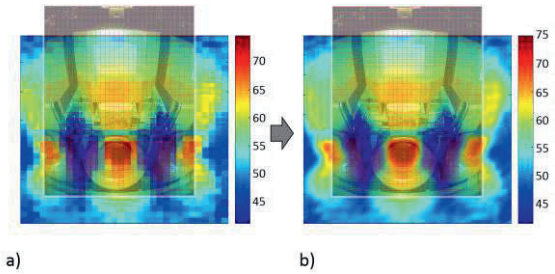


Figure 4: Temperature distribution along the section of the lower breaker spouts of a MV switchgear (IR image and visible light image superimposed): a) Raw image measured by a low-resolution IR camera. b) Spline interpolation of raw image.

Fig. 4 demonstrates exemplary the spline interpolation of the thermal image for the section of the lower breaker spouts with the three phases at a steady state. Performing a spline interpolation at every time step of IR measurement, the hot spot temperature and size are determined as important features from the interpolated images. These features are used as input to the correction algorithm delivering the final temperature estimate of the hot spot. The corrected temperatures show a good agreement with the measurement of reference thermocouples (cf. Fig. 4), additionally installed at the considered hot spots of the switchgear.

We have demonstrated for a real temperature rise test, that the overall accuracy of low-resolution IR cameras can be improved to become better than +-3°C. By applying the interpolation-based correction algorithm, the feasibility has been proved to use low-resolution IR cameras in future applications for condition monitoring of switchgears.

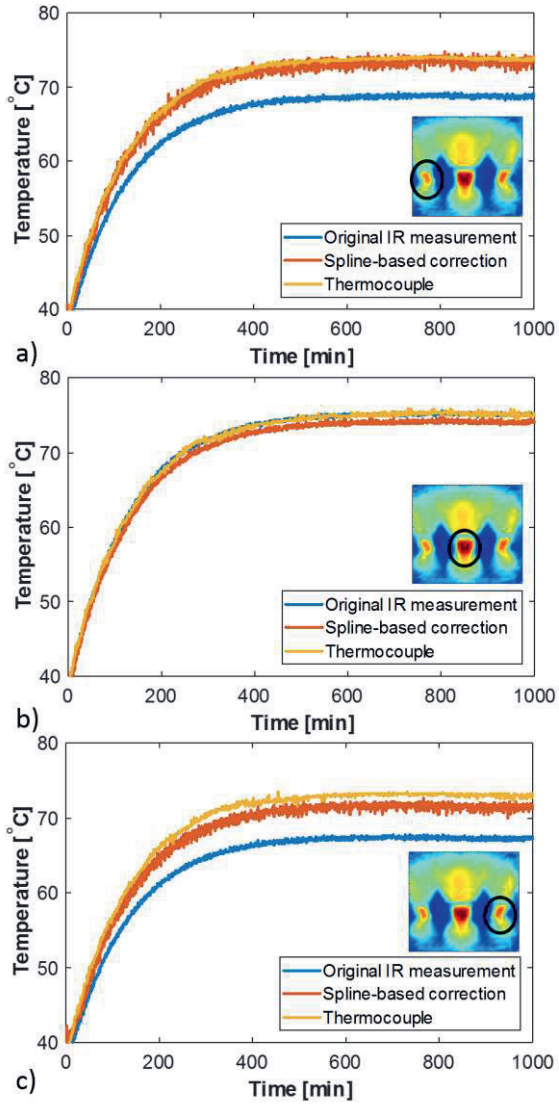


Figure 4: Comparison of original IR-measured temperature, spline-based correction and reference thermocouple temperature along the section of the lower breaker spouts of a MV switchgear at different phases: a) Left phase. b) Middle phase. c) Right phase.

Conclusion

The combination of the proposed temperature correction algorithm and a spline interpolation offers more accurate thermographic measurement by low-resolution IR cameras and makes these systems applicable for an even wider range of online condition monitoring applications in electrical assets.

Literature

[1] J.C. Andle, S. Sabah, D. S. Stevens, S. J. Jumani, M. Baier, B. W.A. Wall, T. Martens, R. Gruenwald, “Temperature Monitoring System Using Passive Wireless Sensors for Switchgear and Power Grid Asset Management”, MNC-CIGRE/CIRED Malaysia, T&D Asset Management Workshop (2010)

[2] S. Wildermuth, P. Szasz, J. Gebhardt, H. Kaul, K. König, “Infrared temperature sensing in electrical equipment by low-cost IR cameras”, VDE-Hochspannungstechnik, Berlin (2010)

DOI 10.5162/sensoren2019/4.4.1

Zustandsüberwachung in der Automatisierungstechnik mittels maschinellem Lernen

Steffen Klein[1], *Tizian Schneider*[1], *Andreas Schütze*[2]
[1]*ZeMA – Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Saarbrücken, Deutschland*
[2]*Universität des Saarlandes, 66123 Saarbrücken, Deutschland*

Zusammenfassung

Ziel des vom BMBF geförderten Verbundprojekts MoSeS-Pro war die Entwicklung eines modularen Sensorsystems für die Echtzeitprozesssteuerung und die smarte Zustandsbewertung. Basis dieses Sensorsystems ist ein modularer Hard- und Softwarebaukasten, welcher je nach Anforderung der spezifischen Anwendung verschiedenste Kombinationen von Sensoren und Algorithmen gestattet. Der entwickelte Softwarebaukasten kombiniert dabei fünf verschiedene Methoden der Merkmalsextraktion mit drei Ansätzen zur Merkmalselektion, wodurch unabhängig von der Herkunft der Daten ein nahezu ideales Klassifikationsergebnis erreicht wird. Die Leistungsfähigkeit des Baukastens wurde innerhalb des Projektes am Beispiel der Lebensdauerprognose von elektromechanischen Zylindern demonstriert. Hierzu wurde ein Lebensdauerprüfstand für elektromechanische Zylinder entwickelt, wodurch Signale verschiedenster Sensoren über die gesamte Lebensdauer des Zylinders aufgezeichnet werden können. Die Auswertung der Daten zeigt, dass für einzelne Zylinder eine hochaufgelöste Verschleißdetektion möglich ist, die eine Zustandsbewertung mit einer Auflösung von einem Prozent der Gesamtlebensdauer ermöglicht. Auch bei einem kombinierten Modell, gebildet aus Daten von mehreren baugleichen elektromechanischen Zylindern, können Merkmale identifiziert werden, welche charakteristische Informationen über den Verschleißfortschritt liefern.

Keywords: Zustandsüberwachung, vorausschauende Instandhaltung, maschinelles Lernen, Big Data, Industrie 4.0

Einführung und Motivation

Mit steigender Leistungs- und Funktionsdichte technischer Systeme verknüpfen sich besondere Anforderungen bezüglich Verfügbarkeit, Wirtschaftlichkeit und Sicherheit. Diese Anforderungen können durch klassische Instandhaltungsstrategien, wie präventive Wartungen oder starre Inspektionsintervalle zunehmend nicht mehr erfüllt werden. Dadurch könne hohe Kosten infolge von Produktionsausfällen oder nicht vollständig ausgenutzten Abnutzungsvorräten entstehen. Für Anlagenbetreiber werden deshalb Strategien der vorausschauenden, zustandsbasierten Instandhaltung von immer höherem Interesse [1].

Die Umsetzung dieser neuen Strategien basiert auf der Entwicklung zuverlässiger Diagnosesysteme, welche Aussagen über den Betriebszustand einzelner Komponenten komplexer Maschinen erlauben. Neben der geeigneten Sensorik zur Messung maschinenspezifischer Parameter bestehen solche Systeme aus Methoden und Algorithmen zur Bewertung von Schädigungs- und Degradationsmechanismen. In diesem Zusammenhang kommen neuerdings bevorzugt Verfahren der signalanalytischen und modellbasierten Zustandsdiagnose aus dem Bereich des maschinellen Lernens zur Anwendung. Diese sollen eine Prognose über die verbleibende Restlaufzeit erlauben, wodurch Ausfälle prognostiziert und Instandhaltungsmaßnahmen in ohnehin vorhandene Stillstandszeiten verlegt werden können [2].

Die Entwicklung eines solches Sensorsystems für die Echtzeit-Prozesssteuerung und smarte Zustandsbewertung für die Industrie 4.0 war Inhalt des BMBF Projekts MoSeS-Pro. Basis dieses Systems ist ein modularer Hard- und Software-Baukasten, welcher je nach Anforderungen der spezifischen Anwendung verschiedenste Kombinationen von Sensoren und Algorithmen gestattet. Die dabei umgesetzte sensornahe Signalverarbeitung vereinfacht eine direkte Integration in industrielle IT-Infrastrukturen durch Übermittlung weniger, hoch verschleißrelevanter Merkmale. Der Einsatz automatisierter, maschineller Lernalgorithmen erlaubt gleichzeitig eine schnelle Anpassung an verschiedene Anwendungsszenarien. Im Rahmen des Projektes wurde die Leistungsfähigkeit der Algorithmen am Beispiel der Zustandsüberwachung von elektromechani-

schen Zylindern demonstriert. Anwendung finden diese hauptsächlich in Positionier-, Montage- und Handhabungssystemen in der Automatisierungstechnik, wobei durch die Kombination von hohen Vorschubkräften und hoher Positioniergenauigkeit auch ein Einsatz in umformenden Prozessen (z. B. Pressaufgaben) möglich ist [3].

Lebensdauerprüfstand

Um die für das maschinelle Lernen benötigte Datenbasis zu generieren, wurde ein Lebensdauerprüfstand entwickelt, welcher das Verfahren des elektromechanischen Zylinders unter kontrollierten Bedingungen ermöglicht. In Abb. 1 a) ist der Prüfstand dargestellt. Hierbei können Verfahrprofile anhand verschiedener Parameter, wie Geschwindigkeit, Beschleunigung oder zu leistender Hub über eine LabVIEW-Software konfiguriert werden. Gleichzeitig besteht die Möglichkeit, die Kolbenstange der Spindelachse mit Kräften in Vorschubrichtung zu belasten, wodurch Einfluss auf die Ausbreitungsgeschwindigkeit der Verschleißmechanismen genommen werden kann. Die axiale Belastung des Zylinders ist über einen Pneumatikzylinder realisiert, welcher, wie in Abb. 1 b) gezeigt, mit der Kolbenstange des elektromechanischen Zylinders gekoppelt ist.

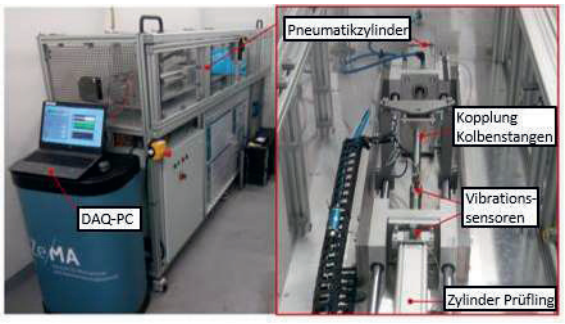


Abb. 1: Gesamtansicht Lebensdauerprüfstand für elektromechanische Zylinder (links), Kopplung von Prüfling und variabler Last in Form eines Pneumatikzylinders (rechts) [4].

Prüfstandssensorik

Um den Verschleißfortschritt des Prüflings über den Lebensdauerversuch zu erfassen, werden einerseits zu Regelzwecken ohnehin vorhandene Sensoren sowie antriebsbasierte Signale des Motorcontrollers verwendet, andererseits wird auf dedizierte Sensoren in Orientierung an die Wälzlagerdiagnose zurückgegriffen. Dabei erfolgt zunächst eine bewusste Überinstrumentierung in Bezug auf die Anzahl und Abtastrate der Sensoren, um den Informationsgehalt der verschiedenen Signale im Hinblick auf fortschreitende Verschleißmechanismen an der Spindelachse zu identifizieren. In der praktischen Anwendung muss dann eine Reduzierung der verwendeten Sensorik auf ein wirtschaftliches Maß stattfinden.

Bei den verwendeten Körperschallsensoren handelt es sich um uniaxiale IEPE (Integrated Electronic Piezo Electric) Beschleunigungsaufnehmer (Kistler) mit einer Bandbreite von 24 kHz. Um einen möglichst steifen Körperschallpfad zwischen Schwingungsursache und Sensor zu erreichen, werden die Beschleunigungsaufnehmer nah an ausfallrelevanten Komponenten des elektromechanischen Zylinders, also Kugellager, Gleitlager und Kugelgewindetrieb befestigt, Abb. 2. Da kein Sensor direkt an den Kugelgewindetrieb gekoppelt werden kann, wird dieser an einem Adapterstück an der Kolbenstange angebracht, welche durch eine Schraubverbindung im Inneren des elektromechanischen Zylinders eine starre Verbindung zum Kugelgewindetrieb aufweist.

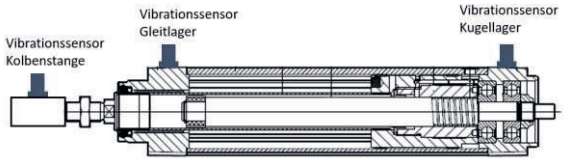


Abb. 2: Positionierung der Beschleunigungssensoren am Prüfling [3]

Neben den Beschleunigungssensoren wird zusätzlich ein Ultraschallmikrofon (G.R.A.S) eingesetzt, um höherfrequente Anteile im Luftschall bis 100 kHz zu erfassen.

Um antriebsbasierte Signale, wie z. B. die Motorströme auf ihre Tauglichkeit zur Zustandsbewertung von elektromechanischen Zylindern zu untersuchen, werden die drei Phasen des Antriebsmotors mittels externer Stromsensoren des Typs CMS205 (Sensitec) gemessen.

Tab. 1 gibt eine Übersicht über alle zur Verfügung stehenden Sensorsignale sowie die dazugehörigen Abtastraten der Datenerfassung.

Tab. 1: Prüfstandssensorik und zugehörige Abtastraten der Datenerfassung

Sensorart	Messgröße	Abtastrate
Dedizierte Sensorik	Beschleunigung Gleitlager	100 kHz
	Beschleunigung Kugellager	
	Beschleunigung Kolbenstange	
	Luftschall	
Antriebsbasierte Signale	Motorstrom (U, V, W)	1 MHz
	Verfahr-geschwindigkeit	10 kHz
	Wirkstrom	10 kHz
Prozesssensorik	Axialkraft	10 kHz
	Druck Pneumatik-zylinder	10 kHz

Lebensdauerversuche

Die im Rahmen des Projektes durchgeführten Lebensdauerversuche, dienen der realitätsnahen Nachbildung von Schadens- und Verschleißbildern elektromechanischer Zylinder über ihre Lebensdauer. Der Ablauf der Versuche orientiert sich an einer für elektromechanische Zylinder typischen zyklischen Arbeitsweise. In Abb. 3 ist exemplarisch ein Arbeitszyklus des elektromechanischen Zylinders anhand der Verfahrgeschwindigkeit dargestellt.

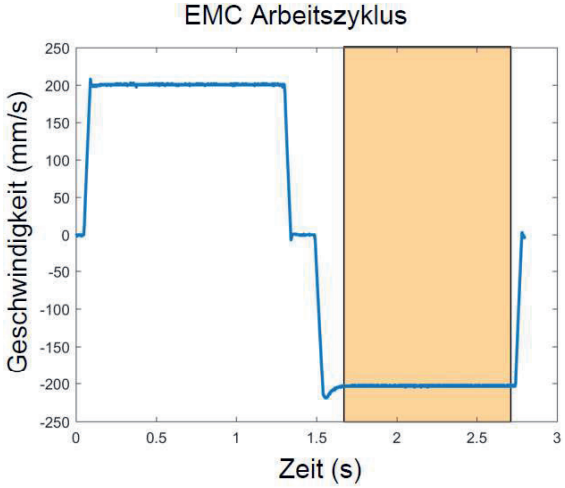


Abb. 3: Arbeitszyklus des elektromechanischen Zylinders dargestellt anhand der Verfahrgeschwindigkeit; markiert ist die Gleichlaufphase des Rückhubs, die für die Zustandsbewertung ausgewertet wird.

Während der Versuche sind die Prüflinge durch sie Kombination von hohen Verfahrgeschwindigkeiten (200 mm/s), Beschleunigungen (5 m/s²) und Axialkräften (ca. 7 kN) im Bereich der Spezifikationsgrenzen hohen Belastungen ausgesetzt. Dadurch werden ein möglichst schneller Verschleißfortschritt und ein daraus resultierender Ausfall provoziert. Ein Überschreiten von festgelegten Grenzen, z. B. der Schleppfehlergrenze bei einer durchgeführten Lageregelung, stellt das Ausfallkriterium für die Spindelachsen dar. Wird ein solcher Grenzwert während des Betriebs überschritten, wird der Versuch automatisch gestoppt und die Spindelachse gilt als defekt. Um eine bestmögliche Vergleichbarkeit der Versuche zu erreichen, wurden die hier betrachteten Lebensdauerversuche unter identischen Bedingungen und Konfigurationen durchgeführt. Tab. 2 gibt eine Übersicht über die Ergebnisse der hier betrachteten Lebensdauerversuche anhand der Laufzeit und der erreichten Arbeitszyklen. Durch die starke Überinstrumentierung des Prüfstandes fallen enorme Datenmengen für die Lebensdauerversuche an, was sich in einem Datenvolumen von ca. 12 TB pro Lebensdauerversuch äußert.

Tab. 2: Vergleich der betrachteten Lebensdauerversuche anhand Laufzeit und erreichter Arbeitszyklen

Lebensdauerversuch Nr.	Laufzeit [Tage]	Anzahl Arbeitszyklen
1	20,4	629.151
2	19.6	603.627
3	18,6	572.750

Modularer Softwarebaukasten

Die Auswertung bzw. Modellbildung erfolgt mittels eines am Zentrum für Mechatronik und Automatisierungstechnik entwickelten vollautomatisierten Methodenbaukastens zum überwachten maschinellen Lernen. Dieser kombiniert fünf Methoden der Merkmalsextraktion (ME) mit drei Ansätzen zur Merkmalsselektion (MS). Durch anschließende Dimensionsreduktion und Klassifikation wird mittels 10-fach Kreuzvalidierung der minimale Klassifikationsfehler für sämtliche Kombinationen bestimmt und die bestmöglichen Methoden für den jeweiligen Datensatz identifiziert [5].
Um ein nahezu ideales Klassifikationsergebnis unabhängig von der Herkunft des Datensatzes zu erreichen, wurden komplementäre Methoden für die ME ausgewählt. Ziel der ME ist es dabei stets, das Rohsignal mit möglichst wenigen Merkmalen und minimalem Rekonstruktionsfehler abzubilden. Welche Methode dabei erfolgreich ist, hängt stark von der Herkunft des betrachteten Signals ab. So bildet beispielsweise die Adaptive Linear Approximation (ALA) lokale Details in Zeitreihen ab, wohingegen periodische Sensorsignale häufig besser durch Beste Fourier Koeffizienten (BFC) im Frequenzbereich repräsentiert werden. Genauere Information zu den implementierten ME-Methoden können [5] entnommen werden.
Neben der möglichst generalisierbaren Anwendbarkeit wurde bei der Auswahl der Methoden für die ME vor allem darauf geachtet, dass sie später sensornah auf einem Mikroprozessor bzw. einer FPGA implementiert werden können. Dadurch entfällt bei der Anwendung der Algorithmen in der Zustandsüberwachung die Notwendigkeit, umfangreiche Rohdaten auf höhere Systemebenen der Prozesssteuerung transferieren zu müssen. Stattdessen werden nur noch wenige charakteristische, in der Trainingsphase identifizierte Merkmale übertragen.
Auch bei der MS kommen komplementäre Algorithmen zum Einsatz. Neben der Pearson Korrelation werden recursive feature elemination support vector machines (RFESVM) sowie RELIEFF für die MS eingesetzt. RFESVM

eignet sich vor allem bei linearer Trennbarkeit der Klassen, wobei auch die Merkmalsinteraktion berücksichtigt wird. Im Gegensatz dazu basiert RELIEFF auf einer radialen Klassifikation, was sie für nichtlineare Klassifikationsprobleme nutzbar macht. Auch zu den MS Algorithmen können weiterführende Informationen aus [5] entnommen werden.

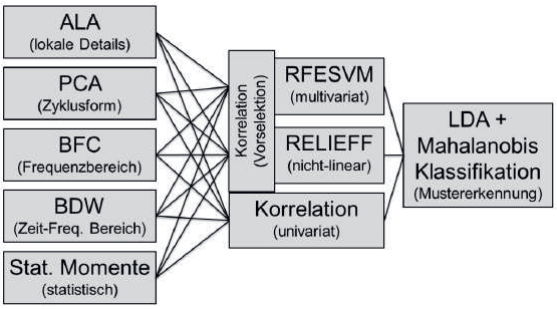


Abb. 4: Vollautomatisierter Methodenbaukasten mit komplementären Methoden für Merkmalsextraktion (links), -selektion (Mitte) und Klassifikation (rechts)

Die abschließende Dimensionsreduktion und Klassifikation erfolgt mittels Linearer Diskriminanzanalyse (LDA) in Kombination mit einem Mahalanobis-Distanz Klassifikator. Durch den Einsatz vergleichsweise einfacher, linearer Verfahren im Schritt der Mustererkennung bleibt im Gegensatz zu Blackbox-Verfahren, wie z. B. neuronalen Netzen, die Möglichkeit erhalten, eine spätere physikalische Plausibilisierung des Modells und der selektierten Merkmale durchzuführen.

Zustandsbewertung für einzelne elektromechanische Zylinder

Im Folgenden werden die Ergebnisse der Zustandsbewertung einzelner elektromechanischer Zylinder exemplarisch anhand des ersten Lebensdauerversuchs vorgestellt. Die Daten des Lebensdauerversuchs werden dazu in 1% Schritte der Gesamtlebensdauer des elektromechanischen Zylinders unterteilt und mittels des beschriebenen Methodenbaukastens ausgewertet, wobei jeweils nur die Gleichlaufphase des Rückhubs, wie in Abb. 3 dargestellt, verwendet wird. In Abb. 5 sind die kreuzvalidierten Klassifikationsfehler der einzelnen FE-Methoden nach Merkmalsselektion mittels Pearson-Korrelation über die Anzahl der zur Modellbildung verwendeten Merkmale dargestellt. Ausgewertet werden dabei die Einzelzyklen, was einer Bewertung des Zustandes des elektromechanischen Zylinders alle 2,8 Sekunden, entsprechend der Dauer des konfigurierten Arbeitszyklus, entspricht.

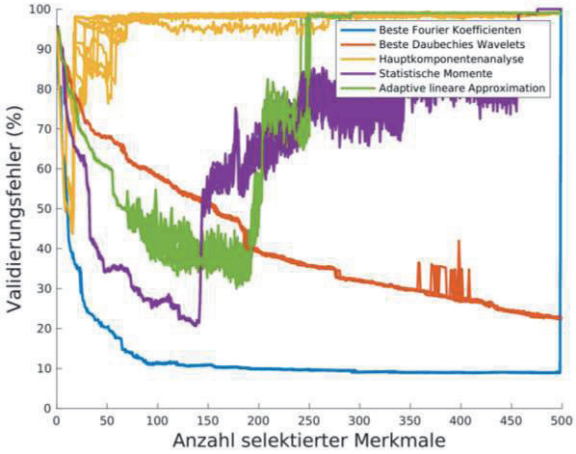


Abb. 5: Kreuzvalidierte Klassifikationsergebnisse der einzelnen ME-Methoden; Merkmalsselektion mittels Pearson-Korrelation, Klassifikation mittels LDA und Mahalanobis-Distanz

Der minimale erreichte Klassifikationsfehler beträgt 9% unter Verwendung von 499 Merkmalen. Erreicht wird dies über die Extraktion der besten Fourierkoeffizienten (BFC), wobei Frequenzen mit dem höchsten Mittelwert über alle Arbeitszyklen ausgewertet werden [6]. Diese beinhalten die meiste Signalenergie und tragen deshalb am meisten zur Minimierung des Rekonstruktionsfehlers bei.
Im Sinne einer vorausschauenden Instandhaltung wird eine Abschätzung der verbleibenden Restlaufzeit des elektromechanischen Zylinders alle 2,8 Sekunden nicht benötigt. Abb. 6 zeigt aus diesem Grund die 10-fach kreuzvalidierte Vorhersage bei einer Zustandsbewertung durch einen Mehrheitsentscheid über die letzten 15 Minuten. Dadurch kann der verbleibende Klassifikationsfehler auf 5% reduziert werden, wobei der Großteil der Fehlklassifikationen an den Quantisierungsstufen auftritt, was einer fehlerhaften Zustandsbewertung von ±1% der Gesamtlaufleistung des elektromechanischen Zylinders entspricht. Lediglich am Ende der Lebensdauer erfolgt eine zu frühe Prognose des Ausfalls, was auch durch den Mehrheitsentscheid nicht korrigiert werden kann.
Wie bereits erläutert, erlauben die hier verwendeten Methoden eine Analyse der zur Modellbildung verwendeten Merkmale. In Abb. 7 ist die Herkunft der 499 besten Merkmale nach den zugrundeliegenden Sensoren aufgeschlüsselt. Hier wird deutlich, dass die bei der Wälzlagerdiagnose bereits etablierten Sensoren auch bei der Zustandsüberwachung von elektromechanischen Zylindern den größten Beitrag liefern. Neben Merkmalen, die aus dem Signal des Ultraschallmikrofons gewonnen werden, liefern vor allem die Beschleunigungs-

sensoren hoch mit dem Verschleiß korrelierte Merkmale. Die Analyse dieser Merkmale zeigt, dass ebenfalls eine physikalische Interpretation und Plausibilisierung dieser Merkmale erfolgen kann. So stammt ein Großteil der selektierten Merkmale aus Frequenzen, welche mechanische Antriebsfrequenzen, Überrollfrequenzen oder deren Harmonische als Ursache haben.

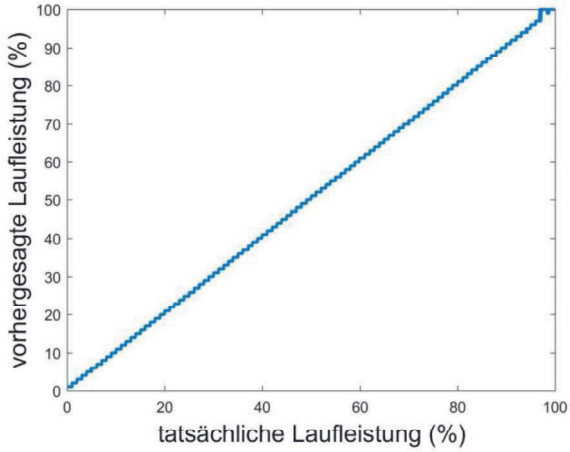


Abb. 6: Lebensdauervorhersage des elektromechanischen Zylinders durch Mehrheitsentscheid über die letzten 15 Minuten

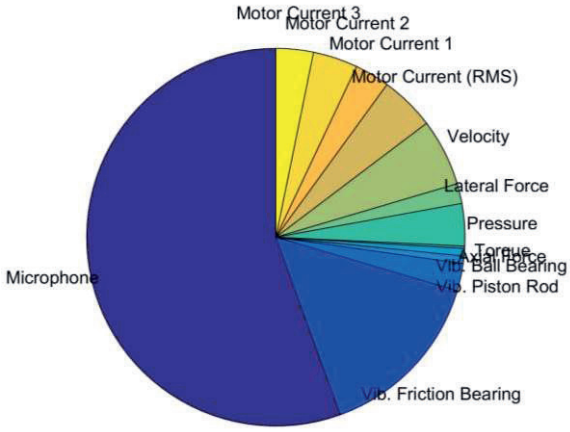


Abb. 7: Herkunft der zur Modellbildung verwendeten Merkmale aufgeschlüsselt nach den Sensoren

Eine weitere Art der Plausibilisierung des Modells ist in Abb. 8 dargestellt. Hier erfolgt die Modellbildung in 20% Schritten (Trainingsdaten 1%, 21%, 41%, 61%, 81%) der Gesamtlaufleistung des elektromechanischen Zylinders. Verwendet wurden dabei die 60 höchstkorrelierten Merkmale. Anschließend werden Testdaten (10%, 30%, 50%, 70%, 90%) projiziert. Diese liegen, wie in Abb. 8 gezeigt, an den zu erwartenden Positionen zwischen den Trainingsdaten, obwohl die Reihenfolge der Klassen nicht als Eingangsinformation bei der Modellbildung vorgegeben wurde. Diese Art der Gruppen-basierten Validierung zeigt deutlicher als die übliche Kreuzvalidierung das Potential des Modells zur korrekten Interpolation von Verschleißdaten [6].

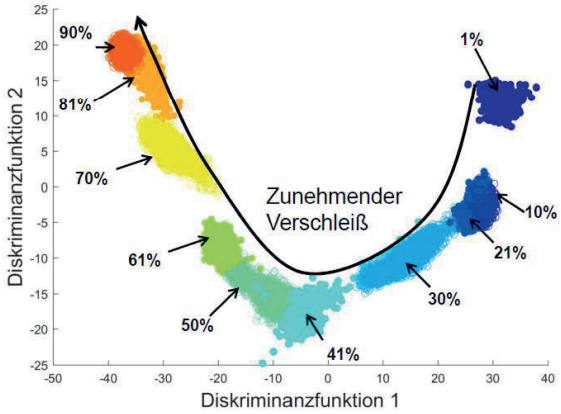


Abb. 8: Verlauf der LDA-Projektion bei zunehmendem Verschleiß des elektromechanischen Zylinders, bei Bildung des Modells in 20% Schritten der Lebensdauer und Projektion von Testdaten

Zustandsbewertung einzelner elektromechanischer Zylinder auf Basis des Motorstroms

Um das Potential rein antriebsbasierter Signale für die Zustandsüberwachung elektromechanischer Zylinder zu überprüfen, wurde eine Auswertung auf Basis der externen Stromsensoren durchgeführt. Im Vergleich zur restlichen Sensorik wird der Motorstrom zu Regelzwecken im Motorcontroller ohnehin erfasst, was die Integration einer Zustandsüberwachung in den Motorcontroller ohne großen Mehraufwand ermöglichen würde. Abb. 9 zeigt wieder die 10-fach kreuzvalidierte Vorhersage durch Mehrheitsentscheid über die letzten 15 Minuten rein auf Basis der Stromsensoren.

Der verbleibende Klassifikationsfehler beläuft sich hierbei auf 6%, was eine nur minimale Verschlechterung der Vorhersagegenauigkeit des Modells im Vergleich zur gesamten Prüfstandssensorik mit Beschleunigungssensoren und Mikrofon bedeutet. Auch hier treten neben dem zu früh vorhergesagten Ausfall gegen Ende der Lebensdauer die Fehlklassifikationen hauptsächlich an den Quantisierungsstufen auf. Die selektierten Merkmale stammen wieder vorwiegend aus Frequenzen mit mechanischen Ursachen, welche sich durch den gesamten Antriebsstrang vom elektromechanischen Zylinder bis hin zum Stator des Antriebsmotors fortsetzen und somit von den Stromsensoren erfasst werden können [7].

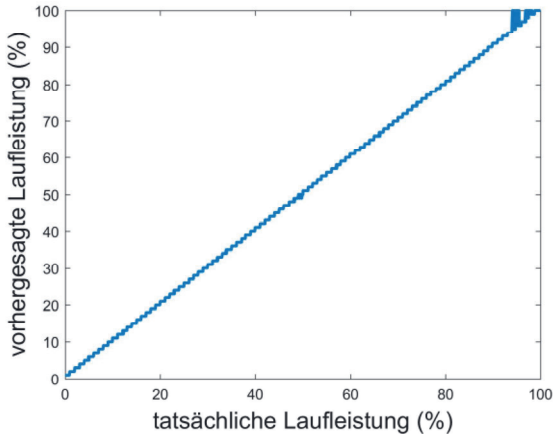


Abb. 9: Lebensdauervorhersage des elektromechanischen Zylinders durch Mehrheitsentscheid über die letzten 15 Minuten nur auf Basis der Stromsensoren

Kombiniertes Modell baugleicher elektromechanischer Zylinder auf Basis des Motorstroms

Um die gebildeten datengetriebenen Modelle im Rahmen einer Lebensdauerprognose nutzbar zu machen, muss eine Übertragbarkeit der Modelle zwischen baugleichen elektromechanischen Zylindern gegeben sein. Im ersten Schritt muss dazu ein Modell bzw. Merkmale gefunden werden, welche zylinderübergreifend charakteristisch den Verschleiß über die Lebensdauer wiederspiegeln. Um diesen Sachverhalt zu untersuchen, wurde ein gemeinsames Modell aus den Motorstromsignalen der drei in Tab. 2 gezeigten Lebensdauerversuche gebildet. Bei der Analyse der Daten hat sich gezeigt, dass die Einlaufphasen nach Inbetriebnahme sowie die Signale kurz vor dem Ausfall der Achse sehr spezifisch für die einzelnen Zylinder sind. Aus diesem Grund wurden für die Bildung des kombinierten Modells die Einlaufphasen (jeweils erste 10% der Lebensdauer) sowie die Ausfallphasen (jeweils letzte 4%) der einzelnen Zylinder bei der Modellbildung herausgelassen. In Abb. 10 ist die 10-fach kreuzvalidierte Vorhersage für das kombinierte Modell auf Basis der Stromsensorsignale für die drei Lebensdauerversuche dargestellt. Wie hier zu sehen ist, kann ein konsistentes Modell für alle drei Achsen bestimmt werden, wobei die Vorhersagequalität für die einzelnen Zylinder durchaus unterschiedlich ausfällt: am besten wird die Restlebensdauer für den 2. Versuch vorhergesagt.

Diskussion und Ausblick

Im Rahmen des Projektes MoSeS-Pro wurde eine flexibel einsetzbare modulare Softwaretoolbox für das automatisierte maschinelle Lernen entwickelt, welche neben dem hier gezeigten Anwendungsfall schon auf einer Vielzahl an Datensätzen erfolgreich getestet wurde [8]. Die Implementierung dieser Algorithmen auf einer ebenfalls in MoSeS-Pro entwickelten modularen Hardwarebasis erlaubt eine sensornahe Signalverarbeitung, was eine drastische Reduzierung der zu übertragenden Datenmengen bedeutet, da nur noch einzelne aussagekräftige Merkmale in höhere Systemebenen übertragen werden müssen. Die infolgedessen deutlich geringeren benötigten Datenraten ermöglichen z. B. die Vernetzung der Sensoren mittels Wireless-Schnittstellen.
Im Falle der Zustandsbewertung für elektromechanische Zylinder konnte für einzelne Zylinder eine hochaufgelöste Verschleißdetektion bzw. Vorhersage der Restlebensdauer erreicht werden. Vor allem eine auf den Stromsignalen des Antriebsmotors basierende Modellbildung erscheint vielversprechend, da im Gegensatz zur verwendeten externen Sensorik keine zusätzliche Sensorik, z. B. ein Ultraschallmikrofon, in das System integriert werden muss.

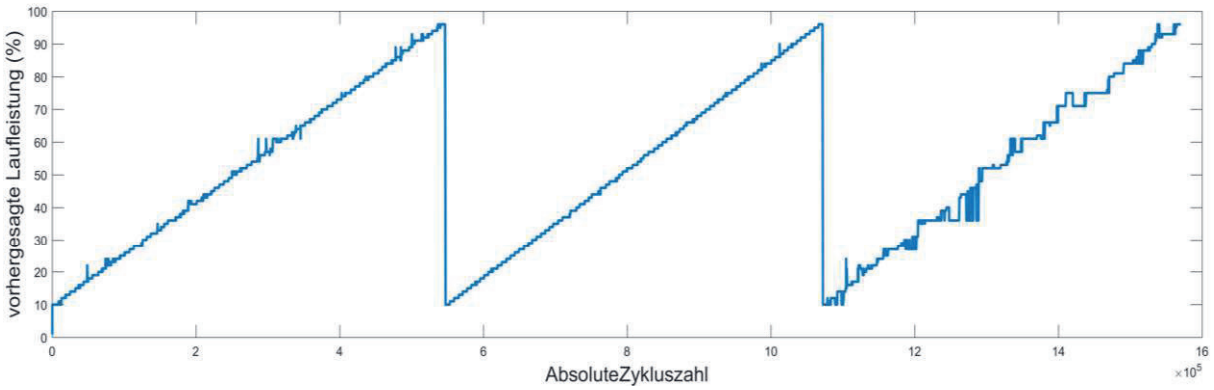


Abb. 10: Lebensdauervorhersage des kombinierten Modells der drei Lebensdauerversuche durch Mehrheitsentscheid über die letzten 15 Minuten auf Basis der Stromsensoren

Um die vorgestellten datengetriebenen Modelle für die Zustandsbewertung nutzbar zu machen, ist vor allem im Bereich der Übertragbarkeit noch Arbeit zu leisten. Wie gezeigt, ist die Bildung eines konsistenten, zylinderübergreifenden Modells möglich, allerdings auf Kosten der Vorhersagegenauigkeit für die einzelnen Zylinder. Für eine praktische Zustandsbewertung müsste allerdings eine generelle Übertragbarkeit der Modelle gegeben sein. Hierfür würde in erster Linie eine breitere Datenbasis benötigt, um Fertigungstoleranzen individueller elektromechanischer Zylinder zu unterdrücken. Des Weiteren müssten Experimente mit anderen Arbeitszyklen durchgeführt werden, um die Auswirkungen z. B. von veränderten Last- oder Geschwindigkeitsprofilen festzustellen.
Neben dem hier präsentierten Methodenbaukasten für das überwachte maschinelle Lernen müssen für eine zuverlässige Zustandsbewertung Methoden der Novelty Detection bzw. des unüberwachten maschinellen Lernens ergänzt werden. Dadurch sollen in der Praxis unbekannte System(fehler)zustände detektiert werden, welche nach Identifikation ihrer Ursache nach und nach in das Zustandsüberwachungsmodell eingelernt werden.

Danksagungen

Die Arbeiten wurden mit Mitteln des Bundesministeriums für Bildung und Forschung im Rahmen des Verbundprojekts MoSeS-Pro unter dem Förderkennzeichen 16ES0419K gefördert. Die elektromechanischen Zylinder sowie die Komponenten des Prüfstands wurden von der Firma Festo im Rahmen der Zusammenarbeit zur Verfügung gestellt.

Literaturnachweis

[1] VDMA, Am Puls der Maschine: Condition Monitoring, VDMA Verlag, 2015.
[2] M. Walther, Antriebsbasierte Zustandsdiagnose von Vorschubantrieben, Dissertation, Universität Stuttgart, Fakultät für Maschinenbau, Stuttgart, 2011
[3] Festo AG & Co KG, Datenblatt ESBF elektromechanischer Zylinder, Online verfügbar: https://www.festo.com/cat/en-gb_gb/data/doc_DE/PDF/DE/ESBF_DE.PDF [Zugriff am 08.01.2019].
[4] T. Schneider, N. Helwig, S. Klein, A. Schütze, Influence of sensor network sampling rate on multivariate statistical condition monitoring of industrial machines and processes, Eurosensors 2018, Volume 2, doi: 10.3390/proceedings2130781
[5] T. Schneider, N. Helwig, A. Schütze, Industrial condition monitoring with smart sensors using automated feature extraction and selection, Measurement Science and Technology, Volume 29, Number 9 (2018)
[6] S. Youssef, C. Zimmer, K. Szielasko, A. Schütze, Bewertung subjektiver und automatisierter Merkmalsextraktion periodischer Zeitsignale am Beispiel des 3MA-X8-Verfahrens, tm - Technisches Messen, Online erschienen: 18.01.2019, doi: 10.1515/teme-2018-0074.
[7] D. Miljković, Brief Review of Motor Current Signature Analysis, HDKBR INFO Magazin, vol.5, no. 1, pp. 14-26, 2015. [Online]. Available: https://hrcak.srce.hr/148715. [Accessed: 07 May 2019]
[8] T. Schneider, N. Helwig, A. Schütze, Automatic feature extraction and selection for condition monitoring and related datasets, IEEE Xplore 218, doi: 10.1109/I2MTC.2018.8409763

DOI 10.5162/sensoren2019/4.4.2

Robuste xMR-Sensoren für die Automatisierungstechnik

Jörg Traute[1], Dr. Frederick Casper[2], Dr. Johannes Paul[2]
[1]Sensitec GmbH, Lahnau, Deutschland
[2]Sensitec GmbH, Mainz, Deutschland

Zusammenfassung

Im Rahmen des vom BMBF geförderten Verbundprojekts MoSeS-Pro hat sich die Sensitec GmbH mit zwei Aufgaben beschäftigt. Zum einen wurde ein xMR-Sensor mit Self-X Funktionen entwickelt und hergestellt. Ähnlich wie Lebewesen sich selbst untersuchen und reparieren, kann dies auch auf MR-Sensoren übertragen werden. Im Projekt wurde ein TMR-Feldsensor entwickelt, der durch geeignetes Design und Zusatzfunktionen in Verbindung mit einem Steuerungs-Chip (z. B. einem ASIC) über Self-X Funktionen verfügt. Weiter wurde ein Spindelachsenprüfstand mit insgesamt vier verschiedenen Sensortypen zur Zustandsüberwachung ausgestattet. Damit konnten an den elektromechanischen Spindelachsen Motorströme sowie rotative und lineare Bewegungen der Kolbenstange gemessen werden. Zudem konnte eine Kugeldetektion direkt am Kugelgewindetrieb der Spindelachse umgesetzt werden. Die MR-Sensoren zeigten dabei durch die hohe Bandbreite und Ortsauflösung auch höhere harmonische Signalanteile, die wichtige Informationen für die Zustandsüberwachung der Achse liefern.

Keywords: xMR-Sensor, Self-X Funktion, Zustandsüberwachung, Kugeldetektion

1. Einleitung

Im Zukunftsprojekt Industrie 4.0 bietet sich heute die Chance, über eine intelligente Steuerung und Vernetzung die Flexibilität, die Energie- und die Ressourceneffizienz von Produktionsprozessen auf eine neue Stufe zu heben. Elektronik und Sensorik, die zu den Stärken gerade auch kleiner und mittlerer Unternehmen (KMU) in Deutschland zählen, spielen dabei eine Schlüsselrolle. Eingebaut in Antriebe und Positionierungssysteme ermöglichen sie schnelle und präzise Fertigungsprozesse, deren Ablauf in Echtzeit überwacht und angepasst werden kann. Die Sensorsysteme müssen hierzu jedoch auf die genauen Anforderungen der Anwendung hin entworfen und aufgebaut werden [1]. Die Sensitec GmbH hat beim BMBF Projekt MoSeS-Pro zwei Hauptaufgaben übernommen. Zum einen wurde ein xMR-Sensor mit Self-X Funktionen entwickelt und hergestellt, zum anderen wurde ein Prüfstand für elektromechanische Spindelachsen mit verschiedenen xMR-Sensoren zur Zustandsüberwachung ausgestattet. xMR steht als Platzhalter für die verschiedenen magnetoresistiven (MR) Effekte wie anisotroper MR-Effekt (AMR), Giant Magneto Resistance (GMR) und Tunneling Magneto Resistance (TMR), die derzeit für Sensoranwendungen industriell genutzten werden. Tab. 1 und Abb.1 geben einen Überblick.

Tab. 1: Industrielle XMR-Technologien im Vergleich.

	AMR	GMR Spin Valve	TMR
Signal-stärke	< 4 %	< 15 %	< 200 %
Strom-richtung	parallel zu aktiver Schicht	parallel zu aktiver Schicht	senkrecht zu aktiver Schicht

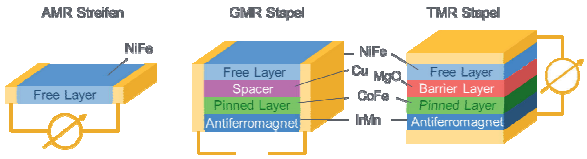


Abb. 1: xMR-Technologien in vereinfachter Darstellung.

Bei magnetoresistiven Sensoren ändert sich der Widerstand linear zum angelegten Feld.

2. Self-X Funktionen

Bei MR Sensoren ist das Ausgangssignal temperaturabhängig, d.h. es sinkt mit steigender Temperatur. Dieser Temperatureffekt ist nicht von einem Feldeffekt zu

unterscheiden, weswegen entweder mit einem zusätzlichen Temperatursensor oder einer temperatur-unabhängigen Kalibration gearbeitet werden muss. Außerdem könnte sich das Signal aufgrund von Langzeit-Drift verändern, was wieder die Messunsicherheit vergrößert. Durch das Anlegen eines bekannten Magnetfelds am Sensor könnte dieser in Echtzeit kalibriert werden und wäre dadurch unabhängig von oben genannten Einflüssen.

Sensordesign

Ein neues Design für TMR-Sensoren wurde entworfen, um ein definiertes Test-Feld am Sensor anlegen zu können. Dies wird realisiert über eine Kupferspule, die oberhalb der sensitiven Elemente verläuft. Ein definierter Strompuls führt dadurch zu einem definierten Feld am Sensor. Abbildung 2 zeigt das prinzipielle Design des Sensors.

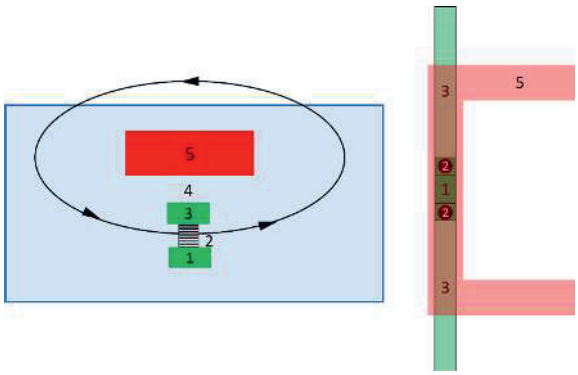


Abb. 2: Schematischer Blick von der Seite (links) und von oben (rechts) des Sensors. Der Sensor besteht aus einer unteren und oberen Elektrode (1 und 3), einem TMR Element (2) und einer durch ein dielektrisches Material (4) abgetrennten Stromspule (5). Die Pfeile geben die Richtung des Magnetfelds an, das durch die Spule erzeugt wird.

Hierbei wird eine Stromspule über die Sensorelemente platziert, wobei beide durch ein dielektrisches Material getrennt sind. Für die Spule werden auch zwei zusätzliche Versorgungspads benötigt. Das durch die Stromspule erzeugte magnetische Feld ist parallel zur Chipoberfläche. Im Falle einer Temperaturänderung ändert sich das durch die Stromspule erzeugte Feld nicht, da dieses nur vom Strom abhängt und dem Abstand zwischen der Spule und der Sensorbrücke. Die durch die Temperatur verursachte Abstandsänderung ist dabei in dem Temperaturbereich von -40°C - 125°C nicht signifikant weswegen dieser Effekt vernachlässigt werden kann (<0,01%).

Durch das Anlegen eines AC-Stroms mit bekannter Amplitude wird ein entsprechendes AC-Feld mit bekannter Feldamplitude generiert, dass dann zu einem Sensorsignal führt, das mit der Feldstärke korreliert werden kann. Dadurch kann der Sensor unabhängig von Temperatur und Langzeit-Drift sehr einfach kalibriert werden.
Zusätzlich kann über die Spule auch eine Self-Repair-Funktion installiert werden, wenn z.B. der Sensor nach einem Störfeldpuls in einen Hysteresezustand geht und ein falsches Signal liefert. Zum Testen dieser Funktion des Sensors wurde ein Versuchssetup aufgebaut, um folgende Messsequenz durchzuführen:

1. Sensor im hysteresefreien Zustand betreiben (B ≤ 20 mT)
2. Störfeld >> H_{exp} (>> 20 mT) anlegen
   ⇒ Sensor mit Hysterese
3. Störfeld wieder entfernen, weiter Feld um 10 -12 mT anlegen
   ⇒ Sensor immer noch mit Hysterese
4. Reset-Puls auf Sensorspule geben
   ⇒ Sensor nach Reset wieder im hysteresefreien Zustand

Abb. 3 zeigt den Verlauf des Sensorsignals nach den einzelnen Versuchsschritten. Dabei ist deutlich zu erkennen, dass der Sensor nach dem Störpuls ein deutlich anderes Signal liefert. Dieses wird jedoch durch den Resetpuls wieder auf den ursprünglichen Signalverlauf gebracht.

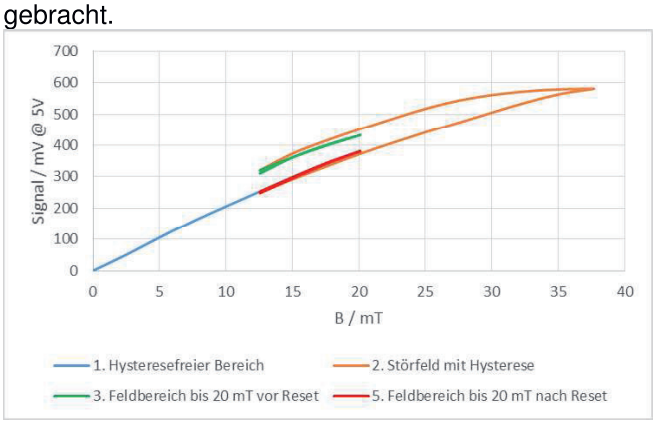


Abb. 3: Sensorsignal vor/ nach Störfeldpuls und Resetpuls

3. Zustandsüberwachung

Innerhalb des MoSeS-Pro Projekts wurde ein Prüfstand zur Zustandsbewertung von elektromechanischen Spindelachsen neben anderen mit verschiedenen MR-Sensoren von Sensitec ausgestattet. Hauptziel war es, dass der Projektpartner ZeMA Saarbrücken auf diesem Prüfstand unter beschleunigten Test-bedingungen Lebensdauer-Tests durchführen und dabei den modularen Sensorbaukasten erproben konnte [2].

Tab. 2: MR-Sensoren im Prüfstand für Spindelachsen

Nr.	Messgröße	Sensor
1	Motorstrom	CMS 2005
2	rotative Bewegung der Kolbenstange	AL780 mit Stützmagnet
3	lineare Bewegung der Kolbenstange	GL715
4	Kugel-Detektion am Kugelgewindetrieb	GL713

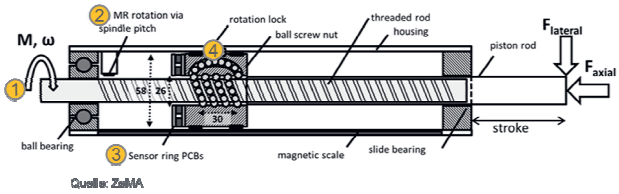


Abb. 4: vereinfachter Querschnitt einer elektromechanischen Spindelachse mit Messstellen für MR-Sensoren [2].

Motorstrom

Anforderungen an den Stromsensor waren Messung von 3 Phasen im Strombereich 0 - 20 A und Frequenzbereich 0 - 1000 Hz. Abb. 5 zeigt erste Messergebnisse im Zeitbereich mit dem MR-Stromsensor CMS2005. Dabei wurden auf eine im Prüfstand montierte Achse veränderliche axiale Lasten eingeprägt, die zu einer Änderung im Stromsignal führen (Bild 5a). Es zeigt sich, dass das Strommessprinzip mit MR Sensoren bedingt durch die hohe Bandbreite der Sensoren auch in der Lage ist, höhere harmonische Anteile im Signal abzubilden, die wichtige Informationen für die Zustandsüberwachung der Achse liefern. Dies ist auch in Abb. 6 zu sehen, die das Frequenzspektrum für eine Phase bei konstanter Geschwindigkeit zeigt.

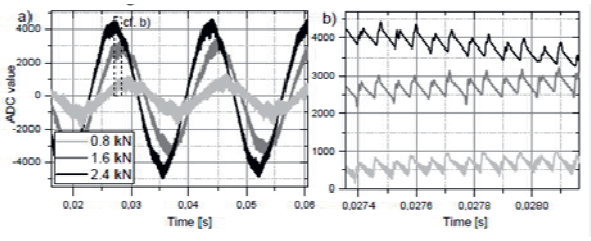


Abb. 5: a) Sinusförmiges Signal des Motorstroms für eine Phase bei verschiedenen eingeprägten axialen Lasten auf der Achse, b) Ausschnitt Stromsignal, kurzperiodische Anteile erzeugt durch Puls-Weiten-Modulation des Frequenzumrichters [2].

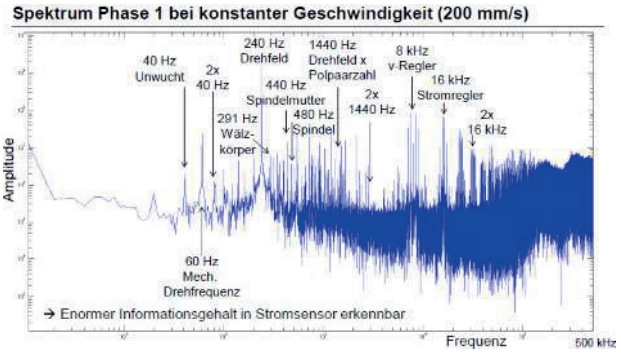


Abb. 6: Amplitudenspektrum des Motorstroms mit Zuordnung der mechanischen und elektrischen Frequenzen beim Betrieb der Spindelachse [3].

Rotative Bewegung der Kolbenstange

Das Konzept für diese Messung war, die bereits auf der Spindelachse vorhandene Zahnstruktur als Maßverkörperung zu nutzen. Die Zahnstruktur hat eine Teilung von 5 mm und eine Steigung von 10 mm, d. h. eine Umdrehung der Achse führt zu einer Auslenkung der Kolbenstange um 10 mm. Der ausgewählte polangepasste AMR-Längensensor AL780 wird normalerweise mit aktiven hartmagnetischen Maßstäben eingesetzt, auf die 5 mm lange Pole mit wechselnder Ausrichtung geschrieben sind. Kombiniert man den Sensor mit einem Hilfs- oder Stützmagnet, ist er auch in der Lage, weichmagnetische Zahnstrukturen abzutasten. Die Zähne modulieren das Magnetfeld des Hilfsmagneten, der Sensor wertet die durch die Zähne verursachten Feldänderungen aus. Maßverkörperungen dieser Art bezeichnet man als passive Maßstäbe.

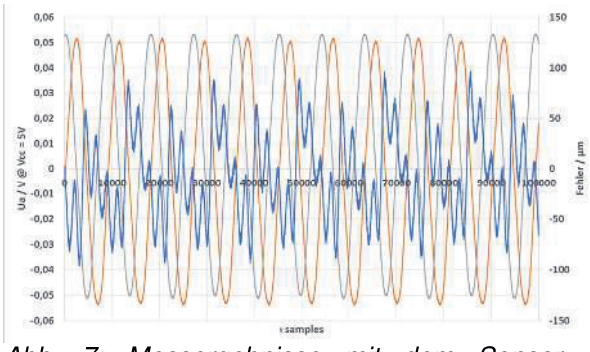


Abb. 7: Messergebnisse mit dem Sensor AL780 bei rotativer Bewegung.

Abb. 7 zeigt typische Messkurven bei einem Arbeitsabstand von 100 µm zwischen Sensor und Zahnstruktur. Die Sensorsignale haben eine Amplitude von 50 mV bei 5 V Versorgungsspannung, der Fehler liegt im Bereich ±100 µm.

Um den Sensor im Gehäuse der Achse zu integrieren, wurde ein Sensorträger konstruiert. Mit diesem Träger kann die Sensor-Position relativ zum Magnet und zur Spindelachse eingestellt werden, um die Sensor-Signale bezüglich Amplitude und Fehlerverhalten zu optimieren.

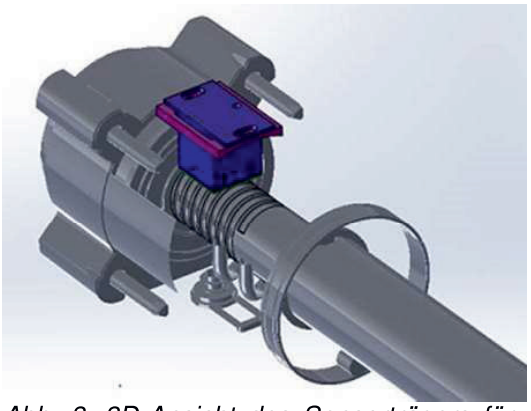


Abb. 8: 3D-Ansicht des Sensorträgers für die rotative Messung.

Lineare Bewegung der Kolbenstange

Zur Detektion der linearen Bewegung wurde eine ferromagnetische Zahnstange mit einer Zahnteilung von Modul 0,5 mm in eine vorhandene Nut des Spindelgehäuses eingesetzt. Der Zahnsensor GLM715 ist in einem SMD-Gehäuse aufgebaut und hat einen integrierten Stützmagnet zum Magnetisieren der Zahnstruktur und Vorhalten eines Stützfelds für den Sensor (Abb. 9). Der Sensor wurde beim Projektpartner ZeMA auf einer externen Sensorplatine in der elektro-mechanischen Spindelachse eingebaut (Abb. 10). Messergebnisse zeigt Abb. 11.

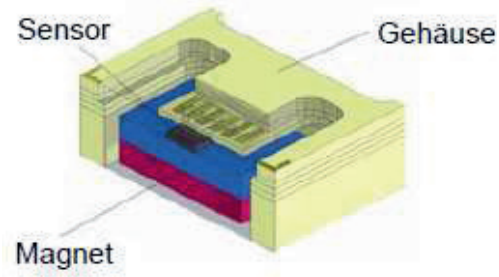


Abb. 9: Ansicht GMR-Zahnsensor GLM715 im SMD-Gehäuse mit Stützmagnet.

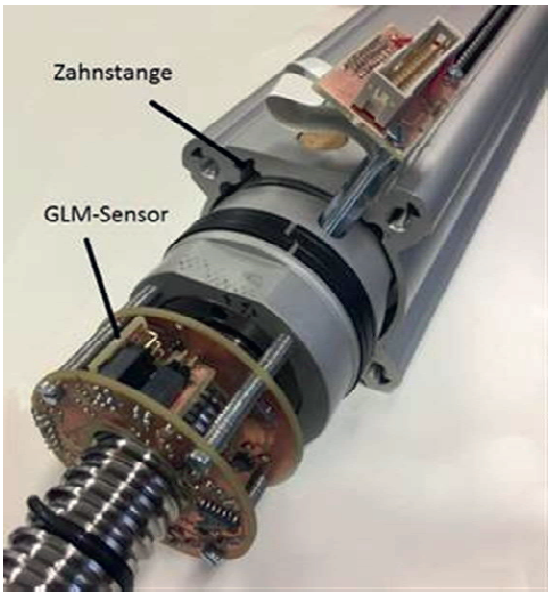


Abb. 10: Messung der linearen Bewegung mit Zahnsensor GLM715 und Zahnstange für Modul 0,5 mm [3].

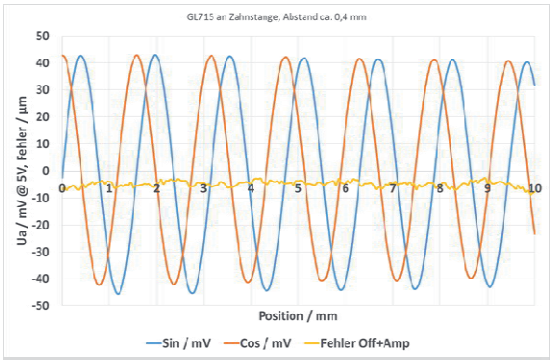


Abb. 11: Sensorsignale und Positionsfehler mit dem Zahnsensor GLM715 bei linearer Bewegung und einem Arbeitsabstand von 0,4 mm. Der Positionsfehler beträgt ± 10 µm.

Kugel-Detektion am Kugelgewindetrieb

Ein weiterer Bestandteil für die Zustands-überwachung ist der Kugelgewindetrieb (KGT).

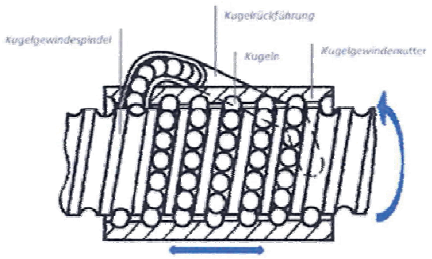


Abb. 10: Schnittbild eines Kugelgewindetriebs mit Umlenkrohr für die Kugelrückführung.

Am Umlenkrohr eines KGT sollen mit MR-Sensoren die durchlaufenden Kugeln erfasst werden nach Volumen (Form), Geschwindigkeit, Anzahl und Frequenz. Die Kugeln in der ausgewählten Spindel-Achse haben einen Durchmesser von 3,5 mm. Über Lebensdauer wird eine Verringerung des Durchmessers durch Verschleiß um 0,2 mm erwartet. Stand der Technik bei der Zustands-überwachung von KGT sind akustische Untersuchungen.

Idee zum Detektieren der Kugeln war, die im Umlenkrohr aufgereihten Kugeln als Zahnstruktur aufzufassen. Für den Test wurde ein GMR-Zahnsensor vom Typ GLM713 ausgewählt. Der Sensor ist ausgelegt für eine 3 mm Zahnteilung, was nah am Durchmesser der Kugeln liegt. Der Sensor ist genauso wie der bereits vorgestellte GLM715 in einem SMD-Gehäuse aufgebaut mit integriertem Stützmagnet.

Für einen ersten Test wurde der Sensor auf das Umlenkrohr eines KGT gesetzt (Abb. 11), die Achse bewegt und Messkurven aufgenommen. Die Signale liegen im Bereich 100 µV bei 5 V Sensor-Versorgungsspannung. Grund für die niedrige Signalstärke ist das Material und die Wandstärke des Umlenkrohrs. Es besteht aus magnetischem Werkzeugstahl, die Wandstärke liegt bei 0,5 mm. Damit schirmt das Rohr die Modulation des Magnetfelds durch die Kugeln stark ab, die Signale haben nur sehr kleine Amplituden (Abb. 12).

Nach Rücksprache mit dem Hersteller der KGT wurde die Wandstärke des Rohrs im Bereich, wo der Sensor platziert wird, von 0,5 mm auf 0,1 mm abgedünnt. Eine typische Messkurve mit einem bearbeiteten Umlenkrohr zeigt Abb. 13. Zu sehen ist ein Sensorsignal beim Durchgang einer Kugel. Die Signalamplitude ist mit ca. 20 mV bei 5 V Sensor-Versorgungsspannung um Faktor 200 höher im Vergleich zur Messung am KGT mit nicht bearbeitetem Umlenkrohr.

Die prinzipielle Eignung, Kugeln in der Spindelachse zu detektieren, wurde damit nachgewiesen. In weiteren Versuchen soll evaluiert werden, ob weitere Merkmale wie Änderungen bei Form und Durchmesser der Kugeln, Frequenz und Geschwindigkeit erkannt werden können.

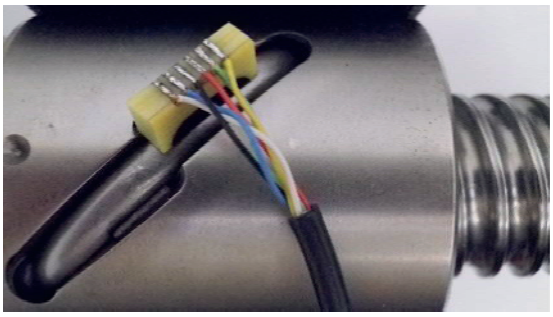


Abb. 11: GMR Zahnsensor GLM713 am Umlenkrohr des Kugelgewindetriebs.

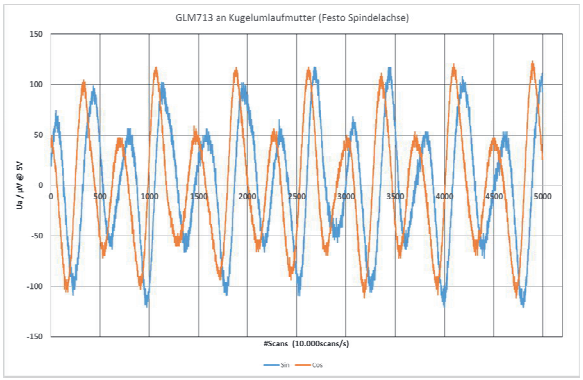


Abb. 12: Messkurven mit GLM713 am unbearbeiteten Umlenkrohr. Signalamplitude 100 µV bei Ub = 5 V.

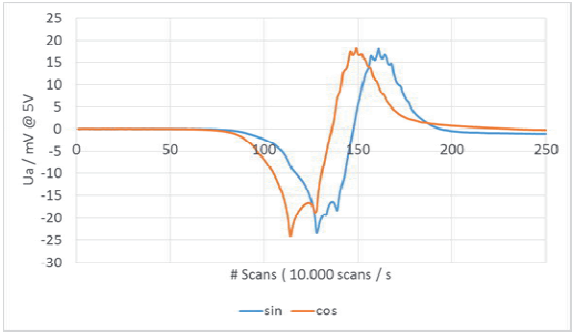


Abb. 13: Messkurven mit GLM713, Wandstärke des Umlenkrohrs auf 0,1 mm abgedünnt. Signalamplitude ca. 20 mV bei Ub = 5 V.

Abb. 14 und 15 zeigen Messungen mit dem Zahnsensor GLM713 am Kugelumlenkrohr, die auf dem Spindelachsenprüfstand von ZeMA aufgenommen wurden. Beim Vergleich der beiden Bewegungsrichtungen im Frequenzbereich in Abb. 15 sind Unterschiede zwischen den Bewegungsrichtungen zu sehen. Dies deutet auf ein unterschiedliches Laufverhalten der Kugeln bei Vor- und Rückhub hin.

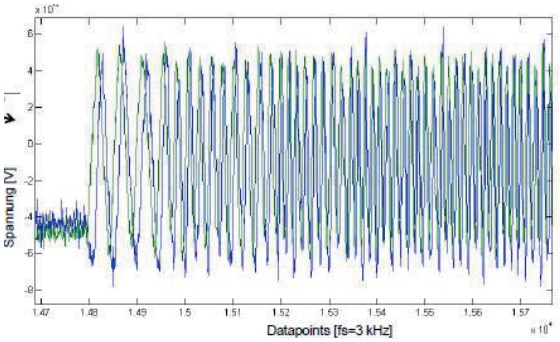


Abb. 14: Messergebnisse im Zeitbereich auf dem Spindelachsenprüfstand. Signale 5 mV mit abgeglichenem Offset bei Rohrwandstärke 0,1mm [3].

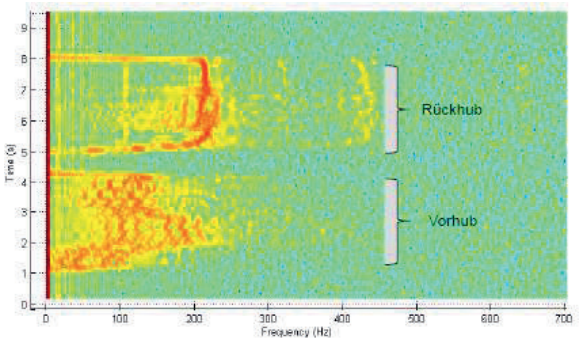


Abb. 15: Messergebnisse im Frequenzbereich auf dem Spindelachsenprüfstand getrennt nach Vor- und Rückhub [3].

Das diesem Bericht zugrundeliegende Vorhaben Moses-Pro wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 16ES0423 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Literaturnachweis

[1] Verbundprojekt MoSeS-Pro, Modulare Sensorsysteme für Echtzeit-Prozesssteuerung und smarte Zustandsbewertung, https://moses-pro.de/, abgerufen am 15.04.2019

[2] N. Helwig, T. Schneider und A. Schütze, MoSeS-Pro: Modular Sensor Systems for Real Time Process Control and Smart Condition Monitoring using XMR-Technology, 14th Symposium Magnetoresistive Sensors and Magnetic Systems, Wetzlar, 2017

[3] ZeMA gGmbH, Präsentation Projekttreffen MoSeS-Pro, Duisburg, 2017

DOI 10.5162/sensoren2019/4.4.3

Modularer Messsystembaukasten für FPGA-basierte Signalverarbeitung

Dominik Laumann, Manuel Selke, Christian Nienhaus, Joachim Doerr
CANWAY TECHNOLOGY GMBH, Graf-Zeppelin-Ring 13, D-48346 Ostbevern

Zusammenfassung

Das vorliegende Dokument betrachtet die aus der Digitalisierung, der Industrie 4.0 und hochgradig vernetzten Systemen resultierenden Anforderungen an aktuelle und zukünftige Sensor- und Messsysteme für industrielle Produktionsprozesse und zeigt einen Lösungsansatz unter Verwendung eines modularen Baukastensystems basierend auf FPGA Technologie. Systemanforderungen, wie hohe Verfügbarkeit, Zuverlässigkeit sowie Flexibilität und eine intensive Vernetzung von Sensoren und Systemen werden mit verschiedenen Lösungsansätzen von Zustandsüberwachung über Self-X Funktionen bis zu "Soft-Sensoren" adressiert. Diese Lösungsansätze erfordern eine leistungsfähige Signalverarbeitung und hohe Flexibilität auf der Ebene der Sensorschnittstelle, welche optimal durch den Einsatz von FPGAs als Technologie umgesetzt werden kann. Im Folgenden werden die Anforderungen und die Zieltechnologie detaillierter erläutert. Im Anschluss wird das Systemkonzept des Modulbaukastens und die konkrete Umsetzung verschiedener Elektronikmodule vorgestellt. Beispielhafte Anwendungen zeigen die Einsatzmöglichkeiten und Flexibilität des Systems.

Die beschriebenen Konzepte sind u. A. im Rahmen des BMBF geförderten Verbundprojekts MoSeS-Pro (FKZ 16ES0422) erarbeitet worden.

Keywords: Modulbaukasten, FPGA, Signalverarbeitung, Sensorschnittstelle

Einführung

Alle bekannten Bestrebungen für eine zukunftsorientierte Auslegung industrieller Produktionsprozesse, wie z. B. der Industrie 4.0 oder der Digitalisierung, haben gemeinsame Anforderungen auf der Systemebene. Zur Umsetzung dieser "high-level" Anforderungen werden verschiedene Lösungsansätze verfolgt, die entsprechende Anforderungen an Sensorelektronik, Sensorschnittstellen und Messsystemebene bedeuten. Diese Anforderungen werden im Folgenden betrachtet, um im Anschluss ein Lösungskonzept auf Basis von FPGA Technologie vorzustellen.

Eine der Systemanforderungen an industrielle Produktionsprozesse ist die hohe Verfügbarkeit und Zuverlässigkeit. Ein allgemein bekannter Ansatz ist die Zustandsüberwachung von Systemen, die eine Verschleiß- oder Ausfallprädiktion zur Einplanung von Wartungsphasen ermöglicht. Weitere Ansätze sind die frühe Fehlerdetektion und Fehlerkompensation z. B. durch Self-X Eigenschaften von Sensoren, die eine Selbstüberwachung, Selbstkalibrierung oder Selbstheilung ermöglichen [1], und sog. "Soft-Sensoren", die den Ausfall eines Sensors durch intelligente Algorithmen mit Daten noch intakter alternativer Sensoren kompensieren können. Diese Methoden erfordern es, einen möglichst hohen Informationsgehalt aus bestehenden Sensorsignalen zu gewinnen, der häufig über die eigentliche Messaufgabe im Rahmen einer Prozessteuerung hinaus geht. Somit ergeben sich auf Ebene der Sensorschnittstelle zum Teil hohe Anforderungen an die Abtastrate und Signalauflösung, folglich aber auch die Notwendigkeit einer leistungsfähigen Signalverarbeitung und intelligenter Algorithmen zur sensornahen Datenreduktion, um die Kommunikationsinfrastruktur nicht zu überlasten [2].

Eine weitere Systemanforderung ist die Flexibilität von Fertigungs- und Montageprozessen, die eine Adaption von Fertigungssystemen an sich ändernde Aufgaben, sowohl durch Parametrierung des Systems beim Anlagenstillstand als auch durch Rekonfiguration im laufenden Betrieb erfordert. Diese Flexibilität erfordert eine Adaption der Sensor- und Schnittstellenelektronik durch Modularisierung (Flexibilität durch Kombination von Hardwaremodulen) und durch Parametrierung und Rekonfiguration im Feld (Flexibilität durch Kombination von Softwaremodulen). Leistungsreserven dienen dabei zur Vorbereitung

auf noch nicht bekannte, zukünftige Anwendungsfälle.
Die dritte hier betrachtete Systemanforderung ist die hochgradige Vernetzung, auch unterschiedlicher hierarchischer Ebenen, in industriellen Fertigungsprozessen. Der Ansatz, ein Maximum an Informationsgehalt aus den Prozessen zu gewinnen, erfordert hohe Datenraten und Echtzeitfähigkeit auf verschiedenen Feldbussen und Schnittstellen. Eine zeitliche Synchronisierung der Messstellen ist für eine sensorübergreifende Signalauswertung zwingend erforderlich. Sensorschnittstellen müssen demnach verschiedene aktuelle und zukünftige Kommunikationsprotokolle und Schnittstellen unterstützen, die hohe Datenraten, Echtzeitfähigkeit und eine Zeitsynchronisierung ermöglichen.
Auf Basis dieser Anforderungen an zukünftige Sensorschnittstellen und Messsysteme wurde die Field-Programmable-Gate-Array (FPGA) und System-on-a-Chip (SoC) Technologie als geeignete Zieltechnologie der Elektronik identifiziert. Im Folgenden werden das Konzept und die Umsetzung eines modularen Messsystembaukastens für FPGA-basierte Signalverarbeitung vorgestellt.

FPGA Technologie

Um die Anforderungen eines vollständig flexiblen Hardwaresystems zu erfüllen, ist die Verwendung von FPGAs derzeit die marktführende Technologie. Dies sind hochkomplexe integrierte digitale Schaltungen zur Adaption an viele verschiedene Anwendungen. FPGAs sind eine Weiterentwicklung von programmierbaren Logikbausteinen. Sie enthalten eine Vielzahl von programmierbaren Logikblöcken (PLB), die komplexe kombinatorische Funktionen ausführen. Diese PLBs sind durch rekonfigurierbare Verbindungen zu einer Gesamtschaltung verbunden. Die Konfiguration wird durch eine Hardwarebeschreibungssprache (HDL) beschrieben. Üblicherweise werden VHDL oder Verilog verwendet. Die HDL beschreibt das Verhalten der Anwendung und wird verwendet, um eine synthetisierbare digitale Schaltung zu erstellen.
Die meisten FPGAs enthalten auch Speicherblöcke zum Zwischenspeichern von Daten. Sie werden oft als SRAM-Speicher implementiert. Ein FPGA Design kann leicht in einen ASIC (Application Specific Integrated Circuit) überführt werden. ASICs sind kundenspezifische Schaltungen, die für eine bestimmte Anwendung ausgelegt sind. Die Produktionskosten sind niedrig, die Einmalkosten jedoch hoch. Die Funktion eines ASICs kann im Betrieb nicht beeinflusst werden. Daher werden FPGAs häufig verwendet, um die Funktionalität vorher zu verifizieren oder sie werden dort eingesetzt, wo hohe Flexibilität erforderlich ist [3].
Moderne Chips kombinieren die FPGA Technologie mit eingebetteten Mikroprozessoren als System-on-a-Chip. Diese Technologie ermöglicht die Kombination leistungsfähiger serieller und paralleler Datenverarbeitung in einem integrierten Schaltkreis [4].

Modulbaukasten Systemkonzept

Eine Sensorschnittstelle oder generell ein Messsystem lässt sich in vier wesentliche Funktionen unterteilen, die durch ihr Zusammenwirken das Systemverhalten abbilden. Im Sinne des Baukastengedankens wird eine Sensorschnittstelle in vier Subkomponenten mit eben diesen Eigenschaften modularisiert. *Abb. 1* zeigt den schematischen Aufbau einer Sensorschnittstelle und die Auftrennung der vier Teilaufgaben auf die einzelnen Module.

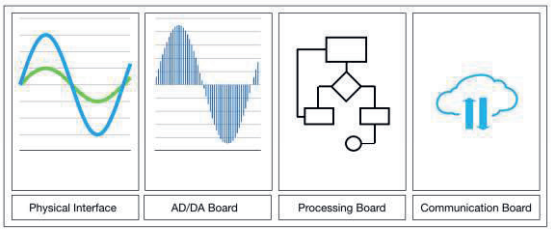


Abb. 1: Schematischer Aufbau einer modularen Sensorschnittstelle

Das Physical Interface dient der Signalkonditionierung des Sensorsignals und stellt notwendige Versorgungsspannungen für die Sensorik zur Verfügung. Der Übergang zwischen analoger und digitaler Welt wird durch ein AD/DA Board realisiert. Die Signalverarbeitung erfolgt auf einem Processing Board. Ein Communication Board dient als Schnittstelle zu einer übergeordneten Steuerung oder Auswertungseinheit. Die Modularisierung auf diese vier Teilaufgaben ermöglicht eine anwendungsspezifische Adaption des Systems durch Kombination der Module. Die Komplexität einzelner Module ist dabei wesentlich reduziert im Vergleich zum Gesamtsystem und generische Module ermöglichen einen hohen Grad der Wiederverwendung.
Im Folgenden werden verschiedene Konzeptbeispiele zum modularen Messsystem erläutert und verschiedene entwickelte Module vorgestellt. *Abb. 2* zeigt den Einsatz des Messsystems als Sensorschnittstelle zur sensornahen Signalvorverarbeitung z. B. bei der Zustandsüberwachung.

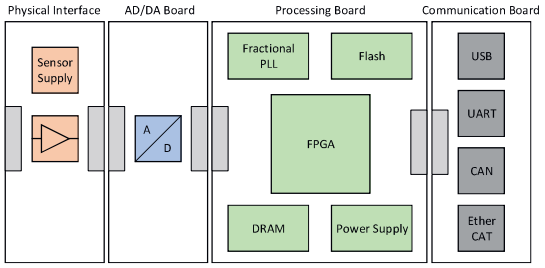


Abb. 2: Sensorschnittstelle für sensornahe Signalvorverarbeitung

Das Physical Interface bildet die analoge Sensorschnittstelle. Hier werden die Signalpegel der Sensoren für eine nachfolgende Verarbeitung angepasst und notwendige Versorgungsspannungen zum Betrieb der Sensorik bereitgestellt. Ein Board mit Analog-Digital-Wandler dient der Digitalisierung des Sensorsignals, das Processing Board stellt einen FPGA für die sensornahe Signalverarbeitung zur Verfügung und das Communication Board ermöglicht die Übertragung von vorverarbeiteten, reduzierten Daten über verschiedene Feldbusschnittstellen. *Abb. 3* zeigt ein erweitertes System mit sensornahem Feedback für z. B. Self-X Eigenschaften.

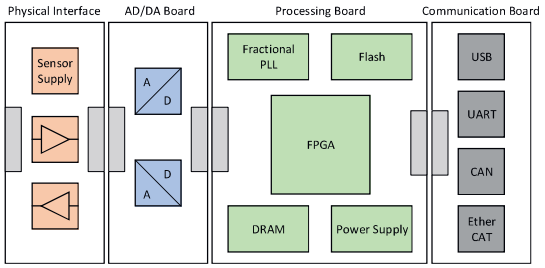


Abb. 3: Sensorschnittstelle mit sensornahem Feedback

Aufgrund der modularen Bauweise müssen zu diesem Zweck nur das Physical Interface und das AD/DA Board ersetzt werden, die Processing und Communication Boards können hardwareseitig identisch verwendet werden und über ein Softwareupdate an die neue Funktion angepasst werden. Das in *Abb. 4* dargestellte Konzept erweitert dieses System für TCP/IP basierte Kommunikationsprotokolle und softwarelastige, für die Verarbeitung auf einem Mikroprozessor geeignete Algorithmen.

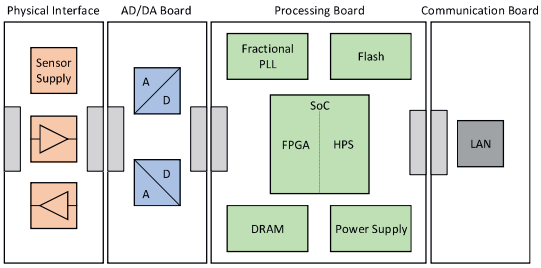


Abb. 4: Sensorschnittstelle mit SoC für TCP/IP-basierte Kommunikation

Dazu wird das FPGA Processing Board durch ein SoC-basiertes Board ersetzt und ein Communication Board mit Ethernet Schnittstelle verwendet. In diesem Fall können das Physical Interface und das AD/DA Board weiterverwendet werden.

Umsetzungsbeispiele der Module

Im Rahmen des Projekts wurden die Grundlegenden Module, wie FPGA und SoC Processing Boards, sowie verschiedene Interface- und Kommunikationsmodule umgesetzt und in ausgewählten Anwendungsfällen erprobt. *Abb. 5* zeigt ein Processing Board mit einem FPGA der Cyclone V Familie der Firma Intel [5]. Diese FPGA Familie befindet sich im unteren bis mittleren Preis- und Leistungssegment von FPGAs und erfüllt sehr gut die Anforderungen an verfügbare Ressourcen und Performance für die sensornahe Signalverarbeitung.



Abb. 5: FPGA Board mit Cyclone V

Das Processing Board bringt neben dem FPGA auch die gesamte für den Betrieb des FPGAs notwendige Peripherie mit sich, wie Arbeitsspeicher zur Speicherung von Daten und Zwischenergebnissen, einen Flash-Speicher für die FPGA Konfiguration und anwendungsspezifische Parameter und eine konfigurierbare Taktquelle für eine möglichst jitterfreie Abtastung von Sensorsignalen. Neben diesem Processing Board wurde ein weiteres Modul auf SoC-Basis entwickelt, das schnittstellenkompatibel zu dem vorgestellten FPGA-basierten Modul ist. Als SoC wurde hier

ebenfalls die Cyclone V Familie der Firma Intel gewählt. Die Peripherie ist sehr ähnlich zu der des FPGA Boards, ergänzt um eine Micro SD Karte bzw. einen eMMC Speicher, in denen typischerweise das Dateisystem eines Betriebssystems für den Mikroprozessor Teil des SoCs gespeichert wird.
Für eine exemplarische Anwendung, der Zustandsüberwachung für Verschleißprädiktion an einer Spindelachse, wurde ein AD Board mit zwei verschiedenen Kanaltypen entwickelt.



Abb. 6: AD Board

Das Board stellt insgesamt vier analoge Eingänge mit einer Abtastrate von 5 MS/s und einer Auflösung von 16 Bit zur Verfügung, die z. B. zur Überwachung von synchron abgetasteten Motorströmen oder Messung von Luftschall eingesetzt werden. Darüber hinaus stehen insgesamt acht analoge Eingänge mit einer Abtastrate von 192 kS/s und einer Auflösung von 24 Bit ebenfalls mit einer synchronen Abtastung zur Verfügung, mit denen typischerweise Encodersignale und Beschleunigungssensoren gemessen werden können. Ein passendes Physical Interface Board bildet die notwendige Signalkonditionierung und Spannungsversorgung für die Sensoren.
Ein zugehöriges Communication Board ermöglicht die Übertragung von vorverarbeiteten Merkmalen für eine Visualisierung, weitere Verarbeitung oder Speicherung über Ethernet. Auch die Konfiguration des Systems kann so über TCP/IP-basierte Protokolle erfolgen.

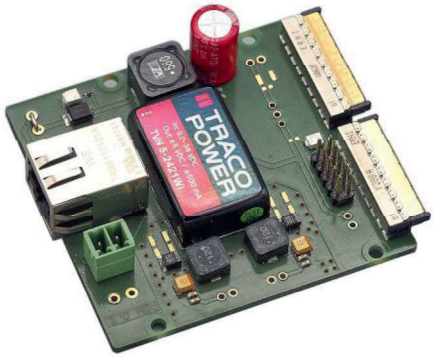


Abb. 7: Communication Board mit Ethernet Schnittstelle

Um auch komplexere Messsysteme realisieren zu können, lassen sich mehrere FPGA Processing Boards über eine Backplane mit einem SoC Processing Board koppeln. Die FPGA Boards dienen dabei der leistungsfähigen Signalvorverarbeitung, das SoC Board zur übergeordneten Konfiguration, Koordination und Kommunikation. Dazu sind die Boards über die Backplane mittels schneller LVDS Links verbunden. *Abb. 8* zeigt die Kombination eines FPGA und eines SoC Processing Boards in dem zu diesem Zweck entwickelten Tischgehäuse, das auch die Spannungsversorgung und eine Ethernet Schnittstelle breitstellt.



Abb. 8: Komplexes Messsystem im Tischgehäuse

Anwendungsbeispiele

Im Rahmen des Projekts MoSeS-Pro wurde mit den vorgestellten Elektronikmodulen eine Sensorschnittstelle für die Zustandsüberwachung an einer Spindelachse entwickelt. *Abb. 9* zeigt das Gesamtsystem bestehend aus Physical Interface, AD Board, FPGA-basiertem Processing Board und Communication Board.

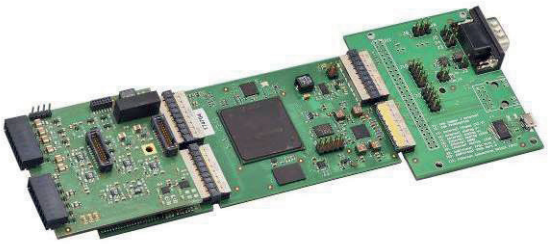


Abb. 9: Sensorschnittstelle für Zustandsüberwachung an Spindelachse

Dieses System bietet die Möglichkeit der Überwachung von Signalen, wie Motorströme, Encoder, Vibration oder Luftschall. Diese Signale können mit Hilfe von Methoden der Signalverarbeitung zur Zustandsüberwachung für eine Verschleißprädiktion genutzt werden. Da diese Algorithmen teilweise hohe Ansprüche an die Signalvorverarbeitung stellen, ist ein FPGA hier optimal geeignet [6].
Die Flexibilität des Systems und die Leistungsfähigkeit der FPGA-basierten Signalverarbeitung kann eindrucksvoll an einem weiteren Beispiel, der Umsetzung eines Netzwerkanalysators, demonstriert werden. Hierzu wurde ein System zur vektoriellen Messung von Übertragungsfunktionen an elektrischen Vierpolen mittels fourier-basierter Technik realisiert. Zu diesem Zweck kann ein zu untersuchendes System mittels verschiedener Signale breitbandig angeregt werden und sowohl der Stimulus als auch die Systemantwort zurückgemessen werden. Die Übertragungsfunktion wird durch Diskrete Fourier Transformation (DFT) im Frequenzbereich ermittelt. *Abb. 10* zeigt den schematischen Aufbau des Netzwerkanalysators.

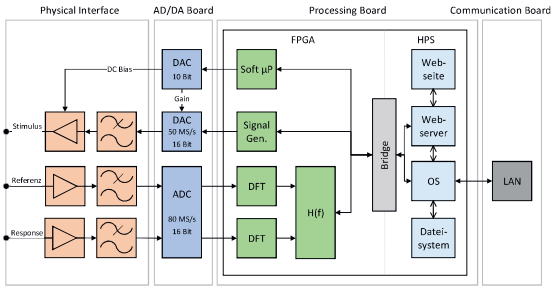


Abb. 9: Netzwerkanalysator (Schematischer Aufbau)

Die Signalverarbeitung zur Anregung und Auswertung wird im FPGA-Teil eines SoCs durchgeführt. Ein Signalgenerator ermöglicht die Erzeugung verschiedener Anregungssignale (z. B. Pseudozufallsfolgen, Frequenzsweeps oder Multisinus). Das digital generierte Signal wird über einen DA-Wandler mit einer Abtastrate von 50 MS/s und einer Auflösung von 16 Bit an eine Filter- und Verstärkerstufe ausgegeben und bildet somit den Stimulus der Messung. Ein 2-kanal AD-Wandler mit einer Abtastrate von bis zu 80 MS/s und einer Auflösung von 16 Bit tastet das Stimulussignal als Referenz und die Systemantwort ab und stellt es dem FPGA zur Verarbeitung zur Verfügung. Ein im FPGA implementierter sog. Softprozessor steuert dabei die Abläufe bei einer Messung und dient der Einstellung von Bias und Gain des Stimulussignals. Die Konfiguration des Systems und die Visualisierung von Messdaten erfolgt über den Mikroprozessorteil (HPS) des SoCs. Auf dem Prozessor kommt ein Embedded Linux Betriebssystem zum Einsatz, welches die Verwendung eines Dateisystems, eines Webservers und die TCP/IP basierte Kommunikation über Ethernet ermöglicht. Eine Webseite dient als Benutzerschnittstelle und ermöglicht die Konfiguration des Systems sowie die Messdatenvisualisierung über einen herkömmlichen Internetbrowser.

Fazit und Ausblick

Im Rahmen der Untersuchungen wurde ein Konzept für ein modulares Messsystem zur leistungsfähigen, sensornahen Signalvorverarbeitung erstellt und die Funktion an verschiedenen Modulen und Anwendungsfällen demonstriert. Die Modularisierung ermöglicht eine einfache Adaption an verschiedene Anwendungsfälle. Die FPGA Technologie dient hierbei als leistungsstarke, flexible Plattform für sensornahe Signalverarbeitung und ermöglicht eine Rekonfiguration im Feld. Zukünftige Anforderungen an Algorithmik oder Schnittstellen können aufgrund von vorhandenen Leistungsreserven ergänzt werden. Die SoC Technologie bietet darüber hinaus die Ausführung von komplexen, softwarelastigen Algorithmen, die Vorteile eines Betriebssystems und die Kommunikation über TCP/IP-basierte Protokolle.
Im Rahmen dieser Arbeit standen das Systemkonzept und die Entwicklung der Elektronikmodule im Vordergrund. Themen der Softwareinfrastruktur und Konfiguration des Systems, Miniaturisierung und Energieeffizienz müssen Bestandteil weiterer Untersuchungen sein. Dabei soll der bereits in Hardware umgesetzte Baukastengedanke konsequent auf der Ebene der Firmware- und Softwareinfrastruktur weiter verfolgt werden, um die Modularität und Leistungsfähigkeit des Baukastens einem breiteren Anwenderkreis, ohne Erfahrung in der Programmierung von FPGAs, zu eröffnen. Im Rahmen eines „Downsizing“ sollen weitere, anwendungsspezifische Module mit reduzierter Größe, geringeren Kosten und minimalem Energiebedarf umgesetzt werden.

Literaturnachweis

[1] Andreas Schütze, Nikolai Helwig, Tizian Schneider, *Sensors 4.0 – smart sensors and measurement technology enable Industry 4.0* (2018), doi: 10.5194/jsss-7-359-2018

[2] *Modulare Sensorsysteme für Echtzeit-Prozesssteuerung und smarte Zustandsbewertung (MoSeS-Pro),* https://www.moses-pro.de/wp-content/uploads/2016/05/Projektsteckbrief_MoSeS-Pro.pdf, Zugriff: 25.04.2019

[3] M. Selke, C. Nienhaus, D. Laumann, J. Doerr, CANWAY TECHNOLOGY GMBH, Ostbevern (Germany), *Universal Sensor Interface for High Performance Signal Processing Based on FPGA Technology* (2017), doi: 10.5162/sensor2017/B2.2

[4] Prof. Dr.-Ing. Frank Kesel, Dr. Ruben Bartholomä, *Entwurf von digitalen Schaltungen und Systemen mit HDLs und FPGAs, Oldenbourg Verlag München*, 3. Auflage 2013, ISBN 978-3-486-73181-1

[5] Intel/Altera Inc., San Jose 2018, *Cyclone V Device Handbook*

[6] U. Meyer-Baese, *Digital Signal processing with Field Programmable Gate Arrays,* Springer Verlag, 4. Auflage 2014, ISBN 978-3-642-45308-3

DOI 10.5162/sensoren2019/5.1.1

Mikromagnetische Detektion von Härteinhomogenitäten an Grobblechen

Sargon Youssef[1], Martin Schuppmann[1], Werner Bähr[1], Sergey Pushkarev[1], Hendrik Theado[1], Philipp Stopp[1], Jens Wagner[1], Daniel Molenda[2], Olaf Stawicki[2]

[1] Fraunhofer Institut für Zerstörungsfreie Prüfverfahren IZFP
Campus E3 1
66123 Saarbrücken, Deutschland

[2] ROSEN Technology and Research Center GmbH
Am Seitenkanal 8
49811 Lingen (Ems)

Zusammenfassung

Die hohen Qualitätsansprüche von Stahlerzeugnissen und Komponenten im Maschinen-, Anlagen- und Fahrzeugbau erfordern eine hohe Qualität des Ausgangsmaterials. Beispielsweise stellt die Homogenität der mechanisch-technologischen Materialeigenschaften wie die Härte bei Grobblechen, ein wichtiges Qualitätsmerkmal dar. Zur Detektion von Härteinhomogenitäten in Grobblechen wurde das 3MA-X8 Verfahren weiterentwickelt und in einen Handwagen implementiert. In einer umfangreichen von der Firma ROSEN beauftragten und begleitenden Qualifizierungsstudie wurden verschiedene Einflussfaktoren wie Zundervariationen, magnetische Restfelder und Blechtemperatur auf die zuverlässige Detektion von Härteinhomogenitäten im industriellen Umfeld untersucht. Die positiven Ergebnisse resultierten in der Implementierung des Prüfverfahrens in eine automatisierte Prüfanlage. Im Rahmen dieses Beitrages werden die Ergebnisse der Qualifizierungsstudie und die aktuellen Fortschritte bei der Implementierung in die automatisierte Prüfanlage vorgestellt. Neben dem mechanischen Konzept wird ein Überblick über die verfahrensseitige FPGA-basierte Implementierung gegeben.

Keywords: Zerstörungsfreie Prüfverfahren, Materialcharakterisierung, Maschinelles Lernen, Grobblech, Hardspot,

Einleitung

Die hohen Qualitätsansprüche von Stahlerzeugnissen und Komponenten im Maschinen-, Anlagen- und Fahrzeugbau erfordern eine hohe Qualität des Ausgangsmaterials. Beispielsweise stellt die Homogenität der mechanisch-technologischen Materialeigenschaften wie die Härte bei Grobblechen, ein wichtiges Qualitätsmerkmal dar. Hieraus resultiert ein in den letzten Jahren steigender Bedarf an der zerstörungsfreien Charakterisierung der Materialeigenschaften von Grobblechen für die Pipelineindustrie. Nach ISO15156, NACE MR0175 oder DNV-OS-F101 dürfen die Stähle für den Sauergaseinsatz eine Härte von 250 HV10 nicht überschreiten. Bei Härtewerten ab ca. 300 HV10 reduziert sich die Dehnbarkeit bei kohlenstoffhaltigen Stählen und die Anfälligkeit auf Materialfehler beim Verlegen der Rohre, oder auch während des Betriebs erhöht sich. Aus wirtschaftlichen Gründen werden von den Betreibern oft hochwertige Stahlgüten (z.B. X60, X65, etc.) auch für Pipelines mit saurem Medium bevorzugt eingesetzt. Diese Stahlgüten weisen eine höhere Grundhärte des Materials auf. Beispielsweise X65 besitzt eine Grundhärte von ca. 220 HV10. Die Differenz zur erlaubten Maximalhärte von Sauergasanwendungen beträgt demnach ca. 30 HV10. Deshalb suchen Grobblechhersteller und deren Kunden, entweder als Rohrlieferant oder als Endanwender nach geeigneten Möglichkeiten, diese Härteinhomogenitäten (engl. Hardspots) sicher zu erfassen. Dadurch kann die Qualität des gesamten Blechs bereits nach dem Walzprozess beurteilt werden. Für diese Aufgabenstellung sind Härteinhomogenitäten von +/- 30 HV10 bei einem minimalen Durchmesser von 10 mm zu detektieren.

3MA-X8 Prüfverfahren

Das am Fraunhofer IZFP entwickelte 3MA-X8 Verfahren (3MA = Mikromagnetische Multipa-

rameter Mikrostruktur und Spannungs-Analyse) nutzt Effekte in der magnetischen Hysterese aus [1] und kombiniert basierend auf diesen Effekten drei unabhängige Prüfmethoden mit unterschiedlicher Eindringtiefe und Sensitivität, um im Hinblick auf Störeinflüsse (z.B. Zunder, Restmagnetfelder, Blechtemperatur) im Material eine zuverlässige und stabile Charakterisierung des Werkstoffzustandes zu ermöglichen. Daraus werden bis zu 21 mikromagnetische Merkmale zur Charakterisierung von Materialeigenschaften wie z.B. der Härte extrahiert [2]. Einzelne Merkmale sind bereits von einem Materialkennwert (z.B. Härte) abhängig, sofern keine weiteren überlagerten Einflüsse wie Gefügevariation, Spannungseinfluss, magnetische Restfelder, etc. vorhanden sind. In der Praxis liegt aber oft eine Überlagerung der Einflussgrößen vor. Die Oberflächenhärte führt beispielsweise häufig zu Veränderungen des Eigenspannungszustandes und variierender Einhärtetiefe, hinzu kommen fertigungsbedingte Schwankungen im Bearbeitungsprozess oder chargenbedingte Gefügeschwankungen. Diese Störeinflüsse beeinflussen wiederum die mikromagnetischen Merkmale, so dass sich beim Betrachten eines einzelnen Merkmals Mehrdeutigkeiten ergeben können. Diese Mehrdeutigkeit kann durch eine Verknüpfung mehrerer Merkmale mit Hilfe von maschinellen Lernalgorithmen (sog. Kalibrierung) minimiert werden. Bei der Kalibrierung erfolgt zunächst eine Zuordnung der mikromagnetischen Merkmale zu den entsprechenden Zielgrößen (z.B. Härte) [1]. Wichtig für die Qualität der Kalibrierung ist in der Anlernphase jedoch die valide Kenntnis der Zielgröße an repräsentativen Kalibrierproben. Diese sollten auch die in der Praxis relevante Variation an zusätzlichen Einflussgrößen abbilden. Die Kalibrierung kann hier abhängig von der Aufgabenstellung qualitativ (z.B. Hart / Weich) oder quantitativ (z.B. Härte in HV10) durchgeführt werden.

Im Rahmen einer von der Firma ROSEN beauftragten und begleiteten Qualifizierungsstudie wurden für die Praxis relevante Untersuchungen durchgeführt.

Ein wesentlicher Punkt betraf die Untersuchung der im Walzwerk vorhandenen Einflussfaktoren und der damit erforderlichen Robustheit und Stabilität des Systems um Hardspots zuverlässig unter Praxisbedingungen zu detektieren.

Qualifizierungsstudie

Zur Qualifizierung des Prüfsystems wurden unterschiedlich bearbeitete Bleche von der Firma ROSEN zur Verfügung gestellt, um verschiedene Mechanismen der Entstehung von Hardspots (z.B. Aufkohlung oder thermische Behandlung) unter Variation von Zunder und magnetischer Restfelder und Blechtemperatur (-20 °C – 80 °C) zu untersuchen (vgl. Abb. 1 oben links und rechts).

Magnetische Restfelder und Zundervariation

Magnetische Restfelder und Zundervariationen treten im Herstellungsprozess bei Stahlblechen als Störgrößen auf. Die genaue Variation der Zunderschicht ist im Stahlherstellungsprozess oft unbekannt. Deshalb wurde der Einfluss des Zunders auf die mikromagnetischen Merkmale bzw. auf das Ergebnis der Kalibrierung untersucht (vgl. entfernter Zunder Abbildung 1 oben links und oben rechts). Analog verhält es sich mit Restmagnetfeldern (vgl. lokale Aufmagnetisierung von ca. 15 A/cm an zwei Stellen auf dem Blech, Abb. 1 unten links). Dies ist vor dem Hintergrund relevant, die Bleche im Stahlwerk oft mit Magnetkränen transportiert werden. Die Kalibrierung wurde mit Hilfe der multiplen linearen Regressionsanalyse durchgeführt.

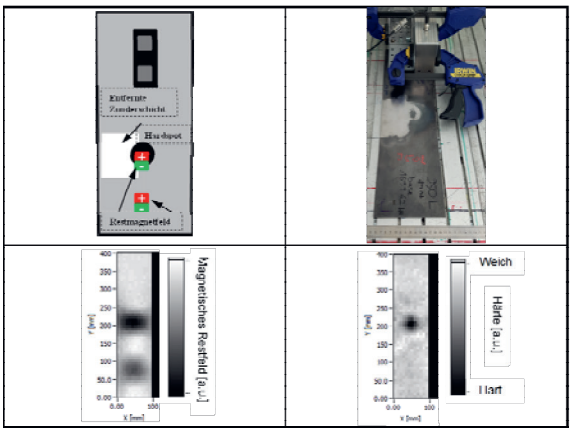


Abb. 1: Darstellung (schematisch) eines Blechs mit partiell entferntem Zunder, magnetischen Restfelder und einer lokalen Aufhärtung von 30 HV (oben links), Foto des Messaufbaus (oben rechts), restfeldabhängiges 3MA-X8 Merkmal Rem (unten links), Ergebnis nach Kalibrierung (unten rechts)

Der Einfluss der Störeinflüsse in den mikromagnetischen Merkmalen wird im Ergebnisbild nach der Kalibrierung kompensiert (vgl. Abb. 1 unten rechts).

Blechtemperatur

Einige der mit dem 3MA-X8 Ansatz extrahierten Merkmale zeigen eine lineare,

signifikante Abhängigkeit von der Blechtemperatur. Zur Untersuchung des Temperaturverhaltens wurde je ein Sensor im Klimaschrank auf einer weichen und auf einer um 30 HV10 aufgehärteten Stelle eines Testblechs positioniert und die Temperatur während der Messung von -20 °C bis 80 °C in 20 K Schritten mit einer Haltezeit von vier Stunden variiert (vgl. Abb. 2 oben). In einer getrennten Messreihe wurden Sensor und Blech getrennt den Temperatureinflüssen in der Klimakammer ausgesetzt, um den Einfluss der Temperatur auf beide Einheiten getrennt zu untersuchen mit dem Ergebnis, dass kein signifikanter Temperatureinfluss der Sensoren vorliegt.

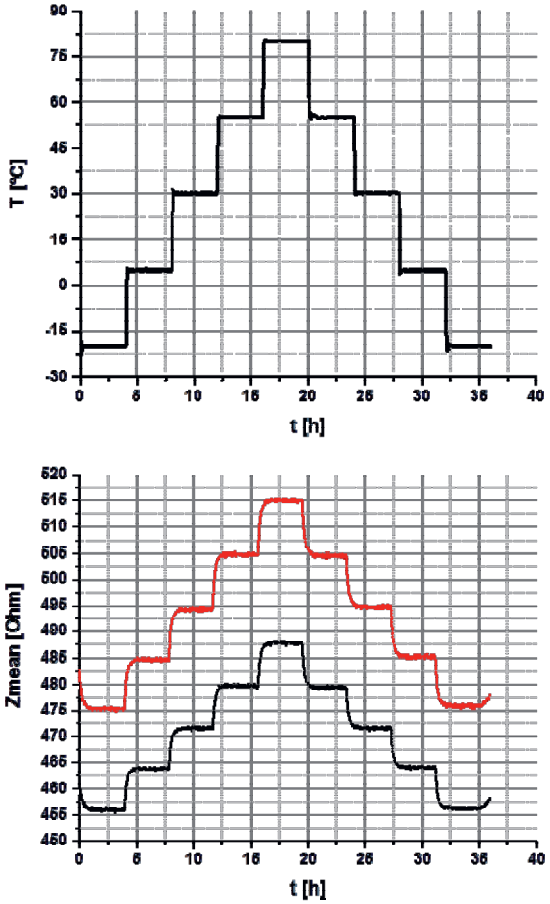


Abb. 2: Temperaturprofil (oben) und Temperaturabhängigkeit von Zmean (unten)

Bei Variation der Blechtemperatur sind verschiedene Merkmale aber stark temperaturabhängig. Das Merkmal Zmean beispielsweise zeigt eine signifikante Temperaturabhängigkeit (vgl. Abb. 2 unten und Abb. 3 oben). Diese Abhängigkeit kann durch aktive Messung der Temperatur mit einem vom magnetischen Verhalten unabhängigen Sensor auf zwei Arten kompensiert werden. Die erste Möglichkeit ist die Erweiterung des Merkmalsraums um die Temperaturinformation. Statt maximal 21 mikromagnetischer Merkmale erhält man ein weiteres ausschließlich von der Blechtemperatur abhängiges Merkmal, mit deren Hilfe der maschinelle Lernalgorithmus den Temperatureinfluss berücksichtigt und diesen passiv kompensiert. Alternativ zur Erweiterung des Merkmalsraums kann die Information der Blechtemperatur zur aktiven Kompensation des Temperatureinfluss auf die mikromagnetischen Merkmale verwendet werden. Hierzu wird das Verhalten im untersuchten Temperaturbereich ausgenutzt und die betroffenen Merkmale mit Hilfe einer linearen Regression vom Temperatureinfluss kompensiert (vgl. Abb. 3 oben).

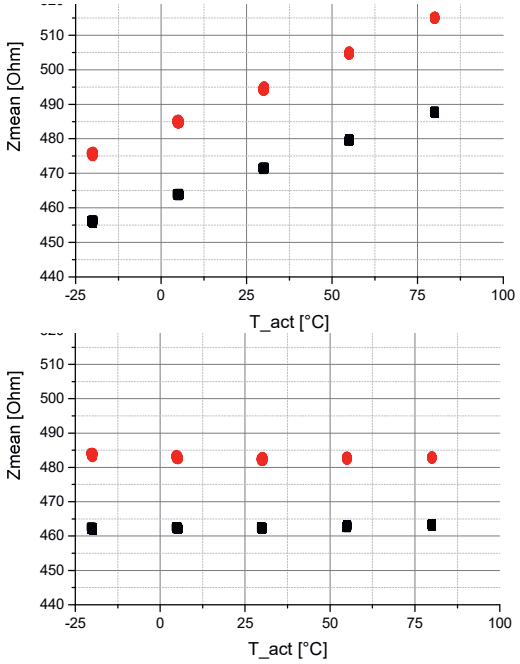


Abb. 3: Temperaturabhängigkeit von Zmean (oben) und Zmean nach der Kompensation

Beide Möglichkeiten erlauben eine Kompensation des Temperatureinflusses, die Auswahl der Kompensationsmethode ist im Anwendungsfall zu prüfen.

Kalibrierung

Bei der Prüfung von Grobblechen ist der erforderliche Aufwand zum Kalibrieren zu berücksichtigen. Neben den untersuchten Einflussfaktoren variieren zusätzlich noch die Mechanismen, die zu Hardspots auf dem Grobblech führen. Dies führt unter Umständen zu Mehrdeutigkeiten in den Daten der mikromagnetischen Merkmale. Um diese zu kompensieren, ist ein repräsentativer

Probensatz mit in der Praxis vorkommender Variation ggf. für unterschiedliche Stahlqualitäten und Stahlzusammensetzungen erforderlich. Die Erstellung dieses Probensatzes und die Ermittlung der Referenzwerte ist aufgrund der Variationsvielfalt mit einem unverhältnismäßig hohen Aufwand für die Stahlhersteller verbunden, weshalb ein alternativer und im Messbetrieb schneller und wirtschaftlicher Kalibrieransatz erforderlich ist. Als alternativer Kalibrieransatz kann die Kalibrierung des 3MA-X8-Prüfsystems (statt wie bisher im Vorfeld) im laufenden Betrieb stattfinden. Durch Vorgabe einer Schwelle (z.B. 250 HV10) erfolgt eine qualitative Kalibrierung (OK = Härtewert im spezifizierten Bereich / NOK = Härtewert über dem spezifizierten Bereich / ERROR = Materialzustand unbekannt).

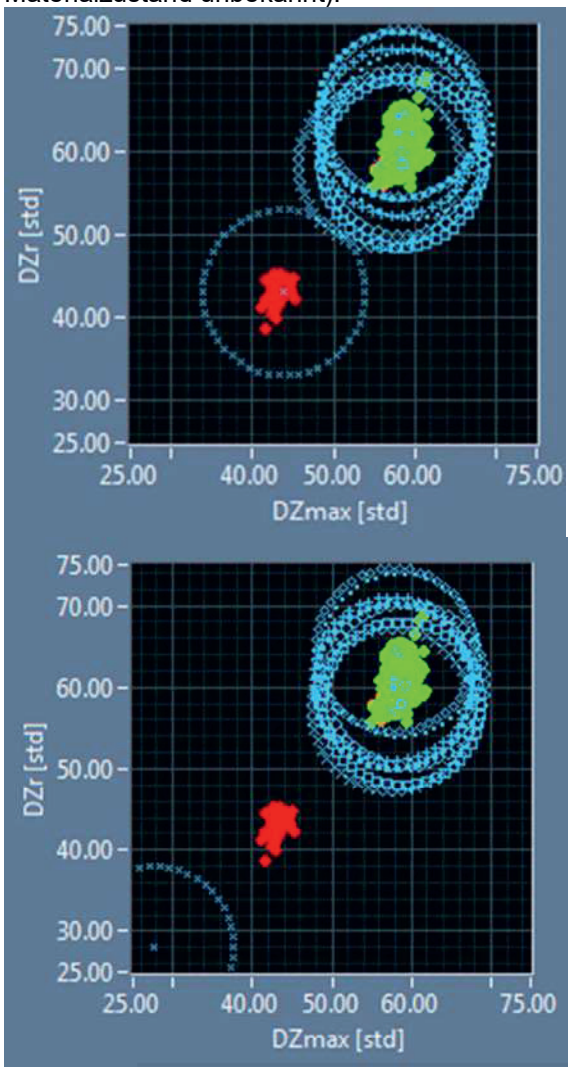


Abb. 4: Einkalibrierte Zustände, OK in grün, NOK in rot (oben) und Zustand ERROR (unten)

Zunächst wird der Sensor auf einer Stelle des Bleches mit einem Härtewert unterhalb des spezifizierten Bereichs positioniert und als „OK“ kalibriert. Zusätzlich kann das Prüfsystem optional auf einer Stelle mit einem Härtewert oberhalb des spezifizierten Bereichs platziert werden und dieser Zustand als „NOK“ kalibriert werden. Durch Aufnahme verschiedener Positionen auf dem auf Blech können weitere Zustände einkalibriert werden (vgl. Abb. 4). Die Klassifizierung „OK“ / „NOK“ erfolgt mit Hilfe eines „K-Nächste-Nachbar-Klassifikators“. Dieser Klassifikator bewertet den Abstand vom aktuellen Messwert zu den einkalibrierten Daten in der Datenbank. Messwerte nahe den kalibrierten Daten (Kalibrierdaten innerhalb Kreis in Abb. 4 oben) werden entsprechend als „OK“ / „NOK“ klassifiziert, Messwerte ohne ausreichend viele Kalibrierdaten innerhalb des Kreises werden als ERROR klassifiziert (vgl. Abb. 4 unten) und sind mit einem Referenzverfahren nachzuprüfen und in der Kalibrierdatenbank aktualisieren. Liegt die Härte im spezifizierten Bereich, wird diese Position als „OK“ kalibriert. Liegt die Härte über dem spezifizierten Bereich, wird diese Position als „NOK“ kalibriert (vgl. Abb. 5).

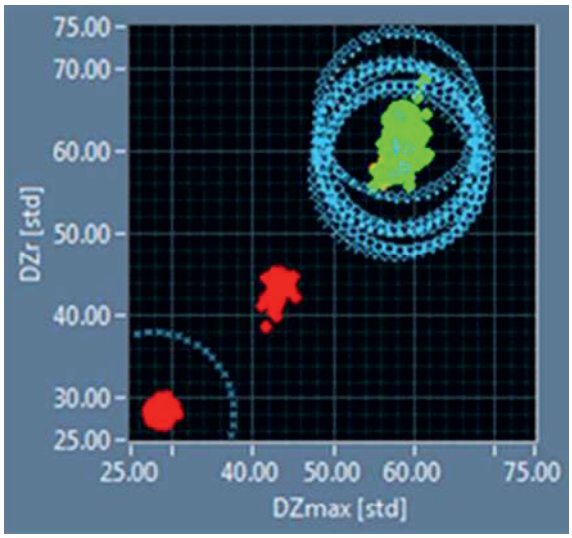


Abb. 5: Aktualisierte Kalibrierdatenbank

Die Visualisierung erfolgt in Form eines Messteppichs, der sich in Echtzeit auf dem Auswerterechner aufbaut. Als „OK“ klassifizierte Messwerte werden in grün, „NOK“ klassifizierte Messwerte werden in orange und als „ERROR“ klassifizierte Messwerte werden in rot dargestellt (vgl. Abb. 6).

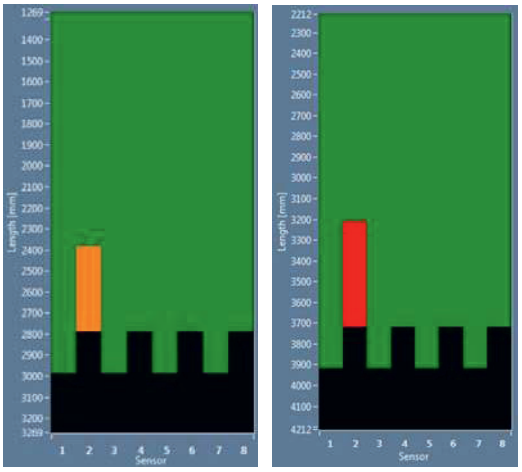


Abb. 6: Messteppich mit OK (grün), NOK (orange) und ERROR (rot)

Dadurch lässt sich mit geringem Aufwand eine flexible und stetig erweiterbare Kalibrierung umsetzen, die unterschiedliche Stähle mit unterschiedlichen Aufhärtungsmechanismen in einer einzigen Kalibrierdatenbank bewerten kann.

3MA-X8 Trolley

Das Prüfverfahrens wurde in einen achtkanaligen Handwagens zur flächendeckenden Detektion von Hardspots auf Grobblechtafeln implementiert (vgl. Abb. 7).



Abb. 7: 3MA-X8 Handwagen

Anforderungen an den Handwagen sind die Detektion lokaler Härteinhomogenitäten mit einem Durchmesser von mindestens 10 mm und einer Aufhärtung von 30 HV10 bei einer Prüfgeschwindigkeit von maximal 1 m/s im Temperaturbereich von 20 °C bis 50 °C. In diesem Temperaturbereich ist eine keine Kompensation des Temperatureinflusses erforderlich. Wesentliche Baugruppen dieses Handwagens sind zwei versetzt angeordnete vierkanalige Sensorzeilen mit einer Prüfbreite von 240 mm, ein Weggeber zur positionsgetreuen Zuordnung der Messungen und zur Detektion von Prüfgeschwindigkeiten oberhalb 1 m/s. Die Ergebnisse werden in der Bediensoftware (vgl. Abb. 8) als Messteppich in Form einer OK (grün) / NOK (orange) Verteilung dargestellt, auf Fehlfunktionen der Sensoren oder einer zu schnellen Bewegung über das Blech wird in der Software hingewiesen.

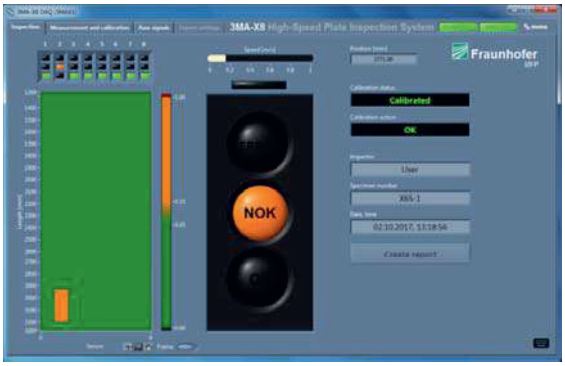


Abb. 8: 3MA-X8 Bediensoftware

3MA-X8 Prüfanlage

Aktuell wird das Prüfverfahren in Kooperation mit der Firma ROSEN in eine online laufende Hardspot-Prüfanlage implementiert. Die Prüfanlage basiert auf einem modularen Konzept, dass sich für konventionelle Ultraschall-Blechprüfanlagen bewährt hat. Je nach Applikation und zu verwendender Technologie können unterschiedliche Module konfiguriert werden. Abb. 9 zeigt einen Überblick über die allgemeine Konfiguration der Prüfanlage.



Abb. 9: Allgemeine Konfiguration der Prüfanlage

Das Grundkonzept der Hardspot-Prüfanlage basiert auf zwei stabilen Rahmen für die Flächenprüfung. Entsprechende Prüfwagen können zur Wartung oder bei Nutzung des Rollgangs ohne Prüfung aus dem Rollgang herausgefahren werden. Die Flächenprüfwagen sind positionierbar ausgeführt. Damit ist eine optimale Ausrichtung der Prüfkopfhalterungen in Abhängigkeit von der Blechbreite möglich. Die Prüfwagen bestehen aus modularen Prüfgruppen mit jeweils 8 Sensorhalterungen. Entsprechend der zu prüfenden Blechbreite wird die benötigte Anzahl an den Prüfwagenrahmen angebaut.
Die Sensorhalterungen sind über Pneumatikzylinder anstellbar und sind beweglich ausgeführt, so dass sie der Blechoberfläche folgen können (vgl. Abb. 10 oben).

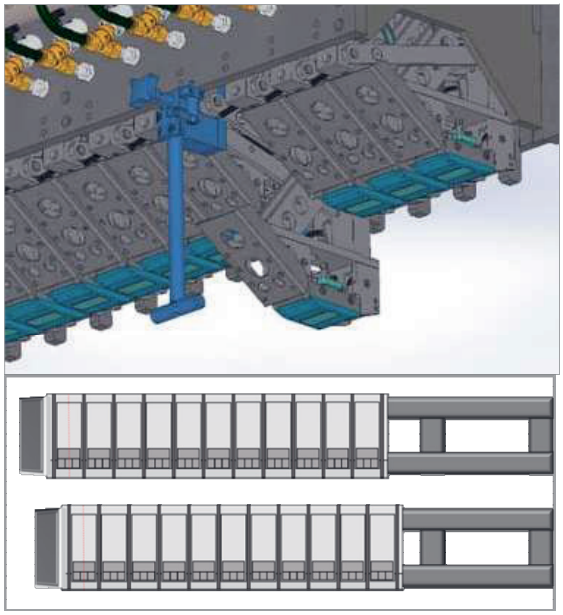


Abb. 10: Sensorhalterung (oben) und versetzte Prüfwagen (unten)

Zwischen den einzelnen Halterungen eines Prüfwagens entsteht eine Prüflücke, die durch eine versetzte Anordnung auf dem zweiten Prüfwagen aufgefüllt wird (vgl. Abb. 10 unten).

Parallel erfolgt die verfahrensseitige Implementierung. Hierfür wurde am Fraunhofer IZFP ein FPGA-basierter IP-core (IP = Intellectual Property) entwickelt der die Extraktion der 21 mikromagnetischen Merkmale in Echtzeit aus bis zu acht parallel arbeitenden Prüfkanälen ermöglicht. Diese Variante eignet sich insbesondere zur Integration in kundenspezifische, automatisierte Prüfanlagen, da hier die besonders zeitkritischen Verarbeitungsschritte und die Reduktion der Daten unmittelbar hinter dem Sensor stattfinden und nur die Ergebnisse an eine Leitstelle übermittelt werden müssen. Zusammen mit dem streng deterministischen Zeitverhalten der hardwarebasierten Verarbeitung können, unter Nutzung mehrerer Instanzen der IP, automatisierte Prüfanlagen mit einer großen Anzahl an parallel arbeitenden Prüfkanälen realisiert werden. Die IP kann von Kunden leicht in seine eigene FPGA Infrastruktur eingebunden und verwendet werden (vgl. Abb. 11).

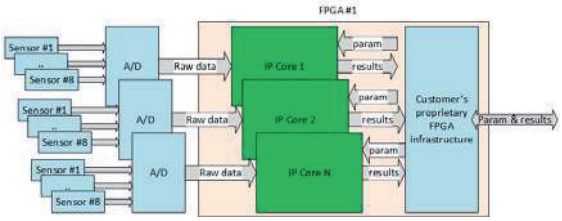


Abb. 11: FPGA-Infrastruktur

Die hardwarebasierte Merkmalsextraktion wird ergänzt von einer speziell für den industriellen Einsatz entwickelten Signalanregungsendstufen- und Signalaufnahmeelektronik (vgl. Abb. 12). Bei der Entwicklung standen hierbei Platzersparnis sowie Energieeffizienz im Vordergrund. Um insbesondere die Abwärme bei der Erzeugung der Magnetisierungsströme zu reduzieren wurde ein achtkanaliges Klasse-D Endstufenkonzept entwickelt und implementiert. Das energieeffiziente Design ermöglicht die vollständig passive Kühlung aller Elektronikkomponenten und damit deren Integration in sensornahe, hermetisch verschlossene Gehäuse.

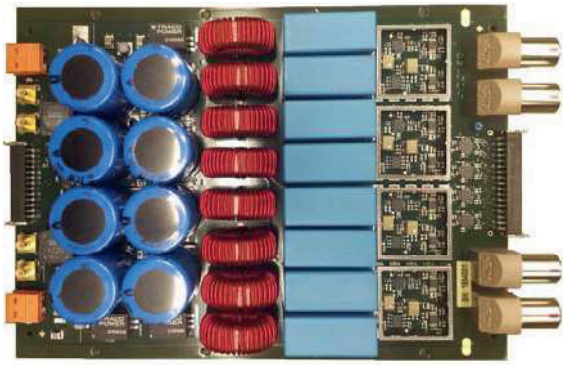


Abb. 12: FPGA-Infrastruktur

Literaturnachweis

[1] Kneller, E.: Ferromagnetismus. Springer-Verlag Berlin, Göttingen, Heidelberg. 1962.
[2] Youssef, Sargon, Cyril Zimmer, Klaus Szielasko, Andreas Schütze: “Bewertung subjektiver und automatisierter Merkmalsextraktion periodischer Zeitsignale am Beispiel des 3MA-X8-Verfahrens“ tm - Technisches Messen, 0.0 (2019): Retrieved 24 Jan. 2019, from doi:10.1515/teme-2018-0074

DOI 10.5162/sensoren2019/5.1.2

A DC bias approach to the characterisation of non-linear material parameters of piezoelectric ceramics

Dmitrij Dreiling[1], Nadine Feldmann[1], Bernd Henning[1]
[1] Measurement Engineering Group, Paderborn University, Warburger Straße 100, 33098 Paderborn, Germany

Abstract

Piezoelectric materials are used for various sensor and actuator purposes. Especially in power acoustic applications piezoceramics are loaded with electrical DC voltage or mechanical pre-load, in addition to the excitation signal. Due to the lack of knowledge of material parameters for piezoceramics, particularly in this operating mode, the design of a power transducer is demanding and often strongly empirically driven. A simulation-supported design process requires that a non-linear model for the piezoelectric material behaviour is known. A first step towards this material model is the description of linear approximated material properties in different operating points. To shift this operating point a DC bias voltage is applied to various piezoelectric discs and the frequency dependent impedance is determined. The results indicate a shift of all characteristic resonance frequencies as well as changes of piezoelectric parameters directly dependent on the applied bias voltages.

Keywords: piezoelectric materials, piezoelectric properties, DC bias field, non-linear material parameters

1 Introduction

Recently, simulation-driven procedures for the determination of material parameters of piezoelectric materials in the linear domain have been published [1, 2]. Nevertheless, there are methods to estimate the linear material properties for simulation-driven design processes. However those provide just rough estimate and deviate strongly regarding power acoustic applications or under external influences like a mechanical/electrical pre-load. Under these circumstances, non-linearities play a non-negligible role. There are different approaches to characterise and model those non-linearities, for example by taking non-linear second order material parameters into account [3, 4] or by extending the Mason model [5] with non-linear voltage sources [6]. An accurate non-linear model of piezoelectricity and a better understanding could be used to increase the efficiency of ultrasonic transducers or to avoid undesired effects such as depolarisation.

A common way to determine electromechanical properties is to linearise in an operating point, to estimate the values by measuring the impedance of a specimen and to detect characteristic frequencies [7]. A DC bias field applied to a piezoceramic shifts the operating point and influences the material parameters [3, 4, 8]. Li et al. investigate the relation between the influence of external bias field or temperature on material parameters and draw the conclusion that this effect can be advantageously used for the tuning of piezoelectric properties [9]. By one-dimensional considerations for a plate resonator with finite thickness, [3, 4] provide a possibility to determine material parameters from a resonance frequency shift. In addition, they derive equations for the determination of non-linear parameters by changing the operating point. This paper follows the same approach but for piezoelectric discs with adapted assumptions and modified equations. DC bias fields of various strengths are used to shift the operating point of the samples and estimate different linear material parameters (figure 1). Finally, these parameters can be used to estimate the non-linear material parameters.

2 Mathematical description

2.1 Linear approximation for material properties

Commonly, in the small signal range the relationship between the mechanical quantities, mechanical stress $\boldsymbol{T}$ and strain $\boldsymbol{S}$, and the electrical quantities, electric field $\boldsymbol{E}$ and dielectric displacement $\boldsymbol{D}$, can be regarded as approximately linear. Therefore the following constitutive equations (in Voigt notation) can be assumed

$$\begin{aligned} \boldsymbol{S} &= \boldsymbol{s}^{E}\boldsymbol{T} + \boldsymbol{d}^{\mathrm{t}}\boldsymbol{E}, \\ \boldsymbol{D} &= \boldsymbol{d}\boldsymbol{T} + \varepsilon^{\boldsymbol{T}}\boldsymbol{E}, \end{aligned} \tag{1}$$

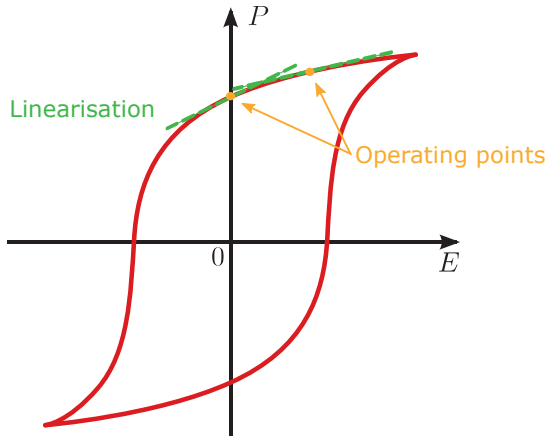


Figure 1: *Exemplary hysteresis curve shows the relation between the polarisation P and the electric field strength E with a linearisation for the small-signal behaviour in two operating points.*

where $\boldsymbol{d}$, $\boldsymbol{s^E}$ and $\boldsymbol{\varepsilon^T}$ are the coupling, the elastic compliance and the permittivity parameter matrices, respectively [7]. Taking into account the transversal isotropy and the resulting symmetry conditions, a total of ten independent material parameters are necessary, to describe the piezoelectric behaviour. In summary, three piezoelectric (d_{15}, d_{31}, d_{33}), five elastic ($s_{11}^{\boldsymbol{E}}, s_{12}^{\boldsymbol{E}}, s_{13}^{\boldsymbol{E}}, s_{33}^{\boldsymbol{E}}, s_{44}^{\boldsymbol{E}}$) and two dielectric ($\varepsilon_{11}^{\boldsymbol{T}}, \varepsilon_{33}^{\boldsymbol{T}}$) material parameters are required. In order to determine all ten parameters, according to the IEEE Standard [7], several specimens of different shapes are required. The use of several specimens, however, entails uncertainties because the material properties are not consistent due to different production and polarisation processes. Since the specimens are piezoelectric discs poled in the thickness direction and with a much larger radius than thickness, the piezoceramic can be reduced to a 1D-model which is only sensitive to $s_{33}^{\boldsymbol{E}}, \varepsilon_{33}^{\boldsymbol{T}}$ and d_{33}. Mathematically this leads to a reduced set of equations:

$$\begin{aligned} S_3 &= s_{33}^{\boldsymbol{E}} T_3 + d_{33} E_3, \\ D_3 &= d_{33} T_3 + \varepsilon_{33}^{\boldsymbol{T}} E_3. \end{aligned} \tag{2}$$

If a piezoelectric ceramic is exposed to an alternating electric field, the electrical impedance Z or admittance Y is dependent on the frequency. The following procedure to determine the desired material parameters $s_{33}^{\boldsymbol{E}}, \varepsilon_{33}^{\boldsymbol{T}}$ and d_{33} is based on the investigation of impedance measurements assuming a free resonator in thickness mode and of the fundamental analysation and the harmonics [7]. According to IEEE [7], at alternating electric fields, the admittance can be described as:

$$Y(\omega) = \frac{1}{Z(\omega)} = \left(\frac{\mathrm{j}\omega \varepsilon_{33}^{\boldsymbol{S}} \varepsilon_0 A}{h}\right) \frac{1}{1 - k^2 \frac{\tan(\omega/4 f_\mathrm{p})}{\omega/4 f_\mathrm{p}}} \tag{3}$$

Here, ω is the angular frequency, ε_0 the dielectric constant of vacuum, A the electrode area and h the thickness of the specimen. f_p is the parallel resonance frequency of the first thickness mode, also known as the frequency of the maximum resistance and $\varepsilon^{\boldsymbol{S}}$ is the permittivity at constant strain (clamped permittivity). The electromechanical coupling factor k can be determined by

$$k = \sqrt{\frac{\pi}{2} \frac{f_\mathrm{s}}{f_\mathrm{p}} \cot\left(\frac{\pi}{2} \frac{f_\mathrm{s}}{f_\mathrm{p}}\right)}, \tag{4}$$

with f_s as the series resonance frequency of the first thickness mode or frequency of the maximum conductance. By measuring the impedance of a piezoceramic it is possible to locate the characteristic frequencies f_s and f_p as well as to determine the free capacity C_0. In addition, the density ρ is required which can be determined by measuring the mass and volume of the sample. Using these the linear parameters, $s_{33}^{\boldsymbol{E}}, \varepsilon_{33}^{\boldsymbol{T}}$ and d_{33} can be estimated for different operating points.

2.2 Non-linear approximation for material properties

A DC bias approach to a characterisation of non-linear material parameters of piezoelectric discs demands an extension of the constitutive equations by second order material parameters [4, 10]. Once again for a one-dimensional approximation, equations 2 can be rewritten as

$$\begin{aligned} S_3 &= s_{33}^{\boldsymbol{E}} T_3 + s_{333}^{\boldsymbol{E}} T_3^2 + d_{33} E_3 + d_{333} E_3 T_3 + \zeta_{33} E_3^2, \\ D_3 &= d_{33} T_3 + d_{333} T_3^2 + \varepsilon_{33}^{\boldsymbol{T}} E_3 + \zeta_{33} E_3 T_3 + \varepsilon_{333}^{\boldsymbol{T}} E_3^2, \end{aligned} \tag{5}$$

where $\varepsilon_{333}^{\boldsymbol{T}}$, d_{333} and ζ_{33} are the non-linear dielectric, piezoelectric and electrostrictive material parameters, respectively. In Voigt notation, the subscript indices describe the direction of the inducing and resulting state variables. Since several inducing forces act on one non-linear material parameter, an additional index is necessary to describe the relation of the second order state variables (e.g. E_3 and T_3 act on d_{333} resulting in S_3). Due to the superposition of a DC bias field $\overline{E}$ with a small AC field $\widetilde{E}$, the mechanical and electrical quantities are

$$\begin{aligned} E &= \widetilde{E} + \overline{E}, \quad S = \widetilde{S} + \overline{S}, \\ D &= \widetilde{D} + \overline{D}, \quad T = \widetilde{T}. \end{aligned} \tag{6}$$

It is noticeable that, because of the assumption of a mechanically free resonator, the stress T has no contribution from a static component $\overline{T}$. According to [3, 4], the extended quantities in equation 6 can be inserted into equation 5 and split into two dynamic equations

$$\widetilde{S}_3 = (s_{33}^{\boldsymbol{E}} + d_{333}\overline{E}_3)\widetilde{T}_3 + (d_{33} + \zeta_{33}\overline{E}_3)\widetilde{E}_3, \quad (7)$$

$$\widetilde{D}_3 = (d_{33} + \zeta_{33}\overline{E}_3)\widetilde{T}_3 + (\varepsilon_{33}^{\boldsymbol{T}} + \varepsilon_{333}^{\boldsymbol{T}}\overline{E}_3)\widetilde{E}_3, \quad (8)$$

and two static equations

$$\overline{S}_3 = d_{33}\overline{E}_3, \quad (9)$$

$$\overline{D}_3 = \varepsilon_{33}^{\boldsymbol{T}}\overline{E}_3, \quad (10)$$

for a DC field biased resonator. These static equations describe the mechanical extension and the dielectric displacement of the material as a result of an applied electrical DC bias field. However, it is important that the dynamic equations are now DC bias dependent and the relationship between the linear and non-linear parameters can be described as follows:

$$s_{33}^{\text{eff}} = (s_{33}^{\boldsymbol{E}} + d_{333}\overline{E}_3) = s_{33}^{\boldsymbol{E}}\left(1 + \frac{d_{333}}{s_{33}^{\boldsymbol{E}}}\overline{E}_3\right), \quad (11a)$$

$$d_{33}^{\text{eff}} = (d_{33} + \zeta_{33}\overline{E}_3) = d_{33}\left(1 + \frac{\zeta_{33}}{d_{33}}\overline{E}_3\right), \quad (11b)$$

$$\varepsilon_{33}^{\text{eff}} = (\varepsilon_{33}^{\boldsymbol{T}} + \varepsilon_{333}^{\boldsymbol{T}}\overline{E}_3) = \varepsilon_{33}^{\boldsymbol{T}}\left(1 + \frac{\varepsilon_{333}^{\boldsymbol{T}}}{\varepsilon_{33}^{\boldsymbol{T}}}\overline{E}_3\right). \quad (11c)$$

These material parameters, s_{33}^{eff}, d_{33}^{eff} and $\varepsilon_{33}^{\text{eff}}$, have been determined by the linear approximation at different operating points. Therefore, the equations can be solved for the non-linear parameters.

2.3 The DC bias field dependence of the resonance frequency

Furthermore, the dependence of the resonance frequency on a DC bias field can be derived [3, 4]. A longitudinal acoustic wave velocity influenced by an applied DC bias field can be described as follows:

$$v_{\text{t}}(\overline{E}_3) = \sqrt{\frac{1}{s_{33}^{\text{eff}}\rho}}. \quad (12)$$

Not only the elastic compliance s_{33}^{eff} but also the density ρ depend on the applied voltage. With the volume V of the piezoelectric disc expanding in the same direction as the DC voltage assuming the 1D approximation, the mass m of the disc remains constant. Thus ρ can be written as

$$\rho(\overline{E}_3) = \frac{m}{V(\overline{E}_3)} = \frac{m}{V_0[1 + d_{33}\overline{E}_3]}, \quad (13)$$

with V_0 the volume without an applied DC bias field.
According to [7] the wave velocity propagating in thickness direction of a disc shaped piezoceramic depends on the serial resonance frequency f_{s}. Solving the equation for f_{s} and using the DC bias field dependent equations 13, 12 and 11 yields

$$f_{\text{s}}(\overline{E}_3) = \frac{v_{\text{t}}(\overline{E}_3)}{2h} = f_{\text{s},0}\left[1 + \left(\frac{d_{33}}{2} - \frac{d_{333}}{2s_{33}^{\text{eff}}}\overline{E}_3\right)\right], \quad (14)$$

where the change of the thickness $h = h_0(1 + d_{33}\overline{E}_3)$ has been included. The index 0 indicates values without an applied DC bias field. By equation 14 a direct dependence of the resonance frequency on the applied electrical field is given. Accordingly, a change of the frequency with and without applied DC voltage can be described as follows:

$$\frac{\Delta f_{\text{s}}}{f_{\text{s},0}} = \frac{d_{33}}{2} - \frac{d_{333}}{2s_{33}^{\text{eff}}}\overline{E}_3. \quad (15)$$

The first term describes the shift of the series frequency due to the mechanical deformation of the piezoceramic disc. The second term describes the part of the resonance shift caused by the non-linear behaviour.

3 Measurement system

The measurement setup depicted in figure 2 is used to determine the frequency dependent impedance of various piezoelectric specimens under high electric fields. A DC bias test fixture connects the sample to a DC voltage source (EA-PSI 9200-04T power supply) and the impedance analyser (Keysight E4990A). The design of the fixture is based on a protection circuit for the impedance analyser recommended by the manufacturer [11]. It supports an external voltage as well as a frequency range of 35 kHz to 20 MHz. The DC source provides voltages up to 200 V which is sufficient to apply an electric field strength up to $1\,\frac{\text{kV}}{\text{mm}}$, depending on the thickness of the samples.
All examined piezoceramics are disc-shaped with full-surface silver electrodes manufactured by **P**hysik **I**nstrumente (PI) Ceramic GmbH [12]. In total, three different types of piezoceramics were investigated. PIC255 is a piezoceramic based on soft PZT with a radius of $r = 8\,\text{mm}$ and thickness of $h = 0.2\,\text{mm}$. PIC181 is a hard PZT piezoelectric ceramic with equal dimensions. PIC700 is a lead-free piezoceramic material based on bismuth sodium titanate with coupling factors and permittivities comparable to piezoceramics based on barium titanate, according to the manufacturer [12]. During the investigations, PIC700 was not available with the same dimensions as the other specimen, therefore the lead-free ceramic has a thickness of $h = 0.5\,\text{mm}$ and a radius of $r = 5\,\text{mm}$.

4 Results

4.1 Measurement results

As described in section 2, impedance spectra of the specimen measured near the resonance frequencies f_{s} and f_{p} can be used to determine the linear material parameters $s_{33}^{\boldsymbol{E}}, \varepsilon_{33}^{\boldsymbol{T}}$ and d_{33}. All samples are mounted as shown in figure 2 and the applied electric field strength is gradually changed from $1\,\frac{\text{kV}}{\text{mm}}$ to $-1\,\frac{\text{kV}}{\text{mm}}$

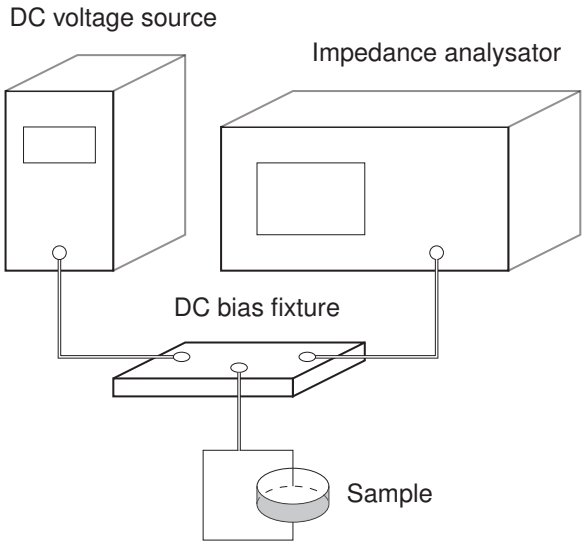


Figure 2: *Schematic of the measurement setup.*

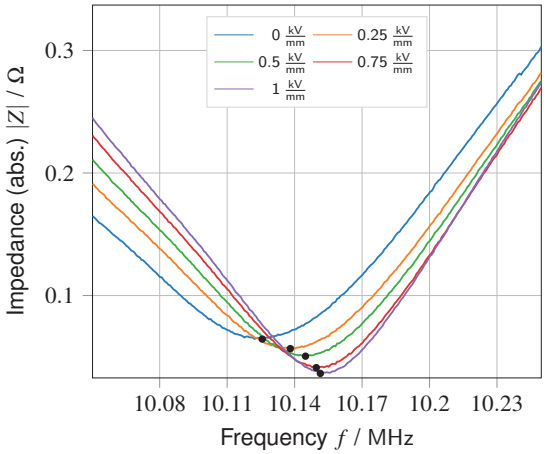


Figure 3: *The relation between the absolute value of the impedance $|Z|$ and frequency f of PIC255 at various positive DC bias fields applied. Series resonance frequency is marked by black dots.*

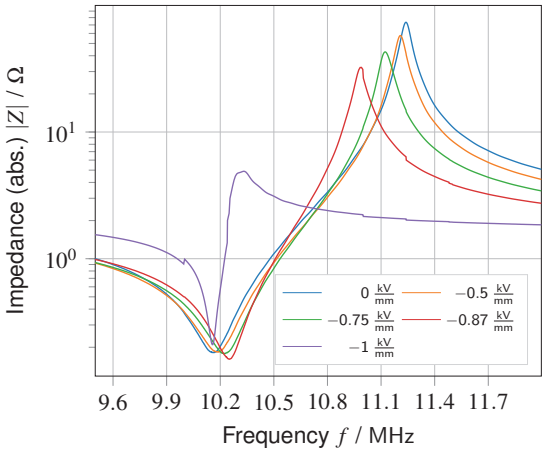


Figure 4: *The relation between the absolute value of the impedance $|Z|$ and frequency f of PIC255 at various negative DC bias fields applied. Depolarisation occurring at $-1\ \frac{\mathrm{kV}}{\mathrm{mm}}$.*

and back again. Figure 3 depicts the measured relation between the magnitude of the impedance $|Z|$ and the resonance frequency f of the PIC255 at different operating points. The black markers indicate the detected series resonance frequency at different bias field strengths and visually underline the shift of f_s. Moreover a general trend of the shift to higher frequencies with an electric field applied along the polarisation direction is recognizable. And vice versa the frequencies f_s and f_p shift to lower frequencies with negative (against the polarisation direction) electric field strengths applied. The maximum Δf_s of PIC255, measured with an applied electric field strength of $1\ \frac{\mathrm{kV}}{\mathrm{mm}}$, is ≈37 kHz. It should be mentioned that for correct mathematical estimation of the linear material parameters, f_s is detected as the maximum weighted conductance $\max(\mathrm{Re}\{Y(\omega)/\omega\})$ and f_p as the maximum weighted resistance $\max(\mathrm{Re}\{\omega Z(\omega)\})$ [3, 13].

Figure 4 depicts the relation between $|Z|$ and f for PIC255 with applied negative electric field strengths. Using E greater than $-800\ \frac{\mathrm{V}}{\mathrm{mm}}$, the piezoceramic material depolarises. This state remains, even if the negative field is switched off. Yang et al. observed a similar material depoling of soft PZT piezoelectric samples EC-65 and EC-76 at field strengths of $-800\ \frac{\mathrm{V}}{\mathrm{mm}}$ and $-500\ \frac{\mathrm{V}}{\mathrm{mm}}$, respectively [14]. However, the piezoelectric ceramic can be re-polarised by applying a positive electric field strength of $1\ \frac{\mathrm{kV}}{\mathrm{mm}}$. This is evident when considering the material parameters as a function of the applied DC bias field in subsection 4.3.

In general, PIC181 shows the same frequency shift behaviour as PIC255 when an electric field is applied, although the effect is smaller ($\Delta f_\mathrm{s} \approx 27$ kHz). Moreover this piezoceramic disc doesn't depolarise at negative bias field strengths up to $-1\ \frac{\mathrm{kV}}{\mathrm{mm}}$. The lead-free piezoceramic PIC700 also appears to be less susceptible to the applied field. However, it should be noted that due to the geometry ($h = 0.5$ mm) only maximum DC bias field strengths of $\pm 400\ \frac{\mathrm{V}}{\mathrm{mm}}$ are applicable. In the aforementioned field range, a maximum Δf_s of ≈1 kHz can be measured for all examined PIC700 ceramics and is comparable to the magnitude of the shift of PIC181 at $\pm 400\ \frac{\mathrm{V}}{\mathrm{mm}}$, but thinner samples are necessary for a reliable statement. The general shift shows an opposite trend in the aforementioned field strength range, thus f_s and f_p shift to lower frequencies at greater DC bias fields and vice versa.

4.2 Frequency shift by deformation

At this point it is important to verify that the frequency shift is caused by the non-linear behaviour and not the material deformation. For this purpose, equation 15 can be used to determine a $\Delta f_{\mathrm{s,mech}}$ which is solely

dependent on mechanical deformation.

$$\Delta f_{\mathrm{s,mech}} = f_{\mathrm{s},0}\frac{d_{33}}{2}\overline{E}_3 \tag{16}$$

The calculated value should be lower than the measured value $\Delta f_{\mathrm{s}}(\overline{E}_3)$. For PIC255 the $\Delta f_{\mathrm{s,mech}} = 12.1\,\mathrm{kHz}$ and for PIC181 $\Delta f_{\mathrm{s,mech}} = 8.7\,\mathrm{kHz}$. Both values are below the aforementioned frequency shift, so it can be claimed that the present effect is caused primarily by the DC bias field and non-linear behaviour of the piezoceramics.

4.3 Linear parameter estimation

So far, only the influence of the applied DC bias field on the absolute value of the impedance has been analysed and compared. The change of the elastic compliance as a function of the applied DC bias field of PIC255 and PIC181 is depicted in figure 5. Both piezoelectric samples were exposed to the same electrical stress cycle. As expected, the piezoceramic PIC181 seems to be less influenced by a bias field, because the changes in the elastic compliance are small compared to PIC255. Nevertheless the shift of material parameter of both piezoceramics shows a trend in the same direction. As the DC bias field decreases, the elastic compliance of both piezoceramics increases.

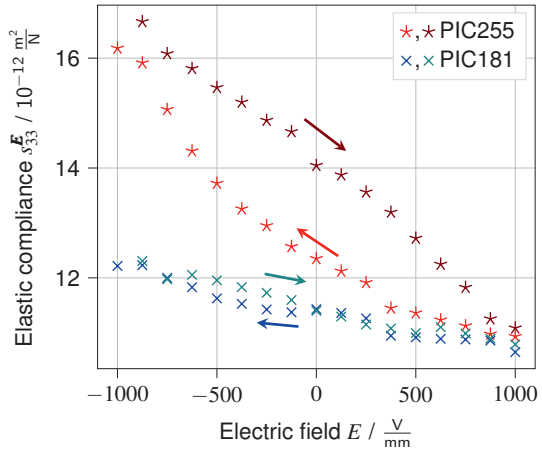


Figure 5: *Elastic compliance s_{33}^{E} of PIC255 and PIC181 at different operating points.*

A similar behaviour for d_{33} and ε_{33}^{T} is shown in figure 6 and 7, respectively. While the piezoelectric parameter and the permittivity of PIC255 change significantly, they remain relatively constant in PIC181. It is not surprising that the trend of both aforementioned material parameters is proportional because the permittivity is necessary to estimate the piezoelectric constant [7]. Another effect which occurs with all material parameters of the PIC255 is a discontinuity of the parameters at high negative field strengths. Due to depolarisation of PIC255, the characteristic resonance frequencies are only poorly indicated and a reliable determination of d_{33}, s_{33}^{E} and ε_{33}^{T}, which is based on the detection of f_{s} and f_{p}, is impossible. Moreover the observed significant decrease of d_{33} and ε_{33}^{T} at $E = -1\,\frac{\mathrm{kV}}{\mathrm{mm}}$ remain low until a positive field strength of more than $500\,\frac{\mathrm{kV}}{\mathrm{mm}}$ is applied to the soft PZT sample.

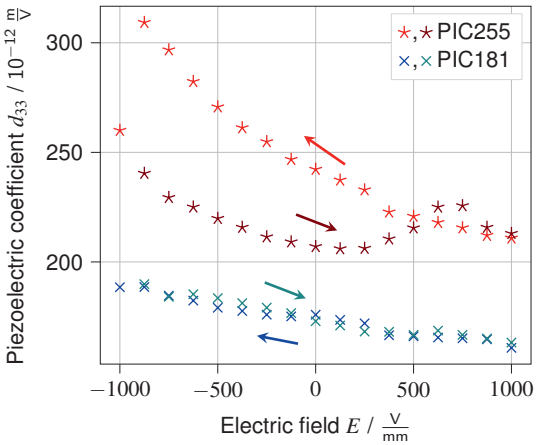


Figure 6: *Piezoelectric coefficient d_{33} of PIC255 and PIC181 at different operating points.*

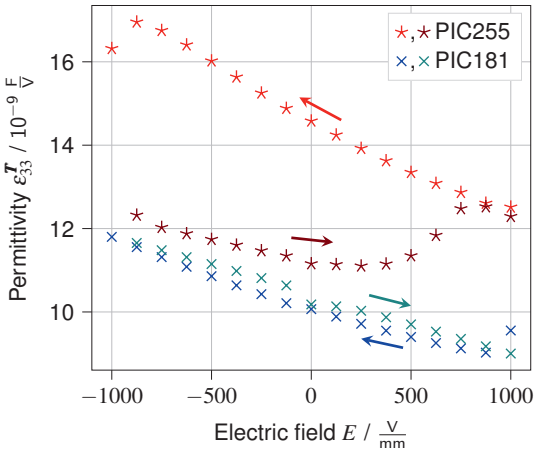


Figure 7: *Permittivity ε_{33}^{T} of PIC255 and PIC181 at different operating points.*

Although the frequency shift direction of PIC700 is opposite to the other PZT based piezoceramics, the material parameters show the same characteristics in the considered field strength range. The relation between the compliance s_{33}^{E} and the applied DC bias field of three different PIC700 samples is shown in figure 8. The parameter change is comparably small to that of PIC181. Moreover if the DC bias field decreases, s_{33}^{E} increases. This behaviour is also observed for the piezoelectric material parameter and the permittivity.

The application of a DC bias field in the direction of polarisation reinforces the remenant polarisation, which leads to a decrease in elastic compliance and the piezoelectric coefficient. According to [3] this can be explained considering the $\boldsymbol{P}$-$\boldsymbol{E}$ hysteresis (figure 1). When a DC bias field is superimposed with an

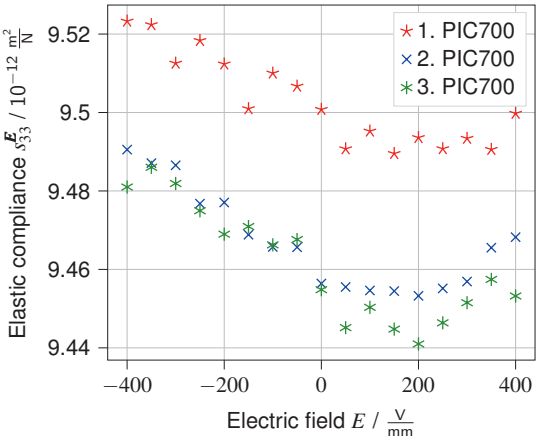


Figure 8: *Elastic compliance $s_{33}^{\boldsymbol{E}}$ of three PIC700 samples at different operating points.*

AC alternating field and applied to a piezoceramic disc, $\varepsilon_{33}^{\text{eff}}$ describes the slope of the $\boldsymbol{P}$-$\boldsymbol{E}$ field curve ($d\boldsymbol{P}/d\boldsymbol{E}$) at a variable operating point. As the DC bias field strength increases, the slope decreases, causing $\varepsilon_{33}^{\text{eff}}$ to decrease as well. Since the piezoelectric parameter $\boldsymbol{d}$ is approximately proportional to $\boldsymbol{\varepsilon^T}$, the DC bias field causes $\boldsymbol{\varepsilon^T}$ to decrease.

4.4 Non-linear parameter estimation

By using equations 5 the non-linear parameters can be estimated. Considering the maximum applicable electrical field strength of $1\,\frac{\text{kV}}{\text{mm}}$ the following values result for the piezoelectric, electrostrictive and dielectric parameter of PIC255:

$$d_{333} = -7.2 \cdot 10^{-20}\,\text{m}^3/\text{VN}$$

$$\zeta_{33} = -17.7 \cdot 10^{-19}\,\text{m}^2/\text{V}^2$$

$$\boldsymbol{\varepsilon}_{333}^{\boldsymbol{T}} = -206.2 \cdot 10^{-17}\,\text{F/V}$$

The following parameters can be estimated for PIC181:

$$d_{333} = -6.1 \cdot 10^{-20}\,\text{m}^3/\text{VN}$$

$$\zeta_{33} = -19.6 \cdot 10^{-19}\,\text{m}^2/\text{V}^2$$

$$\boldsymbol{\varepsilon}_{333}^{\boldsymbol{T}} = -51 \cdot 10^{-17}\,\text{F/V}$$

Because of the different thickness of the lead-free piezoceramic specimen a direct comparison is not possible. Moreover the observed parameter changes in the DC bias range from $-400\,\frac{\text{V}}{\text{mm}}$ to $400\,\frac{\text{V}}{\text{mm}}$ are too small for a reliable assumption. The measured frequency shift of all piezoceramic discs is comparable with similar experiments [3, 8, 9, 14].

5 Conclusions

Impedance measurements of three different piezoceramic materials with various DC bias field strengths applied were performed. A series and parallel resonance frequency shift is recognisably caused by the applied field strength. A frequency shift trend to higher frequencies with an electric field applied along the polarisation direction is detected. The general trend of the frequency shift of the lead-free piezoceramic disc is opposite to the shift of the leaded specimens. Nevertheless the parameter change of PIC700 is comparable to that of the hard PZT piezoceramic PIC181. The investigations of the material parameters shows an almost linear dependence of the parameters and the applied DC bias field, comparable with other publications in this field [3, 4, 8, 9, 14]. The linear elastic parameter $s_{33}^{\boldsymbol{E}}$, the permittivity $\varepsilon_{33}^{\boldsymbol{T}}$ and the piezoelectric constant d_{33} decrease proportionally with increasing electric field strength E. Furthermore, depolarization effects are observed at the soft PZT PIC255 by applying a DC bias field of greater than $-800\,\frac{\text{V}}{\text{mm}}$. A 1D-approximation is described to determine non-linear parameters of piezoelectric discs. Second-order constitutive equations provide results for the non-linear piezoelectric coefficient d_{333}, the permittivity $\varepsilon_{333}^{\boldsymbol{T}}$ and the electrostrictive coefficient ζ_{33}, comparable to [3].

6 Literature

[1] M. Alguero, C. Alemany, L. Pardo, A. M. Gonzalez, "Method for Obtaining the Full Set of Linear Electric, Mechanical, and Electromechanical Coefficients and All Related Losses of a Piezoelectric Ceramic", *Journal of the American Ceramic Society* 87.2, pp. 209–215 (2004); doi: 10.1111/j.1551-2916.2004.00209.x

[2] Y. Dong, Z. Wu, H. Hu, B. Wu, G. Xu, "A novel method for characterization of piezoelectric material parameters by simulated annealing optimization", *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 57.12, pp. 2613–2615 (2010); doi: 10.1109/TUFFC.2010.1735

[3] Q.-M. Wang, T. Zhang, Q. Chen, X.-H. Du, "Effect of DC bias field on the complex materials coefficients of piezoelectric resonators", *Sensors and Actuators A: Physical* 109.1, pp. 149–155 (2003); doi: 10.1016/j.sna.2003.08.008

[4] K. S. Aleksandrov , M. P. Zaitseva , A. M. Sysoev, Yu. I. Kokorin, "The piezoelectric resonator in a dc electric field", *Ferroelectrics* 41.1, pp. 3-8 (1982); doi: 10.1080/00150198208210604

[5] W. P. Mason, "Electromechanical transducers and wave filters", New York: D. Van Nostrand Co., 1948.

[6] D. S. Shim, D. A. Feld, "A general nonlinear Mason model of arbitrary nonlinearities in a piezoelectric film", *IEEE International Ultrasonics Symposium*, pp. 295-300(2010); doi: 10.1109/ULTSYM.2010.5935834

[7] American National Standards Institute, "IEEE Standard on Piezoelectricity", New York, 1987; doi: 10.1109/IEEESTD.1988.79638

[8] R. Simkovics, H. Landes, M. Kaltenbacher, R. Lerch, "Nonlinear finite element analysis of piezoelectric transducers", *IEEE Ultrasonics Symposium*

2, pp. 1057-1060 (1999); doi: 10.1109/ULTSYM.1999.849182

[9] F. Li, Z. Xu, X. Wei, X. Yao, "Temperature- and DC Bias Field-Dependent Piezoelectric Effect of Soft and Hard Lead Zirconate Titanate Ceramics", *Journal of Electroceramics* 24, pp. 294–299; doi: 10.1007/s10832-009-9571-1

[10] R. Simkovics, "Nichtlineares piezoelektrisches Finite-Elemente-Verfahren zur Modellierung piezokeramischer Aktoren", Düsseldorf: VDI-Verl, 2002.

[11] "Impedance Measurement Handbook", Santa Rosa, USA: Keysight Technologies, Nov. 2016

[12] "Piezoelectric Ceramic Products: fundamentals, characteristics and applications", Physik Instrumente (PI) GmbH & Co.KG, Dez. 2016

[13] S. Sherrit, H D. Wiederick und B.K. Mukherjee, "Non-Iterative Evaluation of the Real and Imaginary Material Constants of Piezoelectric Resonators", *Ferroelectrics* 134, pp. 111–119 (1992); doi: 10.1080/00150199208015574

[14] G. Yang, W. Ren, S.-F. Liu, A. J. Masys, B. K. Mukherjee, "Effects of uniaxial stress and DC bias field on the piezoelectric, dielectric, and elastic properties of piezoelectric ceramics", *IEEE Ultrasonics Symposium* 2, pp. 1005-1008 (2000); doi: 10.1109/ULTSYM.2000.921494

DOI 10.5162/sensoren2019/5.1.3

Torque calibration in the MN m range under rotation in a nacelle test bench

Paula Weidinger M.Sc.[1], Dr.-Ing. Gisa Foyer[2], Dipl.-Ing. Stefan Kock[3], Jonas Gnauert M.Sc.[4], Dr.-Ing. Rolf Kumme[5]

[1]Physikalisch-Technische Bundesanstalt, Department 1.2, Bundesallee 100, 38116 Braunschweig, Germany, tel.: +49 531 592-1231; e-mail: paula.weidinger@ptb.de

[2]Physikalisch-Technische Bundesanstalt, Department 1.1, Bundesallee 100, 38116 Braunschweig, Germany, tel.: +49 531 592-1145; e-mail: gisa.foyer@ptb.de

[3]RWTH Aachen University, Chair for Wind Power Drives, Campus-Boulevard 61, 52074 Aachen, Germany, tel.: +49 241 80-95654; e-mail: stefan.kock@cwd.rwth-aachen.de

[4]RWTH Aachen University, Chair for Wind Power Drives, Campus-Boulevard 61, 52074 Aachen, Germany, tel.: +49 241 80-90241; e-mail: jonas.gnauert@cwd.rwth-aachen.de

[5]Physikalisch-Technische Bundesanstalt, Department 1.2, Bundesallee 100, 38116 Braunschweig, Germany, tel.: +49 531 592-1200; e-mail: rolf.kumme@ptb.de

Abstract

Mechanical torque M is the most important parameter for the power transmission of wind turbines in general and on nacelle test benches. However, torque measurement in nacelle test benches is not yet traceable to national standards, consequently neither its precision nor its measurement uncertainty is known. To rectify this, a calibration method for torque measurement under varying rotational speed n using a torque transfer standard was developed. The result of this calibration procedure is the relative indication deviation $\bar{q}(M_L)$ between the torque measuring device in the nacelle test bench and the torque transfer standard, which amounts to about 4.3 %, and the associated expanded measurement uncertainty of less than 0.26 %.

Keywords: Torque calibration, torque transfer standard, calibration procedure, nacelle test benches

Introduction

The European commission and the European countries work together to accelerate the transition of energy production with a shift to cleaner and more efficient energy systems [12]. In 2017, the wind energy sector has overtaken coal in power generation capacity with a coverage of 11.6 % of the EU's electricity demand [8]. With a capacity installation of 15.6 GW, wind power covered 55.2 % of all renewable power installations one year later and was, thus, the leading renewable energy source [8].

To support this tendency, Nacelle Test Benches (NTBs) have been taken into operation. The crucial parameter for the power transmission of wind turbines is mechanical torque M_x, called M in the following. With wind turbines having scaled up to 4 to 10 MW, the torque input rose to a new level in the multi-MN m range. However, traceable torque measurement is only feasible up to 1.1 MN m so far [11]. Moreover, the existing calibration standards, such as DIN 51309 [4] and EURAMET cg-14 [6], are not applicable to calibrate torque measurement under rotation.

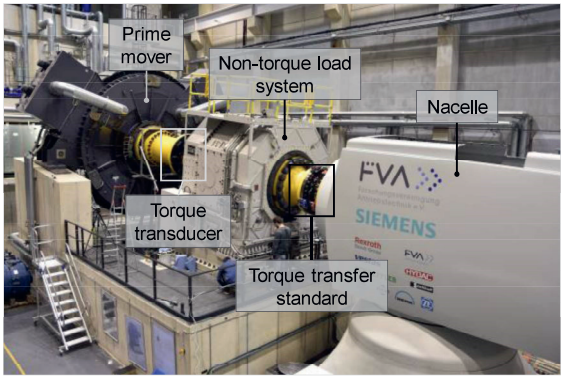


Fig. 1: Experimental set-up to calibrate the torque measurement in the NTB of the Center for Wind Power Drives (CWD) at RWTH Aachen by a TTS.

To rectify this, a procedure for torque calibration under rotation was developed [13–15]. The procedure considers two correlated input variables; torque M and rotational speed n. This paper presents the successful implementation of the calibration procedure using a 5 MN m Torque Transfer Standard (TTS) on the NTB of

the Center for Wind Power Drives (CWD) at RWTH Aachen (Fig. 1).

Experimental set-up

The calibration procedure was tested and validated on the renowned NTB of CWD Aachen. As a reference standard, the 5 MN m torque transducer, which was established to be a TTS, was used.

Nacelle test bench

The main set-up of the NTB consists of a prime mover or Load Application System (LAS), the Device Under Test (DUT), and electrical power supply equipment [2]. In the field, the wind acts on the wind turbine's rotor and causes reaction loads on the drive train of the wind turbine. In NTBs, these loads are simulated by so-called Non-Torque Load systems (NTLs).

The NTB at CWD Aachen has a nominal capacity of 4 MW generated by a direct drive with a permanent-magnet motor. Moreover, the NTB is directly coupled with the DUT. The torque is measured by an internal torque transducer with a nominal range of M_{nom} = 2.7 MN m, which is permanently installed between the prime mover and the NTL. The DUT, that is required as a counter reactor to generate torque during the calibration process, was the research nacelle owned by Forschungsvereinigung Antriebstechnik e. V. [1] with a rated capacity of P_{nom} = 2.75 MW. The maximum applicable torque load of this specific set-up amounts to $M_{max} \approx 1.5$ MN m at a minimum and a maximum rotational speed of n_{min} = 6.5 min^{-1} and n_{max} = 17.5 min^{-1} respectively.

Torque transfer standard

To calibrate the torque measurement in an NTB, it is to be compared to the torque value measured by a well-known torque transducer, which is also called TTS. A TTS transfers the precision of a torque calibration laboratory into another facility, here an NTB, and, therefore, traces the torque measurement to the national torque standard.

The TTS deployed here (Fig. 2) is a hollow shaft, multi-component transducer based on the strain gauge principle to measure forces and moments in all six degrees of freedom. As it is utilised to calibrate torque measurement, the main measuring bridge is the torque bridge measuring torque M around the main axis of the NTB's drive train (around the x-axis).

For an application under rotation, the TTS is equipped with a self-sufficient Data Acquisition (DAQ) system including batteries and a precise amplifier (MX238B, 225 Hz carrier frequency) to digitise the measuring signals and assign them with time stamps directly at the TTS.

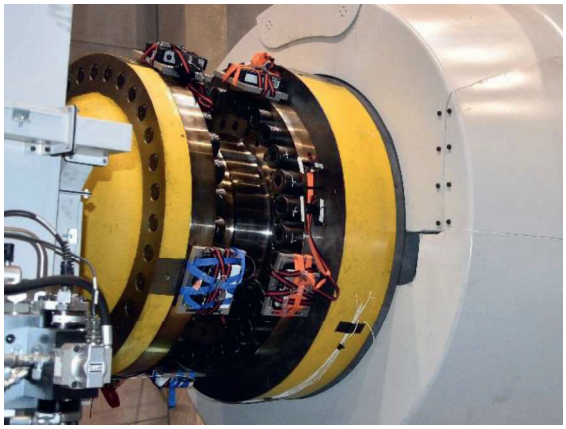


Fig. 2: TTS including DAQ and telemetry system mounted at the rotor hub flange of the nacelle under test on the NTB of CWD Aachen.

Afterwards, the signals are transmitted via WLAN through two access points to a computer. To avoid force and torque shunt and any imbalances caused by the additional platforms to mount the DAQ on the TTS, all components are disseminated symmetrically around the TTS's flanges. For an adequate resolution of 1° per revolution, the sampling frequency was set to f_{sample} = 150 Hz and the Bessel filter to f_{filter} = 50 Hz.

The 5 MN m torque transducer became a TTS by calibrating it up to 1100 kN m statically and without rotation acc. to DIN 51309 [4]. Defined within this calibration process, the coherence between the exactly applied torque steps M_p and the output signals S_p amounts to

$$M_p = 3851.1 \text{ kN m} \cdot (\text{mV/V})^{-1} \cdot S_p. \quad (1)$$

Because of the very small non-linearities and reversibility of the TTS, a linear regression curve combined for increasing and decreasing torque load was calculated. Above 1100 kN m up to the demanded 1500 kN m, the calibration results were extrapolated using a simple extrapolation method based on the calibration data gathered [10] to estimate the measurement uncertainty of the TTS (Tab. 1).

Tab. 1: Relative measurement uncertainty (k = 1) of the TTS for the relevant torque load steps.

Torque load in kN m	Rel. Uncert. $u_{cal}(M_L)$ (k = 1) in %
375	0.055
750	0.045
1125	0.044
1500	0.049

As mechanical torque is to be measured to determine the power transmission of a nacelle,

it must be measured directly at the nacelle's drive train input, which is the rotor hub flange.

Calibration procedure

The calibration procedure including the load cycle and the zero signal determination is based on common torque calibration standards, such as DIN 51309 [4] and EURAMET cg-14 [6]. However, these standards provide only methods for calibrations in a static manner without rotation. For that reason, a new procedure for torque calibration under various rotational speeds was developed [13].

Zero signal determination

The zero signal S_{zero} is the first signal taken before each load cycle without any load applied, as in common calibration standards, but under rotation. Its correctness and quality have an effect upon all measurement signals of the following load cycle as it is used to eliminate the offset of these signals.

Here, the zero signal was recorded with no load applied (M = 0 kN m), at the minimum possible rotational speed n_{min} = 6.5 min^{-1} and averaged over six full rotations (l = 6). To this end, full access to the nacelle control system was required. If this is not feasible, a static zero signal determination via relatively rotated positions could be performed (cf. [13]).

Load cycle

The sequences for the torque calibration are supplemented by varying rotational speed steps in order to analyse the coherence between both input parameters torque M and rotational speed n. Every DUT has device-specific operating points. As many operating points of DUTs tested on the NTB to be calibrated as possible must be mapped during the calibration process. However, the range of the so-called characterisation maps is confined by the boundary conditions of the DUT installed during the calibration measurements.

The characterisation maps are comparable to characterisation maps for electrical motors. In a characterisation map, several segments of the torque-rotational-speed-plane are passed through during the calibration procedure in order to issue a grid of reference-operating points, which are calibrated. This grid is stipulated by the minimum (n_{min}, M_{min}) and maximum (n_{max}, M_{max}) applicable torque M and rotational speed n. Ideally, the measurement points are equally spread (Δn and ΔM = const.) but hitting points of resonance must be prevented.

Moreover, to investigate the influence of rotational speed on the torque measurement and the mechanical hysteresis of the transducer to be calibrated, rotational speed and torque are combined in multiplex ways. As an example, the measurement values of one kind *cm2a* of a characterisation map are plotted in Fig. 3.

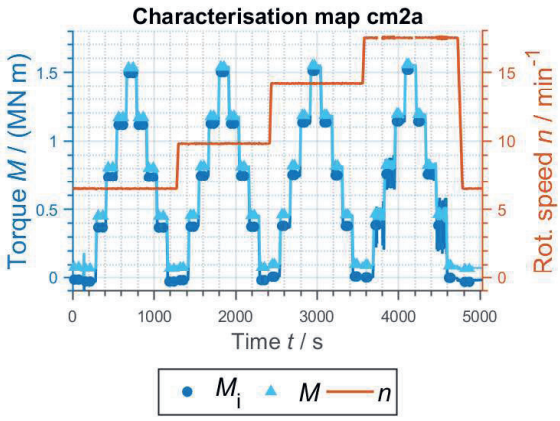


Fig. 3: Tared measurement signals of one sequence (cm2a) of the calibration procedure with varying torque (left ordinate, blue curves with symbols) and varying rotational speed (right ordinate, straight line orange curve).

The performance of the characterisation maps in the present experimental set-up was only possible due to full access to the DUT to control the torque applied. The rotational speed in turn was controlled by the NTB's prime mover.

Signal evaluation

To ensure a monitoring of all events occurring during the measurements, the data was recorded permanently. As the NTB is still calibrated per operating point (step), these points were denoted and a mean value per step over l = 6 revolutions of the drive train was build. Before doing so, a suitable dwell time per step t_{step} was satisfied, which endured for approximately 70 s per step here. As the signals were only averaged for a stationary state of the set-up, which means both M and n were constant in the aftermath of an alteration in M and/or n, t_{step} consists of different components:

$$t_{step} = t_{ramp} + t_{settling} + t_{meas}(n_{min}, l), \quad (2)$$

the time to ramp up to the predefined step t_{ramp}, the time to let the system settle down to reach a stationary state $t_{settling}$, and the actual measuring time t_{meas} depending on the minimum rotational speed n_{min} and the revolutions to average over l. The ramping up can be done fully or partially and in different ways; either in a unit step, a ramp or a sine and the ramping time varies accordingly. Moreover, the settling time varies vastly contingent on the NTB control system, which was here rather quick.

All signals per load step $\overline{S}_{meas}$ were average over an integer number l of revolutions (based

on [3]), which was here $l = 6$, a compromise between the required measurement time and a sufficient number of revolutions, once a stationary state was reached.

$$\overline{S}_{meas} = m^{-1} \cdot \sum_{i=1}^{m} S_{meas,i} \quad (3)$$

$$\text{with } m = \frac{60\ \text{s}}{n_{NTB} \cdot 1\ \text{min}}, \quad (4)$$

$l \in Z_{>0}$, here $l = 6$.

The arithmetic mean is required to minimise influences due to the horizontal installation position and due to parasitic and purposefully applied loads. In a horizontal installation position, the dead weight of the transducers themselves causes permanent load in form of bending moments on the transducer. Furthermore, misalignments along the drive train, such as eccentricities, non-parallelism of flanges, and the inclination of the entire drive train, evoke parasitic loads. These parasitic loads in combination with the loads pointedly generated by the NTL to actively control the set-up to minimise unwanted vibrations and thereby protect the drive train affect all measurements and alter for each relative rotational position.
As mentioned above, each signal $S_{NTB/TTS}$ of both transducers is tared by the zero signal of the respective load cycle:

$$S_{NTB/TTS} = \overline{S}_{meas,\,NTB/TTS} - S_{zero,\,NTB/TTS}. \quad (5)$$

Before gathering any measurement data, all electrical components were pre-heated to avoid creep due to thermal alterations in the signal amplifiers.

Calibration results

The evaluation of the calibration results is based on ISO 7500-1 [5], which standardises the calibration of force measuring systems in uniaxial material testing machines by reference force transducers. This standard was assimilated to the set-up at hand: a uniaxial torque calibration with parasitic loads. The main difference between the cases of application besides the measurement parameter calibrated, force vs. torque, is the measurement under static conditions vs. stationary rotation respectively. The content of the following section is based on ISO 7500-1.
The calibration result of torque calibration in NTBs consists of:

- the relative indication deviation $\overline{q}(M_L)$ per load step M_L,
- the expanded measurement uncertainty $U(M_L)$ for the indication deviation, and
- the relative reversibility.

Here, the calibration result is presented in form of a look-up table that can be consulted to correct the measurement result in a post-processing.

Measurement conditions

In general, this measurement set-up is very complex due to the additional mechanical loads exerted on the drive train as stated in [7], therefore, the measurement conditions to be considered are wide. To avert signal creep due to thermal effects, all electrical components were pre-heated prior to the measurements. However, general ambient conditions needed to be taken into account. Humidity influences can be neglected as the strain gauges are encapsulated inside the transducers, so can the air pressure for strain-gauge-based measurement tools. Common calibration standards stipulate a temperature stability of ± 1 K and within 18 to 28 °C during the measurements. Because of the manifold heat sources in an NTB such as gearbox, generator, prime mover, and NTL, the temperature around the NTB can fluctuate. These fluctuations were permanently monitored to detected that the temperature was stable to $\Delta T_{loadcycle} = \pm 1$ K during the measuring process. Compared to the conditions under which the TTS was calibrated, the temperature in the NTB was + 5 K higher. However, the TTS signal was not corrected for this temperature difference, but the temperature dependence of the TTS's sensitivity was considered in the measurement uncertainty.
Another condition to be considered are emergency brakes, which can result in a torque inversion of up to 80 % of the maximum torque applied beforehand. To minimise the hysteresis effect evoked thereby, the NTB was operated at maximum torque load for about 5 min after an emergency brake before carrying on with the measurements.
Moreover, parasitic and additional loads, as mentioned before, can have a crosstalk effect on the torque signal. Crosstalk means, the applied loads do not only have an influence on the intended measuring bridges, but also on other bridges [9].

Relative indication deviation

The relative indication deviation $q_j(M_L)$ is the deviation of the torque value M_i measured directly in the NTB from the torque M indicated by the TTS relative to the TTS torque:

$$q_j(M_L) = \frac{M_{i,j}(M_L) - M_j(M_L)}{M_j(M_L)} \cdot 100\ \%. \quad (6)$$

here $j = 1,2$

The indication deviation is calculated separately for each load step M_L and for each of the two repetitions $j = 1,2$.

For the result of the torque calibration, the relative indication deviation $\bar{q}(M_L)$, the indication deviation is averaged over the two repetitions *j*:

$$\bar{q}(M_L) = n^{-1} \sum_{j=1}^{n} q_j(M_L), \quad (7)$$

here for *cm2a* $n = 2$.

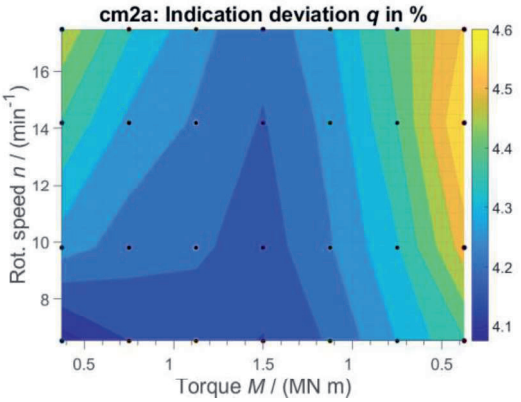


Fig. 4: Calibration result in form of a relative indication deviation $\bar{q}(M_L)$ between the torque transducer in the NTB and the TTS.

Fig. 4 and Tab. 2 reveal a dependence of the relative indication deviation not only on the torque applied, but also on the rotational speed prevailing. For the higher torque load steps, the relative indication deviation varies less with the rotational speed due to the better resolution and the advantages accompanied by that. However, extreme caution

Tab. 2: Relative indication deviation $\bar{q}(M_L)$ in % between the torque value measured by the transducer in the NTB and the torque value measured by the TTS.

Rel. indication deviation $\bar{q}(M_L)$ in %				
$M_L\downarrow$ $n\rightarrow$	**6.5**	**9.8**	**14.2**	**17.5**
375_{up}	4.075	4.275	4.395	4.488
750_{up}	4.150	4.224	4.287	4.355
1125_{up}	4.167	4.210	4.244	4.266
1500	4.148	4.173	4.196	4.216
1125_{down}	4.207	4.246	4.266	4.293
750_{down}	4.302	4.344	4.399	4.424
375_{down}	4.396	4.526	4.608	4.556

Relative reversibility

When increasing M_i and decreasing M_i' torque are observed separately, the relative reversibility *v* for decreasing load can be calculated. Necessary to that end, the TTS must be calibrated for increasing and decreasing load, which was the case here. For the reversibility measurement, one decreasing load cycle is functionally adequate. However, the reversibility highly depends on the maximum torque applied during the load cycle. Because of a non-linear scaling of the reversibility, it cannot be extended to arbitrary load cycles.

The relative reversibility is calculated based on the divergence between the increasing and decreasing load steps (cf. Tab. 3), which were interpolated to matchable full load steps for a comparability.

$$v(M_L) = \frac{M_{i,interpol}'(M_L) - M_{i,interpol}(M_L)}{M_{interpol}(M_L)} \cdot 100\ \%. \quad (8)$$

The relative reversibility can be vaguely divined by the colours in Fig. 4. Tab. 3 shows that the relative reversibility is more significant for a rotational speed of $n = 6.5$ min^{-1}. Moreover, the absolute reversibility decreases distinctly for increasing rotational speed.

Tab. 3: Relative reversibility $v(M_L)$ in % between increasing and decreasing torque load based on interpolated load steps.

Relative reversibility $v(M_L)$ in %				
$M_L\downarrow$ $n\rightarrow$	**6.5**	**9.8**	**14.2**	**17.5**
1125_{down}	0.040	0.000	0.000	0.000
750_{down}	0.151	0.001	0.001	0.001
375_{down}	0.321	0.003	0.002	0.001

Measurement uncertainty

The measurement uncertainty $u_c(M_L)$ consists of three uncertainty contributions:

- the uncertainty contribution of the resolution $u_{res}(M_L)$,
- the uncertainty contribution of the repeatability $u_{rep}(M_L)$, and
- the uncertainty contribution of the transfer standard including additional uncertainty contributions $u_{std}(M_L)$.

The uncertainty contribution of the relative resolution per load step $u_{res}(M_L)$ is the square root of the sum of the squares of the NTB transducer's relative resolution per load step $a_M(M_L)$ and its relative resolution after load release $a_Z(M_L)$. The relative resolution after load release is to be considered, as every torque value per load step is tared using the zero signal. For both contributors, a rectangular distribution ($2\sqrt{3}$) is assumed.

$$u_{res}(M_L) = \sqrt{\left(\frac{a_M(M_L)}{2\sqrt{3}}\right)^2 + \left(\frac{a_Z}{2\sqrt{3}}\right)^2}. \quad (9)$$

The relative resolution is the indication fluctuation of the NTB transducer divided by the respective load step. Hereby, the indication fluctuation is the difference between the minimum and maximum mean per revolution for the averaging sequence of six revolutions.

As a second uncertainty contributor, the relative repeatability is calculated in form of the

standard deviation of the best expected value q_j and the relative mean indication deviation $\bar{q}$, where n (here n = 2) is the number of repetitions per load step:

$$u_{rep}(M_L) = \sqrt{(n(n-1))^{-1} \sum_{j=1}^{n} (q_j(M_L) - \bar{q}(M_L))}. \quad (10)$$

As a last component, the uncertainty contribution of the TTS $u_{cal}(M_L)$, which is assigned in the calibration certificate, is to be considered in the total measurement budget:

$$u_{std}(M_L) = \sqrt{(u_{cal}(M_L))^2 + A^2 + B^2 + C^2}. \quad (11)$$

Moreover, other uncertainty contributors were allowed for at this point: the composition of the measuring chain in form of the measurement uncertainty of the amplifier deployed as contributor *A*, the relative drift of the TTS over time as contributor *B*, and the temperature deviation between the temperature in the NTB and the temperature in the laboratory per 5 K, where the TTS was calibrated, as contributor *C*. These contributors are listed in Tab. 4.

Tab. 4: Additional uncertainty contributors considered in the total measurement uncertainty (cf. [14]).

Uncertainty contributions		Value in %
A	Amplifier calibration	1.20×10^{-5}
B	Drift	8.85×10^{-3}
C	Temperature influence	1.52×10^{-2}

The expanded absolute measurement uncertainty $U(M_L)$ for the relative indication deviation is the product of the coverage factor k (here k = 2, p = 95.45 %, normal distribution) and the combined absolute uncertainty $u_c(M_L)$:

$$U(M_L) = k \cdot u_c(M_L), \quad (12)$$

$$U(M_L) = k \cdot \sqrt{(u_{res}(M_L))^2 + (u_{rep}(M_L))^2 + (u_{std}(M_L))^2}.$$

Tab. 5: Absolute expanded measurement uncertainty U (k = 2) for the relative indication deviation $\bar{q}(M_L)$ in %.

Exp. rel. uncert. $U(M_L)$ (k = 2) in %				
$M_L \downarrow$ $n \rightarrow$	6.5	9.8	14.2	17.5
375$_{up}$	0.199	0.235	0.249	0.260
750$_{up}$	0.180	0.191	0.206	0.223
1125$_{up}$	0.178	0.185	0.178	0.201
1500	0.182	0.183	0.182	0.188
1125$_{down}$	0.182	0.186	0.183	0.183
750$_{down}$	0.196	0.189	0.203	0.188
375$_{down}$	0.231	0.241	0.255	0.241

As expected, the expanded measurement uncertainty (Tab. 5) decreases for increasing torque load up to 1100 kN m due to a better signal-noise-ratio, and a poor resolution and a worse repeatability in the lower measurement ranges. Above 1100 kN m, the measurement uncertainty increases because of the soared measurement uncertainty of the TTS evoked by the extrapolation of the calibration results.

Influence of rotational speed and misalignments

The influence of the rotational speed can be seen not only in the torque signals, but also in the additional loads on the TTS. In general, the signals of the additional loads are periodically recurrent.

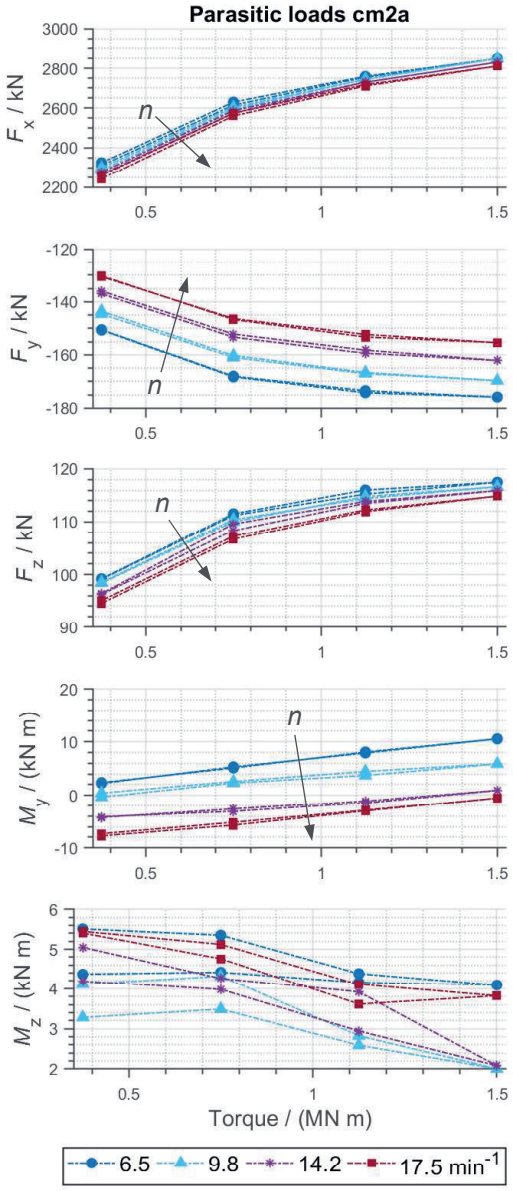


Fig. 5: Averaged parasitic loads on the TTS during the calibration measurements separately plotted for the different rotational speeds applied.

To analyse the influence of the torque alterations, which allows for investigations of misalignments, and of the rotational speed, the

additional loads are averaged over the same sequence of six revolutions as the torque signals. These non-tared mean signals depending on the torque load step in combination with the rotational speed step are depicted in Fig. 5. Due to the fact, that no other measuring bridges than the torque bridge of the TTS are calibrated, only qualitative evaluations can be performed. For example, the thrust (axial force F_x) increases with increasing torque load and decreases slightly with increasing rotational speed. Moreover, the significant thrust offset of F_x = 2757.87 kN is a pointer to lateral bracing in the drive train. However, caution must be taken when analysing these signals any further, as the signals are very small measured by bridges conceptualised for much broader measuring ranges of several thousand kN and kN m. Therefore, further investigations with specific additional loads are to be performed in the future.

Unbalances in the drive train, that might affect the torque signals, were analysed by an FFT. Fig. 6 shows the power spectral density of both torque signals, M taken by the TTS and M_i recorded by the NTB transducer. The FFT is performed on the averaging sequence of the torque signal at the decreasing torque load step of M_L = 375 kN m at a rotational speed of n = 17.5 min^{-1}, where vibrations can already be seen in Fig. 3. These vibrations appear in form of a peak at 22.75 Hz in the power density spectrum in Fig. 6. In all other signals sequences per load steps, no peaks were revealed, which implies that there are no unbalances relevant at the prevailing rotational speeds.

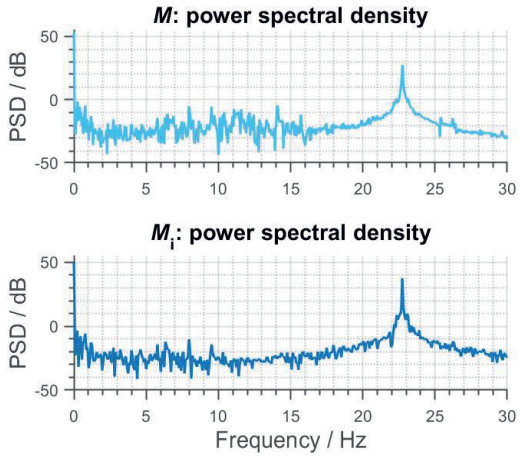


Fig. 6: Power spectral density of the two torque signals measured by the TTS (M) and by the NTB (M_i) at 375 kN m decreasing and 17.5 min^{-1}.

Conclusion and outlook

This paper presents the first attempt of calibrating the torque measurement in an NTB using one type of a characterisation map. It was found that the position of the reference transducer, the TTS, is important: it must be placed directly at the input of the nacelle, which is the nacelle rotor hub flange. Moreover, the measurement procedure for the calibration and its evaluation including the calculations of the expanded measurement uncertainty are introduced.

The calibration result is the relative indication deviation $\bar{q}(M_L)$ between the torque value measured by the NTB transducer and the torque indicated by the TTS, which is the reference torque, relative to this reference torque value. This work has revealed that the torque value and, therefore, the relative indication deviation, depend on the rotational speed, which is why the torque is to be calibrated in combination with rotational speed. The averaged relative indication deviation amounts to $\bar{q}(M_L)$ = 4.3 % and an associated expanded measurement uncertainty of $U(M_L)$ (k = 2) ≤ 0.26 % for this set-up. The results are available in form of a look-up table. In the future, it could also be possible to calculate multiple regression curves that can be used to correct the measurements in-situ after a calibration. Additionally, further future work will consider a representative amount of repetitions for each load step.

Acknowledgements

Many thanks are due to Andreas Brüge and Dirk Röske for the constructive discussions about torque calibration. Moreover, all authors would like to thank Kai Geva, Stefan Augustat, Sebastian Reisch, and Michael Pagitsch for their support in carrying out the measurements.

All of the authors would like to acknowledge the funding of the Joint Research Project "14IND14 MNm Torque – Torque measurement in the MNm range". This project has received funding from the EMPIR programme, which is co-financed by the European Union's Horizon 2020 research and innovation programme, and from EMPIR Participating States.

The lead author gratefully acknowledges the support of the Braunschweig International Graduate School of Metrology B-IGSM.

References

[1] https://www.cwd.rwth-aachen.de/projekte/fva-gondel/. Accessed 22 February 2018.

[2] Averous, N. R., Stieneker, M., Kock, S., Andrei, C., Helmedag, A., De Doncker, R. W., Hameyer, K., Jacobs, G., and Monti, A. 2015. Development of a 4 MW full-size wind-turbine

test bench. In 6th International Symposium on Power Electronics for Distributed Generation Systems, Aachen.
[3] Brüge, A. and Pfeiffer, H. A standard for rotatory power measurement. not yet published.
[4] DIN 51309:2013-09. Material testing machines - Calibration of static torque measuring devices. Technical Committee Drehmoment of Deuscher Kalibrierdienst.
[5] DIN 7500-1:2014. Metallische Werkstoffe - Prüfung von statischen einachsigen Prüfmaschinen - Teil 1: Zug und Durckprüfmaschinen - Prüfung und Kalibrierung der Kraftmesseinrichtung.
[6] EURAMET cg-14. 2011. Guidelines on the Calibration of Static Torque Measuring Devices.
[7] Foyer, G. and Kock, S. 2017. Measurement uncertainty evaluation of torque measurement in nacelle test benches. In IMEKO. TC3, TC5 and TC22 International Conference.
[8] Fraile, D. and Mbistrova, A. 2018. Wind power 2017. Annual combined onshore and offshore wind energy statistics. https://windeurope.org/wp-content/uploads/files/about-wind/statistics/WindEurope-Annual-Statistics-2017.pdf. Accessed 10 January 2019.
[9] Keil, S., Ed. 2016. Dehnungsmessstreifen. 2. Auflage. Springer Fachmedian, Wiesbaden.
[10] P Weidinger, G Foyer, J Ala-Hiiro, and C Schlegel and R Kumme. 2018. J. Phys.: Conf. Ser. 1065 042057, doi: 10.1088/1742-6596/1065/4/042057.
[11] Peschel, D., Mauersberger, D., Schwind, D., and Kolwinski, U. 2005. The new 1.1 MN·m torque standard machine of the PTB Braunschweig/ Germany. In IMEKO TC3 International Conference on Force, Mass and Torque.
[12] Quental, N., Buttle, D., Abrar, S., Firkaviciuté, K., Jimenez Mingo, J., Sofianopoulos, D., Kontoudakis, A., and Sales Agut, C. 2017. The strategic energy technology (SET) plan. At the heart of energy research and innovation in Europe. 2007-2017 SET plan 10th anniversary. https://publications.europa.eu/en/publication-detail/-/publication/771918e8-d3ee-11e7-a5b9-01aa75ed71a1/language-en/format-PDF/source-51344538. Accessed 10 January 2019.
[13] Weidinger, P. and Foyer, G. 2019. Traceable Torque Measurement under Rotation in Nacelle Test Benches. A Good Practice Guide, doi: 10.7795/530.20190111.
[14] Weidinger, P., Foyer, G., Kock, S., Gnauert, J., and Kumme, R. Calibration of torque measurement under constant rotation in a wind turbine test bench. Journal of Sensors and Sensor Systems, doi: 10.5194/jsss-8-149-2019.
[15] Weidinger, P., Foyer, G., Kock, S., Gnauert, J., and Kumme, R. 2018. J. Phys.: Conf. Ser. 1037, 52030, doi: 10.1088/1742-6596/1037/5/052030.

DOI 10.5162/sensoren2019/5.1.4

Bestimmung des Brechungsindex des Schmiermittels Beruforge 152DL unter Verwendung eines dafür entwickelten Dünnschichtnormals

Sebastian Metzner[1], Tamara Reuter[1], Elisa Wirthmann[1], Tino Hausotte[1]
[1]Lehrstuhl für Fertigungsmesstechnik, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nägelsbachstraße 25, 91052 Erlangen, Deutschland

Zusammenfassung

Für eine Messung von blechmassivumgeformten Bauteilen möglichst nahe am Fertigungsprozess spielen systematische Abweichungen durch den darauf noch verbliebenen Schmierstoff eine signifikante Rolle. Die erwartete Schichtdicke dieses Schmierstoffs liegt bei den betrachteten Bauteilen in einem Bereich von unter 35 µm. Für die Bestimmung des Brechungsindex des Schmiermittels und somit die Auswirkungen des Schmierstoffs auf die interferometrische Messung, wird ein Dünnschichtnormal vorgestellt, welches in einem Bereich von 6 bis 100 µm einen kontinuierlichen Messbereich abbildet. Mit dem eingesetzten Messverfahren, einer koaxialen interferometrischen Schichtdickenmessung, können so in verschiedenen Abständen die optische Dicke bestimmt werden. Daraus lässt sich im Anschluss der Brechungsindex näherungsweise bestimmen. Zudem können die Abweichungen der angezeigten Schichtdicke bei bekanntem Brechungsindex nun bewertet werden.

Keywords: Brechungsindex, Interferometrie, Schmiermittel Beruforge 152DL, Dünnschichtstufennormal

Einleitung

Die Blechmassivumformung (BMU) als neuartiges Fertigungsverfahren stellt die optische Messtechnik vor Herausforderungen [1]. Vor allem bei Messungen möglichst nahe am Fertigungsprozess beeinflussen Faktoren wie das verwendete Schmiermittel das Messergebnis. Typischerweise wird bei der BMU das Schmiermittel Beruforge 152DL (BF152DL) verwendet. Da die Auswirkungen des Schmiermittels auf optische Messsysteme meist die Kenntnis des Brechungsindex voraussetzen, ist dessen Bestimmung für weitere Untersuchungen notwendig.

Aktueller Forschungsstand zur Messung dünner Schichten

Bisher werden hauptsächlich Messsysteme, welche auf Ellipsometrie oder Reflektometrie basieren, für die Schichtdickenmessung und somit auch für eine mögliche Brechungsindexbestimmung eingesetzt. Die Grundlagen hierfür werden von McCrackin [2] und Hlubina [3] beschrieben. Aufgrund der langen Messdauer dieser Methoden werden jedoch zunehmend interferometrische Messsysteme verwendet [4]. Grundlage für die Bestimmung des Brechungsindex ist die geometrische Schichtdickenmessung. Bei der interferometrischen Schichtdickenmessungen wird zunächst zwischen Einschicht- und Mehrschichtsystemen unterschieden. Bei Mehrschichtsystemen erfolgt die Trennung der Signale der einzelnen Schichten meist mittels Fourier Analyse unter Anwendung des Algorithmus der Fast Fourier Transformation (FFT). Quinten [5] beschrieb die Grenzen der FFT bei Schichtdickenmessungen durch Alias- und Leakage-Effekte. Diese Einflüsse können durch eine Korrektur reduziert aber nicht vollständig vermieden werden. Ghim und Kim [6] zeigten eine verbesserte Version der dispersiven Weißlichtinterferometrie, die es ermöglicht, den Brechungsindex ohne Kenntnis der geometrischen Dicke zu messen. Die Schichtdicke wird dabei aus den spektral aufgelösten Interferenz-Signalen gewonnen. Cardin und Leduc [7] nutzten zur Brechungsindexbestimmung unter anderem die Prismen-FilmKoppler-Spektroskopie, welche auf der Verwendung von Transfermatrizen basiert und gleichzeitig die Bestimmung des Brechungsindex, der Schichtdicke und die optischen Verluste der gemessenen Schicht zulässt. Die von Metzner in [8] vorgestellten Themen zur Auswirkung von dünnen Schmierstoffschichten auf Streifenlichtprojektionssysteme zeigen eine Abweichung der Messergebnisse in einem betrachteten Bereich von

10 µm bis 100 µm. Ab einer Schichtdicke von 40 µm konvergiert die gemessene systematische Abweichung gegen diesen Wert, da die Strahlen der Streifenlichtprojektion im inneren der Schicht teilweise diffus reflektiert oder absorbiert werden.
In der vorliegenden Arbeit werden neben der Bestimmung des Brechungsindex und der Abweichungen zur geometrischen Schichtdicke auch die Grenzen des Systems bei dem eingesetzten Schmierstoff aufgezeigt.

Verwendetes Messsystem und entwickeltes Dünnschichtnormal

Für die eigentliche Dünnschichtmessung wird das Prinzip der koaxialen interferometrischen Schichtdickenmessung verwendet. Der verwendete Sensor CHRocodile 2S der Firma Precitec erreicht bei einem Messfleckdurchmesser von 13 µm eine vertikale Auflösung von 10 nm und ist für einen Messbereich von 3 µm bis 180 µm optischer Dicke ausgelegt. Die Strahlen treffen bei diesem Messprinzip lotrecht auf die transparente Schicht und werden dort teilweise reflektiert, teilweise durchdringen sie die Schicht und werden an der darunterliegenden Fläche reflektiert. Durch das polychromatische Licht variiert die Phasenverschiebung zwischen den reflektierten Teilstrahlen mit der Wellenlänge, wodurch konstruktive und destruktive Interferenz entsteht. Wird die Intensität über die Wellenzahl aufgetragen entsteht eine harmonische Funktion. Die Frequenz dieser Funktion ist proportional zur optischen Weglänge der Schicht. Durch die bereits beschriebene Anwendung der Fourier Analyse ist es möglich, bei mehreren Schichten die Funktionen zu trennen und die Frequenzen zu bestimmen. Daraus kann anschließend bei bekannter Brechzahl die einzelnen Schichtdicken bei einem Mehrschichtsystem errechnet werden [9].

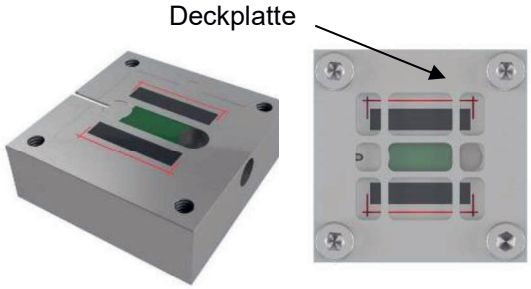


Abb. 1: Dünnschichtnormal als CAD-Modell mit den farblich markierten Funktionsflächen

Zur Abbildung der erwarteten Schichtdicken in einem definierten Bereich wurde angelehnt an DIN 32567 und DIN 5436 für die Messung von Dünnschichten ein Dünnschichtnormal entwickelt. Das Dünnschichtnormal, welches in Abbildung 1 dargestellt ist, wurde aus Edelstahl mit einer Abmessung von 30 mm x 30 mm x 10 mm realisiert. Zur Messung der Schichtdicken wurde in der Mitte des Modells eine Keilform gefräst (grün). Der Keil deckt einen Dickenbereich von 6 µm bis 100 µm ab. Die absoluten Schichtdicken werden im Vergleich zu den beiden Bezugsebenen (schwarz) bestimmt, die neben dem Keil angeordnet sind. Die Zuleitung des Schmierstoffes erfolgt über ein Loch an der Stirnseite des Keils, in das ein kurzer Zuleitungsschlauch mit Verschluss angebracht wird. Referenzlinien (rot) gewährleisten eine reproduzierbare Messung auf den verschiedenen Positionen des Keils.
In Abbildung 2 ist das reale Dünnschichtnormal, das mit dem Schmierstoff BF152DL gefüllt ist, dargestellt. Eine Deckplatte drückt eine dünne Glasscheibe auf das Modell, wodurch ein Austreten des Schmierstoffs an den Seiten des Keils verhindert wird.

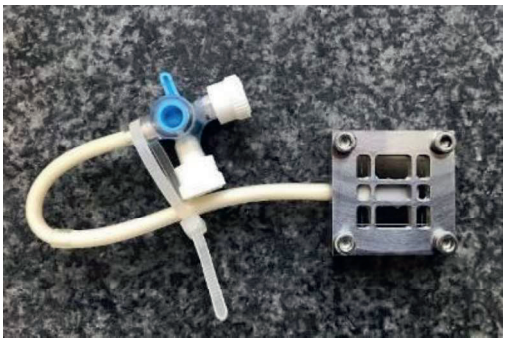


Abb. 2: Mit Schmierstoff gefülltes Dünnschichtnormal mit montierter Glas- und Deckplatte sowie Zuflussschlauch.

Durch die Verstrebungen und der Verwendung einer 100 µm dicken Deckplatte wird außerdem ein Aufwölben der Deckschicht verhindert. Vergleichsmessungen mit einer dünnen Folie zeigen eine deutliche Aufwölbung und dadurch eine Verfälschung der Messergebnisse.

Untersuchter Schmierstoff Beruforge 152DL

Der untersuchte Schmierstoff ist ein Einschichtschmierstoffsystem welcher aus anorganischen und organischen Komponenten besteht. Der Schmierstoff der auf Basis von Wasser ist, ist mit Wachspartikeln versehen, um die Reibung während der Umformung positiv zu beeinflussen. Bei ausreichendem Druck platzen diese Partikel auf und verringern dabei die Reibung während des Umformprozesses. Vor allem bei großen Prozesskräften und unterschiedlichen

Umformgraden wie es bei der BMU der Fall ist, ist eine druckabhängige Steuerung der tribologischen Eigenschaften sinnvoll. Nach der Umformung wird der Schmierstoff wieder zähflüssiger und führt bei der optischen Messung zu systematischen Messabweichungen in Form der Schichtdicke des Schmierstoffs. Die erwarteten Schichtdicken des Schmierstoffs liegen in der hier vorgestellten Anwendung der BMU in einem Bereich von 1 µm bis 35 µm.

Vorgehen zur Bestimmung des Brechungsindex und eingesetzter Messaufbau

Die Bestimmung des Brechungsindex erfolgt in Anlehnung an DIN 32567-5. Die geometrische Dicke d_{geo} ergibt sich dabei aus der Division aus der geometrischen Dicke d_{opt}.und dem Brechungsindex n:

$$d_{geo} = \frac{d_{opt}}{n} \quad (1)$$

Bei bekannter optischer und geometrischer Dicke kann der Brechungsindex berechnet werden. Die Bestimmung der optischen Schichtdicke mit dem wellenlängenabhängigen Brechungsindex wird dabei mit folgender Formel realisiert:

$$d_{opt} = \frac{m}{2 \cdot (v_{max} - v_{min})} \quad (2)$$

mit minimaler und maximaler Wellenzahl v_{min}/ v_{max} und Pixelnummer der Spektralanalyse m. Die Wellenzahl ist als Inverse zur Wellenlänge definiert. Um die geometrische Dicke an dem entwickelten Dünnschichtnormal zu bestimmen, wird das Stufennormal zunächst mit dem kalibrierten Fokus-Variationsmesssystem Alicona InfiniteFocus gemessen. Abbildung 3 zeigt das geometrische Modell. Die Referenzlinien an den Seitenflächen des Modells sind dabei gut zu erkennen. Dadurch ist ein wiederholtes Messen der Schichtdicke an verschiedenen Positionen des Keils möglich.
Die möglichen Schichtdicken auf dem Keil liegen in einem Bereich von 6 µm bis 100 µm. Die geometrische Dicke des Deckglases wird mit 100 µm angegeben.
Die Versuche werden auf der am Lehrstuhl vorhandenen Nanopositionier- und Nanomessmaschine SIOS NMM-1 durchgeführt. Mithilfe dieses Systems kann der Sensor mit einer Positionierunsicherheit von wenigen Nanometern zu dem Modell positioniert werden. Zudem ist eine Synchronisierung der Maschine mit dem Messsignal möglich, wodurch eine Abtastung der Oberfläche im Positionierbereich denkbar ist.

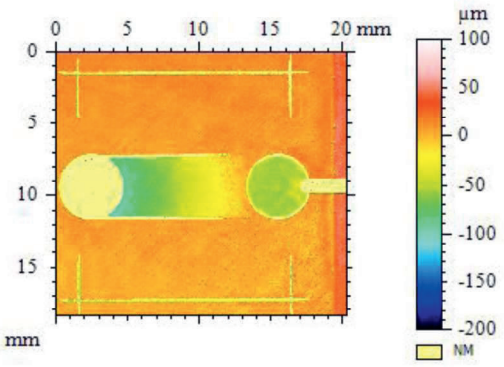


Abb. 3: Oberflächenmodell des Stufennormals gemessen mit dem Alicona InfiniteFocus

Abbildung 4 zeigt den Messaufbau auf der Nanomessmaschine mit dem Dünnschichtnormal. Durch die Referenzmarker ist eine wiederholbare Positionierung des Messsystems auf den verschiedenen Keilpositionen möglich.

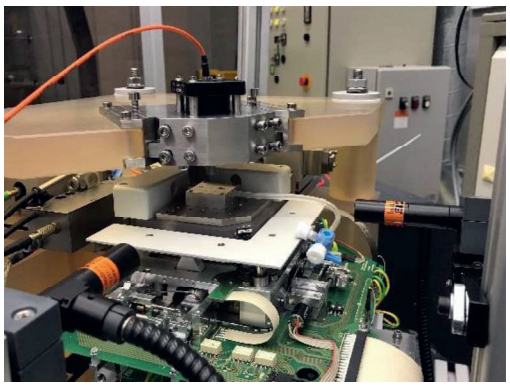


Abb. 4: Nanomessmaschine NMM-1 mit integrierten interferometrischen Messkopf und Stufennormal

Ergebnisse der Brechungsindexbestimmung bei der Dünnschichtmessung

Da der Brechungsindex noch unbekannt ist, werden zunächst Wiederholmessungen zur Bestimmung der optischen Dicke durchgeführt. Bei der Beurteilung des Messgerätes werden Messungen mit Luft und mit Wasser als Referenzschichten mit jeweils 5 Wiederholungen in einem Bereich von 6 µm bis 30 µm durchgeführt. Abbildung 6 zeigt die Abweichung der Schichtdicken in Bezug zur geometrischen Dicke an. Der Brechungsindex wird dabei mit 1,33 bei Wasser und mit 1,00 bei Luft angenommen. Die Standardabweichung ist dabei bei Wasser, unabhängig von der Schichtdicke größer als bei der Messung der Luftschicht.

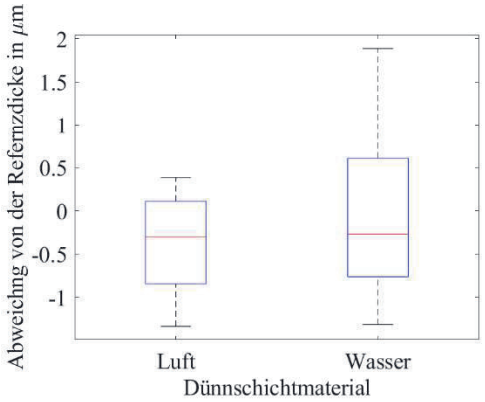


Abb. 6: Abweichungen zur absoluten Schichtdicke bei Luft und Wasser als Medium

Gründe für die erhöhte Standardabweichung ist die Wellenlängenabhängigkeit des Brechungsindex. Dieser wird aufgrund der Anwendung der FFT als konstant angenommen, ist aber abhängig von der Wellenlänge. Dies kann zu Abweichungen in der Bestimmung der Schichtdicke führen.
Die Bestimmung der optischen Schichtdicke mit dem wellenlängenabhängigen Brechungsindex würde mit folgender Formel erfolgen:

$$d_{fft} = \frac{m}{2 \cdot \left(\frac{n(\lambda_{min})}{(\lambda_{min})} - \frac{n(\lambda_{max})}{(\lambda_{max})}\right)} \qquad (3)$$

mit $\lambda_{min}/\lambda_{max}$ als minimale und maximale Wellenlänge des Weißlichtspektrums. Durch das verwendete Weißlichtinterferometer ist jedoch eine Berechnung des wellenlängenabhängigen Brechungsindex nicht möglich.
Bei intransparenten Schmierstoffen ist zusätzlich der Effekt der Absorption zu beachten. Wird von einem absorbierenden Material ausgegangen, so wird der komplexe Brechungsindex N mit dem Extinktionskoeffizienten k berechnet [10]:

$$N = n + jk \qquad (4)$$

Die insgesamt sehr geringen Abweichungen zeigen jedoch unabhängig von der Wellenlängenabhängigkeit des Brechungsindex die Fähigkeit des Messsystems in dem definierten Messabstand die genauen Höhenwerte zu ermitteln.
Abbildung 7 zeigt den nach (1) berechneten Brechungsindex für Luft, Wasser und dem Schmierstoff BF152DL in Abhängigkeit der absoluten Schichtdicke.

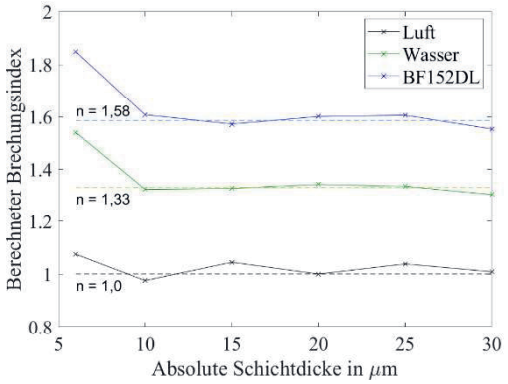


Abb. 7: Abhängigkeit des berechneten Brechungsindex in Abhängigkeit der gemessenen Schichtdicke

Dabei fällt die systematisch erhöhte Abweichung bei einer absoluten Schichtdicke von 6 µm auf. Eine Bestimmung des Brechungsindex kann nach DIN 32567-5 bei kleinen Schichtdicken zu einer Abweichung des Brechungsindex vom wahren Wert führen. Die Ergebnisse der hier aufgezeigten Messungen bestätigen diese These. Die Brechzahl des Schmiermittels konvergiert erst ab einer Schichtdicke von 10 µm gegen die erwartete Brechzahl. Der Brechungsindex des Schmierstoffes wird daher mit 1,585 ± 0,079 bestimmt.

Ergebnisse bei der Messung der absoluten Schichtdicke des Schmierstoffes

Nachdem der Brechungsindex bestimmt wurde, ist eine Messung des gesamten Dickenbereichs des Keils möglich. Dazu werden Wiederholmessungen mit den drei Schichtarten Luft, Wasser und Schmierstoff durchgeführt. Abbildung 8 zeigt die Abweichungen der gemessenen und berechnetet Schichtdicken von der absoluten Schichtdicke.

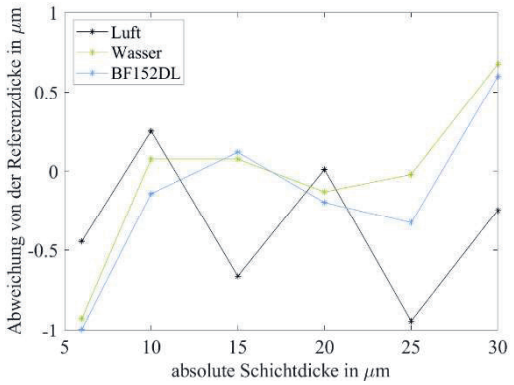


Abb. 8: Abweichungen zur absoluten Schichtdicke bei Luft, Wasser und BF152DL als Medium

In der Abbildung ist eine leicht negative Abweichung zur absoluten Schichtdicke zu erkennen. Die Schichtdicken Wasser und BF152DL haben dabei einen ähnlichen Verlauf mit einer großen Abweichung bei einer Schichtdicke von 6 µm. Diese Abweichung entspricht bei einem Brechungsindex > 0 aber der erwartenden Abweichung bei dünnen Schichten. Abbildung 9 stellt die Standardabweichungen bei den einzelnen Widerholmessungen im gesamten Dickenbereich des Keils dar.

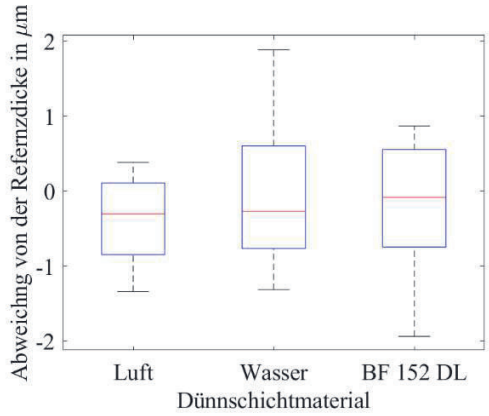


Abb. 9: Abweichungen zur absoluten Schichtdicke bei Luft und Wasser als Medium

Dabei sind die Standardabweichungen bei Wasser und BF152DL im Vergleich zur Messung der Schichtdicke mit Luft erhöht. Die erhöhten Streuwerte sind durch die unterschiedlichen Reflexionen der Lichtstrahlen an der Grenzfläche zwischen den Schichten zu erklären. Grundsätzlich weisen die Ergebnisse die erwartete Abhängigkeit der Schichtdicke von der geometrischen Höhe des Keils mit einer mittleren Abweichung von unter 1 µm auf. Die Standardabweichungen der Schichtart Luft liegen unter 0,5 µm. Im Vergleich dazu liegen die Standardabweichungen der Schichtart Wasser und BF152DL bei ca. 1,0 µm.
Die erhöhten Standardabweichungen sind zum einen durch ein stärkeres Umgebungsrauschen zu begründen. Zudem ist eine eindeutige Detektion der Peaks im Leistungsspektrum bei zunehmender Schmierstoffschichtdicke oft nicht möglich. Abbildung 10 zeigt beispielhaft ein aufgenommenes Leistungsspektrum des Schmierstoffs bei einer geometrischen Schichtdicke von 10 µm. Hierbei ist eine Bestimmung der Schichtdicke lediglich durch die Differenz der Schichtdicke der Deckschicht von der Gesamtschichtdicke möglich.

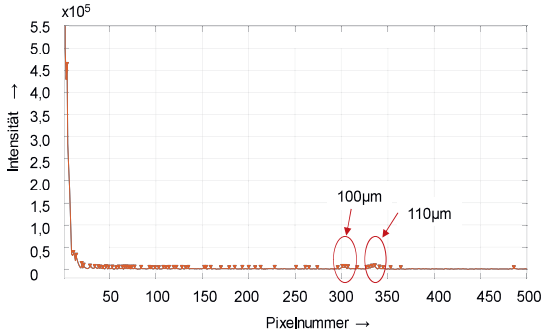


Abb. 10: Leistungsspektrum der Schmierstoffs BF152DL bei einer geometrischen Schichtdicke von 10 µm

Erst bei Betrachtung von lokalen Maxima des Leistungsspektrums (Abbildung 11) ist ein Peak der Schichtdicke erkennbar.

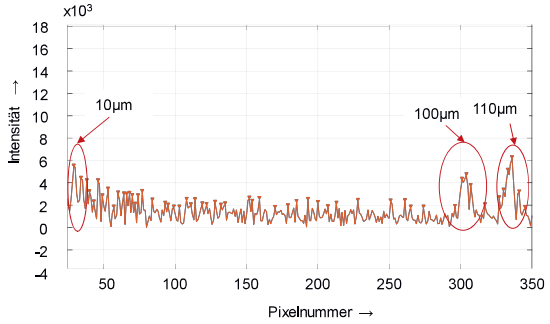


Abb. 11: Lokale Maxima des Leistungsspektrums zeigen keine eindeutige Ausprägung von Peaks

Dies ist mit der Absorption der Strahlung in der Schmierstoffschicht zu begründen. Durch die erhöhte Absorption des Signals ist das Signal-Rauschverhältnis deutlich schlechter, was zu einer verminderten maximal messbaren Schichtdicke führt. Zudem werden an den Wachspartikeln in dem Schmierstoff die einfallenden Strahlen diffus reflektiert. Dies führt dazu, dass die reflektierten Strahlen nicht mehr vom Detektor aufgenommen werden, und, dass das Erkennen der Peaks schlechter wird

Danksagung

Die Arbeit wird durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen des Sonderforschungsbereichs SFB/TR73 in dem Teilprojekten A6 gefördert und wurde an der Friedrich-Alexander-Universität Erlangen-Nürnberg erstellt.

Literaturnachweis

[1] Metzner S., Ramesh R., Müller A., Hausotte T.: Evaluation of reflection properties of sheet bulk metal formed parts by pixel wise analysis of camera images provided for triangulation measurement. ITG-Fachbericht. Bd. 281 Sensoren und Messsysteme, S. 203-206, 2018

[2] McCrackin F., Passaglia E., Stromberg S., Steinberg H.: Measurement of the thickness and refractive index of very thin films and the optical properties of surfaces by ellipsometry. Journal of Research of the National Bureau of Standards Section A: Physics and Chemistry, 1963. doi: 10.6028/jres.067a.040

[3] Hlubina P., Luňáček J., Ciprian D., Chlebus R.: Spectral interferometry and reflectometry used to measure thin films. Applied Physics B, Bd. 92, Nr. 2, S. 203-207, 2008.

[4] Ghim Y., Kim S.-W.: Fast, precise, tomographic measurements of thin films. Applied Physics Letters, Bd. 91, Nr. 9, 2007. doi: 10.1063/1.2776015

[5] Quinten M.: Optische Schichtdickenbestimmung. Optik & Photonik, Bd. 9, Nr. 1, S. 53-57, 2014. doi: 10.1002/opph.201300037

[6] Ghim Y., Kim S.-W.: Thin-film thickness profile and its refractive index measurements by dispersive white-light interferometry. Optics Express, Bd. 14, Nr. 24, 2006. doi: 10.1364/oe.14.011885

[7] Cardin J., Leduc D.: Determination of refractive index, thickness, and the optical losses of thin films from prism-film coupling measurements, Applied Optics, Bd. 47, S. 894, 2008. doi: 10.1364/ao.47.000894

[8] Metzner S., Hausotte T.: Evaluation of influences of thin lubricant on fringe projection measurements. Optics & Photonics Yokohama, OPTM 2019, 2019.

[9] Jakob, G.: Koaxiale interferometrische Schichtdickenmessung. Photonik, Nr. 3, 2000

[10] Quinten, M.: Optische Schichtdickenmessung mit miniaturisierten Spektrometern. Books on Demand, 2015 S.32. ISBN: 978-3734783869

DOI 10.5162/sensoren2019/5.2.2

Entwicklung eines Hochdruckprüfstands für NDIR-Messungen zur Verunreinigungsbestimmung in Wasserstoff für Drücke bis 900 bar

Marco Schott[1], *Andreas Schütze*[1,2]
[1]*ZeMA - Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Eschberger Weg 46, 66121 Saarbrücken, Deutschland*
[2]*Universität des Saarlandes, Lehrstuhl für Messtechnik, Campus A5.1, 66123 Saarbrücken, Deutschland*

Zusammenfassung

Um die Ziele der Energiewende und somit der globalen Reduktion von CO_2 zu erreichen, rückt Wasserstoff (H_2) als Energieträger immer stärker in den Fokus. Mit Hilfe einer Brennstoffzelle wird die chemische Energie von H_2 direkt in elektrische Energie ohne Ausstoß von CO_2 umgewandelt. Dieses Antriebsverfahren wird bereits in Automobilen, Zügen oder auch Schiffen eingebaut. Kleinste Verunreinigungen des H_2, je nach Schadstoff im ppm- oder sogar sub-ppm Bereich, können Schäden verursachen oder die Leistung der Brennstoffzelle reduzieren. Um die H_2-Qualität im Feld zu überprüfen, wird derzeit ein online-fähiges Sensorsystem entwickelt, dass solche Verunreinigungen detektieren und identifizieren soll. Das Messsystem arbeitet in einem Druckbereich bis 900 bar und beruht auf der nicht-dispersiven Infrarot (NDIR)-Absorption. Der hohe Druck hat den Vorteil, dass sich die Empfindlichkeit verbessert, so dass eine optische Messweglänge von etwa 1 m auch für sehr kleine Schadstoffkonzentrationen ausreicht. Für die Auslegung des Systems spielen die einzelnen Absorptionsbanden der Verunreinigungen sowie des H_2 eine entscheidende Rolle. Auf Grund des hohen Drucks verschieben und verbreitern sich die einzelnen Absorptionsbanden, zudem wird H_2 selbst durch Stoßprozesse infrarotaktiv und kann die Erfassung der Schadstoffe stören. Um dieses Verhalten zu untersuchen, wurde ein Prüfstand entwickelt, bei dem eine Gasküvette in einem FTIR-Spektrometer mit definierten Schadstoff-H_2-Gemischen bis zu einem Druck von 900 bar befüllt werden kann.

Keywords: Wasserstoff (H_2), Infrarot, Gasqualität, Hochdruckprüfstand, IR-Aktivität.

Motivation

Um die Ziele des Pariser Übereinkommens aus dem Jahr 2015 und somit die Reduktion der weltweiten CO_2-Emissionen zu realisieren, spielt Wasserstoff als Energieträger eine immer stärkere Rolle. Mit Brennstoffzellen kann die chemische Reaktionsenergie des Wasserstoffs direkt in elektrische Energie umgewandelt werden. Diese kann wiederrum durch einen Elektromotor zur Fortbewegung genutzt werden. Der wesentliche Vorteil liegt darin, dass kein CO_2-Ausstoß auftritt. Bereits jetzt stellen Brennstoffzellen-betriebene Elektrofahrzeuge eine interessante Alternative für den ÖPNV sowie den Individualverkehr dar. Am häufigsten werden Polymerelektrolyt-Brennstoffzellen (PEMFC) eingesetzt, die eine hohe Leistungsdichte besitzen und aufgrund ihres großen Kostensenkungspotenzials favorisiert werden.
Aus einer Studie von Toyota aus dem Jahr 2015 geht hervor, dass Brennstoffzellen-betriebene Autos eine deutlich günstigere Gesamt-CO_2-Bilanz aufweisen als Benzin- oder auch Hybrid-Fahrzeuge, siehe Abbildung 1 [1].

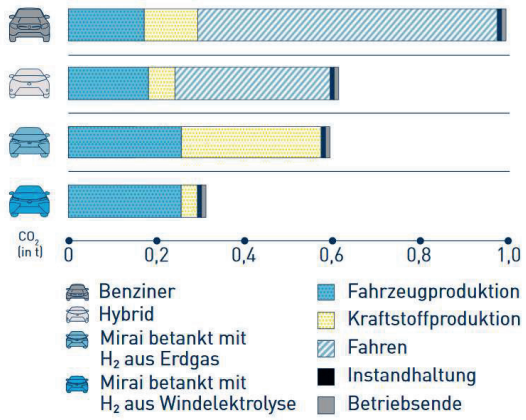


Abb. 1: Vergleich von verschiedenen Antrieben bzgl. ihrer CO_2-Bilanz [1].

In Deutschland steigt die Zahl der Windenergieanlagen seit Jahren kontinuierlich an und aufgrund der temporär hohen Stromproduktion bei entsprechenden Wetterlagen steht zeitweise mehr Strom zur Verfügung als benötigt wird [2]. Dieser Strom kann zur Wasserstoffproduktion mit Hilfe eines Elektrolyseurs genutzt werden, der Wasserstoff kann dann in Hochdrucktanks gespeichert werden. Bei Bedarf kann daraus wieder elektrische Energie erzeugt werden. Wird ein Fahrzeug direkt mit H_2 aus der Windelektrolyse betankt, reduziert das im Vergleich zu einer H_2-Betankung aus Erdgas den CO_2-Ausstoß nochmals um 50 %, siehe Abbildung 1.
Die Nachfrage nach Wasserstoff ist in Deutschland über die vergangenen Jahre kontinuierlich gestiegen. Abbildung 2 zeigt, dass sich die H_2-Nachfrage von Mitte 2016 bis März 2019 ungefähr vervierfacht hat [3]. Eine ähnliche Entwicklung zeigt sich auch beim Ausbau der Wasserstoffinfrastruktur. So hat sich die Zahl der H_2-Tankstellen in den letzten 2,5 Jahren verdreifacht und liegt bei aktuell 66 öffentlichen Tankstellen in Deutschland [3]. Bis Ende 2019 ist ein weiterer Ausbau auf insgesamt 100 Tankstellen geplant [3].



Abb. 2: Wasserstoffnachfrage in Deutschland über die letzten drei Jahre [3].

Auch die Anzahl der entwickelten bzw. am Markt angebotenen Fahrzeuge zeigt eine deutlich steigende Tendenz. Diese positive Entwicklung wird auch in den kommenden Jahren weiter anhalten.
Vor allem im Hinblick auf die Funktion und Langzeithaltbarkeit der Brennstoffzellen ist die Reinheit des Wasserstoffs ein entscheidender Faktor. Je höher die Qualität ist, desto länger ist die Lebensdauer und desto größer die Leistung der Brennstoffzelle. Bereits kleinste Verunreinigungen können Membranschäden versuchen [4]. Kohlenwasserstoffe wie beispielsweise Methan, Ethan, Ethylen oder Propan führen zu einer starken Leistungsreduzierung und vorzeitigem Altern der Membran [4]. Kohlenmonoxid oder Ammoniak können eine Katalysatorvergiftung verursachen [4]. Daher wurde in der SAE-Norm J 2719 [4] festgehalten welche Verunreinigungen zu überwachen und welche Grenzwerte dabei einzuhalten sind, siehe Tabelle 1.

Tab. 1: Schadstoffgrenzwerte nach Norm SAE J 2719 [4].

Stoff		Verunreinigungs-grenzwert (ppmv)
Wasser		5,000
Kohlenwasserstoffe		2,000
	Methan	0,100
	Ethan	0,100
	Ethylen	0,100
	Propan	0,010
Sauerstoff		5,000
Stickstoff		100,000
Argon		100,000
Helium		300,000
Kohlenstoffdioxid		2,000
Kohlenstoffmonoxid		0,200
Schwefelverbindungen		0,001
Hydrocarbonsäure		0,200
Ammoniak		0,100
Formaldehyd		0,010
Halogenierte Stoffe		0,050

Für die Sicherstellung der H_2-Qualität ist daher ein online-fähiges Messsystem gefragt, das die relevanten Verunreinigungen im entsprechenden Konzentrationsbereich detektieren kann.
Einzelne Stoffe lassen sich zwar über verschiedene Messverfahren wie IMS (Ionen-Mobilitäts-Spektrometer), PED (Plasma-Emissions-Detektor) oder WLD (Wärmeleitfähigkeitsdetektor) jeweils in Verbindung mit einem Gaschromatograph (GC) zur Erzielung einer ausreichenden Selektivität bzw. mittels IR-Laser messen [5]. Derzeit ist jedoch kein Messsystem am Markt verfügbar, das die meisten Verunreinigungen (abgesehen von Schwefelverbindungen) selektiv und mit einer sehr hohen Empfindlichkeit erfassen kann.
Theoretische Überlegungen und Voruntersuchungen haben gezeigt, dass ein Messsystem basierend auf der nicht-dispersiven Infrarot (NDIR)-Absorption prinzipiell geeignet ist, die meisten relevanten Verunreinigungen zu detektieren [6]. Für die Auslegung des Messsystems sind Lage und Intensität der gasspezifischen Absorptionsbanden sowie deren Linienverbreiterung und -verschiebung aufgrund des hohen Drucks entscheidend. Zudem weisen bei derartig hohen Drücke auch normalerweise nicht-IR-aktive Gase wie der

Wasserstoff selbst, aber auch Stickstoff (N_2) Absorptionsbanden auf, die berücksichtigt werden müssen [7]. Um dies zu untersuchen, wurde ein Hochdruckprüfstand konzipiert, der Drücke bis zu 900 bar und definierte Schadstoffkonzentrationen im (sub-)ppm-Bereich einstellen kann.

Grundlagen

In [8] und [9] wurde bereits ein Messsystem entwickelt, dass die Qualität von Schmierölen untersucht. Auch dieses System basiert auf NDIR-Absorption und soll nun auf die Wasserstoffqualitätsüberwachung übertragen werden. Obwohl die Grenzwerte für die Verunreinigungen im (sub-)ppm-Bereich liegen, können durch den hohen Druck, der z.B. in H_2-Tankstellen vorliegt, bereits mit einer Küvettenlänge von weniger als 1 m die erforderlichen Nachweisgrenzen erzielt werden.
Abbildung 3 zeigt eine Prinzipskizze des NDIR-Messsystems. Die IR-Quelle sendet Energie durch die IR-durchlässigen Fenster und die Gasküvette hindurch. Innerhalb der Küvette befindet sich das zu untersuchende Gasgemisch. Die absorbierenden Gasmoleküle nehmen Energie bei bestimmten Wellenlängen auf. Die nicht absorbierte Strahlung erreicht den IR-Detektor, wodurch die Transmission bestimmt wird. Durch passende statische oder auch variable Filter kann ein eingeschränkter Wellenzahlbereich betrachtet werden und einem bestimmten Gasmolekül zugeordnet werden. Je mehr Strahlung absorbiert wird, desto höher ist die jeweilige Konzentration. Üblicherweise können nicht alle Gasmoleküle erfasst werden, sondern nur solche, die ein permanentes oder induziertes Dipolmoment aufweisen und somit von der IR-Strahlung zu Schwingungen angeregt werden können.

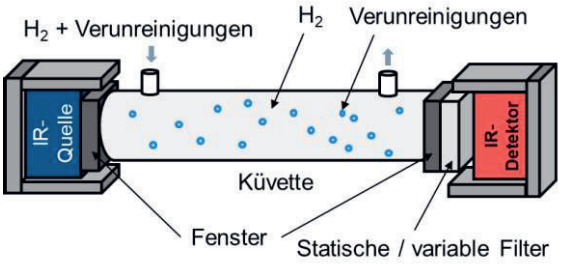


Abb. 3: Prinzipskizze des NDIR-Messsystems zur H_2-Qualitätsüberwachung [7].

Wasserstoff und Stickstoff sind homonukleare Moleküle und unter Normalbedingungen, d.h. bei Umgebungsdruck und Normaltemperatur, nicht IR-aktiv. Durch eine Druckerhöhung steigt jedoch die Wahrscheinlichkeit von Stoßvorgängen, durch die kurzzeitig ein Dipol induziert und letztendlich eine IR-Absorption angeregt wird. Die IR-Aktivität von H_2 wurde bisher bis zu einem Druck von maximal 65 bar untersucht [7]. Für höhere Drücke finden sich auch in der Literatur keine Messergebnisse.
IR-Absorptionsspektren bestehen eigentlich aus einzelnen diskreten Peaks, da die einzelnen Molekülschwingungen nur durch ganz bestimmte Energien angeregt werden können. Neben der Schwingungsquantenzahl ändert sich auch die Rotationsquantenzahl, so dass sich die Absorptionsspektren aus einer Vielzahl eng zusammenliegender Peaks zusammensetzen, die jeweils eine sehr geringe Peakbreite besitzen.
Die Intensitätsverteilung um eine bestimmte Wellenzahl lässt sich durch Höhe, Breite, Form sowie eingeschlossene Fläche (Integral der Absorbanz über Wellenzahl) der einzelnen Linien charakterisieren. Zur Beschreibung wird insbesondere die Halbwertsbreite, d.h. die Breite der Linie bei halbem Maximum, als Beschreibungsmerkmal verwendet. Mit steigendem Druck erhöht sich die Halbwertsbreite. Dieser Effekt wird als Druckverbreiterung $\Delta\nu_c$ beschrieben, die proportional zum Druck ist. Ein weiterer Effekt ist die Druckverschiebung $\Delta\nu_0$. Die Lage des Intensitätsmaximums der einzelnen Peaks verschiebt sich ebenso mit dem Druck. In Abbildung 4 sind beide Effekte dargestellt. Um eine genaue Identifikation und Quantifizierung von Verunreinigungen zu ermöglichen, müssen daher Referenzspektren bei hohen Drücken aufgenommen werden.

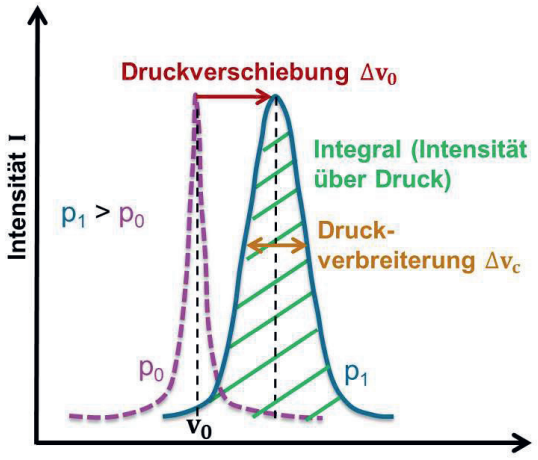


Abb. 4: Druckverbreiterung und Druckverschiebung bei erhöhtem Druck p_1 nach [10].

Messaufbau des Hochdruckprüfstands

Der Hochdruckprüfstand, siehe Abbildung 6, ist so konzipiert, dass zunächst der komplette Aufbau mit Argon gespült wird. Hierdurch werden Umgebungsluft bzw. andere Verun-

reinigungen entfernt. Danach wird H_2 mit definierten Verunreinigungen in die Hochdruckküvette eingebracht. Hierzu wird die Küvette mit einem Prüfgas mit der gewünschten Verunreinigung auf einen definierten Druck, z. B. 9 bar, befüllt. Die im Prüfgas vorliegende Konzentration ist dabei typisch 100-mal so hoch wie der Grenzwert der Verunreinigung gemäß Norm. Anschließend wird die Küvette mit einer zweiten Gasleitung mit reinem H_2 auf den gewünschten Enddruck von z. B. 900 bar gebracht und dabei gleichzeitig die Verunreinigung verdünnt. Hierzu wird ein Druckverstärker (GPS 144, Firma Hydratron) verwendet. Die Hochdruckküvette mit einer optischen Weglänge von 4 cm und einem Gasvolumen von 10 ml wird in ein hochauflösendes FTIR-Spektrometer (Vertex 80v, Bruker) eingebracht. Die an beiden Enden der Küvette angebrachten IR-transparenten Germanium-Fenster sind 2 mm dick. Die Fenster sind im Wellenlängenbereich zwischen 1,85 bis 23 µm transparent und decken damit den IR-Spektralbereich der meisten der in der SAE-Norm genannten Verunreinigungen ab. Die Germanium-Fenster haben ein Bruchmodul von 110 MPa und sind in der Küvette so eingebaut, dass diese einem maximalen Druck von über 900 bar standhält. Zusätzlich ist an der Küvette ein Temperatursensor (PT100) angebracht, um Temperatureinflüsse erfassen und später kompensieren zu können.

Am Ausgang der Küvette befinden sich zwei Drucksensoren für Maximaldrücke bis 1000 bar bzw. 16 bar. Der Hochdrucksensor misst dauerhaft, der 16 bar Sensor wird über zwei Ventile (V3, V4) zur genauen Messung kleinerer Drücke zugeschaltet. Am Ende des Leitungssystems befindet sich schließlich eine kaskadierte Ventilschaltung (V6, V7) mit der ein definiertes Volumen aus der Küvette abgelassen werden kann, um den Druck schrittweise zu verringern.

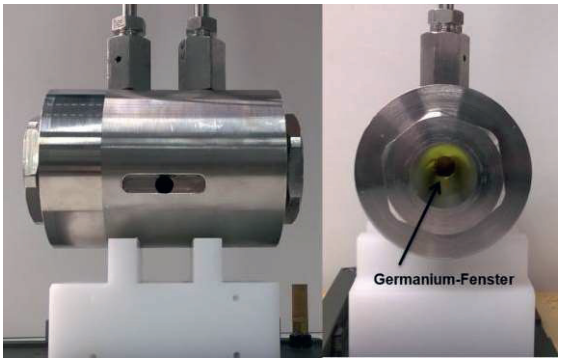


Abb. 5: Hochdruckküvette mit Germanium-Fenstern. Links Seitenansicht, rechts Frontalansicht.

Für Messungen, bei denen keine Verdünnung vorgesehen ist, z. B. bei reinem H_2 oder N_2 bzw. bei höheren Konzentrationen, kann die Prüfgasflasche direkt am Druckverstärker angeschlossen und die zweite Linie mit V1 abgeriegelt werden. Der Druckverstärker befüllt auf diese Weise die Küvette und bringt das Gasgemisch auf hohen Druck, ohne dass sich die Konzentrationen ändern. Die eingesetzten H_2-Prüfgasgemische weisen eine ca. 100-mal höhere Konzentration der Verunreinigungen im Vergleich zum Grenzwert gemäß SAE-Norm auf. Damit können die Absorptionspeaks zunächst eindeutig den Verunreinigungen zugeordnet werden. Abbildung 7 zeigt einige Fotos des Prüfstands.

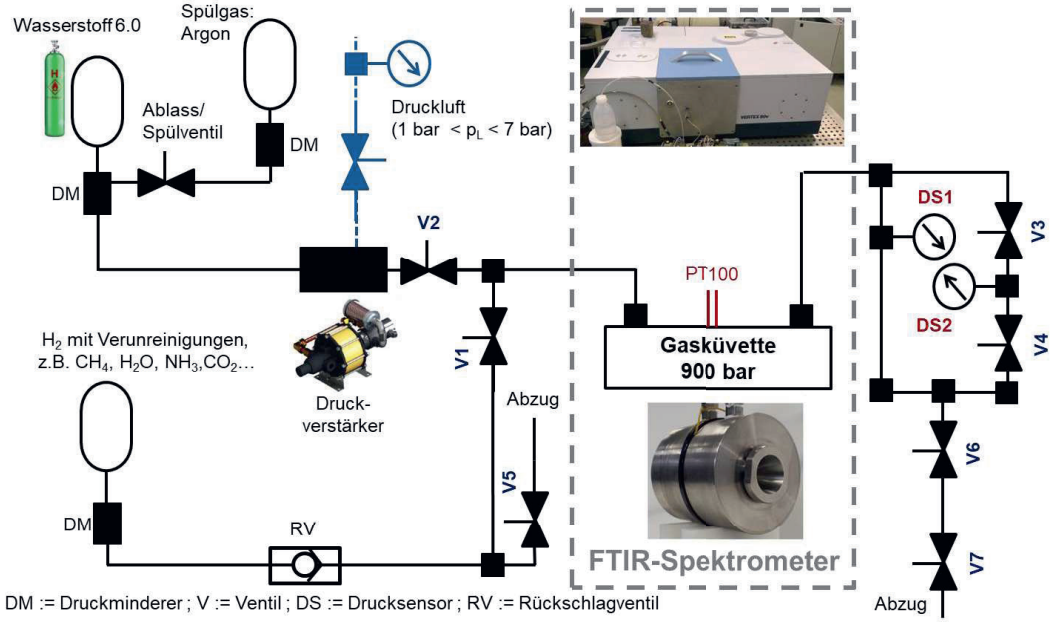


Abb. 6: Prinzipskizze des Hochdruckprüfstands.

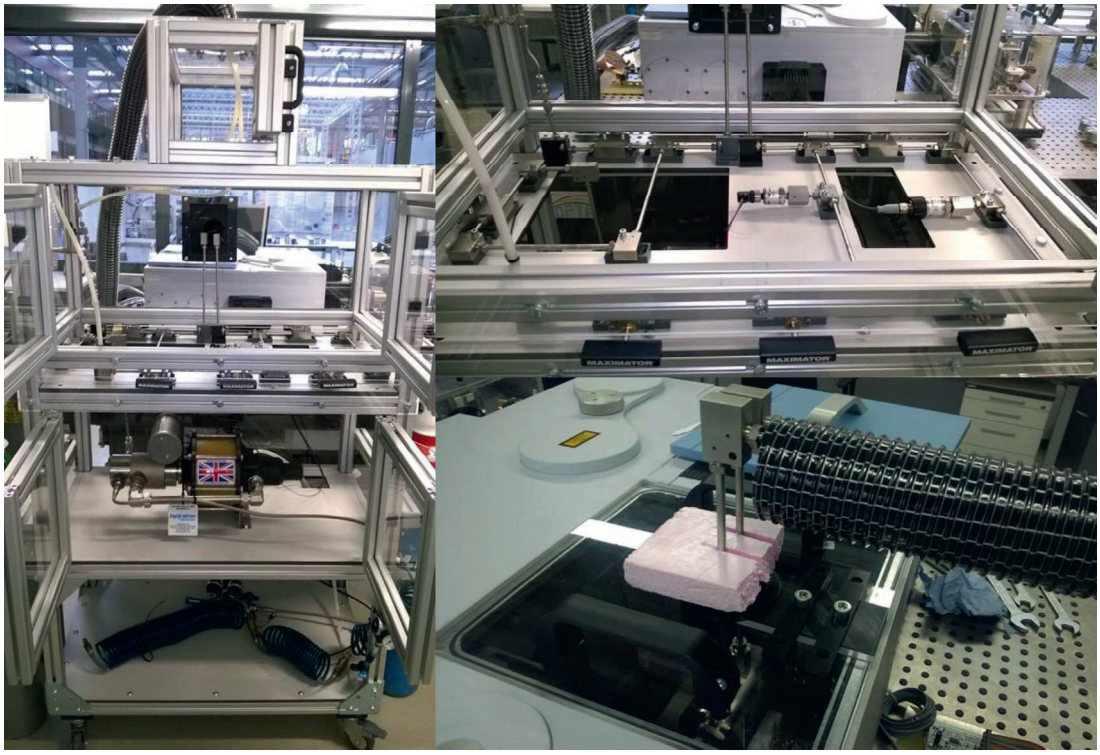


Abb. 7: Hochdruckprüfstand, links: Gesamtansicht, rechts oben: Ventilgruppe, rechts unten: Verbindung zum FTIR-Spektrometer.

Messergebnisse und Analyse für Stickstoff

Untersucht wurde zunächst Stickstoff 6.0, um die IR-Aktivität bei hohem Druck zu prüfen sowie um Verunreinigungen zu charakterisieren. Dabei wurde das Gas mit Hilfe des Hochdruckprüfstands in die Küvette gebracht und mit dem FTIR-Spektrometer untersucht. Laut Datenblatt können die in Tabelle 2 aufgeführten Verunreinigungen auftreten.

Tab. 2: Datenblatt für Stickstoff 6.0 [11]

Bestandteil	Konzentration
CH_4 (Methan)	≤ 0,05 ppm
CO_2 (Kohlendioxid)	≤ 0,05 ppm
H_2O (Wasser)	≤ 0,50 ppm
O_2 (Sauerstoff)	≤ 0,20 ppm
CO (Kohlenmonoxid)	≤ 0,05 ppm
H_2 (Wasserstoff)	≤ 0,10 ppm

Die Küvette wurde zunächst mit Stickstoff 6.0 bis zu einem Druck von 710 bar befüllt und dann das Absorptionsspektrum gemessen. Anschließend wird wiederholt etwas Druck aus der Küvette abgelassen und wieder gemessen, bis die Küvette schließlich wieder drucklos ist. Die Messergebnisse sind in den Abbildungen 8 bis 10 dargestellt.

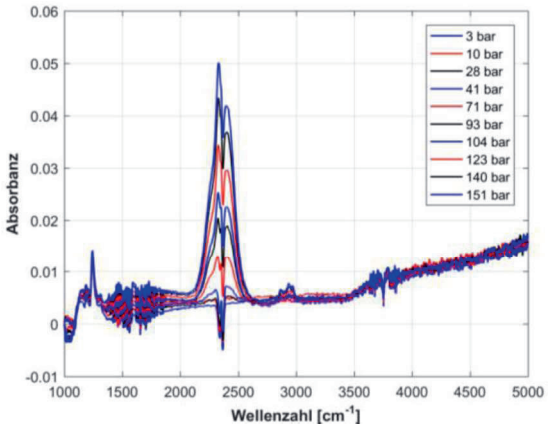


Abb. 8: N_2-Absorptionsspektren (3 - 151 bar).

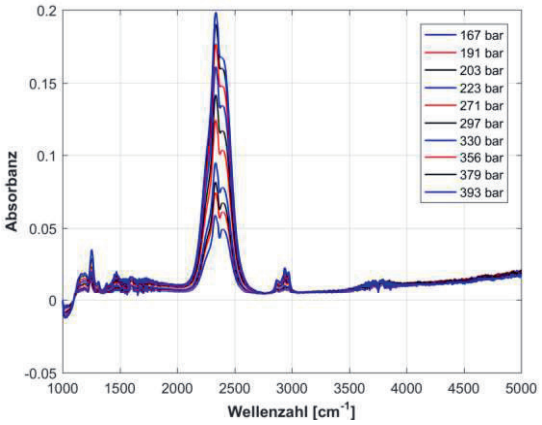


Abb. 9: N_2-Absorptionssp. (167 - 393 bar).

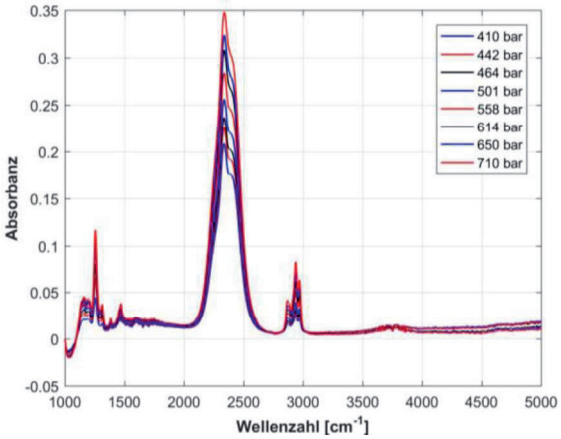


Abb. 10: Stickstoff 6.0-Absorptionsspektren im Druckbereich von 410 bis 710 bar.

Für die Zuordnung der Banden wurden die Datenbanken HITRAN on the Web [12], die IR-Spektrensammlung von Gasmet Ansyco [13] sowie das NIST Chemistry WebBook [14] verwendet. Bei der Analyse der Kurven lässt sich feststellen, dass sich bereits ab ca. 41 bar ein leichter Peak im Wellenzahlbereich ~2330 cm^{-1} zeigt. Nach [15] handelt es sich um N_2. Dies bedeutet, dass auch Stickstoff bei höherem Druck IR-aktiv wird. Zudem sind in den Wellenzahlbereichen zwischen ~1630 cm^{-1} und ~1830 cm^{-1} sowie zwischen ~3500 cm^{-1} und ~3900 cm^{-1} Absorptionsbanden zu erkennen, die die Wasser-Bande darstellen. Die Banden zwischen ~1200 cm^{-1} und ~1300 cm^{-1} sowie ~2840 cm^{-1} und ~3000 cm^{-1} sind den Kohlenwasserstoffen zu zuordnen.
Das Integral der Absorbanz für die N_2-Bande (Abb. 11) zeigt zu Beginn ein anwachsendes Verhalten und geht bei einem Druck von 100 bar in ein lineares Verhalten über. Ähnlich zeigt sich dies auch bei der Halbwertsbreite (Abb. 12). Das gleiche Verhalten zeigt sich auch für die Kohlenwasserstoff- und die H_2O-Banden.

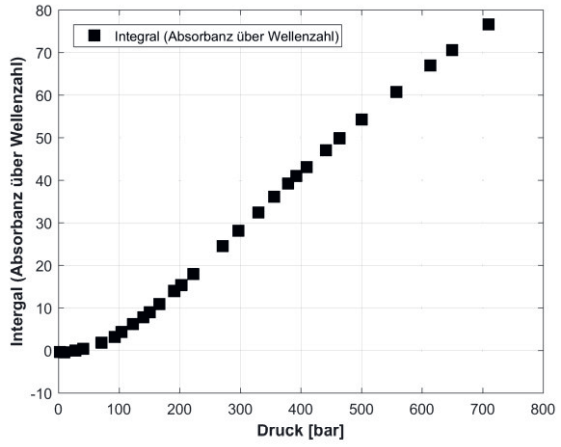


Abb. 11: Integral der Absorbanz für N_2 über den Druck von 0 bis 710 bar.

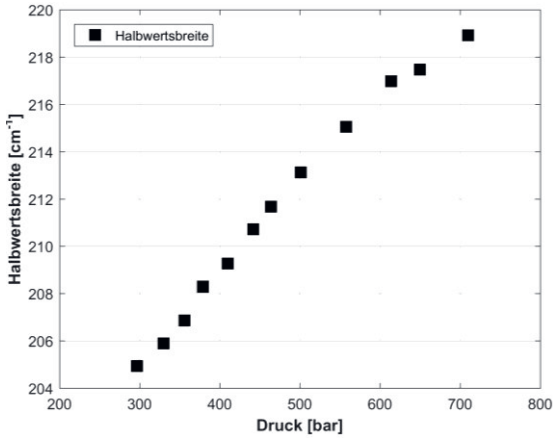


Abb. 12: Halbwertsbreite der N_2-Bande über den Druck von 300 bis 710 bar.

Messergebnisse und Analyse für Wasserstoff

Anschließend wurden Messungen mit Wasserstoff 6.0 mit 10 ppm CO bis zu einem maximalen Druck von 555 bar durchgeführt. Vor den Messungen wurde das Rohrleitungssystem mit Stickstoff 6.0 gespült. Danach wurde das System mehrfach mit Wasserstoff 6.0 mit 10 ppm CO bis zu einem Druck von 50 bar befüllt und wieder auf einen Druck von rund 3 bar abgelassen, um weitere Verunreinigungen zu vermeiden. Der Druck wurde danach schrittweise bis zu einem Druck von 555 bar hochgefahren. Bei verschiedenen Drücken wurde dann eine FTIR-Messung durchgeführt. Die Ergebnisse sind in den Abbildungen 13 und 14 dargestellt.
Dabei ist zu erkennen, dass es zwei markante Wasserstoffbanden gibt. Wie bereits in [7] gezeigt, befindet sich eine breite Absorptionsbande in einem Wellenzahlbereich von ~3550 cm^{-1} bis ~5000 cm^{-1}. Diese H_2-Bande tritt bei einem Druck von 33 bar zum ersten Mal deutlich auf. Dies deckt sich auch mit [15].
Eine weitere H_2-Bande beginnt unterhalb von ~2000 cm^{-1} und geht über 1000 cm^{-1} hinaus. Diese zweite H_2-Bande tritt erst bei einem Druck von 63 bar deutlich auf und wird auch bei steigendem Druck in Ihrer Intensität nicht so stark wie die erste Bande im Bereich ~3550 cm^{-1} bis ~5000 cm^{-1}.
Positiv ist zu bewerten, dass die zwei relativ breiten H_2-Banden die Absorptionsbanden von Methan, Kohlenmonoxid, Kohlendioxid und Ammoniak nicht überdecken, so dass diese Verunreinigungen mit dem Sensorsystem empfindlich gemessen werden können.
Ebenso wie bereits bei den Stickstoffmessungen befinden sich auch zwei leichte Wasserbanden im gemessenen Spektrum.

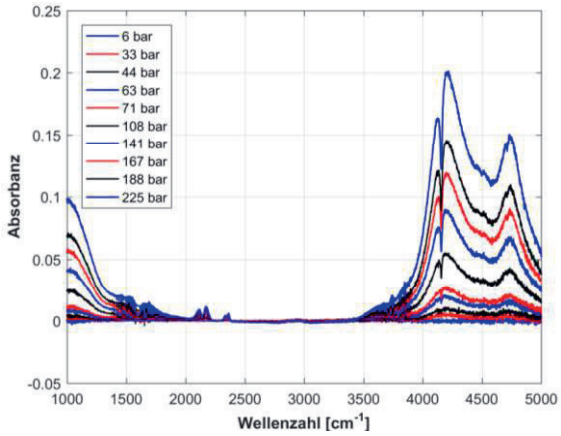


Abb. 13: Absorptionsspektrum von Wasserstoff mit 10 ppm CO im Druckbereich von 6 bis 225 bar.

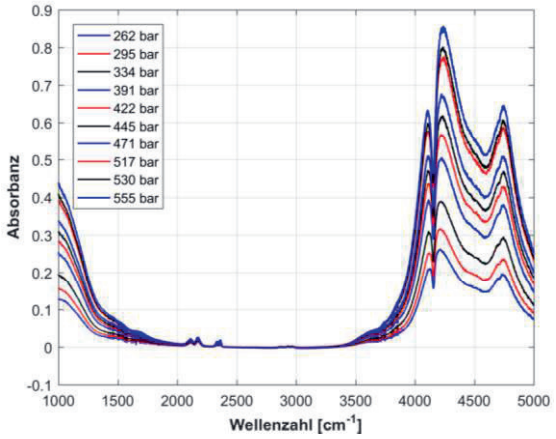


Abb. 14: Absorptionsspektrum von Wasserstoff mit 10 ppm CO im Druckbereich von 262 bis 555 bar.

Der CO-Peak ist im Bereich zwischen 2050 cm^{-1} und 2225 cm^{-1} zu erkennen und in Abbildung 15 dargestellt. Bei der Betrachtung der maximalen Intensität der CO-Banden aufgetragen über den Druck, zeigt sich, dass auch hier eine Proportionalität vorliegt.

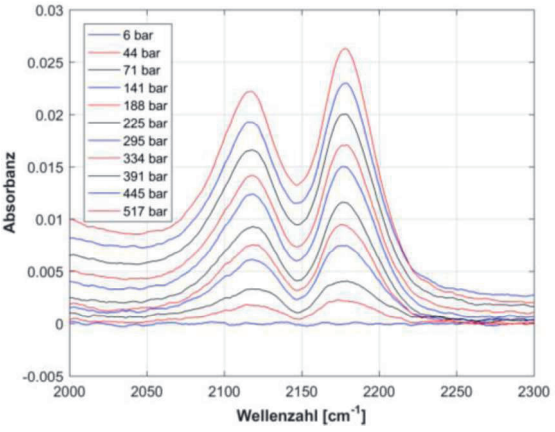


Abb. 15: CO-Bande für 6 bar ≤ p ≤ 517 bar.

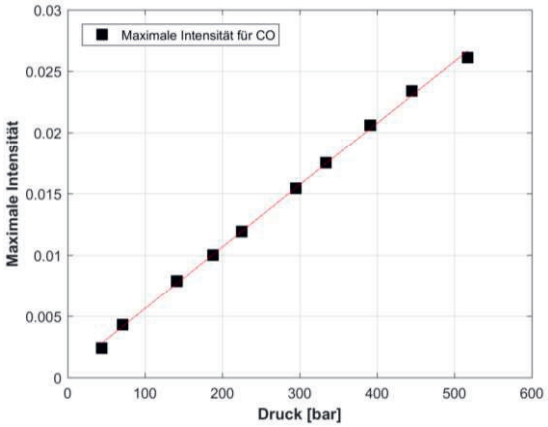


Abb. 16: Maximale Intensität der CO-Bande für Drücke von 44 bis 517 bar.

Zusammenfassung und Ausblick

Es konnte gezeigt werden, dass Verunreinigungen in Stickstoff und Wasserstoff mit einem Hochdruckprüfstand in einem FTIR-Spektrometer charakterisiert werden können.

Im Versuch wurde Stickstoff 6.0 auf einen Maximaldruck von 710 bar gebracht, der anschließend schrittweise heruntergefahren wurde, um die Absorptionsspektren bei verschiedenen Drücken zu bestimmen. Dabei zeigt sich, dass im Wellenzahlbereich von 2150 cm^{-1} bis 2540 cm^{-1} eine breite, druckinduzierte IR-Bande von Stickstoff ersichtlich ist. Diese ist bei ~41 bar zum ersten Mal zu erkennen und steigt mit zunehmendem Druck an. Auch zwei Kohlenwasserstoff-Banden sind in den Diagrammen zu erkennen. Sichtbar sind diese ab ca. 141 bar und steigen ebenfalls mit zunehmendem Druck weiter an. Die Intensität sowie die Halbwertsbreite dieser Banden steigen proportional zum Druck.

Bei den Wasserstoffmessungen wurden zwei intensive und spektral sehr breite H_2-Banden festgestellt. Diese überdecken allerdings keine Banden, die relevanten Verunreinigungen wie CO, CH_4 oder CO_2 zuzuordnen sind, so dass die angestrebte Funktion des IR-Messsystems nicht gestört wird. Insgesamt zeigt sich für alle Verunreinigungen, dass die Intensität, d.h. das Integral der Absorbanz, sowie die Halbwertsbreite der Banden ab einem gewissen Druck proportional mit dem Druck ansteigen.

Im nächsten Schritt sollen weitere Verunreinigungen untersucht werden. Zudem soll der Druck noch weiter gesteigert werden, um den Maximaldruck von bis zu 900 bar auszureizen. Damit soll untersucht werden, ob sich die Banden auch bei höheren Drücken weiterhin proportional zum Druck verhalten. Zudem sollen Konzentrationen entsprechend der in der SAE-Norm genannten Grenzwerte untersucht werden.

Danksagung

Die zugrundeliegenden Arbeiten wurden vom Bundesministerium für Wirtschaft und Energie (BMWi) über den Projektträger PTJ im Rahmen des Projekts HAIMa (H_2- und Kationen-Kontamination: Alterungseffekte, Material und Sensorentwicklung; Teilvorhaben: Entwicklung und Erprobung von Messprinzipien zur H_2-Qualitätsüberwachung) gefördert.
Ein besonderer Dank gilt Herrn Eliseo Pignanelli, der die Forschungsarbeiten für dieses Projekt maßgeblich vorbereitet hat.

Literaturnachweis

[1] Life Cycle Assessment Report Toyota, 2015.
[2] H.-W. Sinn, Buffering volatility: A study on the limits of Germany's energy revolution, European Economic Review 99, 2017.
[3] https://h2.live/, Stand 22.04.2019.
[4] SAE J 2719: Hydrogen Fuel Quality for Fuel Cell Vehicles.
[5] Das H_2-Messsystem zur Outdoormessung in Laborqualität, Produktflyer, Emcel GmbH, Hannover Messe 2019.
[6] M. Schott, E. Pignanelli, S. Gratz-Kelly, A. Schütze: Modellierung und Simulation eines IR-Messsystems zur Optimierung der Strahlungsintensität für die Wasserstoffqualitätsüberwachung, ITG-Fachbericht 281: Sensoren und Messsysteme, VDE-Verlag Berlin (2018), S. 315-318.
[7] M. Schott, E. Pignanelli, S. Gratz-Kelly, A. Schütze: Modulares IR-Messsystem zur Überwachung der Wasserstoffqualität für Brennstoffzellenfahrzeuge, Tagungsband 13. Dresdner Sensor-Symposium (2017), S. 34-39.
[8] T. Bley, E. Pignanelli, A. Schütze: COPS - Combined Oil quality and Particle measurement System, Proceedings ICST 2012 – The Sixth International Conference on Sensing Technology, Kolkata, India, December 18-21, 2012; doi: 10.1109/ICSensT.2012.6461742
[9] E. Pignanelli, T. Bley, A. Schütze: Modulares Condition Monitoring System zur Überwachung von Fluidqualität und Partikelkontamination am Beispiel von Öl, 11. Dresdner Sensor Symposium, Dresden, 9.-11. Dezember 2013; doi: 10.5162/11dss2013/E8
[10] A. T. Brysch, Laser-Emissionsspektrometrie für die Partikelanalyse von Prozessgasen bei der Roheisenerzeugung, Dissertation, Rheinisch-Westfälische Technische Hochschule Aachen, Fakultät für Maschinenwesen, 2004.
[11] Produktdatenblatt für Stickstoff 6.0, Nippon Gases.
[12] http://hitran.iao.ru/bands/simlaunch?mol=5
[13] http://www.ansyco.de/technologie/ir-spektren
[14] https://webbook.nist.gov/chemistry/form-ser/
[15] J. H. Odhner, D. A. Romanov, R. J. Levis, Filament-based stimulated Raman spectroscopy, Proc. SPIE 7582, Nonlinear Frequency Generation and Conversion: Materials, Devices, and Applications IX, 75820M (17 February 2010); doi: 10.1117/12.842129

DOI 10.5162/sensoren2019/5.2.3

Zinkoxid als photostabiler Luminophor zur optischen Sauerstoffdetektion

Michael Poeplau[1], Stephan Ester[1], Bernd Henning[2], Thorsten Wagner[3]

[1]Wöhler Technik GmbH, Wöhler-Platz 1, 33181 Bad Wünnenberg, E-Mail: m.poeplau@woehler.de
[2]Elektrische Messtechnik, Universität Paderborn, Warburger Straße 100, 33098 Paderborn
[3]Department Chemie, Universität Paderborn, Warburger Straße 100, 33098 Paderborn

Zusammenfassung

In der vorgestellten Arbeit untersuchen wir Zinkoxid als ein photostabiler Luminophor zur Sauerstoffdetektion. Bisher verfügbare organische Luminophore, üblicherweise Metallkomplexe, weisen gute Eigenschaften bezüglich Sensitivität und Selektivität auf, altern jedoch im Betrieb durch die optische Anregung (Photobleichung). Zinkoxid hingegen ist bekannt als langzeitstabiles lumineszierendes Material und wird unter anderem in Leuchtmitteln verwendet (Bspw. in Leuchtdioden). Wir zeigen die Gasabhängigkeit der Zinkoxid-Lumineszenz mit einer eigens konzipierten Messzelle. Zusätzlich wurden durch simultane Messungen der Lumineszenzemission und des Leitwerts Hinweise auf den Lumineszenz-Quenching-Mechanismus gefunden. Dieser unterscheidet sich qualitativ von den Quenching-Mechanismen an Metallkomplexen.

Keywords: Gassensor, Zinkoxid, Lumineszenz, Luminophor, Photobleichung

1. Einleitung

Lumineszenzbasierte Gassensorik, z.B. für Sauerstoff (O_2), besitzt hervorragende Eigenschaften hinsichtlich Empfindlichkeit und Reproduzierbarkeit. Zusätzlich weisen die als Luminophore verwendeten Metallkomplexe, z.B. basierend auf Ruthenium, eine hohe, inhärente Selektivität auf. Nach Bestrahlung mit kurzwelligem ultraviolettem oder blauem Licht relaxieren die angeregten Elektronen über strahlende Übergänge unter Abgabe einer molekülspezifischen Lumineszenz. Die Reaktion mit Sauerstoff ändert die Molekülorbitale des Luminophors, so dass nun nichtstrahlende Übergänge bevorzugt werden; die Intensität der Lumineszenz nimmt ab (Lumineszenz-Quenching). Die Lumineszenzintensität bzw. deren Abklingverhalten wird als gasabhängige Messgröße ausgewertet [1] und erlaubt so einen einfachen Messaufbau mit kostengünstigen optischen Komponenten [2].
Nachteilig ist jedoch die geringe Stabilität der Metallkomplexe gegenüber kurzwelliger Strahlung. Auf Grund der sogenannten Photobleichung, d.h. die Zersetzung der Moleküle unter Beleuchtung, eignen sie sich daher bisher nicht für kontinuierliche Messaufgaben und werden üblicherweise nur für punktuelle Messungen, beispielsweise zur Messung der Sauerstoffkonzentrationen in Transportverpackungen verwendet [3].

Um die prinzipbedingten Vorteile der lumineszenzbasierten Gassensorik für kontinuierliche Messaufgaben zu nutzen, gilt es, neue langzeitstabile und zugleich gasempfindliche Luminophore zu finden.
Eine interessante Alternative ist halbleitendes, photostabiles Zinkoxid. Dieses weist bei entsprechender kristalliner Beschaffenheit ebenfalls Lumineszenzemissionen auf, welche von der umgebenden Gasatmosphäre beeinflusst werden [4]. Im Gegensatz zu den zuvor beschriebenen Metallkomplexen ist der Mechanismus des Lumineszenz-Quenchings für die Metalloxide jedoch nicht abschließend geklärt. Optimierungen der Materialsynthese und der Signalauswertung können deshalb bisher nur rein empirisch vorgenommen werden.
Ziel dieser Arbeit ist es, ausgehend von bekannten Mechanismen aus dem Bereich der resistiven Gassensorik mit Metalloxiden, Hinweise auf den Quenching-Mechanismus zu gewinnen. Bei resistiven Sensoren wird üblicherweise der Leitwert von thermisch aktiviertem Zinkoxid als Sensorgröße verwendet. Durch simultane Messungen von Lumineszenz und Leitwert an Zinkoxid-Partikeln sollen bei variierender Sauerstoffkonzentration Indizien für die Quenching-Mechanismen gefunden werden.

Zinkoxid

Zinkoxid (ZnO) ist ein halbleitendes Metalloxid mit einer Bandlücke von ca. 3,2 eV. Bei hochenergetischer optischer Anregung, z.B. mit einer Wellenlänge von 280 nm (4,43 eV), zeigt Zinkoxid ausgeprägte Lumineszenzemissionen. Ursache dafür sind strahlende Übergänge der angeregten Elektronen vom Leitungs- in das Valenzband (Abbildung 1). Direkte Relaxation der Elektronen führt zur Emission mit einer Wellenlänge von ca. 384 nm (3,2 eV), dem Bandabstand von ZnO (Band-Band-Lumineszenz). Je nach Dotierung bzw. Defekten kommt es zudem zur indirekten Relaxation über Defektniveaus, welche im gezeigten Beispiel zu einer breiten Emissionsbande bei ca. 510 nm (2,43 eV) führt (Defekt-Lumineszenz). Der Anteil der Band-Band- und Defekt-Lumineszenz hängt dabei von der Kristallinität des verwendeten Zinkoxids ab [5].

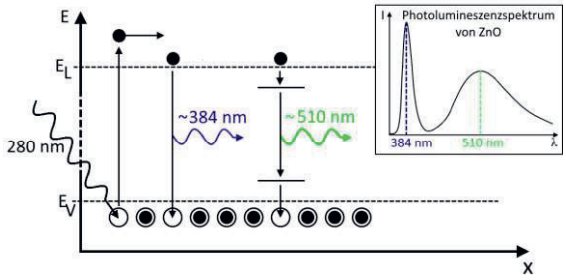


Abb. 1: Schematische Darstellung der Energieniveaus von Zinkoxid. Links: UV-Anregung bei 280 nm. Mitte: Band-Band-Lumineszenz bei ca. 384 nm. Rechts: Defekt-Lumineszenz bei ca. 510 nm. Einschub: Charakteristisches Photolumineszenzspektrum von ZnO.

Neben der Lumineszenz, kann zudem die Leitfähigkeit des Zinkoxids optisch aktiviert werden. Die Elektronen können dabei, je nach Energie, aus leitungsbandnahen Fehlstellen oder aus dem Valenzband in das Leitungsband gehoben werden und dort zur Leitfähigkeit beitragen.

2. Experimentelles

2.1 Messzelle

Um simultan den Leitwert und die Lumineszenz der Zinkoxid-Partikel bei variierenden Gaskonzentrationen zu erfassen, wurde eine Messzelle entworfen. Diese besteht aus einer Standard-Fluoreszenzküvette und einem Küvetteneinsatz (Abbildung 2). Der Küvetteneinsatz schließt die Küvette gasdicht ab und fixiert eine Keramikplatine, die als Sensorsubstrat dient. Auf der Keramikplatine ist eine Interdigital-Elektrodenstruktur aus Kupfer aufgebracht, die mit einer Schicht aus Chemisch Gold (ENIG) passiviert ist. Der Elektrodenabstand und die Elektrodenbreite betragen 100 µm. Durch den Küvetteneinsatz können die Drähte zur Kontaktierung der Interdigitalstruktur nach außen geführt und die Messzelle mit Prüfgasen beaufschlagt werden. Die Messzelle ermöglicht es, die Probe optisch anzuregen und unter einem Winkel von 90° zur Anregung die Lumineszenzemission zu beobachten. Zeitgleich kann der Leitwert der Probe erfasst und die Gaskonzentration innerhalb der Messzelle variiert werden.

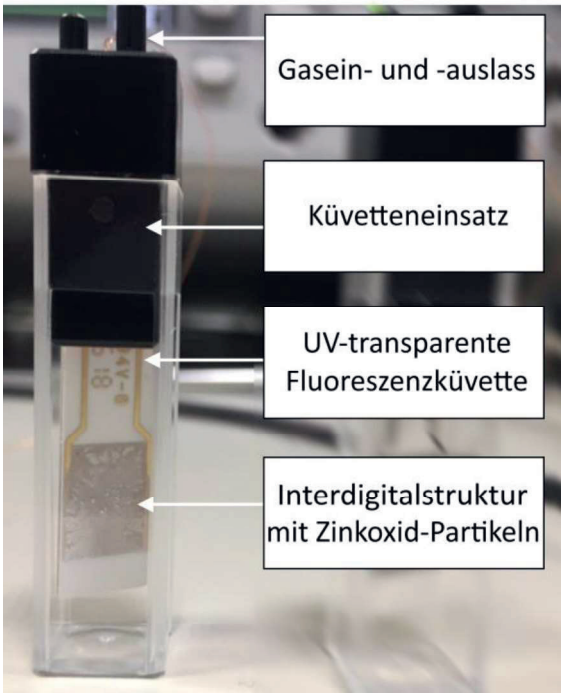


Abb. 2: Verwendete Messzelle, bestehend aus einer UV-transparenten Laborküvette, dem Küvetteneinsatz sowie der Keramikplatine mit Zinkoxid-Partikeln.

2.2 Sensorpräparation

Für die Messungen in dieser Arbeit wurden kommerziell verfügbare Zinkoxid-Nanopartikel verwendet (abcr.de - Art.-Nr.: AB354609). Zur Präparation wurden 50 mg der Zinkoxid-Partikel und 2 ml Ethanol in ein Schnappdeckelglas gegeben und in einem Ultraschallbad für 5 min dispergiert. Die Keramikplatinen wurden auf einer Labor-Heizplatte auf 60°C erwärmt. Anschließend wurden mit einer Mikroliterpipette 20 µl der Dispersion entnommen und in einem Zug auf die Interdigitalstruktur getropft. Nachdem das Ethanol verdampft war, wurde dieser Vorgang zwei weitere Male wiederholt. Das Substrat wurde anschließend mit zwei Kupferlackdrähten kontaktiert und in der Messzelle montiert.

2.3. Messaufbau

Abbildung 3 zeigt schematisch die Anordnung der elektrischen und optischen Komponenten des Messaufbaus. Das Zinkoxid wird über eine Leuchtdiode mit einem Emissionsmaximum bei 280 nm angeregt, welche mit einer Konstantstromquelle betrieben wird. Die optische Leistung beträgt ca. 1 mW. Die Lumineszenzemission wird über eine Kollimatorlinse erfasst und über eine optische Faser an ein Faserspektrometer geleitet (Ocean Optics, Typ: Flame S). Um die reflektierten Anteile der intensiven anregenden Strahlung zu filtern, ist vor der Kollimatorlinse ein optischer Langpassfilter mit einer Grenzwellenlänge von 375 nm angebracht.

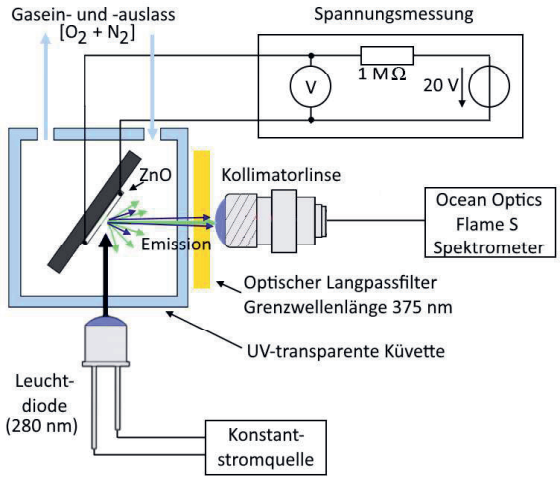


Abb. 3: Schematische Darstellung des Messaufbaus.

Um den Leitwert unter definierten Bedingungen zu messen, wird die Interdigitalstruktur in Reihe zu einem 1 MΩ Widerstand an einem Labornetzteil mit einer konstanten Ausgangsspannung von 20 V betrieben. Der Leitwert der Zinkoxid-Partikel wird anschließend aus der gemessenen Spannung über der Interdigitalstruktur berechnet (Multimeter: HP 34401A). Die Beaufschlagung mit Gas erfolgt über eine PC-gesteuerte Gasmischanlage. Die Ausgangsgase sind Stickstoff (Fa. Linde, Reinheit 5.0), sowie synthetische Luft (Fa. Linde, 20 % Sauerstoff, KW-frei). Um den Einfall von Umgebungslicht zu vermeiden, wurden die optischen Komponenten in einem Aluminiumblock angeordnet. Abbildung 4 zeigt den Messaufbau mit eingesetzter Messzelle.

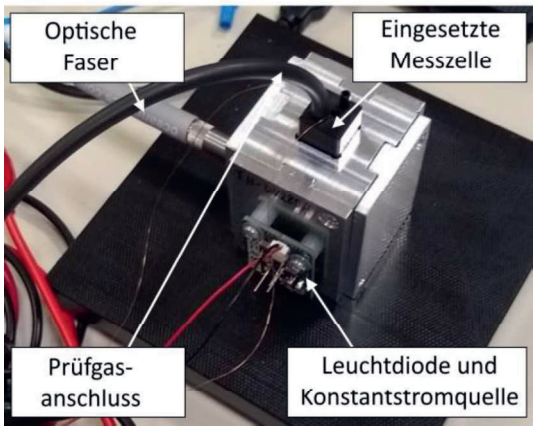


Abb. 4: Messaufbau.

2.4. Messdatenerfassung

Um thermisch stabile Bedingungen zu schaffen, wurde die Leuchtdiode zwei Stunden vor den eigentlichen Messungen in Betrieb genommen und die Messzelle während dieser Zeit mit Stickstoff gespült. PC-gesteuert wurde die Sauerstoffkonzentration in der Messzelle anschließend schrittweise erhöht (60 Minuten je Konzentrationsänderung). Der Volumenstrom beträgt dabei 100 ml / min. Der Spannungsabfall über der Zinkoxidprobe, sowie das optische Spektrum wurden über eine PC-Schnittstelle erfasst. Aus dem Spannungsabfall lässt sich anschließend der Leitwert der Probe berechnet. Die Messdaten des Spektrometers wurden auf die maximale Lumineszenzintensität normiert. Eine Korrektur der Lumineszenzspektren, um beispielsweise systembedingte Messartefakte zu kompensieren, erfolgte nicht.

3. Messergebnisse und Diskussion

Im ersten Schritt soll gezeigt werden, dass das verwendete Zinkoxid ein gasabhängiges Lumineszenz-Verhalten aufweist. Dazu wurden mit dem zuvor beschriebenen

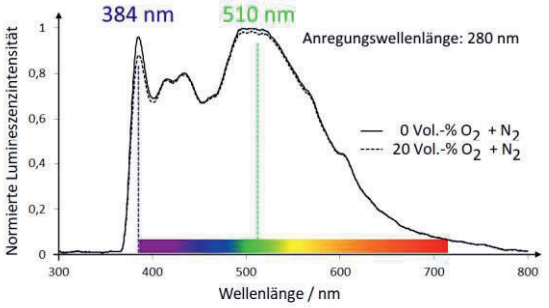


Abb. 5: Photolumineszenzspektrum von Zinkoxid bei Abwesenheit (durchgezogene Linie) und bei Anwesenheit (gestrichelte Linie) von Sauerstoff.

Aufbau Lumineszenzspektren unter unterschiedlichen Gasatmosphären aufgenommen (Abbildung 5). Die Spektren weisen die für Zinkoxid charakteristischen Band-Band-Lumineszenzen bei 384 nm, sowie die defektbedingten Lumineszenzen bei ca. 510 nm auf.
Bei Variation der Gasatmosphäre kann eine Veränderung der Lumineszenzintensität festgestellt werden. In Anwesenheit von Sauerstoff nimmt die Lumineszenzintensität ab (Lumineszenz-Quenching). Bei einer Variation der Sauerstoffkonzentration von 0 auf 20 Vol.-% wird ein Rückgang der ZnO-Lumineszenzintensität (384 nm) von ca. 10 % beobachtet (Abbildung 6). Der Einfluss des Sauerstoffs auf die Defekt-Lumineszenz (510 nm) ist mit einer Abschwächung von ca. 2 % wesentlich geringer.
Die Nebenmaxima bei 410 nm und 430 nm sind Messartefakte, die auf die Komponenten des Messaufbaus zurückzuführen sind und zeigen dementsprechend kein gasabhängiges Verhalten. So sind den gemessenen Spektren zum einen, trotz des eingesetzten UV-Filters, Anteile des Spektrums der UV-Leuchtdiode überlagert; diese emittiert geringe Anteile im sichtbaren Spektralbereich. Zum anderen weisen alle optischen Komponenten (Kollimatorlinse, opt. Faser, Detektor) ein wellenlängenabhängiges Übertragungsverhalten auf, sodass der tatsächliche Anteil der Band-Band- und Defekt-Lumineszenz ggfs. anders gewichtet ist.

Ein Vergleich des zeitlichen Verlaufs der Intensität der Band-Band-Lumineszenz (384 nm) mit dem Verlauf des simultan gemessenen Leitwerts bei variierender Sauerstoffkonzentration zeigt eine starke Korrespondenz der beiden Signale (Abbildung 6).

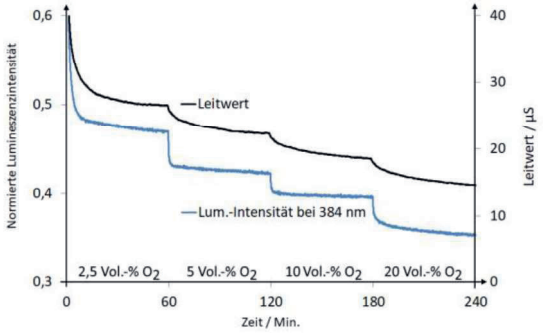


Abb. 6: Zeitliche Änderung der Lumineszenzintensität (384 nm) und des Leitwerts bei Variation der Sauerstoffkonzentration.

Sowohl die Lumineszenzintensität als auch der Leitwert nehmen mit steigender Sauerstoffkonzentration ab. Die Abnahme des Leitwertsignals erfolgt durch die Lokalisierung freier Elektronen an der Oberfläche der ZnO-Partikel.
Zusätzlich zur Bandverbiegung an den Korngrenzen trägt auch die Verringerung der freien Ladungsträger zur Abnahme des Leitwertes bei (Abbildung 7). Die beobachtete Korrespondenz der beiden Signale legt nahe, dass diese Abnahme der freien Ladungsträger auch die Ursache für das Lumineszenz-Quenching darstellt. Die optisch angeregten Elektronen werden durch Sauerstoff-Adsorbate an der Grenzfläche der ZnO-Partikel lokalisiert und können nicht strahlend relaxieren.

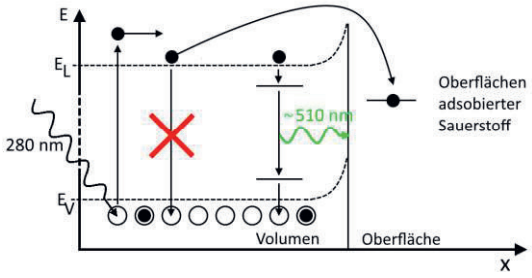


Abb. 7: Schematische Darstellung des Einflusses von Sauerstoff auf der Partikeloberfläche auf die Lumineszenzemission.

Somit ist davon auszugehen, dass Zinkoxid als Luminophor ähnliche gassensitive Eigenschaften wie im Fall der resistiven Signalauswertung aufweisen wird. Bemerkenswert ist jedoch, dass sich die Dynamik der Signale bei Variation der Gaskonzentration deutlich unterscheidet.
Die Lumineszenzintensität weist eine steilere Flanke auf und stabilisiert sich schneller als das Leitwertsignal. Die Ursache dafür liegt wahrscheinlich in der Überlagerung parasitärer Effekte durch Partikel-Partikel-Übergänge im Zinkoxid-Pulver sowie in Umladungseffekten auf der Oberfläche. Die damit verbundenen Variationen in den Raumladungsschichten bzw. den Bandverbiegungen an den Korngrenzen beeinflussen das Leitwertsignal stark. Der beschriebene Mechanismus zum Lumineszenz-Quenching jedoch ist in gewissen Grenzen unabhängig von der Verbiegung der Bänder und reagiert primär auf die Zahl der freien Elektronen. Dieser Ansatz eröffnet somit eine Möglichkeit zur Signalstabilisierung.

4. Zusammenfassung und Ausblick

In dieser Arbeit wurde gezeigt, dass die Lumineszenz von Zinkoxid-Partikeln grundsätzlich zur Detektion von Sauerstoff genutzt werden

kann und es wird ein Quenching-Mechanismus für Zinkoxid vorgeschlagen. Die beobachtete Korrespondenz des Lumineszenzsignals mit dem Leitwert legt nahe, dass das Quenching auf der Lokalisierung von Leitungselektronen durch Adsorbate an der Oberfläche der Zinkoxidpartikel beruht. Damit unterscheidet sich der Mechanismus qualitativ vom Quenching der Metall-Komplexe; eine inhärente Selektivität bzgl. des Sauerstoffs ist im Fall des Zinkoxids nicht zu erwarten. Das Lumineszenzsignal zeigt allerdings ein deutlich schnelleres Ansprechverhalten als das Leitwertsignal. Weitere Studien müssen zeigen, wie diese Dynamik von den Materialparametern, wie z.B. der Kristallitgröße des Zinkoxids, beeinflusst wird.

Literaturnachweis

[1] X. Wang, O. Wolfbeis, Chem. Soc. Rev., 3666-3761 (2014); doi: 10.1039/C4CS00039K
[2] O. Wolfbeis, Bioessays. 37(8):921-8 (2015); doi: 10.1002/bies.201500002
[3] P. Hartmann, M. Leiner, P. Kohlbacher, Sensors and Actuators B: Chemical. 51. 196-202; doi: 10.1016/S0925-4005(98)00188-9
[4] V. M. Zhyrovetsky, D.I. Popovych, S. S. Savka, A. S. Serednytski, Nanoscale Research Letters 12:132 (2017) doi: 10.1186/s11671-017-1891-5
[5] O. Marin, M. Tirado, N. Budini, E. Mosquera, C. Figueroa, D. Comedi, Materials Science in Semiconductor Processing, 56, 59-65 (2016); doi: 10.1016/j.mssp.2016.07.007

DOI 10.5162/sensoren2019/5.2.4

Photoakustisches Gasmesssystem zur Bestimmung des Schwefeldioxidgehaltes in Schiffsabgasen

Mahmoud El-Safoury[1], Miguel Dufner[1], Christian Weber[1], Katrin Schmitt[1,2], Hans-Fridtjof Pernau[1], Bert Willing[3] und Jürgen Wöllenstein[1,2]

[1] Fraunhofer Institut für Physikalische Messtechnik IPM, Heidenhofstr. 8, 79110 Freiburg, Deutschland

[2] Institut für Mikrosystemtechnik – IMTEK, Lehrstuhl für Gassensoren, Universität Freiburg, Georges-Köhler-Allee 102, 79110 Freiburg, Deutschland

[3]RÜEGER SA, Ch. de Mongevon 9, 1023 Crissier, Schweiz

Zusammenfassung

Ab dem Jahr 2020 gelten weltweit auf den Ozeanen strengere Grenzwerte für die Emission von SO_2 in Schiffsabgasen. Diese können entweder über schwefelarmen Treibstoff oder durch den Einbau von Filteranlagen erreicht werden, wobei in letzteren der SO_2-Ausstoß permanent gemessen werden muss. Wir präsentieren erste Feldtest-Messungen eines photoakustischen SO_2-Messsystems, das seit August 2018 auf einem Schiff auf der Ostsee unterwegs ist.

Keywords: SO_2-Sensorik, Gassensorik, photoakustischer Sensor, maritime Sensorsysteme

Motivation

Der Transport von Waren per Schiff ist nur auf die CO_2 Bilanz bezogen sehr umweltfreundlich. Schaut man aber auf andere Emissionen wie NO_x, Feinstaub oder SO_2 ist der Schiffsverkehr für einen Löwenanteil der im Transportwesen emittierten Schadstoffe verantwortlich. Bereits 2005 wurden daher strengere Abgasrichtwerte für Schiffe festgelegt die final 2020 in Kraft treten sollen, um weltweit auch auf hoher See den SO_2-Austoß deutlich zu begrenzen. Speziell in einigen Küstenregionen Europas und den USA gelten diese strengeren Regeln bereits jetzt und werden 2020 nochmals verschärft. War 2012 noch bis zu 4.5% (Massenprozent) Schwefel im Kraftstoff erlaubt, werden es 2020 nur noch 0,1% sein. Schiffseigner müssen also bis 2020 auf deutlich teureren schwefelarmen Kraftstoff umstellen oder sie installieren eine Abgas-reinigungsanlage, die das Schwefeldioxid aus dem Abgas wäscht. Der SO_2–Gehalt im gefilterten Abgas, das aus diesen Scrubbern genannten Gaswäschern emittiert wird, muss aber permanent gemessen werden [1]. Benötigt wird also eine robuste, günstige und dennoch genaue Messmethode, um diese Aufgaben zu erfüllen.

Die hier vorgestellten Arbeiten entstanden im Eurostars Projekt MASUM, das Ende März 2019 abgeschlossen wurde.

Resonante Photoakustik

Der photoakustische Effekt wurde bereits 1880 von Alexander Graham Bell beschrieben [2] und kann dazu verwendet werden, um ein robustes und kostengünstiges Messsystem für die Messung von SO_2 zu realisieren. Genutzt wird dabei, dass Festkörper, Gase und Flüssigkeiten bei charakteristischen Wellenlängen Licht besonders stark absorbieren. Durch die Absorption von Photonen mit diesen charakteristischen Wellenlängen wird der Stoff intern angeregt, relaxiert dieser aus dem angeregten Zustand über Rotations- und Vibrationsübergänge erwärmt sich dabei der Stoff und dann seine Umgebung. Bei einem in einem festen Volumen eingeschlossenen Gas kann dies als Druckerhöhung registriert werden. Regt man das Gas mit einer modulierten Lichtquelle der Frequenz *f* an, kann das periodisch entstehende Drucksignal als Schallwelle derselben Frequenz über ein Mikrophon nachgewiesen werden. Die Amplitude des gemessenen Schallsignals ist dabei proportional zur Konzentration des Zielgases im Volumen [3]. Wie in einem Musikinstrument kann das akustische Signal durch ein gezieltes Design des Messvolumens (Messzelle) verstärkt werden. Stimmt die Anregungsfrequenz *f* mit der akustischen Resonanzfrequenz der Messzelle überein, ist eine Verstärkung des akustischen Signals um mehrere Größenordnungen möglich [4–6]. Der prinzipielle Aufbau einer solchen resonanten photoakustischen Messzelle ist in Abbildung 1 dargestellt.

Durch den Einsatz von Finite-Elemente Simulationswerkzeugen kann die Messzelle für den jeweiligen Einsatzzweck optimiert werden

[6]. Leistungsstarke Silizium-Mikrophone, wie sie in Mobiltelefonen zum Einsatz kommen, und LED's als Lichtquellen ermöglichen den Bau von photoakustischen Messsystemen für Gase, für die bisher teure laserbasierten Messsysteme zum Einsatz kamen [7,8].

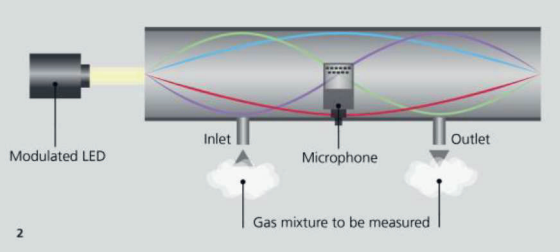


Abb. 1: Prinzip einer resonanten photoakustischen Messzelle. Die Modulationfrequenz der LED als Lichtquelle stimmt mit der akustischen Resonanzfrequenz der Messzelle überein, dadurch bilden sich stehende akustische Wellen im Resonator, welche zu einer Verstärkung des akustischen Signals um mehrere Größenordnungen führen können.

Das Messsystem

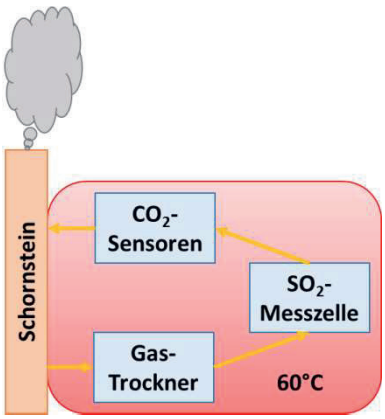


Abb. 2: Schematischer Aufbau des MASUM-Messsystems auf dem Testschiff. Das Abgas wird aus dem Schornstein entnommen und durch einen Gastrockner geleitet, dabei kühlt es auf 60°C ab. Es durchläuft dann die SO_2-Messzelle und zwei CO_2-Sensoren bevor es wieder in den Schornstein zurückgeleitet wird.

In Abbildung 2 ist der schematische Aufbau des MASUM-Messsystems dargestellt, Abbildung 3 zeigt den realen Aufbau. Nach der Entnahme aus dem Schornstein durchläuft das Gas einen Gastrockner (MircoGass von Perma Pure), bevor es erst durch die SO_2-Messzellle und dann durch zwei CO_2-Sensoren [9] geleitet wird. Das gemessene Gas wird dann wieder in den Schornstein zurückgeleitet. Die Temperatur des Gases wird innerhalb des gesamten Messprozesses auf 60°C gehalten.
Das gesamte Messsystem befindet sich im Schornsteinturm des Testschiffes. Dies ist ein halboffenes Treppenhaus neben dem Schornstein (siehe Abbildung 6) und bietet nur bedingt Schutz vor Wind und Wetter. Die Temperaturen am Aufstellort schwanken von -30°C bis 50°C, bei laufendem Motor herrschen starke Vibrationen und eine Umgebungslautstärke von bis zu 80 dBa. Zum Schutz des Gesamtaufbaus wurden alle Einzelkomponenten in einen Schaltschrank verbaut, der die Komponenten vor Regen und Vogelkot schützt. Der installierte Schaltschrank ist in Abbildung 3 zu sehen.



Abb. 3: MASUM-Messsystem im Schaltschrank auf dem Testschiff. Oben links im Schaltschrank befindet sich der Gastrockner von Perma Pure, daneben die Steuerung des beheizten Gasentnahmeschlauches sowie die 24V Versorgung der Elektronik, darunter sieht man B2 und B3 die CO_2-Sensoren; unten links befindet sich das SO_2-Messsystem und die Auswerteelektronik mit einem Display zur direkten Überprüfung des Messsignals.

Oben links im Schaltschrank befindet sich der Gastrockner daneben die Steuerung des beheizten Gasentnahmeschlauches sowie die 24V Versorgung der Elektronik, darunter sieht

man B2 und B3 die CO_2-Sensoren; unten links befindet sich das SO_2-Messsystem und die Auswerteelektronik mit einem Display zur direkten Überprüfung des Messsignals.
Die ermittelten Daten werden mittels RS458 Schnittstelle gemäß dem NMEA 183 Protokoll an einen zentralen Rechner übertragen, der die Speicherung der Daten des MASUM-Systems sowie eines installierten Referenzsystems IMR 7500 und der Antriebszustandsdaten des Schiffes übernimmt.

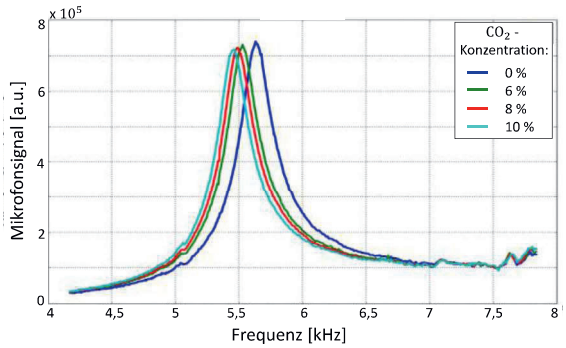


Abb. 4: Mikrofonsignal in der Messzelle über die Modulationsfrequenz der Lichtquelle. Mit steigender CO2-Konzentration sinkt das Messsignal ebenso wie die Resonanzfrequenz der Messzelle.

Die verwendeten Komponenten wurden vor dem Einbau im Schiff ausführlich im Labor getestet [7], trotzdem ist ein aufwendiger Feldtest notwendig ist.
Das Abgas eines Schiffes lässt sich nur rudimentär im Labor aus den Hauptkomponenten Sauerstoff, Stickstoff, Kohlendioxid und Spurengasen wie Schwefeldioxid nachstellen. Bereits im Laborversuch mit ternären Mischungen auf Basis von Stickstoff hatte sich gezeigt, dass Sauerstoff- und Kohlendioxidgehalt eine Auswirkung auf das Messsignal hatten. Mit steigender Kohlendioxidkonzentration nimmt bei konstantem Schwefeldioxidgehalt die Amplitude des Messsignals sowie dessen Resonanzfrequenz ab (siehe Abbildung 4). In Abbildung 5 sieht man den Einfluss des Sauerstoffgehaltes auf das Messsignal, hier steigt die Amplitude leicht an, die Resonanzfrequenz sinkt dagegen.
Tests in der realen deutlich komplexeren Mischung des Abgases sind daher zwingend notwendig. Die bereits genannten rauen Umgebungsbedingungen am Aufstellort des Systems lassen sich genauso wenig im Labor simulieren. Ein weiterer zu klärender Punkt war der tatsächliche notwendige Messbereich für Bestimmung des CO_2-Gehaltes. Der Motor eines Autos wird so betrieben, dass der gesamte Sauerstoff in Kohlendioxid umgewandelt wird, der Motor eines Schiffes dagegen wird zur Reduzierung von NO_x Emissionen mit einem variablen Luftüberschuss betrieben, dieser reduziert die Effizienz der Verbrennung und führt zu einem nicht vorhersehbaren Sauerstoff- und Kohlendioxidgehalt im Abgas. Parallel zum hier entwickelten Messsystem wurde auf dem Schiff ein Referenzsystem von IMR (IMR 7500) installiert. Das IMR 7500 misst den O_2, CO_2, SO_2, NO und NO_x-Gehalt des Abgases. Die Daten beider Systeme werden zentral von einem Rechner gespeichert, von diesem wird außerdem der Zustand des Motors (Leistung, Drehzahl, Verbrauch, usw.) und die Position und Geschwindigkeit des Schiffes mitaufgezeichnet.

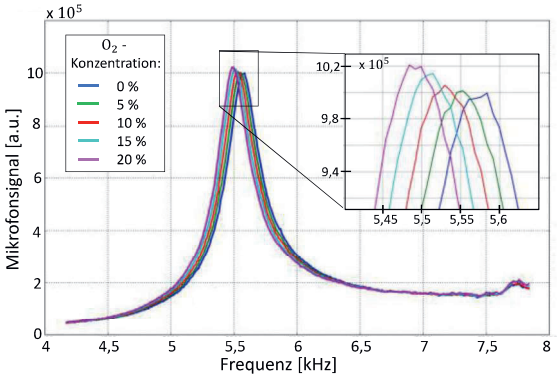


Abb. 5: Mikrofonsignal in der Messzelle über die Modulationsfrequenz der Lichtquelle. Mit steigende Sauerstoff-konzentration steigt das Messsignal bei sinkender Resonanzfrequenz der Messzelle.

Feldtest

Der Feldtest des MASUM-Messsystem fand auf dem Containerschiff MS Priamos statt. Dieses Containerschiff fährt überwiegend im Verkehr auf der Nord- und Ostsee und verbrennt daher standardmäßig schwefelarmen Kraftstoff, was zu Abgaskennzahlen wie bei einem vergleichbaren Schiff mit Abgasreinigungs-anlage führt. Dieses Schiff, aus dem Jahr 2011, ist mit einer Gross Tonnage von 11550, einer Länge von 158 m und einer Breite von 23,5 m eher klein, aber genau die Kategorie von Schiff, die zukünftig von den strengeren Abgasnormen betroffen sein wird. Ein Bild der MS Priamos mit dem markierten Einbauort des Messsystems ist in Abbildung 6 zu sehen.

Die Installation des Messsystems erfolgte in mehreren Etappen zwischen Juli und August 2018, da diese im laufenden Betrieb des Schiffes durchgeführt werden musste. Seit August 2018 wurden dann kontinuierlich Daten aufgezeichnet die ausgewertet werden konnten.
Abbildung 7 und 9 zeigen typische Tage im Fahrbetrieb des Schiffes. Die Graphen sind zweigeteilt, im oberen Bereich sind die Daten den MASUM-System dargestellt, der untere Bereich zeigt die Messdaten des IMR-Systems und die Fahrdaten des Schiffes. Aufgetragen sind diese Daten über der Sweep- bzw. Messungs-Nummer, mit den derzeitigen Einstellungen misst das System etwa 120mal pro Tag.
In Abbildung 7 sind in die direkt auf dem System ausgewertet Daten rot dargestellt, dazu wurden online aus den Messkurven, wie sie in Abbildung 4 und 5 zu sehen sind, die Höhe des Resonanzpeaks sowie die Resonanzfrequenz ermittelt. Die schwarze Kurve zeigt die Ergebnisse der nachträglichen Auswertung der Rohdaten, die aus dem System übertragen wurden, wie man sieht sind diese identisch, sodass im Folgenden nur noch die direkt ausgewerteten Daten betrachtet werden.



Abb. 6: Das Feldtestschiff MS Priamos; ca. 5 m unterhalb des Schornsteinendes ist das Messsystem im Schornsteinturm am Hauptschornstein verbaut.

Ist die CO_2 und O_2 Konzentration bekannt, kann man aus der Stärke des Mikrofonsignals die SO_2-Konzentration ermitteln. Die Abbildungen 8 und 9 zeigen, dass die gemessenen SO_2-Konzentrationen im Bereich von 0-30 ppm lagen.

In den dargestellten Messkurven aus den Abbildungen 7 und 9 sieht man einen deutlichen Versatz der MASUM-Messsignale gegenüber den Fahrdaten. Diese sind reproduzierbar und kommen aus dem Zusammenspiel von Gesamtinnenvolumen des Messsystems und der Fördermenge der Pumpe, mit der das Abgas aus dem Schornstein entnommen wird. Insbesondere der lange Probenentnahmeschlauch ist hier für diese Verzögerung verantwortlich. In der nächsten Entwicklungsstufe des Systems soll dieser ersatzlos entfallen.

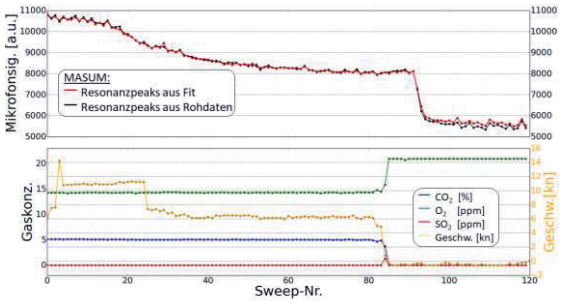


Abb. 7: Messdaten im Fahrbetrieb. Im oberen Graphen sind zwei Auswertungsvarianten der Messdaten des MASUM-Systems dargestellt. Rot die direkt auf dem System erfolgte Fit-Auswertung und schwarz die nachträgliche Auswertung der Rohdaten. Der untere Graph zeigt die vom IMR System aufgezeichneten Messdaten für O2, CO2 und SO2 sowie Fahrdaten des Schiffes. Es besteht ein zeitlicher Versatz des SO2-Messignals des MASUM-Systems und dem Motorzustand.

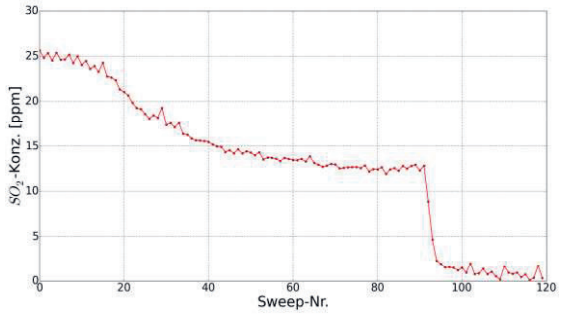


Abb. 8: Auswertung des SO2-Messignals aus Abbildung 7. Das Schiff hatte kurz vor dem Log-Beginn seine Leistung und Geschwindigkeit verringert, wodurch während der Aufzeichnung der SO2-Gehalt von

25 ppm auf 12 ppm fällt. Nach dem Ausschalten des Motors geht das SO2-Signal etwas verzögert auf 0 ppm.

Alle gemessenen SO_2-Werte im Fahrbetrieb lagen bei unter 30 ppm und damit im Bereich, den man beim Verbrennen von schwefelarmem Marinekraftstoff erwartet. Leider lieferte das Referenzsystem nur selten und dann nicht reproduzierbare SO_2-Werte, zusammen mit dem Hersteller wird derzeit evaluiert warum dies der Fall war. Die vom Referenzsystem ermittelten Daten für Sauerstoff und Kohlendioxid sowie die verschiedenen Stickoxide waren alle plausibel und reproduzierbar.

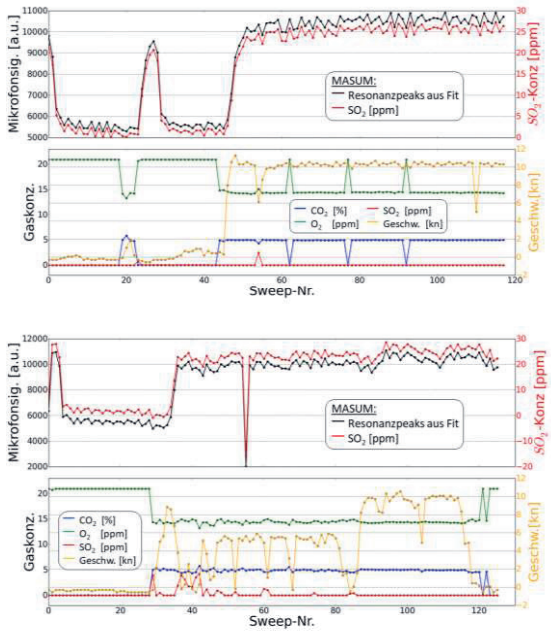


Abb. 9: Weitere Messdaten im Fahrbetrieb. Auch hier besteht ein zeitlicher Versatz des SO2-Messignals des MASUM-Systems und den Fahrdaten, dieser Versatz ist reproduzierbar und lässt sich durch das Gesamtvolumen des Systems und die verwendete Pumpleistung erklären.

Ausblick

Der Feldtest mit dem bestehenden System soll im Mai 2019 abgeschlossen werden, danach ist geplant basierend auf den gewonnenen Daten und Erkenntnissen einen Nullserienprototypen zu entwickeln und diesen mit dem Projektpartner Rüeger SA spätestens 2020 auf den Markt zu bringen.

Danksagung

Diese Arbeit wurde gefördert durch die Europäische Union und das deutsche Bundesministerium für Bildung und Forschung mit der Förderzeichen E! 9727-MASUM.

Wir danken Nadine und Kay Paschen sowie Christian Jessen von TXMarine für die Unterstützung bei der Durchführung des Feldtests.

Literaturnachweis

[1] I.M. Organisation, 2009 GUIDELINES FOR EXHAUST GAS CLEANING SYSTEMS MEPC 59/24/Add.1, New York, 2009.

[2] A.G. Bell, J.G. Dingler, A. G. Bell's Photophon, Polytechnisches Journal 238 (1880) 409–413.

[3] F.J.M. Harren, J. Mandon, S.M. Cristescu, Photoacoustic Spectroscopy in Trace Gas Monitoring, in: Encyclopedia of Analytical Chemistry, John Wiley & Sons, Ltd, 2006; doi: 10.1002/9780470027318.a0718.pub2.

[4] Geometrical optimization of a longitudinal resonant photoacoustic cell for sensitive and fast trace gas detection, Review of Scientific Instruments 67 (1996) 2914–2923; doi: 10.1063/1.1147072.

[5] P. Hess, Resonant photoacoustic spectroscopy, in: Physical and Inorganic Chemistry, Springer Berlin Heidelberg, Berlin, Heidelberg, 1983, pp. 1–32; doi: 10.1007/3-540-12065-3_1.

[6] B. Parvitte, C. Risser, R. Vallon, V. Zéninari, Quantitative simulation of photoacoustic signals using finite element modelling software, Applied Physics B 111 (2013) 383–389; doi: 10.1007/s00340-013-5344-2.

[7] M. El-Safoury, M. Dufner, C. Weber, K. Schmitt, H.-F. Pernau, B. Willing, J. Wöllenstein, Resonant Photoacoustic Gas Monitoring of Combustion Emissions, Proceedings 2 (2018) 962; doi: 10.3390/proceedings2130962.

[8] J. Kapp, C. Weber, K. Schmitt, H.-F. Pernau, J. Wöllenstein, Resonant Photoacoustic Spectroscopy of NO2 with a UV-LED Based Sensor, Sensors 19 (2019) 724; doi: 10.3390/s19030724.

[9] J. Huber, J. Wöllenstein, S. Kolb, A. Dehé, F. Jost (Eds.), E6.3 - Miniaturized Photoacoustic CO2 Sensors for Consumer Applications. AMA Service GmbH, P.O. Box 2352, 31506 Wunstorf, Germany, 2015; doi: 10.5162/SENSOR2015/E6.3.

DOI 10.5162/sensoren2019/5.3.1

Key Design Elements of a Radio-Acoustic-Sounding System for Industrial Applications

Dominik Exel[1], Stefan Schuster[2], Stefan Scheiblhofer[2], Dominik Zankl[2], Vera Ganglberger[2], Johann Reisinger[2], Bernhard G. Zagar[1]
[1] Johannes Kepler University Linz, Institute for Measurement Technology
Altenbergerstraße 69, A-4040 Linz
[2] voestalpine Stahl GmbH
voestalpine-Straße 3, A-4031 Linz

Abstract

This article discusses compact radio-acoustic-sounding systems (RASSs) for industrial applications. The electromagnetic-acoustic interaction enables to measure temperatures of gases without contact and spatially resolved. For this purpose, the speed of a modulated sound pulse, travelling through the gas with a velocity proportional to the temperature, is determined with a Doppler radar. The optimum sound pulse´s in terms of its frequency is discussed. The so called Bragg condition is essential to obtain sufficient receive power. The two most important conditions for an optimal Bragg reflection are the tuning of the wavelengths and an optimal collocation of the sources. These conditions are calculated numerically and verified with real-world measurement data. One purpose of this paper is to summarize physical effects that are key elements to the function of the RASS.

Keywords: Radio-Acoustic-Sounding System, RASS, Doppler radar

1 Introduction

Non-contact, continuous measurement methods for determining the temperature of gases are used to automate industrial processes. There are many requirements for an universally applicable temperature measurement system. For example, the functional guarantee in a dusty, smoky, and aggressive industrial air, accuracy, safety and last but not least a cost-effective realization. The idea of the radio-acoustic-sounding system (RASS), which is known from the meteorology [1], is to measure the speed of sound by Doppler radar. The measurement method is based on the effect that an acoustic wave modulates the density of the air. When an electromagnetic wave passes through the air, in which these acoustically induced changes in density occur, a small fraction of the signal will be reflected at each of the transitions. Utilizing the Bragg effect the signal return power can be improved. Therefore the Bragg condition must be fulfilled

$$\lambda_e = 2 \cdot \lambda_a, \qquad (1)$$

with λ_e the radar wavelength and λ_a the acoustic wavelength. This leads to a constructive interference and afterwards to a measurable signal power [2, 3].
Due to the fact that the pulsed acoustic wave (red drawn in figure 1), which moves with respect to the source and thus enables spatially resolved measurements, represents the Bragg-reflector, the backscattered radar signal (black drawn in figure 1) exhibits a Doppler-shift [4, 5]

$$f_\mathrm{d} = \frac{2(c_\mathrm{a} + v_\mathrm{w})}{c_\mathrm{e}} f_\mathrm{e}. \qquad (2)$$

c_e describes the speed of light ($2997 * 10^5 \ \frac{\mathrm{m}}{\mathrm{s}}$), and f_e the transmit frequency of the radar. This Doppler-shift f_d is proportional to speed of sound c_a and speed of wind v_w in the propagation direction of the electromagnetic wave. To determine the temperature with this system, the relationship between speed of sound and temperature must be known and will be discussed in Section 2.

DOI 10.5162/sensoren2019/5.3.1

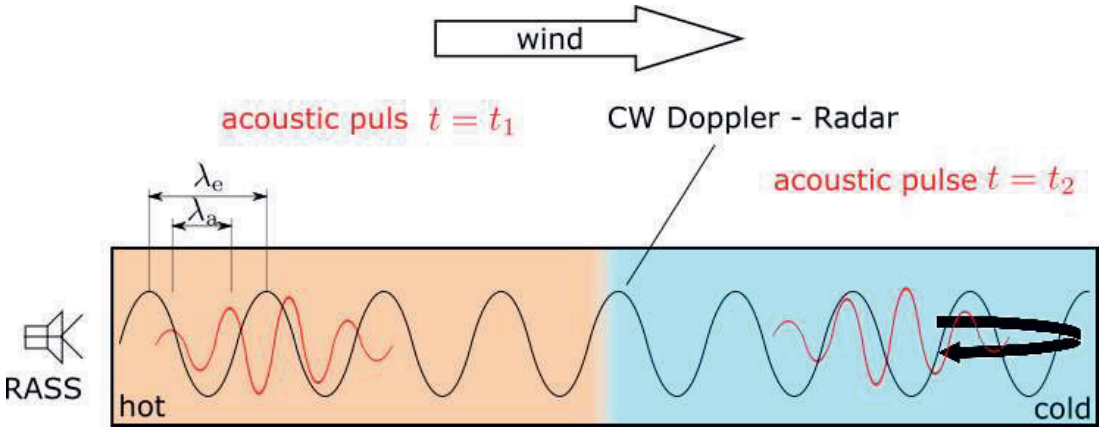


Figure 1: The speed of a transmitted sound pulse is composed of the speed of sound and a wind speed and is measured by a Doppler radar. The acoustic pulse travels through the gaseous medium of interest. By measuring its velocity, spatial temperature measurement is possible.

In industrial applications, the reliability requirements of a measuring system are very high. It is therefore beneficial to characterizes the limits of a system, as it has been done for remote temperature measurement with multispectral camera [6, 7, 8]. Thus it is important to know the backscattering efficiency and this is discussed in Section 3. Section 4 describes the measurement setup with a level plan. Furthermore, the importance of optimal collocation is highlighted. Measurement results of the physical effects are presented and discussed in Section 5.

2 Speed of sound - temperature relationship

The wave velocity c_a is calculated with the rate of change of the pressure according to the density at atmospheric pressure. Feynman describes it in the derivation of the wave equation for the sound [9]

$$c_\mathrm{a}^2 = \left(\frac{\mathrm{d}P}{\mathrm{d}\rho}\right)_0, \tag{3}$$

with the pressure P in Pa and density ρ in $\frac{\mathrm{kg}}{m^3}$. The adiabatic connection between P and ρ is

$$P = const. \cdot \rho^\kappa, \tag{4}$$

with the adiabatic exponent κ. The derivative from (4) to ρ is

$$\frac{\mathrm{d}P}{\mathrm{d}\rho} = \frac{\kappa P}{\rho}. \tag{5}$$

Combining (3) with (5) leads to

$$c_\mathrm{a}^2 = \frac{\kappa P}{\rho}. \tag{6}$$

By extension in the numerator and denominator with the volume V, the relations $PV = NkT$ and $\rho V = \mu$ can be used

$$c_\mathrm{a}^2 = \frac{\kappa N k T}{\mu}. \tag{7}$$

T describes the temperature in K, N the Avogadro-constant $(6.0221 * 10^{23}\ \mathrm{mol}^{-1})$, k the Boltzmann-constant $(1.3806 * 10^{-23}\ \frac{\mathrm{J}}{\mathrm{K}})$ and μ the molecular mass in $\frac{\mathrm{kg}}{\mathrm{mol}}$. Further, the relation $R_u = Nk$ can be used and leads to the known formula for the speed of sound

$$c_\mathrm{a}^2 = \frac{\kappa R_\mathrm{u} T}{\mu}. \tag{8}$$

$\kappa = 1.402$ and $\mu = 0.02896\frac{\mathrm{kg}}{\mathrm{mol}}$ can be used for dry air and thus there is a clear correlation between temperature and sound velocity [1]

$$c_\mathrm{a} = 20.045\sqrt{T}. \tag{9}$$

2.1 Speed of sound of a gas mixture

In industrial applications, it is often necessary to determine the temperature of gas mixtures. For this purpose it is necessary to know the speed of sound for a mixture

$$c_{\mathrm{a_mix}} = \sqrt{\frac{(\sum_i^M x_i \gamma_i) R_\mathrm{u} T}{(\sum_i^M x_i \mu_i)}}, \tag{10}$$

with *M* components. x_i describes the fraction of the i component and γ_i the molecular heat capacity [10].

3 RASS equation

3.1 Fluctuation of the refractive index

To be able to calculate the received power, it must be known how much the electromagnetic waves are scattered by the artificial radar target generated by the modulated acoustic

pulse. Therefore the fluctuation of the refractive index is investigated [11].
Balsley and Gage [12] describes in their work a term for the refractive index n

$$n = 1 + 7.76 * 10^{-7} \frac{p_\mathrm{a}}{T} + \\ +3.73 * 10^{-3} \frac{e}{T^2} - 40.3 \frac{N_\mathrm{e}}{f^2}, \quad (11)$$

for microwave frequencies in the troposphere. p_a describes the atmospheric pressure in Pa, T the temperature in K, e the partial pressure of water vapor in Pa, N_e the density of the free electrons in $\mathrm{m}^{-3}$and f the radar frequency in Hz. The second term on the right side describes the dipole moments due to the polarization field of neutral molecules. The last two terms on the right side contribute only marginally to the refractive index and can be neglected in further calculations. Thus (11) simplifies to

$$n = 1 + 7.76 * 10^{-7} \frac{p_\mathrm{a}}{T} \quad (12)$$

and the relative change of the electromagnetic refractive index can be calculated using

$$\frac{\Delta n}{n} = 7.76 * 10^{-7} \left(\frac{\Delta p}{T} - \frac{p_\mathrm{a}}{T^2} \Delta T \right). \quad (13)$$

Combining the adiabatic temperature gradient $\frac{\Delta T}{\Delta z} = \frac{-\rho}{c_\mathrm{p}}$ with the hydrostatic equation $\Delta p = -\rho g \Delta z$, leads to $\Delta T = \frac{\Delta p}{\rho c_\mathrm{p}}$ [13]. Inserting this result into (13) and using $p = 1.013 * 10^5$ Pa, $T = 273$ K, $\rho = 1.2\ \frac{\mathrm{kg}}{m^3}$ and $c_\mathrm{p} = 1005\ \frac{\mathrm{J}}{\mathrm{kg\,K}}$, the expression for refractive index simplifies to

$$\frac{\Delta n}{n} \approx 2 * 10^{-9} \Delta p. \quad (14)$$

The amplitude of the sound pressure variation for an isotropic radiator at a distance r is given by [14]

$$\Delta p = \sqrt{\frac{\rho c_\mathrm{a} P_\mathrm{a}}{4\pi r^2}}, \quad (15)$$

where P_a is acoustic transmitted power in W. Finally, (15) can be used in (14) and leads to

$$\frac{\Delta n}{n} \approx 1.6 * 10^{-8} \frac{\sqrt{P_\mathrm{a}}}{r}. \quad (16)$$

The change in the refractive index is thus inverse proportional to the distance r and proportional to the square root of the acoustic transmission power P_a.

3.2 Receive power

Marshall [15] and Clifford [16] describe in their work an expression for the receive power P_r of radar waves scattered at sound waves

$$P_\mathrm{r} = \left(\frac{\Delta n}{n} \right)^2 \frac{4\pi^2 P_\mathrm{t} G_\mathrm{t} G_\mathrm{r}}{\lambda_e^2 r^2} \left(\frac{\Delta r}{2} \right)^2 \cdot \\ sinc^2 \left[(2k_e - k_a) \frac{\Delta r}{r} \right]. \quad (17)$$

P_t describes the transmitted radar power in W, G_t and G_r the gain of the transmitting respectively receiving antenna in dBi, $k_\mathrm{e} = \frac{2\pi}{\lambda_\mathrm{e}}$ and $k_\mathrm{a} = \frac{2\pi}{\lambda_\mathrm{a}}$ the radar and acoustic wave number respectively and r the range of the target in m. So r describes the current distance of the sound wave to the measurement system. If the radar is monostatic, then G_t = G_r. $\Delta r = N_\mathrm{BRAGG} \cdot \lambda_\mathrm{a}$ represents the length of the acoustic packet and can be calculated with the number of the acoustic cycles N_BRAGG and the acoustic wavelength λ_a. Combining the relative change in refractive index (16) with (17) leads to the RASS equation

$$P_\mathrm{r} \propto \left(\frac{\lambda_\mathrm{a}}{\lambda_\mathrm{e}} \right)^2 \frac{P_\mathrm{t} G_\mathrm{t}^2 P_\mathrm{a} N_\mathrm{BRAGG}^2}{r^2} \cdot \\ sinc^2 \left[(2k_\mathrm{e} - k_\mathrm{a}) \frac{\Delta r}{r} \right]. \quad (18)$$

The RASS equation shows that the receive power will be proportional to r^{-2} instead of the traditional r^{-4} for the classical radar equation [17]. The r^{-2} dependence is due to the spherically divergent acoustic wave. This is a great benefit of the focused geometry [2].

3.3 Bragg condition

As already mentioned, the optimal Bragg condition is only valid if the radar wavelength is twice as large as the acoustic wavelength see (1). Substituting the Bragg condition into the RASS equation (18) maximizes the sinc function. This means that the backscattered electromagnetic power will be maximized. Due to the small change in the refractive index, or the low backscatter of the electromagnetic power on an acoustic wavefront, it is endeavored to hit the Bragg condition as well as possible. This leads to a constructive interference of the scattered electromagnetic power at the acoustic wavefronts in the propagation direction. In addition, the received power can be improved by increasing N_BRAGG. The received power is proportional to N_BRAGG^2. Thus, doubling the Bragg cycles results in an increase in received power by 6 dB. Therefore, a compromise must

be found between the spatial resolution and sufficient signal-noise ratio. The importance of an optimal Bragg condition is shown with real-world measurement results in Section 5.

4 Realization, measurement setup

There are several possible RASS configurations. The Radar can be monostatic or bistatic. In addition, the radar and the acoustic sources can be either continuous wave (CW) or pulsed [4]. In this work a monostatic CW radar with pulsed sound is presented. This configuration has the advantage that all previously emitted sound pulses have sufficient time to die away and thus there are no significant reflections in enclosed environments.

4.1 Level plan

The RASS system is designed to measure indoor room temperatures. For this purpose, a $10\,\mathrm{GHz}$ synthesizer with $10\,\mathrm{dBm}$ output power was chosen for the radar system.
Fig. 2 shows a block diagram of the setup of the RASS. To get the highest possible output power, the $1\,\mathrm{dB}$ compression point of the $23\,\mathrm{dB}$ amplifier in front of the circulator was approached. The total loss in the transmit path is $\approx 8\,\mathrm{dB}$, which includes the insertion losses, splitter loss, attenuator and cable loss. The antenna has a gain of $20\,\mathrm{dBi}$ with a $-3\,\mathrm{dB}$ beam width of $14\,°$. Thus, the equivalent isotropically radiated power (EIRP) is $43\,\mathrm{dBm}$.
For 100 acoustic cycles ($N_{BRAGG} = 100$) the radar cross section (RCS) is $-60\,\mathrm{dBm^2}$. A description as well as the calculation of the RCS is discussed by the authors in [18]. With the RASS equation (18) the path loss can be calculated. For this radar setup and a distance of $5\,\mathrm{m}$, the path loss is approximately $-90\,\mathrm{dB}$. Thus, the receive power is $-87\,\mathrm{dBm}$. After the mixer, the low-frequency Doppler signal is again amplified by $20\,\mathrm{dB}$ with a programmable gain amplifier ($50\,\Omega$ output). A power of $-75\,\mathrm{dBm}$ is digitalized at the analog-to-digital converter (ADC) with $1\,\mathrm{V}$ reference voltage. With $16\,\mathrm{bit}$ ADC resolution, the signal can be detected.
Due to the Bragg condition and $10\,\mathrm{GHz}$ radar frequency, wavelengths of $\lambda_a = 15\,\mathrm{mm}$ $\lambda_e = 30\,\mathrm{mm}$ result at room temperatures.
The right path of the block diagram shows an ultrasonic loudspeaker connected to an audio amplifier. The bandwidth of the tweeter is $4 - 25\,\mathrm{kHz}$.

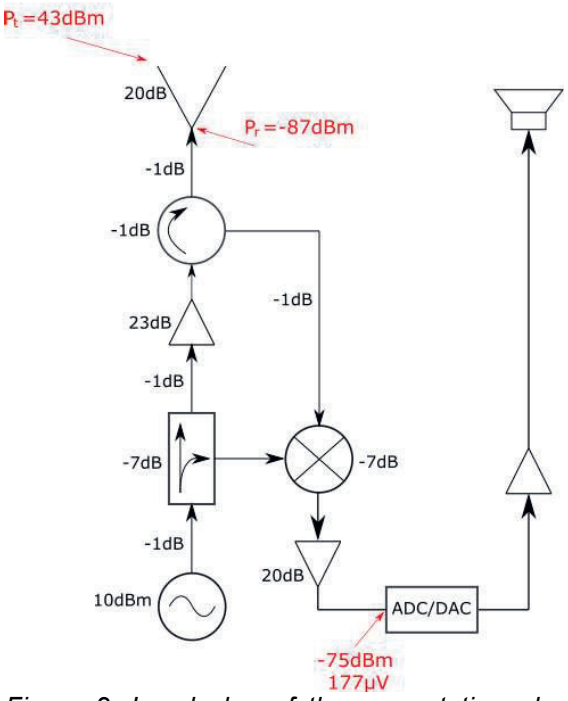


*Figure 2: Level plan of the monostatic radar setup. -75*dBm *input power at the ADC can be detected with a resolution of 16 bit and 1*V *reference voltage.*

4.2 Collocation

Another important prerequisite for RASS functionality is collocation. This means, that the phase center of the radar horn and the loudspeaker should be identical. As proposed by Weiss [4, 5], a fine metal grid can be used to enable this. Figure 3 shows the set-up of the measurement head.

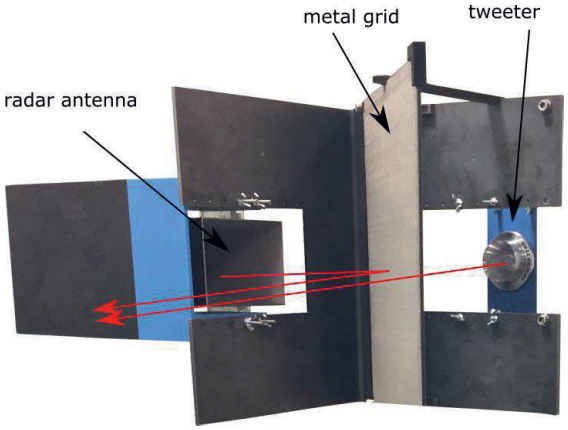


Figure 3: To get the highest possible reflectivity of the radar waves, the phase center of the radar horn and the loudspeaker should be identical, i.e. collocated. Utilizing a metal grid as proposed in [4] enables this.

The metal grid redirects the radar waves by 90° and the sound waves travels straight through the grid. Thaller [19] shows in his work, that the grid has only a little influence of

the sound propagation and can be neglected. If the phase centers are not virtually the same, then there is no optimum overlapping of the wavefronts and thus a lower receive power.
Influence of a horizontal wind on the receive power was theoretically studied by Peters et al. [20], Karyukin [21] and Kon [22]. A horizontal wind causes a virtual displacement of the sound source. The decrease of the receive power can be calculated by

$$I = e^{\frac{-2\rho_v^2}{D^2}}, \tag{19}$$

where ρ_v is the virtual horizontal displacement of the sound source and D is the e^{-1} diameter of the radar antenna. Equation (19) is verified with real world measurement data in the next section.

5 Measurement results

As already mentioned, an optimal Bragg condition is essential to obtain sufficient receive power. The measurements were performed in a closed laboratory room. The room temperature was $25.8\,°\mathrm{C}$ and was measured with a PT100 thermocouple type K. For different sound frequencies, which can be calculated by evaluation of

$$f_\mathrm{a} = \frac{\sqrt{\frac{R_\mathrm{u}\kappa T}{\mu}}}{\lambda_\mathrm{a}} = 1336.33 \cdot \sqrt{T}, \tag{20}$$

temperatures between $12\,°\mathrm{C}$ and $38\,°\mathrm{C}$ were chosen. Figure 4 shows the levels of electromagnetic backscattered power at different transmitted sound frequencies (temperatures). If the Bragg condition is matched exactly, maximum power will be backscattered. This is the case when the temperature of the transmitted frequency coincides with the actual temperature. For the measurement a sound packet length Δr of $1.5\,\mathrm{m}$ was chosen. With a threshold of $-6\,\mathrm{dB}$ a temperature range of $\approx 20 - 32°\mathrm{C}$ can be meaningful covered under these conditions. This corresponds to $\pm 6°\mathrm{C}$ of the temperature to be measured. The longer the sound packet length (more Bragg cycles N_BRAGG), the smaller the temperature range that can be covered. This is discussed in detail in [1] and [18]. With this setup and settings, a signal-to-noise ratio (SNR) of $17\,\mathrm{dB}$ is possible with an optimal Bragg condition.

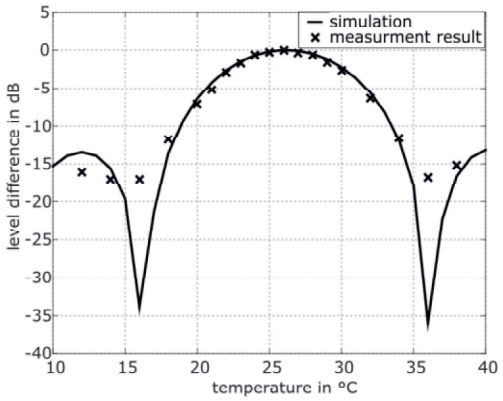


Figure 4: Comparison between the simulation and measurement result of the Bragg condition. Parameter: $N_\mathrm{BRAGG} = 66$, $\lambda_\mathrm{e} = 30\,\mathrm{mm}$, number of averaged measurements $= 200$

To verify the collocation (Section 4.2), the speaker was moved in $5\,\mathrm{mm}$ steps. Figure 5 shows the measurement results compared to the simulation (19). If the collocation is not maintained exactly, there is an exponential loss in the received power. The measurement result shows a good agreement with the simulation.

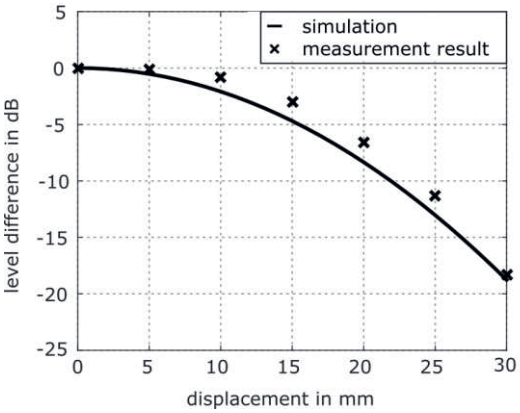


Figure 5: Horizontal displacement of the speaker in 5 mm steps. If the collocation is not maintained exactly, there is an exponential loss in the received power (19). Parameter: $N_\mathrm{BRAGG} = 66$, $\lambda_\mathrm{a} = 15\,\mathrm{mm}$, $\lambda_\mathrm{e} = 30\,\mathrm{mm}$, number of averaged measurements $= 200$

6 Conclusion and future work

This article describes a radio-acoustic-sounding system, which can be used in industrial applications. The two most important conditions for an optimal Bragg reflection were shown and verified with real world measurement data.
It is important to characterize the limits of a system and the results presented in this paper are important steps for a successful and reliable assertion regarding the quality of a RASS measurement.

For industrial applications, it is important to know the measurement uncertainty of a measuring system. The Cramer-Rao lower bound [23, 24] can be used to calculate the lower bound of the temperature variance. These accuracy limits, achievable resolution and methods of signal processing will be discussed in a future work.

Acknowledgement

The authors gratefully acknowledge the partial financial support for the work presented in this paper by the Austrian Research Promotion Agency (FFG) under contract grant 861570 and voestalpine Stahl GmbH.

Literatur

[1] Stuart Bradley, "*Atmospheric Acoustic Remote Sensing*", CRC Press, ISBN 0-8493-3588-4, 2008.

[2] J. A Saffold, F. Williamson, K. Ahuja, L. Stein, M. Muller, "*Radar-acoustic Interaction for IFF Applications*", Radar Conference, Waltham USA, S. 198-202, 1999

[3] Benjamin Moussa Daas, "*Die elektromagnetischakustische Wechselwirkung und ihre Anwendung zur berührungslosen Entfernungs- und Temperaturmessung in geschlossenen Räumen*", Shaker Verlag, Aachen, ISBN 3-8265-2148-X, 1996

[4] M. Weiß, R. Knöchel, "*A Monostatic Radio-Acoustic Sounding System Used as an Indoor Remote Temperature Profiler*", IEEE Transaction on Instrumentation and Measurement, Vol. 50, No. 5, 2001

[5] M. Weiß, R. Knöchel, "*Monostatic Radio-Acoustic Sounding System*", Geoscience and Remote Sensing Symposium, Hamburg, Germany, S. 29-31, 2001

[6] D. Exel, B. Zagar, S. Schuster, V. Ganglberger, J. Reisinger, "*Parameter Estimation and Performance Bounds for Radiometric Model-based Non-contact Temperature Measurement*", IEEE International and Measurement Technology Conference, Turin, Italien, S. 641-646, 2017

[7] D. Exel, B. Zagar, S. Schuster, V. Ganglberger, J. Reisinger, "*Parameterschätzung und Cramér-Rao-Schranken bei der berührungslosen Temperaturmessung"*, tm-Technisches Messen, De Gruyter Oldenbourg, Vol. 86, S. 45-56, 2019

[8] D. Exel, B. Zagar, S. Schuster, V. Ganglberger, J. Reisinger, " *Unsicherheitsbetrachtungen der Temperaturmessung aus Daten einer Multispektralkamera*", Tagungsband der 5. Tagung Innovation Messtechnik, 2017

[9] R. P. Feynman, R. B. Leighton, M. Sands, "*Feynman-Vorlesung über Physik 2 Strahlung und Wärme"*, De Gruyter, 6. Auflage, ISBN 978-3-11-036770-6, 2015

[10] F. Wächter, "*Zur Gasanalyse durch Messung von Schallgeschwindigkeit und Wärmeleitfähigkeit*", Proc. DAGA02, 2002, S. 386-387.

[11] D. Exel, S. Schuster, S. Scheiblhofer, D. Zankl, V. Ganglberger, J. Reisinger, B. Zagar, "*Ein kompaktes Radio-Acoustic-Sounding System zur Temperatur- und Strömungsgeschwindigkeitsmessung*", Tagungsband der 6. Tagung Innovation Messtechnik, 2019

[12] B.B. Balsley, K.S. Gage, "*The MST radar technique: Potential for middle atmospheric studies*", Pure and Applied Geophysics, Vol. 118, S. 452-493, 1980

[13] F. Bernhard, "Handbuch der Technischen Temperaturmessung", Springer Vieweg, ISBN 978-3-642-24506-0, 2014

[14] J. R. Kutia, W. Xu, "*A Portable Radio-Acoustic Sounding System for Temperature Profiling Within the Boundary Layer"*, Industiral Electronics Society, IECON, S. 2308-2313, 2014

[15] J. M. Marshall, A. M. Peterson, A. A. Barnes, "Combined Radar-Acoustic Sounding System", Applied Optics, Vol. 11, S. No.1, 108-112, 1972

[16] S.F. Clifford, T.I. Wang, "The range limitation on radar-acoustic sounding system (RASS) due to atmospheric refractive turbulence", IEEE Transactions on Antennas and Propagation, Vol. 25, S. 319-326, 1977

[17] M. Skolnik, "*Introduction to Radar Systems"*, McGraw-Hill Book Company, ISBN 0-07-057909-1, 1981

[18] D. Exel, S. Schuster, K. Thaller, B. Zagar, D. Zankl, S. Scheiblhofer, V. Ganglberger, J. Reisinger, "*Radio-Acoustic-Sounding System – ortsaufgelöste akustische Temperaturmessung im industriellen Umfeld*", e&i Elektronik und Informationstechnik, Vol. 135, No. 6, S. 389-395

[19] K. Thaller "*Messaufbau zur interferometrischen Schallfeldmessung mit tomografischer Rekonstrunktion*", Linz, 2008

[20] G. Peters, H. Timmermann, H. Hinzpeter, "*Temperature sounding in the planetary boundary layer by RASS-system analysis and results*", Int. J. Remote Sens., 4, S. 49-63, 1983

[21] G. A. Karyukin, "*Influence of wind on operation of radar-acoustic atmospheric sounding system*", Izv. Acad. Sci. USSR Atoms. Oceanic Phys., 18, S. 26-30, 1982

[22] A. I. Kon, "*A bistatic radar-acoustic atmospheric-sounding system*", Izv. Acad. Sci. USSR Atoms. Oceanic Phys., 17, S. 481-484, 1981, Innovation Messtechnik, Shaker Verlag GmbH,

[23] S. Schuster, S. Scheiblhofer, A. Stelzer "*DFT Frequency Estimation Threshold Level with Windowed Data*", International Instrumentation and Measurement Technology Conference, Sorrento, Italy, 2006

[24] S. M. Kay "*Fundermentals of Statistical Signal Processing-Estimation Theory*", Englewood Cliffs, Upper Saddle River, NJ: Prentice Hall, 1993

DOI 10.5162/sensoren2019/5.3.3

Parametric Spectrum Analysis of Backscattered Ultrasound Signals for the Characterization of Particles in Suspensions

Michael Vogt, Michael Deilmann
KROHNE Messtechnik GmbH, Ludwig-Krohne-Str. 5, 47058 Duisburg, Germany

Summary

Suspensions, i.e. heterogeneous mixtures of solid particles in liquids, play an important role in various industries and applications. Especially in-line methods for analyzing the concentration and the size of the solid particles in a given suspensions are of large interest in chemistry and process industry, for example for monitoring and control of sedimentation processes in wastewater treatment plants. Further examples include the analysis of drilling mud, paints, and others. Ultrasound-based methods are a powerful tool for this task, because the ultrasound waves are scattered at the solid particles inside the liquid phase and, different from optical techniques, also allow for measurements with opaque media. In the case of a particles size much smaller than the ultrasound wavelength, as is given in many practical applications, Rayleigh backscattering occurs. Accordingly, the intensity of the scattered ultrasound waves is proportional to the fourth power of the frequency, i.e. the scattering phenomenon is largely frequency-dependent. This can advantageously be utilized for characterizing suspensions by a spectral analysis of the backscattered waves (echo measurement) or of the transmitted waves (transmission measurement). The goal of the work presented here was to develop a robust method for characterizing the concentration and the size of the particles by employing a parametric approach. This is achieved by fitting an analytical model to the spectra of acquired echo signals and by deriving quantitative parameters thereof. The proposed concept has been tested and validated by means of experimental echo measurements with ultrasound in the 2 MHz range.

Keywords: Ultrasound analysis, Rayleigh scattering, parameter extraction, quantitative measurement, model fit.

Introduction

Ultrasound-based methods can very advantageously and quite easily be applied for the characterization of *fluids* (liquids, gases). Especially in the case of *homogeneous* media, various acoustic parameters can be quantitatively measured by means of transmission and reflection measurements with ultrasound [1]. Using the measured speed of sound, acoustic impedance, and attenuation coefficient, it is possible in many applications to estimate further parameters like the mass density or the concentration of a mixture of known liquids [2].

In also interesting other applications, however, *inhomogeneous* media are given. This includes various *aerosols* (liquid particles in a gas, 'wet gas'), and *suspensions* (solid particles in a liquid). The problem in this case is that the propagating ultrasound waves are largely scattered (or reflected) *inside* the medium. For this reason, an ultrasound wave shows large and random amplitude variations after it has been propagated through the scattering medium. On the other hand, echo measurements can be made in this case and allow for analyzing the scattering behavior of the medium. Many works have already been published in this field [3-8]. A fundamental challenge with regard to the characterization of inhomogeneous scattering media remains to derive *quantitative* parameters from the ultrasound measurement signals in a robust manner.

In this contribution, a parametric model approach is used to describe the spectral characteristics of backscattered ultrasound waves. The model is fitted to the magnitude spectra of acquired echo signals, and quantitative parameters are derived from the fitted model. Experiments have been performed in order to evaluate the potential of this concept for a further analysis of different suspensions. Results fg these measurements will be presented and discussed in detail.

Ultrasound Backscattering Measurement

In Fig. 1, a setup for ultrasound backscattering measurements of suspensions is illustrated.

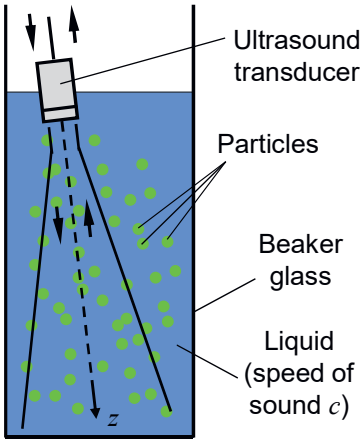


Fig. 1: Ultrasound backscattering measurement setup (illustration).

A single-element ultrasound transducer is directly immersed into the *suspension*, which is a heterogeneous mixture of *solid particles* in a *liquid*. In this *monostatic* configuration, the transducer is connected to a *pulser / receiver electronic* and used for both, the transmission and the reception of ultrasound. Accordingly, ultrasound waves, which are *backscattered* at the particles inside the suspension, are acquired over *time of flight* (TOF) by means of *pulse-echo measurements*.

The positions of the particles in the liquid are *random*, and the echo signal amplitude over TOF is a *random variable* for this reason. To allow for a *statistical analysis* under these conditions, the echo measurement is preferably repeated several times with different spatial arrangements of the particles. For the below described experimental measurements, this has been achieved by continuously *stirring* the suspension (using a magnetic stirrer).

Each given TOF t of the echo signals $x(t)$ can be linked to an *axial distance* $z = c/2 \cdot t$ by taking the *speed of sound* (SOS) c of the suspension into account, see Fig. 1. With a relatively small concentration of the particles, the SOS is approximately equal to that of the liquid (for example water: $c = 1480\,\mathrm{m/s}$).

The goal behind the below presented signal processing strategies is to realize a robust spectral analysis of the digitized echo signals. Special care is taken in this context to only include the echo signals, which are caused by the backscattering at the particles, in the analysis and to exclude other disturbing echoes. In practice, the latter can especially be given by multiple reflections *inside the transducer* ('transducer ringing'). Furthermore it might happen quite easily that also significant echoes from further objects appear. An example could be reflections from the floor or walls of the beaker glass in Fig. 1, which contains the suspension. Each acquired echo signal is *time-gated* over TOF prior to the further signal processing in order to analyze only the backscattering from inside the suspension.

A general drawback of using ultrasound is that the ultrasound waves are also strongly scattered at *gas bubbles*, which might be given in the liquid. Accordingly, the applicability of ultrasound-based analysis systems is generally limited to suspensions with no relevant content of gas bubbles.

Ultrasound Scattering Properties

The focus of this work is on analyzing *small* particles with a *mean particles diameter* d that is *much smaller* compared to the wavelength λ of the utilized ultrasound waves. With the SOS c of the liquid, this condition can be described as follows for a given ultrasound frequency f:

$$d \ll \lambda, \text{ with: } \lambda = c/f \tag{1}$$

As an example, a wavelength $\lambda \approx 740\,\mu\mathrm{m}$ is given in water for $f = 2\,\mathrm{MHz}$. The condition in (1) is fulfilled for many practically relevant suspensions, and *Rayleigh scattering* occurs at the particles in this case. As a consequence, the incident ultrasound wave is *diffusely scattered* into all directions [9].

The acoustic *backscattering coefficient* of the particles depends largely on the *mismatch* between the *acoustic impedances* (AI) of the particles and of the liquid. Furthermore, this parameter also depends on the mean particles diameter d, and it largely increases with *frequency* in the case of Rayleigh scattering. Accordingly, the *amplitude* (and so the *intensity*) of the echo signal as well as its *spectral characteristic* depend on the particles size and concentration. If echo signals from different suspensions but with the same *type* of particles, i.e. consisting of the same *material*, are analyzed, the aforementioned parameters allow for analyzing both, the size and the concentration of the particles.

The *scattering* at the *particles* is a first reason for the *attenuation* of the ultrasound waves, which propagate through the suspension. Generally, a second reason is *viscous losses* inside the *liquid*. In most cases, however, and especially with a large particles concentration and a large acoustic mismatch of the particles, the attenuation given by the scattering is much larger as compared to that given by the viscosity of the liquid.

Specifically for *Rayleigh scattering*, the *intensity* of the scattered ultrasound waves is proportional to the *fourth power* of the frequency. Under the assumption of a *uniform distribution* of the random particles positions, the echo signal amplitude is a random variable, which is described by the Rayleigh distribution [9].

Spectral Analysis of Backscattered Waves

In this section, echo signal processing concepts for a robust and reliable *spectral analysis* of the backscattered ultrasound waves from the suspension are presented and discussed.
In a first step, the *amplitude spectrum* of each digitized and time-gated echo signal is calculated by using the *Fast Fourier Transform* (FFT). Thereafter, the amplitude spectra obtained from multiple repeated measurements with changing spatial arrangement of the particles are averaged to account for the statistical variations of the echo signal amplitude. Next, a *parametric model* is fitted to the *mean amplitude spectrum*. In a final step, *quantitative parameters* are derived from the fitted model to characterize the backscattering properties of the suspension. In Fig. 2, a *Gaussian* model for describing the amplitude spectrum $|X(\omega)|$ over the (angular) frequency ω of an ultrasound echo signal $x(t)$ is shown:

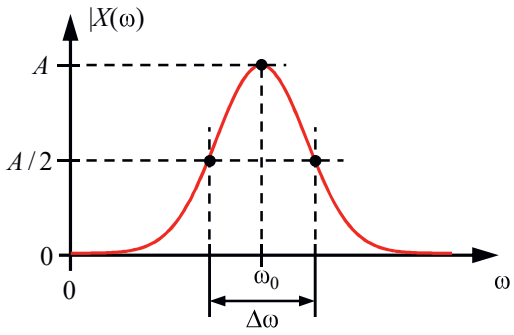


Fig. 2: Gaussian spectrum model; definition of quantitative parameters derived thereof: Amplitude A, center frequency ω_0, bandwidth (-6 dB) $\Delta\omega$.

As will be shown below, the Gaussian model usually fits quite well to amplitude spectra of 'real' ultrasound echo signals obtained in practical experiments. The Gaussian model in Fig. 2 can be analytically described as follows:

$$|X(\omega)| = A \cdot \mathrm{e}^{-k\cdot(\omega-\omega_0)}, \quad \Delta\omega = 2 \cdot \sqrt{\ln(2)/k} \quad (2)$$

The parameter A in (2) denotes the amplitude, ω_0 the (angular) center frequency, and $\Delta\omega$ the (angular) bandwidth (-6 dB) of the echo signal. (The constant k in (2) is only used for an easier writing of the equations.) Accordingly, the amplitude spectra of measured echo signals are completely described by these three parameters after the Gaussian parametric model has been fitted.

Fitting of Parametric Spectrum Model

The concept of the further processing steps is to fit the analytic spectrum model to the mean amplitude spectrum obtained from the echo measurements. The parameters of the fitted model (amplitude, center frequency, and bandwidth) are then used for analyzing the suspension, as they are expected to be closely linked to the particles concentration and size, which are the interesting physical parameters of the suspension.
A further important advantage of using the *Gaussian* model in (2) is that an *analytical* and *unambiguous* mathematical solution exists to calculate the *least square fit* to any given spectral measurement data [10]. This solution can be derived quite easily, and the computational effort for calculating the fit is very low.
As has already been discussed, the extracted model parameters are used below to *characterize* the suspension and, accordingly, to *differentiate* between different suspensions. Reference measurements on suspensions with *known* properties can be utilized in this context in order to *calibrate* the system.

Realized Measurement Setup

We have realized the laboratory measurement setup shown in Fig. 3.

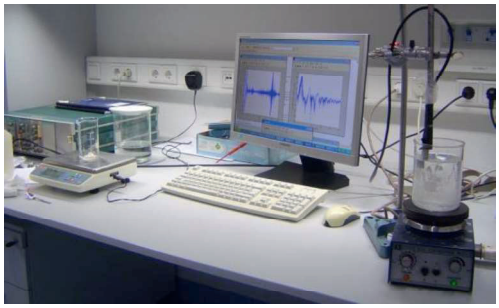


Fig. 3: Realized measurement setup (laboratory equipment).

A single-element *ultrasound piston transducer* with a nominal center frequency $f_0 = 2\,\mathrm{MHz}$ and $D = 10\,\mathrm{mm}$ aperture diameter has been used for all the below presented measurements. A *pulser / receiver* is used in this setup for performing pulse-echo measurements. The echo signals are directly digitized with an *analog-to-digital converter (ADC) board* with $14\,\mathrm{bit}$ amplitude resolution and $10\,\mathrm{MHz}$

sampling frequency (Spectrum Systementwicklung Microelectronik GmbH, Grosshansdorf, Germany). The digitized echo data is transferred into a *personal computer* (PC) for the further digital processing and analysis. The utilized transducer has a relatively large bandwidth (-6 dB) of 700 kHz, what allows for a good time-gating and for analyzing the echo signals over a relatively wide frequency range.

Measurement Samples

The above introduced concepts for processing the echo signals have been evaluated with results of various practical experiments. Measurements were performed using two different samples of suspensions, both containing the same type of *plastic particles* (i.e. consisting of the same material), but with different particles size:

- **Sample #1:** Finely granulated particles
- **Sample #2:** Coarsely granulated particles

Five different suspensions were prepared by using tap water and different particles concentrations (mass-volume fractions of 6, 12, 18, 24, and 30 g/l). A magnetic stirrer has been used to achieve a largely homogeneous spatial distribution of the particles in the suspensions. In Fig. 4, two exemplarily measurement situations with the two different samples #1 and #2, respectively, are shown with the same particles concentration of 12 g/l.

a) b)

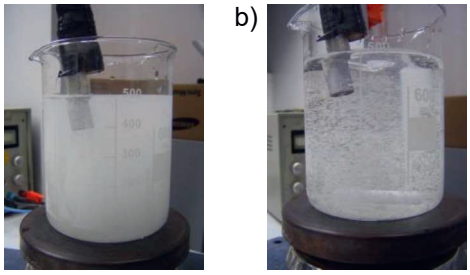
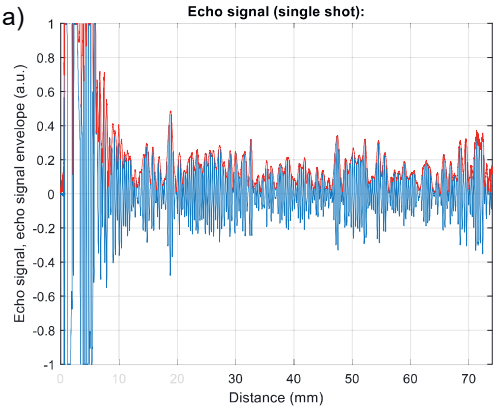


Fig. 4: Suspensions with particles of different granularity and size; particles concentration of 12 g/l: a) Finely granulated particles (sample #1), b) coarsely granulated particles (sample #2).

Examples of Recorded Echo Signals

In Fig. 5, one exemplarily recorded echo signal is shown (blue curves) for each of these two cases. The echo signals have been plotted over the axial distance z, which has been calculated by using a SOS $c = 1480\,\mathrm{m/s}$ for the water. At small distances (up to about 15 mm), echoes with a very large amplitude can be identified, which are caused by multiple reflections inside the transducer. Here, the ADC has actually been saturated on purpose with the transducer ringing in order to keep the quantization noise small and to achieve a large dynamic range for the interesting backscattering echoes. The latter can clearly be identified at larger distances and have much smaller amplitudes compared to the transducer ringing.

a)



b)

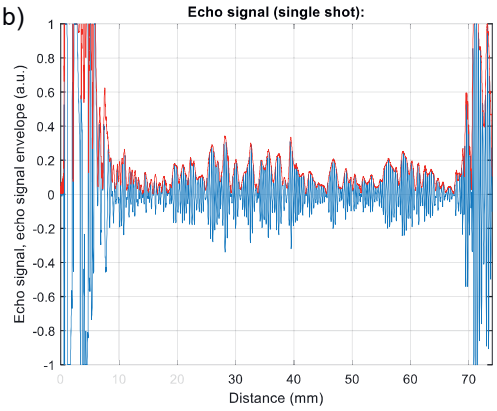


Fig. 5: Examples of acquired echo signals (blue curves) and echo signal envelopes (red curves); particles concentration of 12 g/l: a) Finely granulated particles (sample #1), b) coarsely granulated particles (sample #2).

At axial distances of about 70 mm and larger, echoes from the floor of the beaker glass, see Fig. 4, can be seen. The red curves in Fig. 5 show the echo signal *envelopes*, which have been calculated in each case as the magnitude of the *analytical* echo signal, obtained by using the *Hilbert transform* of the digitized echo signal. The comparison of Figs. 5 a) and b) shows that the *mean* echo signal amplitude is larger in the case of sample #1 (finely granulated particles) as compared to the sample #2 (coarsely granulated particles).
A series of 512 consecutive echo measurements has been performed with each of the investigated samples with different particles concentration. In Fig. 6, the envelopes of the

consecutively acquired echo signals are shown. The utilized gray-scale coding shows the logarithmic echo signal amplitude over a dynamic range of $50\,\mathrm{dB}$. Each column in the images in Fig. 6 represents one single echo signal envelope over the axial distance z.

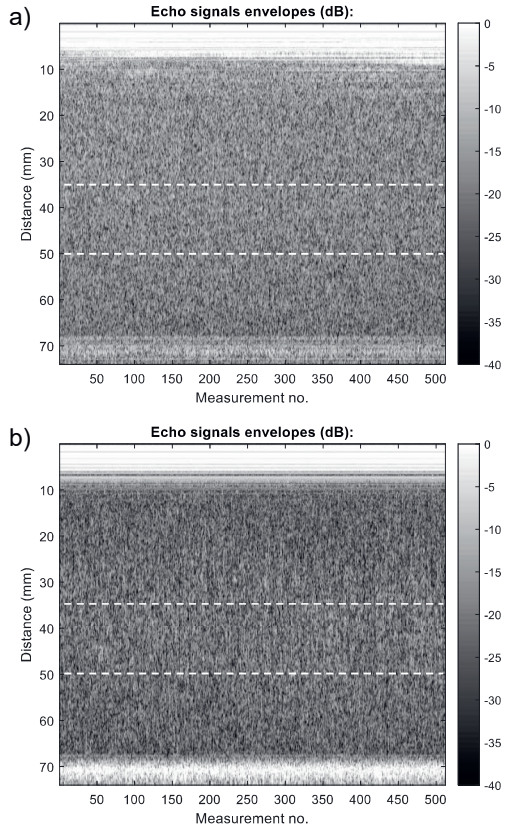


Fig. 6: Examples of acquired echo signals envelopes for consecutively repeated measurements; particles concentration of $12\,\mathrm{g/l}$: a) Finely granulated particles (sample #1), b) coarsely granulated particles (sample #2).

The transducer ringing can now be seen much better, appearing as *constant* echoes over the consecutive measurements (i.e. appearing as horizontal lines). The spatial arrangement of the scattering particles changes from one measurement to the next, because of the stirring. Accordingly, the consecutive echo signals are uncorrelated with each other. The random scatterer arrangement can be seen by the very large fluctuations of the echo signals envelopes, which are quite commonly referred to as '*speckle noise*' [9].

Series of Measurements

As described above, measurements have been performed with the two samples #1 and #2 and with varying particles concentration. As in the example above, echoes from 512 consecutive measurements have been digitized for each concentration. In the following, the acquired echo signals and the echo signal spectra, as well as results of the further spectral analysis will be presented and discussed.

Echo Signals Envelopes

The echo signal envelopes of the individual echoes have been calculated separately and averaged over the 512 measurements. The only purpose of averaging the echo signal envelopes was to show the according results in Fig. 7 and to allow for a discussion of these results.

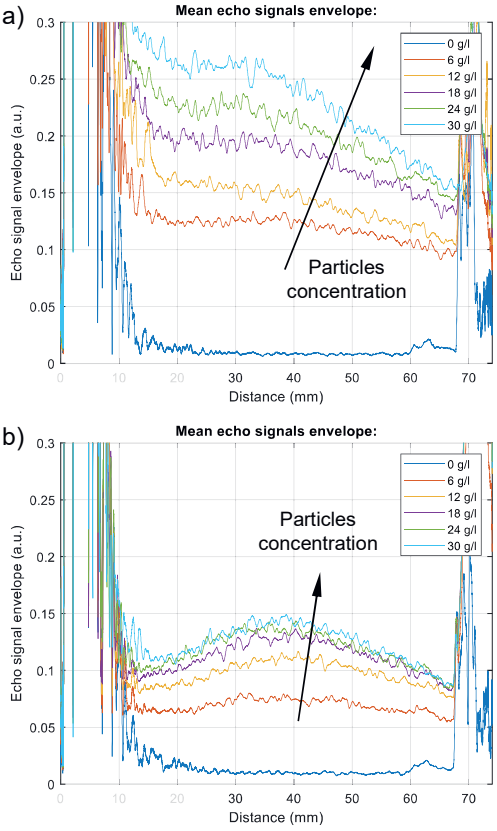


Fig. 7: Mean envelopes of acquired echo signals for varying particles concentration: a) Finely granulated particles (sample #1), b) coarsely granulated particles (sample #2).

The comparison of Fig. 7 with Fig. 5 shows that the echo signal fluctuations are very efficiently suppressed by the performed averaging of the uncorrelated 'observations' of the scattering phenomenon. It can be seen in Fig. 7 that the envelope signal increases with increasing particles concentration, what is in agreement with the expectations. For the finely granulated particles (sample #1), the echo signal amplitude is much larger as compared to the coarsely granulated particles (sample #2) in the case of comparing the same particles concentrations.

Echo Signals Spectra

Each digitized echo signal has been time-gated prior to the spectral analysis in order to include only the backscattering from the suspension. A *Hanning window* function has been applied over the axial distance range from $35\,\mathrm{mm}$ to $50\,\mathrm{mm}$ for this purpose, see the dashed lines in Fig. 6.
In a first step, the amplitude spectrum of each individual time-gated echo signal has been calculated using the FFT. Thereafter, the resulting 512 echo signal amplitude spectra have been averaged. This has been done here again in order to suppress the random echo signal fluctuations prior to fitting the analytical Gaussian spectrum model in each case. The resulting mean amplitude spectra are shown in Fig. 8 (marked by the stars in the figures) for a frequency interval from $1\,\mathrm{MHz}$ to $3\,\mathrm{MHz}$.

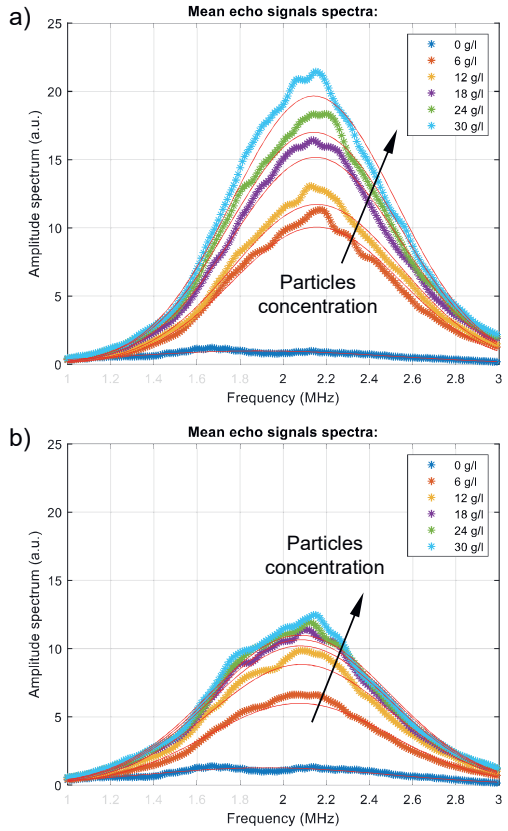


Fig. 8: Mean amplitude spectra of acquired echo signals for varying particles concentration: a) Finely granulated particles (sample #1), b) coarsely granulated particles (sample #2).

The spectra are predominantly given by the passband of the utilized ultrasound transducer, and this is in good agreement with the nominal center frequency $f_0 = 2\,\mathrm{MHz}$ of the transducer. The plotted spectra show a $(-6\,\mathrm{dB})$ bandwidth $\Delta f \approx 0.7\,\mathrm{MHz}$ of the transducer.

In another step, the parametric Gaussian spectrum model according to (2) has been fitted to each of the individual mean amplitude spectra. Only the measured amplitude spectra in the frequency interval from $1\,\mathrm{MHz}$ to $3\,\mathrm{MHz}$ has been used for the least square fit in order to exclude spectral components with a low signal-to-noise ratio (SNR), see Fig. 8.
The fitted spectra have also been plotted in Fig. 8 (solid lines). It can be seen that the Gaussian model fits actually quite well to the amplitude spectra of the experimentally acquired ultrasound echo signals. The amplitudes A of the fitted spectra, see Fig. 2, show a similar result like the mean echo signal envelopes of the backscattered ultrasound signals in Fig. 7.

Quantitative Spectral Parameters

The three parameters describing the Gaussian spectrum model (amplitude, center frequency, and bandwidth) have been extracted from the fitted spectra in Fig. 8 in each case. Based on these quantitative data, the potential of these parameters for a differentiation between the two samples #1 and #2 of suspensions at varying particles concentration has been evaluated and tested.
In Fig. 9, the two estimated parameters, center frequency and bandwidth, are shown over the given particles concentration for the two samples:

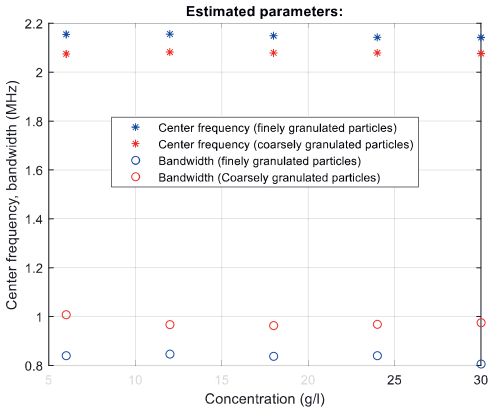


Fig. 9: Estimated center frequency and bandwidth for varying particles concentration: Finely granulated particles (sample #1), coarsely granulated particles (sample #2).

Both parameters are largely constant over the concentration, but they are significantly different for the two samples. This is in a good agreement with the expectation that the spectral backscattering properties change with the mean particles size.
In Fig. 10, the same two parameters (center frequency and bandwidth) are now plotted

against each other. Consequently, this plot represents a two-dimensional *parameter-space* [2]. Different symbols (stars and circles) have been used to indicate to which of the two samples each data point at varying concentration belongs.

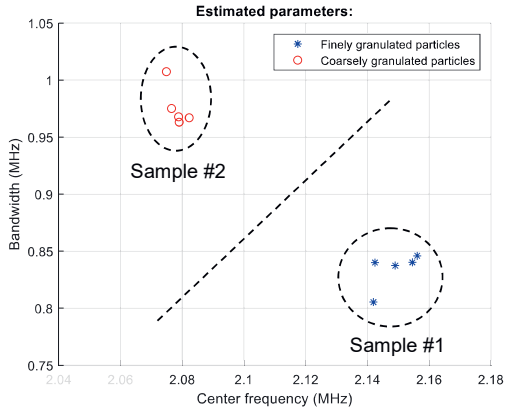


Fig. 10: Two-dimensional parameter space (bandwidth over center frequency): Clustering of data points for the two samples #1 and #2, respectively.

It can be seen much better now that the two parameters (center frequency and bandwidth) are largely different for the two samples, and that they are largely independent of the particles concentration. The mismatch in the acoustic impedances (of the particles and the liquid) and the particles concentration, on the other hand, are expected to have a large influence only on the echo signal amplitude.
The data points in Fig. 10 for the varying concentration are building separate and non-overlapping *clusters* for the two different samples. Accordingly, a *differentiation* between these two samples can be made, i.e. the two parameters allow for a *classification* of the actually given type of sample.
Finally, the *intensity* $I = A^2$ of the backscattered ultrasound waves, has been calculated from the estimated amplitude A (derived from the Gaussian spectrum model) in each single case, see Fig. 11. The plots show that the particles concentration can be derived from the measured intensity by using these results as *reference data* in terms of a *calibration* of the system. In order to estimate the concentration of a given unknown sample, the intensity has to be measured and the functions in Fig. 11 have to be inverted. However, the problem is that the actually given type of suspension (finely or coarsely granulated particles) has to be identified prior to the inverse calculation in order to know which of the two reference curves has to be used. But the according classification of the given suspension can be done as discussed above with Fig. 10.

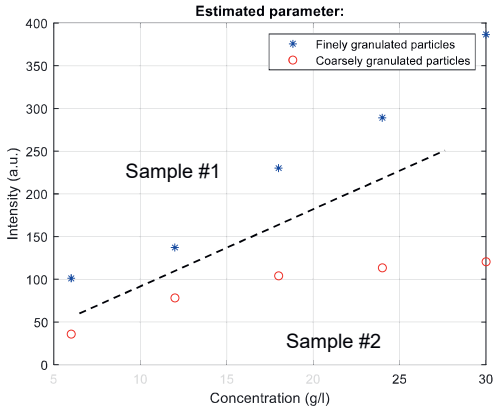


Fig. 11: Estimated intensity of backscattered ultrasound waves for varying particles concentration: Finely granulated particles (sample #1), coarsely granulated particles (sample #2).

Consequently, in a more general approach, a given *unknown* suspension can be characterized with regard to the *particles size* based on the estimated *center frequency and bandwidth*, while the *particles concentration* can be estimated based on the measured echo signal *intensity*. In any case, however, reference data obtained from calibration measurements with well-defined and *known* samples would have to be used ('supervised learning') to eliminate the *system-specific* influences onto the ultrasound echo measurement. Among these effects, the actually given transfer characteristics of the utilized ultrasound transducer is the most important one, which has to be accordingly considered.

Summary and Conclusions

In this contribution, concepts for analyzing suspensions by means of ultrasound echo measurements have been presented. General aspects of ultrasound echo measurement setups and the realization of a system for pulse-echo measurements with a single-element ultrasound transducer have been discussed. The fitting of a parametric Gaussian spectrum model to the measured mean echo signal amplitude spectrum has been proposed as a robust approach. Experiments have been performed with two different kinds of suspensions, consisting of the same type of particles, but with different particles size. It has been shown that the quantitative parameters amplitude, center frequency, and bandwidth, which are derived from the fitted model, have a large potential for assessing the particles size as well as the particles concentration of a given suspension. The proposed and

investigated method might potentially be used for inline analysis of suspensions in various industrial applications.

References

[1] M. Vogt, H. Ermert: Quantitative Analysis of Liquids and Emulsions by means of High-Frequency Ultrasound (15-35 MHz). IEEE Int. Ultrasonics Symp., 68-71 (2011); doi: 10.1109/ULTSYM.2011.0018

[2] M. Vogt, S. Langener, J. Kunze, T. Musch, H. Ermert: Measurement and Multivariate Analysis of Ultrasound Parameters for Quantitative Characterization of Liquids. IEEE Int. Ultrasonics Symp., 1137-1140 (2012); doi: 10.1109/ULTSYM.2012.0283

[3] S. Wöckel, U. Hempel, R. Weser, B. Wessely, J. Auge: Ultraschallstreuverfahren zur Charakterisierung von Dispersionen mit hohem Partikelanteil. 16. GMA/ITG-Fachtagung Sensoren und Messsysteme, 344-351 (2012); doi: 10.5162/sensoren2012/3.3.2

[4] R. Weser, B. Wessely, S. Wöckel, U. Hempel: Partikelmessung an hochkonzentrierten Dispersionen nach dem Ultraschall-Reflexionsprinzip. 16. GMA/ITG-Fachtagung Sensoren und Messsysteme, 344-351 (2012); doi: 10.5162/10dss2011/6.2

[5] S. Wöckel, U. Steinmann, R. Weser, B. Wessely: Inline-Prozessüberwachung von Suspensionen mit hohem Feststoffanteil. 11. Dresdner Sensor Symposium, 463-466 (2013); doi: 10.5162/11dss2013/L17

[6] S. Wöckel, R. Weser, U. Hempel, B. Wessely, J. Auge: Monitoring of Highly Concentrated Suspensions by Ultrasonic Scattering. IEEE Int. Ultrasonics Symp., 1133-1136 (2012); doi: 10.1109/ULTSYM.2012.0282

[7] D.J. McClements: Ultrasonic Measurements in Particle Size Analysis. In: Encyclopedia of Analytical Chemistry, (Sept. 2006); doi: 10.1002/9780470027318.a1518

[8] J. Carlson, A. Greenberg: Ultrasonic Measurements of Particle Concentration in a Multiphase Flow. IEEE Int. Ultrasonics Symp., 757-760 (1999); doi: 10.1109/ULTSYM.1999.849510

[9] R.F. Wagner, S.W. Smith, J.M. Sandrik, H. Lopez: Statistics of Speckle in Ultrasound B-Scans. IEEE Trans. Sonics Ultrasn., vol. 30, no. 3, 155-163 (May 1983) ; doi: 10.1109/T-SU.1983.31404

[10] M. Vogt, H. Ermert: In Vivo Ultrasound Biomicroscopy of Skin: Spectral System Characteristics and Inverse Filtering Optimization. IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol. 54, no. 8, 1551-1559 (August 2007) ; doi: 10.1109/TUFFC.2007.425

DOI 10.5162/sensoren2019/5.4.1

Synchronisationsprobleme innerhalb eines Sensorsystems und deren Auswirkungen auf Ergebnisse des maschinellen Lernens

Tanja Dorst[1], Tizian Schneider[2], Steffen Klein[3], Sascha Eichstädt[1], Andreas Schütze[2,3]
[1]Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin, Deutschland
[2]ZeMA – Zentrum für Mechatronik und Automatisierungstechnik gGmbH, Eschberger Weg 46, 66121 Saarbrücken, Deutschland
[3]Universität des Saarlandes, Lehrstuhl für Messtechnik, Campus A5 1, 66123 Saarbrücken, Deutschland

Zusammenfassung

Prozesssensordaten erlauben nicht nur die Steuerung industrieller Prozesse, sondern auch eine Bewertung des Anlagenzustands zur Erkennung von Fehlerzuständen und Verschleiß mittels Mustererkennung durch maschinelles Lernen. Um den Einfluss von Synchronisationsfehlern innerhalb eines verteilten Sensorsystems abzuschätzen, wird hier ein System zur Zustandsüberwachung, zur Lebensdauerprognose und für End-Of-Line-Tests elektromechanischer Zylinder betrachtet. Die im Prüfstand integrierten Sensoren werden mit 10 kHz bis 1 MHz abgetastet. Betrachtet werden im Folgenden die Auswirkungen von Synchronisationsfehlern von maximal 100 ms zwischen den einzelnen Sensoren auf die Lebensdauervorhersage der elektromechanischen Zylinder. Simuliert werden diese Fehler an Hand eines experimentell erfassten Rohdatensatzes durch Manipulation mit zufälligem Zeitversatz zwischen den unterschiedlichen Sensoren.

Keywords: Maschinelles Lernen, Zustandsüberwachung, multivariate Statistik, Synchronisation, Industrie 4.0

Einleitung

Globalisierte Fertigung und internationaler Handel haben durch die Digitalisierung einen technologischen Strukturwandel erfahren. Neben intelligenten Sensoren, welche nicht nur ihre Umwelt wahrnehmen, sondern auch die Messwerte gleich weiterverarbeiten und kommunizieren, kommen große Sensornetzwerke in der Fabrik der Zukunft zum Einsatz. Maschinelles Lernen (ML) ist dabei eine Schlüsseltechnologie für diese intelligenten Systeme. In der Fabrik der Zukunft stellt das „Industrial Internet of Things“ (IIoT) die Vernetzungsbasis für die Industrie 4.0 dar und dient somit der Verbesserung der betrieblichen Effektivität und Flexibilität [1]. Die stetige Weiterentwicklung intelligenter Sensorsysteme erlaubt die Automatisierung und Verbesserung komplexer Prozess- und Geschäftsentscheidungen in vielfältigen Anwendungsgebieten. Beispielsweise können Sensoren zur Bewertung des Anlagenzustandes, zur Bestimmung optimaler Wartungstermine oder zum Erkennen von Fehlerzuständen ebenso wie zur Steuerung von ganzen Produktionslinien genutzt werden. Ein grundlegendes Problem bereitet dabei die Sensordatenqualität, welche u.a. durch Umwelteinflüsse, Sensorausfälle und – besonders in verteilten Sensornetzwerken – durch Synchronisationsprobleme zwischen den einzelnen Sensoren beschränkt ist. Auch durch Messfehler wird die Qualität der Daten herabgesetzt. Dabei erfordert das Vertrauen in ML-Algorithmen und deren Entscheidungen verlässliche Daten. Die Bewertung der Datenqualität ist daher essentiell, d.h. es wird eine metrologische Infrastruktur benötigt, welche Aussagen über die Qualität der Daten zulässt. Dieser Beitrag zeigt, wie mit Hilfe einer Software-Toolbox für statistische Analysen große Datenmengen unter dem Einfluss von künstlich erzeugten Synchronisationsfehlern ausgewertet werden und wie groß der Einfluss dieser Synchronisationsfehler auf die Ergebnisse des maschinellen Lernens sind.

Der Prüfstand zur Datenerfassung

Am verwendeten Prüfstand können Zustandsüberwachungen, Lebensdauervorhersagen und End-Of-Line (EOL) Tests an elektromechanischen Zylindern (EMZ) mit Kugelumlaufspindeln durchgeführt werden. Diese Dauerhochlast- und Geschwindigkeitstests werden am Prüfstand solange durchgeführt, bis ein Fehler des EMZ auftritt. So können

relevante Sensoren und charakteristische Signalmuster für die Zustandsüberwachung identifiziert werden und eine Restlebensdauerschätzung des EMZs durchgeführt werden. Das Schema des Prüfstandes ist in Abb. *1* zu sehen. Der Aufbau besteht vereinfacht aus dem geprüften EMZ und einem Pneumatikzylinder, welcher die variable Last auf den EMZ in axialer Richtung simuliert. Die Parameter des Arbeitszyklus können mittels eines LabVIEW-GUI eingestellt werden.

Ein typischer Arbeitszyklus dauert 2,8 s. Er besteht aus einer Vorhub- und einer Rückhubbewegung des Zylinders, sowie einer Wartezeit zwischen beiden Bewegungen von 150 ms. Dabei werden die Bewegungen immer mit maximaler Beschleunigung von 5 m/s² und annähernd maximaler Geschwindigkeit ausgeführt. Der Hubbereich des EMZs im Prüfstand liegt zwischen 100 mm und 350 mm. Die Kombination aus hoher Axialkraft (7 kN), hoher Verfahrgeschwindigkeit (200 mm/s²) und hoher Beschleunigung (5 m/s²) führt zu schnellem Verschleiß des EMZ. Als Fehlerkriterium für den Ausfall des EMZ wird eine zu große Abweichung zwischen Positionssoll- und Positionsistwert festgelegt, d.h. der Test ist beendet, sobald die spezifische Positionsgenauigkeit aufgrund erhöhter Reibung nicht mehr erfüllt ist. Um ausreichend Daten für eine umfassende Zustandsüberwachung zu sammeln, werden die folgenden Sensoren verwendet [2]:

- drei Motorstromsensoren mit einer Abtastrate von jeweils 1 MHz,
- drei Beschleunigungssensoren mit einer Abtastrate von jeweils 100 kHz, angebracht am Gleitlager, am Kugellager und an der Kolbenstange
- ein Mikrofon mit einer Abtastrate von 100 kHz
- vier Prozesssensoren (Axialkraft, Geschwindigkeit, pneumatischer Druck und Wirkstrom) mit einer Abtastrate von jeweils 10 kHz

Die damit erfassten Daten spiegeln die Funktionalität des EMZs und deren Abnahme während des Langzeittests wider. Für die Datenanalyse, welche im nächsten Abschnitt genauer beschrieben wird, wurden verschiedene EMZs bis zu deren Ausfall geprüft. Jeder EMZ erzeugte dabei durchschnittlich 12 TB Rohdaten.

Die Software-Toolbox zur Datenanalyse

Für die Datenanalyse wird die automatisierte Software-Toolbox verwendet, welche in [3] vorgestellt wurde. Diese kann voll automatisiert, d.h. ohne Expertenwissen und ohne detailliertes physikalisches Modell des Prozesses, verwendet werden. Nach Erfassen der Rohdaten erfolgen, wie in Abb. *2* zu sehen, die Merkmalsextraktion und -selektion sowie die Klassifikation und Evaluation. Die Merkmalsextraktion findet dabei unüberwacht statt, d.h. sie wird ohne Kenntnis der Gruppenzugehörigkeit, d.h. des aktuellen Alterungszustands, der einzelnen Arbeitszyklen durchgeführt. Aus den zyklischen Rohdaten, welche über einen konstanten Zeitraum periodisch erfasst wurden, werden formbeschreibende Merkmale generiert. Da es keine Methode gibt, die für alle Anwendungsfälle gut funktioniert, werden die Merkmale durch fünf verschiedene, sich gegenseitig ergänzende, Methoden aus verschiedenen Domänen extrahiert: Die Adaptive Lineare Approximation (ALA) unterteilt die Zyklen im Zeitbereich in annähernd lineare Abschnitte, aus denen als Merkmale Mittelwert und Steigung extrahiert werden. Bei der Hauptkomponentenanalyse (PCA) werden die Projektionen auf die Hauptkomponenten als

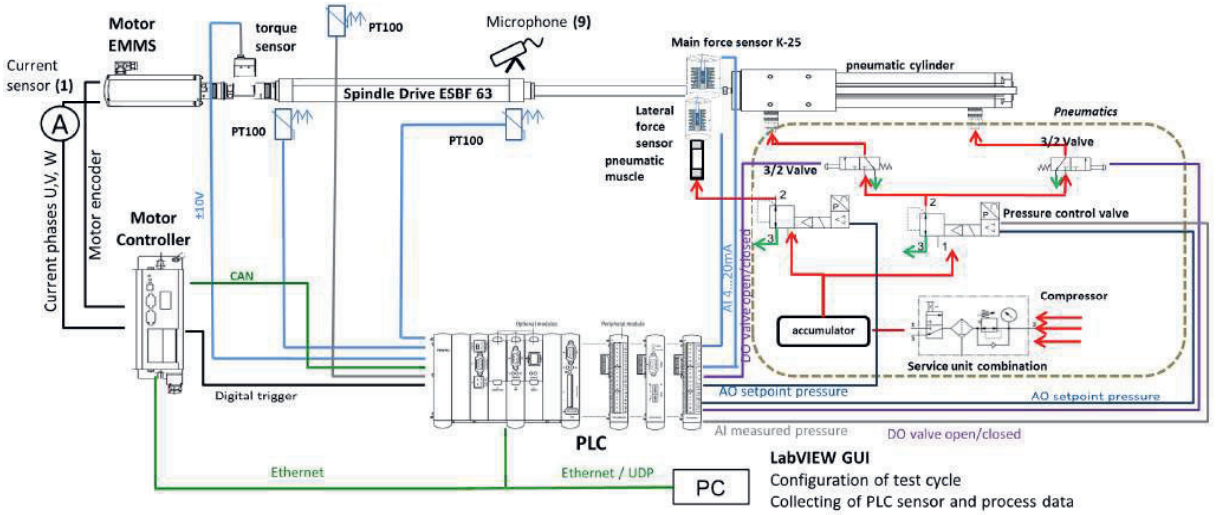


Abb. 1: Schema des Prüfstandes [4]

Merkmale verwendet. Mit Hilfe der Fouriertransformation werden bei der Methode der Besten Fourier-Koeffizienten (BFC) im Frequenzbereich die 10 % der Amplituden mit dem höchsten Betrag und die dazugehörigen Phasen als Merkmale verwendet. Die Beste Daubechies Wavelettransformation (BDW) erfolgt analog zur BFC und liefert dabei Merkmale aus dem Zeit-Frequenz-Bereich. Auch in der statistischen Verteilung der Messwerte ist Information enthalten. Diese wird mit Hilfe der ersten vier statistischen Momente (Mittelwert, Varianz, Schiefe und Wölbung) aus gleich langen Zyklusabschnitten extrahiert.

Ziel der Merkmalsextraktion ist es, Informationen in so wenigen Merkmalen wie möglich zu konzentrieren, allerdings ist die Anzahl der Merkmale nach der Merkmalsextraktion noch zu hoch. Für Big Data Anwendungen, wie z.B. die am vorher beschriebenen Prüfstand durchgeführte Restlebensdauervorhersage eines EMZs, ist die Datenreduktion nach diesem Schritt noch unzureichend. Aus diesem Grund wird im Anschluss an die Merkmalsextraktion eine Merkmalsselektion durchgeführt.

Die Merkmalsselektion ist ein überwachter Schritt, d.h. hier ist die Gruppenzugehörigkeit eines jeden Zyklus ist bekannt. Im konkreten Anwendungsfall bei den EMZ ist die Zielgröße die Lebensdauer mit einer Auflösung von 1 %. In diesem Schritt werden die relevantesten Merkmale bezogen auf die Klassifikationsaufgabe ausgewählt und redundante Merkmale oder Merkmale mit wenig Informationsgehalt entfernt. Auch hier liefert nicht eine einzige Methode für alle Anwendungsfälle die optimale Lösung, sodass drei verschiedene verwendet werden. Bei der Recursive Feature Elimination Support Vector Machine (RFESVM) wird eine lineare SVM verwendet, um rekursiv die Merkmale mit dem kleinsten Beitrag zur Gruppentrennung aus der Menge aller Merkmale zu entfernen [5,6]. Der RELIEFF-Algorithmus kommt zur Anwendung, wenn die Gruppen nicht linear trennbar sind [7,8]. Die Pearson-Korrelation wird aufgrund der geringen Komplexität des Algorithmus als dritte Methode zur Merkmals(vor)selektion verwendet. Dabei werden die Merkmale nach ihrem Korrelationskoeffizienten zur Zielgröße sortiert. Dieser Koeffizient gibt an, wie groß der lineare Zusammenhang zwischen einem Merkmal und der Zielgröße ist. Neben einer Reduktion der Datenmenge wird so gleichzeitig auch eine Überanpassung vermieden, die häufig auftritt, wenn die Anzahl der Datenpunkte zur Entwicklung des Klassifikationsmodells nicht wesentlich größer als die Anzahl der Merkmale ist.

Die Klassifikation läuft in zwei Schritten ab: eine weitere Dimensionsreduktion gefolgt von der eigentlichen Klassifikation. Die Dimensionsreduktion erfolgt mittels Linearer Diskriminanzanalyse (LDA). Sie führt für k Gruppen eine lineare Projektion des Merkmalsraums in einen $k-1$ dimensionalen Unterraum durch. Dabei wird die Intraklassenvarianz, die Varianz innerhalb der Klassen, minimiert und gleichzeitig die Interklassenvarianz, die Varianz zwischen den Klassen, maximiert [9]. Die eigentliche Klassifikation erfolgt dann unter Verwendung der Mahalanobis-Distanz, siehe Gl. (1):

$$d_{Mahal}(\vec{x}) = \sqrt{(\vec{x} - \overline{m})^T S^{-1} (\vec{x} - \overline{m})}. \quad (1)$$

Dabei bezeichnet $\vec{x}$ den Vektor der Testdaten, $\overline{m}$ das komponentenweise arithmetische Mittel und S die Kovarianzmatrix der Gruppe. Die Mahalanobis-Distanz gibt für jeden Datenpunkt an, wie weit er vom Zentrum der Datengruppe unter Berücksichtigung der Streuung entfernt

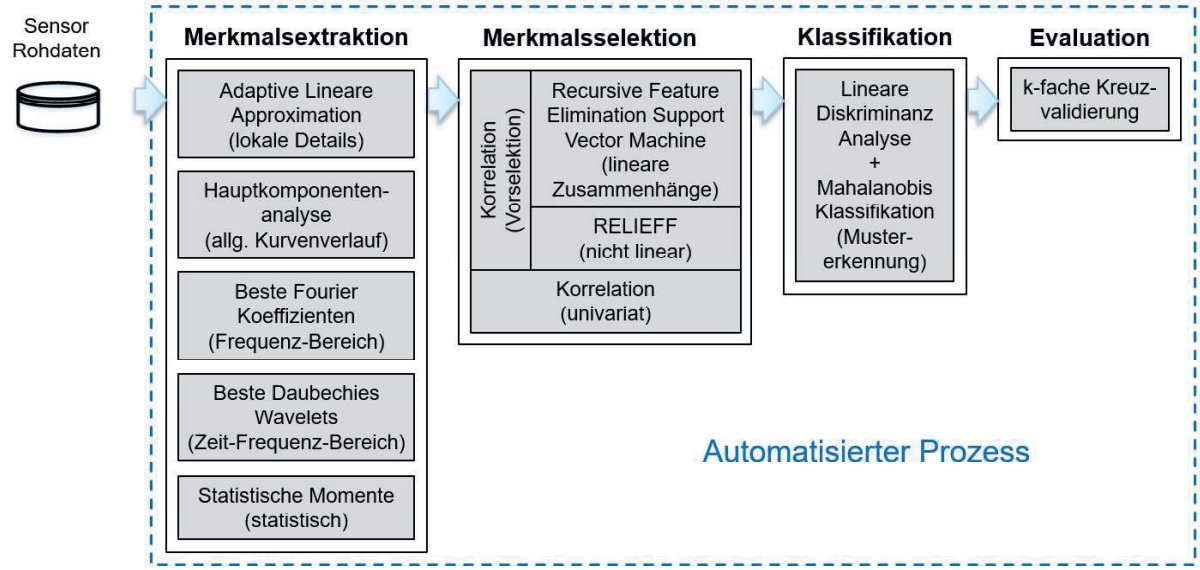


Abb. 2: Schematischer Aufbau des Methodenbaukastens

ist. Punkte gleicher Mahalanobis-Distanz von einem Zentrum bilden im zweidimensionalen Raum graphisch gesehen eine Ellipse.
Zur Evaluation wird die k-fache stratifizierte Kreuzvalidierung verwendet. Dabei wird der Datensatz in k Gruppen unterteilt, mit $k \in \mathbb{N}$. Stratifiziert bedeutet dabei, dass jede der k Teilmengen annähernd dieselbe Verteilung besitzt wie die Gesamtmenge. Gewöhnlich wird k gleich 10 gesetzt. So bildet eine Gruppe den Testdatensatz und $k-1$ Gruppen bilden den Trainingsdatensatz.
Die verwendete Software-Toolbox vergleicht die insgesamt 15 Möglichkeiten, welche durch Kombination je einer Merkmalsextraktions- und einer Selektionsmethode einstehen. Der Kreuzvalidierungsfehler, d.h. der Prozentsatz der anhand der 10-fachen Kreuzvalidierung falsch klassifizierten Zyklen, wird automatisch für jede der 15 Möglichkeiten und jede der 10 Permutationen, welche durch die Kreuzvalidierung entstehen, errechnet. Somit ist das Finden der besten Kombination aus den Merkmalsextraktions- und Merkmalsselektionsmethoden für das jeweils aktuelle Anwendungsszenario ein vollautomatischer Prozess, welcher offline durchgeführt wird. Die eigentliche Klassifikationsaufgabe wird dann online nur noch mit der besten der 15 Möglichkeiten durchgeführt, was zu einem geringeren Rechenaufwand führt.
Zusammen führen Merkmalsextraktion und Merkmalsselektion ungefähr zu einer 60-fachen Datenreduktion, d.h. die anfangs vorhandenen 12 TB an Rohdaten werden auf einen Merkmalsdatensatz von ca. 200 MB reduziert.

Synchronisationsprobleme und deren Auswirkungen

Zur Analyse der Auswirkungen von Synchronisationsproblemen zwischen den am Prüfstand verbauten Sensoren und deren Auswirkungen auf die Lebensdauerprognose am Beispiel elektromechanischer Zylinder wird ein Langzeittest mit einer Dauer von 20,4 Tagen und über 600.000 Zyklen herangezogen. Für die im vorherigen Abschnitt vorgestellte Software-Toolbox zum maschinellen Lernen wird je Arbeitszyklus nur eine Sekunde während der Gleichlaufphase des Rückhubes verwendet. In dieser Zeitspanne ist, wie in Abb. 3 zu sehen, die Geschwindigkeit konstant und die Belastung am größten, da dort gegen eine konstante Last gezogen wird. Simuliert wurden die Synchronisationsfehler durch Manipulation des Rohdatensatzes mit zufälligem Zeitversatz zwischen den unterschiedlichen Sensoren. Der zufällige Zeitversatz beträgt dabei maximal 100 ms, um sicherzustellen, dass für alle Sensoren nur Daten aus der Gleichlaufphase zur Auswertung herangezogen werden. Als Selektor wird bei allen Auswertungen die Pearson-Korrelation verwendet.

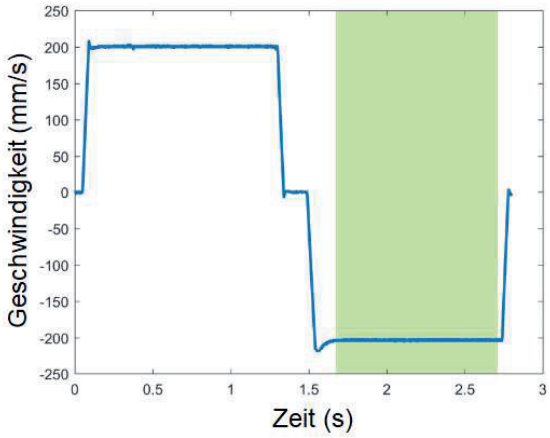


Abb. 3: Arbeitszyklus bestehend aus Vorhub, Wartezeit und Rückhub, sowie der Ausschnitt (grün), welcher für die Auswertung der Lebensdauer verwendet wird

Bei den Daten ohne Zeitverschiebung werden die besten Merkmale, wie in Abb. 4 zu sehen, mit BFC aus dem Frequenzbereich extrahiert. Der geringste Kreuzvalidierungsfehler mit 8,72 % Fehlklassifikation tritt bei der Verwendung von 493 Merkmalen auf, mit 100 Merkmalen liegt der Kreuzvalidierungsfehler bei 11,33 %.

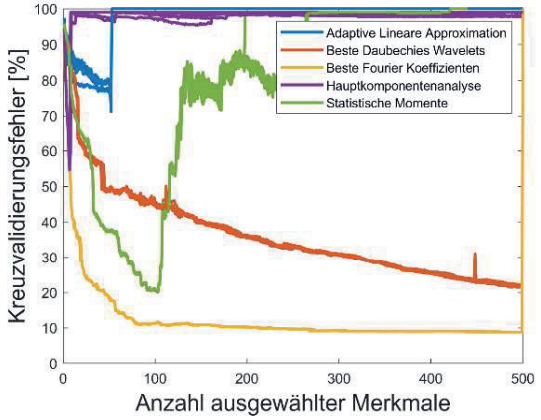


Abb. 4: Kreuzvalidierungsfehler über die Anzahl der Merkmale beim Rohdatensatz mit Verwendung der Pearson-Korrelation als Selektor

Die Klassifikationsfehler treten dabei häufig an Gruppengrenzen auf. Durch eine Mittelung über 15 Minuten (321 Zyklen) mittels Mehrheitsfilter werden diese Klassifikationsfehler entfernt, sodass danach nur noch sehr wenige Fehler am Ende der Laufzeit verbleiben, wie in Abb. 5 gezeigt. Zusätzlich wird in Abb. 6 für BFC als Merkmalsextraktor gezeigt, welche Sensoren bei der Lebensdauerprognose am wichtigsten sind. Dabei ist deutlich zu erkennen, dass 58 %

der Merkmale aus den Daten des Mikrofons stammen, wovon wiederrum 99 % Amplituden sind. Diese Informationen werden bei der Erzeugung einiger künstlicher Zeitversätze berücksichtigt.

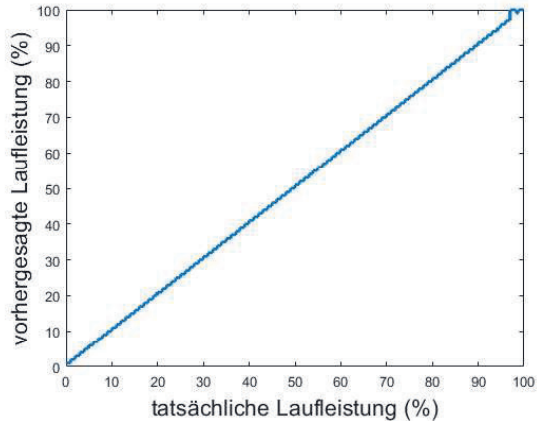


Abb. 5: Vorhergesagte über tatsächliche Laufleistung bei einer Mittelung über 15 Minuten

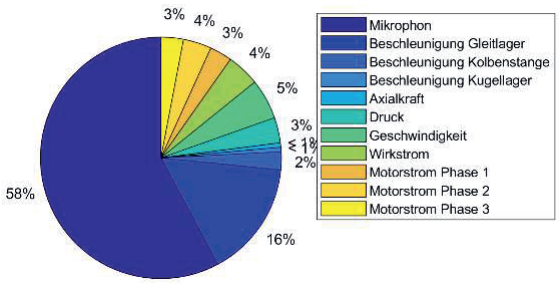


Abb. 6: Prozentsatz der besten Merkmale je Sensor, ausgewählt durch Pearson-Korrelation

Zunächst werden Zeitversätze von maximal 0,1 ms je Sensor bis zu maximal 100 ms je Sensor betrachtet. Dabei ist ein Anstieg des minimalen Klassifikationsfehlers bei steigendem Zeitversatz zu erkennen (siehe Tab. 1). Da die Ergebnisse fast äquivalent zu denen mit den Rohdaten sind, werden diese hier nicht explizit gezeigt. BFC ist immer die beste Extraktionsmethode, die Änderung des Klassifikationsfehlers liegt im einstelligen Bereich und über 50 % der selektierten Merkmale stammen aus den Daten des Mikrofons.

Daher wird im Folgenden nur ein Zeitversatz von maximal ±100 ms betrachtet, um die die Daten jedes einzelnen Sensors zufällig verschoben werden. Dieser Zeitversatz ist zum einen der, welcher bei der Zykluslänge von 2,8 s und Betrachten des Rückhubes maximal möglich ist, zum anderen liefert dieser Versatz die schlechteste Klassifikationsrate. Wie in Abb. 7 zu sehen, werden die besten Merkmale wie beim Rohdatensatz mit Hilfe der BFC generiert. Allerdings liegt der minimale Klassifikationsfehler nun bei 19,3 % und damit deutlich schlechter als bei Verwendung der Rohdaten.

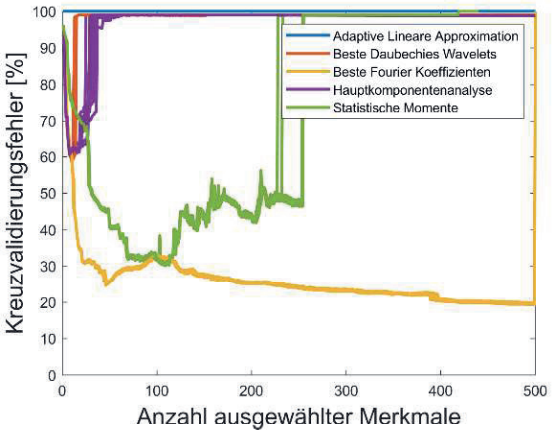


Abb. 7: Kreuzvalidierungsfehler über die Anzahl der Merkmale bei maximal ±100 ms Zeitversatz und Pearson-Korrelation als Selektor

Tab. 1: Klassifikationsfehler der verschiedenen Zeitversätze mit BFC als Merkmalsextrator

Zeitversatz	*Sensoren mit Zeitversatz*	*min. Fehler*	*selektierte Merkmale (davon Amplituden)*	*Merkmale aus Mikrofon-daten*	Fehler bei 100 selektierten Merkmalen
ohne	-	8,72 %	493 (90 %)	58 %	11,33 %
max. ±0,1 ms	alle	9,11 %	497 (93 %)	60 %	11.05 %
max. ±0,5 ms	alle	9,43 %	498 (96 %)	63 %	11,84 %
max. ±1 ms	alle	9,73 %	491 (97 %)	64 %	12,80 %
max. ±10 ms	alle	10,10 %	496 (100 %)	65 %	13,61 %
max. ±50 ms	alle	13,17 %	497 (100 %)	65 %	20,87%
max. ±100 ms	alle	19,30 %	499 (100 %)	66 %	32,55%
konst. +1 ms	nur Mikrofon	8,88 %	498 (90 %)	57 %	11,66 %
konst. +100 ms	nur Mikrofon	8,59 %	499 (90 %)	58 %	11,14 %
max. +100 ms	nur Mikrofon	8,78 %	495 (90 %)	57 %	11,74 %

Die mögliche Ursache für die deutliche Verschlechterung der Klassifikationsrate liegt vermutlich darin, dass nun nicht mehr nur Daten aus der Gleichlaufphase des Rückhubes verwendet werden, sondern auch Daten aus der Beschleunigungs- bzw. Verzögerungsphase des Rückhubes mit in die Auswertung einfließen. Mittels Mehrheitsfilter werden einige Klassifikationsfehler entfernt, allerdings verbleiben während der letzten 20 % der Lebensdauer des EMZs deutliche Klassifikationsfehler, wie in Abb. 9 dargelegt.

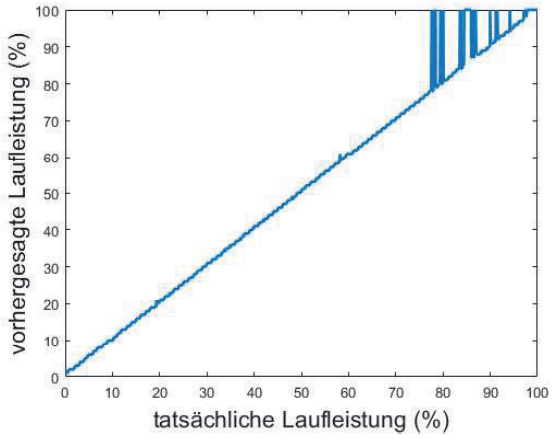


Abb. 8: Vorhergesagte über tatsächliche Laufleistung bei einer Mittelung über 15 Minuten für den manipulierten Datensatz mit max. ±100 ms Zeitversatz zwischen den Sensoren

Abschließend werden die Klassifikationsfehler für die Zeitversätze von 0,1 ms bis 100 ms zwischen allen 11 Sensoren betrachtet. Diese sind in Abb. 9 dargestellt.

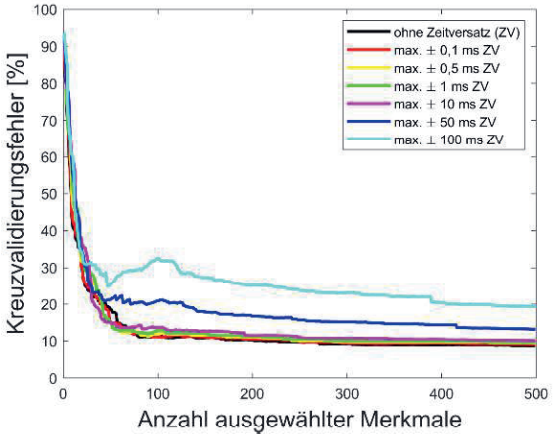


Abb. 9: Kreuzvalidierungsfehler über die Anzahl der Merkmale für unterschiedliche simulierte Synchronisationsfehler

Auch hier ist deutlich zu erkennen, dass ein minimaler Zeitversatz nur eine geringe Auswirkung auf das Ergebnis der Lebensdauerprognose eines EMZs hat, wohingegen größere Zeitversätze die Ergebnisse deutlich verschlechtern.

Fazit und Ausblick

Wie gezeigt, wirken sich minimale Synchronisationsfehler zwischen den einzelnen Sensoren nur sehr gering auf die Klassifikationsraten aus. Ein Zeitversatz nur für die Daten des Mikrofons hat ebenfalls nur sehr geringe Auswirkungen auf die Ergebnisse der Lebensdauervorhersage von EMZs, da aus den Daten des Mikrofons primär Amplituden verwendet werden. Die durchgeführte Analyse gibt somit wichtige Hinweise für die Auslegung verteilter Messsysteme speziell im Hinblick auf die erforderliche Synchronisation zwischen den Sensoren.
Im weiteren Verlauf des Projektes „Metrology for the Factory of the Future“ (Met4FoF) werden mathematische Modelle zur Berücksichtigung von metrologischen Informationen entwickelt. Damit werden die Methoden der Software-Toolbox erweitert, um Messunsicherheiten bei der Vorhersage der Restlebensdauer von EMZ berücksichtigen zu können. Ein weiterer Schritt im Projekt betrifft die Klassifikation innerhalb der Software-Toolbox. Dabei wird die Robustheit der LDA als Klassifikator bei Redundanz der Merkmale überprüft. Eine Frage für die weitere Forschung ist dabei, wie lange sich die Qualität der Ergebnisse der LDA verbessert und ab wann sie sich plötzlich verschlechtert, da die Kovarianzmatrix singulär wird, d.h. ihre Determinante verschwindet.

Danksagungen

Ein Teil dieser Arbeit wurde durch das EU-weite Forschungsprogramm EMPIR finanziert, welches von der EU, den EMPIR-Mitgliedsländern und von Horizont 2020, dem Rahmenprogramm der Europäischen Union für Forschung und Innovation, kofinanziert wird. Der ZeMA-Prüfstand wurde im Rahmen des Forschungsprojekt MoSeS-Pro entwickelt, welches vom Bundesministerium für Bildung und Forschung im Aufruf Sensorbasierte Elektroniksysteme für Anwendungen für Industrie 4.0 (SElekt I4.0), Förderkennzeichen 16ES0419K, im Rahmen der deutschen Hightech-Strategie gefördert wurde.

Literaturnachweis

[1] S. Eichstädt, “Publishable Summary for 17IND12 Met4FoF Metrology for the Factory of the Future,” 2018.
[2] T. Schneider, N. Helwig, S. Klein, and A. Schütze, “Influence of Sensor Network Sampling Rate on Multivariate Statistical Condition Monitoring of Industrial Machines and

Processes," *Proceedings*, vol. 2, no. 13, p. 781, Dec. 2018; doi: 10.3390/proceedings2130781
[3] T. Schneider, N. Helwig, and A. Schütze, "Industrial condition monitoring with smart sensors using automated feature extraction and selection," *Meas. Sci. Technol.*, vol. 29, no. 9, 2018; doi: 10.1088/1361-6501/aad1d4
[4] N. Helwig, T. Schneider, and A. Schütze, "MoSeS-Pro: Modular sensor systems for real time process control and smart condition monitoring using XMR-technology," *Proc. 14th Symp. Magnetoresistive Sensors Magn. Syst.*, 2017.
[5] I. Guyon, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157--1182, 2003; doi: 10.1162/153244303322753616
[6] A. Rakotomamonjy, "Variable Selection Using SVM-based Criteria," *J. Mach. Learn. Res.*, vol. 3, pp. 1357–1370, 2003; doi: 10.1162/153244303322753706
[7] I. Kononenko and S. J. Hong, "Attribute selection for modelling," *Futur. Gener. Comput. Syst.*, vol. 13, no. 2–3, pp. 181–195, Nov. 1997; doi: 10.1016/S0167-739X(97)81974-7
[8] M. Robnik-Šikonja and I. Kononenko, "Theoretical and Empirical Analysis of ReliefF and RReliefF," *Mach. Learn.*, vol. 53, no. 1, pp. 23–69, 2003; doi: 10.1023/A:1025667309714
[9] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2. ed. New York: Wiley, 2001.

DOI 10.5162/sensoren2019/5.4.2

Feature-based analysis of reproducible bearing damages based on a neural network

Andreas Beering, Jakob Döring, Karl-Ludwig Krieger
Institute of Electrodynamics and *Microelectronics (ITEM.ae), University of Bremen, Germany*
Email: {beering, doering, krieger}@item.uni-bremen.de

Abstract

In the interest of detecting damage to tapered roller bearings at an early stage and avoiding further consequential damage in future, an investigation of reproducible damage to bearing inner races, outer races and rolling elements is carried out. The vibration signals generated by the contact of the damaged surfaces during bearing runtime are recorded via a piezoelectric vibration transducer. Different scenarios with regard to the rotational speed and the size of the damage are investigated. Based on a calculation of features, a feed-forward neural network is trained and used to classify the damage. To assess the quality of the neural network, the receiver operator characteristic (ROC) and the area under curve (AUC) are compared for the neural network as well as for other popular classification algorithms such as support vector machine (SVM), decision tree and k-Nearest-Neighbor (KNN). Based on the AUC values, this approach showed that the neural network used has the best performance for the classification of bearing damage with an AUC of 0.994 and an overall classification accuracy of 93.5 %.

Keywords: Bearing faults, feed-forward neural network, piezoelectric vibration transducer, classification performance, condition monitoring

I. Introduction

In the field of condition monitoring, a large number of machine components are already monitored in order to optimize processes or detect damage. However, many areas have not yet been sufficiently investigated, so that avoidable machine failures still occur. This includes undetected bearing or gear damage in agricultural machines, which can result in fatal consequential damage and, due to the repair and downtime of the machine, cause considerable costs. A solution to avoid such costs in future could be the detection of damage by an analysis of vibration data.

Previous approaches to the detection of faults on rotation machine elements have largely been based on envelope analysis of vibration signals. For this analysis, the theoretical damage frequencies of the bearing, which result from the geometry and the rotation speed, are calculated first [1]. Subsequently, the vibration signal generated by the damaged bearing is recorded and the envelope curve of the signal is calculated. From the frequency bands occurring in the amplitude spectrum and the theoretically defined damage frequencies, the damage to the bearing or the gear is concluded [2, 3]. Another approach to the investigation of bearing damage is based on empirical mode decomposition (EMD) to obtain a sum of intrinsic mode functions (IMF) followed by feature extraction and analysis [4]. A further approach to the detection of bearing damage uses an envelope analysis, followed by a classification using a decision tree [5].

This publication presents a study of reproducible artificial damage to tapered roller bearings. The damage is caused to the rolling elements, the inner races and the outer races of the bearings in defined sizes. The resulting vibration signals of the differently damaged bearings are measured on a laboratory test bench. For this purpose, the bearing is inserted into a bearing socket, to which a piezoelectric vibration transducer is attached by a screw connection, and driven at a defined speed via a three-phase motor and a frequency inverter. Feature extraction is performed for the acquired vibration signals. Not only common features from the classification of rotating machines [6], but other acoustic features such as the spectral crest factor or the mel-frequency cepstrum coefficients (MFCCs) are used. Once the features have been calculated, a feed-forward neural network is trained based on a training data set with labelled data. Afterwards, the feed-forward neural network is evaluated and compared with other classifiers using the ROC and AUC.

The remainder of this paper is organized as follows. In Section II the feature extraction is outlined and the theoretical fundamentals of the used feed-forward neural networks are described. Furthermore, the receiver operator characteristic and the area under curve are

described, which are regarded as measures for the performance of a classifier. Section III outlines the structure of the used tapered roller bearing with the reproducible induced damage. Afterwards, the monitoring setup is presented. In Section IV the experimental results are discussed and the used feed-forward neural network is compared to other classification algorithms. Finally, a conclusion and an outlook on future work is drawn in Section V.

II. Classification algorithm and performance

A. *Feature extraction*

Feature extraction describes the representation of characteristic pattern properties for the purpose of differentiating pattern classes. This also ensures that irrelevant information is not represented and the effective amount of data is reduced. The majority of machine learning classification algorithms uses features as input parameters to assign patterns to specific classes [7]. In this approach known features from the analysis of rotating machines such as kurtosis, skewness or variance are calculated [6]. In addition, features from the field of speech processing such as the mel-frequency cepstrum coefficients, the spectral crest factor or various features for the evaluation of the signal power are calculated. Overall, a total of 71 features are calculated from the vibration data of the bearing and used as inputs for the feed-forward neural network.

B. *Feed-forward neural network*

Neural networks can be described as a series of functional transformations, which combine input signals with weights and biases for the classification of several classes. Feed-forward neural networks represent a special case of neural networks. They allow the signals to propagate exclusively from input to output without feedback. The two-layer feed-forward network used in this paper is shown in Fig. 1 and consists of 71 inputs, 10 neurons in the hidden layer, 4 neurons in the output layer and 4 outputs. For the inputs of the neural network, a total of 71 features in the time and frequency domain have been calculated.

The basic neural network model to describe the first layer of the network is given as

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} \qquad (1)$$

where D describes the number of input signals x_i of the network, j the number of neurons in the first layer, $w_{ji}^{(1)}$ the weights and $w_{j0}^{(1)}$ the biases. The resulting values a_j are known as activations.

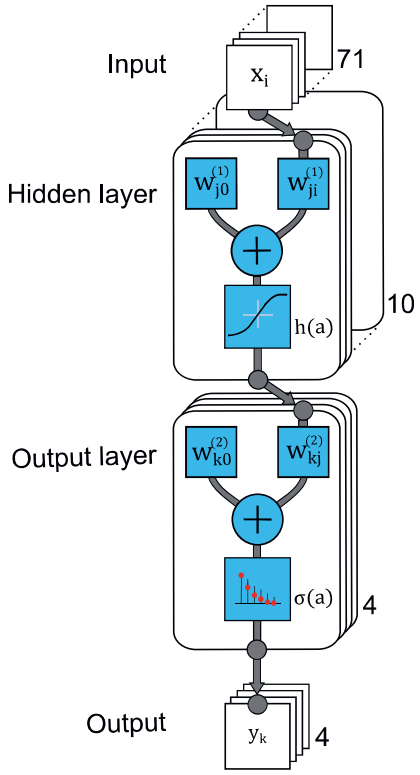


Fig. 1: Feed-forward neural network structure with 71 inputs, a hidden layer with 10 neurons and an output layer with 4 neurons for the classification of 4 classes.

Each activation is transformed using a nonlinear activation function h to form

$$z_j = h(a_j). \qquad (2)$$

In this approach a hyperbolic tangent sigmoid transfer function is used, which is represented by the following term

$$h(a) = \frac{2}{1 + e^{-2a}} - 1. \qquad (3)$$

The second layer of the network can be described by an analogous consideration to form

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)} \qquad (4)$$

where M describes the number of inputs z_j from the first layer, k the number of neurons of the second layer, $w_{kj}^{(2)}$ the weights and $w_{k0}^{(2)}$ the biases. To give a set of network outputs y_k, the activations are transformed using an appropriate activation function to form

$$y_k = \sigma(a_k). \qquad (5)$$

The activation function for the second layer used in this approach is the softmax activation function which can be described as

$$\sigma(a) = \frac{2}{1 + e^{-a}}. \tag{6}$$

The whole feed-forward neural network can be described as a combination of the previous stages and takes the form the following equation

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right). \tag{7}$$

The output y_k of the network is given as a non-linear function of a set of input signals $\mathbf{x}$ controlled by a vector of adjustable weights $\mathbf{w}$ [7].

C. Cross-validation

Cross-validation is a technique used to evaluate the performance of a model in machine learning. With new data sets, which were not used in the training phase, the performance of the prediction is tested. This is done by partitioning a data set into subsets for training and testing the algorithm. Since cross-validation does not use all data when developing a model, it is a commonly used method to prevent over-fitting during training. One of the most common cross-validation techniques is k-fold cross-validation. The data set is divided into k subsets of approximately the same size. Each subset is used to validate the model, while the remaining subset is used for training. In total, this process is repeated k times to use each subset once for validation. The k validation runs allow the mean performance of the algorithm to be determined [9].

D. Receiver operating characteristic

A receiver operating characteristic (ROC) curve illustrates the relative trade-off between benefits (true positives) and costs (false positives) of a classification model at all classification thresholds. A ROC diagram represents the true positive rate over the false positive rate, as shown in Fig. 2.

The true positive rate TPR is given by

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

where TP are the true positive classified samples and FN the false negatives. The false positive rate FPR is given by

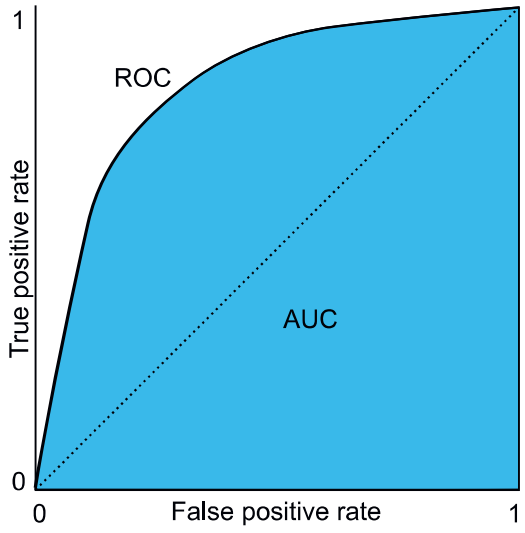


Fig. 2: Receiver operating characteristic (ROC) curve and the area under curve (AUC)

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

where FP are the false positive classified samples and TN the true negatives [10]. The dotted line in Fig. 2 represents the strategy of randomly guessing a class. Classifiers whose ROC curves lay in the lower right triangle are understood as misinterpretation of the data. Whereas classifiers whose ROC curves lay in the upper left triangle are considered good classifications, depending on the size of the area under the ROC curve (AUC) is. Since the AUC is a portion of the unit square, its value will be between 0 and 1. For the strategy of random guessing, the AUC value is 0.5. The height of the AUC value thus describes the quality of the classifier, whereby a value of 1 represents the ideal classifier [11].

III. Experimental Methodology

A. Tapered roller bearing structure

Tapered roller bearings are a special type of roller bearings that can support both axial and radial forces. Due to the rollers being situated perpendicular to the rolling axis, these bearings can transmit higher forces than deep groove ball bearings. The essential components of such a bearing consist of the outer race, the rolling elements, the inner race and the cage. The cage holds the rolling elements on the running surface of the inner race. The outer race is detached from the remaining part of the bearing. During the rotation of the bearing, the greatest forces are transmitted between the running surfaces of the inner race, the outer race and the

rolling elements. The schematic structure of the investigated bearing is shown in Fig. 3.

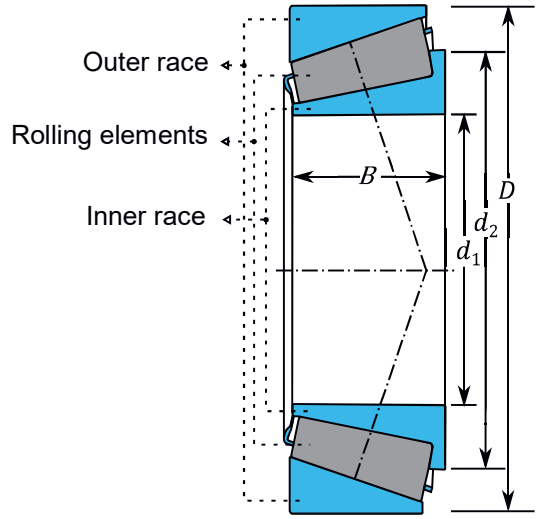


Fig. 3: Schematic structure of the investigated tapered roller bearing.

The bearing used in this paper has an inner diameter of d_1 = 35 mm, a diameter from the center to the rolling elements of d_2 = 51.95 mm, an outer diameter of D = 72 mm, and a width of B = 17 mm. Due to the force applied to the contact surfaces, these surfaces are most susceptible to damage [12]. Within the scope of this paper, damage to the outer race, inner race and rolling element surfaces is specifically investigated.

B. Reproducible bearing damage

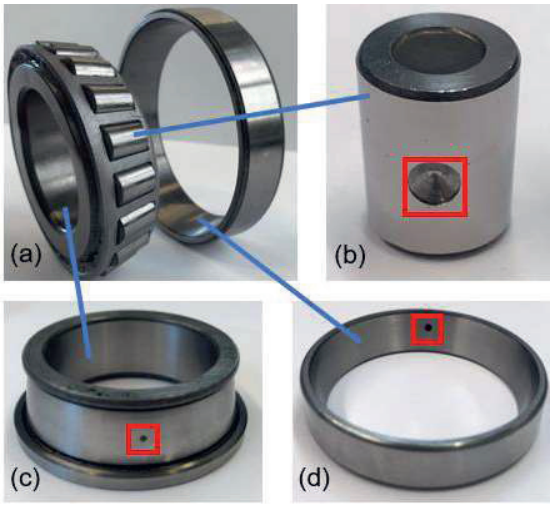


Fig. 4. Entire tapered roller bearing (a) and its individual parts with the artificially induced damage to the rolling element (b), the inner race (c) and the outer race (d).

A total of twelve damaged bearings are being investigated. Four different bearings have been examined for an outer race damage, four for an inner race damage and four for a rolling element damage. The damage is introduced into the inner races, outer races and one rolling elements via erosion. Each class of damage is examined individually. For each damage class a damage of 1.5 mm, 2 mm, 2.5 mm and 3 mm diameter is examined. Fig. 4 (a) shows the tapered roller bearing type examined and the positions of the respective damage for the individual test objects. Fig. 4 (b) shows a rolling element damage, Fig. 4 (c) an inner race damage and Fig. 4 (d) an outer race damage.

C. Experimental setup

In order to investigate the artificial damage, a laboratory test bench has been set up, which is shown in Fig. 5. A bearing socket has been manufactured which can hold different bearing sizes. Piezoelectric vibration transducers can be attached to this by means of screw connections with a defined prestressing force to ensure a good signal transmission [13]. A defined radial force can be applied to the bearing's outer ring by means of a screw attached to the bearing block. The bearing, which is set in the socket, is connected with a 450 W three-phase motor via a low-vibration coupling. A frequency inverter is used to set the speed of the three-phase motor to an adjustable value.

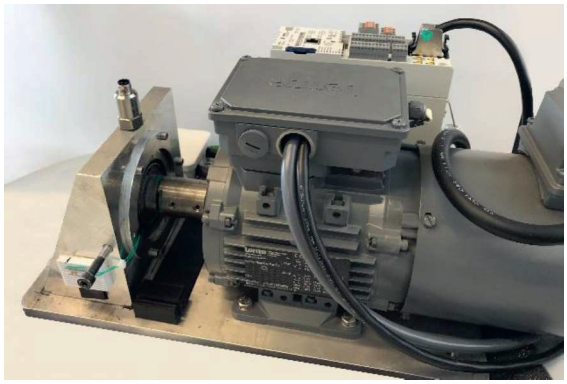


Fig. 5: Test bench for the investigation of bearing damages consisting of a three-phase motor, a frequency inverter, a bearing socket and a piezoelectric vibration transducer.

A total of 4 intact tapered roller bearings and 12 with artificial damage have been examined. Damage in the form of circular surfaces with diameters of 1.5 mm, 2 mm, 2.5 mm and 3 mm have been investigated for each of the three damage classes (outer race, inner race, rolling elements). All tested bearings have been investigated on the laboratory test bench at speeds of 100 min^{-1}, 200 min^{-1}, 300 min^{-1}, 400 min^{-1}

and 500 min^{-1}. The piezoelectric vibration transducer (IDS Innomic, ICS80) mounted on the bearing block via a screw connection has a linear frequency range up to approximately 10 kHz. In order to cover the entire linear range of the sensor, a sampling rate of 20 kHz is used.

IV. Experimental results

With respect to the classification algorithms, the four bearing states are combined into classes according to Tab. 1.

Tab. 1: List of the damage assigned to the classes for the classification

	Class description
Class 1	Intact bearing
Class 2	Inner race damage
Class 3	Outer race damage
Class 4	Rolling element damage

For each class, a data set of 3440 time windows with a length of 2048 sample points and a sampling rate of 20 kHz is examined. Each measurement series for a class contains measurements of five different rotational speeds for four bearings each. For the investigation of the classification algorithms a 10-folded cross-validation has been used.

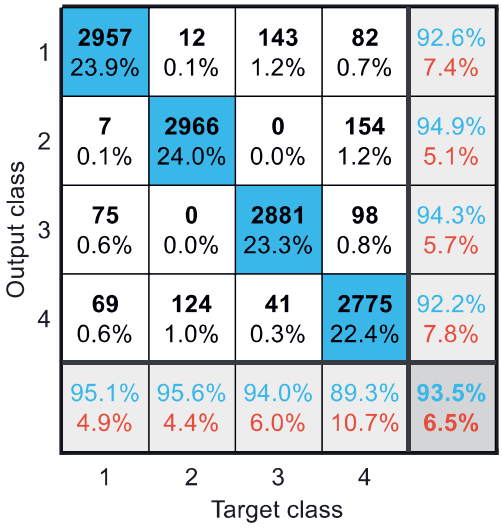


Fig. 6: Confusion matrix of the test data set classified via the feedforward neural network.

Fig. 6 shows the average confusion matrix for the classification of the test data set via the feed-forward neural network after the 10-folded cross-validation. The rows correspond to the predicted class of the neural network and the columns to the true class. Accordingly, the diagonal fields show the correctly classified time windows and the non-diagonal fields show the incorrectly classified time windows. The column on the right shows the percentages of all time windows belonging to each class that are correctly or incorrectly classified. The bottom row shows the percentages of all time windows belonging to each class that are correctly and incorrectly classified. The field at the bottom right shows the total accuracy.

The classification of the neural network is very accurate for all classes with an accuracy for each class of at least 89.3 %. The overall accuracy of the network is 93.5 %, which can be interpreted as a very good prediction of the damage investigated via the network.

In order to compare the classification of the feed-forward neural network with other classification algorithms, a linear support vector machine, a fine decision tree and a k-nearest neighbor (k = 10) classification have been evaluated. For each of these classification algorithms the same data set and features have been used.

For the comparison of the classifier performance the receiver operating characteristic (ROC) curve for each classification algorithm has been determined.

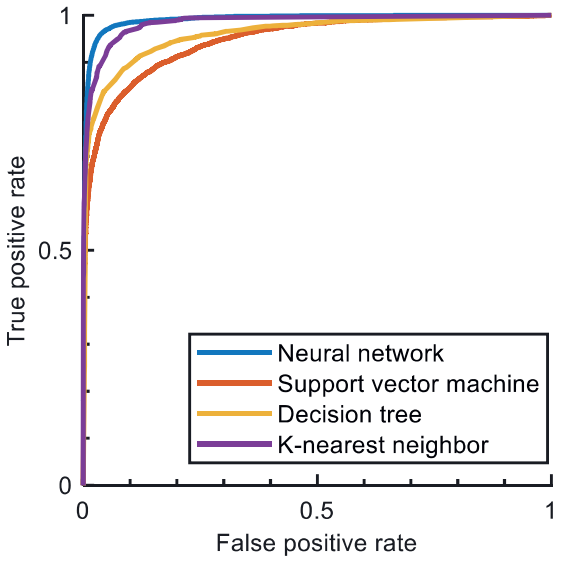


Fig. 7: Receiver operating characteristic (ROC) for the investigated bearing classes determined for four different classifiers.

The averaged results after the 10-folded cross-validation for the ROC curves are shown in Fig. 7. The x-axis shows the false positive rate and the y-axis the true positive rate of the classification. For each individual classification algorithm, the mean value of all four ROC curves for the respective classes is considered. The linear SVM achieves the worst ROC curve in a visual comparison. Followed by the fine decision tree and the k-nearest neighbor algorithm. The

best result is achieved by the feed-forward neural network.
In order to quantify the performance of the ROC curves and thus of the classifiers more precisely, the average AUCs are calculated and listed in Tab. 2.

Tab. 2: Area under curve for the different classification algorithms

Classification algorithm	AUC
Neural network	0.994
Support vector machine	0.968
Decision tree	0.978
K-Nearest neighbors	0.989

Overall, the neural network achieves the best result with 0.994, followed by KNN with 0.989, Decision tree with 0.978 and SVM with 0.968. However due to the high AUC values, all classifiers can therefore be rated as suitable for the detection of bearing damage.

V. Conclusion

This paper presents a study of reproducible damage to inner races, outer races and rolling elements of tapered roller bearings. In the first part of the study, damage caused by erosion in the form of circular surfaces with a diameter of 1.5 mm to 3 mm has been introduced into the various components of the bearings. By varying the speed, several series of measurements have been recorded for each bearing using piezoelectric transducers. A total of 71 features have been calculated for the measurement data and different classification algorithms have been compared using ROC curves and AUC. The best classification result has been achieved by the feed-forward neural network.
In future investigations, the influence of the damage size on the classification accuracy will be investigated. At the same time, the artificial damage will be compared with real damage of bearings from agricultural machines in order to derive a more robust classification. In addition, more damaged bearings will be investigated to build the test and training data set from different bearings. Beyond that wear phenomena, such as pitting, will be investigated in order to predictively determine the remaining useful lifetime of the bearing.

References

[1] G. Brian and K. Starry, “Rolling Element Bearing Analysis,” *Materials Evaluation, Vol. 70, No. 1 pp:78-85,* 2011.

[2] A. Beering, R. Klemm and K.-L. Krieger, “Analyse und Lokalisation von Zahnbrüchen in mehrstufigen Getrieben mit Hilfe von vibroakustischen Signalen,” in *DAGA 2019 - 45. Jahrestagung für Akustik*, Rostock, 2019.

[3] W. Smith and R. Randall, “Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study,” *Mechanical Systems and Signal Processing 64-65 pp: 100–131,* 2015, doi: 10.1016/j.ymssp.2015.04.021.

[4] C. Luo, M. Jia and Y. Wen, “The diagnosis approach for rolling bearing fault based on Kurtosis criterion EMD and Hilbert envelope spectrum,” in *2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference* , Chongqing, 2017, doi: 10.1109/ITOEC.2017.8122439.

[5] H. Van Khang, J. Sri Lal Senanayaka and K. G. Robbersmyr, “Towards Online Bearing fault detection using Envelope Analysis of vibration signal and Decision Tree Classification Algorithm,” in *2017 20th International Conference on Electrical Machines and Systems* , Sydney, 2017, doi: 10.1109/ICEMS.2017.8056146.

[6] Verein Deutscher Ingenieure, Körperschallmessungen zur Zustandsbeurteilung von Wälzlagern in Maschinen und Anlagen, Beuth Verlag GmbH, 2013.

[7] C. M. Bishop, Pattern Recognition and Machine Learning, Cambridge: Springer Science & Business Media, 2009.

[8] D. Baughman and Y. Liu, “Fundamental and Practical Aspects of Neural Computing,” *Neural Networks in Bioprocessing and Chemical Engineering,* pp. 3-31, January 1995, doi: 10.1016/B978-0-12-083030-5.50008-4.

[9] I. H. Witten, E. Frank and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, Elsevier, 2011, doi: 10.1145/2020976.2021004.

[10] J. Davis and M. Goadrich, “The Relationship Between Precision-Recall and ROC Curves,” in *International Conference on Machine Learning*, Pittsburgh, 2006, doi: 10.1145/1143844.1143874.

[11] T. Fawcett, “ROC Graphs: Notes and Practical Considerations for Researchers,” *Pattern Recognition Letters,* pp. 1-38, 2004.

[12] SKF Group, “Railway technical handbook,” Göteborg, 2012, pp. 122-135.

[13] W. Kuttner, Praxiswissen Schwingungsmesstechnik, Wiesbaden: Springer Fachmedien, 2015, doi: 10.1007/978-3-658-04638-5.

DOI 10.5162/sensoren2019/5.4.3

Ultraschallsensoren für die Instandhaltung

Peter Holstein[1], Nicki Bader[1], Christian Probst[1], Steffen Seitz[2]
[1]SONOTEC GmbH, Nauendorfer Str. 2, 06102 Halle
[2]TU Dresden, Mommsenstr. 12 01062 Dresden

Zusammenfassung

Im Beitrag wird ein Konzept und dessen Realisierung vorgestellt, dass die Erfassung von Ultraschall in einem breiten Frequenzbereich ermöglicht. Die physikalischen Grundlagen der akustischen Vorgänge im Ultraschallbereich wurden analysiert, um darauf aufbauend ein neues Sensor- und Messgerätekonzept zu entwickeln und technologisch umzusetzen. Sensorische Grundlage hierfür sind neu entwickelte Sensoren, die für einen breiteren Frequenzbereich einsetzbar sind. Damit sind Vorgänge neben den bisher hauptsächlich bewerteten zeitlichen Fluktuationen auch spektral bewertbar. Dies erweitert die Möglichkeiten für den Einsatz komplexerer Datenbewertungen erheblich.
Die Körperschallsensoren basieren dabei auf sogenannten piezoelektrischen Kompositmaterialien. Der Frequenzgang dieser Sensoren ermöglicht die spektrale Nutzung der aufgezeichneten Signale. Damit steht eine breite Palette an Signalverarbeitungsoptionen zur Verfügung, die sowohl in Echtzeit als auch für aufwendigere Postprocessing-Verarbeitung genutzt werden können. Die Vorteile der Sensorik kommen auch dadurch zum Tragen, dass viele der gegenwärtig verfügbaren modernen Technologien integriert werden. Dies betrifft Elektronik, embedded IT, Software zur Visualisierung und Kommunikation und industriell orientierte Mathematik aber auch Design, Ergonomie und Produzierbarkeit. Auf Basis dieser Sensoren wurde deshalb eine neue Plattform für Prüftechnik für die Instandhaltung konzipiert, die neue Features und Möglichkeiten bietet– insbesondere im Hinblick auf die Kundenspezifik von Lösungen.
Im Beitrag wird der Nutzen des neuen Sensorkonzepts für die Erschließung neuer Anwendungen vorgestellt. Insbesondere werden die Daten genutzt, um mit Verfahren des maschinellen Lernens bessere Voraussagen zu Lebensdauern, zur Bewertung von Betriebszuständen und Maschinenteilen zu generieren.

Keywords: Ultraschall, Körperschallsensor, Instandhaltung, maschinelles Lernen

Motivation

Ultraschall spielt im Kontext der Methoden der Technischen Akustik in der Instandhaltung eine sehr spezialisierte Rolle. Dies widerspiegelt sich darin, dass nur relativ einfache Verfahren auf wenige ausgewählte Anwendungsfälle ausgerichtet sind. Beispiele sind die Suche von Druckluftlecks oder einfache Bewertungen von Fehlern in Kugellagern. Das Finden weiterer Anwendungen wird dabei vor allem durch sehr konservative technologische Lösungsansätze eingeschränkt, die im wesentlichen darauf beruhen, dass schmalbandige Resonanz-Sensoren eingesetzt werden. Dies sichert zwar eine hohe Sensitivität, schränkt aber andererseits durch die Bandpassbegrenzung die Bewertung komplexer physikalisch-technischer Prozesse erheblich ein. Dies ist symbolisch in Abbildung 1 skizziert. Die rote Linie steht für die Verwendung der Ultraschallverfahren in einem, engen Frequenzbereich (analoge Heterodyn-technologie). Damit ist nur eine sehr begrenzte Einsicht in die technologischen in die physikalischen und technischen Prozesse möglich. Die Grenzen zwischen Verfahren sind oft willkürlich. Der Balken zwischen 10 und 100 kHz steht für die im Beitrag beschrieben Technik.
Die Erweiterung des Frequenzbereich scheint folgerichtig, wenn man grundlegende physikalische Phänomene wie:

- Turbulenz (Strömungsakustik)
- Reibung (Rauigkeit von Materialpaarungen)
- Mikrorisse und Rissfortschritt
- Kavitation
- Entladungen
- Impulskontakte (lose Bauteile …)
- Eigenschwingungen kleiner Strukturen

betrachtet. Für jedes dieser Phänomene kann die Signalverarbeitung angepasst und optimiert werden, um (im Sinne des Einsatzes in der Instandhaltung) Verschleißbewertungen, Trendanalysen, Einsatzentscheidungen und Prognosen sichere Daten zu liefern.

Es sei darauf hingewiesen, dass der Einsatz breitbandiger Sensoren die zu verarbeitenden und anfallenden Datenmengen erheblich vergrößern. Entsprechende intelligente Verfahren zur Reduktion der Datenmenge – möglichst in Echtzeit – wie Mustererkennung, automatische Fehlerklassifizierung u.ä. bieten sich im Kontext der aktuellen Entwicklungen der Behandlung großer Datenmengen an.

Gerätekonzept

Das Gerätekonzept ist unmittelbar mit der Sensorfamilie gekoppelt. Für Prüfaufgaben im industriellen Umfeld – insbesondere in der Instandhaltung – steht ein portabler Messdatensammler zur Verfügung, der für Anwendungen im Frequenzbereich des für die Instandhaltung relevanten Bereich ausgelegt ist und aufgrund der internen Struktur (Schnittstellen, Prozessor, Speicher, Echtzeitmodul, Betriebssystem) äußerst flexibel an verschiedene Messaufgaben angepasst werden kann [1].

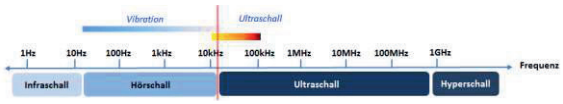


Abb. 1: Einordung der Ultraschallmethoden in der Instandhaltung in das akustische Spektrum.



Abb. 2: Basiseinheit für die Ankopplung digitaler Sensoren. Die DAQ-Einheit verfügt über einen integrierten FPGA für Echtzeit.

Die Kombination aus digitaler Prüftechnik und digitaler Sensorik in Verbindung mit der breitbandigen Betrachtung physikalischer Vorgänge ist in der Instandhaltung relativ neu [2,3]. Entsprechend sind Verfahren, Algorithmen und Normen usw. erst im Entstehen.

Luftschallsensoren

Luftultraschallanwendungen werden in der Instandhaltung in zwei generellen Betriebsmodi eingesetzt. Zum einen wird direkt von einem technischen System erzeugter Ultraschall nachgewiesen und bewertet. Zum anderen kann Ultraschall aktiv von einem Ultraschallsender abgestrahlt und dann entsprechend der Prüfaufgabe in einiger Distanz nachgewiesen werden. Eine typische Anwendung ist die Prüfung technischer Systeme auf Dichtheit. Die Gründe für die Verwendung der Ultraschallfrequenzen sind unterschiedlich. Oft ist in industriellen Umgebungen der Umgebungslärm zu groß um akustische Methoden (zumindest mit einfachen Verfahren) sinnvoll einzusetzen zu können. Ein „Ausweichen“ zu höheren Frequenzen kann in wichtigen Fällen Ziel führend sein. Für die Luftschallanwendungen wurde ein ähnliches Konzept für digitale Ultraschallmikrofone entwickelt. Als Wandlerelemente dienen hierbei ¼” Mikrofonkapseln, die im hochfrequenten Bereich über einen geeigneten Frequenzgang verfügen. In die digitalen Ultraschallmikrofone sind Frequenzfilter, Verstärkungsstufen, A/D-Wandlung und Digitalschnittstelle integriert. Aufgrund der Analogie zu Messmikrophonen sind damit im Rahmen der möglichen Genauigkeit Messungen im Ultraschallbereich möglich, die zu den im Hörschall üblichen Bezugswerten und Berechnungen kompatibel sind. Die Empfindlichkeit dieser digitalen Mikrofone ist mit der von in der Instandhaltung weit verbreiteten einfachen resonanten Empfängern vergleichbar. Allerdings eröffnet der breite Frequenzbereich auch bei Luftultraschall erweiterte Anwendungsmöglichkeiten und bessere Bewertungen [4].



Abb. 3: Links - Digitaler Luftultraschallsensor BS10 mit verschiedenen Aufsätzen zur Suche und Bewertung von Lecks in Druckluftsystemen
Rechts - Kompakte Integration des digitalen Luftultraschallsensors in einen Parabolspiegel zur Messung über größere Distanzen.

Körperschallsensoren

Für Anwendungen in der Instandhaltung sind Frequenzen im Bereich von etwa 20 kHz bis 100 kHz von Bedeutung. Damit liegen diese Frequenzen über dem Erfassungsbereich von Schwingungssensoren und unter dem von den für Anwendungen in der akustischen Emission verfügbaren Körperschallsensoren [5].
Ziel der breitbandigen Sensorik ist die Aufwertung der Ultraschalldiagnoseverfahren durch Digitalisierung und Nutzung der spektralen Information. Die Notwendigkeit breitbandiger Körperschallsensorik ergibt sich aus vielen Anforderungen in der Instandhaltung (Maschinendiagnose, Strömungsprobleme).
Kernstück der Körperschallsensoren ist ein Transducerelement aus piezoelektrischem Kompositmaterial. Das Herstellungs- und Funktionsprinzip ist in Abbildung 4 veranschaulicht. Bei monolithischen Transducerelementen tritt eine Vielzahl von Resonanzen auf [2,6], die spektrale Bewertungen erschweren, wohingegen bei der stäbchenförmigen Anordnung der piezoaktiven Elemente in wesentlichen die Dickenschwingung dominiert. Damit ist die Frequenzcharakteristik [2,6] wesentlich „glatter“ und Auswertealgorithmen numerisch besser beherrschbar.

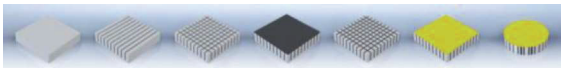


Abb. 4: Technologische Schritte bei der Herstellung des Sensorelements als piezo-elektrische Compositmaterial - Piezokeramikplatte, erster Schnitt, zweiter Schnitt, Einfüllen der Matrix, Schleifen auf Enddicke, Metallisieren und Polarisieren, Zuschnitt der Endgeometrie

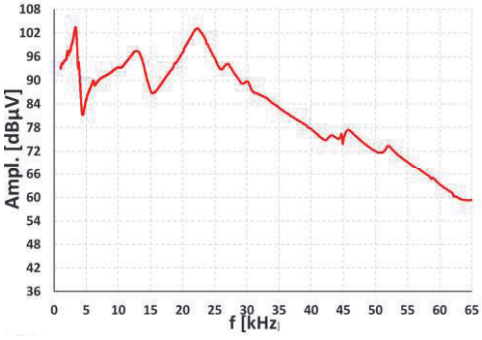


Abb. 5: Frequenzkennlinie des kompletten Sensors T10



Abb. 6: Ausführungsform des Sensors

Die digitalen Bauformen der Körperschallsensoren sind für wichtige Anwendungen in der Instandhaltung optimiert. Diverse Waveguides erlauben leichte Kontakte zu Prüflingen und damit schnelle Prüfungen.



Abb. 7: Digitale Variante des Körperschallsensor BS20. Die unterschiedlichen Kopplungen sind an die jeweiligen Anwendungen angepasst.

Allerdings ist die Frequenzkennlinie von den Vorläufen beeinflusst. Deshalb werden diese vor allem bei qualitativ orientierten Prüfungen (z. B. an heißen oder schlecht zugänglichen Stellen) eingesetzt.

Kalibrierung der Sensoren

Die für die Kalibrierung von Schwingungssensoren eingesetzten Systeme sind nicht für Frequenzen ausgelegt, die bis in die niederen Ultraschallfrequenzen reichen. Vibrations- und Ultraschallsensoren unterscheiden sich im Aufbau und der Funktionsweise [5]. Ein wichtiger Unterschied ist die bei den Körperschallsensoren fehlende seismische Masse. Damit ist eine direkte Kalibrierung bezüglich üblicher Bezugsgrößen für Vibrationen schwierig.
Ziel ist es, unter Anwendung und Modifizierung geeigneter Kalibrierverfahren, die Sensoren in einem möglichst großen Frequenzbereich zu kalibrieren und die Reproduzierbarkeit der Kennlinien bei der Herstellung und im Einsatz zu beherrschen. Oft werden die Sensoren mit der Hand für Prüfzwecke gehalten. Dies macht Vergleiche (und Kalibrierungen) besonders herausfordernd. Nach bisherigem Kenntnisstand sind Kalibrierungen bis etwa 60 kHz realisierbar [7]. Kern der Verfahren sind gewisse Optimierungen der Referenzen, die verbunden sind mit entsprechenden

konstruktiven Anpassungen. Darüber hinaus sind weitere Untersuchungen notwendig. Auf dieser Basis können die Zielparameter für die Modifikation der Kalibrierungsverfahren festgelegt werden.

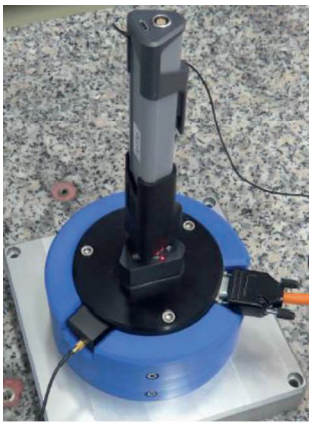


Abb. 8: Schwingregelsystem, hier mit Körperschallsensor BS20 [7]

Anwendungsbeispiel

Die Körperschallsensoren wurden bei der Instrumentierung von Langzeitversuchen an einen Maschinenprüfstand eingesetzt. Ziel war dabei die Nutzung breitbandiger Ultraschalldaten für die Bewertung entstehender Schäden in Kugellagern. Dafür wurden zum einen Kugellager mit „künstlichen“ Schäden präpariert und zum anderen in Dauerversuchen hoch belastet. Als Input für Lernverfahren wurden die hochfrequent (200 kHz) abgetasteten Daten der Körperschallsensoren verwendet. Der Prüfstand wurde mit mehreren Sensoren instrumentiert. Die Belastungsversuche dauerten mehrere Wochen bis zum völligen Versagen der Lager.



Abb. 9: Präparierte Kugellager; Im Außenring von Kugellagern wurden mit einem Laser definierte Schäden mit wachsendem Ausmaß eingebracht.

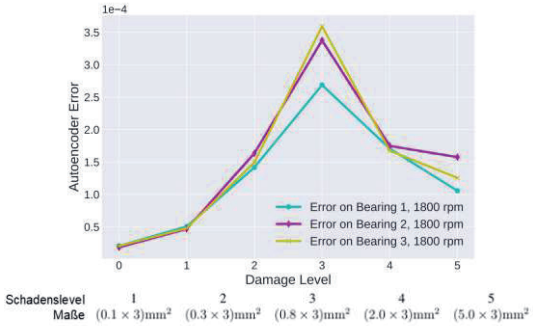


Abb. 10: Darstellung des Autoencoder-Fehlers über dem Schadenslevel

Wird der Schaden ähnlich groß wie die Abmessung der Wälzlager (4,7 mm) kommt es zu einer Stoßentlastung, welche vom Autoencoder wieder als geringerer Fehler wahrgenommen wird (siehe Abb. 10) Die Möglichkeit eines solchen Verhaltens wird auch in der VDI3832 [8]. Dies demonstriert, dass vereinfachende Prüfdaten in vielen Fällen nicht zielführend sein werden.

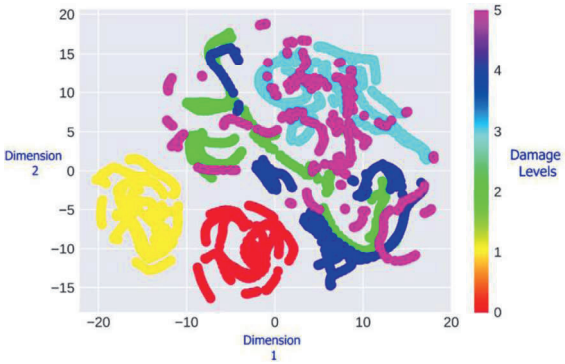


Abb. 11: Plot der im Autoencoder gruppierten Samples der verschiedenen Schadensklassen

Die Gruppierung für Schadensklassen mit niedrigem bzw. keinem Fehler ist deutlich erkennbar (Abb. 11). Damit ist eine Unterscheidung auch von geringen Abweichungen (1) vom unbeschädigten Fall (0) möglich. Die Gruppierung der höheren Schadenszustände ist erkennbar. Die Klassifikation wird aber komplizierter. Es ist allerdings zu erwarten, dass mit der Erhöhung der Datenmenge (Trainigsdaten) die automatischen Auswertungen besser werden [9].

Zusammenfassung

Die Einführung breitbandiger Sensoren für Prüfaufgaben in der Instandhaltung ermöglicht eine deutliche Erweiterung der Anwendungsbreite von Ultraschall-verfahren und damit der akustischen Methoden. Deren Anwendung erfordert sowohl für Luftultraschall als auch für Körperschall neue Kalibriermethoden sowie leistungsfähige spezifische Messtechnik und angepasste Algorithmen.

Danksagung

Die Arbeiten wurden teilweise durch das Bundesministerium für Bildung und Forschung (BMBF-Programm *zwanzig20, Förderkenn-zeichen 03ZZ0518C*) und das Bundes-ministerium für Wirtschaft und Energie (*Projekt sUSe - Förderkennzeichen ZF4327503PO8*) gefördert.

Literaturnachweis

[1] *https://www.sonotec.de/produkte/vorbeugende-instandhaltung/produkte/sonaphone/*

[2] C. Probst, P. Holstein, Application of ultrasound technology in condition monitoring, *Condition Monitor, 372(2018)5-10*,

[3] P. Holstein, C. Probst, A. Tharandt, G. Werner, T. Werner, Condition Monitoring with Ultrasound – new Approaches*, Proceed. 3rd International Rotating Equipment Conference (IREC), Pumps, Compressors and Vacuum Technology*, Düsseldorf, Germany, 14 – 15 September 2016, ISBN- 978-3-8163-0697-9

[4] P. Holstein, M. Barth, C. Probst, Acoustic methods for leak detection and tightness testing, *Proceedings, 19th World Conference on Non-Destructive Testing*; 13 - 17 June 2016 in Munich, Germany

[5] T. Petzsche, Sensing Structural Borne Noise in Solid State Materials and Related Applications, *Proceed. 21. Koll. Schallemission,* Fulda 2017, *https://www.dgzfp.de/Portals/schallemission2017/BB/20.pdf*
L. Soares, M. Dumont, T. Petzsche, Sensing Structural Borne Noise in Solid State Materials and Related Applications, *Proceed. EuroRegio2016*, 13.-15.06.2016, Porto, Portugal *http://www.sea-acustica.es/fileadmin/-Oporto16/223.pdf*

[6] P. Holstein, C. Probst, H. Münch, A. Klepel, A. Tharandt, Application of Acoustic Emission in Machinery and Process Monitoring*, Proceed. 31st Conference of the European Working Group on Acoustic Emission* (EWGAE), 3.-5.9.2014, Dresden, Germany,

[7] P.Holstein, C. Probst, P. Begoff, H. Schulze Kalibrierung von Körperschallsensoren, *Fortschritte der Akustik - 44. Deutsche Jahrestagung für Akustik,* 19.-22.03. 2018 in München, ISBN-10: 3939296139

[8] VDI 3832, *https://www.vdi.de/nc/richtlinie/vdi_3832-koerperschallmessungen-_zur_zustandsbeurteilung_von_waelzlagern_in_maschinen_und_anlagen/*

[9] S. Seitz, R. Tetzlaff, P. Holstein, C. Probst, Towards Fault Diagnosis with Attention, *Proceedings 26th Int. Congr. on Sound and Vibration*, 7-11.07.2109, im Druck

DOI 10.5162/sensoren2019/5.4.4

Homodyne Vector Network Analysis as a Tool for the Real-Time Measurement of Electrical Material Parameter Distributions in the Field

Ronny Peter, Gerhard Fischerauer
Universität Bayreuth – Chair of Measurement and Control Systems
Universitätsstraße 30, 95447 Bayreuth, Germany

Abstract

The spatial distribution of electrical material parameters (e.g. permittivity and conductivity) can be a valuable indicator of the state of a chemical process or the condition of the processing system. Unfortunately, there are very few measurement systems that could handle this task in the field at low cost and with small shape factors. A potential solution currently under research for the determination of material parameter distributions is based on electromagnetic resonances inside chemical reactors. In the laboratory, vector network analyzers (VNA) and personal computers are used, which is expensive. This contribution reports on an application-specific stand-alone homodyne VNA with integrated data processing as an effort to transfer the laboratory-proven method to the field. The quality of the approach, validated by comparison with commercial VNAs, is shown to suffice for typical field applications.

Keywords: Material parameters, in-process measurement, vector network analyzer, field application.

Introduction

Synthetic fertilizers, processing of crude oil, pharmaceuticals, and exhaust gas treatment are just a few important application examples for the use of chemical reactions [1]. To run the processes with high efficiencies, one needs to know the process state. This includes the spatial distribution of the reactants and of thermodynamic variables such as temperature.

By way of an example, consider a three-way-catalytic (TWC) converter. It consists of a ceramic honeycomb carrier coated with a catalytically activated oxygen storage component, all placed in a metallic housing. When lean (oxygen-rich) exhaust gases flow through the TWC, its oxygen storage component stores the excess oxygen to prevent the generation of nitrogen oxides. In rich (oxygen-deficient) atmospheres, the storage component releases oxygen to burn excess hydrocarbon and carbon monoxide molecules. It is evident that the relative oxygen storage level (ROSL) in the TWC is an important process parameter. Knowing it allows one to better control the efficiency of the internal combustion engine (ICE) and its emissions by running it with alternately lean and rich air-fuel mixtures.

The state of the art is to indirectly monitor the integral ROSL by lambda probes mounted upstream and downstream of the TWC and measuring the air-fuel ratio in the exhaust gas [2]. A direct and spatially resolved ROSL measurement can be based on the Raman scattering of light, but such systems are expensive and limited to point measurements [3]. Another method is microwave imaging which reconstructs a 2- or 3-dimensional distribution of material parameters by evaluating the electric field inside a metallic enclosure [4]. However, it is hard to sample the electric field at sufficiently many points in a running process and the reconstruction requires substantial computational time.

Another microwave-based method is the use of resonances in a microwave cavity. A metallic enclosure (e.g., the housing of a TWC) serves as the resonant cavity [5] (Fig. 1). The frequencies at which resonances occur are influenced by the electrochemical material parameters inside the cavity. We have dealt with several aspects of this method [6], [7], but the availability of low-cost equipment suitable for field (in-process) use remains an open issue. Potentially viable approaches include the use of mobile-phone equipment [8] and of application-specific scalar network analysis [9].

In the following, we describe a third approach, viz., an application-specific VNA. This includes the background of the cavity perturbation method, our measurement system concept, its implementation, and validation results.

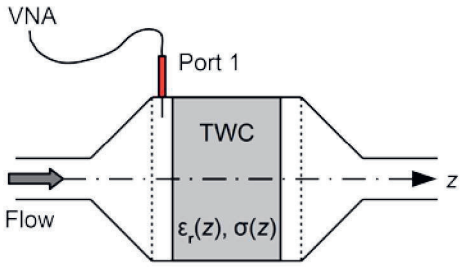


Fig. 1: Schematic drawing of a TWC used as resonant cavity.

The Cavity Resonator Method

Resonant cavities have been used for decades to determine material parameters. The cavity is either completely filled with a homogeneous sample (whole-medium perturbation, WMP) or loaded with a small sample (small-sample perturbation, SSP) [10]. The latter approach does not lend itself for in-process measurements as the sample needs to be prepared and precisely placed inside the cavity. The WMP method is better suited for in-process monitoring, but only so if the process provides homogeneous conditions. This is not always the case.

Axially varying material parameters inside a cavity may be resolved by evaluating several cavity resonances [6]. (Note that in many tubular reactors the material parameters are inhomogeneous in the direction of the flow but can be assumed to be homogeneous over the cross-section with little error [1], [2].) During the reconstruction of the material parameter distribution from the measured resonance parameters, the cavity is approximated by a discretized model with *N* homogeneous sections (Fig. 2). The reconstruction algorithm optimizes the material parameters such that the model predicts the cavity resonances with the least possible error in comparison with the measured data. This approximates the continuous material parameter variation in the axial direction by a stepwise constant distribution.

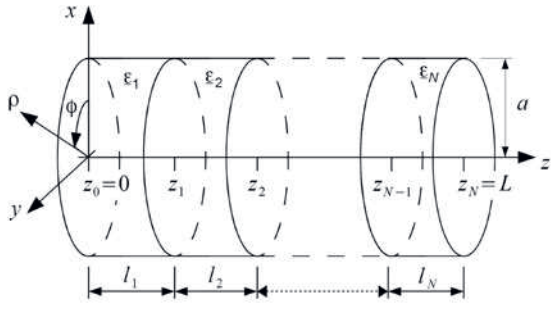


Fig. 2: Discretized cavity model used for the reconstruction of the material parameter distribution inside a cavity.

The method only requires the solution of an (over-)determined system of linear equations. The algorithm is simple and fast enough to be executed on a microcontroller in real-time.

When comparing simulated and measured cavity resonance properties, care has to be exercised that the simulated and measured objects are the same. The cavity is connected to a VNA by a coupling structure. The VNA outputs scattering (S-)parameter spectra, from which resonance curves are constructed. These curves lead to resonance parameters, such as resonance frequencies or Q-factors, by curve fitting [11]. As the coupling influences the resonance parameters – the cavity is said to be loaded –, this influence must be removed from the measured data by de-embedding before the results can be compared with the predictions of the simulation. The distinction between loaded and unloaded resonance parameters is especially important in field applications, which restrict one's control over sample sizes and coupling conditions.

Application-Specific Homodyne VNA

S-parameters can be scalar (magnitude only) or "vectorized", i.e., complex-valued (magnitude and phase). A good de-embedding of coupling effects is usually only achieved with methods based on complex S-parameters [7]. Commercial VNAs are too expensive and bulky for in-process applications in vehicles and similar environments. Measurement solutions for such applications have to meet the following requirements:

- measure complex S-parameters accurately enough for the extraction of resonance parameters
- operate in the low GHz range
- allow measuring times below 1 s
- be low-cost
- be low-profile
- have computational power commensurate with real-time material parameter extraction

Modern commercial VNAs use a (super-) heterodyne receiver design (Fig. 3) [12]. It mixes the RF signal produced by a voltage-controlled oscillator (VCO) and then reflected from the device under test (DUT; here: the cavity resonator) with a local-oscillator (LO) signal to obtain an intermediate frequency (IF). The latter is then sampled with a fast analog-to-digital-converter (ADC). The architecture is rather expensive because it requires two phase-synchronized RF sources, a fast ADC, and a field programmable gate array (FPGA) or a digital signal processor (DSP) to process

the large amounts of data produced by the ADC in real-time. This processing consists mainly of filtering and a fast Fourier-transformation (FFT).

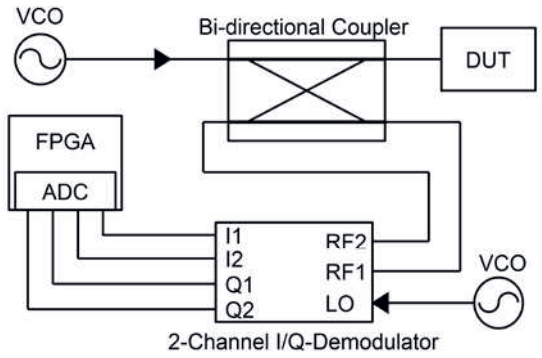


Fig. 3: Schematic drawing of a heterodyne VNA.

A less expensive architecture uses a homodyne receiver (Fig. 4). It mixes the reflected RF signal down to the baseband ($f_{IF} = f_{RF} - f_{LO} = 0$). Hence, a single oscillator can act as both RF and LO source, one can use a slower (less expensive) ADC, one needs no FFT, and a single measurement of the baseband signal suffices. A rather simple microcontroller (MCU) is suited for this task.

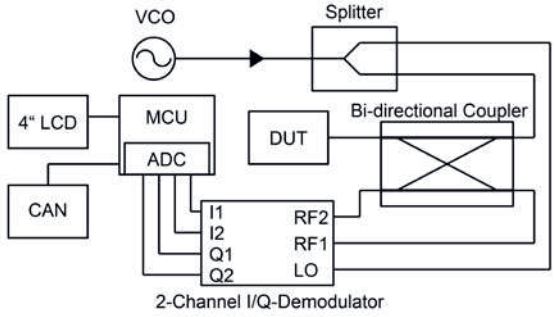


Fig. 4: Schematic drawing of a homodyne VNA.

The main disadvantage of the homodyne design is the LO leakage [13]. The self-mixing of the LO signal with the leaked LO signal produces a DC offset which will be superimposed on the DC signal resulting from the measurement. Another disadvantage is the higher pink noise level, compared to a heterodyne receiver. The resulting difficulty of a homodyne VNA to cope with very small signals is, however, inconsequential for resonance parameter measurements.

We have designed and implemented an application-specific homodyne VNA (Fig. 5). It comprises a digital programmable phase-locked loop (PLL) with a VCO as RF source, a 16-bit ADC with SPI-Interface for sampling, a bi-directional coupler, and an IQ-demodulator IC. The heart of the system is the STM32F7 development board running at 216 MHz. The parts list is completed by several amplifiers, filters, and power splitters. The only part costing more than 10 euros in large quantities is the MCU.

The system meets the following specifications:

- frequency range from 1 to 2.5 GHz (limited only by bi-directional coupler)
- measuring time ca. 311 ms for a 101-point spectrum
- resolution 10 kHz
- maximum of 10,001 frequency points
- built-in resonance parameter extraction

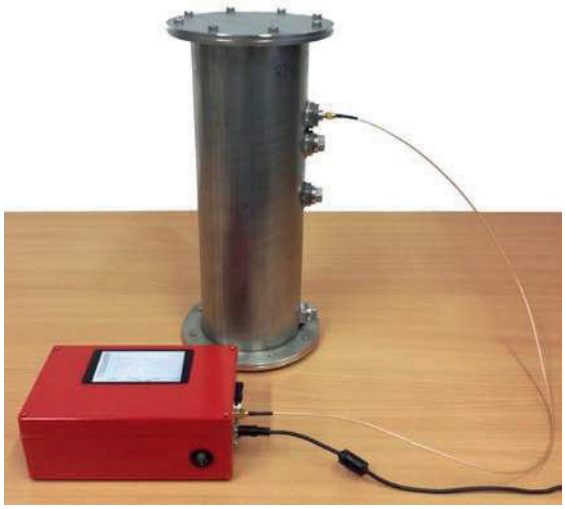


Fig. 5: Application-specific VNA coupled to a circular cylindrical cavity resonator.

Comparison with Commercial VNAs

The application-specific VNA (A) was validated by comparison with an Agilent N5230A PNA-L 300kHz-20GHz (B) and a Rohde&Schwarz ZVRe 9kHz-4GHz (C). Unless noted otherwise, the tests were performed with 6,001-point spectra from 1.4 to 2 GHz (frequency step size 100 kHz). All investigated cavities had at least three resonances in this frequency range. The VNAs were one-port calibrated with known calibration elements (short, open, match) [14, p. 484]. The calibration elements used for calibration of VNA A (Rohde&Schwarz ZV-Z132) were assumed to be ideal (except for offset). All measurements were taken with an averaging factor of 10.

Fig. 6 shows the measured negative return loss of various shorted attenuators. The error-less results would be flat lines at twice the nominal attenuation.

It is obvious that the application-specific VNA A performs the better, the higher the input signal level. Return-loss spectra at levels of about 20 dB show a somewhat higher ripple than spectra measured by the commercial VNA B,

and at 30 to 40 dB, the spectra are corrupted by considerable noise.

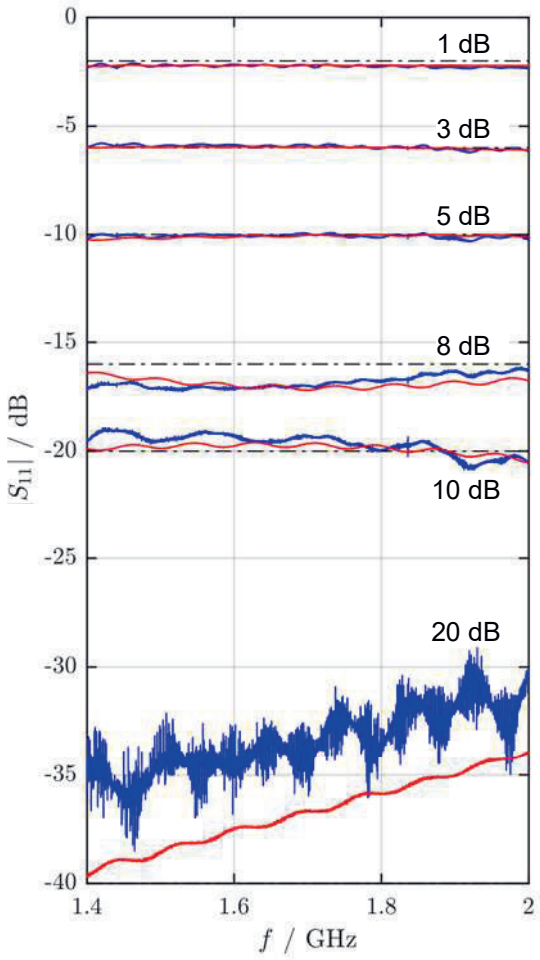


Fig. 6: Return-loss measurements on various shorted attenuators (attenuation as noted in the chart). Blue lines: VNA A. Red lines: VNA B.

What is also evident, however, is the fact that the application-specific VNA is very well suited for the extraction of resonance parameters from typical resonance peaks. Such typical peaks are associated with return losses below about 10 dB for most coupled cavities.

Fig. 6 is an indication of the magnitude error in the application-specific VNA A. To also quantify the phase error, we computed the error vector magnitude (EVM), i.e., the magnitude of the difference between the complex-valued phasors output by VNAs A and B. This was done for every attenuator measured for Fig. 6. The worst-case value at each frequency is shown in Fig. 7. The average worst-case EVM was -38.3 dB, which convincingly demonstrates the practical usability of VNA A.

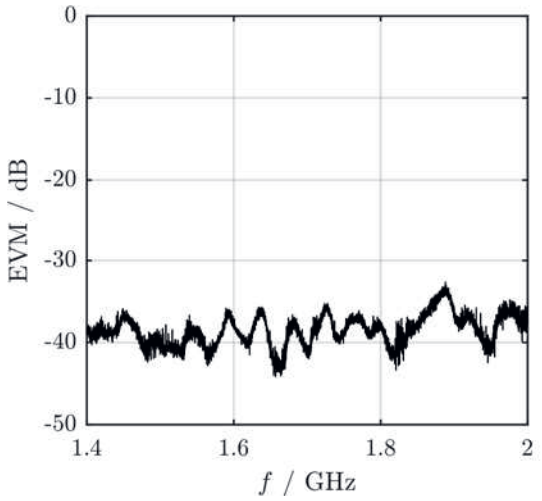


Fig. 7: Worst-case EVM of the VNA-A results for the shorted attenuators measured for Fig. 6.

Further tests were conducted with various cavity resonators as DUTs (Fig. 8). The cavities were of different shape (circular cylindrical and rectangular), of different length (365 mm to 412 mm), coupled with different coupling structures, and filled with different materials.
Fig. 9 shows the negative return-loss spectra from an air-filled cavity with circular cross-section (cavity 4). The three resonance curves produced by VNAs A, B, and C look very similar. From these curves, the resonance parameters were calculated as described in [7, eqn. 1]. This was done by post-processing on a PC for VNAs B and C, whereas VNA A computed the results directly via the implemented MCU routines.



Fig. 8: Cavities used for the characterization of VNA A.

By way of an example, the resonance frequencies and Q-factors of the fundamental mode (TE_{101}) as determined by the three VNAs are listed in Table 1. The good agreement of the results further confirm the practical usability of the application-specific VNA A.

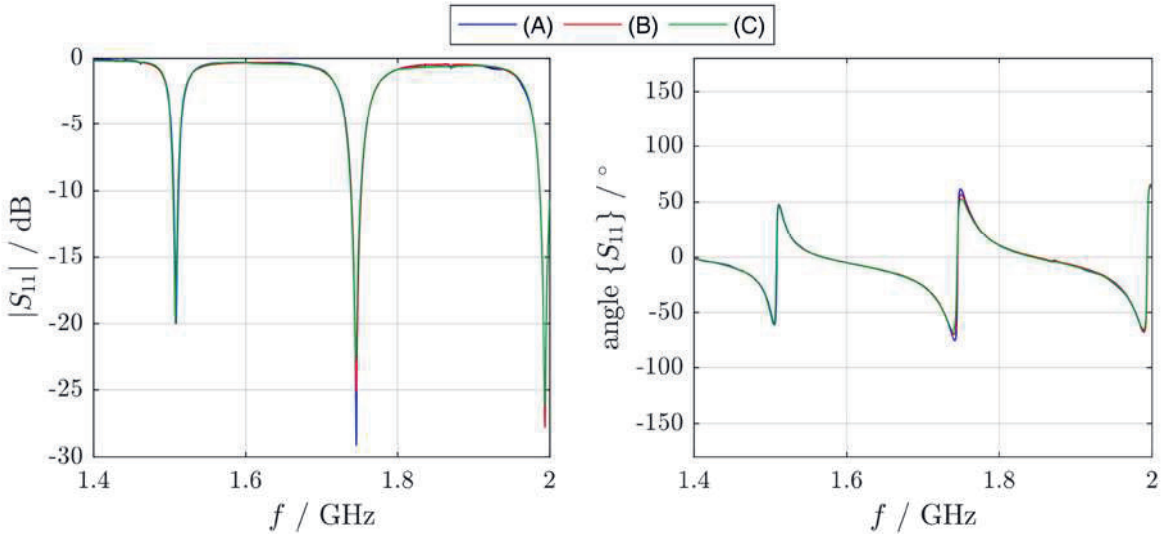


Fig. 9: Results of return-loss measurements with cavity no. 4.

Tab. 1: Resonance parameters of the fundamental mode (TE_{101}) in cavity 1 as determined with VNAs A to C.

VNA	f_{res} / GHz	Q_0
A	1.550874	1635.7
B	1.550687	1735.6
C	1.550785	1692.3

Similar comparisons were made for other modes (in the frequency range from 1.4 to 2 GHz) and other cavities. In the following, we restrict ourselves to the cavities 1 through 4 shown in Fig. 8. The average resonance parameters for each mode of each cavity were used as reference.

The fractional deviations of the resonance frequencies in cavity 1 from the average of the values returned by VNAs A, B, and C are presented in Fig. 10. It looks as if VNA A consistently returns resonance frequencies that are slightly too high, which may be a consequence of an offset error in its frequency base. As the fractional frequency deviation is only about 50 ppm (corresponding to 100 kHz for a 2-GHz resonance), the error may be neglected in practice, as reconstruction algorithms can cope with it [6]. Otherwise the error can be compensated in software by simply adding or subtracting an offset to or from the resonance frequency. Besides, the commercial VNAs B and C do not agree better with each other than with VNA A.

Whereas Fig. 10 is an indication of the accuracy of VNA A (measure of offset from the true value), Fig. 11 serves to judge its precision (measure of scattering around the mean result when an experiment is repeated). Fig. 11(a) shows the empirical standard deviation of the measured fractional resonance frequency deviations of all modes between 1.4 – 2 GHz in cavities 1 through 4, the reference being the average of the values returned by VNAs A, B, and C. Fig. 11(b) is an analogous representation of the standard deviation of the measured fractional Q-factor deviations for the same modes, the references being the average Q-factors returned by VNAs A, B, and C.

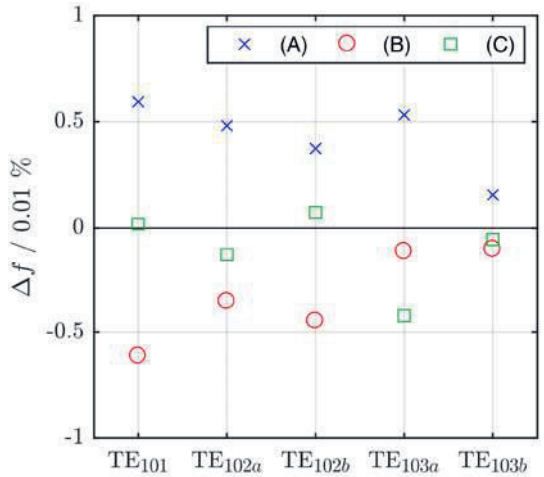


Fig. 10: Selected measured resonance frequencies in cavity 1, expressed as fractional deviation from the average of the values returned by VNAs A, B, and C. "a" and "b" denote split resonances associated with modes that are degenerate in the unloaded case.

Fig. 11 corroborates the notion that the application-specific VNA A allows one to determine resonance parameters with a performance comparable to that of commercial VNAs. This paves the way for in-situ field measurements at a price two orders of magnitude below the price of the laboratory equipment which has often been used to demonstrate the value of the cavity resonance method for material parameter measurements and process monitoring tasks.

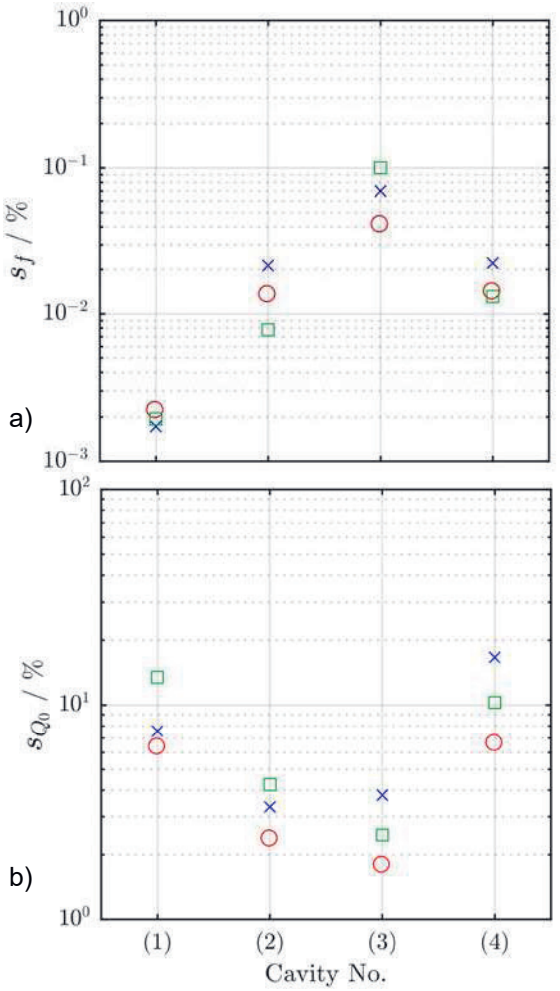


Fig. 11: Standard deviation of measured resonance parameters for all modes between 1.4 – 2 GHz in cavities 1 to 4. a) Fractional resonance frequency deviation. b) Fractional Q-factor deviation.

Measurement Time

The broadband measurements discussed above for validation purposes involve sweep times that may not be compatible with real-time applications. It is more likely that narrow bandwidths in the vicinity of known resonance frequencies would be used for process monitoring tasks to reduce the measurement time. Table 2 lists the required measurement time for ten repeated measurements with 101 discrete frequency points. Again, VNA A performs well in comparison with the commercial VNAs.

Tab. 2: Time required for ten measurements of a 101-point spectrum.

VNA	Measurement time / ms
A	311
B	25
C	520

The built-in parameter extraction algorithm of VNA A takes about 128 ms per resonance. This results in an overall time of 440 ms per resonance. An inhomogeneous cavity with four sections would require the characterization of four resonances, which would take 4×440 ms = 1.76 s. The time needed for the reconstruction of the material parameters of the four sections from the resonance parameters is difficult to state as the convergence of the iterative inversion algorithm depends on the estimated start values. In process monitoring applications, one may assume that the process state varies smoothly. In this case, the previous material parameter values may be taken as rather good start values for the subsequent estimation step. It is then safe to state that the overall time to measure the material parameter distribution in the four-section cavity would not exceed 5 s. This time scale suffices for all chemical processes the state changes of which are limited by mass flow, diffusion, temperature effects, and other slow phenomena.

An Application Example: Sand

As a practical application example, we demonstrate the characterization of sand. To this end, cavity 3 was uniformly filled with sand, and the resonance of the TE_{211} mode was evaluated. No attempt has been made to remove the coupling-stub influence by de-embedding.

Table 3 lists the real and imaginary parts of the relative permittivity estimated from the resonance parameters measured with VNA A, B, and C. The agreement is obvious.

To judge if the results are near the “true” parameter values, we compared the results to literature values. Unfortunately, the permittivity of sand depends on many parameters such as the composition of the sand, the grain size, the grain size distribution, wetness, etc. Table 3 cites two literature results which are respectively lower and higher than our permittivities. In any case, our measurements are in agreement with known literature data.

Tab. 3: Calculated permittivity of sand from extracted resonance parameters (cavity 3, TE_{211}).

VNA	ε'_r	ε''_r
A	2.7292	0.021
B	2.7353	0.018
C	2.7530	0.021
Ref. [15]	2.55	0.025
Ref. [16]	3.2	0.026

Conclusion

We have described the architecture and the performance of an application-specific low-cost homodyne VNA. It was demonstrated to be well suited for resonance-parameter determination and, by built-in software routines, for the reconstruction of material parameter distributions in cavities, which may serve to monitor the state of chemical processes.

The system was validated by comparison with commercial VNAs (which are more expensive by at least two orders of magnitude). Numerically, the resonance frequencies and Q-factors determined by the application-specific system typically deviated by less than 0.1 % and 10 % from the results of the commercial VNAs. Material parameters extracted from these resonance parameters agreed just as well.

The typical measurement time required for the reconstruction of a material parameter distribution in a process reactor used as cavity resonator is on the order of a few seconds.

Acknowledgment

Part of this work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 389867475.

References

[1] A. Jess and P. Wasserscheid, *Chemical Technology*. Weinheim: Wiley-VCH, 2013.

[2] T.S. Auckenthaler, “Modelling and Control of Three-Way Catalytic Converters,” Ph.D. dissertation, ETH Zurich, Zurich, 2005.

[3] J. R. Ferraro, K. Nakamoto, and C. W. Brown, *Introductory Raman Spectroscopy.* 2nd Edition, Boston, MA, USA: Academic Press, 2003.

[4] M. Pastorino, *Microwave Imaging*. Hoboken, NJ, USA: Wiley, 2010.

[5] S. Schödel, R. Moos, M. Votsmeier, and G. Fischerauer, „SI-engine control with microwave-assisted direct observation of oxygen storage level in three-way catalysts“, *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 6, pp. 2346–2353, Nov. 2014. DOI: 10.1109/TCST.2014.2305576.

[6] R. Peter and G. Fischerauer, “Measurement of Axially Inhomogeneous Permittivity Distributions in Resonant Microwave Cavities,” *IEEE Trans. Microw. Theory Techn.*, *in print*. DOI: 10.1109/TMTT.2019.2910177.

[7] R. Peter and G. Fischerauer, “De-Embedding Method for Strongly Coupled Cavities,” *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 4, pp. 2025–2033, Apr. 2018. DOI: 10.1109/TMTT.2018.2791934.

[8] L. F. Chen, C. K. Ong, C. P. Neo, V. V. Varadan, and V. K. Varadan, *Microwave Electronics: Measurement and Materials Characterization*. Chichester, U.K.: Wiley, 2004.

[9] I. Motroniuk, R. Królak, R. Stöber, and G. Fischerauer, “Wireless Communication-Based State Estimation of Automotive Aftertreatment Systems,” *Measurement*, vol. 106, pp. 245–250, Aug. 2017. DOI: 10.1016/j.measurement.2016.08.004.

[10] R. Królak and G. Fischerauer, “Scalar Reflectometry as a Means of Catalyst Monitoring Suitable for Field Applications,” *Measurement*, vol. 137, pp. 278–286, Apr. 2019. DOI: 10.1016/j.measurement.2019.01.030.

[11] D. Kajfez, “Q-Factor Measurement with Network Analyzer”, *IEEE Trans. Microw. Theory Techn.*, vol. 32, no. 7, pp. 666–670, Jul. 1984. DOI: 10.1109/TMTT.1984.1132751.

[12] N. B. Carvalho and D. Schreurs, *Microwave and Wireless Measurement Techniques*. New York, NY, USA: Cambridge Univ. Press, 2013.

[13] M. B. Steer, *Microwave and RF Design: A Systems Approach.* Herndon, VA: SciTech Publishing, 2010.

[14] S. A. Dyer (Ed.), *Survey of Instrumentation and Measurement*. New York, NY, USA: Wiley-Interscience, 2001.

[15] R. F. Harrington, *Time-Harmonic Electromagnetic Fields*. New York, NY, USA: IEEE Press, 2001.

[16] H. M. Al-Rizzo and H.T Al-Hafid, “Measurement of the Complex Dielectric Constant of Sand and Dust Particles at 11 GHz”, *IEEE Trans. Instrum. Meas.*, vol. 37, no. 1, pp. 110–113, Mar. 1988. DOI: 10.1109/19.2677.

DOI 10.5162/sensoren2019/6.1.3

Chancen und Herausforderungen von Kooperationen zwischen Forschungseinrichtungen und bürgerwissenschaftlichen Initiativen am Beispiel des Forschungsprojekts „CitizenSensor“

Urban Kaiser[1], Katrin Tina Möbius[2], Dr. Tina Klages[3]
[1]Fraunhofer-Zentrum für Internationales Management und Wissensökonomie IMW, Neumarkt 9-19, 04109 Leipzig, Deutschland
[2]Fraunhofer-Einrichtung für Mikrosysteme und Festkörper-Technologien EMFT Hansastr. 27d, 80686 München, Deutschland
[3] Fraunhofer-Informationszentrum Raum und Bau IRB Nobelstraße 12, 70569 Stuttgart, Deutschland

Zusammenfassung

Aktuelle Erkenntnisse der Innovationsforschung zeigen, dass eine offene, lernorientierte und kooperative Innovationskultur in Zukunft an Bedeutung gewinnen wird. Dahinter steht auch das Bestreben, nicht nur für die Gesellschaft, sondern auch mit der Gesellschaft zu forschen und damit Forschungs- und Innovationsprozesse weiter zu demokratisieren. Die Einbindung von Bürgerinnen und Bürgern in den Forschungs- und Innovationsprozess wird daher auch für etablierte Forschungsorganisationen wie die Fraunhofer-Gesellschaft wichtiger werden. Der vorliegende Beitrag ordnet zunächst den Begriff und die Stellung von Citizen Science in den größeren Kontext Open Science ein, bevor ein Überblick über aktuelle Kooperationsformen der Fraunhofer-Gesellschaft mit bürgerwissenschaftlichen Initiativen folgt.
Hierbei wird insbesondere das aktuelle BMBF-Förderprojekt „CitizenSensor – Umweltanalytik für jedermann“ näher vorgestellt. In diesem Projekt arbeitet die Fraunhofer-Einrichtung für Mikrosysteme und Festkörper-Technologien EMFT zusammen mit BürgerwissenschaftlerInnen des FabLab München e.V. an einem einfachen Messverfahren für Nitrat. Hobbygärtnerinnen und -gärtner sollen damit schnell und unkompliziert den Düngestatus in ihren Beeten oder die Wasserqualität im Gartenteich messen können. Valide Messungen dieser Art konnten bisher nur in Fachlaboren vorgenommen werden. Das Team arbeitet gemeinsam daran, ein elektrochemisches Messverfahren für Nitrat im Boden aufzubauen und zu testen. Im Laufe des Jahres 2019 sollen interessierte Hobbygärtnerinnen, Landwirte oder Urban Gardening Initiativen das neue Messverfahren testen und dem Entwicklerteam ihr Feedback geben.
Neben den vielseitigen technischen Herausforderungen präsentiert der Beitrag Erkenntnisse über Chancen und Herausforderungen einer heute noch eher außergewöhnlichen Zusammenarbeit, die zum einen auf Ergebnissen von semi-strukturierten Interviews mit Erfahrungsträgern der Fraunhofer-Gesellschaft und bürgerwissenschaftlichen Initiativen und zum anderen auf der intensiven Begleitung des Projekts „CitizenSensor“ beruhen.

Keywords: Elektrochemische Messverfahren, Nitrat, Urban Gardening, Verwertung, Citizen Science

Begriffliche und historische Einordnung von Open Science und Citizen Science

Die Digitalisierung als Megatrend verändert nicht nur die Wirtschaft, sondern hat einen ebenso weitreichenden Effekt auf die Wissenschaft, der direkt zu dem Konzept Open Science führt. Open Science beschreibt einen Paradigmenwechsel von geschlossener zu offener Forschung. Diese Entwicklung wurde innerhalb der letzten Jahre von weiteren sozioökonomischen Trends begleitet, die sich mit den Themen Öffnung, Sharing und Nachhaltigkeit befassen. Letztlich hat Open Science auch im forschungspolitischen Kontext eine steigende Bedeutung zu verzeichnen. Auf europäischer Ebene ist Open Science eng verknüpft mit dem Konzept der Responsable Research and Innovation (RRI), die zentraler Bestandteil der Forschungsprogramme ist [1].
Open Science bezeichnet verschiedene Strategien und Verfahren, die die Möglichkeiten der Digitalisierung nutzen, um wissenschaftliche Prozesse und Ergebnisse zu öffnen. Diese Öffnung bietet die Möglichkeit, Forschungsergebnisse (Daten, Software und Publikationen) interdisziplinär nachnutzbar zu machen und schafft Nachvollziehbarkeit und Transparenz im Forschungsprozess. Es lassen sich grundsätzlich zwei Dimensionen von Open Science unterscheiden: Der Forschungsprozess an sich und

dessen Ergebnisse. In Bezug auf die Öffnung von Forschungsergebnissen sind drei Formen zu nennen: Open Access (kostenfreier Zugang zu wissenschaftlichen Veröffentlichungen aller Art – meist vereinfacht Publikationen genannt), Open Data (offener Zugang zu Forschungsdaten – hier sind im engeren Sinne meist die den veröffentlichten Publikationen zugrundeliegenden konkreten Datensätze gemeint) und Open Source (offener Zugang zu Software, die im Forschungsprozess entsteht und genutzt wird). Im Kontext der möglichen Öffnung des Forschungsprozesses sind zu nennen: Open Peer Review (offene Qualitätsbegutachtung von Publikationen), Open Notebooks (offene Laborbücher), Open Methodology (offene Diskussion von Forschungsmethoden vor Projektstart), Open Educational Resources (offen zugängliche Lehrmaterialien) und schließlich Citizen Science (Integration von Bürgerwissenschaftlerinnen und Bürgerwissenschaftlern in Forschungsprozesse).

Das grundlegende Konzept von Citizen Science ist allerdings keine Entwicklung der letzten Jahre. Interessierte und motivierte Bürgerinnen und Bürger haben bereits vor Jahrhunderten signifikant zu wissenschaftlichen Erkenntnissen beigetragen, beispielsweise in der Botanik oder der Astronomie [2]. Auch findet bereits seit dem Jahr 1900 jährlich eine von der *National Audubon Society* in Zusammenarbeit mit Bürgerinnen und Bürgern organisierte Vogelzählaktion statt. Jedoch hat das Aufkommen von digitalen Technologien das Spektrum und das Potenzial von bürgerwissenschaftlichen Aktivitäten im aktuellen Jahrzehnt deutlich erweitert, auch ortsunabhängig an Forschungsprozessen aktiv mitzuwirken.

Zugleich ist der aktuelle wissenschaftliche Diskurs zum Themenfeld Citizen Science noch stark von unterschiedlichen disziplinären Zugängen, Ansätzen und Begrifflichkeiten geprägt. Citizen Science wird folglich als *umbrella term*, also als Sammelbegriff für eine große Bandbreite von Aktivitäten verwendet [3]. Verschiedene Initiativen haben in den vergangenen Jahren gemeinsame Arbeitsdefinitionen entwickelt. Während *die League of European Research Universities (LERU)* – ein Zusammenschluss von 23 forschungsstarken europäischen Universitäten – Citizen Science in einem Diskussionspapier recht allgemein als „Aktive Beteiligung von Bürgerinnen und Bürgern in der Forschung“ definiert [3], verwenden die Autoren des Grünbuchs „Citizen Science Strategie 2020 für Deutschland“ eine weitergehende Arbeitsdefinition. Zentrale Eckpunkte hierbei sind: 1) Die Beteiligung von Bürgerinnen und Bürgern erfolgt in verschiedenen Phasen des Forschungsprozesses -von der Entwicklung von Fragestellungen über die Sammlung und wissenschaftliche Auswertung von Daten bis hin zur Kommunikation von Ergebnissen, 2) Die Zusammenarbeit zwischen Forschungseinrichtungen und institutionell ungebundenen Personen kann sich sehr unterschiedlich gestalten, findet aber immer auf Augenhöhe statt, in dem hauptamtliche und ehrenamtliche Expertinnen und Experten voneinander und miteinander lernen können, 3) Gemeinsames Ziel aller Citizen Science-Projekte ist ein Erkenntnisgewinn für die Wissenschaft und möglichst auch für Praxis und Politik [4].

Diese zwei Arbeitsdefinitionen lassen verschiedene Lesarten von Citizen Science vermuten: Während die LERU-Definition eher einem „Top-Down-Ansatz“ folgt, in dem interessierte Laien die Forschenden bei ihrer Arbeit unterstützen, folgt die Arbeitsdefinition der Autoren des Grünbuchs einem Idealbild der „Demokratisierung von Wissenschaft“ im Sinne Irwins [5][2]. Dieser Ansatz wurde im nachfolgend beschriebenen Forschungsprojekt zwischen der Fraunhofer-Gesellschaft und dem FabLab München gelebt.

Citizen Science in der wissenschaftlichen Praxis

Vor dem Hintergrund der zunehmenden Bedeutung von Citizen Science stellt sich für Forschungsorganisationen die Frage, wie die aktive Beteiligung von Bürgerinnen und Bürgern in die verschiedenen Phasen des Forschungs- und Innovationsprozesses im Sinne der Arbeitsdefinition des oben genannten Grünbuchs organisiert werden kann. Dieser Herausforderung stellt sich auch die Fraunhofer-Gesellschaft.

Dieser Beitrag präsentiert am Beispiel eines vom Bundesministerium für Bildung und Forschung BMBF geförderten gemeinsamen Forschungsprojektes der Fraunhofer-Gesellschaft mit dem FabLab München e.V. Erkenntnisse sowohl über die vielseitigen technischen und nichttechnischen Herausforderungen in der wissenschaftlichen Praxis als auch die damit verbundenen Chancen. Das Team aus hauptamtlichen und ehrenamtlichen Expertinnen und Experten arbeitet im Projekt „CitizenSensor – Umweltanalytik für jedermann“ (FKZ BF1711A-C) gemeinsam daran, ein elektrochemisches Messverfahren für Nitrat im Boden aufzubauen und zu testen.

Zwar sind elektrochemische Messmethoden nicht ganz trivial, sie haben jedoch die Vorteile, einen direkten digitalen Wert der Nitratkonzentration im Boden zu liefern und nur sehr wenig Strom zu benötigen. Damit eignet sich das Messverfahren prinzipiell gut für Vor-Ort-Messungen, die tatsächlich valide Messwerte produzieren, was bisher so nur in Fachlaboren möglich war. Hobbygärtnerinnen und -gärtner

sollen damit in Zukunft schnell und unkompliziert den Düngestatus ihrer Beete oder die Wasserqualität ihres Gartenteichs messen können.

Aufbau und Funktionsweise des Citizen-Sensors

Das Messkit für Nitrat des Projekts Citizen-Sensor basiert auf einem Aufsteckmodul für den Raspberry Pi („Hat"), der mit seiner Firmware die Signale chemischer Sensoren aufnehmen kann. Herzstück dieses Raspberry Pi-Hats ist ein Meter-on-a-chip Mikrocontroller mit hochauflösendem 16-bit Analog-Digital-Converter (ADC) und Frequenzgenerator bis 75 kHz (Analog Devices ADuCM350).

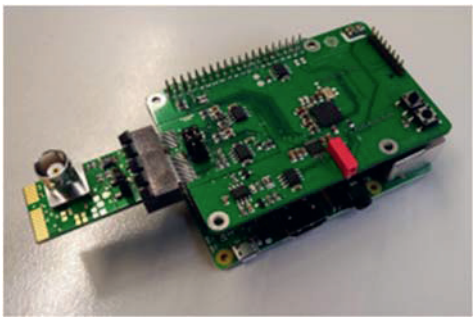


Abb. 1: Der Raspberry Pi-Aufsatz. Foto: Fraunhofer EMFT.

Ein eigenentwickelter Transimpedanzverstärker (Elektrometerverstärker mit Ladungspumpe) erlaubt auch hochohmige Spannungssignale von pH- und ionenselektiven Elektroden aufzunehmen.

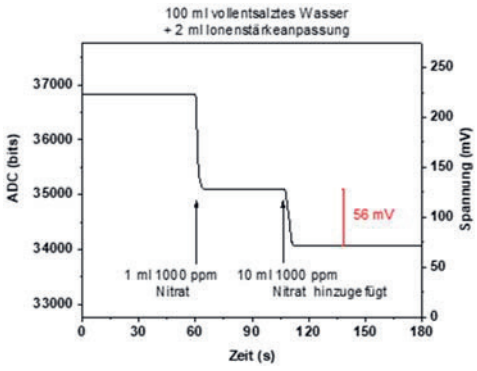


Abb. 2: Die Kurve zeigt den Verlauf einer Messung, in der zuerst 1 ml und anschließend 10 ml einer Nitratlösung zugefügt wurden. Die 56 mV Spannungsänderung, die sich aufgrund der 10-fachen Konzentrationsänderung ergeben (Nernstgleichung), werden von der Elektronik sehr schnell aufgezeichnet.

Dieser Hat erlaubt, angesteuert über eine serielle Schnittstelle, alle für die chemische Sensorik nötigen Messmethoden durchzuführen: Potentiometrie, Amperometrie, Zyklovoltammetrie und Impedanzspektroskopie. Ein Umweltsensor (Bosch BME680) auf dem Hat misst die Laborbedingungen (Luftdruck, Temperatur und Luftfeuchte) während einer Messung automatisch. Weiteres zentrales Element für die chemische Analytik ist ein am Raspberry Pi angeschlossenes digitales Wägemodul (Sparkfun Loadcell mit HX711 AD-Wandler). Damit ist die elektronische Basis für eine valide Analytik vorhanden.

Zur Detektion unterschiedlicher Nitratkonzentrationen in wässriger Lösung wird eine ionenselektive Elektrode (Halbzelle) mit austauschbaren Ionophor-Membran für Nitrat (NO3) zusammen mit einer AgCl-Referenzelektrode (World Precision Instruments DRIREF-2SH) angeschlossen, deren Detektionsgrenze derzeit bei etwa 4 mg/l liegt. Alternativ kann mit einer pH-Elektrode am selben Anschluss auch der pH-Wert der Flüssigkeit aufgenommen werden.

Da nahezu alle kommerziellen ionensensitiven Elektroden (Halbzellen bzw. Kombielektroden) zum Hat kompatibel sind, ist mit geringem Aufwand die Umrüstung auf andere Ionen möglich. Wesentliches Element der chemischen Messung ist die regelmäßige Kalibrierung der Sensoren und die genaue Einwaage aller verwendeten Stoffe. Die integrierte Waage erlaubt die einfache Herstellung der für die Kalibrierung notwendigen Lösungen von Nitratsalz in destilliertem Wasser. Mit Hilfe des integrierten Temperatursensors oder eines zusätzlichen Flüssigkeitsthermometers kann der Einfluss der Temperatur auf die Messergebnisse kompensiert werden. Die integrierte Datenbank zur Rückverfolgung und Qualitätskontrolle für alle Messungen und Kalibrierungen ist der letzte Baustein für eine valide Messung.

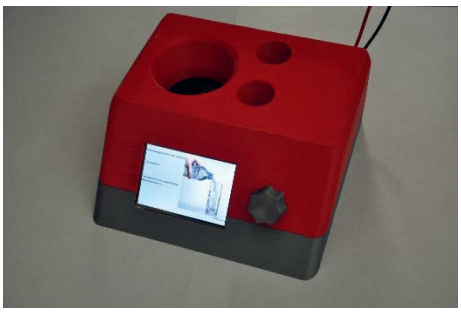


Abb. 3: Gehäuse-Prototyp mit integrierter Waage, Display und Raspberry Pi. Foto: Fraunhofer EMFT.

Als Benutzerinterface dient ein Farbbildschirm. Ein Drehencoder und eine zusätzliche Tastatur

ermöglichen eine schnelle Bedienung der Software, die den ganzen analytischen Prozess steuert. Der Raspberry Pi kann über den Bildschirm bei Bedarf auch kleine Erklärfilme abspielen, über einen Internetanschluss können zusätzliche Informationen und Vergleichswerte bereitgestellt werden.
Sowohl bei der Hard- wie auch Software wurde von Anfang an Wert auf große Benutzerfreundlichkeit gelegt. Zu diesem Zweck wurde ein Gehäuse entworfen, das sowohl die Waage, den Raspberry Pi und das Hat für die elektrochemischen Messungen integriert, als auch ein Display und einen Drehgeber zur Steuerung. Das Gehäuse aus dem 3D-Drucker ist spezifisch auf die Erfordernisse der Messung zugeschnitten und kann unkompliziert an neue Methoden angepasst werden.
Die elektrochemische Messung des Nitratgehalts ist mit diesem energiesparenden Device vor Ort besonders einfach möglich, da die Ergebnisse bereits in digitaler Form vorliegen, direkt angezeigt werden und digital verbreitet werden können. Der Messablauf wird nicht von einem Automaten durchgeführt. Stattdessen legen die Nutzenden selbst Hand an, was diese Methode insbesondere für den Privatgebrauch sowie kleinere Betriebe prädestiniert.
Nach der Entnahme der Bodenprobe, die im Projekt von der Landesanstalt für Landwirtschaft Bayern (LfL) angeleitet wurde, wird die Probe zunächst kräftig durchmischt und auf die korrekte Feuchte überprüft. Damit vermeidet man Fehler, die die Bodenwerte nicht ausreichend gut wiedergeben. Anschließend wird die Probe zerkleinert und mit dem wässrigen Lösungsmittel vermischt. Während sich das Nitrat langsam im Lösungsmittel löst, kann die ionensensitive Elektrode kalibriert werden, um Fehler durch schlechte Kalibrierungen zu vermeiden. Dazu wird die Kalibrierlösung mit Hilfe der Waage in einigen Stufen verdünnt. Da auch gleichzeitig die Temperatur gemessen wird, vermeidet man Verfälschungen durch den Einfluss der Probentemperatur. In einem letzten Schritt wird dann die Probenlösung abfiltriert und durch das Eintauchen der Elektrode der Nitratgehalt bestimmt. Wird die Probe ständig gerührt, stellt sich der Gleichgewichtswert in der von der Elektrode abgegebenen Spannung schnell ein, die Nitratwerte können sofort vom Display abgelesen werden.
Auf längere Sicht sollen die Projektergebnisse auch die Grundlage für die Entwicklung elektrochemischer Nitratsensoren auf Basis von Transducern dienen. Ziel ist es, eine ionenselektive Membran sowie einen geeigneten Transducer für eine Do-It-Yourself (DIY) ionenselektive Elektrode zu finden, um einen selektiven, sensitiven, wiederverwendbaren und kostengünstigen High-Speed-Sensor zu entwickeln.

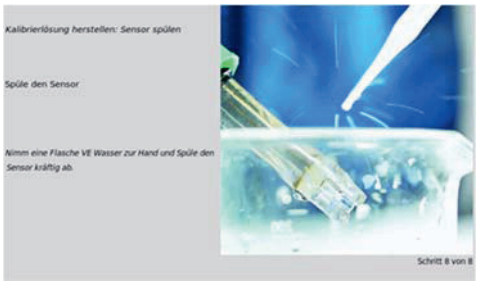


Abb. 4: Die Software führt den Benutzer Schritt für Schritt durch den Prozess einer elektrochemischen Messung. Foto: Fraunhofer EMFT.

Was das Messkit noch besonders macht, ist sein ganzheitlicher Ansatz: Im Gegensatz zu bisher am Markt erhältlichem Messequipment und den in der Literatur beschriebenen Schnelltests [7], bietet das integrierte System des CitizenSensor-Teams eine detaillierte Schritt-für-Schritt-Anleitung für den analytischen Prozess. Diese Software führt in Einzelschritten von der korrekten Probenahme bis hin zur Dokumentation der Messergebnisse, erläutert die einzelnen Schritte bei Bedarf, führt die elektrischen Messungen und weitere Messungen zur Kompensation von Störgrößen durch, prüft die einzelnen Schritte auf Plausibilität und dokumentiert den analytischen Prozess.
Eine eigens entwickelte Python-Software erstellt für einen ausgewählten Prozess (z.B. die Herstellung der Kalibrierlösungen) die nötigen Einzelschritte, zeigt diese auf dem Display an und erledigt für die Benutzerinnen und Benutzer sämtliche Rechenaufgaben und Korrekturen, falls diesen eventuell ein Bedienungsfehler unterläuft. Letztlich ist diese Methodik der Schlüssel zu validen, qualitätsgesicherten Messungen, die auch einer externen Überprüfung standhalten können. Letztlich ist diese Methodik der Schlüssel zu validen, qualitätsgesicherten Messungen, die auch einer externen Überprüfung standhalten können.
Die Thematik der Messvalidität von der Probennahme bis hin zur Ermittlung der Endergebnisse wird in Kooperation mit der Abteilung Qualitätssicherung und Untersuchungswesen der Bayerischen Landesanstalt für Landwirtschaft angegangen. Als Ergebnis eines gemeinsamen Workshops wurden viele Anregungen für das Messkit entwickelt, z.B. die Aufnahme der Wägezelle und die Methodik der Probennahme, die Basis für die Vergleichbarkeit von Messergebnissen ist.
Der einfache DIY-Analyseprozess mit seiner detaillierten Schritt-für-Schritt-Anleitung, dem

synergetischen Zusammenwirken verschiedener elektrischer Komponenten, der integrierten Dokumentation in einer Datenbank und der Plausibilitätskontrolle ist im Vergleich zu kommerziellen Einzelmessgeräten oder automatischen Laboreinheiten ein entscheidender Mehrwert für die Nutzer, um auch als Laien aussagekräftige und korrekte Messungen durchführen zu können.
Bis zum Sommer 2019 hofft das Projektteam, der Gardening-Community ein einfaches, aber funktionales Hilfsmittel für die Nitratmessung zur Verfügung stellen zu können. Gleichzeitig kann und soll der entwickelte Raspberry Pi-Aufsatz aber auch Ausgangspunkt für verschiedenste neue Maker-Projekte sein. Ob es darum geht, ein pH-Meter zu bauen, weitere chemische Analyten zu messen oder eine elektrochemische Gasmessstation zu errichten – der Raspberry Pi-Aufsatz liefert die nötige Hardware dafür. Die bewusste Fokussierung auf Bauteile und Materialien, die allgemein zugänglich sind und eine „hands-on"-Funktionalität aufweisen, sollen dabei die Hürden geringhalten, um innovative neue Ideen mit möglichst unkomplizierten Mitteln zu realisieren.

Erwartungen und Motivation der Beteiligten

Die Erwartungen und die Motivation der Beteiligten an diese neue Form der Kooperation wurden zu Beginn des Projektes in zwei getrennten Workshops mit den Fraunhofer-Wissenschaftlerinnen und Wissenschaftlern und dem Projektteam des FabLab München durchgeführt.
Hierzu wurde jeweils mithilfe eines moderierten Kreativprozesses, intensiv an der Frage gearbeitet, was für den Einzelnen gute Zusammenarbeit ausmacht und was dabei mit Blick auf die eigene Motivation besonders wichtig ist. Die Teilnehmenden stellten dazu mit speziellen Bausteinen und weiteren Elementen wie Figuren oder Tieren ihre Vorstellung einer guten Zusammenarbeit dar. Zu den einzelnen Modellen wurde von jedem Beteiligten in einer kurzen Geschichte die Bedeutung der einzelnen Bauteile erläutert, so dass über das aktive Bauen und das anschließende Sprechen darüber viele Aspekte von Kooperation und der eigenen Motivation thematisiert werden konnten. Die Methodik wurde in beiden Teilgruppen sehr positiv aufgenommen.
Im nächsten Schritt wurde über die Dimensionen einer guten Kooperation in Bezug auf das Projekt CitizenSensor diskutiert und wesentlichen Aspekte und Themen durch Mindmapping festgehalten. Übereinstimmend wichtig war allen Beteiligten die Definition eines gemeinsamen Projektziels, wofür aber zunächst eine gemeinsame Sprache bzw. ein gemeinsames Verständnis notwendig sei. Ein hieraus entwickelter Leitgedanke für das gesamte Projekt war, dass es für eine gute Kooperation wichtig ist (an-)zu erkennen, dass im Projektteam eine große Diversität vorhanden ist und dies gleichzeitig eine Chance und Herausforderung darstellt und in jedem Fall für jeden Beteiligten einen Perspektivwechsel notwendig macht.

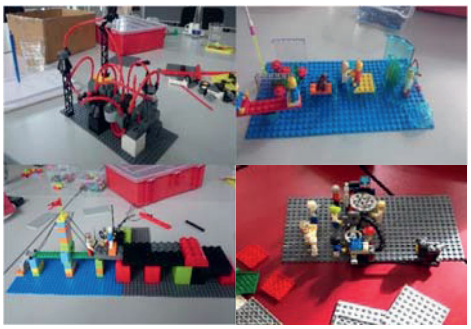


Abb. 5: Beispiel-Modelle der guten Zusammenarbeit der Workshop-Teilnehmenden. Fotos: Fraunhofer IMW.

10 Leitsätze für die Kooperation

Aufbauend auf diesem Leitgedanken sowie unter Berücksichtigung der Arbeitsdefinition des Grünbuchs wurden gemeinsam folgende 10 Leitsätze formuliert, die den interessierten, wertschätzenden Umgang auf Augenhöhe in den Mittelpunkt der Kooperation stellten: 1) Wir forschen und entwickeln gemeinsam an innovativen Lösungen mit und für die Gesellschaft; 2) Dabei respektieren wir die Besonderheiten der Beteiligten und suchen bei unterschiedlichen Interessenslagen nach Lösungen im Konsens. Auch gehen wir bewusst mit unseren unterschiedlichen Wissensständen und Vorkenntnissen um. Bei großen Unterschieden stellen wir uns auf unser Gegenüber ein, und holen uns dort ab wo wir stehen. Verständnis-Schwierigkeiten sprechen wir offen an; 3) Das Erreichen der Ziele der Zusammenarbeit ist nur mit Hilfe aller Beteiligten möglich. Dabei streben wir an, dass für alle Beteiligten ein Mehrwert entsteht; 4) Bei der Generierung und Weitergabe von Wissen streben wir Lösungen im Sinne des Open-Science-Gedankens an; 5) Unser zentrales Kommunikations- und Dokumentations-Tool ist ein Wiki das den E-Mail-Verkehr weitgehend ersetzt. Hierzu haben alle Projektbeteiligten Zugang. Wir legen hier alle wesentlichen Informationen als Arbeitsversionen in transparenter Weise ab und halten diese jederzeit aktuell; 6) Alle Projektbeteiligten werden dazu ermutigt, sich aktiv an der Erstellung von Arbeitsdokumenten zu beteiligen; 7) Darüber hinaus sind wir uns einig, dass persönliche Treffen der Beteiligten ein wertvolles Instrument sind, um aktuelle Herausforderungen vertrauensvoll zu diskutieren und zu guten Lösungen zu kommen; 8)

Zwischen den Beteiligten findet mindestens monatlich ein Jour-Fixe statt, wenn möglich wöchentlich. Die wesentlichen Ergebnisse werden im Wiki dargestellt. Dazu gehört ebenso sich aktiv um Information zu bemühen und selbst als Informationsgeber daran zu denken, dass ich für andere Information bereitstelle; 9) In einem Glossar werden zentrale Begriffe definiert, damit alle „die gleiche Sprache“ sprechen; 10) In einem Arbeitsplan werden die einzelnen Aufgaben und Arbeitsschritte für alle transparent festgehalten. Dieser ist als lebendes Dokument zu verstehen, der mindestens einmal monatlich den aktuellen Entwicklungen angepasst wird.
Im Folgenden wird nun näher auf die drei oben erwähnten zentralen Herausforderungen in der Projektkooperation eingegangen.

Kommunikation

Transparenz und ein guter Informationsfluss über entsprechende Prozesse und Schnittstellen sind von essenzieller Bedeutung, um Missverständnisse und Unmut im Projektteam zu vermeiden und anhaltendes Vertrauen zwischen allen Beteiligten aufzubauen. Im Projekt CitizenSensor erwies sich vor allem die Einrichtung eines gemeinsamen Wikis als zentrale Kommunikationsplattform als äußert hilfreich, um diesem Idealbild in der Praxis nahezukommen. Ein großer Vorteil für alle Projektbeteiligte ist insbesondere, dass durch das Forum-Format ersichtlich wird, wer wann was zu welchem Thema geschrieben hat. Für die konkrete Zusammenarbeit bedeutet dies aber auch, dass sich dieser Tatsache alle Beteiligten immer wieder vergegenwärtigt werden müssen, also die einen in der „Bringschuld“ sind und erklären müssen, die anderen in der „Holschuld“ sind und bei Unklarheiten aktiv nachfragen müssen.

Ressourcen

Arbeitskultur und das jeweilige Hintergrundwissen beider Partner sind sehr unterschiedlich. Während die Fraunhofer-Forschenden in ihrer regulären, ergo bezahlten Arbeitszeit am Projekt mitarbeiten, engagieren sich die FabLab-Mitglieder in der Freizeit. Arbeitstreffen im Projekt finden daher oft am Freitag bzw. am Wochenende oder abends statt. Um die technologisch anspruchsvollen Arbeiten inhaltlich und zeitlich effektiv zu koordinieren, floss viel Zeit und Mühe in einen kleinteilig strukturierten Projektplan. In Form einer Mindmap unterstützt er das heterogene Team beim Schnittstellenmanagement, bei der Ablaufplanung und bei der Fortschrittskontrolle. Allerdings führte diese Vorgehensweise zwischenzeitlich durchaus zu kontroversen Debatten innerhalb des Projektteams. Dabei verlief die Konfliktlinie nicht an der Schnittstelle zwischen Fraunhofer-Forschenden und FabLab-Mitgliedern, sondern quer zu diesen. Während einige bezweifelten, dass eine seriöse Zeitplanung überhaupt über einen langen Zeitraum möglich ist, da die FabLab-Mitglieder das Projekt in ihrer Freizeit bestreiten müssen, sahen andere den Vorteil, dass diese Art der Projektplanung in gewisser Weise notwendig sei, damit alle mitziehen können.

Verwertung

Eine generelle Geheimhaltungsverpflichtung der beteiligten Projektmitarbeitenden, wie sie in Industriekooperationen der Fraunhofer-Gesellschaft in der Regel angewandt wird, ist im bürgerwissenschaftlichen Kontext kaum umzusetzen ist. Dies ist zum einen operativ begründet, da das FabLab auch während der Projektarbeitstage seinen anderen Mitgliedern offensteht. Zum anderen widerspricht so ein Vorgehen auch der Kernphilosophie der Maker-Bewegung „Make, Work, Share“: Verfahren und Prozesse, die am FabLab entstehen, sollen prinzipiell geteilt werden. Eine entsprechende Haftungsklausel erscheint hier als nicht möglich. Eine „Klassifizierung“ in Projektmitarbeitende, die Zugang zu vertraulichen Informationen erhalten, aber eine Geheimhaltungserklärung unterzeichnen müssen und anderen Projektbeteiligten könnte wiederum als Ausgrenzung verstanden werden und sich demotivierend auswirken. Die Frage des Umgangs mit den Projektergebnissen erwies sich im Projekt CitizenSensor als größte zu überbrückende Differenz zwischen den Beteiligten. Während der Open-Source-Gedanke in der Maker-Szene fest verankert ist, ist das Fraunhofer-Modell dagegen im Kern darauf ausgelegt, neues Wissen über Patente oder Lizenzen zu verwerten, um so die eigene Forschungstätigkeit zu refinanzieren.
Um für beide Seiten einen Mehrwert zu generieren, arbeiteten die Partner eine differenzierte Lizensierung aus, die viele Open-Source-Elemente enthält, aber auch den Weg einer Kommerzialisierung durch Fraunhofer offenhält

Fazit

Für das CitizenSensor Team steht fest: Der Blick über den Tellerrand lohnt sich. Die Kooperation hat beiden Seiten neue Blickwinkel eröffnet. Ein gutes Netzwerk zu Universitäten, Erfahrung in der Akquise und Durchführung von Forschungsprojekten, spezialisiertes Wissen zu komplexen, technologischen Themen und die dafür vorhandene Technologie und Labore sind der Mehrwert, den das Fraunhofer-Team in die gemeinsame Projektarbeit mit einbringen konnte. Das breit gefächerte Know-how der beteiligten FabLab-Mitglieder, eine nutzenorientierte Sichtweise und der Fokus auf praktikable und effiziente Lösungen verliehen der Projektarbeit eine große Dynamik und einen engen Anwendungsbezug.

Danksagung:

Das Projekt CitizenSensor wird unter dem Förderkennzeichen BF1711A-C vom BMBF gefördert. Die Autoren bedanken sich bei dem gesamten Projektteam: M. Steinmaßl, J. Boudaden, E. Karakoc, C. Hochreiter, H.-E. Endres – Fraunhofer EMFT München; A. Kahler, P. Weiß, H. Hottarek, R. Meisenecker, B. Kahler, N. Bobrich-Draxler – Fablab München; N, Pelke, L. Quiring – Fraunhofer IMW Leipzig und dem LfL Bayern, Dr. G. Strauß, G. Henkelmann.

Literaturnachweis

[1] Europäische Kommission: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation, letzter Zugriff am 06.05.2019

[2] R. Bonney et al., *Science* 343:6178, 1436–1437 (2014); doi: 10.1126/science.1251554

[3] C. Arendes, *Heidelberger Jahrbücher Online* Band 2, 19-58 (2017), doi: 10.17885/heiup.hdjbo.2017.0.23691

[4] League of European Research Universities (2016). Citizen science at universities: Trends, guidelines and recommendations. Leuven, Belgium: LERU Office.

[5] A. Bonn et al., Grünbuch Citizen Science Strategie 2020 für Deutschland, o.O. (2016), http://www.buergerschaffenwissen.de/sites/default/files/assets/dokumente/gewiss-gruenbuch_citizen_science_strategie.pdf (letzter Zugriff am 03.05.2019).

[6] A. Irwin, Citizen Science – A study of people, expertise, and sustainable development, London, New York (1995, 2002 (e-book), doi: 10.4324/9780203202395

[7] R. Shaw et al., *Agriculture* 3(3):327-341 (2013); doi:10.3390/agriculture3030327

DOI 10.5162/sensoren2019/6.1.5

Citizen Science für Schüler*innen: Durchführung von Umweltstudien mit Smartphone und mobiler Messtechnik

S. Höfner[1], A. Schütze[1], B. Brück[2], M. Hirth[3], J. Kuhn[3]
[1]Universität des Saarlandes, Lehrstuhl für Messtechnik, Saarbrücken (Deutschland)
[2]Schülerforschungszentrum Saarlouis (Deutschland)
[3]TU Kaiserslautern, AG Didaktik der Physik, Kaiserslautern (Deutschland)

Zusammenfassung

Im Rahmen des DBU Projekts „SUSmobil“ (**S**chüler**U**mwelt**S**tudien mit **mobil**en Endgeräten) wird Schülerinnen und Schülern die Möglichkeit geboten, mit Hilfe von Wissenschaftlern Umweltstudien durchzuführen. Dabei sind sie angehalten, eigene Fragestellungen zu entwickeln. Zur Vorbereitung und Unterstützung können Schülerinnen und Schüler im Alter von 15 bis 18 Jahren zudem Selbstlernkurse (sog. Module) besuchen, in denen sie die Funktionsweise von Halbleiter-Gassensoren kennenlernen, eine Kalibrierung durchführen und Innenraumluftqualität messen. Der Beitrag stellt die entwickelten Module vor und präsentiert verschiedene Beispiele von Schülerumweltstudien, die im letzten Jahr durchgeführt wurden.

Keywords: Gassensorik, Citizen Science, Internet of Things, Umweltstudien, Schülerinnen und Schüler

Einleitung

Citizen Science (CS) hat als gemeinschaftliches Bestreben von Bürger*innen und Bürgern , zusammen mit Wissenschaftlern, neue wissenschaftliche Erkenntnisse zu schaffen, einen zunehmend beliebten Weg der wissenschaftlichen Forschung und Praxis eröffnet [1].
Bei der Planung und Durchführung von CS-Projekten ergeben sich jedoch mehrere Herausforderungen. Zum einen muss es leicht sein, eigene Daten beizutragen, damit die Motivation erhalten bleibt. Zum anderen muss die Qualität der aufgenommenen Daten gesichert werden. Um dies zu erreichen, ist es wichtig, dass Sensorsysteme nicht einfach nur verwendet werden, um Daten zu sammeln, sondern auch zu vermitteln, wie diese funktionieren und wie die aufgenommenen Daten zu interpretieren sind [2]. Je mehr diese Punkte erfüllt sind, umso wahrscheinlicher ist es auch, Bürger transformativ weiterzubilden und so eine nachhaltige Wirkung z.B. auf deren eigenes Umweltverhalten zu erreichen.
Im von der DBU geförderten Projekt „SUSmobil“ werden diese Ansprüche in insgesamt vier Modulen verfolgt. Das Projekt hat das Ziel, das Umweltbewusstsein von Schülerinnen und Schülern im Alter zwischen 12 und 18 Jahren zu stärken. Dies soll durch die Sensibilisierung für umweltschädliche Schadstoffe und vor allem durch die Durchführung eigener Umweltstudien erreicht werden [3].

1. Selbstlernmodule

1.1 Funktionsweise von Halbleiter-Gassensoren

Im ersten Modul wird den Schülerinnen und Schülern zunächst die Funktionsweise von Halbleiter-Gassensoren nähergebracht. Dazu wurde eine Platine entwickelt, die es erlaubt, sowohl die Temperatur des Sensors über Potentiometer zu kontrollieren als auch die Sensorreaktion (die Änderung des elektrischen Widerstands) mit einem Multimeter auszulesen, siehe Abbildung 1.

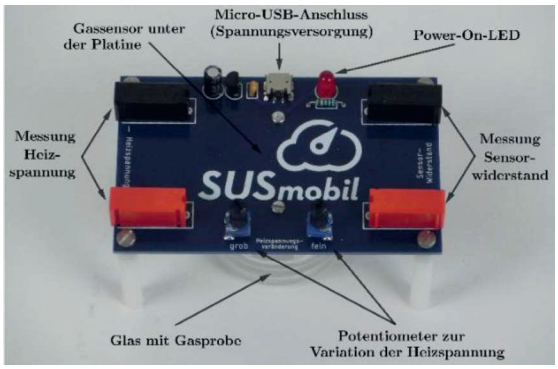


Abb. 1: Versuchsplatine zu Modul 1: Funktionsweise eines Halbleiter-Gassensors.

Verschiedene Flüssigkeitsproben, z.B. Apfelsaft, alkoholfreies und alkoholhaltiges Bier, können in Gläser gefüllt und unter die Platine geschraubt werden. Dabei lernen die Schülerinnen und Schüler, dass der Sensor auf verschiedene Stoffe unterschiedlich reagiert und dass diese Reaktion ebenfalls temperaturabhängig mit. Diese Beobachtungen werden anschließend durch ein schülergerechtes Modell erklärt. Die Erkenntnisse sollen die Grundlage und Motivation zur Nutzung eines temperaturzyklischen Betriebs zur Verbesserung von Sensitivität und Selektivität der Sensoren bilden [4].

1.2 Kalibrierung von Halbleiter-Gassensoren

Im zweiten Modul geht es um die Wichtigkeit einer korrekten Kalibrierung von Gassensoren zur Sicherung der Datenqualität. Die Schülerinnen und Schüler haben die Aufgabe, einen Gassensor, der durch einen Mikrocontroller temperaturzyklisch betrieben wird, auf verschiedene Ethanol-Konzentrationen zu kalibrieren. Dazu erzeugen sie zunächst durch Verdampfen flüssigen Ethanols eine definierte Atmosphäre in einer Gasvorratskammer, siehe Abbildung 2.



Abb. 2: Messkammer mit zu kalibrierenden Sensor (links) und Gasvorrats-kammer zur Erzeugung einer definierten Ethanol-Konzentration (rechts). Diese kann anschließend in einer Messkammer weiter verdünnt und zur Kalibrierung verwendet werden.

Durch Entnahme einer gewissen Menge des Gasgemischs und Überführen in eine Messkammer kann dort die gewünschte Konzentration eingestellt und durch wiederholtes Zudosieren schrittweise erhöht werden.
Für jede Konzentration werden nun Temperaturzyklen als Trainingsdaten aufgenommen und aus diesen Mustern besondere Merkmale extrahiert, siehe Abbildung 3.

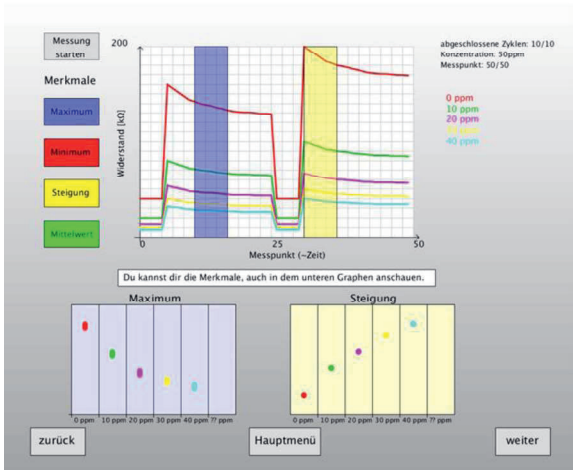


Abb. 3: Software zur Aufnahme von Trainingsdaten, sowie zur Bestimmung besonderer Merkmale der Zyklen bei verschiedenen Gaskonzentrationen.

Schließlich erstellt ein neuronales Netz ein mathematisches Modell, das einen Zusammenhang zwischen den aufgenommenen Merkmalen und der jeweiligen Konzentration bestimmt. Das Modell wird abschließend zur Validierung mit einer unbekannten Ethanolkonzentration getestet.
Die Vorgehensweise der Kalibrierung wird hierbei schrittweise erklärt, gleichzeitig werden Analogien zur Kalibrierung bekannter Sensoren wie z. B. einem Thermometer aufgezeigt.

1.3 Auswirkung verschiedener Gase auf den Menschen und die Umwelt

Neben dem Verständnis der Funktionsweise von Gassensoren und deren Kalibrierung ist es wichtig, Wissen über die Gefahren und Auswirkungen schädlicher Stoffe auf den eigenen Körper und die Umwelt zu vermitteln [5]. Im dritten Modul, das sich zurzeit noch in der Entwicklung befindet, wird diese Thematik aufgegriffen.
Analog zu Modul 1 wurde dafür eine weitere Platine entwickelt, die es ermöglicht, die Ausgasungen verschiedener Stoffe wie Plastikspielzeug, Lösemittel, Lacke, Zigarettenstummel und Bodenbeläge zuuntersuchen. Die Messwerte (Gasart und Konzentration) werden direkt auf einem

integrierten Display angezeigt. Der Versuchsaufbau ist in Abbildung 4 dargestellt.



Abb. 4: Platine zur Untersuchung der Ausgasung schädlicher Gase aus verschiedenen Materialien.

Der Fokus liegt hierbei auf der Messung von flüchtigen organischen Verbindungen (engl. VOCs, volatile organic compounds), CO_2 und Kohlenmonoxid, die maßgebloch die Innenraumluftqualität beeinflussen. Gleichzeitig sollen den Schülerinnen und Schülern verschiedene Umweltlabel wie „Blauer Engel“ oder die „Euroblume“ nähergebracht werden. Die Bedeutung der Innenraumluftqulität ist in unserem Alltag besonders hoch, da viele Menschen über 80% der Zeit in Innenräumen verbringen [6].

2. Schülerumweltstudien

Neben den Modulen zur Funktionsweise, Kalibrierung und Messung von Gasen in der Innenraumluft sollen die Schülerinnen und Schüler animiert werden, eigenständig wissenschaftliche Fragestellungen zu entwickeln, um beispielsweise bei „Jugend forscht“ Wettbewerben anzutreten. Unterstützt durch Wissenschaftler werden die Fragestellungen formuliert und bearbeitet.
Durch die Verbindung von formellen Lernanlässen in Schülerlaboren (Module 1 bis 3) und informellen Lernanregungen außerhalb dieser Einrichtungen (Schülerumweltstudien) kann ein Mobile Assisted Seamless Learning initiiert werden, wodurch die Initiierung von „Citizen Science“-Projekten ermöglicht wird [7,8].

2.1 Datenübertragung auf das Smartphone und Webseiten

Neben dem Verständnis der Funktionsweise von Gassensoren und der korrekten Art und Weise der Datenaufnahme ist die Visualisierung der Daten ein wichtiger Aspekt.

Das Smartphone ist aus der heutigen Zeit nicht mehr wegzudenken und bietet sich durch seine Mobilität perfekt an, um Daten zu visualisieren. Mittlerweile gibt es zahlreiche „Internet of Things“ Plattformen, die sich auf die Darstellung von Messdaten auf dem Smartphone oder im Internet spezialisiert haben. Ein Beispiel für eine solche Visualisierung ist in Abbildung 5 unter Verwendung der App ThingView [9] gezeigt.

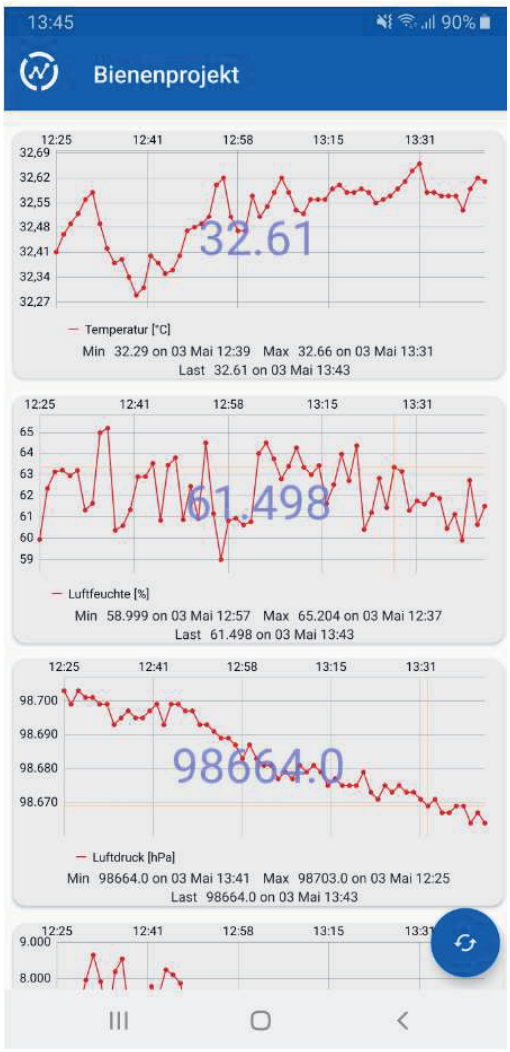


Abb. 5: Visualisierung der Messergebnisse auf der Smartphone App „ThingView“.

Neben der Visualisierung per Smartphone-App gibt es auch die Möglichkeit, alle Messergebnisse verschiedener Umweltstudien zentral auf einer Webseite darzustellen. Dies hat den Vorteil, dass die Ergebnisse für jeden in Echtzeit einsehbar sind, so dass die Projekte eine Multiplikatorwirkung über die direkt beteiligten Schülerinnen und Schüler hinaus erzielen können.
Im Folgenden werden einige der aktuellen Schülerumweltstudien kurz vorgestellt.

2.2 Schadstoffkarte

Im Rahmen des Nachwuchswettbewerbs „Jugend forscht“ wurde eine „Google Maps Schadstoffkarte“ am Schülerforschungszentrum in Saarlouis entwickelt. Mit Hilfe einer 3D-gedruckten Messkammer (siehe Abbildung 6), in der sich verschiedene Gas- und Luftqualitätssensoren sowie ein GPS-Modul zur Lokalisierung befinden, können (unkalibrierte) Luftqualitätsmessungen der Umgebungsluft gemacht werden, um damit relative Schadstoffkonzentrationen zu visualisieren.



Abb. 6: 3D-gedruckte Messkammer mit integrierten Sensoren und GPS-Empfänger.

Die Messdaten werden zusammen mit GPS-Daten an einen Server im Internet gesendet und von einem IoT-Dienst gespeichert. Dieser ermöglicht es, die Daten als Heatmap über Dienste wie „OpenStreetMap“ oder „Google Maps“, wie in Abbildung 7 gezeigt, darzustellen. Die Schüler erreichten mit ihrer Idee den dritten Platz beim Wettbewerb im Saarland.

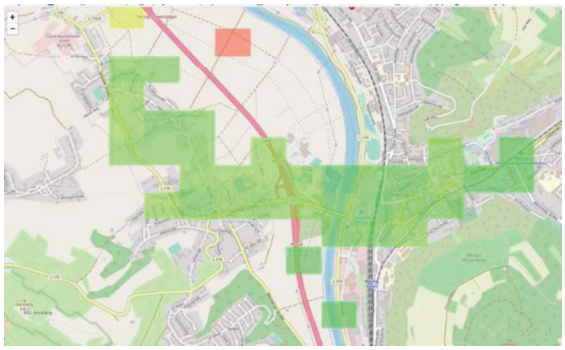


Abb. 7: Darstellung von relativen Luftqualitätsdaten mittels einer Heatmap.

2.3 Untersuchung von Bienenstöcken

In Zusammenarbeit mit Dr. Martin Wörner, einem Mitglied im Landesverband saarländischer Imker, wurde ein „Jugend forscht“-Projekt zur Untersuchung von Bienenstöcken auf Gase, Temperatur, Luftfeuchte und Luftdruck entwickelt. Schülerinnen und Schülern wird dabei die Möglichkeit geboten, an einem authentischen und hochaktuellen Kontext wissenschaftlich zu arbeiten, indem Daten aufgenommen, ausgewertet und interpretiert werden müssen [10].
Das für den Einsatz im Bienenstock realisierte Sensorsystem umfasst die folgenden Sensoren:

- BME680 (Temperatur, Luftfeuchte, Luftdruck)
- SCD30 (CO_2)
- SGP30 (Multigassensor für Gesamt-VOC-Konzentration (TVOC) und äquivalente CO_2-Konzentration, eCO_2)
- SPS30 (Feinstaub PM2.5 und PM10)

Zur Kontrolle und Vergleich der Daten wurden zwei Messsysteme realisiert, die parallel laufen. Ein System befindet sich dabei in einem bevölkerten Bienenstock (siehe Abbildung 8), das andere in einem leeren Stock. Ein Auszug der bisherigen Messergebnisse ist in Abbildung 9 zu sehen.



Abb. 8: Einsetzen des Messsystems in einen voll besetzten Bienenstock.

Auffällig ist die periodische Änderung der TVOC bzw. CO_2 Konzentrationen, die jeden Tag um etwa 11 Uhr morgens ein Höchstwerte von etwa 1600 ppb (TVOC) bzw. 14000 ppm (CO_2) erreichen und danach sehr schnell abfallen. Dies ist damit zu erklären, dass die Bienen um diese Uhrzeit den Stock verlassen.

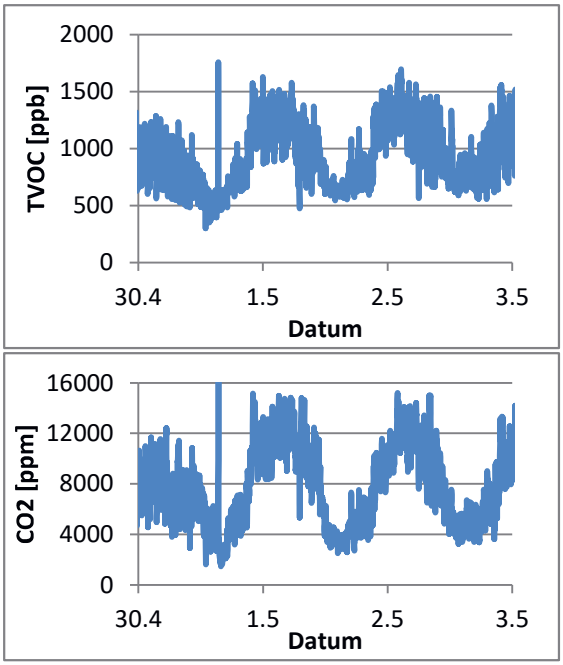


Abb. 9: TVOC-Messwert (oben) und CO_2-Messwert (unten) in einem Bienenstock.

2.4 Auswirkung von Pflanzen auf die Innenraumluftqualität

Ein weiteres Schülerprojekt ist die Untersuchung der Innenraumluftqualität in Abhängigkeit davon, ob sich Pflanzen im Zimmer befinden. Dazu wurden zwei Messboxen konstruiert. In beiden wurden nacheinander verschiedene VOCs verdampft und so eine definierte Gasatmosphäre erzeugt. In eine Messbox wurden die Pflanzen Einblatt und Bogenhanf gestellt, die zweite blieb leer, siehe Abbildung 10.



Abb. 10: Messkammern zur Untersuchung der Auswirkung von Pflanzen auf die Luftqualität.

Anschließend wurde die VOC-Konzentration über einen längeren Zeitraum von 5-7 Stunden gemessen. Es zeigte sich ein deutlicher Abfall der gemessenen TVOC-Konzentration in der Messkammer mit Pflanzen, wie in Abbildung 11 zu sehen ist. Daraus kann die Schlussfolgerung gezogen werden, dass Pflanzen die Luft in Bezug auf die getesteten VOCs, in Abb. 11 Aceton, reinigen können.

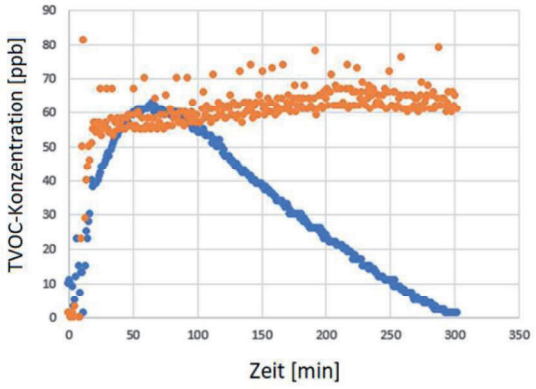


Abb. 11: Verlauf der Aceton-Konzentration in den Messkammern. Orange: ohne Pflanzen. Blau: Mit Pflanzen.

2.5 Feinstaubmessungen

Die Diskussionen um ein Verbot von Pyrotechnik an Silvester aufgrund der erhöhten Feinstaubbelastung werden jedes Jahr mehr, weshalb sich einige Schüler dazu entschlossen haben, selbst eine Feinstaubmessung an Silvester durchzuführen. Um Mitternacht wurde ein PM2.5-Wert von ca. 1800 µg/m³ erfasst. Dies würde dem 72-fachen des täglichen Grenzwerts entsprechen [11], allerdings sind die Sensoren nicht kalibriert und auch nur für Messungen bis max. 70% Luftfeuchtigkeit geeignet. Allerdings ist deutlich erkennbar, dass der Schadstoffpeak weit über den Messwerten vor Mitternacht lag und auch nur wenige Minuten anhielt.

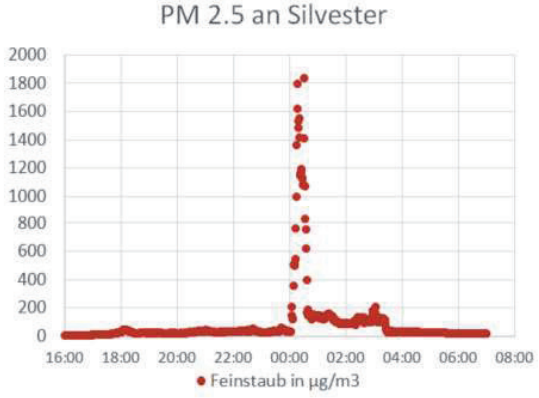


Abb. 12: Unkalibrierte Feinstaub-Messwerte für PM2.5 an Silvester 2018

Ausblick

Ein Ziel des Projekts ist die Nachhaltigkeit, weshalb die verschiedenen Selbstlernkurse bis zum Ablauf der Projektzeit (Ende 2020) anderen Schülerlaboren, Schülerforschungszentren und Schulen zur Verfügung gestellt werden sollen. Zudem werden Fortbildungen für Lehrkräfte entwickelt und angeboten, um mit ihnen gemeinsam Schülerinnen und Schüler an die Nutzung von Gassensoren heranzuführen und für Schülerumweltstudien zu motivieren. Auch eine Übersetzung ins Englische wurde bereits für Modul 1 durchgeführt und ist für die weiteren Module in Planung, um die entwickelten Ansätze auch international anbieten zu können.
Schließlich sollen weitere Schülerinnen und Schüler für eigene Umweltstudien gewonnen werden. Die Ergebnisse dieser Studien sollen in Schulen präsentiert und im Internet zur Verfügung gestellt werden. Auch besteht die Möglichkeit, dass diese im LeLa-Magazin veröffentlicht werden, um über die direkt beteiligten Schülerinnen und Schüler hinaus ein gesteigertes Umweltbewusstsein zu erzielen.

Literaturnachweis

[1] Storksdieck, M; Shirk, JL; Cappadonna, JL; Domroese, M; Göbel, C; Haklay, M; Miller-Rushing, AJ; Roetman, P; Sbrocchi, C; Vohland, K. (2016) Associations for Citizen Science: Regional Knowledge, Global Collaboration. Citizen Science: Theory and Practice**,** 1 (2) , Article 10. 10.5334/cstp.55

[2] Aceves-Bueno, E.; Adeleye, A.; Feraud, M.; Huang, Y.; Tao, M.; Yang, Y.; Anderson, S. (2017). *The Accuracy of Citizen Science Data: A Quantitative Review,* Bulletin of the Ecological Society of America, 98(4), pages 278ff.

[3] Kaiser, F. G., Oerke, B., & Bogner, F. X. (2007). Behavior-based environmental attitude: Development of an instrument for adolescents. *Journal of Environmental Psychology*, *27*(3), 242-251.

[4] Baur, T., Schütze, A., Sauerwald, T. (2015). Optimization of temperature cycled operation of semiconductor gas sensors. tm - Technisches Messen, Band 82, Heft 4, Seiten 187–195, ISSN (Online) 2196-7113, ISSN (Print) 0171-8096, DOI: https://doi.org/10.1515/teme-2014-0007.

[5] Moriske, H. J. & Szewzyk, R. (2008). *Leitfaden für die Innenraumhygiene in Schulgebäuden*. Umweltbundesamt.

[6] SOER 2015 — The European environment — state and outlook 2015, European Environment Agency, 2015, http://www.eea.europa.eu/soer [Stand: 11/2016]

[7] Wong, L. H., & Looi, C. K. (2011). What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. *Computers & Education*, *57*(4), 2364-2381.

[8] Hirth, M., Kuhn, J., Müller, A., Rohs, M., & Klein, P. (2016). iMobilePhysics: Seamless Learning durch Experimente mit Smartphones & Tablets in Physik. *Zeitschrift für Hochschulentwicklung*, *11*(4), 17-37.

[9] ThingView App: Android: https://play.google.com/store/apps/details?id=com.cinetica_tech.thingview&hl=de; iOS: https://itunes.apple.com/de/app/thingview/id1284878579?mt=8

[10] Human, H., Nicolson, S.W. & Dietemann, V. Naturwissenschaften (2006) 93: 397. https://doi.org/10.1007/s00114-006-0117-y

[11] Umweltbundesbamt http://www.umweltbundesamt.at/pm25

DOI 10.5162/sensoren2019/6.2.1

Unterscheidung und Quantifizierung von Geruchsstoffen im ppb-Bereich mit temperaturzyklisch betriebenen MOS-Sensoren

Julian Joppich, Andreas Schütze, Tilman Sauerwald
Lehrstuhl für Messtechnik, Universität des Saarlandes, 66123 Saarbrücken
j.joppich@lmt.uni-saarland.de

Zusammenfassung

Eine Methode zur Messung von zwei geruchsintensiven Gasen in Luft, Schwefelwasserstoff und Dimethylsulfid, wird präsentiert. Die verwendeten Verdünnungen im ppb-Bereich entsprechen Geruchsstoffkonzentrationen von ca. 20-250 GE/m³. Die Messung erfolgt mit einem temperaturzyklisch betriebenen Halbleiter-Gassensor. Mithilfe des DSR-Modells (differential surface reduction) werden Merkmale des transienten Leitwertsignals extrahiert. Die Merkmale werden bezüglich ihrer Temperaturabhängigkeit und der Möglichkeit, die Geruchsstoffkonzentration zu quantifizieren, untersucht. Über die unterschiedliche Temperaturabhängigkeit der Merkmale können die Zielgase voneinander unterschieden werden. Neben einer stoffspezifischen oder gemeinsamen Quantifizierung unter Zuhilfenahme aller Merkmale ist unter den gegebenen Bedingungen auch eine gemeinsame Quantifizierung für beide Gase bei ausschließlicher Nutzung des Merkmals einer bestimmten Temperaturstufe des Temperaturzyklus möglich.

Keywords: Gassensor, Temperaturzyklus, Schwefelverbindungen, Quantifizierung, Geruch

Einleitung

Gerüche haben einen großen Einfluss auf die (empfundene) Luftqualität und das Wohlbefinden der Menschen. Im Gegensatz zu den meisten anderen Messgrößen werden Geruchsstoffkonzentration und Geruchsintensität immer noch ausschließlich auf eine humansensorische Messung, also eine Bewertung durch menschliche Prüfer, zurückgeführt [1]. In vielen Anwendungsfeldern (z. B. der Bewertung von geruchsintensiven Anlagen [2, 3]) wäre eine kontinuierliche Bewertung der Luftqualität mit einem Sensorsystem als Ergänzung dieser gesetzlichen Methode vorteilhaft und würde den Anwendungsbereich von Geruchsmessungen erheblich erweitern. Hierzu müssen die relevanten Geruchsstoffe in sehr geringen Konzentrationen selektiv vor variablem geruchsirrelevantem Hintergrund erkannt und die modellbasierte Vorhersage aus Messdaten mit humansensorischen Bewertungen zur An-/Abwesenheit von Geruch und zur Geruchsintensität korreliert werden. Ein wichtiger Anwendungsfall sind Anlagen zur Abwasserreinigung, in denen primär Schwefelverbindungen zur Geruchsemission beitragen [4]. Zur Bewertung müssen geringe Konzentrationen im ppb-Bereich innerhalb von 10 s gemessen werden, weshalb Metalloxid-Halbleiter-Gassensoren (MOS) mit temperaturzyklischem Betrieb („virtuelles Sensorarray“) verwendet werden, die eine sehr hohe Empfindlichkeit und Selektivität aufweisen [5].

Experimentelle Methoden

Die Gasangebote werden mit einer Gasmischanlage erzeugt. Als Basis wird katalytisch gereinigte trockene Nullluft verwendet, der eine definierte Feuchte (50 %) sowie ein realistischer Hintergrund an Wasserstoff (500 ppb) und Kohlenmonoxid (150 ppb) hinzugefügt werden („Normalluft“). Der Normalluft werden die Probengase Schwefelwasserstoff (H_2S) und Dimethylsulfid (DMS) mittels eines zweistufigen Mischers [6] hinzudosiert, sodass sich definierte Konzentrationen zwischen 8,2 und 102,5 ppb (H_2S) beziehungsweise zwischen 7,6 und 95,0 ppb (DMS) ergeben. Es wird ein kommerziell erhältlicher MOS-Gassensor (AS-MLV-P2, ams Sensor Solutions Germany) verwendet, der mit einer speziellen Elektronik [7] temperaturzyklisch betrieben wird. Der Temperaturzyklus ist für die Auswertung nach der DSR Methode optimiert [8]. Dabei wird der Sensor mehrfach in Phasen hoher Temperatur betrieben, in denen die Sensoroberfläche oxidiert wird. Nach einer

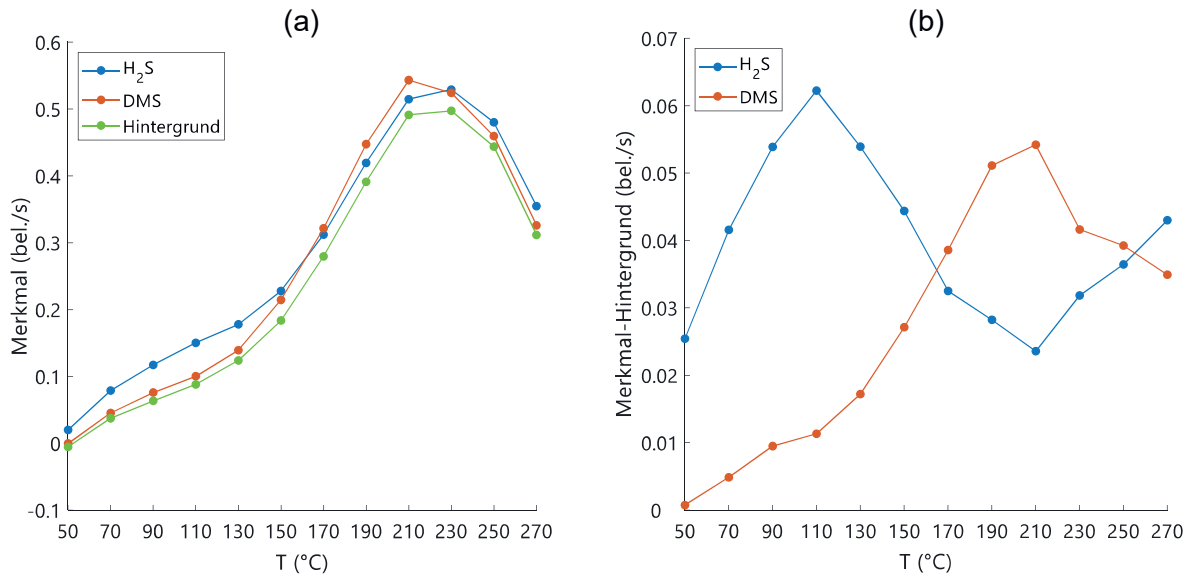


Abb. 1: Temperaturabhängigkeit des Merkmals (Steigung der verschiedenen Temperaturstufen) für H_2S und DMS ohne Abzug des Hintergrunds (verursacht v. a. durch H_2 und CO) (a) und nach Abzug des Hintergrunds (b) bei 100 GE/m³. Durch Abzug des Hintergrunds wird der unterschiedliche Temperaturverlauf der beiden Gase deutlich. Es handelt sich jeweils um Mittelwerte aus mehreren Temperaturzyklen.

schnellen Abkühlung auf eine geringere Temperatur bildet sich die Oxidation der Oberfläche nur langsam zurück. In dieser (kurzen) Phase nach einer Abkühlung ist die Reaktion von reduzierenden Gasen mit dem Oberflächensauerstoff bevorzugt, wodurch der Sensorleitwert steigt [5]. Auf jede der zwölf Hochtemperaturstufen von 450 °C folgt eine der zwölf äquidistanten Niedertemperaturstufen mit Temperaturen von 50 °C bis 270 °C (vgl. Tabelle 1). Die Dauer der Hochtemperaturstufe beträgt 1 s bis 3 s und ist so bemessen, dass der Sensor einen stabilen Gleichgewichtswert erreichen kann. Die Dauer der Niedertemperaturstufen beträgt je nach Temperatur 5 s bis 30 s. Der Sensorleitwert wird bei konstanter Spannung über eine Strommessung mit einem Logarithmier-Verstärker [7] mit einer Abtastrate von 1 kHz gemessen

Tab. 1: Übersicht über die Temperaturstufen und die Zeiten der extrahierten Merkmale

t(450 °C)	T	$t(T)$	t_{fit} (ab Sprung)
1 s	50 °C	30 s	0,91-1,64 s
1 s	70 °C	30 s	0,46-0,94 s
1 s	90 °C	25 s	0,50-0,83 s
1 s	110 °C	25 s	0,47-1,01 s
2 s	130 °C	20 s	0,50-1,46 s
2 s	150 °C	20 s	0,51-3,22 s
2 s	170 °C	15 s	0,53-2,43 s
2 s	190 °C	15 s	0,55-1,28 s
3 s	210 °C	10 s	0,54-0,87 s
3 s	230 °C	10 s	0,46-0,89 s
3 s	250 °C	5 s	0,36-0,71 s
3 s	270 °C	5 s	0,37-0,75 s

Ergebnisse und Diskussion

Mit Hilfe der Datenauswertungs-Toolbox DAV³E [9] werden aus den Messdaten die Geruchsstoffe identifiziert und ihre Konzentrationen bestimmt. Für die Vorverarbeitung wird das DSR-Modell [7] verwendet, bei dem die Steigung des logarithmierten Leitwerts kurz nach dem Temperatursprung berechnet und als Merkmal extrahiert wird (vgl. Tabelle 1, t_{fit} beschreibt den Zeitbereich nach dem Temperatursprung, für den die Steigung berechnet wird). In dieser Phase dominiert die Reaktion des Sauerstoffs mit reduzierenden Gasen, die Signalsteigung ist in erster Näherung proportional zur Ratenkonstante [5].
Im Folgenden werden die Konzentrationen in Relation zur Geruchsschwelle des jeweiligen Gases gesetzt, um eine Einordnung der entsprechenden zu bewertenden Geruchsstoffkonzentration zu erhalten. Geruchsstoffkonzentrationen können eigentlich nur durch olfaktorische Messungen bestimmt werden, für die hier verwendeten riechenden Gasmischungen aus jeweils einem verdünnten Reinstoff wird die Geruchsstoffkonzentration jedoch abgeschätzt als Quotient der Volumenkonzentration und der Geruchsschwelle des jeweiligen geruchsintensiven

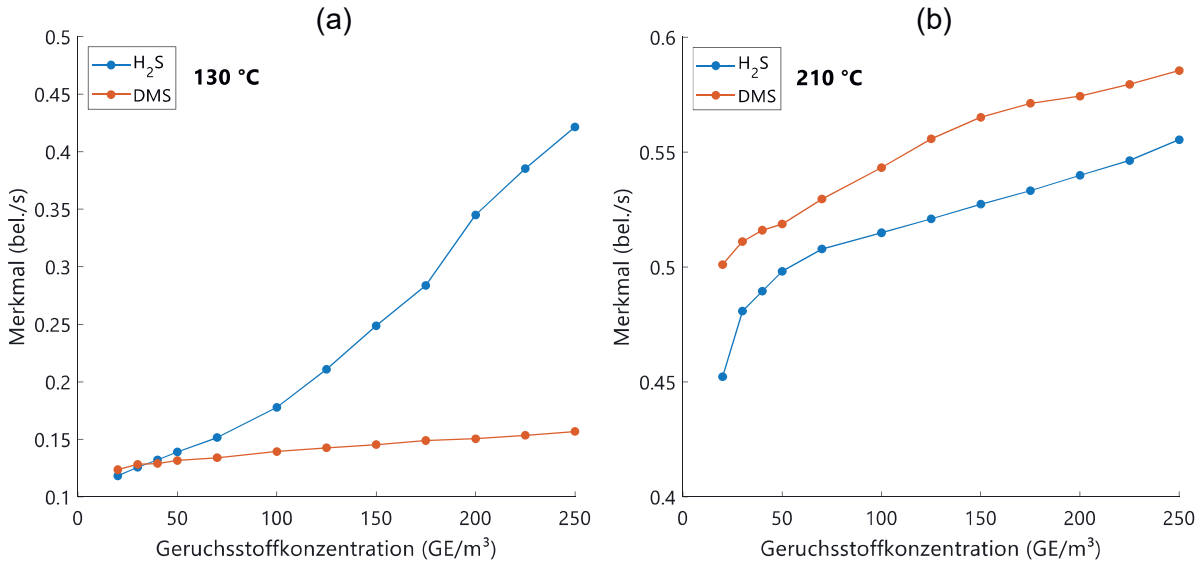


Abb. 2: Beispiele der Konzentrationsabhängigkeit der DSR-Steigung beider Gase bei 130 und 210 °C (a bzw. b). Auch die meisten anderen Temperaturen zeigen für beide Gase einen deutlich unterschiedlichen Verlauf des Merkmals über die Konzentration. Nur für die Temperaturstufe bei 230 °C ist die Empfindlichkeit für beide Gase ähnlich (vgl. Abbildung 3). Auch hier wurden Mittelwerte aus mehreren Temperaturzyklen berechnet.

Gases. Sie wird in GE/m³ angegeben (GE: Geruchseinheiten). Die Geruchsschwelle (per Definition 1 GE/m³) liegt für H_2S bei einer Konzentration von 0,41 ppb [10] und für DMS bei einer Konzentration von 0,38 ppb [11].

Das Temperaturverhalten der über mehrere Durchläufe des Temperaturzyklus gemittelten Merkmale von H_2S und DMS ist in Abbildung 1 a für eine Geruchsstoffkonzentration von 100 GE/m³ gezeigt. Der Untergrund ist durch die ebenfalls reduzierenden Hintergrundgase H_2 und CO dominiert. Dieser Anteil ist in Abbildung 1 b abgezogen. Es zeigt sich ein deutlich unterschiedlicher Verlauf für beide Gase. Während H_2S ein Maximum bei 110 °C besitzt, liegt dieses bei DMS erst bei einer Temperatur von circa 210 °C.

Die einzelnen Merkmale zeigen über die Geruchsstoffkonzentration ein in guter Näherung lineares Verhalten und eignen sich damit gut zur Quantifizierung von Einzelkomponenten. Bei den meisten Temperaturen sind deutliche Unterschiede zwischen den Zielgasen festzustellen, wie in Abbildung 2 beispielhaft für 130 und 210 °C dargestellt. Für das DSR-Merkmal bei 230 °C zeigt sich hingegen eine ähnliche Empfindlichkeit für H_2S und DMS im Bereich von 40 bis 250 GE/m³ (Abbildung 3), nur für die beiden kleinsten Geruchsstoffkonzentrationen (20 und 30 GE/m³) weichen die Merkmale deutlich voneinander ab.

Nach Abbildung 3 können H_2S und DMS durch Auswertung nur des Merkmals bei 230 °C und mithilfe einer linearen Anpassung quantifiziert werden (Abbildung 4). Die Verwendung ausschließlich eines Merkmals beziehungsweise einer Temperaturstufe zur gemeinsamen Quantifizierung beider Gase ist jedoch wahrscheinlich nur unter Laborbedingungen möglich. Variationen der Hintergrundkonzentrationen haben möglicherweise einen Einfluss auf die gleiche Konzentrationsabhängigkeit der Merkmale für beide Probengase bei 230 °C

In der Praxis wird daher eine Temperaturstufe nicht ausreichen und wie in diesem Beitrag ein

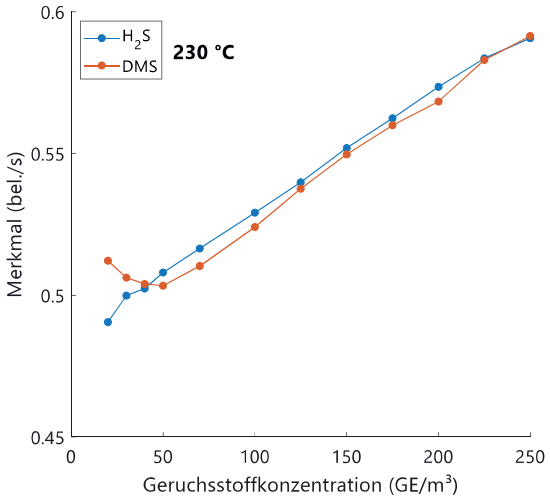


Abb.3: Konzentrationsabhängigkeit der DSR-Steigung (gemittelt über mehrere Temperaturzyklen) beider Gase bei 230 °C.

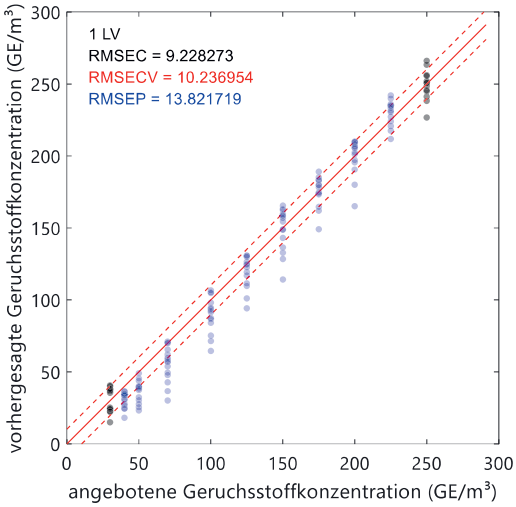


Abb. 4: Modell für H_2S und DMS gemeinsam über eine lineare Anpassung des Merkmals bei 230 °C. Die mittleren Konzentrationen (blaue Datenpunkte) wurden hierbei nach Training des Modells mit den äußersten Konzentrationen (schwarze Datenpunkte) vorhergesagt. Es zeigt sich eine gute Interpolierbarkeit.

Temperaturzyklus verwendet werden müssen, der eine selektive Messung ermöglicht und für die Quantifizierung mehr Merkmale zur Verfügung stellt. Die Quantifizierung kann dann mittels partial least squares regression (PLSR) erfolgen. Abbildung 5 zeigt die entsprechende Regression eines Einzelgases (Beispiel: H_2S in Abbildung 5 a) und die Regression beider Gase gemeinsam (Abbildung 5 b). Die Fehler (RMSE, root mean square error) des Modelltrainings (RMSEC), der zehnfachen Kreuzvalidierung (RMSECV) und der Vorhersage nicht beim Training oder der Validierung verwendeter Datenpunkte (RMSEP) sind in den Abbildungen enthalten. Für die Einzelgase beträgt der RMSECV jeweils etwa 2 GE/m³ und der RMSEP jeweils circa 10 GE/m³. Für die gemeinsame Quantifizierung beider Gase liegen beide Werte bei etwa 11 GE/m³. Die Vorhersage der PLSR ist somit noch etwas besser als die der linearen Anpassung des Merkmals bei 230 °C (RMSEP in etwa 14 GE/m³, Abbildung 4). Zudem ist zu erwarten, dass die Regression stabiler ist gegenüber Schwankungen des Hintergrundgases als ein einzelnes Merkmal.

Für die Quantifizierungen (Abbildungen 4 und 5) wurden nur die kleinste und die größte Konzentration zum Training verwendet, die Konzentrationen dazwischen wurden anschließend entsprechend des Modells vorhergesagt. Es zeigt sich hierbei eine sehr gute Interpolierbarkeit, was auch der kleine RMSEP verdeutlicht. Es können somit auch Konzentrationen gemessen werden, die nicht explizit kalibriert wurden.

Abbildung 5 b ist zusätzlich entsprechend eines gewählten Warnwertes von 100 GE/m³ farblich unterlegt. Grün kennzeichnet hierbei den Bereich unter 100 GE/m³, in dem eine Geruchsstoffkonzentration als ausreichend niedrig klassifiziert werden könnte. Oberhalb

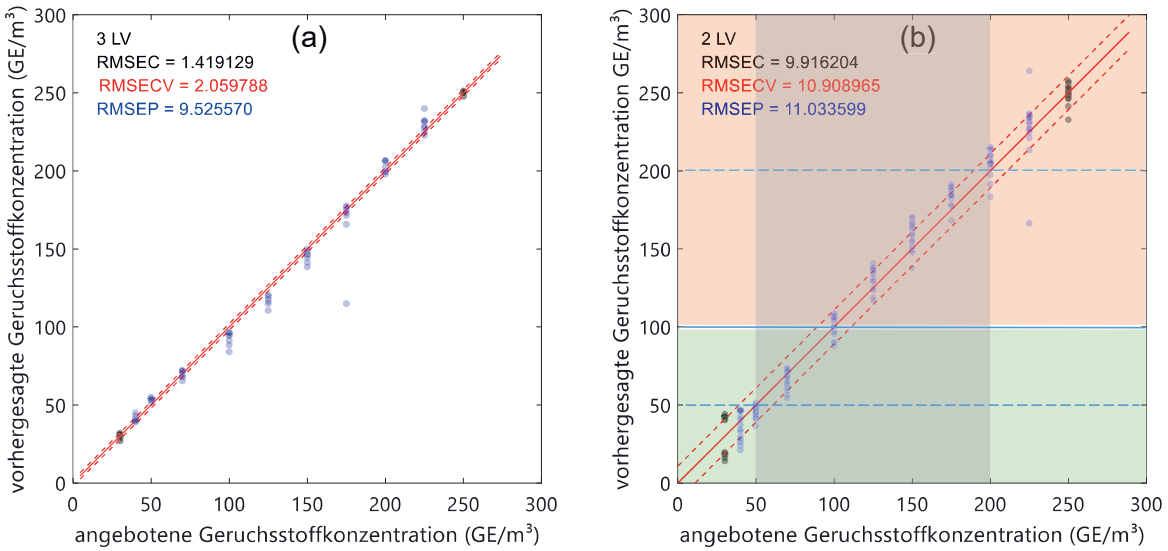


Abb. 5: PLSR für H_2S einzeln (a) sowie für H_2S und DMS gemeinsam (b). In beiden Fällen wurden alle Merkmale, d. h. alle Temperaturstufen, für die Regression verwendet. Zusätzlich ist in (b) der Bereich der erlaubten Unsicherheit bei der Bewertung bezüglich des Warnwertes von 100 GE/m³ grau unterlegt. Eine Fehlbewertung sehr kleiner (unter 50 GE/m³, grün) bzw. sehr großer Geruchsstoffkonzentrationen (über 200 GE/m³, rot) ist ausgeschlossen.

des Warnwertes würde hingegen eine zu starke Geruchsemission erkannt werden (rot). Zwischen 50 und 200 GE/m³ (erlaubte Unsicherheit von Faktor zwei des Warnwertes, grauer Bereich) werden Fehlbewertungen erlaubt, außerhalb dieses Bereichs sollen falsche Vorhersagen jedoch ausgeschlossen werden. In der präsentierten Messung wird dies sicher erreicht.
Über die Merkmale sind die beiden Gase im untersuchten Konzentrationsbereich (20-250 GE/m³) außerdem mit einer linearen Diskriminanzanalyse (LDA) gut voneinander trennbar (Abbildung 6). Für Geruchsstoffkonzentrationen unter 50 GE/m³ liegen sie allerdings nahe beieinander, und H_2S ist für diese Konzentrationen nicht von Normalluft zu unterscheiden. Mit einem 5-Nächste-Nachbarn-Klassifikator und einer zehnfachen Kreuzvalidierung ergibt sich ein Klassifikationsfehler von etwa 12 %, zu dem vor allem die H_2S-Datenpunkte unter 50 GE/m³ beitragen, die der Normalluft zugeordnet werden.

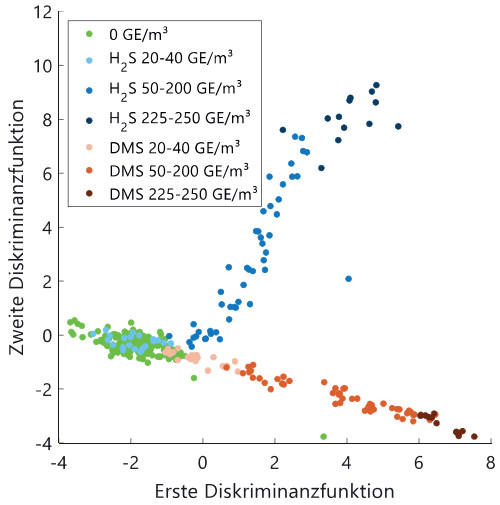


Abb. 6: Lineare Diskriminanzanalyse bezüglich der beiden Zielgase und des Hintergrunds.

Zusammenfassung und Ausblick

Es wurde gezeigt, dass durch geeignete Wahl des Temperaturzyklus sowohl die beiden Zielgase unterschieden werden können als auch eine gemeinsame Quantifizierung möglich ist. Außerdem reicht in diesem Fall schon ein Merkmal (230 °C) für eine gute Quantifizierung, mit einer Messunsicherheit kleiner 15 GE/m³. Das Überschreiten eines Schwellwertes von 100 GE/m³ mit einer erlaubten Unsicherheit von Faktor 2 kann somit sicher angezeigt werden.
In der vorliegenden Messung wurde der Hintergrund nicht variiert. Ob der Hintergrund einen Einfluss auf die Übereinstimmung der Empfindlichkeit für beide Gase bei 230 °C hat, muss noch untersucht werden.
Bei der präsentierten Untersuchung waren darüber hinaus die ähnlichen Geruchsschwellen der beiden Probengase von Vorteil. Eine Erweiterung des Verfahrens auf andere Gase, wie Methanthiol oder Dimethyldisulfid, deren Geruchsschwellen eine Größenordnung kleiner oder größer sind als die von H_2S und DMS, ist Gegenstand aktueller Untersuchungen.

Danksagung

Die Arbeiten wurden mit Mitteln des Bundesministeriums für Bildung und Forschung im Rahmen des Verbundprojekts SEPEG unter dem Förderkennzeichen 01IS17087 gefördert. Wir danken den Projektpartnern 3S – Sensors, Signal Processing, Systems GmbH, Saarbrücken, sowie Olfasense GmbH, Kiel, für die wichtigen Diskussionsbeiträge und die gute Zusammenarbeit.

Literaturnachweis

[1] DIN EN 13725:2003-07, Luftbeschaffenheit - Bestimmung der Geruchsstoffkonzentration mit dynamischer Olfaktometrie; Deutsche Fassung EN 13725:2003; doi: 10.31030/9364305
[2] DIN EN 16841-1:2017-03, Außenluft - Bestimmung von Geruchsstoffimmissionen durch Begehungen - Teil 1: Rastermessung; Deutsche Fassung EN 16841-1:2016; doi: 10.31030/2503315
[3] DIN EN 16841-2:2017-03, Außenluft - Bestimmung von Geruchsstoffimmissionen durch Begehungen - Teil 2: Fahnenmessung; Deutsche Fassung EN 16841-2:2016; doi: 10.31030/2503331
[4] E. C. Sivret, G. Parcsi, R. M. Stuetz: Sampling and Analysis Methodology Concerns for Volatile Organo-Sulfur Compounds (VOSCs), *Air and Waste Management Association Annual Conference and Exhibition* 103, 5131 (2010)
[5] T. Baur, A. Schütze, T. Sauerwald: Optimierung des temperaturzyklischen Betriebs von Halbleitergassensoren, *Tech. Mess.* 82(4), 187-195 (2015); doi: 10.1515/teme-2014-0007
[6] M. Leidinger, C. Schultealbert, J. Neu, A. Schütze, T. Sauerwald: Characterization and calibration of gas sensor systems at ppb level – a versatile test gas generation system, *Meas. Sci. Technol.* 29(1), 015901 (2018); doi: 10.1088/1361-6501/aa91da
[7] T. Baur, C. Schultealbert, A. Schütze, T. Sauerwald: Device for the detection of short trace gas pulses, *Tech. Mess.* 85(7-8), 496-503 (2018); doi: 10.1515/teme-2017-0137
[8] T. Sauerwald, T. Baur, M. Leidinger, W. Reimringer, L. Spinelle, M. Gerboles, G. Kok, A. Schütze: Highly sensitive benzene detection with metal oxide semiconductor gas sensors-an inter-laboratory comparison, *J. Sens. Sens. Syst.* 7, 235-243 (2018); doi: 10.5194/jsss-7-235-2018

[9] M. Bastuck, T. Baur, A. Schütze: DAV3E – a MATLAB toolbox for multivariate sensor data evaluation, *J. Sens. Sens. Syst.* 7, 489-506 (2018); doi: 10.5194/jsss-7-489-2018

[10] Y. Nagata: Measurement of odor threshold by triangle odor bag method, *Odor Meas. Rev.* 118-127 (2003)

[11] Persönliche Mitteilung, Olfasense GmbH: Interne Datenbank zu Geruchsschwellen

DOI 10.5162/sensoren2019/6.2.3

Impedanzspektroskopie eines auf metallorganischen Netzwerken basierenden Sorptionssensors

H. Lensch[1], N. Baskent[2], I. Wilhelm[3], J. Huerttlen[3], M. Voelker[2], A. Schütze[1], T. Sauerwald[1]
[1]Lehrstuhl für Messtechnik, Campus A5.1, Universität des Saarlandes, 66123 Saarbrücken
E-Mail-Adresse, h.lensch@lmt.uni-saarland.de
[2]Fraunhofer-Institut für Integrierte Schaltungen (IIS) Am Wolfsmantel 33 91058 Erlangen
[3]Fraunhofer-Institut für Chemische Technologien (ICT) Joseph-von-Fraunhofer-Str. 7 76327 Pfinztal

Zusammenfassung

Es wird ein auf metallorganischen Netzwerken (MOF) basierender Sensor präsentiert, auf welchem flüchtige organische Verbindungen (VOC) aufgrund von hohen Verteilungskoeffizienten sehr gut sorbieren können. Der Sensor besteht aus einem keramischen Substrat, auf dem Interdigital-Elektroden aufgebracht sind. Auf diesen Elektroden wird das mit Bindern versetzte MOF UIO-66 mittels Airbrush Verfahren aufgesprüht. Die Kapazität des Sensors im Bereich von 14,9 pF bis 15,4 pF bei einem Ethanolangebot von 0 ppm bis 1.000 ppm wird mittels Impedanzspektroskopie gemessen. Weiterhin wird die Querempfindlichkeit auf Feuchte (40 %rH und 60 %rH) untersucht, welche zusätzlich zu einer starken Änderung der Kapazität (auf bis zu 24,3 pF) auch noch einen zusätzlichen resistiven Signalanteil aufweist. Mit Hilfe von Impedanzspektroskopie, können daher die kapazitiven und resistiven Beiträge der Feuchte separat bestimmt werden. Dies eröffnet die Möglichkeiten, um den Einfluss der Feuchte zu kompensieren.

Keywords: Gassensoren, VOC-Messung, Metal-organic frameworks, Impedanzspektroskopie

Einleitung

Das Messen von flüchtigen organischen Verbindungen (engl. Volatile organic compounds – VOCs) gewinnt aufgrund von der von ihnen ausgehenden Gesundheitsrisiken zunehmend an Bedeutung. Eine Herausforderung beim Messen dieser Stoffe ist die niedrige Konzentration der Einzelkomponenten und die große Vielfalt an möglichen Zielgasen [1], [2]. Metallorganische Netzwerke (engl. Metal-Organic Framework – MOF) eignen sich ausgesprochen gut zum sorbieren von VOCs, da die erzielten Verteilungskoeffizienten um Größenordnungen oberhalb denen klassischer Sorptionsmaterialien liegen können [3]. Darüber hinaus stellen MOFs sterisch und chemisch auf bestimmte VOC zugeschnittene Adsorptionsplätze bereit [4]. Dies ermöglicht ein neues Sensorkonzept für VOCs basierend auf einer direkten Sorptionsmessung.
Die kapazitive Auslesung von Sorptionssensoren für Feuchte ist gut beschrieben, da diese eine hohe Änderung der Permittivität verursacht [5]. Neben diesem kapazitiven Signalanteil führt die Feuchtesorption auch zu einer Änderung der ionischen Leitfähigkeit [6], was für VOC-Sensoren zur Kompensation der Feuchtequerempfindlichkeit genutzt werden kann, die eine Messung sonst erheblich stört. In dieser Arbeit wird eine erste Charakterisierung eines MOF-basierten Sorptionssensor mittels Impedanzspektroskopie vorgestellt und im Rahmen eines vereinfachten Leitungsmodells interpretiert, das als Basis für eine Kompensation der Querempfindlichkeit auf Feuchte dienen kann.

Sensordesign

Der Sorptionssensor weist drei Interdigital-Elektroden (*Ana, Ref, Com*) auf, welche nebeneinander auf einem Substrat aufgebracht sind. Dabei bildet jeweils eine Elektrode zusammen mit der gemeinsamen Elektrode (*COM*) einen Kondensator (Abbildung 1). Eines der Elektrodenpaare (*Ana - Com*) wird mit einer sensitiven Schicht versehen und das andere Paar bleibt unbeschichtet (*Com - Ref*), um für eine spätere differentielle Messanordnung zur Verfügung zu stellen. Als Substrat wird eine keramische Leiterplatte verwendet (Rogers4350B, rel. Permittivität 3.48±0.05 @ 10 GHz).

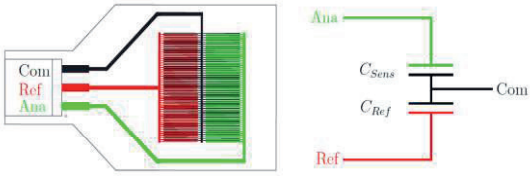


Abbildung 1: Sensoraufbau (links) mit elektrischen Ersatzschaltbild der Anordnung (rechts).

Als sensitive Schicht wurde die mikroporöse metall-organische Gerüstverbindung (MOF) UiO-66 verwendet, welche mit Bindern zu einer viskosen Paste verarbeitet und via Airbrush-Verfahren auf Interdigital-Elektroden aufgetragen wurde. Durch die Anzahl der Beschichtungsvorgänge lässt sich die Dicke des sensitiven Materials einstellen, wobei pro pro Beschichtung ca. 1 µm aufgetragen wird. Für erste Untersuchungen wurde eine Schichtdicken von ca. 2 µm bis 20 µm eingestellt, die auf die Elektroden mit einer Metalldicke von 35 µm sowie einer Breite und einem Abstand von 75 µm aufgebracht wurden.



Abbildung 2: Sensor mit 2 (links), 10 (Mitte) und 20 (rechts) MOF-Schichten

Experimenteller Aufbau

Um den Sensor mit verschiedenen Testgasen zu beaufschlagen wurde dieser in eine luftdichte Messkammer aus Aluminium eingebaut und mit Hilfe einer Flachdichtung aus Viton auf der Platine gedichtet, so dass die elektrischen Anschlussbuchsen herausragen. An die Pins *Ana* und *Com* wurde direkt das Agilent 4294A bzw. Solartron 1260 Impedanzspektrometer angeschlossen.
Die Gasanschlüsse der Messkammer wurden mit einer Gasmischanlage verbunden, mit der eine Gasmischung aus verschiedenen Testgasen bei verschiedenen relativen Feuchten angeboten werden kann.

Modell

Durch Sorption eines Gases an der Gerüstverbindung ändert sich die Permittivität ε_r der Sensorschicht.
Diese Änderung führt zu einer messbaren Erhöhung der Kapazität [7]. Für nichtleitende Adsorbate ändert sich nur der Realteil der Permittivität und der Sensor verhält sich im gesamten Impedanzspektrum als idealer Kondensator. Dies entspricht im Bode-Diagramm einer Steigung von -1 in der Amplitude und mit einem Phasenwinkel von -90°.
Die Adsorption von Feuchte führt zu einer ausgesprochen hohen Änderung der Permittivität bzw. Erhöhung der Kapazität, da Feuchte sehr polar ist, ein Umstand der auch in kapazitiven Feuchtesensoren genutzt wird. Trotz der Verwendung von Materialien, die unpolare VOC bevorzugt adsorbieren, führt ein rein kapazitives Auslesen der Sensoren zu einer sehr ausgeprägten Querempfindlichkeit auf Feuchte. Zur selektiven Bestimmung zwischen Gas und Feuchte muss also zumindest ein weiterer Sensorparameter verwendet werden. In [6] wurde gezeigt, dass adsorbierte Feuchte auf TiO_2 zusätzlich zu dem kapazitiven Verhalten noch weitere Signalanteile besitzt, u.a. ein sogenanntes Constant Phase Element, welches eine Überlagerung von rein realen und rein imaginären frequenzabhängigen Impedanzen mit konstantem Phasenwinkel ergibt. Dieser Signalanteil wird mit der ionischen Leitfähigkeit der Wassermoleküle auf der Oberfläche gedeutet (Grothuss Mechanismus). Da es sich hierbei also um einen Effekt vom physisorbiertem Wasser handelt, erwarten wir, dass dieser Signalanteil bei einer Vielzahl von Materialien, z.B. den hier verwendeten MOFs auftritt. Mittels Fit an ein passendes Ersatzschaltbild können CPE und Kapazität aus der Gesamtimpedanz getrennt werden, da die ionischen Effekte in niedrigeren Frequenzbereichen dominanter sind als der für VOC Detektion interessante kapazitive Anteil [7],[8]. So entstehen weitere Merkmale im Spektrum, die selektiv auf Feuchte sind. Diese können genutzt werden, um den Feuchteanteil der Impedanz im rein kapazitiven Bereich zu kompensieren.

Ergebnisse

Alle Messungen wurden mit dem Sensor mit 20 Schichten durchgeführt, da anzunehmen ist, dass dort, durch die größte Zahl an Adsorptionsplätzen die Änderung der Kapazität am größten ist. Alle Messungen wurden bei Raumtemperatur durchgeführt.

Abbildung 3 zeigt ein Bode Diagramme des Sensors aufgenommen mit dem Agilent 4294A Impedanzspektrometer im Frequenzbereich von 40 Hz bis 125 MHz. Die Messung wird in Stickstoff mit Reinheit 5.0. Die Amplitude hat bei allen aufgenommenen Spektren eine Steigung von –1 und einen Phasenwinkel von –90°, was auf ein rein kapazitives Verhalten hinweist. Unter der Annahme eines idealen Kondensators wurde anschließend die Kapazitäten des Sensors bestimmt (Tabelle 1).

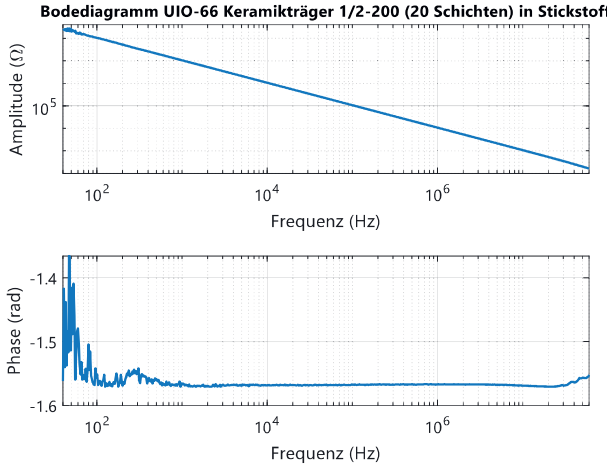


Abbildung 3: Sensor mit Schicht in Stickstoff 5.0

Um das zeitliche Verhalten zu charakterisieren wurde eine Messung in Stickstoff 5.0 und verschiedenen Ethanolkonzentrationen (0 ppm, 200 ppm, 500 ppm, 1000 ppm) durchgeführt. Die Konzentration wurden jeweils nach 30 min erhöht. Die Spektren wurden mit dem Agilent 4294A Impedanzspektrometer aufgenommen, da dieses pro Spektrum weniger Zeit als das Solartron benötigt und so das dynamische Verhalten besserabgebildet werden kann. Anschließend wurde aus den Spektren die Kapazität bestimmt und für 520 kHz den zeitlichen Verlauf der Kapazität aufgetragen (Abbildung 4).

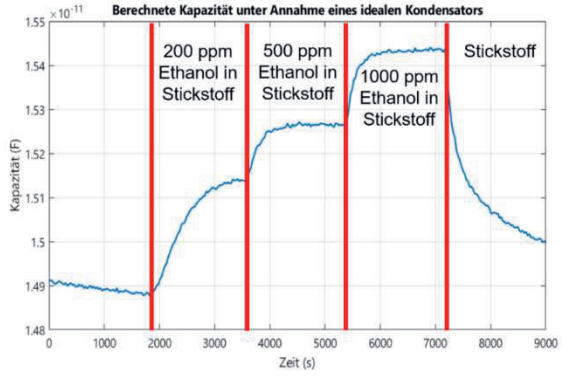


Abbildung 4: Zeitlicher Verlauf der Kapazität des Sensors mit MOF-Schicht bei unterschiedlichen Ethanolangeboten bei 520 kHz

Nachdem der Sensor bei trockener Umgebung charakterisiert wurde, wird im Folgenden der Einfluss auf Feuchte untersucht. Um Messung mit einer niedrigeren Grenzfrequenz (1 Hz) zu ermöglichen wurden die Messungen mit einem Solatron 1260 durchgeführt. In Abbildung 5 sind zwei Spektren bei jeweils 40 %rH und 60 %rH abgebildet. Diese zeigen im oberen Frequenzbereich (ab 4 kHz) ein rein kapazitives Verhalten. Im mittleren Bereich (5 Hz bis 200 Hz) bildet sich bei 60 %rH ein deutlich resistiver Bereich aus und im unteren Bereich wird die Impedanz leicht kapazitiver.
Bei 40 %rH wird der mittlere resistive Bereich erst nahe der unteren Grenzfrequenz ausgebildet, was an dem Abflachen der Amplitude und dem Rückgang der Phase Richtung 0° erkennbar ist. Für detailliertere Untersuchungen ist es allerdings notwendig den untersuchten Frequenzbereich weiter nach unten zu erweitern.

Tabelle 1 Kapazitäten bei verschiedenen Gasangeboten

Sensor	**Kapazität [nF]**
Trocken in Stickstoff	14.9
200 ppm Ethanol in Stickstoff	15.11
500 ppm Ethanol in Stickstoff	15.25
1000 ppm Ethanol in Stickstoff	15.4
40 %rH in Stickstoff	20.71
60 %rH in Stickstoff	24.30

Diskussion & Fazit

Die Messungen in Abbildung 3 und Abbildung 4 zeigen ein rein kapazitives Verhalten des Sensors bei verschiedenen Konzentrationen von Ethanol. Obwohl sich der Sensor nicht linear verhält, ist eine klare Quantifizierung der Konzentration möglich. Die Empfindlichkeit von 15 pF zwischen 0 und 1000 ppm Ethanol ist zwar ausreichend, allerdings sind die hier untersuchten Konzentrationen relativ hoch und liegen z.B. oberhalb derer, die in eine Messung der Innenluft potenziellen Anwendung relevant sind. Es kann außerdem aufgrund der zu wählenden Gitterstruktur des MOF keine Unterscheidung von Gasen mit ähnlicher oder kleinerer Molekülgröße erfolgen. Aus diesem Grund zeigen MOFs ohne weitere Funktionalisierung eine geringe Selektivität. Auch optisch ausgelesene Sorptionssensoren wie in [9] beschrieben zeigen für leichtflüchtige VOC eine niedrigere Nachweisschwelle als alternative Sensoren, wie Halbleitergassensor oder Photoionisationsdetektor. Allerdings nimmt die Nachweisempfindlichkeit mit sinkender Flüchtigkeit sehr stark zu. Wir erwarten, dass kapazitiv ausgelesene

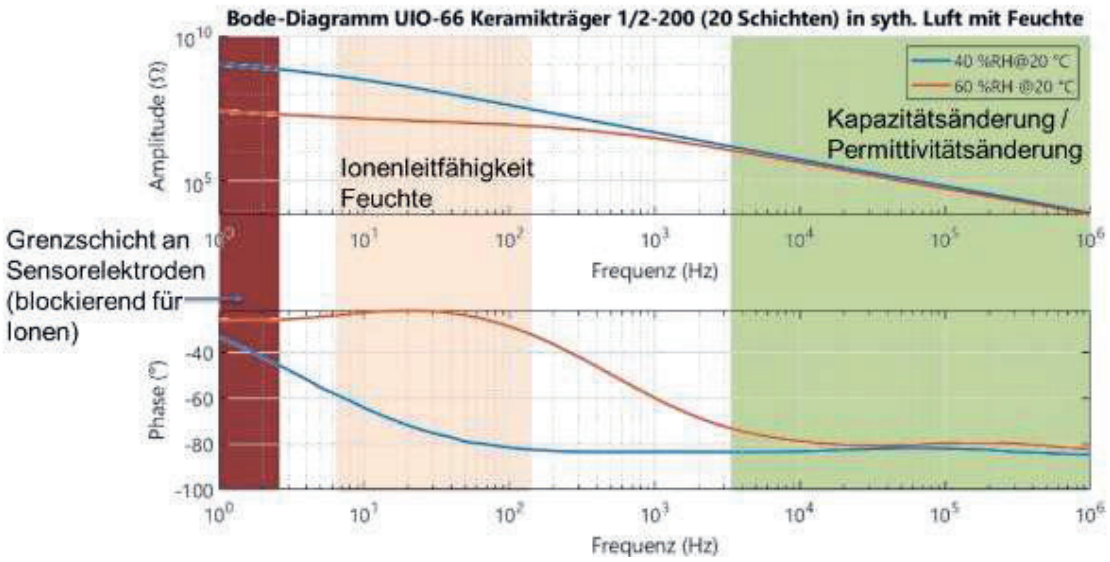


Abbildung 5: Bode Diagramm der Sensorschicht bei 40 %rH (blau) und 60 %rH (rot)

Sensoren v.a. für schwerflüchtige VOC eine Ergänzung zu bestehenden Konzepten sind, da für diese Stoffgruppen sehr hohe Verteilkoeffizienten zu erwarten sind. Für VOC in niedrigen Konzentrationen können auch kombinierte System verwendet werden [3]. In denen der MOF als Präkonzentrator benutzt wird und ein hochempfindlicher Halbleitergassensor als Detektor verwendet wird. In einem solchen System kann eine Messung des Präkonzentrators mit Impedanzspektroskopie zusätzliche Informationen über die Menge an sorbierten Stoffen liefern. Der Einfluss von Feuchte auf den Sensor ist vergleichbar mit dem in [6] beschriebenen Feuchtesensoren, obwohl es sich um ein völlig anderes Material handelt. Dies stützt die Vermutung, dass der Sensormechanismus in beiden Fällen auf eine Leitfähigkeit von physisorbiertem Wasser zurückzuführen ist. In Abbildung 5 hat der Feuchteeffekt soweit im Frequenzbereich abgebildet dem Modell aus [6] ähnelt. Impedanzspektroskopie als Messprinzip zum Auslesen des Sensors funktioniert sowohl bei der Sorption von Ethanol als auch von Feuchte. Feuchte liefert zusätzliche Merkmale in anderen Frequenzbereichen als Ethanol im Spektrum. Die Sensitivität auf Feuchte ist in dem untersuchten Messbereich deutlich größer als die auf die Zielgas, was eine Kompensation erschwert.

Danksagung

Die Entwicklung und Fertigung der Sensoren erfolgte im Rahmen des Forschungsvorhabens „Technologien für das Internet der Sinne“ gefördert durch das Bundesland Bayern.

Literaturnachweis

[1] J. A. Bernstein *et al.*, “The health effects of nonindustrial indoor air pollution,” *J. Allergy Clin. Immunol.*, vol. 121, no. 3, pp. 585–591, 2008.

[2] A. Schütze *et al.*, “Highly Sensitive and Selective VOC Sensor Systems Based on Semiconductor Gas Sensors: How to?,” *Environments*, vol. 4, no. 1, p. 20, 2017.

[3] M. Leidinger, M. Rieger, T. Sauerwald, C. Alépée, and A. Schütze, “Integrated pre-concentrator gas sensor microsystem for ppb level benzene detection,” *Sensors Actuators, B Chem.*, vol. 236, pp. 988–996, 2016.

[4] J. L. C. Rowsell and O. M. Yaghi, “Metal-organic frameworks: A new class of porous materials,” *Microporous and Mesoporous Materials*, vol. 73, no. 1–2. pp. 3–14, 2004.

[5] T. Wagner *et al.*, “A high temperature capacitive humidity sensor based on mesoporous silica,” *Sensors*, vol. 11, no. 3, pp. 3135–3144, 2011.

[6] H. Lensch, B. Manuel, A. Schütze, and T. Sauerwald, “Impedance model for a high temperature ceramic humidity sensor,” *J. Sensors Sens. Syst.*, vol. 8, no. Special issue Sensors and Measurement Systems 2018, pp. 1–14, Apr. 2019.

[7] C. Sapsanis *et al.*, “Insights on capacitive interdigitated electrodes coated with MOF thin films: Humidity and VOCs sensing as a case study,” *Sensors (Switzerland)*, vol. 15, no. 8, pp. 18153–18166, Jul. 2015.

[8] E. Barsoukov and J. R. Macdonald, *Impedance Spectroscopy: Theory, Experiment, and Applications*, 2nd ed. Wiley-VCH-Verlag, 2005.

[9] DIN Deutsches Institut für Normung e. V., “Indoor air — — Part 29 : Test methods for VOC detectors,” no. 40, Berlin: Beuth Verlag, 2013, p. 44.

DOI 10.5162/sensoren2019/6.2.4

Einsatz der Cyclovoltammetrie zur Steigerung der Selektivität von Festelektrolytsensoren

Anastasiya Ruchets[1], N. Donker[2], D. Schönauer-Kamin[2], R. Moos[2], J. Zosel[1], U. Guth[3], M. Mertig[1,3]
[1]Kurt-Schwabe-Institut für Mess- und Sensortechnik e. V. Meinsberg, Kurt-Schwabe-Straße 4, 04736 Waldheim, Germany
[2]Lehrstuhl für Funktionsmaterialien, Universität Bayreuth, 95440 Bayreuth, Germany
[3]Professur für Physikalische Chemie, Mess- und Sensortechnik, Technische Universität Dresden, 01062 Dresden, Deutschland

Zusammenfassung

Mit einem kommerziellen Festelektrolytsensor (FES) auf der Basis von stabilisiertem Zirconiumdioxid wurden unter Verwendung der Cyclovoltammetrie die redoxaktive Gase wie H_2, O_2 und H_2O, verdünnt in N_2, bestimmt. Die erzielten Ergebnisse zeigen, dass parallel sowohl H_2 als auch O_2 quantitativ und mit hoher Selektivität bei verschiedenen H_2O-Konzentrationen mittels Cyclovoltammetrie gemessen werden können. Die wasserstoffabhängigen Peaks treten in den cyklischen Voltammogrammen (CV) in anodischer Scan-Richtung bei Potentialen zwischen 0,4 und 0,2 V gegen eine Pt/Luft-Referenzelektrode bei Scangeschwindigkeiten zwischen 10 und 100 mV/s im Temperaturbereich 700 – 750 °C auf. Die untere Nachweisgrenze für H_2 liegt unter 5 Vol.-ppm während die obere Nachweisgrenze etwa 100 Vol.-ppm beträgt. Dieser Messbereich wurde für O_2-Konzentrationen 0 – 200 Vol.-ppm und H_2O-Konzentrationen 0 – 600 Vol.-ppm ermittelt.

Keywords: Cyclovoltammetrie, Gassensoren, Wasserstoff, Yttriumoxid-stabilisiertes Zirconiumdioxid (YSZ), Pt-Elektroden.

Einführung

Festoxid-Elektrolyte wie Yttriumoxid-stabilisiertes Zirconiumdioxid (YSZ) mit hoher Sauerstoffionenleitfähigkeit und geringer elektronischer Leitfähigkeit sind ausgezeichnete Materialien für Festelektrolyt-Sauerstoffsensoren bei mittleren und hohen Temperaturen.
Solche Festelektrolytsensoren (FES) können im stationären oder dynamischen Modus betrieben werden. Je nachdem, ob es sich bei dem Ausgangssignal um eine Spannung oder einen elektrischen Strom handelt, können die im stationären Modus betriebenen elektrochemischen Gassensoren als potentiometrisch oder amperometrisch klassifiziert werden [1]. Da Nichtgleichgewichts-Gasgemische an den heißen und katalytisch hochaktiven Pt-Elektroden dieser Sensoren sofort equilibriert werden, sind diese Sensoren im stationären Modus nicht für einzelne Komponenten selektiv, was für manche Applikationen einen gravierenden Nachteil darstellt [2].
Ein dynamischer Modus umfasst alle Methoden, bei denen die Messelektrode zeitabhängig polarisiert und die Antwort des Sensors ebenfalls zeitabhängig gemessen wird. Beispiele für diesen Modus sind die Pulspolarisation [3], die Cyclovoltammetrie [4] oder die Impedanzspektroskopie (EIS) [5]. Es wurde bereits gezeigt, dass bei dynamisch betriebenen Sensoren eine Verbesserung der Selektivität erreicht werden kann, weil unterschiedliche Kinetiken der Elektrodenreaktionen genutzt und so Einzelkomponenten unabhängig voneinander erfasst werden können [6].
In dieser Arbeit wird die Eignung der Cyclovoltammetrie beim Betrieb eines rohrförmigen FES für die Steigerung der Selektivität in Nichtgleichgewichts-Gasgemischen untersucht. Der FES ist mit zwei Pt-Elektroden ausgestattet, wobei die Arbeitselektrode auf der Rohrinnenseite in der Messgasatmosphäre und die Referenz- und Gegenelektrode auf der Rohraußenseite an Luft positioniert sind. Es konnte bereits gezeigt werden [4, 6], dass mit Hilfe der Cyclovoltammetrie an solchen FES eine selektive Messung in redoxaktiven Gasgemischen möglich wird. Der vorliegende Beitrag hat zum Ziel, die Messbereichsgrenzen für diese Methode genauer zu quantifizieren.

Experimenteller Aufbau

Der gelb hinterlegte Teil des in Abb. 1 gezeigten Versuchsaufbaus ist für die gleichzeitige und unabhängige Einstellung der Konzentrationen von O_2, H_2 und H_2O in Stickstoff-Gasgemischen vorgesehen.

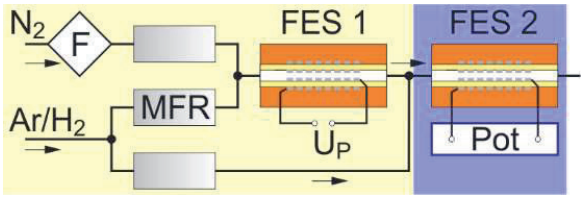


Abb 1: Schema des Versuchsaufbaus, FES = Festelektrolytsensor, Pot = Potentiostat, MFR = Masseflussregler, F = Filter für O_2, H_2O und Kohlenwasserstoffe.

O_2 und H_2O wurden durch Sauerstoffelektrolyse in den wasserstoffhaltigen Stickstoff im ersten Festelek-trolytsensor (FES 1) fixiert. Der blau hinterlegte zweite Festelektrolytsensor (FES 2, beide FES vom Typ O2-DF 28.0, ZIROX - Sensoren & Elektronik GmbH, Greifswald, Deutschland) wurde für dynamische elektrochemische Messungen verwendet. Beide Sensoren vom Typ "Messgas, Pt|YSZ|Pt, Luft" weisen zwei 4 cm lange zylindrische Elektroden auf, die aus einem reinem Pt-Netz (250 mesh) bestehen und mit poröser YSZ-Paste auf den Innen- und Außenflächen des YSZ-Rohres (8 mol-% Yttriumoxid, d_{innen} = 3,9 mm (Arbeitselektrode) und $d_{außen}$ = 5,9 mm (Referenz-/Gegenelektrode) gesintert wurden. Alle Gasversorgungsleitungen sind mit Edelstahlrohren und Verschraubungen (Swagelok, Solon, OH, USA) ausgestattet. Die Anschlüsse beider FES an die Gasversorgung wurden so modifiziert, dass die He-Leckrate der FES von 10^{-6} auf 10^{-9} mbar·L/s verringert wurde.

Da die signalrelevante Kinetik in hohem Maß von der Elektrodenmorphologie beeinflusst wird, wurde diese mittels Rasterelektronenmikroskopie (REM) untersucht. Die in Abb. 2 aufgeführten REM-Aufnahmen belegen die Anbindung der Elektrodendrähte an den Festelektrolyten durch eine poröse YSZ-Schicht. Diese poröse Schicht bedeckt große Teile der Mess- und Referenzelektrode und bietet eine erweiterte Dreiphasengrenze (DPG), an welcher der Sauerstofftransfer zwischen Gas und Festelektrolyt erfolgt. Das Messgas diffundiert teilweise durch diese poröse Schicht, um die DPG zu erreichen.

Die Durchflussmengen von Stickstoff 5,0 und Wasserstoff-Testgas (0,1 Vol.-% H_2 in Ar) wurden mit Masseflussreglern (Brooks Instrument Company, Hatfield, USA) eingestellt. Das Verdünnungsgas Stickstoff wurde durch einen Dreifachfilter für Wasserdampf, Sauerstoff und Kohlenwasserstoffe (Restek GmbH, Bad Homburg) geleitet, wodurch dessen Reinheit auf einen Grad > 6,0 angehoben wird.

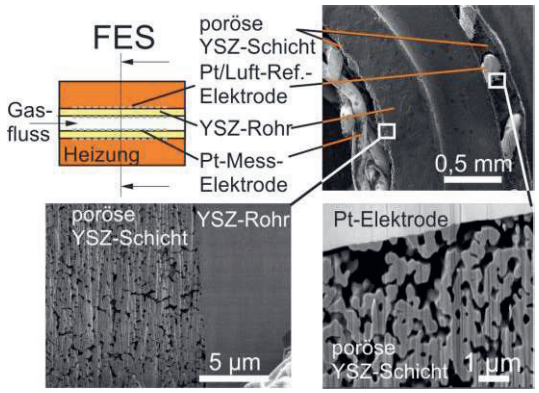


Abb 2: Schema und REM-Aufnahmen von Schnitten durch ein Festelektrolytrohr (8 mol-% YSZ) eines FES mit beidseitig aufgesinterten Pt-Elektroden in einer porösen YSZ-Schicht.

Die elektrochemischen Messungen an FES 2 wurden mit einem elektrochemischen Mehrzweck-Messsystem Reference 600™ (Gamry Instruments, Warminster, USA) durchgeführt.

Ergebnisse und Diskussion

Einfluss der Scangeschwindigkeit

Die in Abb. 3 dargestellten CV wurden in einem nicht-equilibrierten Gasgemisch mit $\varphi(H_2)$ = 50 Vol.-ppm, $\varphi(O_2)$ = 10 Vol.-ppm und $\varphi(H_2O)$ < 10 Vol.-ppm bei unterschiedlichen Scangeschwindigkeiten und der FES-Temperatur 750 °C gemessen. Sie bestätigen die in [4] dargestellten Ergebnisse, die ebenfalls belegen, dass wasserstoffbedingte Peaks in anodischer Scanrichtung zwischen -0,4 und -0,2 V bei Scangeschwindigkeiten zwischen 10 und 100 mV/s auftreten.

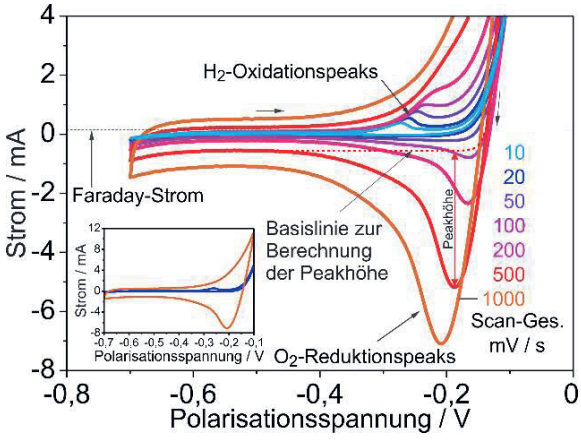


Abb 3: CV verschiedener Scangeschwindigkeiten; Durchfluss = 10 ml/min, $\varphi(H_2)$ = 50 Vol.-ppm, $\varphi(O_2)$ = 10 Vol.-ppm, $\varphi(H_2O)$ < 1 Vol.-ppm, Sensortemperatur = 750°C, Inset-Diagramm: CV bei 20 und 1000 mV/s im gesamten Strommessbereich.

Im Gegensatz dazu entwickeln sich die sauerstoffbedingten Peaks bei Scangeschwindigkeiten über 100 mV/s in kathodischer Scanrichtung zwischen -0,1 und -0,45 V.
Diese Unterschiede zwischen beiden Reaktionskinetiken ermöglichen einen unabhängigen und selektiven Nachweis beider Gase im untersuchten Konzentrationsbereich.
In Anbetracht der Tatsache, dass die katalytisch unterstützte Wasserstoffoxidation an heißen Pt-Oberflächen mit sehr hohen Geschwindigkeiten abläuft, war es überraschend, bei diesen vergleichsweise niedrigen Scangeschwindigkeiten einen ausgeprägten wasserstoffbedingten Oxidationspeak zu finden. Aus diesem Ergebnis lässt sich vermuten, dass dieser Analyt in der Nähe der Elektrode während der kathodischen Polarisation gespeichert und aus diesem Speicher beim Rücklauf in anodische Richtung verzögert freigegeben wird.
Um die verschiedenen gemessenen Peaks mit den Versuchsbedingungen zu korrelieren, wurden die entsprechenden Peakhöhen berechnet, indem der Strom einer Basislinie vom Peakstromwert abgezogen wurde. Dies ist in Abb. 3 für den Sauerstoffpeak bei 10 – 1000 mV/s dargestellt. Die Basislinien wurden berechnet, indem eine Exponentialfunktion an die kathodische CV-Kurve in den Potentialbereichen (-0,1 – -0,15 V) und (-0,5 – -0,55 V) für die O_2-Peaks sowie in den anodischen Bereichen (-0,5 – -0,45 V) und (-0,15 – -0,1 V) für die H_2-Peaks angepasst wurde.
Um mehr über die kinetischen Limitierungen der Elektrodenreaktionen zu erfahren, wurde die Abhängigkeit der Peakhöhe von der Scangeschwindigkeit untersucht.

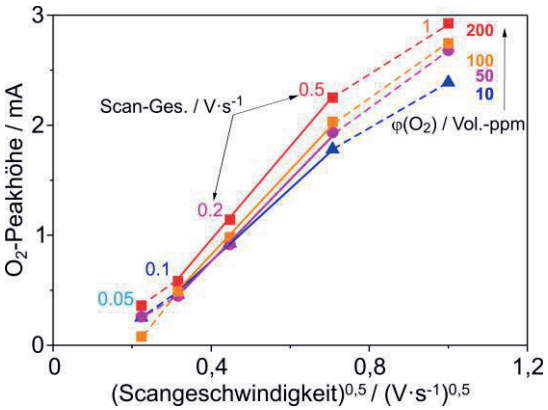


Abb. 4: Abhängigkeit der O_2-Peakhöhen von der Quadratwurzel der Scangeschwindigkeit;, $\varphi(H_2)$ = 50 Vol.-ppm, $\varphi(O_2)$ = 10, 50, 100, 200 Vol.-ppm, $\varphi(H_2O)$ < 1 Vol.-ppm, Sensortemperatur = 650 °C, Vol.str. = 10 ml/min.

Wenn die Peakhöhe linear mit der Quadratwurzel der Scangeschwindigkeit zunimmt, wie es die Randles-Ševčik-Gleichung (RSG) [7, 8] beschreibt, liegt eine kinetische Limitierung der elektrochemischen Reaktion durch Diffusion vor. Diese Abhängigkeit wurde bereits für die H_2-bezogenen Peaks gefunden [4].
In Abb. 4 ist dargestellt, dass die RSG diese Abhängigkeit auch für die O_2-bezogenen Peaks bei Scangeschwindigkeiten zwischen 100 und 500 mV/s beschreibt. Außerhalb dieses Bereichs wirken andere Prozesse wie der Ladungstransfer begrenzend und führen so zu einer Abweichung vom durch die RSG beschriebenen linearen Verlauf.
Mit der RSG ist eine Abschätzung des O_2-Diffusionskoeffizienten aus dem Geradenanstieg möglich. Die größte Unsicherheit dieser Schätzung betrifft die Elektrodenfläche A, die aus den REM-Aufnahmen der in eine poröse YSZ-Schicht eingebetteten Elektrode auf A = 3,5 cm^2 geschätzt wurde. Der resultierende Diffusionskoeffizient D = 2,6 cm^2/s ist etwas höher als der O_2-Diffusionskoeffizient in der an die Elektrode angrenzenden Gasphase, der sich aus der Chapman-Enskog-Gleichung [9] zu $D(O_2 \text{ in } N_2)$ = 1,4 cm^2/s bei 650 °C und 1 bar berechnen lässt. Die O_2-Umsatzlimitierung erfolgt demnach vor allem durch Gasdiffusion an der Arbeitselektrode und weniger durch Oberflächendiffusion.
Durch Änderung der Scangeschwindigkeit lassen sich somit selektiv und nahezu parallel Wasserstoff und Sauerstoff im Gasgemisch nachweisen. Im Folgenden werden die Bereichsgrenzen dieser Methode beschrieben.

Einfluss von Volumenstrom und Temperatur

In Abb. 5 ist der Einfluss des Volumenstroms auf die CV bei drei FES-Temperaturen sowie $\varphi(H_2)$ = $\varphi(O_2)$ = 50 Vol.-ppm dargestellt. Mit zunehmendem Volumenstrom steigt die Menge an überschüssigem Sauerstoff, der die Elektrode pro Zeiteinheit erreicht. Dies führt zu einer Verschiebung der CV zu negativeren Stromwerten. Bei Temperaturen zwischen 700 und 750 °C wird am kathodischen Wendepunkt ein zunehmender Reduktionsstrom beobachtet, der bei geringen Volumenströmen am deutlichsten ausgeprägt ist. Dies deutet darauf hin, dass freier Sauerstoff den H_2-Speicherprozess beeinflusst.
Mit diesem Anstieg des Reduktionsstroms bei abnehmendem Volumenstrom und steigender Temperatur steigt auch die Höhe der H_2-Oxidationspeaks.

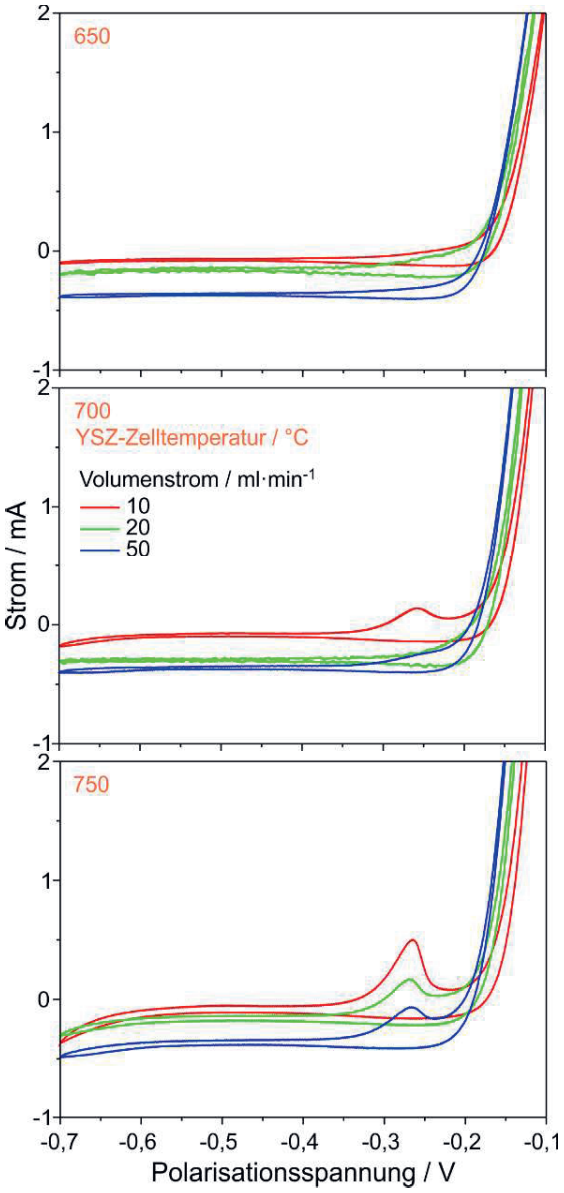


Abb. 5: CV bei verschiedenen Volumenströmen und Sensortemperaturen; Scangeschwindigkeit = 20 mV/s, φ(O_2) = 50 Vol.-ppm, φ(H_2) = 50 Vol.-ppm, φ(H_2O) < 1 Vol.-ppm.

Wiederholgenauigkeit der Methode

Um die Wiederholgenauigkeit des H_2-Peaks zu untersuchen, wurden 30 identische CV mit einer Minute Pause zwischen zwei aufeinanderfolgenden CV durchgeführt. Aus dieser Messung sind Beispiele in Abb. 6 dargestellt. Diese zeigen, dass der Wasserstoffpeak mit zunehmender Zyklenzahl abnimmt, bis nach etwa 20-25 Zyklen ein konstanter Wert erreicht wird. Weiterhin wurde gefunden, dass sich diese Änderungen zwischen aufeinanderfolgenden CV dadurch verringern lassen, dass der Sensor zwischen zwei CV kurzzeitig mit der vor dem ersten CV gemessenen Gleichgewichts-Spannung U_{eq} polarisiert wird.

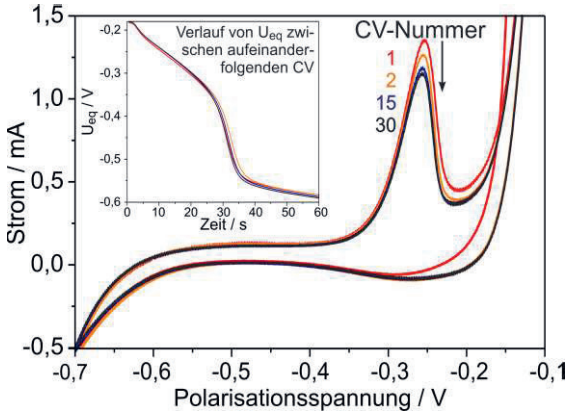


Abb. 6: Beispiele von 30 CV bei konstanten Bedingungen mit einer Minute Pause zwischen zwei aufeinanderfolgenden CV; Sensortemperatur = 750 °C, Volumenstrom = 10 ml/min, Scangeschwindigkeit = 20 mV/s, φ(O_2) = 0,2 Vol.-ppm, φ(H_2) = 50 Vol.-ppm, φ(H_2O) < 1 Vol.-ppm, Inset-Diagramm: zwischen zwei CV gemessene Gleichgewichtsspannung U_{eq}.

Um die Reproduzierbarkeit der Methode zu verbessern, wurden automatisierte Abläufe programmiert und immer die CV mit der gleichen Zyklusnummer verglichen.

Randbedingungen für die Wasserstoffdetektion

Messbereich der H_2-Konzentration

Ausgewählte CV in Messgasen unterschiedlicher H_2-Konzentration mit φ(O_2) = 0,2 Vol.-ppm sind in Abb. 7 dargestellt. Die Peakauswertung erfolgte wie oben für Sauerstoffpeaks beschrieben durch Subtraktion der exponentiell ansteigenden Basislinie.

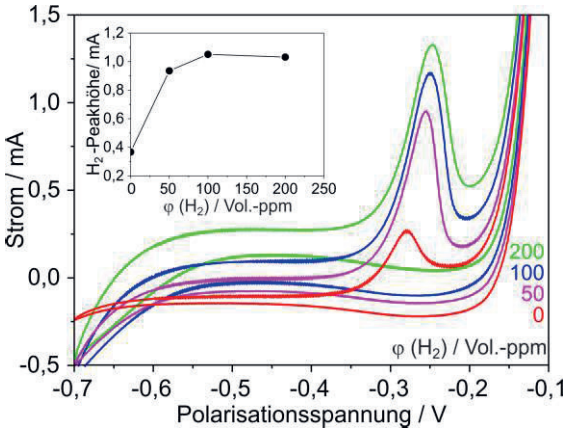


Abb 7: CV bei verschiedenen H_2-Konzentrationen; Sensortemp. = 750 °C, Volumenstrom = 10 ml/min, Scangeschwindigkeit = 20 mV/s, φ(O_2) = 0,2 Vol.-ppm, Inset-Diagramm: H_2-bezogene Peakhöhe.

Die im Inset von Abb. 7 dargestellten resultierenden Peakhöhen steigen bis zu einer H_2-Konzentration von etwa 100 Vol.-ppm an und verlaufen danach nahezu konstant. Weiterhin verschieben sich die CV mit zunehmender H_2-Konzentration in anodische Stromrichtung, da im gesamten Scanbereich ein potentialunabhängiger Faradayscher Stromfluss überlagert ist. Ähnliche Experimente zur Charakterisierung der O_2-Querempfindlichkeit der H_2-Peaks sind in [6] für den Temperaturbereich 700 – 750°C bei $\varphi(O_2)$ = 0,2 – 200 Vol.-ppm beschrieben.

Einfluss der H_2O-Konzentration

In Abb. 8 sind Ergebnisse zum Einfluss der H_2O-Konzentration auf die H_2-Detektion bei niedrigen Scangeschwindigkeiten dargestellt. Im untersuchten H_2O-Konzentrationsbereich zwischen 0 – 600 Vol.-ppm wurden zunächst ansteigende H_2-Peakhöhen gefunden, die später mit weiter zunehmender H_2O-Konzentration wieder abnehmen. Der bei kleinen H_2O-Konzentrationen <400 Vol.-ppm gefundene Anstieg der H_2-Peakhöhe könnte auf eine Intensivierung der Oberflächendiffusion von H_2 in den Zwischenspeicher zurückzuführen sein. Weiterhin tritt bei steigender H_2O-Konzentration eine leichte Peak-Verbreiterung auf und die Stromwerte im Potentialbereich zwischen -0,4 und -0,7 V werden nach unten verschoben. Die bei H_2O-Konzentrationen > 400 Vol.-ppm beobachtete Abnahme der H_2-Peakhöhe wird dagegen wahrscheinlich von anderen Vorgängen verursacht, z.B. könnte eine erhöhte Konkurrenz-Adsorption von Wassermolekülen an der Elektrode den H_2-Übertritt hemmen.

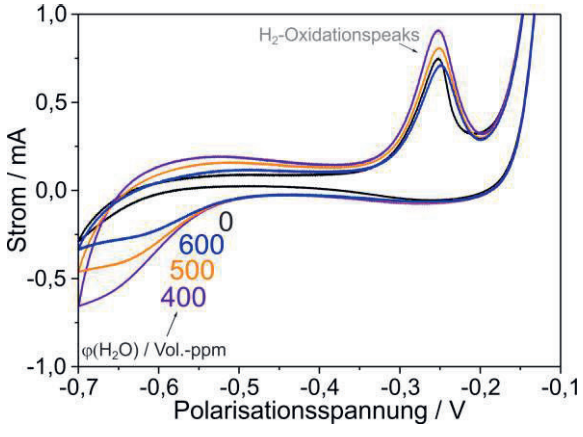


Abb 8: CV bei verschiedenen H_2O-Konzentrationen; $\varphi(H_2)$ = 50 Vol.- ppm, $\varphi(O_2)$ = 10 – 15 Vol.-ppm, Sensortemperatur = 750 °C, Scangeschwindigkeit = 20 mV/s, Volumenstrom = 10 ml/min.

In [10] wird beschrieben, wie die Oberfläche einer Pt-Mikroelektrode mit H_2O-Molekülen bedeckt ist, wenn die H_2O-Konzentration im Gas ansteigt. Darüber hinaus beeinflusst Wasser auch die Wasserstofflöslichkeit im Festelektrolyt [11, 12]. Wasser wandert auch selbst in den Elektrolyt ein [13].
Dennoch lässt sich der H_2-bezogene Peak auch bei relativ hohen H_2O-Konzentrationen mit hoher Sensitivität nachweisen. Der Anstieg des Reduktionsstroms im Potentialbereich zwischen -0,6 und -0,7 V eröffnet die Möglichkeit, neben H_2 und O_2 auch H_2O bei höheren Konzentrationen in den Gasgemischen messtechnisch zu erfassen.

Einfluss der O_2-Konzentration

Der in [6] beschriebene Einfluss der Sauerstoffkonzentration im Bereich $\varphi(O_2)$ = 0 – 200 Vol.-ppm auf die Peakhöhe der gemessenen H_2-Peaks ist auch bei hohen Sauerstoffüberschüssen so gering, dass die H_2-Peaks sicher nachweisbar bleiben. Steigt die O_2-Konzentration über diesen Bereich hinaus an, sollten sich die H_2-Peaks verringern und schließlich verschwinden, da die H_2-Oxidation in der Gasphase immer mehr beschleunigt wird. Diese Vermutung wird durch die in Abb. 9 gezeigten CV bestätigt. Im Bereich $\varphi(O_2)$ = 500 – 1000 Vol.-ppm sind keine H_2-Peaks mehr nachweisbar.
Möglicherweise lassen sich auch bei höheren Sauerstoffkonzentrationen als 200 Vol.-ppm noch Wasserstoffpeaks finden, wenn der negative Umkehrpunkt der CV zu noch negativeren Potentialen verschoben wird. Diese Untersuchungen sind geplant.

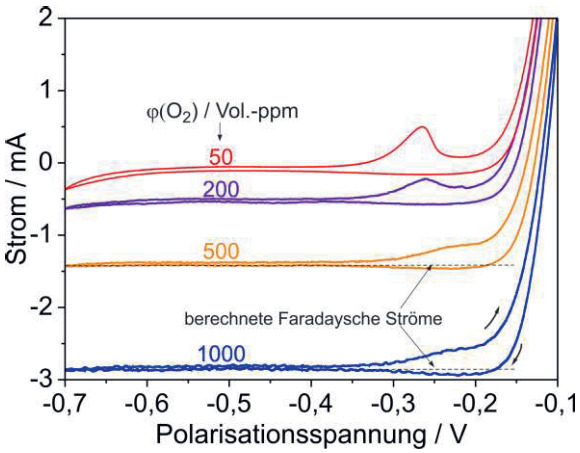


Abb 9: CV bei erhöhten O_2-Konzentrationen; $\varphi(H_2)$ = 50 Vol.-ppm, $\varphi(H_2O)$ < 1 Vol.-ppm, Sensortemperatur = 750 °C, Scangeschwindigkeit = 20 mV/s, Volumenstrom = 10 ml/min.

Schlussfolgerungen

Die Cyclovoltammetrie als dynamische Polarisationsmethode wurde an einem kommerziell verfügbaren Festelektrolytsensor angewendet, der im Durchflussmodus betrieben wurde. Durch den Einsatz der Cyclovoltammetrie und die Verwendung von Gasgemischen mit unterschiedlichen Wasserstoff-, Sauerstoff- und Wasserdampfkonzentrationen wurde nachgewiesen, dass es möglich ist, zwischen den untersuchten Redoxkomponenten mit dieser dynamischen Methode in nicht-equilibrierten Gemischen durch Variation experimenteller Parameter wie Scangeschwindigkeit und Sensortemperatur zu unterscheiden. Eine unabhängige quantitative Bestimmung kleiner H_2- und O_2-Konzentrationen ist auch bei höheren Überkonzentrationen des jeweiligen Reaktionspartners mit hoher Selektivität und Sensitivität möglich. In dieser Arbeit wurden die Messbereichsgrenzen einer solchen selektiven Messung näher bestimmt. Dabei ließ sich zeigen, dass der H_2-Nachweis bis zu 100 Vol.-ppm bei Volumenströmen zwischen 10 und 20 ml/min und O_2-Konzentrationen im Bereich von 0 bis 200 Vol.-ppm sowie in Gegenwart von Wasserdampf mit 0 – 600 Vol.-ppm im untersuchten Potentialbereich zwischen -0,7 und -0,1 V möglich ist.

Danksagung

Die Arbeit wird von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Projekts "Dynamische Methoden für elektrochemische Gassensoren (DynaSens)" (MO 1060/30-1, ZO 139/3-1) gefördert.
Die Autoren danken J. Posseckardt vom KSI für die Anfertigung der REM-Aufnahmen. Darüber hinaus danken die Autoren Herrn F. Altmann und seinem Team am Fraunhofer-Institut für Mikrostrukturierung von Materialien und Systemen (IMWS), Halle, für die Durchführung von materialbezogenen Untersuchungen und Analysen.

Literatur

[1] S. Zhuiykov, *Electrochemistry of Zirconia Gas Sensors*. CRC Press, Boca Raton (2007)
[2] H.-H. Möbius, *Solid-State Electrochemical Potentiometric Sensors for Gas Analysis*, in: W. Göpel et al. (Eds.), Sensors A comprehensive survey, Vol. 3, VCH, Weinheim (1991)
[3] S. Fischer, R. Pohle, E. Magori, M. Fleischer, R. Moos, *Solid State Ionics* 262, 288–291 (2014); doi: 10.1016/j.ssi.2014.01.022
[4] M. Schelter, J. Zosel, W. Oelßner, U. Guth, M. Mertig, *Journal of Sensors and Sensor Systems* 5, 319-324 (2016); doi: 10.5194/jsss-5-319-2016
[5] H. Pöpke, E. Mutoro, C. Raiß, B. Luerßen, M. Amati, M.K. Abyaneh, L. Gregoratti, J. Janek, *Electrochim. Acta* 56, 10668–10675 (2011); doi: 10.1016/j.electacta.2011.04.057.
[6] A. Ruchets, N. Donker, D. Schönauer-Kamin, R. Moos, J. Zosel, U. Guth, M. Mertig *Sensors & Actuators: B. Chemical* 290, 53–58 (2019); doi: 10.1016/j.snb.2019.03.063
[7] J. E. B. Randles, *Trans. Faraday Soc.,* 44, 327-338 (1948); doi: 10.1039/TF9484400327
[8] A. Sevcik, *Collect. Czech. Chem. Commun.* 13, 349-377 (1948); doi: 10.1135/cccc19480349
[9] E. l. Cussler, Cambridge University Press, Cambridge 104-108 (1997); ISBN 0 521 56477
[10] C. Vonau, Dissertation, Technische Universität Dresden, 151 (2015); http://nbn-resolving.de/urn:nbn:de:bsz:14-qucosa-183965.
[11] K. Park, D.R. Olander, *J. Am. Ceram. Soc.* 74 (1) (1991); 72–77, doi: 10. 1111/j.1151-2916.1991.tb07299.x.
[12] S. Yamanaka, T. Nishizaki, M. Uno, M. Katsura, *J. Alloys Compd.* 293-295, 38–41 (1999); doi: 10.1016/S0925-8388(99)00396-5.
[13] C. Wagner, *Berichte der Bunsengesellschaft für physikalische Chemie* 72 (7), 778–781 (1968); doi: 10.1002/bbpc.19680720709.

DOI 10.5162/sensoren2019/6.2.5

Niedertemperatur-Pellistoren aus mesoporösem Au-Pd@Co_3O_4

Xuemeng Lyu[1,], Dr. Haitao Gao[1,*], Dr. Katrin Schmitt[1,2], Dr. KarinaTarantik[2], Prof. Dr. Jürgen Wöllenstein[1,2]*

[1]Institut für Mikrosystemtechnik – IMTEK, Universität Freiburg 79110 Freiburg i. Br.
[2]Fraunhofer Institut für physikalische Messtechnik IPM, 79110 Freiburg i. Br.
** Beide Autoren haben im gleichen Umfang zu dieser Arbeit beigetragen*
Telefon: +49 761 203-97632, Email: haitao.gao@imtek.uni-freiburg.de

Zusammenfassung

Durch die Verwendung von mesoporösem Co_3O_4 als Trägermaterial für einen bimetallischen Au-Pd-Katalysator wird die Arbeitstemperatur von sogenannten Pellistoren für den Methannachweis von ca. 450 °C bei herkömmlichen Al_2O_3-Sensoren mit Pd-Katalysator auf 300 °C reduziert. Hierdurch werden die möglichen Vergiftungen der verwendeten Katalysatoren bei höheren Temperaturen, die bei herkömmlichen Pellistoren inhärent sind, verringert. Die erzielten Ergebnisse zeigen, dass mesoporöses Co_3O_4 vorteilhaft gegenüber kommerziell erhältlichem polykristallinem Co_3O_4 in Bezug auf die katalytische Umsetzung ist. Die Au-Pd-Bedeckung des Co_3O_4 beeinflusst das Verhalten des Sensors. Außerdem hat die Anzahl der mit einer Mikropipette (0,2 µL/Schicht) aufgebrachten Schichten Einfluss auf das Pellistorsignal. Die Entwicklung bimetallischer katalytischer Nanomaterialien auf Spinell-basierten Trägern für den Niedertemperaturnachweis brennbarer Gase verbessert den Explosionsschutz entscheidend, verringert die Heizleistungsaufnahme und verhindert die „Vergiftung“ des Katalysators durch Silane.

Keywords: Katalytischer Sensor, Pellistor, Methan, Niedertemperaturnachweis, Au-Pd@meso-Co_3O_4, mesoporös, Au-Pd

Einführung

Die frühzeitige Detektion explosiver Gasgemische ist hochrelevant und vermeidet die Gefährdung von Menschen und Schäden an Infrastrukturen. Millionen von katalytischen Sensoren für die Anwendung sogenannter Pellistoren werden jährlich neu installiert. Herkömmliche Pellistoren arbeiten bei hohen Temperaturen (> 500 °C), um sicherzustellen, dass die katalytische Reaktion stattfindet. Dies führt jedoch zu einigen Nachteilen, wie z. B. schlechter Stabilität, hoher Anfälligkeit gegenüber Katalysatorvergiftung und einem kleinen Messbereich [1]. Die katalytische Vergiftung ist die Hauptursache bei Sensorversagen, wobei vor allem Si-haltige Gase eine irreversible Hemmung der katalytischen Oxidation verursachen [2]. Als repräsentative organische Silikonverbindung ist Hexamethyldisiloxan (HMDSO) bis zu einer Temperatur von 300 °C stabil, d. h. weniger als 3,5 % der Sensoren mit dieser (oder einer geringeren Arbeitstemperatur) werden innerhalb eines Jahres vergiftet [3]. Um eine dauerhafte Vergiftung zu minimieren, ist es somit notwendig, die Betriebstemperatur der katalytischen Sensoren auf bis zu 300 °C zu senken. Dies erfordert die Untersuchung neuer Materialsysteme, die eine höhere katalytische Aktivität und zudem eine große Oberfläche oder Porosität aufweisen und gleichzeitig robust gegenüber Vergiftungen sind. Kürzlich wurde über mesoporöses Co_3O_4 als Trägermaterial für einen bimetallischen Au-Pd-Katalysator (Au-Pd@Co_3O_4) zur Verbesserung der katalytischen Eigenschaften bei der Methanverbrennung berichtet [4]. In der Veröffentlichung wurden einige Materialkombinationen beschrieben, bei denen ein Methanumsatz von 50 % bei 280 °C und von 90 % bei 324 °C erreicht wurde. Tatsächlich liegen diese Ergebnisse sehr nahe an der Temperatur, bei der siliziumhaltige Gase stabil bleiben. Sie sind deshalb aussichtsreiche funktionelle Materialien für die Realisierung von Pellistoren mit deutlich niedrigeren Betriebstemperaturen.
Nachfolgend stellen wir unsere Forschungsergebnisse zu mesoporösem Au-Pd@Co_3O_4 vor, das für Pellistoren mit niedriger Arbeitstemperatur eine der vielversprechendsten Materialkombinationen darstellt.

Herstellung

Abbildung 1 veranschaulicht schematisch den Herstellungsprozess von Au-Pd@Co_3O_4. Dreidimensionales mesoporöses KIT-6 Silizium–dioxid wurde im Autoklaven als Matrize hergestellt [5]. $Co(NO_3)_2 \cdot 6H_2O$ wurde mit dem KIT-6 in Toluol gemischt. Nach der Filtration wurde das feste Gemisch vier Stunden bei 600 °C kalziniert. Das Siliciumdioxid-Template wurde durch heißes NaOH (70 °C) entfernt. Das verbleibende Meso-Co_3O_4 wurde zu einer Au-Pd-Legierungskolloidlösung (Au-Pd@PVA) gegeben, die durch die Reduktion von Au- und Pd-Salzen (Au:Pd = 1:2) mittels $NaBH_4$ in verdünnter Polyvinylalkohol-Lösung (PVA, als Oberflächenstabilisator) hergestellt wurde. Nach der Filtration wurde das Material Au-Pd@PVA@meso-Co_3O_4 bei 400 °C getempert, wodurch Au-Pd@meso-Co_3O_4 nach der thermischen Entfernung von PVA gebildet wurde. Das Endprodukt wurde in Wasser mit einem Feststoffgehalt von 15 g/L dispergiert, um die Tinte des Funktionsmaterials zu erhalten. Dann wurde die hergestellte Tinte auf einem Keramiksubstrat (Abb. 2(a)) mittels einer Mikropipette (0,2 µL/Schicht) abgeschieden. Je nachdem, wie viele Schichten gewünscht wurden, wurde die Abscheidung wiederholt.

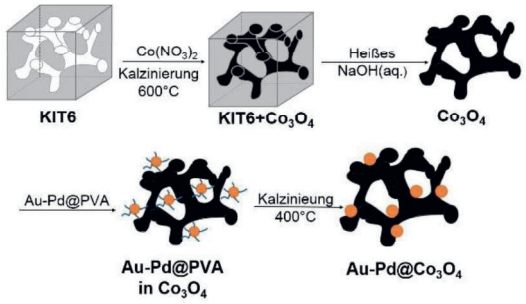


Abb. 1: Schematische Darstellung des Synthesevorgangs.

Der fertige Chip (Abb. 2(b)) wurde schließlich auf einen TO39-Sockel montiert. Der fertiggestellte Chip ist in Abb. 2(c) gezeigt.

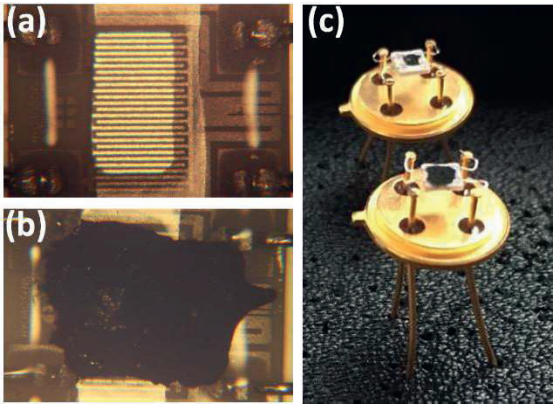


Abb. 2: (a) Unbeschichtetes Keramiksubstrat bestehend aus Pt-Heizer und Interdigitalelektroden. (b) Keramik–substrat, bedeckt mit dem Funktionsmaterial Au-Pd@Co_3O_4. (c) Die fertigen Chips wurden auf einem TO39-Sockel montiert.

Ergebnisse und Diskussion

Die synthetisierten Au-Pd-Partikel mit PVA-Oberflächenstabilisator wurden mittels dynamischer Lichtstreuung (DLS) charakterisiert. Die durchschnittliche Größe der Au-Pd-Partikel liegt bei 9,4 ± 1,1 nm (Abb. 3).

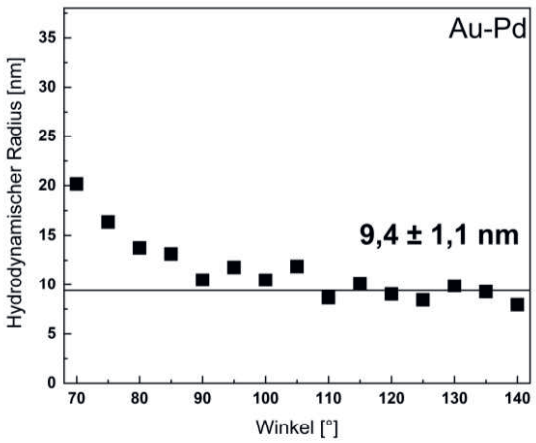


Abb. 3: DLS-Ergebnis für Au-Pd-Legierung zeigt, dass der durchschnittliche Durchmesser bei 9,4 ± 1,1 nm liegt.

Die Pulver-XRD-Messung wurde für Meso-Co_3O_4 und Au-Pd(12 Gew.%)@meso-Co_3O_4 (nominaler Gehalt an Au-Pd) durchgeführt. Die Pulver-XRD Daten (Abb. 4) zeigen, dass Meso-Co_3O_4 (JCPDS 42-1467) phasenrein ist und eine durchschnittliche Größe von 12,1 nm aufweist, die mit Hilfe der Scherrer-Gleichung geschätzt wurde. Die Au-Pd-Phase ist in Au-Pd@meso-Co_3O_4 nicht nachweisbar, sowohl

aufgrund der geringen Menge (7,5 % Au-Pd, Gew.% ermittelt aus der REM-EDX Analyse) als auch aufgrund des breiten Reflexes, der sich aus seiner Nanogröße ergibt. Daher ist nur Co_3O_4 in der XRD-Gesamtpulvermessung sichtbar.

Das REM-Bild in Abb. 5(a) zeigt, dass das Material im Allgemeinen hochporös ist. Im vergrößerten Bild Abb. 5 (b) liegt die Partikelgröße unter 1 µm. Die Größe der Edelmetallpartikel liegt im Bereich von wenigen Nanometern [4]. Daher ist es schwierig, sie mittels REM-Untersuchungen aufzulösen. Die EDX-Analyse zeigt, dass der tatsächliche Gehalt an Au-Pd in Au-Pd(12 Gew.%)@meso-Co_3O_4 7,5% beträgt.

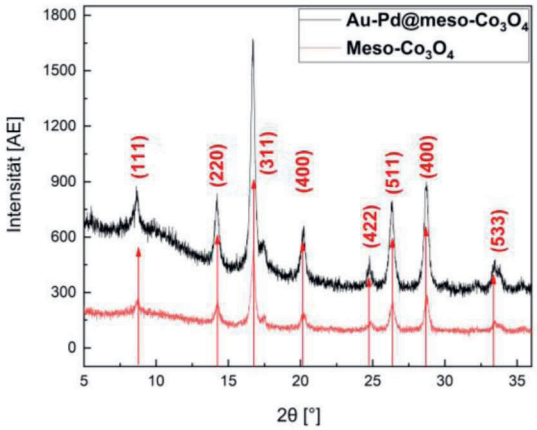


Abb. 4: XRD-Ergebnisse für Au-Pd@meso-Co_3O_4 (schwarz) und Meso-Co_3O_4 (rot). Die Au-Pd-Partikel werden aufgrund eines geringen Anteils (4,9 Gew.% nach REM-EDX Analyse) nicht detektiert.

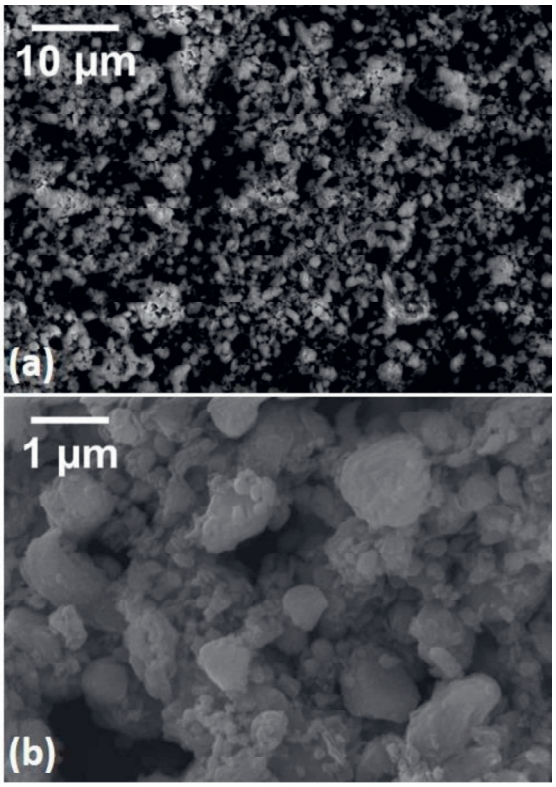


Abb. 5: REM-Aufnahmen der Au-Pd (12 Gew.%)@meso-Co_3O_4-Schicht (a) Übersicht (in 2000-facher Vergrößerung), (b) in 18000-facher Vergrößerung.

Um die Struktur von Au-Pd@meso-Co_3O_4 und die Verteilung der Au-Pd-Partikel zu untersuchen, wurden TEM-Messungen durchgeführt (Abb. 6). Drei Tinten wurden verglichen: Au-Pd (4 Gew.%) auf kommerziellen Co_3O_4 Partikeln, Au-Pd(4 Gew.%)@meso-Co_3O_4 und Au-Pd(12 Gew.%)@meso-Co_3O_4. Die Bedeckung von kleinen Au-Pd-Partikeln auf der Co_3O_4-Oberfläche ist in allen drei Tinten deutlich zu sehen. Die Größenverteilung ist ähnlich, die Größe der Au-Pd-Teilchen ist durchschnittlich 3,18 ± 0,55 nm für Au-Pd(4 Gew.%)@kommerzielles Co_3O_4 und 2,91±0,66 nm Au-Pd(4 Gew.%)@meso-Co_3O_4 und 3,13 ± 0,62 nm für Au-Pd(12 Gew.%)@meso-Co_3O_4. Diese Werte sind aufgrund der Oberflächenstabilisatorwirkung von PVA auf die Lichtstreuung geringer, als die mit der DLS-Messung ermittelten. Die Morphologie des kommerziellen Co_3O_4 unterscheidet sich von der des Meso-Co_3O_4. Das kommerzielle Co_3O_4 besteht aus großen "bulk" Co_3O_4-Partikeln mit einer Größe von mehr als 50 nm und seine Oberfläche ist mit kleinen Au-Pd-Partikeln bedeckt (Abb. 6(a) und (b)). Das Meso-Co_3O_4 bildet dahingegen ein dreidimensionales

Netzwerk, bestehend aus mesoporösen Co_3O_4- und Au-Pd-Partikeln (Abb. 6(c) und (d)). Die dunkleren Edelmetallpartikel sitzen entweder auf der Oberfläche des Co_3O_4 oder in den Mesoporen des Netzwerks. Abb. 6(e) und (f) zeigen Meso-Co_3O_4 mit nominal 12 Gew.% Au-Pd. Die Bedeckung mit Au-Pd-Partikeln im Au-Pd(12 Gew.%)@meso-Co_3O_4-System ist hoch im Vergleich zum Au-Pd(4 Gew.%)@meso-Co_3O_4-System (Abb. 6(c) und (d)). Es wird angenommen, dass sich die zusätzlichen Au-Pd-Partikel hauptsächlich in den Poren befinden, was zu einer höheren Füllung führt.

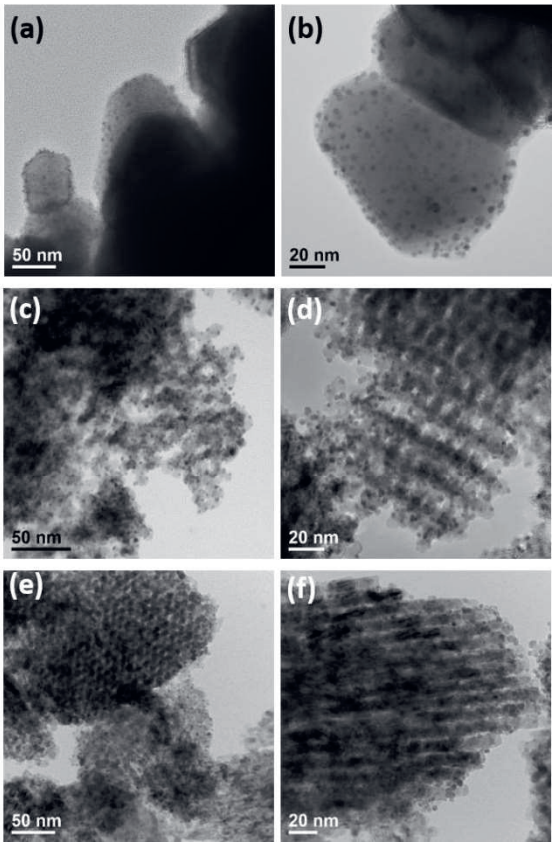


Abb. 6: TEM-Aufnahmen für drei verschiedene Tinten: (a) und (b) zeigt Au-Pd(4 Gew.%)@kommerzielles-Co_3O_4. (c) und (d) Au-Pd(4 Gew.%)@meso-Co_3O_4. (e) und (f) Au-Pd(12 Gew.%)@meso-Co_3O_4.

Die Elementverteilung von Au, Pd, Co und O im Au-Pd(12 Gew.%)@meso-Co_3O_4-System wurde unter Verwendung eines STEM-EDX analysiert (Abb. 7). Die Elementverteilungen von Au und Pd überlappen sich, was auf eine scheinbare Phasentrennung hindeutet. Darüber hinaus sind die Au-Pd-Partikel auf der Oberfläche des Meso-Co_3O_4 ungefähr gleichmäßig verteilt, woraus man schließen kann, dass das Au-Pd keine Mischung, sondern eine Bimetalllegierung aus Au und Pd ist.

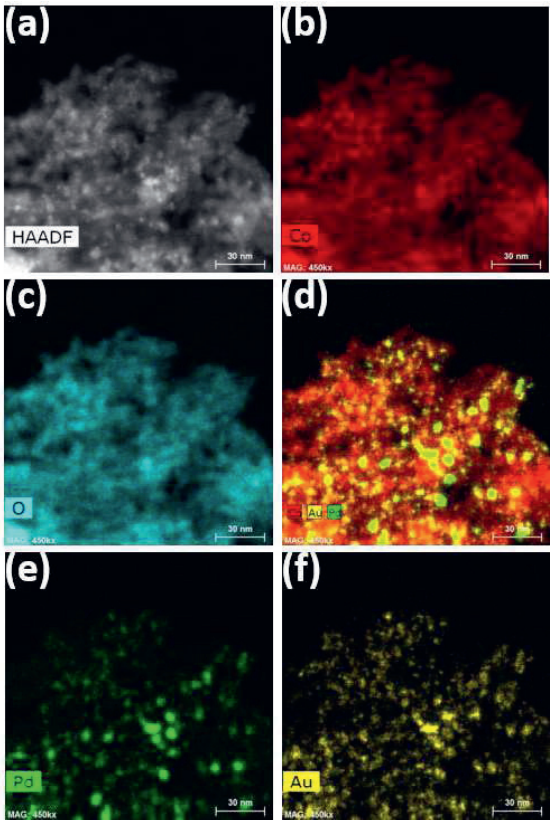


Abb. 7: EDX-Analyse des Au-Pd (12 Gew.%)@meso-Co_3O_4. (a) HAADF-Bild. In (b), (c), (e) und (f) sind die Verteilungen von Co, O, Pd bzw. Au zu sehen. (d) zeigt die Verteilung der drei Metalle: Co, Au und Pd.

Für die gasabhängigen Charakterisierungen wurden Pellistoren mit 15 Lagen aus Au-Pd(12 Gew.%)@meso-Co_3O_4 ausgewählt. Die Messungen fanden mittels einer Wheatstone-Brücke statt. Die Messgröße ist die Brückenspannung. Als Referenzsensor wurde dabei ein leeres Keramiksubstrat mit Pt-Heizer verwendet. Ein weiteres Keramiksubstrat, das mit den katalytischen Funktionsmaterialien beschichtet ist, dient als aktiver gassensitiver Sensor. Die Messungen wurden bei Methankonzentrationen von 1 %, 0,7 %, 0,5 %, 0,3 %, 0,1 % in synthetischer Luft bei 300 °C bzw. 250 °C durchgeführt. Die Gasmess–ergebnisse sind in Abb. 8 dargestellt. Mit Ausnahme eines kleinen Signals bei 250 °C für 0,1 % sind sämtliche Sensorantworten auf die unterschiedlichen Konzentrationsstufen bei beiden Temperaturen deutlich zu sehen. Da bei einer Arbeitstemperatur von 300 °C schon ein deutliches Sensorsignal zu beobachten ist, ist

von einem so aufgebauten Pellistor zu erwarten, dass er erheblich weniger anfällig gegenüber Vergiftungen ist und somit eine höhere Lebensdauer besitzt.

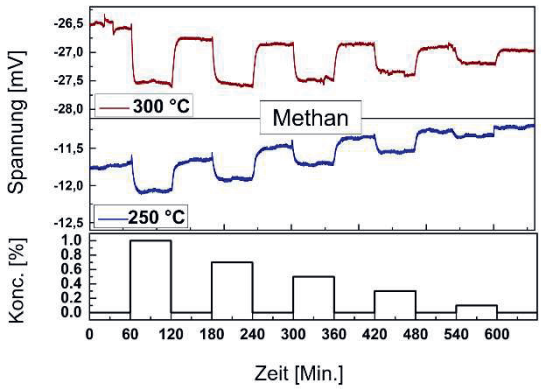


Abb. 8: Gasmessungen eines Pellistors mit 15 Lagen Au-Pd(12 Gew.%)@meso-Co_3O_4. Der Pellistor wurde mit Methankonzentrationen von 1 %, 0,7 %, 0,5 %, 0,3 % und 0,1 % in synthetischer Luft beaufschlagt. Die Betriebstemperaturen betrugen 300 °C und 250 °C.

Eine Analyse von Gasmessungen mit Pellistoren, die mit allen Tinten beschichtet wurden, welche auch mittels TEM analysiert wurden, ist in Abb. 9 dargestellt. Die Sensorreaktion des Pellistors wird anhand der Differenz der Brückenspannungen mit und ohne CH_4-Beaufschlagung bewertet. Das Sensorsignal, des mit Meso-Co_3O_4 beschichteten Sensors im Vergleich zu dem mit kommerziellen Co_3O_4 beschichteten Sensors ist ca. eine Größenordnung größer. Die höhere Bedeckung mit Au-Pd-Partikeln auf dem Meso-Co_3O_4 führt zu einer starken Verbesserung der Pellistoreneigenschaften bei verschiedenen Methankonzentrationen.

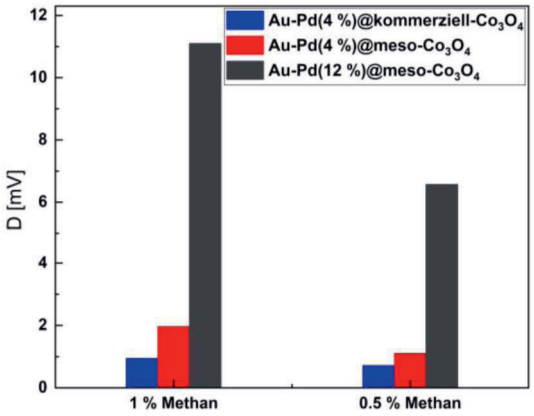


Abb. 9: Datenanalyse der Ergebnisse von Gasmessungen bei 400 °C Betriebstemperatur mit Pellistoren, die mit drei verschiedenen Tinten hergestellt wurden: Au-Pd(4 Gew.%)@kommerzielle-Co_3O_4 (schwarz), Au-Pd(4 Gew.%)@meso-Co_3O_4 *(rot) und Au-Pd(12 Gew.%)@meso-Co_3O_4 (blau).*

Der Einfluss der Anzahl der abgeschiedenen Schichten auf die Sensoreigenschaften wurde anhand von Pellistoren aus Au-Pd(12 Gew.%)@meso-Co_3O_4 und der Reaktion mit 1 % CH_4 untersucht (Abb. 10). Das Ergebnis zeigt, dass bei einer Schichtzahl von 15 das Sensorsignal am höchsten ist. Das Sensorsignal bei 15 Schichten ist deutlich größer als bei 5 Schichten, da mehr Schichten mehr aktive Stellen bieten, die eine effizientere Methanverbrennung gewährleisten. Im Gegensatz dazu verringert sich das Sensorsignal bei 30 Schichten im Vergleich zu 15 Schichten, ist aber noch deutlich größer als bei 5 Schichten. Dies ist wahrscheinlich auf den Temperaturgradienten vom Boden des Substrats zur Oberfläche der Schichten zurückzuführen, was zu einer verringerten Oberflächentemperatur führt. Der Effekt des Temperaturgradienten dominiert gegenüber dem der erhöhten aktiven Zentren. Der Einfluss der Anzahl der funktionalen Schichten auf die Sensoreigenschaften ist temperaturabhängig. Bei niedrigeren Temperaturen ist der Einfluss der Schichtanzahl auf die Sensorreaktion geringer.

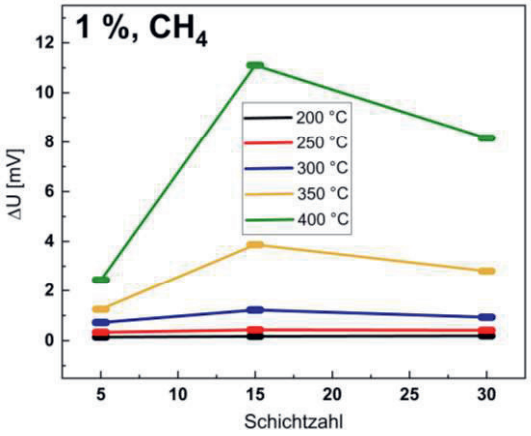


Abb. 10: Einfluss der Anzahl der abgeschiedenen Schichten aus Au-Pd(12 Gew.%)@meso-Co_3O_4 auf die Pellistoreigenschaften.

Zusammenfassung und Ausblick

Das Materialsystem mesoporöses Co_3O_4 als Trägermaterial mit einem bimetallischen Au-Pd-Katalysator wurde auf seine Eignung als funktionales Material von Pellistoren zum Niedertemperaturnachweis von Methan untersucht. Es konnte gezeigt werden, dass eine Reduktion der Arbeitstemperatur des Pellistors auf 300 °C möglich ist, wodurch mögliche dauerhafte Vergiftungen durch Si-haltige organische Verbindungen vermieden werden können. Das untersuchte Material besitzt einen dreidimensionalen mesoporösen Aufbau, was im Vergleich zu Au-Pd@kommerziell-Co_3O_4 zu verbesserten Pellistoren-Eigenschaften führt. Ein hoher Anteil an Au-Pd-Legierungspartikeln wirkt sich ebenfalls positiv auf das Pellistorenverhalten aus. Die Anzahl der abgeschiedenen Schichten bzw. die Schichtdicke hat ebenfalls Einfluss auf die Sensoreigenschaften. Die Absenkung der Arbeitstemperatur unter Verwendung einer mesoporösen 3D-Basis in Kombination mit einem Edelmetall-Nanokatalysator verbessert zudem den Explosionsschutz entscheidend und verringert die Heizleistungsaufnahme der Pellistoren.
Langzeitmessungen und Untersuchungen mit anderen Materialsystemen werden derzeit durchgeführt. Es wird erwartet, dass dies den Weg zur Entwicklung robuster Pellistoren mit verbesserten Eigenschaften ebnen wird.

Danksagung

Die Autoren danken Dr. Yuechen Jia für die REM-Analyse, Dr. Marcel Werner für DLS-Messungen, Dr. Thilo Ludwig für XRD, Andreas Kürzinger für die Montage der Pellistoren und Laura Engel für die Unterstützung bei der Tintenabscheidung. Die TEM-Untersuchungen wurden von Patrick Diehle am Fraunhofer IMWS in Halle durchgeführt. Die durchgeführten Arbeiten wurden zum Teil von der Fraunhofer Gesellschaft im Rahmen des Projektes FLUMEMS finanziert.

Literaturverzeichnis

[1] Y. Wang, M. M. Tong, D. Zhang, and Z. Gao, *Sensors*, vol. 11, no. 1, pp. 19–31, (2010); doi: 10.3390/s110100019
[2] J. B. Miller, *IEEE Sens. J.*, vol. 1, no. 1, pp. 88–93, (2001); doi: 10.1109/JSEN.2001.923591
[3] M. Preißinger, and D. Brüggemann, "Thermal Stability of Hexamethyldisiloxane (MM) for High-Temperature Organic Rankine Cycle (ORC)," *Energies*, vol. 9, no. 3, p. 183, (2016); doi: 10.3390/en9030183
[4] Z. Wu *et al.*, "Three-dimensionally ordered mesoporous Co_3O_4-supported Au–Pd alloy nanoparticles: High-performance catalysts for methane combustion," *J. Catal.*, vol. 332, pp. 13–24, (2015); doi: 10.1016/j.mcat.2017.09.002
[5] F. Kleitz, S. Hei Choi, and R. Ryoo, "Cubic *Ia3d* large mesoporous silica: synthesis and replication to platinum nanowires, carbon nanorods and carbon nanotubes," *Chem. Commun.*, vol. 0, no. 17, p. 2136, (2003); doi: 10.1039/B306504A

DOI 10.5162/sensoren2019/6.3.1

Precise Viscosity and Density Sensing in Industrial and Automotive Applications

Thomas Voglhuber-Brunnmaier[1], Alexander O. Niedermayer[2], Friedrich Feichtinger[1], Bernhard Jakoby[1]
[1]Johannes Kepler University, Altenbergerstr. 69, 4040 Linz, Austria
[2]Micro Resonant Technologies, Hafenstr. 47-51, 4020 Linz, Austria

Abstract

A resonant sensor based system for measuring viscosity and density is shown and applied to two measurement tasks comprising oil characterization in hydraulic systems and measurement of diesel and soot contamination in engine oil. It is demonstrated that from measuring over varying temperature and pressure, a variety of additional physical properties can me exploited for online condition monitoring. This encompasses also the heat expansion coefficient and the bulk modulus, which are difficult to measure otherwise.

Keywords: condition monitoring, resonant sensor, bulk modulus, engine oil dilution

Introduction

Besides the measurement of liquids in production processes, sensors for monitoring lubricating liquid properties are used in order to retrieve information on the condition of the lubricant and the machinery. With such condition monitoring processes, it is possible to shift maintenance policies from a scheduled towards a predictive maintenance. Doing this reduces the maintenance costs by minimizing machine shutdowns for inspection and service, and furthermore, significantly reduces the risk of catastrophic component failure during regular operation and resulting repair downtime [1].

The traditional approach of analyzing oil samples in a laboratory on regular schedule typically gives reliable results but implies a lag in information due to long sampling intervals, shipping and processing time and furthermore, possible irregularities in taking or handling the sample. Another aspect of the bottle sampling approach that is hardly considered is that the sampling schedule often is not synchronized with the duty cycles on the assets, resulting in systematic errors due to violation of the Nyquist-Shannon theorem. Online sensors for monitoring oil properties are capable of increasing the frequency of the tests significantly while reducing total cost of ownership. Furthermore, quick and adequate action can be taken in case of rapid changes due to failures (like water leaking into the oil system).

Compared to offline analyses in a lab, online sensors can only determine a rather limited number of different parameters. That is one of the reasons why some manufacturers try to overcome this drawback by implementing data based oil-ageing models that are configured with reference data from experiments with the very same oil. The data that is provided by such sensors may yield good results in applications where the type of oil is precisely known and the operating conditions of the machinery do not differ from that of the test machines. Under these circumstances the reliability of the condition monitoring is mainly determined by the training data; hence, the confidence of information obtained with this type of sensors will be relatively poor with respect to unexpected changes in the fluid (i.e. caused by failure of components, contamination with alien substances or erroneous refill of wrong fluid).

With respect to the claim that online condition monitoring should increase the reliability of an asset, the confidence of the sensor data as well as the reliability of the condition monitoring system are crucial parameters. Unreliable sensors may provide inconsistent data that trigger false alarms, resulting in unnecessary troubleshooting or even shutdowns and, if occurring frequently, may very likely lead to personnel ignoring alarms. Ultimately, this may result in even higher costs than without any online condition monitoring system at all [2].

The measurement system shown in this contribution fulfills stringent requirements in measurement precision and long-term stability. A high level of confidence can be achieved by implementing methods to constantly check sensor data against physical models of liquid properties, such as temperature dependence of viscosity and density, for example.

Measurement Setup

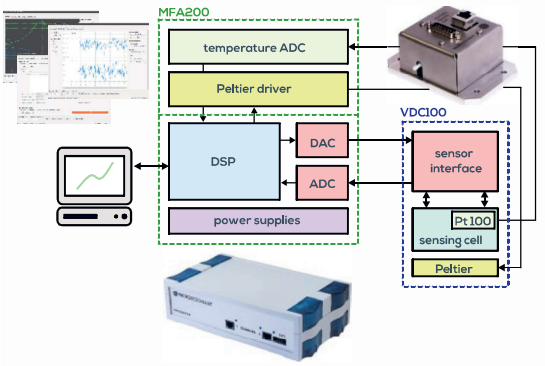


Fig. 1: The measurement setup consists of the temperature controlled measurement cell VDC100 and the universal resonance analyzer MFA200.

The measurement system that was used for the experiments shown in this contribution comprises a flow-through measurement cell VDC100 in combination with the universal resonance analyzer MFA200. Measured viscosity and density parameters are sent to a PC and analyzed in real-time (see Fig. 1). The system was developed by the University Spin-off company Micro Resonant in collaboration with the Institute for Microelectronics and Microsensors at the Johannes Kepler University.

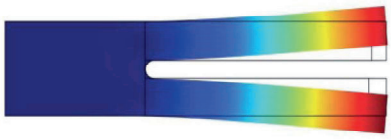


Fig. 2: Fundamental vibration mode of the quartz tuning fork (QTF) used for measurement.

The temperature of the sample volume in the cell is controlled with a thermoelectric element (Peltier) and a PT100 resistive temperature sensor. The physical liquid properties are determined using a piezoelectric quartz tuning fork (QTF) resonator as sensor. In order to obtain the resonance characteristics (i.e. resonance frequency and the *Q*-factor) of the sensor, the QTF is excited with a digitally generated waveform. The frequency components in this waveform are chosen such that the information gathered from the resonator at these frequencies is maximized. This approach enables the use of dedicated signal processing algorithms for suppressing unwanted influences (like offset signals and gain errors) of the analog signal path [3]. For the QTF the Butterworth-Van Dyke equivalent circuit is a usable approximation. The shunt capacitance C_0 is formed by the quartz material and the electrodes, and the RLC series network represents the electromechanical interaction and is termed *motional branch*. We consider only the fundamental vibration model shown in Fig. 2, but in principle for each mode of vibration, another set of component values would be required.

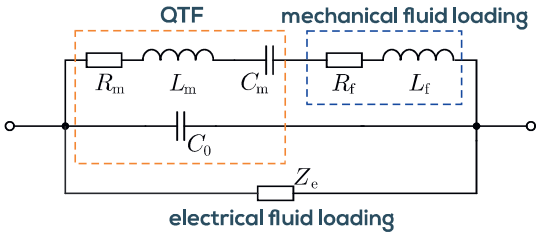


Fig. 3: The extended Butterworth-Van Dyke model describing the fluid loaded quartz tuning fork.

When the QTF vibrates in liquid, the prongs of the QTF periodically displace fluid mass, and therefore, the liquid density adds mass to the equivalent mass, which is reflected in the additional inductance L_f. The dissipated energy due to viscous drag is accounted for by a resistor R_f in series to the motional branch. The permittivity and conductivity of the fluid also influence the measured impedance. If highly conductive fluids are measured, an insulating layer must be applied to avoid high shunt currents which would reduce the signal-to-noise ratio. This layer, as well as electrochemical double layer effects are summarized in the impedance Z_e. The effects of C_0 and Z_e are separated from the measured impedance spectra by the signal processing in the resonance analyzer instrument shown in Fig. 1. Only the resonance parameters of the extended motional branch are analyzed, which enhances the stability of and accuracy of the density and viscosity measurement substantially.

Advantages of using tuning fork sensors

Compared to alternative viscosity sensors, e.g., based on thickness shear mode resonators, shear horizontal or Love wave sensors (see

e.g., [4]), using a tuning fork has distinct advantages which are relevant especially in field applications. Due to the low operation frequency of the QTF (here: 25-30 kHz), the results are more comparable to laboratory instruments, because viscoelastic effects of fluids are less pronounced [5]. The penetration depth of shear waves scale with $f^{-1/2}$ and is therefore larger at low frequencies. Furthermore, the penetration depth δ, which usually limits the spatial sensing range of pure shear sensors [6], is not a limiting factor when using QTFs. Given the vibration mode in Fig. 2, the flow profile mainly consists of a potential flow field due to the fluid displacement. The spatial decay of the flow field of the QTF follows approximately r^{-2} instead of $\exp(-r/\delta)$ of a plane shear wave. As was demonstrated in [7], watery suspensions containing spherical particles of 40 µm diameter can be measured equally well as smaller particles or clear fluids, although δ is 3.4 µm and therefore much less than the particle size. Consequently, a good spatial average is obtained when measuring fluids with certain microstructure such as emulsions (see e.g. [8]) or suspensions, giving results closer to larger laboratory instruments. Furthermore, also the impact of surface contaminations due to oil additives, soot, or air microbubbles is reduced significantly. Another benefit of QTF resonators is the fact, that viscosity and density can be measured simultaneously. When using pure shear wave sensors only the viscosity-density product can be measured (except by implementing a more complex design such as adding structured surfaces [9]).

Condition Monitoring in Industrial Applications

The viscosity value of an oil used in a hydraulic or lubricating system is one of the dominant design parameter and therefore, monitoring oil condition by evaluating viscosity is an advantageous approach. Furthermore, viscosity reacts to contamination very sensitively, as is shown in [10,11], where an engine-oil dilution with 0.05% diesel is reliably detected within seconds. In the described application, the density and viscosity functions of hydraulic oil versus temperature and pressure are measured and additional physical monitoring parameters are derived from the characteristics.

Fig. 4 shows the viscosity of a hydraulic oil versus temperature at various pressures measured with the system in Fig. 1. The high sensitivity of viscosity to temperature is apparent and underpins the need for an accurate temperature control. The temperature characteristics as a new monitoring parameter can be is obtained by measuring at various defined temperatures.

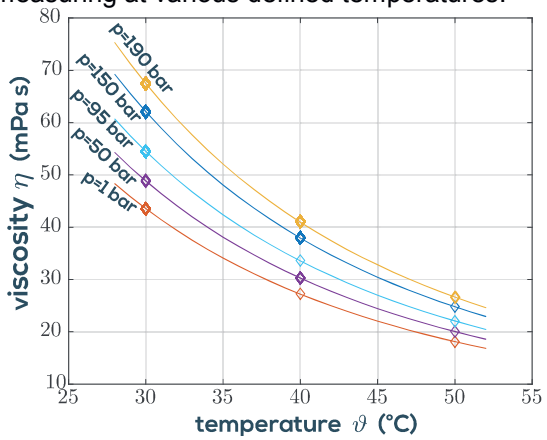


Fig. 4: Viscosity of hydraulic oil over temperature for various pressures.

The solid lines in Fig. 4 follow the Vogel-Cameron Model (DIN 53017) [12] which is given by

$$\eta(\vartheta) = A\, e^{\left(\frac{B}{\vartheta + C}\right)}, \qquad (1)$$

where ϑ is the temperature in °C, C=95°C for mineral oils and A and B are fluid specific constants. Besides the strong temperature dependence, liquid viscosity is influenced by pressure. The sensitivity of viscosity to pressure is demonstrated in Fig. 5. Here, the measurement cell (equipped with Minimess connectors) was implemented in industrial environment at a high pressure supply rail of an operating machine.

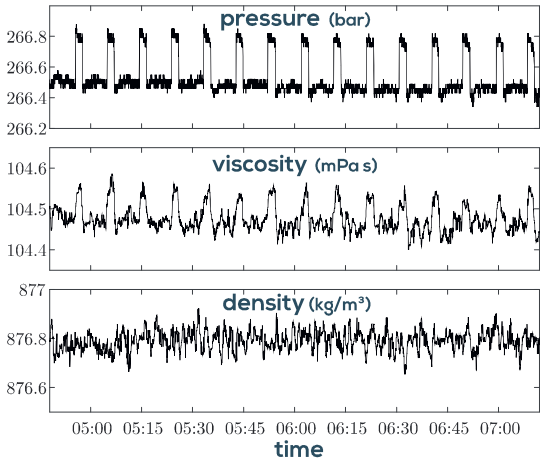


Fig. 5: Small pressure pulses influence viscosity noticeably, but not density.

The cyclic pressure drops of 0.3 bar at 266 bar caused by that machine are clearly visible in the

viscosity. The viscosity versus pressure is shown in Fig. 6. The dashed lines follow the exponential law given in Eq. (2) with the fitting parameters in Tab. 1.

$$\eta(p) = \eta_0 \, e^{\alpha \, p} \qquad (2)$$

Here, η_0 is the viscosity at ambient pressure and the viscosity-pressure coefficient α is an additional monitoring parameter.

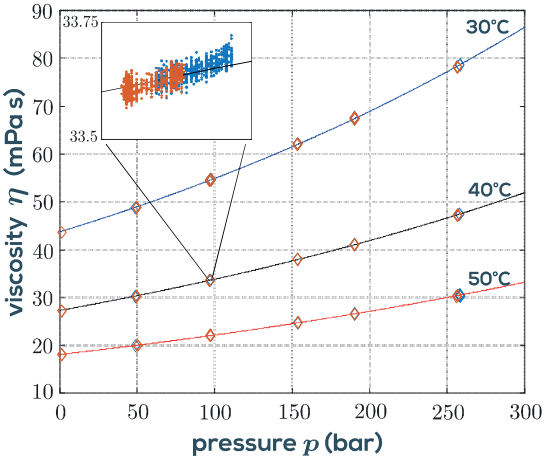


Fig. 6: Viscosity of hydraulic oil over pressure at three different temperatures. The inset shows the noises of two measurement cycles performed at separate days.

Tab. 1: Calculated ambient viscosity and viscosity-pressure coefficient following Eq. (2).

ϑ (°C)	η_0 (mPas)	α (10^{-3}/bar)
30	43.75	2.272
40	27.33	2.139
50	18.14	2.016

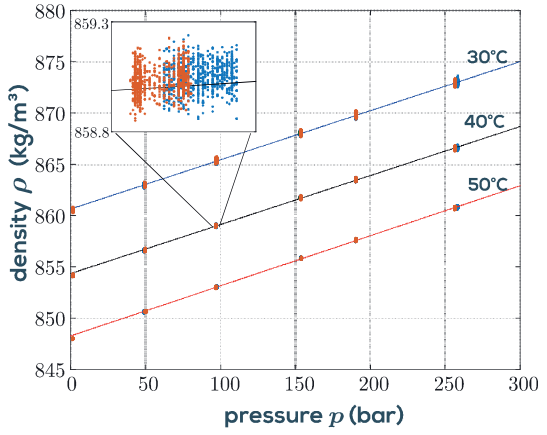


Fig. 7: Density of hydraulic oil over pressure. at three different temperatures, simultaneously measured with viscosity in Fig. 6.

While temperature sweeps can directly be performed by the measurement cell, for pressure variation an external pressure regulator is required. Small variations in temperature (ΔT) and pressure (Δp) also cause a change in density

$$\frac{\Delta\rho}{\rho} \approx \underbrace{\frac{1}{\rho}\left(\frac{\partial\rho}{\partial p}\right)_T}_{\beta_T} \Delta p + \underbrace{\frac{1}{\rho}\left(\frac{\partial\rho}{\partial T}\right)_p}_{-\gamma_p} \Delta T \qquad (3)$$

The two additional monitoring parameters in Eq. (3) are the isothermal compressibility β_T and the isobaric heat expansion coefficient γ_p. The inverse of the isothermal compressibility is also known as the isothermal bulk modulus K_T. In Tab. 1, the approximate ranges of bulk modules for three pressure fluids (from [12]) are listed.

Tab. 2: The range of the bulk modulus K_T for oils measured [12].

$(1.4 \ldots 1.6) \cdot 10^4$	air-free mineral oil
$(1.0 \ldots 1.2) \cdot 10^4$	mineral oil with air
$(2.3 \ldots 3.5) \cdot 10^4$	non-mineralic high-pressure fluids.

In Tab. 3 and 4 these coefficients derived from the measurements in Fig. 7 are shown. It has to be emphasized that the derivatives in Eq. (3), are sensitive to the fitting model. It is therefore advisable to record more temperature points than are shown in Fig. 4. Due to the short thermal time constants of the measurement cell of around 1-2 minutes per temperature point, this is not a limiting factor.

As indicated in Tab 1, the amount of dissolved air can be estimated on basis of the bulk modulus. Its measurement is therefore of interest for hydraulic system engineers. In this field three approaches for measuring the bulk modulus are commonly used [13]: i) the volume change method, ii) the flow change method and iii) sound wave propagation. Measurement of the bulk modulus by evaluating the local density variation using the micro sensor has certain advantages over these methods. In the volume change method (i), where the reduction of a defined volume is measured versus applied pressure, or for the flow change method (ii) where a small volume is injected into the chamber and pressure is measured, the finite rigidity of the chamber introduces errors. Furthermore, measurements are meaningful only at high pressures where all entrapped air bubbles are dissolved [14]. The technical realizations of these principles are therefore generally complicated,

expensive and require significant calibration efforts [15]. Due to the larger observed volumes, sustaining thermal equilibrium in the entire volume is also more demanding. From acoustic propagation measurements (iii) the speed of sound and consequently the adiabatic bulk modulus K_A can be determined using Eq. (4).

$$K_A = c^2 \rho \tag{4}$$

It is emphasized that the adiabatic bulk modulus K_S differs from K_T by ca. 10-20% for mineral oil. They can be interrelated, but specific heat capacitance, heat expansion coefficient, density and temperature are required [15]. Nevertheless, speed of sound is a parameter of interest on its own, because it characterizes pressure wave propagation in high speed switching hydraulics [16].

Tab. 3 isobaric heat expansion coefficient at ϑ = 40°C for different pressures

p (bar)	1	50	100	190
γ_p (10^{-4} K^{-1})	7.34	7.19	7.1	7.095

Tab. 4 isothermal compressibility and bulk modulus at p=100 bar for different temperatures.

ϑ (°C)	30°C	40°C	50°C
β_T(GPa^{-1})	0.565	0.575	0.588
K_T (GPa)	1.77	1.74	1.70

Oil condition monitoring in automotive applications.

The diesel particulate filters (PDF) used in vehicles require periodic regeneration to prevent clogging. For this purpose, one approach is by exhausting late post-injected unburnt diesel is to the filter to burn off the accumulated soot. Oil dilution occurs when part of the late injected fuel is wiped into the crankcase by the piston rings and thus into the engine oil (see Fig. 8). The addition of biodiesel aggravates this issue, because the boiling point of biodiesel is higher than that of conventional diesel [17]. Therefore, biodiesel evaporates at the cylinder walls at a lower rate. Furthermore, biodiesel has strong chemical interactions with various additives in engine oil and thus hinders, for example, the cleaning effect of the oil. Therefore, there are efforts by engine developers to minimize fuel entry by optimizing late-injection schemes. This happens, for example, by adjustments of the number and duration of the individual injections, as is shown in 5. In order to be able to assess the effect of different settings faster than by using off-site laboratory analyzes of oil samples, an online monitoring system measuring viscosity and density can be used.

Although, the densities of diesel and engine oil are very similar, the viscosities are distinctly different. While diesel-addition lowers the viscosity of the oil, it is increased by an unavoidable entry of combustion residues (soot).

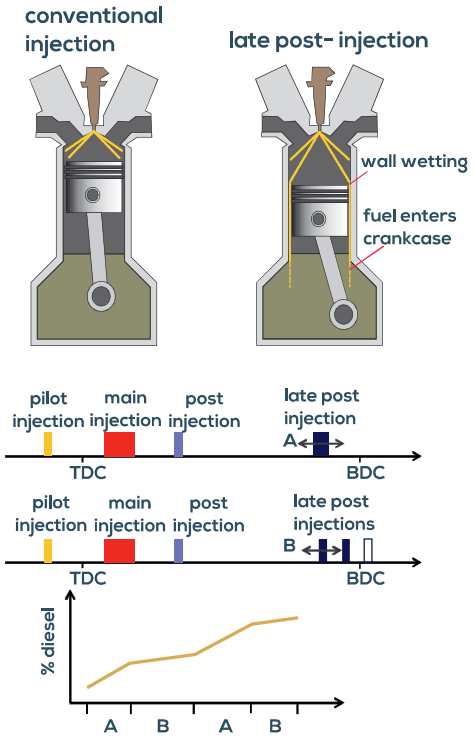


Fig. 8: Late post injections for particle filter regeneration causes gradual oil dilution [17]. Optimization of injection timings can reduce this effect, but requires close monitoring of diesel content.

It is therefore required to separate the effects of soot and diesel which requires an additional parameter to be measured. As soot increases the density of the oil, it is possible to calculate the diesel and soot content by simultaneously measuring density and viscosity. Fig. 9 shows the effects of diesel and soot entry into engine oil for two different operating cycles A and B. Fig. 10 shows the test setup, which is used to assess the general performance of the measurement system. For the experiment, an oil sample of 200 ml was prepared consisting of 50 ml used oil (containing 5 % diesel and an unknown soot content) and 150 ml new oil of the

same batch. The overall diesel content is therefore 1% and the soot content is 0.2 used-oil-soot equivalents. Computer controlled syringe pumps add two mixtures A and B at a rate of 100 µl/min to the mixture which is continuously stirred to guarantee homogeneity. A sample is pumped into the measurement cell in intervals of 8 minutes and measurements are taken after reaching thermal equilibrium.

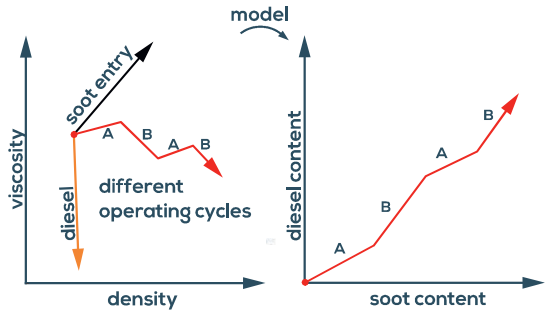


Fig. 9: Diesel and soot have different effects on viscosity and density and can therefore be calculated by a mixture model.

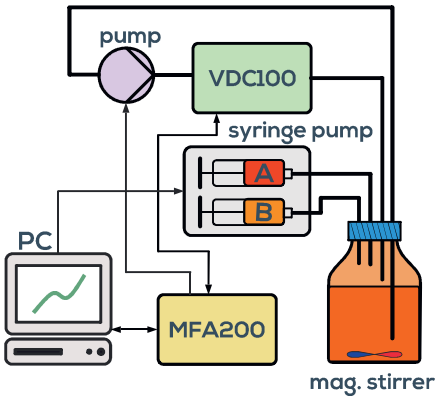


Fig. 10: Schematic of the laboratory test setup used to asses to achievable performance of the system.

The concentrations are therefore known throughout the experiment and can be compared with the estimated values in order to assess the achievable performance. Fig. 11 and Fig. 12 show viscosity and density. A linearized mixing model, valid for low foreign substance concentrations is used to determine the diesel and used-oil soot equivalents in Fig. 13 and Fig. 14, respectively. The entry of substances is stopped for approximately 3 h indicated by the plateaus. No significant drifts are observed and the actual diesel and soot equivalents are shown as dashed lines. The agreement and sensitivity of the system suggests that the system is suitable for this task and reaching an accuracy of 0.1% diesel content is feasible.

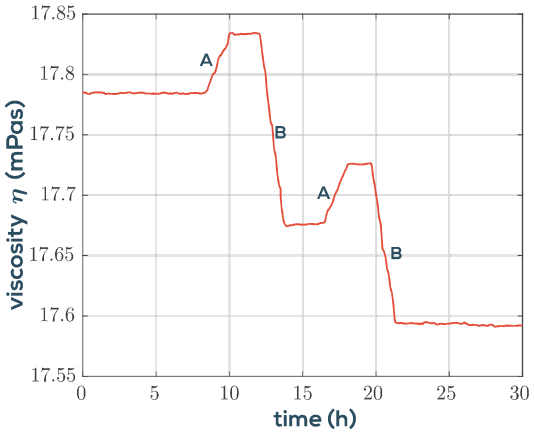


Fig. 11: Viscosity vs. time obtained with the setup in Fig. 10.

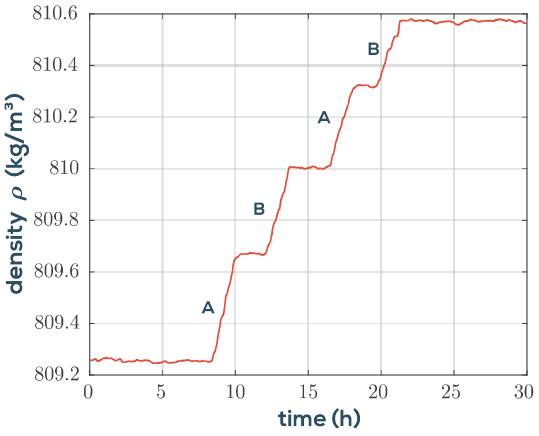


Fig. 12: Density vs. time.

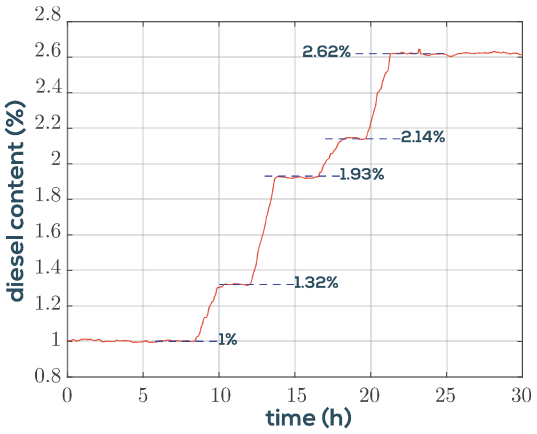


Fig. 13: Measured diesel content vs. time. The dashed lines represent the real values known throughout the experiment.

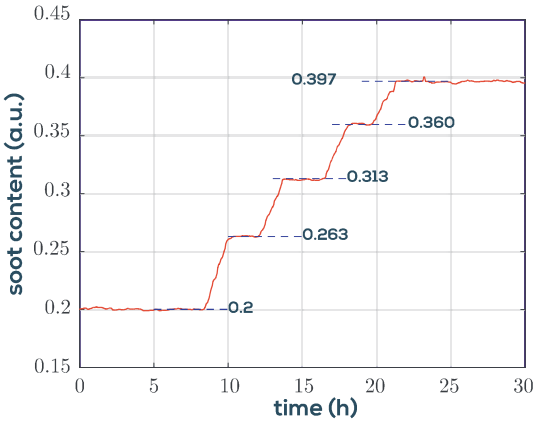


Fig. 14: Soot equivalents vs time. The dashed lines represent the real values.

Conclusions

The ability of the described online condition monitoring system to measure viscosity and density over a large temperature and pressure range yields additional monitoring parameters such as the bulk modulus, the heat expansion coefficient, but also the V-I index and the pressure-viscosity functions. These parameters are for instance relevant for the determination of aging of lubricating oils or dissolved air in hydraulic systems. Furthermore, the applicability of the system to measure diesel and soot concentrations in engine oil was demonstrated on basis of simultaneous viscosity and density sensing. The results indicate high sensitivity and accuracy.

References

[1] A. Davies, Handbook of Condition Monitoring: Techniques and Methodology, Springer, 1989 doi: 10.1007/978-94-011-4924-2

[2] M. Blanke et.al., Diagnosis and Fault-Tolerant Control, 3rd Edition, Springer, (2016) doi: 10.1007/978-3-662-47943-8

[3] A.O. Niedermayer, T. Voglhuber-Brunnmaier, M. Heinisch, B. Jakoby, Accurate Determination of Viscosity and Mass Density of Fluids using a Piezoelectric Tuning Fork Resonant Sensor, Proceedings SENSOR 66-71. (2015), doi: 10.5162/sensor2015/A2.4

[4] B. Jakoby, M. J. Vellekoop, Viscosity sensing using a Love-wave device. Sensors and Actuators A: Physical, 68(1-3), 275-281 (1998) doi: 10.1016/S0924-4247(98)00017-X

[5] L. Matsiev, J. Bennett, E. McFarland, Application of low frequency mechanical resonators to liquid property measurements. IEEE Ultrasonics Symposium. Proceedings IEEE, 1, 459–462. (1998) doi: 10.1109/ULTSYM.1998.762188

[6] B. Jakoby, R. Beigelbeck, F. Keplinger, F. Lucklum, A.O. Niedermayer, ... B. Weiss, Miniaturized sensors for the viscosity and density of liquids-performance and issues. IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 57(1), 111-120. (2010) doi: 10.1109/TUFFC.2010.1386

[7] T. Voglhuber-Brunnmaier, E. K. Reichel, E. K., A. O. Niedermayer, F. Feichtinger, J.K. Sell, B. Jakoby, Determination of particle distributions from sedimentation measurements using a piezoelectric tuning fork sensor. Sensors and Actuators A: Physical, 284, 266-275. (2018), doi: 10.1016/j.sna.2018.10.038

[8] B. Jakoby, M. J. Vellekoop, Physical sensors for water-in-oil emulsions. Sensors and Actuators A: Physical, 110, (1-3), 28-32. (2004). doi: 10.1016/j.sna.2003.08.005

[9] S. J. Martin, K. O. Wessendorf, C. T. Gebert, G. C. Frye, R. W. Cernosek, L. Casaus, M. A. Mitchell, Measuring liquid properties with smooth-and textured-surface resonators. IEEE international frequency control symposium, 603-608. (1993), doi: 10.1109/FREQ.1993.367450

[10] A.O. Niedermayer, T. Voglhuber-Brunnmaier, F. Feichtinger, M. Heinisch, B. Jakoby, Online Condition Monitoring of Lubricating Oil based on Resonant Measurement of Fluid Properties. Sensors and Measuring Systems; 19th ITG/GMA-Symposium VDE. 1-4 (2018), isbn: 978-3-8007-4683-5

[11] T. Voglhuber-Brunnmaier, A.O. Niedermayer, F. Feichtinger, M. Heinisch, B. Jakoby, High Precision Resonant Sensor Evaluation with Application to Fluid Sensing, Sensors and Measuring Systems. Sensors and Measuring Systems; 19th ITG/GMA-Symposium VDE. 355-358 (2018), isbn: 978-3-8007-4683-5

[12] D. Will, N. Gebhardt, (Eds.). Hydraulik: Grundlagen, Komponenten, Systeme. Springer-Verlag. (2015), doi: 10.1007/978-3-662-44402-3

[13] C. Ding, C., H. Murrenhoff, Accuracy estimation of a novel setup for online bulk modulus measurement in hydraulic systems. O+P Ölhydraulik und Pneumatik 3, (2016)

[14] J. Wang, G. Gong, H. Yang, Control of bulk modulus of oil in hydraulic systems. IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 1390-1395 (2008), doi: 10.1109/AIM.2008.4601865

[15] R. Haas, Material behavior of hydraulic fluids under the influence of thermoelasticity and entrained air. Trauner. (2013), isbn: 978-3990331774

[16] R. Scheidl , H. Kogler, B. Winkler, Hydraulic switching control-objectives, concepts, challenges and potential applications. Hidraulica, 1, (2013), issn: 1453-7303

[17] D. Uy, R. J. Zdrodowski, A. E. O'Neill, S. J. Simko, A. K. Gangopadhyay, M. Morcos, ... & G. Parsons. Comparison of the effects of biodiesel and mineral diesel fuel dilution on aged engine oil properties. Tribology Transactions, 54(5), 749-763. (2011), doi: 10.1080/10402004.2011.597545

DOI 10.5162/sensoren2019/6.3.2

Accuracy Study of a Torsionally Oscillating Pipe Viscometer

S. Clara[1]*, F. Feichtinger[1], T. Voglhuber-Brunnmaier[1,2], A.O. Niedermayer[1] und B. Jakoby[1]
[1]Institute for Microelectronics and Microsensors
[2]Linz Institute of Technology
Altenbergerstr. 69, 4040 Linz, Österreich

Abstract

We recently presented a robust viscosity measurement system based on a torsionally oscillating pipe. This resonant sensor consists of two concentrically aligned pipes. The outer pipe acts as an oscillating torsion spring while the inner one remains at rest. At both ends of the pipe flywheels are located, oscillating against each other in the torsional mode. The pipe is filled with a fluid, which is causing additional damping related to its viscosity. In this contribution we present the sensor setup, describe the measurement principle and analyze the viscosity measurement accuracy achieved in test measurements for different viscous fluids.

Keywords: viscosity, flow-through, resonant, balanced, torsional oscillation.

Introduction

Several resonant viscosity sensors have been presented over the last years, see, e.g., [1] and [2]. Most of these sensors are working in the higher frequency regime above 20 kHz or even in the MHz regime (see [3] and [4]). These high resonance frequencies result in a low penetration depth and therefore a very thin fluid layer being sensed directly on the surface of the sensor which leads to spurious influences due to surface contamination. In this contribution we present an accuracy study of our recently presented viscosity sensor based on a torsionally oscillating pipe [5] and [6] with a low resonance frequency around 7.7 kHz. The sensor setup is almost completely machined in aluminum and is actuated and read out electromagnetically. This relatively large and robust sensor can be used as a flow-through sensor due to its special aligned pipe system.

Setup

The main components of this sensor-system are two concentrically aligned pipes (Figure 1.3 and Figure 1.4) suspended with a mounting spring in the center (Figure 1.2). This mounting spring is located at the position of the nodal point of the torsional oscillation mode, which reduces the influence of the mounting structure on the oscillation drastically. The outer pipe has two flywheels (Figure 1.10) at its ends.

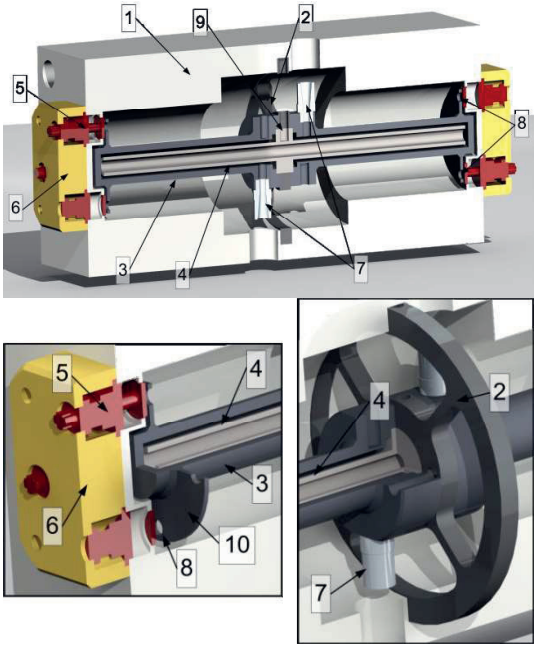


Figure 1: Sensor setup: 1) housing, 2) mounting spring, 3) outer pipe, 4) inner pipe, 5) actuation coils, 6) coil housing, 7) in- outlet, 8) drills for permanent magnets, 9) drill for temperature sensor and 10) fly wheels.

The dimensions of these flywheels combined with the thickness of the outer pipe allow easy control of the resonance frequency of the oscillator (Note: by changing these dimensions also the sensitivity of the sensor is influenced).

Each of these flywheels has four holes (Figure 1.8) fitted with permanent magnets. Next to the permanent magnets, electro-magnetic coils (Figure 1.5) are attached to the sensor housing (Figure 1.1). These coils have an angular offset to the position of the permanent magnets and are used to actuate a torsional oscillation along the outer pipe. If the sensor is filled with a fluid through the two Luer-Lok (Figure 1.7) connectors, placed close to the mounting spring, the periodically moving inner surface of the outer pipe leads to a shear wave traveling into the fluid. Due to this interaction the oscillation of the resonator is damped. This damping is related to the viscosity η and density ρ of the fluid in the sensor. The inner pipe is not oscillating and is only used to guide the fluid through the setup and guarantee homogeneous filling with the fluid. Due to this concentric pipe arrangement, the sensor is also suitable as a flow-through sensor. The whole demonstrator is machined of aluminum except from the in-outlets and the actuation coils. Aluminum was chosen as prototype material due to its low intrinsic damping and its easy machinability.
The housing of the sensor is 70mm x 70mm x 140mm, the outer pipe D_a= 12mm, D_i= 8mm, L_a= 134mm, the inner pipe d_a= 6mm d_i= 4mm, l_i= 128mm and the flywheels are d= 30mm and h= 1mm.

Readout and Actuation

The resonator is actuated using two coils shown in Figure 1.5. Also a four coil actuation, reducing the stimulation of other modes, is possible but not necessary in the present configuration because the resonance frequencies of the different modes are well separated from the resonance frequency of the torsional mode. These coils are opposing the permanent magnets placed in the flywheels (with an angular offset). If the coils are driven accordingly a moment of force is acting on the flywheel on one side of the sensor. In the torsional oscillation mode the second flywheel is rotating against the first one, the outer pipe is twisted, but the center of the pipe, where also the mounting spring is located, represents a nodal point and is ideally not oscillating.
The oscillating flywheel, with the permanent magnets, on the other side is moving in front of the readout-coils inducing a voltage which is proportional to the angular velocity of the flywheel. This signal is used to detect the motion of the flywheel. The polarity of the permanent magnets on the readout side and the connection of the coils are carried out in a way to suppress induced voltages due to the magnetic field of the actuation-coils.

As measurement device for the readout and actuation of the sensor the MFA200 from Microresonant (www.micro-resonant.at) is used. This system stimulates the sensor with eight frequencies simultaneously and measures the response of the sensor [7]. This allows to monitor the quality factor Q and the resonance frequency f_r at a very high repetition rate of around five seconds per sample (also much higher sample rates are possible depending on the quality factor of the resonator). The used frequencies are distributed around the resonance frequency and allow the reconstruction of the frequency response of the sensor.

Modeling

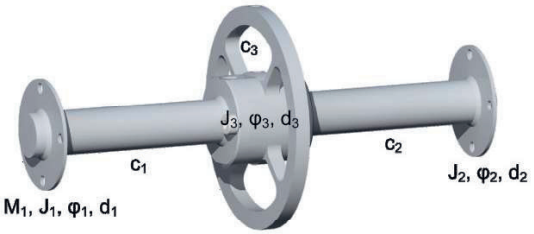


Figure 2: Sketch of a simple 1d model of the mechanical system.

To model the mechanical behavior of the system a simple one dimensional model is sufficient. Figure 2 shows the necessary parameter of a simple 1d mechanical model of the system. This model consists of three flywheels (moments of inertia J_1, J_2 and J_3), two torsional springs with the spring constants c_1 and c_2 and the mounting spring with the spring constant c_3. The deflection angles are defined as φ_1, φ_2 and φ_3.
The three resonances are clearly visible in Figure 3. The lowest resonance is the symmetric mode. In this mode all three flywheels are oscillating symmetrically against the mounting spring in the middle with the spring constant c_3. In the mode at 10.2 kHz the outer flywheels oscillate against the one in the middle of the system. The mode used for the viscosity measurements is the mode at 7.7 kHz, in this mode the outer flywheels oscillate against each other. The flywheel in the middle is in the nodal point of the oscillation and is therefore not moving. All measurements were made in this torsional mode.

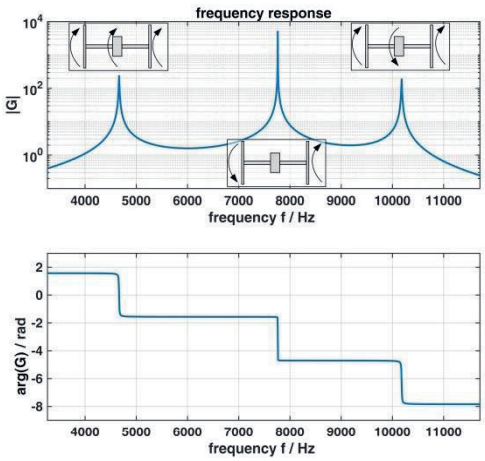


Figure 3: Resonance behavior of the 1d model. The resonance frequencies of the three possible modes are clearly visible. The torsional mode used in the measurements is located in the middle at 7700 Hz.



Figure 4: FEM simulation result for the displacement of the mechanical structure oscillating in the torsional mode (red: high deflection, blue low deflection). The resonance frequency of this mode is at 7.7 kHz.

The result of the simple model was validated using COMSOL™ a FEM program, were an eigenmode analyses for the oscillator were performed. Figure 4 shows the result of the FEM simulation of the mechanical structure oscillating in the torsional mode. The increasing deflection along the outer pipe is clearly visible. The mounting spring placed in the middle of the oscillator is not moving.
A more complex model including the interaction with the fluid was presented in [6].

Measurement Setup

Different viscosity standard oils were measured at 20±0.01°C to characterize the sensor setup and study its behavior for high and low viscosities. The used oils are listed in *Table 1*, the viscosity range is from 19.3 mPa.s to 7 Pa.s. The range was chosen to provide a large viscosity range but does not represent the limits of the sensor. As shown in [6] also lower viscous fluids with viscosities of 1.96 mPa.s and below are measurable and also higher viscous fluids with viscosities up to 21.5 Pa.s were already measured. The measurements were performed using the MFA200 analyzer from Microresonant (presented in [7] and [8] www.micro-resonant.at) and an automated sampling system. The whole setup was placed in an isolated temperature box were the temperature was controlled by a water-bath. A calibrated PT100 was used to measure the temperature of the setup; this sensor was placed in a drilling close to the center of the setup (see Figure 1.9). The setup was placed in the temperature box (Figure 5) two hours before the sampling was started to guarantee sufficient thermal settling. As the viscosity is highly temperature dependent a precise defined temperature is required for accurate viscosity measurements, e.g. a change of 0.1°C causes a viscosity change of ca. 0.53% for the standard oil N44 at 20°C. After the temperature settling time a fluid sample was pumped into the sensor and was monitored for one hour. The used measurement system provides one value for the quality factor Q and the resonance frequency f_r every five seconds. After 720 measurements the next sample of the same fluid was pumped into the sensor and the measurements were restarted. For each fluid at least four samples were measured.
To switch to the next fluid the measurement setup was cleaned with cleaning solvent and after wards flushed with air for two minutes. Due to the flow-through possibility it was not necessary to disassemble the sensor for the cleaning process.

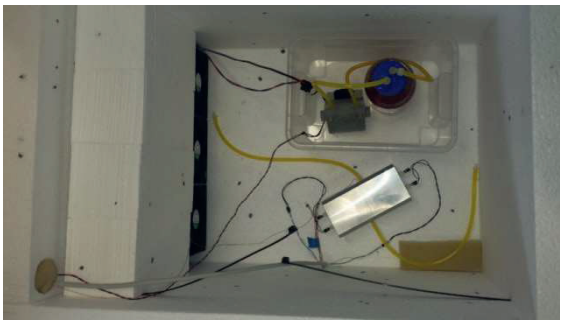


Figure 5: Isolated temperature box containing the viscosity sensor and the automated sampling system (in this picture the sensor is not connected to the sampling system).

Table 1: Viscosity and density values of the measured viscosity standard oils (measured at 20° C).

Fluids	*viscosity η / mPa.s*	*density ρ / kg / m³*	*damping factor mean(D)*	*Relative standard deviation σ(D) / mean(D)*
S2000	*7000*	*882.8*	*2.500e-03*	*3.138e-02*
N415	*1166*	*845.9*	*1.285e-03*	*1.672e-02*
S200	*602.2*	*887.0*	*9.937e-04*	*1.289e-02*
N44	*91.47*	*827.9*	*4.422e-04*	*5.764e-03*
N10	*19.29*	*870.3*	*2.542e-04*	*7.162e-03*

Measurement Results

Figure 6 shows a histogram of the measured damping *D*=1/2Q for N44 at 20±0.01°C for all fluid sample measurement points. The mean value is *D*=4.422e-04 and the standard deviation amounts to 0.5764%.

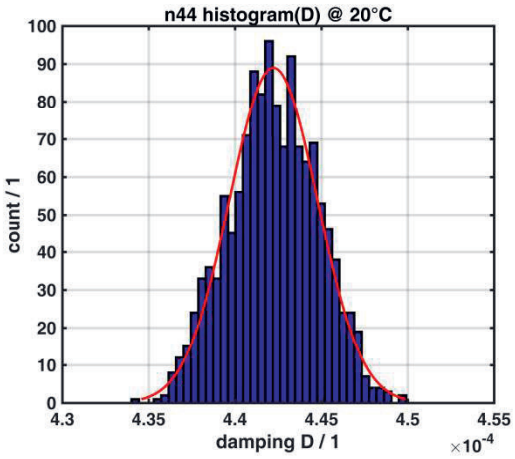


Figure 6: Histogram of the measured damping factor D for N44 at 20 C, for all measured samples.

Figure 7 shows the 3d-histogram where also the temperature is visible. For this diagram only measurements with a temperature in the range of 20±0.01°C were considered. This narrow temperature range is necessary due to the strong nonlinear temperature dependency of the viscosity. The temperature is measured as close as possible to the sample, see Figure 1.9.

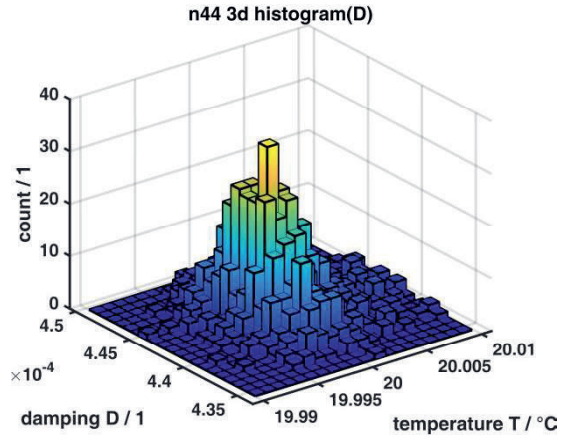


Figure 7: 3d-histogram of the damping factor D and the temperature. To guarantee thermal settling, only temperatures in the range of 20±0.01°C are considered.

Figure 8 shows the Damping *D*=1/2*Q* versus the √ηρ. The solid line represents a simple second order polynomial function fitted to the measurement values. It is clearly visible that for low viscosities the damping *D* is nearly linearly related to √ηρ. In this region the penetration depth δ (see [9]) of the shear wave is small compared to the fluid layer height (h_{fl}=(D_i-d_a)/2=1 mm) between the outer and the inner pipe. The penetration depth δ is defined as

$$\delta = \sqrt{\frac{2\eta}{\rho\omega}}, \qquad (1)$$

where ρ, η and ω denote fluid density, fluid viscosity and angular oscillation frequency, respectively. The linear relation for lower viscosities is not valid for higher viscosities, the penetration depth of the shear wave is for 1 Pa.s in the range of 0.2 mm (√ηρ approximately 30), and therefore the influence of the inner pipe is not negligible any more.

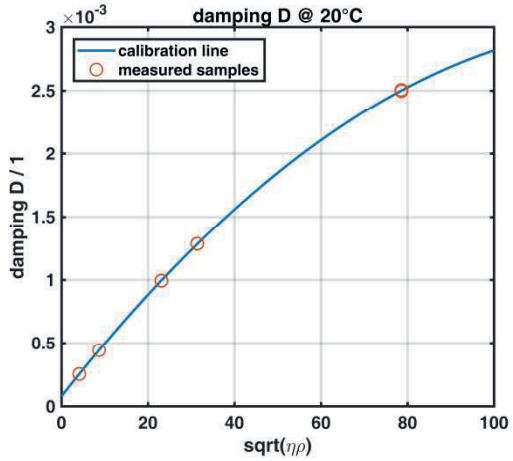


Figure 8: Measured damping D for all measured samples for all fluids versus √ηρ calculated from the calibration values of the oils (see Table 1).

Discussion

Figure 9 shows the deviation of the calculated viscosity from the reference viscosity. The blue circles represent the mean value of the calculated viscosity values (calculated from the measured damping factor *D* for known fluid density ρ), for each sample. For the calculation of the viscosity η the simple polynomial calibration curve, shown in Figure 8, was used. The error for the whole range from 19.3 mPa.s to 7 Pa.s is smaller than 1.5% which, in view of the simplicity of the used model, is a very satisfying result.

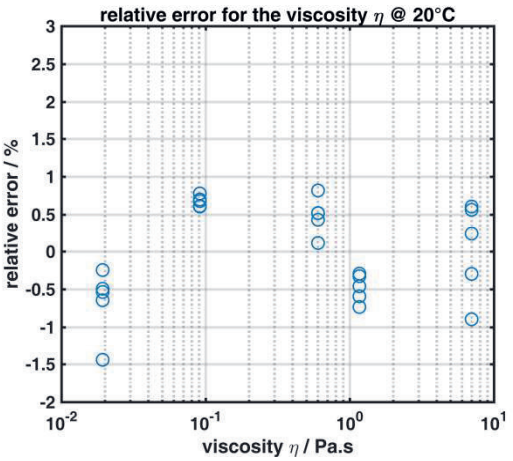


Figure 9: Relative error for the calculated viscosity η for a known density ρ. For the value of the density ρ the nominal values of the calibration oils @ 20°C were used.

Conclusion

We presented a robust viscosity sensor based on a torsionally oscillating pipe. Five different viscosity standard oils were measured at a very well defined temperature of 20±0.01°C. A sampling system was used to automatically change the sample-fluid in the sensor. More than four samples were measured for each standard oil. Each sample was measured for two hours, with a value for the resonance frequency f_r and the quality factor *Q* every five seconds. A simple second order polynomial calibration curve was used to calibrate the whole measured viscosity range from 19 mPa.s to 7 Pa.s. With this calibration curve the viscosity η was calculated for all measured samples and a relative error of less than 1.5% was achieved. For the calculation the density of the fluid was assumed as known.

Acknowledgments

Financial support was provided by the Austrian research funding association (FFG) under the scope of the COMET programme within the research project “Photonic Sensing for Smarter Processes (PSSP)” (contract # 871974).

References

[1] B. Jakoby, R. Beigelbeck, F. Keplinger, F. Lucklum, A. Niedermayer, E. K. Reichel, C. Riesch, T. Voglhuber-Brunnmaier and B. Weiss, "Miniaturized sensors for the viscosity and density of liquids-performance and issues," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 57, no. 1, pp. 111-120, Jan. 2010. doi: 10.1109/TUFFC.2010.1386

[2] S. J. Martin, Victoria Edwards. Granstaff, and Gregory C. Frye, “Characterization of a quartz crystal microbalance with simultaneous mass and liquid loading”, Analytical Chemistry 1991 63 (20), 2272-2281, doi: 10.1021/ac00020a015.

[3] E. K. Reichel, C. Riesch, F. Keplinger, C. E.A. Kirschhock, B. Jakoby, „Analysis and experimental verification of a metallic suspended plate resonator for viscosity sensing”, Sensors and Actuators A: Physical, Volume 162, Issue 2, 2010, Pages 418-424, ISSN 0924-4247, doi: 10.1016/j.sna.2010.02.017.

[4] A. Rahafrooz, S. Pourkamali, "Characterization of rotational mode disk resonator quality factors in liquid", Frequency Control and the European Frequency and Time Forum (FCS), 2011 Joint Conference of the IEEE International, 2011, doi: 10.1109/FCS.2011.5977863

[5] S. Clara, H. Antlinger, F. Feichtinger, A. O. Niedermayer, T. Voglhuber-Brunnmaier and B. Jakoby, "A balanced flow-through viscosity sensor based on a torsionally resonating pipe,"

2017 IEEE SENSORS, Glasgow, 2017, doi: 10.1109/ICSENS.2017.8234137

[6] S Clara and F Feichtinger and T Voglhuber-Brunnmaier and A O Niedermayer and A Tröls and B Jakoby, “Balanced torsionally oscillating pipe used as a viscosity sensor”, Measurement Science and Technology, 2018, doi: 10.1088/1361-6501/aae755.

[7] A.O. Niedermayer , T Voglhuber-Brunnmaier, J Sell and B Jakoby, "Methods for the robust measurement of the resonant frequency and quality factor of signifcantly damped resonating devices", 2012 Meas. Sci. Technol. 23 085107, doi: 10.1088/0957-0233/23/8/085107.

[8] A.O. Niedermayer, T. Voglhuber-Brunnmaier, M. Heinisch, F. Feichtinger, B. Jakoby, “Monitoring of the Dilution of Motor Oil with Diesel Using an Advanced Resonant Sensor System”, Procedia Engineering, Volume 168, 2016, Pages 15-18, ISSN 1877-7058, doi: 10.1016/j.proeng.2016.11.116.

[9] L.D. Landau, E.M. Lifshitz, "Fluid Mechanics", Butterworth-Heinemann, 1987

DOI 10.5162/sensoren2019/6.3.3

Bestimmung des Mischungsverhältnisses eines strömenden Gasgemisches

Jaqueline Stauffenberg[1], Steve Durstewitz[1], Martin Hofmann[1], Tzvetan Ivanov[1], Mathias Holz[1], Waleed Ehrhardt[2], W.–Ulrich Riegel[3], Jens-Peter Zöllner[1], Martin Ziegler[1] und Ivo W. Rangelow[1]

[1] Fachgebiet Mikro- und nanoelektronische Systeme, TU Ilmenau, Gustav-Kirchhoff-Str.1, 98693 Ilmenau; jaqueline.stauffenberg@tu-ilmenau.de; steve.durstewitz@tu-ilmenau.de; hofmann.martin@tu-ilmenau.de; tzvetan.ivanov@tu-ilmenau.de; mathias.holz@tu-ilmenau.de; jens-peter.zoellner@tu-ilmenau.de; martin.ziegler@tu-ilmenau.de; ivo.rangelow@tu-ilmenau.de; Tel-Nr. +49-3677-69 3717;
[2] Electronic & Sensors, Am Vogelherd 10 (Tor 1), 98693 Ilmenau; info@electrosens.de; Tel-Nr. +49-3677-667529;
[3] ETG – Entwicklungs- und Technologie Gesellschaft mbH, Am Eichicht 1 A, 98693 Ilmenau; wur@etg-ilmenau.de; Tel-Nr. +49-3677-46120.

Zusammenfassung

Mikrocantilever bieten ein breites Anwendungsspektrum in der Sensorik und der Messtechnik. In dieser Arbeit werden Cantilever als Strömungssensoren verwendet. Die meisten konventionell erhältlichen Strömungssensoren sind oft nur auf eine Gasart kalibriert und lassen eine Analyse von Gasgemischen nur mit erhöhtem Aufwand zu. Bei den hier angewendeten Sensoren wird ein Cantilever senkrecht im Strömungskanal positioniert. Dabei ist es möglich, den Sensor im dynamischen und im statischen Modus zu betreiben. Im dynamischen Modus wird der Cantilever in Schwingung versetzt. Dazu werden Resonanzfrequenz, Resonanzamplitude und Phase als Messgrößen erfasst. Im statischen Modus wird die Verbiegung des Cantilever registriert. Die Kombination der Modi lassen es zu, gleichzeitig die verschiedenen Messgrößen zu ermitteln. Durch eine Strömung wird der Sensor in seinem Bewegungsverhalten beeinflusst, wodurch es möglich ist, auf die Strömungsgeschwindigkeit zu schließen. Zusätzlich zur Bestimmung der Strömungsgeschwindigkeit besteht die Möglichkeit, unterschiedliche Gasarten zu detektieren. Jedes Medium hat bestimmte Eigenschaften (Dichte und Viskosität), die sich verschieden auf die Bewegung des Sensors auswirken. Demzufolge können Konzentrationen eines bekannten Gasgemisches ermittelt werden. Am Beispiel des Luft-Kohlenstoffdioxid-Gemisches wird in dieser Arbeit untersucht, wie sich das Mischungsverhältnis bei unterschiedlichen Strömungsgeschwindigkeiten auf das Cantileververhalten auswirkt.

Keywords: Gassensorik, Gasgemisch, Strömungssensoren, Cantilever, Piezoresistiv

Einleitung in die Thematik

Mikrocantilever bilden eine wichtige Gruppe unter den MEMS-Bauelementen. Sie sind einige µm große freistehende Zungen, die infolge von Krafteinwirkungen und Masseanlagerungen Änderungen in ihrem mechanischen Verhalten erfahren. Sie bieten ein weites Spektrum für sensorische Aufgaben: Raster-Kraft-Sensoren mit atomarer Auflösung, Kraft- und Drucksensoren, Beschleunigungssensoren, Massesensoren, Temperatursensoren, IR-Strahlungssensoren, bio- und chemische Sensoren, Strömungssensoren [1-3].
Auf Grund der Möglichkeit diese resonanten MEMS-Sensoren extrem klein zu bauen, sind sie in der Lage Kräfte im nN-Bereich, Massen im fg-Bereich und Abmessungen im sub–nm-Bereich erfassen zu können. Neben Metallen ist Silizium der dominierende Werkstoff für diese MEMS-Sensoren. Damit lassen sich erfasste physikalische Größen in elektrische Signale wandeln. Ihre Herstellung erfolgt mit Verfahren, die äquivalent den Prozessschritten der Mikro- und Nanoelektronikfabrikation sind.
Wissenschaftliches Interesse besteht an der Untersuchung von vibrierenden Mikrocantilevern in viskosen Medien. Diese Untersuchungen sind u.a. mit der Atom-Kraft-Mikroskopie und der Sensorik mit mikromechanischen Oszillatoren verbunden. Zur Beschreibung des Einflusses des umgebenden Mediums ist die Voraussage der auf den Oszillator wirkendenden Kräfte von entscheidender Bedeutung. In verschiedenen Arbeiten wurde dieser Einfluss untersucht [4-7]. Die wirkenden Kräfte werden mit dem Massezuwachs durch das mitbewegte

Fluid und der hydrodynamischen Dämpfung beschrieben. Beide Effekte wirken als Trägheits- und Viskositätskräfte auf die Bewegung des Körpers im Fluid. Sie sind durch die charakteristischen Abmessungen des Körpers und der Dicke der viskosen Grenzschicht bestimmt und bewirken eine Verschiebung der Resonanzfrequenz f_0 und der Güte Q. Im allgemeinen Fall ist das Problem der Berücksichtigung der aerodynamischen Kräfte auf einen vibrierenden Cantilever wegen des 3-dimensionalen Gasflusses sehr kompliziert. Die wirkenden Kräfte hängen vom Verhältnis der Dicke T zur Breite W der Platte, dem Größenverhältnis von Oszillationssamplitude $\hat{z}$ zur charakteristischen Plattenabmessung $\frac{KC}{2\pi} = \frac{\hat{z}}{W}$ (KC – Keulegan-Carpenter-Zahl) und der frequenzabhängigen normierten Reynoldszahl (siehe Gl. (16)) ab [7]. In [4] wurde dieses Problem für eine im ruhenden Fluid mit kleiner Amplitude schwingenden dünne Platte behandelt und der Einfluss des Fluids auf die Resonanzfrequenz f_0 und der Güte Q abgeleitet. Bardalis u.a. [8] untersuchten unter Nutzung dieser Beziehungen die Bestimmung von Dichte und Viskosität von Gasen mittels Mikrocantilever. In [9] wurde gezeigt, dass das Resonanzverhalten von Mikrocantilevern durch die Gasströmung beeinflusst wird und dieser Einfluss zur Messung des Gasflusses genutzt werden kann.

In strömenden Gasen ist das Resonanzverhalten von der Gasart bzw. dem Gasgemisch und der Strömungsgeschwindigkeit abhängig. Am Beispiel des Luft-Kohlenstoffdioxid-Gemisches wird in dieser Arbeit untersucht, wie sich das Mischungsverhältnis bei unterschiedlichen Strömungsgeschwindigkeiten auf das Cantileververhalten auswirkt. Diese Untersuchung ist eingebettet in die Entwicklung eines kompakten, autokalibrierenden, fernabfragbaren Strömungsmesssystems mit Gasarterkennung.

Grundgleichungen zum Resonanzverhalten

Gegenstand der Untersuchungen ist ein vibrierender Cantilever in einem strömenden Fluid. Diese elastische Platte der Länge L, der Breite W und der Dicke T ($L > W \gg T$) wird quer zur Strömung platziert (Abb.1). Die Bewegung dieses dünnen Biegebalkens lässt sich im Rahmen der Bernoulli-Näherung durch folgende Differentialgleichung beschreiben

$$EI\,\frac{\partial^4 \mathrm{z(x,t)}}{\partial x^4} + \rho \cdot W \cdot T\,\frac{\partial^2 z(x,t)}{\partial t^2} = f(x,t)\,, \qquad (1)$$

wobei das Elastizitätsmodul E und das Flächenträgheitsmoment I längs des Balkens als konstant angenommen wurden. $f(x,t)$ ist die längenbezogene Krafteinwirkung (diese enthält die Trägheits- und Dämpfungskräfte des umgebenden Mediums sowie die externe Anregung), W – die Breite des Balkens, T - die Dicke und ρ - die mittlere Dichte des Balkens. Für diesen Cantilever mit rechteckigem Querschnitt gilt für das Flächenträgheitsmoment

$$I = \frac{T^3 \cdot \mathrm{W}}{12} \qquad (2)$$

Für einen einseitig am Ort x=0 eingespannten Cantilever mit den Randbedingungen

$$\mathrm{z(x=0,t)} = 0,\ \frac{\partial z(\mathrm{x=0,t})}{\partial x} = 0,$$
$$\frac{\partial^2 z(\mathrm{x=L,t})}{\partial x^2} = 0,\ \frac{\partial^3 z(\mathrm{x=L,t})}{\partial x^3} = 0\,. \qquad (3)$$

und für die Näherung $f(x,t) \approx 0$, was für Vakuumbedingungen angenommen werden kann, erhält man die natürlichen Eigenmoden

$$z(x,t) = \left\{ \begin{matrix} [\cos(k_n x) - \cosh(k_n x)] + \\ \frac{[-\cos(k_n L) - \cosh(k_n L)]}{[\sin(k_n L) + \sinh(k_n L)]} \\ \cdot [\sin(k_n x) - \sinh(k_n x)] \end{matrix} \right\} \cdot A \sin(\omega_n t - \varphi_0) \qquad (4)$$

und Eigenfrequenzen

$$f_{n,vac} = \frac{1}{2\pi}\omega_n = (k_n L)^2 \frac{1}{2\pi}\frac{T}{L^2}\sqrt{\frac{E}{\rho \cdot 12}} \qquad (5)$$

mit $k_0 L = 1{,}8751$, $k_1 L = 4{,}6941$, $k_2 L = 7{,}8548$, ...

Für die in dieser Arbeit durchgeführten Untersuchungen wird der Cantilever im ersten Resonanzmode (n=0) genutzt. Zur Beschreibung seines Verhaltens nutzen wir das Modell des einfachen harmonischen Oszillators (SHO)

$$\frac{d^2 z}{dt^2} + \frac{D}{m} \cdot \frac{dz}{dt} + {\omega_0}^2 \cdot z = \frac{F sin(\omega t)}{m} \qquad (6)$$

mit z der Auslenkung, D der Dämpfung, m der Masse des Cantilevers und des mitbewegten Fluids, $\omega_0/2\pi$ der Resonanzfrequenz des Cantilevers im Fluid und $F sin(\omega t)$ der periodischen Anregung. Schwingungsamplitude $\hat{z}(\omega)$ und Phasenverschiebung φ_0 berechnen sich nach

$$\hat{z}(\omega) = \frac{\hat{F}/m}{\sqrt{({\omega_0}^2 - \omega^2)^2 + \left(\frac{\mathrm{D}}{m}\omega\right)^2}}\,, \qquad (7)$$

$$\varphi_0 = \arctan\left(\frac{\frac{\mathrm{D}}{m}\omega}{{\omega_0}^2 - \omega^2}\right). \qquad (8)$$

Das Resonanzverhalten ist charakterisiert durch die Güte Q mit der Bandbreite Δf

$$Q = \frac{f_0}{\Delta f} = \frac{\omega_0 \cdot m}{\mathrm{D}}\,, \qquad (9)$$

durch die Resonanzamplitude

$$\hat{z}_{max} = \frac{\hat{F}/m}{{\omega_0}^2} \cdot Q \cdot \frac{1}{\sqrt{1 - \left(\frac{1}{2Q}\right)^2}} \approx \frac{\hat{F}/m}{{\omega_0}^2} \cdot Q \approx \frac{\hat{F}}{\omega_0}\frac{1}{D}\,, \qquad 10)$$

durch die Resonanzfrequenz

$$\omega_{max} = \omega_0 \sqrt{1 - \frac{1}{2Q^2}} \approx \omega_0 \qquad (11)$$

und durch die Phasenverschiebung im Resonanzpunkt

$$\varphi_{max} = \arctan\left(\sqrt{\frac{Q^2}{4} - 2}\right) \approx \frac{\pi}{2} \qquad (12)$$

Die in den Gleichungen (10)-(12) gemachten Näherungen gelten für $Q \gg 1$ und sind für die im Rahmen dieser Arbeit durchgeführten Untersuchungen anwendbar. Die Gütefaktoren der verwendeten Cantilever in Gasen unter Normaldruck sind > 100.
Der Einfluss des Fluids auf die Resonanzfrequenz gegenüber Vakuumbedingungen wird beschrieben durch

$$f_0 = f_{0,vac} \cdot \sqrt{\frac{m_{can}}{m_{can}+m_{fluid}}} \qquad (13)$$

mit der Masse des Cantilevers m_{can} und der mitbewegten Masse des Fluids m_{fluid}.
Die Cantilevermasse erhält man aus seiner mittleren Dichte ρ und den Abmessungen

$$m_{can} = \rho \cdot T \cdot W \cdot L \quad . \qquad (14)$$

Die Masse des mitbewegten Fluids berechnet sich nach [6]

$$m_{fluid} = \frac{\rho_{fluid} \cdot \pi \cdot W^2 \cdot L}{4} \cdot \left[a_1 + a_2 \sqrt{\frac{1}{2\,Re}}\right] \qquad (15)$$

mit ρ_{fluid} der Dichte des Gases, den Konstanten $a_1 = 1{,}0553$ und $a_2 = 3{,}7997$ und der Reynoldszahl

$$Re = 2\pi f_0 \cdot \frac{w^2 \rho_{fluid}}{4 \cdot \eta_{fluid}}. \qquad (16)$$

mit η_{fluid} der dynamischen Viskosität des Gases. Die Güte des im Fluid schwingenden Cantilevers wird durch die intrinsischen, das sind im Wesentlichen der Energieverlust an der Einspannstelle [10] und durch die thermoelastische Dämpfung [11] im Cantilever selbst, und extrinsischen Einfluss, das ist der Energieverlust durch das umgebende viskose Fluid, bestimmt. Für die in dieser Arbeit genutzten Cantilever dominiert der Einfluss des Fluids. Es bestimmt die Güte gemäß

$$Q \approx Q_{fluid} = \frac{\frac{4 \cdot \rho \cdot T}{\pi \cdot W \cdot \rho_{fluid}} + a_1 + a_2 \frac{1}{\sqrt{2Re}}}{b_1 \frac{1}{\sqrt{2Re}} + b_2 \cdot \frac{1}{2Re}} \approx \frac{\frac{4 \cdot \rho \cdot T}{\pi \cdot W}}{b_1 \frac{\sqrt{\eta_{fluid} \rho_{fluid}}}{\sqrt{\pi f_0 w^2}}} \qquad (17)$$

mit den Konstanten $b_1 = 3{,}8018$ und $b_2 = 2{,}7364$.

Cantileversensoren und Messaufbau

Bei einem Mikrocantilever handelt es sich um eine einseitig eingespannte dünne Blattfeder aus Silizium. Die hier verwendeten Cantilever mit rechteckigem Querschnitt haben Abmessungen von L=350µm x W=140µm x T=5µm und besitzen einen integrierten thermomechanischen Aktuator und eine Wheatstone-Messbrücke in einer Vollbrückenanordnung (Abb. 1).

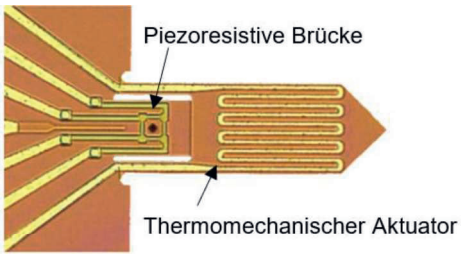


Abb.1: Vergrößerung eines Cantileversensors mit einer piezoresistiven Widerstandsbrücke und einem mäanderförmigen thermomechanischen Aktuator [12]

Das Anlegen eines periodischen Signals $u_h(t)$ (Anregungssignal) an den Aktuator führt zu einer periodischen Temperaturänderung. Aufgrund der unterschiedlichen thermischen Ausdehnungskoeffizienten (Al-Aktuator und Si-Cantilever) führt die periodische Temperaturänderung ΔT zum Schwingen der Anordnung (Bimorph-Effekt):

$$\frac{\partial^2 z(x,t)}{\partial x^2} \sim \Delta T(x,t) \sim \frac{u_h(t)^2}{R_{Aktuator}} \qquad (18)$$

Über die an der Einspannstelle befindliche Wheatstone-Messbrücke kann die Verbiegung detektiert werden. Aus dem Antwortsignal kann die statische und dynamische Verbiegung bestimmt werden und die daraus resultierenden Kräfte auf den Cantilever. In Abbildung 2 sind Frequenzsweeps eines Cantilevers für unterschiedliche Volumenströme (SCCM - Standard Cubic Centimeters per Minute) in Luft dargestellt. Die Kurvenschar zeigt, die Oszillationsamplitude und die Phase in Abhängigkeit von der Frequenz des Anregungssignals.
Mit steigendem Volumenstrom und damit steigender Dämpfung sinkt die Oszillationsamplitude laut Gleichung (7). Die geringe Verschiebung der Resonanzfrequenz (siehe Gl. 11) zu höheren Frequenzen ist durch eine geringere Masse des Fluids durch einen steigenden Volumenstrom laut Gleichung (13) zu erklären [9]. Die Phasenverschiebung (Phase) zwischen Antwort- und Anregungssignal ist im unteren Bild gezeigt.

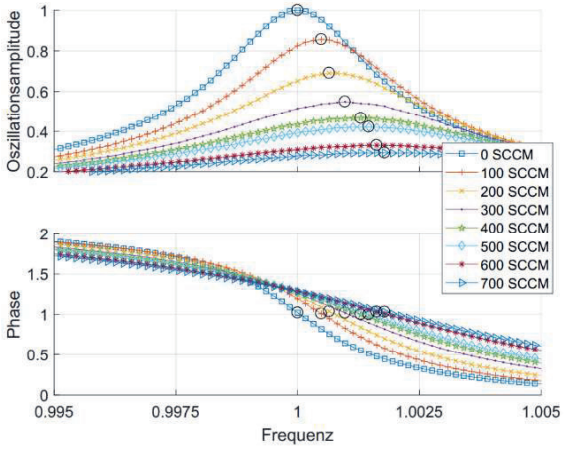


Abb.2: Oszillationsamplitude und Phase in Abhängigkeit der Frequenz bei unterschiedlichen Volumenströmen (Die Kreise markieren die Oszillationsmaxima und die zugehörige Phase)

Durch einen Vergleich der Punkte in Resonanz bezüglich der Phase kann eine konstante Phase gegenüber der Resonanzfrequenz nach Gleichung (12) aufgezeigt werden. In Abbildung 3 ist die Güte der Resonanz über der Resonanzamplitude für konstante Anregung aufgetragen. Sie wurden aus Frequenzsweeps für unterschiedliche Gase (Luft und Kohlenstoffdioxid) und Volumenströme bestimmt.

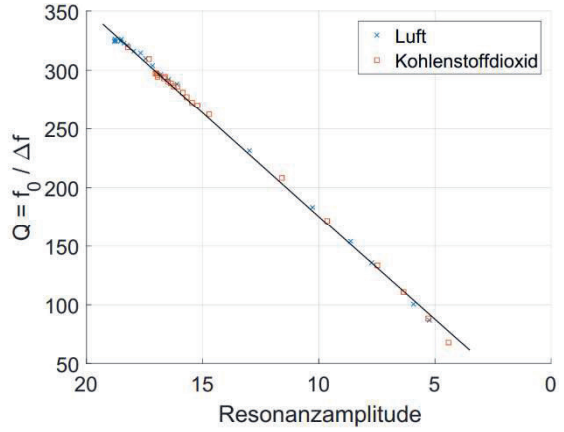


Abb.3: Güte in Abhängigkeit der Resonanzamplitude im Gasmedium Luft und Kohlenstoffdioxid für verschiedene Volumenströme

Sie zeigt, dass eine gasartenunabhängige Proportionalität zwischen dem Gütefaktor und der Resonanzamplitude nach Gleichung (10) besteht und die Güte direkt aus der Resonanzamplitude bestimmt werden kann.

Messprinzip

Zur Auswertung der Mikrocantileversignale wurde ein Messprinzip verwendet, was in Abbildung 4 dargestellt ist. Zur Bereitstellung des Anregungssignales des Mikrocantilevers wird ein Signalgenerator (DDS - Direct Digital Synthesis) genutzt. Das Antwortsignal $u_s(t)$ wird anschließend durch Filterung in ein statisches (Gleichspannung) u_{stat} und dynamisches (Wechselspannung) $\hat{u}$ Signal aufgetrennt. Mithilfe des Lock-In-Verstärker wird die dynamische Amplitude und die Phase bestimmt. Die gemessene Phase wird vom Lock-In-Verstärker auf einen Phasenregler gegeben. Dieser steuert die Frequenz des Signalgenerators so, dass auf eine voreingestellte Phase geregelt wird.

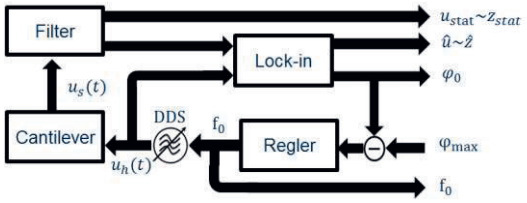


Abb.4: Prinzipielle Darstellung des Messprinzips

Somit kann garantiert werden, dass sich der Mikrocantilever auch bei Änderung der Umgebungseinflüsse (z.B. durch Massenstrom, Gasart) immer in Resonanz befindet. Dieses Messverfahren ermöglicht weiterhin im Gegensatz zum Frequenzsweep eine Echtzeitaufnahme der Messsignale.

Versuchsaufbau

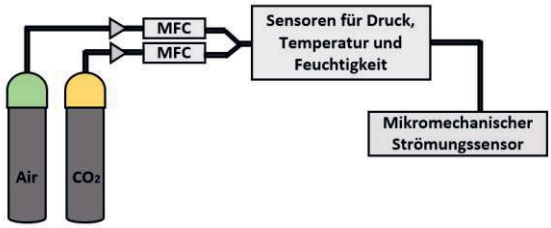


Abb.5: Prinzipielle Darstellung des Messaufbaus zur Analyse eines Luft-Kohlenstoffdioxid-Gasgemisches mit Sensoren zur Detektion der Quereinflüsse und dem verwendeten mikromechanischen Strömungssensor (Cantilever)

In Abbildung 5 ist der Versuchsaufbau dargestellt. Die beiden Gasquellen Luft und Kohlenstoffdioxid sind jeweils an einen Massenflussregler (MFC) angeschlossen. Durch die Steuerung dieser Regler können die Anteile im Gasgemisch variiert werden. Das Gasgemisch wird über eine Sensoreinheit zur Überwachung der Umgebungsgrößen, wie Druck, Temperatur und Feuchtigkeit, geführt, um sicherzustellen, dass keine Beeinflussung der Messgrößen durch Quereinflüsse stattfindet. Danach trifft

das Gasgemisch auf den Cantilever, der senkrecht im Strömungskanal positioniert ist.

Messergebnisse und Auswertung

Bei der Detektion der Gaskonzentration im Luft-Kohlenstoffdioxid-Gasgemisch wurde der Volumenstrom auf einen konstanten Wert eingestellt. Dabei wurde lediglich das Mischungsverhältnis verändert, um ausschließlich eine Detektion der Konzentration zu ermöglichen. Dazu wurden diese Versuche bei unterschiedlichen Volumenströmen durchgeführt.

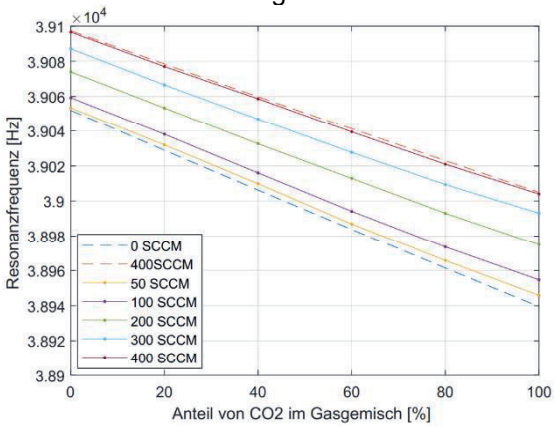


Abb. 6: Resonanzfrequenz in Abhängigkeit der Konzentration von Kohlenstoffdioxid im Luft-Kohlenstoffdioxid-Gasgemisch bei unterschiedlichen Volumenströmen (— - gemessen, ---- - simuliert)

In Abbildung 6 ist die Resonanzfrequenz in Abhängigkeit des Mischungsverhältnisses für verschiedene Gasströme (50, 100, 200, 300 und 400 SCCM) dargestellt. Für ein ruhendes Gas (0 SCCM) und für eine Strömung (400 SCCM) wurden die in Abb.6 aufgetragenen Kurven für die Resonanzfrequenz nach Gl. (13)-(15) berechnet. Die Resonanzfrequenz sinkt bei der Zunahme an Kohlenstoffdioxid im Gasgemisch. Dieser Einfluss lässt sich im Wesentlichen auf den Dichteunterschied von Luft zu Kohlenstoffdioxid zurückführen. Mit zunehmender Dichte sinkt die Resonanzfrequenz des Cantilevers. Anhand der Verschiebung der Resonanzfrequenz lässt sich somit auf das Mischungsverhältnis im Gasgemisch schließen. Mit zunehmender Strömungsgeschwindigkeit wird die Resonanzfrequenz zusätzlich zu höheren Werten verschoben. Beim Übergang von 100 % Luft auf 100 % Kohlenstoffdioxid bei einem Volumenstrom von 50 SCCM ist eine Änderung der Resonanzfrequenz von 107 Hz messbar. Bei einem Volumenstrom von 400 SCCM hingegen beträgt die Änderung 92 Hz. Die Standardabweichung liegt bei ± 1 Hz. Bei der theoretisch berechneten Funktion bei 0 SCCM ist eine Änderung der Resonanzfrequenz von 113 Hz errechnet worden. Die vom Cantilever mitbewegte Masse bewirkt nach [4] eine Reduzierung der Resonanzfrequenz. Sie ist abhängig von der Geometrie, der Dichte und der Viskosität des umgebenden Mediums. Diese mitbewegte Masse nimmt mit steigender Strömungsgeschwindigkeit ab. Dies bedeutet, dass eine Erhöhung der Strömungsgeschwindigkeit des Gases eine Reduzierung der Grenzschichtdicke bewirkt und somit zu einer Zunahme der Resonanzfrequenz führt. Der Einfluss der Dichte ist hingegen zur Viskosität stärker. Dies lässt sich darauf zurückführen, dass der Dichteunterschied deutlich höher ist als der Unterschied zwischen den Viskositätswerten der beiden Gasmedien.

In Abbildung 7 ist die Resonanzsamplitude in Abhängigkeit des Mischungsverhältnisses bei verschiedenen Gasströmen dargestellt. Zusätzlich ist die Abhängigkeit des Faktors $1/\sqrt{\eta_{fluid}\,\rho_{fluid}}$ vom Mischungsverhältnis eingetragen. Im Gegensatz zur Resonanzfrequenz wird die Resonanzamplitude deutlich stärker von der Strömungsgeschwindigkeit (Volumenstrom) beeinflusst.

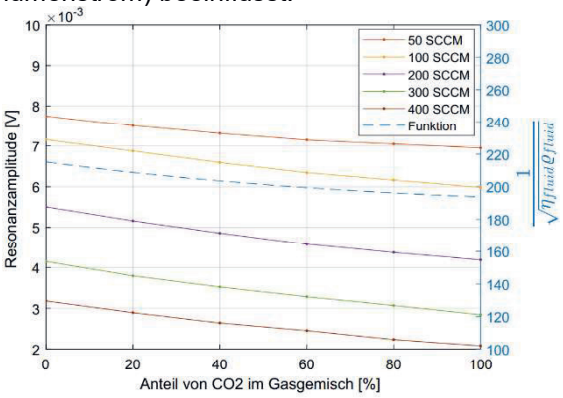


Abb.7: Resonanzamplitude und Funktion zur Beschreibung der Mischeigenschaften (blau) in Abhängigkeit vom Luft-Kohlenstoffdioxid-Gasgemisch bei unterschiedlichen Volumenströmen

Eine Erhöhung des Volumenstromes von 50 auf 400 SCCM bewirkt eine Abnahme der Resonanzamplitude um 60%. Der Einfluss durch das Gasmedium auf die Resonanzsamplitude ist jedoch geringer. Die Änderung der Resonanzsamplitude ist proportional zu $1/\sqrt{\eta_{fluid}\,\rho_{fluid}}$ (siehe Gl. (17)).

Zusätzlich zum dynamischen Messmodus ist es möglich, die statische Verbiegung aufzunehmen. Diese Messgröße wird unabhängig von den dynamischen Messgrößen detektiert. Der Einfluss der Gasart und der Strömungsgeschwindigkeit auf die statische Verbiegung ist in Abbildung 8 dargestellt. Dazu

wurde die statische Verbiegung bei unterschiedlichen Volumenströmen in Luft, Kohlenstoffdioxid und einem Gemisch mit einem Anteil von 50% Kohlenstoffdioxid und 50% Luft aufgenommen. Für die statische Verbiegung gilt folgender Zusammenhang:

$$z_{stat} \sim c_D \cdot \rho_{fluid} \cdot v^2$$

Die statische Verbiegung steigt somit mit der Strömungsgeschwindigkeit als auch durch eine höhere Dichte. Die Empfindlichkeit ist bei kleinen Gasströmungen sehr gering. Hingegen ist bei großen Strömungsgeschwindigkeiten der Einfluss von Volumenstrom und Gasart detektierbar.

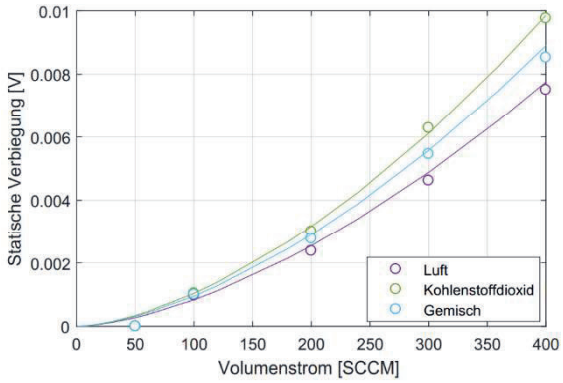


Abb. 8: Statische Verbiegung in Abhängigkeit des Volumenstromes für Luft, Kohlenstoffdioxid und einem Gasgemisch mit der Zusammensetzung 50% zu 50%

Zusammenfassung

Im Rahmen dieser Arbeit wurde gezeigt, dass mit einem selbstaktuierten Cantileversensor die Möglichkeit besteht, die Gaszusammensetzung und die Strömungsgeschwindigkeit eines bekannten strömenden Gasgemisches zu bestimmen. Das phasengeregelte Messprinzip erlaubt die Erfassung der Resonanzfrequenz und Resonanzamplitude in Echtzeit. Die Messungen zeigen eine Genauigkeit von 1 % bei der Bestimmung der Gaskonzentration. Der verwendete Cantilever erlaubt neben der dynamischen Messung gleichzeitig die Erfassung der Verbiegung. In dieser Arbeit konnte gezeigt werden, dass dieses Konzept erfolgreich zur Bestimmung der Flussmenge und der Konzentration eines Gasgemisches angewendet werden kann.

Danksagung

Die Autoren bedanken sich für die Unterstützung durch die Thüringer Aufbaubank im Rahmen des Projekts (Vorhaben-Nr. 2017VF0047).

Literaturnachweis

[1] Kumar A., Periasamy C. and Pant B.D.; Cantilever based MEMS pressure sensors using different Piezoelectric Materials: A comparative study; *IJEDR* 2 (4), 4022-26 (2014); doi: 10.14419/ijet.v7i1.5.9153

[2] Steffanson M., Gorovoy K., Ramkiattisak V., Ivanov Tzv., Kròl J., H. Hartmann and Rangelow I.W.; ARCH-type micro-cantilever FPA for uncooled IR detection; *Microelectronic Engineering Volume* 98, 614-618 (2012); doi: 10.1116/1.2134721

[3] Waggoner P. and Craighead H.; Micro- and nanomechanical sensors for environmental, chemical, and biological detection, *Lab on a Chip* 7, 1238–55 (2007); doi: 10.1039/b707401h

[4] Sader, J.; Frequency responds of cantilever beams immersed in viscous fluids with application to the atomic force microscope; *Journal of Applied Physics* 84(1), 64-78 (1998); doi: 10.1063/1.368002

[5] Hosaka H., Itao K., and Kuroda S.; Damping characteristics of beam-shaped micro-oscillators, *Sens Actuators A Phys.* 49, 87-95 (1995); doi: 10.1016/0924-4247(95)01003-J

[6] Maali A., Hurth C., Boisgard R., Jai C.Cohen-Bouhacina T. and Aime J.-P.; Hydrodynamics of oscillating atomic force microscopy cantilevers in viscous fluids; *Journal of Applied Physics* 97, 074907 (2005); doi: 10.1063/1.1873060

[7] Egorov A.G., Kamalutdinov A.M., Nuriev A.N. and Paimiushin V.N.; Theoretical-experimental method for determining the parameters of damping based on the study of damped flexural vibrations of test specimens, *2. Aerodynamic component of damping; Mechanics of Composite Materials* 50 (3), 267-278 (2014); doi: 10.1134/S0021894416020103

[8] Bardalis, A., Pfau, A. and Kalfas, A.; Measurement and Evaluation of the Gas Density and Viscosity of Pure Gases and Mixtures using a Micro-Cantilever Beam; *Sensors*, 15, 24318-24342 (2015); doi: 10.3390/s150924318 .

[9] Zöllner, J.-P., Durstewitz, S., Stauffenberg, J., Ivanov, Tz., Holz, M., Ehrhardt, W., Riegel, W.-U. and Rangelow, I. W.; Gas-Flow Sensor Based on Self-Oscillating and Self-Sensing Cantilever, *Eurosensors Proceedings*, *2*(13), 846 (2018); doi:10.3390/proceedings2130846

[10] Hao, Z., Erbil, A., and Ayazi, F.; An analytical model for support loss in micromachined beam resonators with in-plane flexural vibrations, *Sensors and Actuators A: Phys.* 109, 156-164 (2003); doi: 10.1016/j.sna.2003.09.037

[11] Zener C.; Internal friction in Solids II. General Theory of Thermoelastic Internal Friction; *Phys. Rev.* 53, 90-99 (1938); doi: 10.1103/PhysRev.53.90

[12] Rangelow, I.W., et al.; Apparatus and method for investigating surface properties of different materials, Patent US8689359B2 (2014);

DOI 10.5162/sensoren2019/6.3.4

Integraler Defekterkennungssensor für Faserkunststoffverbunde mit koppelmittelfreiem luftgekoppelten Ultraschall

Yannick Bernhardt[1], Marc Kreutzbruck[1]

[1]Universität Stuttgart, Institut für Kunststofftechnik, Pfaffenwaldring 32, 70569 Stuttgart, Deutschland
0711 685 62888, Yannick.Bernhardt@ikt.uni-stuttgart.de, Marc.Kreutzbruck@ikt.uni-stuttgart.de

Zusammenfassung

Luftgekoppelter Ultraschall erfreut sich einer zunehmenden Nachfrage für die Prüfung von Faserkunststoffverbunden sowie auch für die Prüfung von Fügeverbindungen an metallischen Bauteilen. Die Vorteile liegen hier im kontaktlosen Prüfen, bei dem gegenüber konventioneller Ultraschallprüfung gleiche Prüfzeiten erreicht werden. Der Auftrag eines Koppelmittels erschwert bis heute eine vollständige Automatisierung in der Serienprüfung. Beim luftgekoppelten Ultraschall muss aufgrund der hohen Impedanzunterschiede von Festkörper zu Luft mit speziell angepassten Prüfköpfen gearbeitet werden, was dann meist mittels Senkrechteinschallung in Transmissionsanordnung erfolgt. Für Prüfaufgaben bei denen nur einseitiger Zugang möglich ist, kann die Prüfung mittels schräger Einschallung und den dadurch erzeugten Plattenwellen durchgeführt werden. Unter einem bestimmten Resonanzwinkel, bei dem der Sinus des Einschallwinkels dem Verhältnis von Luftschallgeschwindigkeit und Plattenwellenausbreitungsgeschwindigkeit entspricht, wird ein höheres Signal-Rausch-Verhältnis im Vergleich zur Normaldurchschallung erzielt. Dieser Winkel wurde bisher bei der Prüfung in Schrägeinschallung manuell iterativ ermittelt, indem unter Beobachtung der Amplitude die Winkel von Ultraschallsender und Empfänger sukzessive angepasst wurden. Über eine motorgesteuerte Prüfkopfhalterung und einer online Auswertung des empfangenen Signals wurde nun eine automatisierte und damit zeitsparende Winkelfindung realisiert. Auch wenn sich die Plattenwellengeschwindigkeit durch eine Wanddickenänderung verändert, ist eine nachführende Anpassung des Einschallwinkels möglich.

Keywords: Ultraschall, zerstörungsfreie Bauteilprüfung, automatisierte Prüfung

Einführung

Durch die hohen Anforderungen von Umweltschutz an die moderne Automobilität, nimmt Leichtbau eine immer wichtiger werdende Rolle bei der Entwicklung neuer Fortbewegungsmittel ein. Masseeinsparung ermöglicht eine Reduzierung des Energiebedarfs und somit eine Verringerung des Kraftstoffbedarfs, der mit einem niedrigeren CO2 Ausstoß einhergeht. In der Luftfahrt ist die Energieeinsparung durch Leichtbau neben dem Wunsch zum Erreichen von Klimazielen, auch ein kostengetriebener Faktor. Eine Gewichtsreduzierung der Flugzeugstruktur um beispielsweise 12 % führt zu einer Reduzierung des Kerosinverbrauchs um 7 % [1]. Die Reduzierung von Strukturgewicht kann durch die Variation verschiedener Parameter erreicht werden. Einerseits kann durch den sogenannten Werkstoffleichtbau Masse eingespart werden, in dem Werkstoff mit hoher Materialdichte durch Werkstoff mit geringer Materialdichte, aber ähnlichem Festigkeits- und Steifigkeitsverhalten ausgetauscht werden. Beispielsweise sei hier die Verwendung von Hochleistungskunstoffen, wie zum Beispiel Faserkunststoffverbunden, statt Aluminium im Flugzeugrumpf genannt. Gleichzeitig wird auch versucht, durch konstruktiven Leichtbau das eingesetzte Materialvolumen vollständig auszunutzen. Dies geht nur, wenn die tatsächlichen, am Bauteil angreifenden Lasten, sich in einer ähnlichen Größenordnung befinden, wie die Auslegungslasten. Um eine Überdimensionierung zu vermeiden, muss das Werkstoffverhalten genau bekannt sein und die Streuung der Werkstoffkennwerte muss eng toleriert sein. Beschädigungen im Bauteil müssen erkannt werden können und die genauen Auswirkungen eines Schadens auf das Bauteil müssen vorhersagbar sein.

Leichtbauwerkstoffe, wie Faserkunststoffverbunde (FKV), haben ganz spezielle Eigenschaften, die bei der zerstörungsfreien Bauteilprüfung beachtet werden müssen. Defekte, die

an diesen Werkstoffen auftreten sind beispielsweise Delaminationen, Impactschäden, Faserbrüche, Matrixbrüche, Harznester, Faserondulationen, Poren oder Fremdeinschlüsse. Aus diesem Grund wurden in den letzten Jahren verschiedene Prüfverfahren weiterentwickelt, um die speziellen Anforderungen der Prüfung von FKV zu realisieren. Einerseits gibt es Verfahren, die nur an der Oberfläche wirksam sind und tieferliegende Defekte nicht erkennen können. Bei oftmals dünnen FKV-Strukturen ist dies aber oft weniger problematisch. Zu diesen Oberflächenverfahren zählt die Shearografie [2] oder die aktive Thermografie [3]. Für dickwandigere Strukturen gibt es Volumenverfahren wie die Durchstrahlungsprüfung (Computertomografie), Kontaktultraschall (Einkanal + Phased-Array [4]) oder auch den luftgekoppelten Ultraschall [5], bei dem auf ein flüssiges Koppelmittel verzichtet wird und eine berührungsfreie Prüfung möglich ist.

1. Luftgekoppelter Ultraschall

Durch den Verzicht auf ein flüssiges Koppelmittel, müssen die Ultraschallwellen direkt durch das Medium Luft übertragen werden, was weitere Herausforderungen mit sich bringt. Einerseits gilt es die hohen akustischen Impedanzunterschiede zwischen Festkörper und Luft an den Grenzschichten Prüfkopf/Luft und Prüfkörper/Luft mehrfach zu überwinden, was zu hohen Reflexionsverlusten führt. Andererseits können wegen der frequenzabhängigen Schalldämpfung in Luft keine Frequenzen höher als ~ 1 MHz eingesetzt werden. Die Problematik der hohen Impedanzunterschiede, kann durch ein spezielles Prüfkopfdesign, hohe Schalldrücke und empfindliche Verstärker angegangen werden [6]. Wegen der starken Reflexionen wird luftgekoppelter Ultraschall gewöhnlich in Transmission angewendet. Möglich ist neben der Normaldurchschallung oder die Erzeugung von Lambwellen durch die Schrägeinschallung (siehe Abb. 1.).

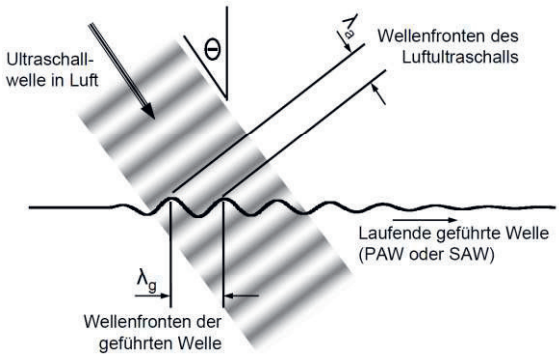


Abb. 1: . Erzeugung von Lambwellen mittels luftgekoppeltem Ultraschall

Bei der Schrägeinschallung erzeugen die einfallenden Luftultraschallwellen, Transversal- und Longitudinalwellen im dünnwandigen Prüfkörper. Diese überlagern sich und bilden sogenannte Lamb- oder Plattenwellen aus. Die Ausbreitungsgeschwindigkeit der Lambwelle ist abhängig von der Anregungsfrequenz, der Plattendicke und der Plattensteifigkeit.

Wenn eine Welle mit einer bestimmten Plattenwellengeschwindigkeit angeregt werden soll, lässt sich der notwendige Einschallwinkel mit dem Brechungsgesetz

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{c_{Luft}}{c_{Lamb}}$$

nach Snell berechnen. Der Winkel θ2 ist in Richtung der sich ausbreitenden Lambwellen cLamb = 0 und sin θ2 nimmt daher den Wert 1 an. Für den Einschallwinkel muss deshalb die Beziehung

$$\theta = \sin^{-1}\frac{c_{Luft}}{c_{Lamb}}$$

gelten. Bei richtig gewähltem Einschallwinkel, tritt am Empfänger eine höhere Schallamplitude auf, als bei Normaldurchschallung. Die sich ausbreitenden Plattenwellen, werden einerseits durch eine werkstoffabhängige Dämpfung abgeschwächt, andererseits verlieren sie Energie an das umgebende Medium. Dieses Phänomen wird als „Leaky-Lamb-Waves“ bezeichnet. Durch die Bewegung aus der Ebene heraus, entstehen Luftultraschallwellen, die auf beiden Plattenseiten empfangen werden können. Auf diese Weise ist auch eine einseitige Prüfung mit luftgekoppeltem Ultraschall möglich.

2. Automatisierte Winkelsteuerung

Der Einschallwinkel für die Prüfung von plattenartigen Prüfkörpern mit Lambwellen und luftgekoppeltem Ultraschall kann simulativ oder analytisch, bei bekannter Luftschallgeschwindigkeit, Werkstoffsteifigkeit, Prüffrequenz und Plattendicke bestimmt werden. In der Praxis erfolgt jedoch die Winkelbestimmung experimentell – durch sukzessives manuelles verändern des Einschallwinkels. Dieses Verfahren ist sehr stark vom Können des Prüfers abhängig und oft auch ein zeitintensiver Prozess. Aus diesem Grund wurde eine Prüfapparatur entwickelt, die ein automatisiertes Drehen und Verschieben der Prüfköpfe mittels Linearführungen und Goniometern zulässt. In Abb. 2. Ist die CAD-Zeichnung der Mechanik dargestellt.

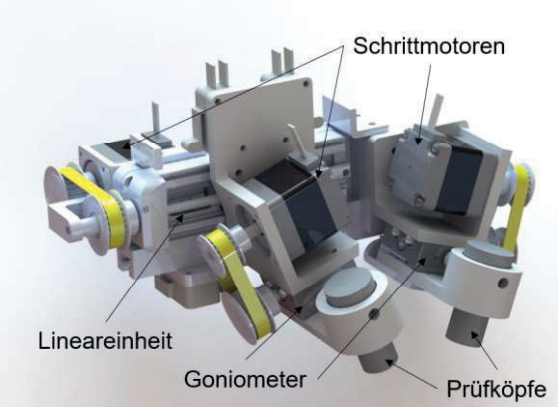


Abb. 2: . Konstruktion der mechanischen Einheit zur Winkelverstellung

3. Winkelsuchalgorithmus

Mittels einem dem Bisektionsverfahren [7] ähnelndem Verfahren, werden im ersten Schritt drei Punkte im gesamten Winkelbereich gleichverteilt und an jedem Punkt das maximale Amplitudensignal gemessen. Wenn einer der beiden äußeren Punkte den größten Amplitudenwert aufweist, liegt der notwendige Winkel außerhalb des Messbereichs. Wenn der mittlere Amplitudenwert am höchsten ist, werden die Amplitudenwerte der äußeren Punkte miteinander verglichen. In die Richtung mit dem höheren Amplitudenwert am Rand erfolgt der nächste Winkelschwenk. Gemessen wird genau zwischen dem äußersten und mittleren Punkt. Schrittweiße wird dieses Verfahren fortgeführt, bis die Maximalamplitude und somit das höchste Signal-zu-Rauschverhältnis erzielt wurde. In Abb. 3. ist der Algorithmus beispielhaft dargestellt. Gestartet wird mit der Messung des Amplitudensignals an den Punkten θ1 und θ2. Diese Winkel werden durch die mechanischen Begrenzungen der Goniometer bestimmt. Die Position von Punkt θ3 bestimmt sich aus der Hälfte des Abstands von θ1 und θ2. Die Amplitudensignale werden nach der Messung miteinander verglichen. Wenn A(θ1)>A(θ3) liegt der nächste Punkt in der Mitte von θ1 und θ3. Wenn A(θ2)>A(θ3) liegt der nächste Punkt in der Mitte von θ2 und θ3. Dieser Ablauf wird kontinuierlich fortgesetzt, bis entweder der Betrag von A(θn)-A(θn+2) und A(θn+1)-A(θn+2) eine zuvor festgelegte Grenze unterschreitet oder eine bestimmte Anzahl an Iterationen abgelaufen ist.

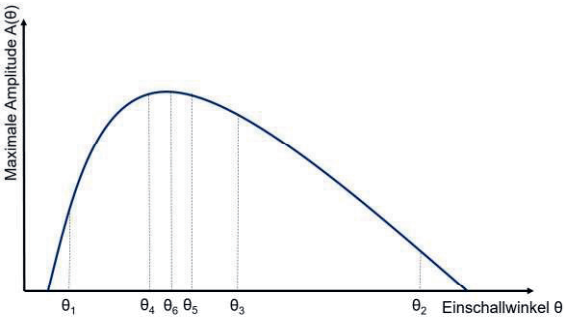
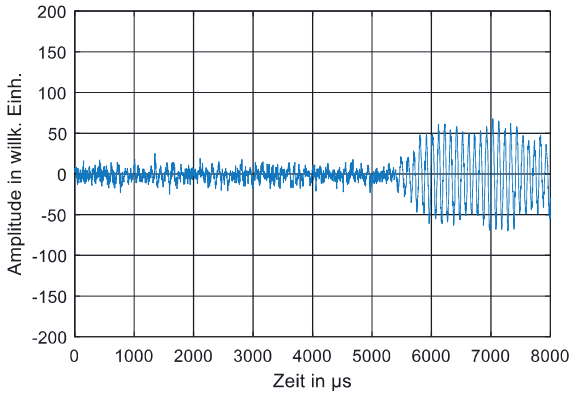


Abb. 3: .Darstellung des Algorithmus zum Finden des Hochpunkts der Amplitudenkurve in Abhängigkeit des Einschallwinkel

4. Ergebnisse

Um den Aufbau zu testen, wurde eine Platte aus kohlenstofffaserverstärktem Kunststoff (CFK) mit einer Dicke von 6 mm unter die Prüfkopfanordnung gelegt. Für die Versuche wurden Prüfköpfe des Herstellers ULTRAN Group, USA, mit einer Mittenfrequenz von 200 kHz eingesetzt (NGC200). Das verwendete Luftultraschallprüfgerät war das Airscope TT des Herstellers Dasel Sistemas, Spanien. Der Algorithmus für die Winkelfindung wurde mithilfe der Software Matlab ©, the MathWorks, inc., USA erstellt. Um den passenden Winkel zu finden, wurde das Ultraschallsignal zeitlich begrenzt. Betrachtet wurde ausschließlich der Zeitbereich von 0 µs bis 8000 µs nach dem Senden des Ultraschallpulses, um das Signal der „Leaky-Lamb-Waves“ zeitlich herauszufiltern. Zur Auswertung wurde dann der Maximalwert vom Betrag des Signals herangezogen. In Abb. 4. sind das Ultraschalsignal in Abhängigkeit der Zeit (Einschallwinkel 1°) und der Winkelsweep von 1° bis 23° dargestellt.



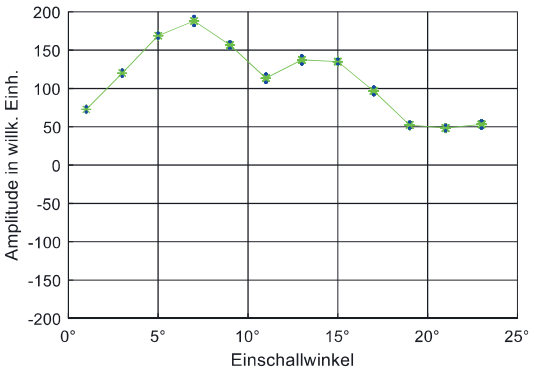


Abb. 4: .Oben: Zeitfenster für das Amplitudensignal (Einschallwinkel 1°), unten: Winkelscan von 1° bis 23°. Die höchste Amplitude wird bei einem Winkel von 7° erzielt

Wie zu erkennen ist, wird auch bei Winkeln größer und kleiner des Einschallwinkels von 7° ein Ultraschallsignal übertragen, das auch zur Prüfung verwendet werden könnte. Allerdings hebt sich dieses nicht so sehr vom Rauschen ab, wie das Signal mit optimiertem Winkel (SNR ~ 138). Gegenüber der Senkrechteinschallung kann also eine um den Faktor 4 (ca. 12 dB) höhere Signalstärke erzielt werden. In die CFK-Platte wurde bei der Herstellung ein Stück eines Handschuhs einlaminiert. Bei der Prepregbearbeitung werden Latexhandschuhe als Arbeitsschutz getragen, ein am klebrigen Harz versehentlich abgerissenes Handschuhstück ist deshalb ein typischer Herstellungsfehler. Ein C-Scan der Platte ist in Abb. 5. dargestellt. Durch das einlaminierte Gummihandschuhstück verändert sich die Steifigkeit der Platte lokal und die Amplitude reduziert sich signifikant. Dieser Herstellungsfehler ist deshalb deutlich in der linken oberen Ecke mit Amplituden kleiner 40 zu erkennen (dunkelblau).

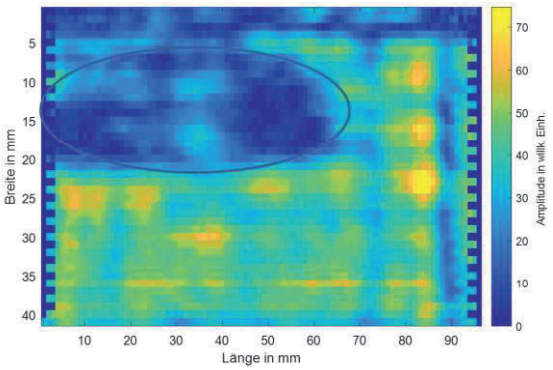


Abb. 5: .Darstellung des Algorithmus zum Finden des Hochpunkts der Amplitudenkurve in Abhängigkeit des Einschallwinkel

Literaturnachweis

[1] AMENDOLA, A., G. IANNUZZO, P. CERRETA und R. PINTO, 2011. Future aerostructure for the next generation green civil aircraft. In: *Aerodays 2011*, Madrid

[2] GREGORY R, 2003. Rapid, Sensitive Inspection of Marine Composites using Laser Shearography. In: *ACMC/SAMPE Conference on Marine Composites Plymouth, 11-12 September 2003.* ISBN: 1-870918-02-09.

[3] RAHAMMER, M, 2018. Nachweis von Impact-Schäden in Faserkunststoffverbunden mittels Resonanter Frequenzsweep Thermografie. Dissertation, Universität Stuttgart, Institut für Kunststofftechnik

[4] HOLMES,C., B.W. DRINKWATER und P.D. WILCOX, 2005. Post-processing of the full matrix of ultrasonic transmit-receive array data for non-destructive evaluation. In: NDT & E International, 38(8), 701-711. ISSN 09638695. Verfügbar unter: doi:10.1016/j.ndteint.2005.04.002

[5] GAAL, M., J. DÖRING, J. BARTUSCH, T. LANGE, W. HILLGER, G. BREKOW und M. KREUTZBRUCK, 2012. Ferroelektret-Prüfköpfe für berührungslose Ultraschallprüfung von faserverstärkten Kunststoffen. In: *DACH-Jahrestagung 2012 in Graz*

[6] BHARDWAJ, M.C., 2001. Non-Destructive Evaluation: Introduction of Non-Contact Ultrasound", In: *Encyclopedia of Smart Materials*, Hg. M. Schwartz, John Wiley & Sons, New York, S. 690-714

[7] MUNZ, C.D. und T. WESTERMANN, 2012. Numerische Behandlung gewöhnlicher und partieller Differenzialgleichungen, Springer-Verlag Berlin Heidelberg, Verfügbar unter: doi:0.1007/978-3-642-24335-6

[8] LEHMANN, N., A. TEGTMEIER, S. JÜTTNER, 2017. Automatisierte Luftultraschallprüfung von Widerstandspunktschweißverbindungen im Karosseriebau. In: *DGZfP-Jahrestagung 2017*, Koblenz

[9] LEHMANN, N., S. JÜTTNER, 2017. Automatisierte Luftultraschallprüfung im Automobilbau. In: *Seminar des FA Ultraschallprüfung 2017*, Berlin

DOI 10.5162/sensoren2019/6.4.1

Separation of locally determined work piece deviations and measurement uncertainties for structured-light scanning of customized polymer gear wheels

Andreas Michael Müller[1], Sebastian Metzner[1], Tino Hausotte[1], Dominik Schubert[2], Dietmar Drummer[2]
[1]Institute of Manufacturing Metrology (FMT), Naegelsbachstrasse 25, 91052 Erlangen, Germany
[2]Institute of Polymer Technology (LKT), Am Weichselgarten 9, 91058 Erlangen-Tennenlohe, Germany

Summary

During the manufacturing of work pieces, geometrical deviations from the intended nominal geometry of the designer are inevitable. The procedure of conformance testing defined in ISO 14253-1:2018-07 is used to ensure the function of a work piece by verifying the geometrical compliance with pre-defined tolerance specifications. Depending on the measurement setup used for the validation step and the accuracy of the manufacturing process, it is possible that the measurement uncertainty is large enough to have a significant influence on the conformance evaluation. The measurement uncertainty for optical measurement systems is influenced by the surface properties of the test specimen. This contribution aims to demonstrate the complete workflow for the determination of the single point uncertainties for a given measurement task in order to separate the local work piece deviations from the systematic and random components of the measurement uncertainty. It could be shown for a demonstration scenario that different necessary coloring methods of polymer gear wheels, which are required to enable measurements using structured-light scanning, have a measureable influence on the local distribution of the measurement uncertainty. This information could then be used for downstream processes in various use cases, e.g. for the improvement of holistic tolerance simulation models or the improvement of geometrical measurements using weighted regression analysis.

Keywords: structured-light scanning, work piece deviations, measurement uncertainty, gear wheel, colored polymer

The concept of the single point uncertainty

Generally, the results of any geometrical measurement of a work piece and the subsequent comparison against the nominal geometry (CAD model) as defined by the designer can be divided into three different contributions: a) the work piece deviations with respect to the nominal geometry, b) the (signed) systematic measurement error of the measurement system and c) the (unsigned) random measurement error of the measurement system. The measurement uncertainty is a positive non-zero parameter associated with every real measurement and can only be reduced but never be completely avoided. The “Golden rule” of metrology states that the measurement uncertainty shall be less than 10 % to 20 % of the tolerance [1–3]. This requires a close observation of the achieved measurement uncertainties for each geometric verification task. Recent research efforts at the Institute of Manufacturing Metrology (FMT) dealt with the question of how to determine the local distribution of the measurement uncertainties over the surface of a work piece and subsequently derive metrological benefits from that knowledge. The developed framework of the single point uncertainty (SPU) requires several inputs:

- A measurement object.
- The nominal geometry of the measurement object (CAD model).
- A measurement system targeted for the uncertainty examination and capable of recording complete areal measurements of the chosen measurement object.
- A statistically significant number of measurement repetitions (here 20, following guideline VDI/VDE 2630 [4] and ISO 15530-3:2011 [5]).
- A single reference measurement of the measurement object (reference geometry) performed by a reference measurement system in the sense of VIM [6] which is also capable of recording areal measurements of the chosen measurement object.

- A suitable sampling strategy dependent on the processed geometries and the computational resources available.

At first, a reference measurement of the measurement object needs to be created using the reference measurement system. The captured geometry (usually represented by a triangle mesh in the STL format) is then geometrically aligned against the nominal geometry (CAD) by means of a geometrical registration routine. One property of a reference measurement system is that the measurement uncertainties for that specific measurement task is known to be low compared to the expected geometrical deviations. Consequently, the observed geometrical differences between the nominal geometry and the reference measurement can be fully assigned to the work piece deviations. Then, a measurement series consisting of *n* repeated measurements of the same measurement object is created, using the measurement system chosen for the uncertainty evaluation. Each measurement is then geometrically registered against the reference measurement. After the alignment of the surfaces, the calculation of the spatial measurement uncertainty distribution can be performed using different sampling strategies (Fig. 1). If the normal vectors of the reference geometry are trustworthy and the expected measurement uncertainty is rather small compared to geometric features on the surface, the sampling strategy “normal vector” can be used (Fig. 1, right). Starting from each sampling point SP_k on the reference geometry, the intersection distance of a constructed search ray in the direction of the vertex normal vector is calculated for each measurement repetition. Thus, the resulting sets of distances $\left(d^k = d_i^k = \{d_1^k, d_2^k, \dots, d_n^k\}\right)$ are effectively assigned to their respective nominal surface vertices SP_k, which effectively maps the computed distances onto the nominal geometry.

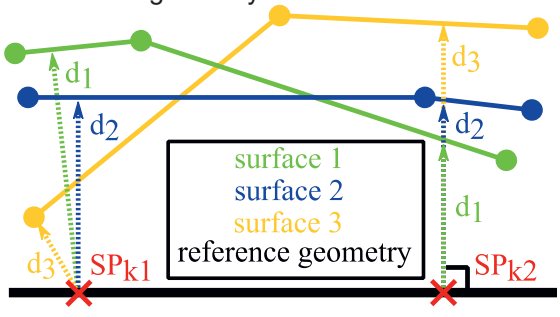


Fig. 1: Visualization of different sampling strategies: “shortest distance” (left) and “normal vector” (right).

The detected distances d^k per definition only represent contributions to the measurement uncertainty of the used measurement system, because the reference geometry by agreement represents the true geometry of the work piece. Consequently, the mean value of d^k represents the local systematic measurement error, while the local distribution of random measurement errors (precision) is determined by the standard deviation of d^k. After that, a suitable sampling method is applied to determine the local work piece deviations for each sampling point SP_k by comparing of the (same) reference surface against the nominal geometry. Summarizing the above, the single point uncertainty (SPU) is a framework to determine the geometric deviations and both components of the associated measurement uncertainty for each sampling point SP_k on the surface of a reference geometry, thus drawing a spatially resolved uncertainty map of a specific measurement task. It is advantageous if the reference surface is pre-processed in such a way that the triangles of the mesh have uniform properties, thus resulting in a homogeneous and preferably very dense sampling point distribution. Compared to the sampling strategy “normal vector”, the method “shortest distance” can result in better results in edge regions or if large distances are expected to be computed (Fig. 1). Nonetheless, in general both methods result in very similar uncertainty characteristics with the main difference being the computation requirements and the easiness with which certain run time sensitive optimization measures can be implemented [7]. In case no reference measurement is available, the method can be used nonetheless to compute the areal distribution of the random measurement error, while the work piece deviations and the systematic measurement error cannot be separated this way. It is important to note, that the determination of the SPU can be rather sensitive to geometrical misalignments resulting from a poor registration routine, which requires a certain diligence when deciding for a registration strategy. Extensive research at the Institute of Manufacturing Metrology (FMT) has revealed numerous different usage scenarios, in which the determination of the SPU can better the understanding of a measurement setup and consequently improve the measurement itself. It was successfully shown in [8] that locally varying systematic measurement errors of a real industrial X-ray computed tomography (CT) system could be corrected by the determination of the SPU of repeated simulated measurements, if the simulation framework is sufficiently adopted to the real CT system. The information about the local measurement precision was successfully used within the implementation of a weighted geometry element regression analysis (also called fit / fitting), resulting in improved

measurement results [9]. Fits are commonly used to assign standard geometry elements to measurement coordinates by minimizing a certain error function in order to determine the size and orientation of a specific geometry element (e.g. fitting of a cylinder to the point cloud representing a measured drilling). The method also allows the transfer of information about the measurement uncertainty associated with a certain measurement task to the downstream process of tolerance management [10]. Here, the measurement uncertainty can then be taken into account during the design of work pieces and determination of the associated tolerances. The determination and evaluation of the SPU yields the best results for measurement systems, which produce an areal representation of the work piece geometry (e.g. computed tomography, structured-light scanning in combination with a rotatory axis and focus variation). Nonetheless, similar evaluations are also possible for tactile coordinate measurement machines, although a different data processing pipeline needs to be implemented and utilized [11].

The gear wheel measurement object

The gears examined in this contribution (Fig. 2, Ø approx. 39 mm) were produced by injection molding using a Polyoxymethylen (POM) of the type Hostaform C9021 by Celanese Services Germany GmbH, Kelsterbach, Germany. POM is often used in gear applications due to its self-lubricating properties and resulting good dry running capabilities and low friction and wear [12]. The semi-crystalline structure of the material leads to a process dependent microstructure. In injection molding processes, an optically amorphous edge layer is formed due to the high cooling rate in the areas with contact to the mold material. This leads to the formation of crystalline nanostructures with size below the wavelength of visible light. As a result, the edge layer appears optically amorphous [13]. This semi-transparent edge layer leads to difficulties in optical measurements, especially with stripe light projection methods, due to the scattering of the projected lines into the material. Thus, the surface cannot clearly be determined. A further factor complicating optical measurements on this material is its gloss. To achieve useful measurements, the optical properties of the material have to be altered.

Using colorants i.e. pigments with or without carrier materials, the optical properties of polymer materials can be altered easily and cost-efficiently. Pigments and other nano-sized particles act as scattering centers, reducing the translucency of the optically amorphous skin layers. They can also decrease the amount of directed reflection, reducing gloss [14]. There are three main methods of colorizing polymers with pigments. Firstly, pigments can be added to the base material to produce pre-colored granulate. This is often used for large batches of material and frequently used colors such as black and white. Secondly, master batches, highly pigment-filled granules, are used to color smaller batches individually. Thirdly, liquid colors can be used instead of master batches, allowing for a more flexible mixing of pigments and a wider range of colors. The cost of coloration depends on choice of pigment, way of coloration and dosage. Master batches and liquid colors are more expensive than pre-dried material, but they are more flexible. Pigment cost can vary widely from inexpensive carbon to more expensive mineral or metal based effect pigments. Due to the complex interaction of the pigment particles, carrier materials and polymer materials, there is no single generalized effect of colorants on the resulting properties of polymers. However, most pigments act as a nucleating agent, increasing the degree of crystallinity, leading to higher shrinkage and therefore smaller parts [15].

To exclude influences of carrier materials, two pigment-based colorants were chosen for this research. The pigments were added by means of direct compounding in order to achieve a non-translucent, low-gloss surface with homogenous pigment distribution to improve the optical properties of the gear specimens for further optical measurements. The following two pigments were used to color the POM gear wheels (Fig. 2):

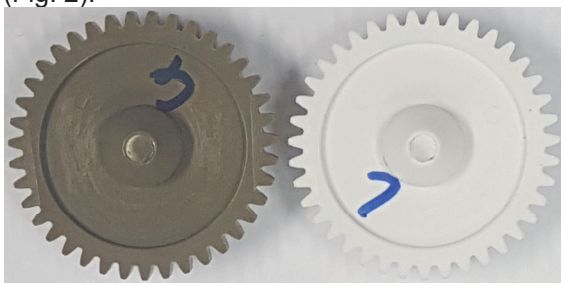


Fig. 2: Colored gear wheels: ID_{DG} (left) and ID_{W} (right)

i. A pigment mixture in the ratio of 1:1 of a yellow chrome antimony titaniumoxide rutile pigment ((Ti,Cr, Sb) O2) of the type Sicotan Gelb K 2001 FG, supplied by BASF Pigment GmbH, Besigheim, Germany, with a Carbon Black (C) pigment of the type Dei®Pow – Spezialschwarz, also provided by Deifel GmbH & Co. KG. The identifier (ID) for the corresponding measurement object is ID_{DG} (DG: dark green, Fig. 2, left).
ii. A white titanium dioxide (TiO2) pigment of the type Dei®Pow - Weiß, provided by Deifel GmbH & Co. KG, Schweinfurt, Germany. The

identifier (ID) for the corresponding measurement object is ID_W (W: white, Fig. 2, right).

An injection molding machine of the type Arburg 370 U-700-30-30 with a plasticizing unit with 18 mm screw diameter by Arburg GmbH & Co. KG, Loßburg, Germany, was used to produce the gear specimens. All gear wheels were produced using the same tool. Table 1 shows the processing parameters used for manufacturing.

Table 1: Overview of processing parameters

Parameter	Unit	Value
Mass temperature	°C	205
Mold temperature	°C	105
Injection speed	mm/s	15
Injection time	s	1.5
Holding pressure	bar	1200
Holding time	s	28.5

To ensure a correct alignment of the measured gear wheel geometries, a drilling Ø1.5 mm was arranged in one of the tooth roots (arbitrarily chosen) of each gear wheel before the measurements. This marker is then also captured by the measurement system and makes the correct alignment of the same teeth of different measurements possible during a manual registration step.

Measurement setup and data processing

Goal of this contribution was to evaluate the locally distributed single point uncertainties for gear wheel measurements using a structured-light scanner. The used system was GOM ATOS CORE 200 5M [16] in combination with a rotatory stage (Fig. 3).

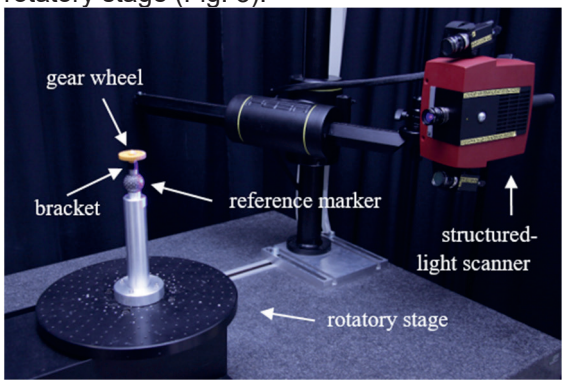


Fig. 3: Figurative representation of the measurement setup, consisting of a structured-light scanner (here GOM ATOS Compact Scan 2M / 300) in combination with a rotatory stage.

The sensor was controlled by the software GOM ATOS Professional 2018 [17]. The internal settings for the sensor were set to the least restrictive options possible, thus making sure that as many surface points as possible were recorded during the measurement. The illumination settings were determined by user-controlled selection of a target of interest on the live image once before each measurement series. During the measurements, all external light sources in the laboratory were dimmed as much as possible in order to ensure the exclusive illumination by the measuring device. The measurement object (gear wheel) was mounted onto the rotatory stage through its accessible rotatory axis. The used bracket was equipped with reference markers, which enabled the software GOM ATOS Professional 2018 to perform the referencing of each image with the angular position of the rotatory stage. The sensor is characterized by the following key data: measuring range 200 x 150 mm², lateral resolution 80 µm, measuring distance 250 mm and native camera resolution 5 million pixels. Before measurement, the sensor was calibrated using the calibration plate CP40 170 40760 consisting of 3 657 calibration points. A single measurement consisted of 200 single images (projections), which were equally distributed over a full rotation (360°) of the rotatory stage, with two different sensor positions each: One position captured the gear wheel sloping from above the gear wheel and the other from below with both angles roughly equaling 45°. As soon as all images were recorded from that position, the scanner was manually moved to the second position and the second set of 200 projections was recorded. The reference markers were visible to the scanner from both senor positions, which made the data fusion of several image stacks possible. Thus, the GOM software merged all of the 400 images together into one surface representation of the gear wheel in a triangulated mesh format (STL). The option “post processing” during the polygonization routine was disabled and thus all explicit filter operations disabled in order to achieve a preferably unaltered measurement result. Each measurement was repeated 20 times to ensure a statistically valid result for the following single point uncertainty calculations. The measurements were performed in a temperature-controlled environment with temperatures set to 20 °C ±1 K. Overall, two different gear wheels were examined, and thus 40 single measurements (each with two positions) were performed.

The reference measurements were captured using the industrial X-ray computed tomography system Zeiss METROTOM 1500 [18]. The gear wheel was positioned on the rotatory stage using the bracket already presented in Fig. 3. The upper part of the bracket consisted of polymer, which was easy to penetrate by X-rays and did not lead to CT-specific sources of errors due to

a possibly high material density. CT systems are in general very well suited for measuring polymer parts with high accuracy. To ensure a high quality measurement, time sensitive device settings were chosen in order to guarantee a very low measurement uncertainty of the reference measurement system. Because of the fact that the polymer coloration had no measureable impact on the absorption properties of the work piece, these following settings were used for all reference measurements: tube voltage 150 kV, tube current 400 µA, resulting nominal focal spot size 60 µm, X-ray filter 0.5 mm Cu, geometrical magnification 9.0 x, resulting voxel size 22.3 µm, integration time 1 000 ms, number of projection averaged 10, number of projections per measurement 2 050. The required surface determination was performed using VGStudio Max 3.2 [19] using the settings: automatic grey value threshold, search distance 16 voxels, remove all particles and voids, iterative surface determination. The measured surface could then be exported as a triangle mesh in the STL format using high quality settings. The measurement was subsequently registered against the nominal geometry (CAD) using VGStudio Max. Finally, the gear sprocket was extracted from the full surface using an own MATLAB program by removing all triangles with an edge point distance to the rotatory axis of smaller than 16 mm. The authors are aware of the fact, that an industrial CT system might not necessarily show improved measurement uncertainty characteristics compared to a structured-light scanner. Nonetheless, these measurement devices were chosen to demonstrate the principle approach for the determination and evaluation of the single point uncertainty framework as a tool in dimensional metrology.

As mentioned above, the sampling strategy “normal vector” requires trustworthy sampling vectors in order to compute the metrologically correct distances. CT measurements are often subject to high frequency noise with location wavelengths considerably higher than the structural resolution of the measurement system. Two important factors, which influence the structural resolution in a relevant way, are the geometrical magnification, which results in a certain voxel size, and the X-ray focal spot size. The structural resolution can reach sub-voxel values, and the actual value depends heavily on the measurement itself. Thus, in order to remove very high spatial frequencies and to improve the triangle mesh properties regarding triangle edge size and homogeneity, the generated surface from the reference measurement was undertaken a post-processing routine using GOM ATOS Professional 2018. First, the mesh was smoothed using the operation “Mesh → Smooth…” with settings “very large filter radius”, “very low detail sharpness” and “surface tolerance 10 µm”. After that, the homogeneity of the surface point distribution was improved using the operation “Mesh → other → Regularize…” with settings “surface tolerance 10 µm and “maximum edge length 50 µm”. Of course, these kinds of operations need to be used with special care, because any filtration of the reference measurement results in altered uncertainty values for the examined structured-light measurement device. Because of that, the effect of the described mesh operations was investigated further by comparing the filtered mesh against the unfiltered mesh using a conventional nominal-actual comparison (VGStudio Max). Figure 4 shows, that the deviations between the filtered and the unfiltered geometry nearly form a perfectly shaped Gaussian distribution with the expected value $\mu = 0$. The comparison against the regression result of the deviations using a Gaussian model confirms the underlying normal distribution. Consequently, we can state, following the central limit theorem [20], that the removed part of the geometry consists of noise, thus it is part of the random measurement errors. Because of that, and because of the fact that the removed parts of the geometry are clearly below the voxel size (22.3 µm) and the X-ray focal spot (60 µm) of the CT system, it is stated that no relevant measurement information of the geometry form was removed by the mesh post-processing operations.

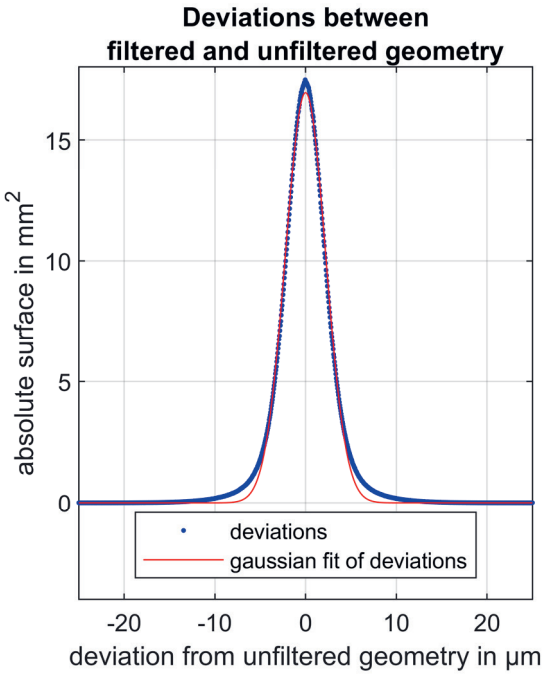


Fig. 4: Examination of the geometry altering mesh operations using a nominal-actual comparison and a Gaussian fit.

After the post-processing of the reference geometry, each of the 20 measurement repetitions was registered against that reference measurement using VGStudio Max. Following a first rough alignment, the rotatory symmetry was dissolved by arrangement of the drilling in one of the tooth roots. Then the gear wheel sprocket was extracted by removing all triangles with an edge point distance to the rotatory axis of smaller than 16 mm (same procedure as for the reference measurement before). The geometrical registration was finalized by a fine registration using the settings “quality level 50”, “consider current transformation”, “improved refinement” and “consider surface orientation”. Figure 5 shows the alignment of each single measurement with the reference geometry. It is clearly observable, that the drillings in each mesh lie on top of each other, thus confirming a valid registration result.

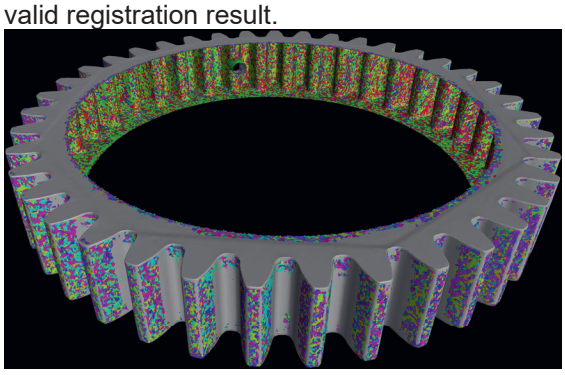


Fig. 5: Alignment of the sprocket of the reference geometry (gray) and the 20 measurement repetitions (various colors) after the registration routine.

After that, the work piece deviations were calculated by determining the distances from each sampling point of the reference geometry to the nominal geometry using the sampling strategy “shortest distance”. Then, the single point uncertainty was calculated by determination of the distances from each of the same sampling points to each of the measurements of the measurement series using the sampling strategy “normal vector” (as described in the first chapter). Finally, each sampling point SP_k was associated with a value describing the local work piece deviations *t* and a vector d^k characterizing the local measurement uncertainty.

Results

First, the geometrical work piece deviations were examined by calculating the local distances between the reference measurement and the nominal geometry (CAD). Fig. 6 shows that the gear exhibits some deviations at the tooth root as well as the tooth tip of around 200 µm, while the tooth flanks only show very small geometrical deviations compared to the nominal geometry. The systematic work piece deviations for the other gear wheel (ID_W) only deviated marginally from the shown results in Fig. 6, which means that the different colored compounds were not affected differently by the used manufacturing process parameters (see also Table 1).

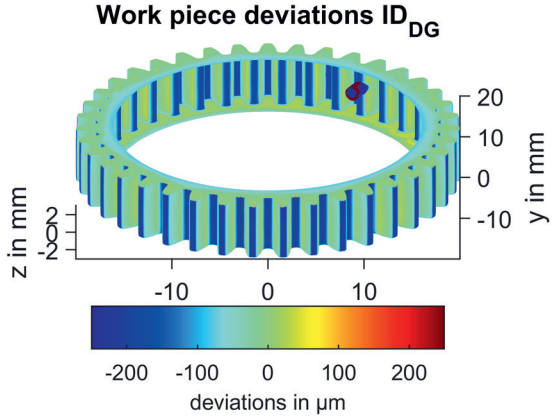


Fig. 6: Visualization of the systematic work piece deviations of a gear wheel

Figure 7 shows the expanded measurement uncertainty with a coverage factor of *k* = 2 for the gear wheel measurement using a structured-light scanner. The expanded uncertainty is calculated by the quadratic sum of the systematic and random measurement errors and is then multiplied with a coverage factor *k*. A coverage factor of 2 results in having a level of confidence of approx. 95 %, assuming a normal distribution [21]. The results in Fig. 7 show that the limited optical accessibility of the gear roots resulted in a substantial increase of the expanded measurement uncertainty over 100 µm.

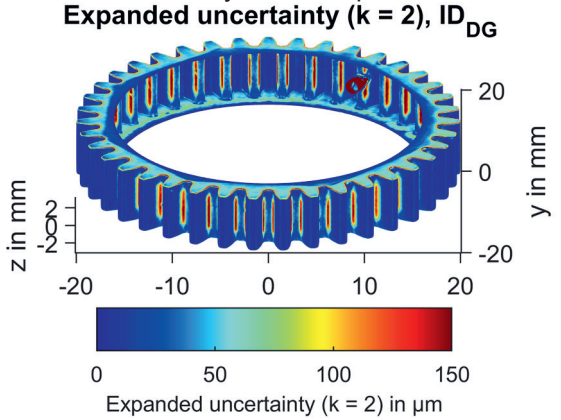


Fig. 7: Visualization of the expanded measurement uncertainty (k = 2) of gear wheel ID_{DG}

The gear wheel side also exhibits a larger measurement uncertainty because of the angled image capturing. The upper part of the gear flanks and the tooth tips were measured with a

comparably low expanded measurement uncertainty significantly below 50 µm. Table 2 presents a statistical view of the calculated deviations, which shows that the precision of the measurement setup was very high (P50 approx. 3 µm).

Table 2: Statistical view of the calculated work piece deviations and measurement uncertainty contributions (ID_{DG})

sys. work piece deviations percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
-206	-193	-72	-23	-3	19	32
rand. measurement error percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
1	1	2	3	3	5	8
sys. measurement error percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
-80	-60	-26	-11	-2	3	6
meas. uncertainty (k = 2) percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
3	4	8	24	53	121	163

Figure 8 shows the expanded uncertainty (k = 2) for the evaluated measurement series of gear wheel ID_W. It is observable, that the measurement uncertainty is higher (P50 nearly twice as large as for gear wheel ID_{DG}). The gear roots are also characterized by a very high measurement uncertainty.

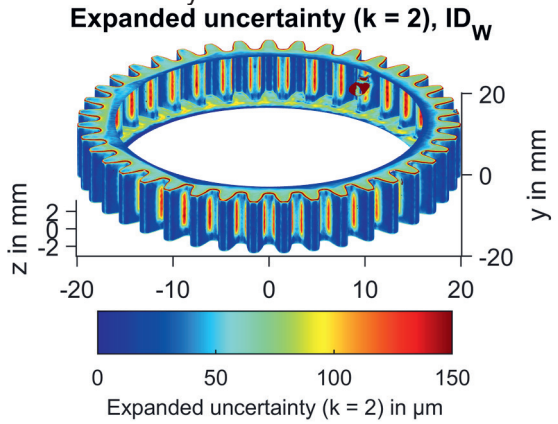


Fig. 8: Visualization of the expanded measurement uncertainty (k = 2) of gear wheel ID_W

Table 3 confirms that the expanded measurement uncertainty is higher for this coloration setting compared to gear wheel ID_{DG}, although the random part of the measurement error is even lower with P50 only reaching 2 µm. The systematic work piece deviations of both gear wheels are nearly identical with only minor differences.

Table 3: Statistical view of the calculated work piece deviations and measurement uncertainty contributions (ID_W)

sys. work piece deviations percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
-211	-199	-74	-26	-4	22	41
rand. measurement error percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
1	1	1	2	2	4	6
sys. measurement error percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
-97	-62	-35	-22	-12	-6	-3
meas. uncertainty (k = 2) percentiles in µm						
P01	P05	P25	P50	P75	P95	P99
8	12	24	44	71	124	201

Discussion

It could be shown that the measurement of POM gear wheels can in principle be realized by coloration during the manufacturing process. This was demonstrated at hand of two prototypes with different colorizations. Nonetheless, it was also observed that the measurement uncertainty varies on the measured surface, which is possibly influenced by the different inclination angles during the measurements. Especially the limited optical accessibility of the tooth roots presumably lead to increased uncertainty values. This shows that a global illumination setting during the measurement is not optimal for all geometrical features of the gear wheel. Although a complete measurement was possible, the achieved expanded measurement uncertainties of 24 µm (P50 of ID_{DG}) and 44 µm (P50 of ID_W) might be insufficient nonetheless. Depending on the defined tolerances for different geometrical features on the polymer gear wheels surface, the “Golden Rule” of metrology might prevent the geometrical inspection of those gear wheels using structured-light scanners. The measurement uncertainty could still be reduced for the presented measurement system by increasing the number of sensor positions and introducing additional illumination settings but the measurement expense would then rise quite substantially. It is worth noting that the calculated precision of the measurements was very good compared to the systematic measurement error. This could be caused by extensive averaging of measurement data due to the high number of recorded projections. Lastly it is noted, that the used reference measurement setup (CT) also exhibits some kind of measurement uncertainty, which was not taken into account within this contribution. A more ac-

curate determination of the single point measurement uncertainty of structured-light scanners using CT as a reference measurement system must also take this error source into account.
The presented work demonstrated the possibility for a better understanding of a certain measurement setup regarding the achieved measurement uncertainty. It is not uncommon in industrial applications that a single measurement setup is operated to monitor the output of a production line. The single point uncertainty provides a flexible tool to verify the measurement results of different measurement setups. Further research efforts could be invested to determine improved coloration options in order to find optimally coordinated optically imaging mechanisms between the gear wheel surface and the projected light of the structured-light scanner.

Acknowledgments

The authors would like to thank the German Research Foundation (DFG) for supporting the research project “FOR 2271 process-oriented tolerance management based on virtual computer-aided engineering tools” under grant numbers HA 5915/9-1 and DR 421/21-1 and for the financial support of the acquisition of the CT system "Zeiss Metrotom 1500" through Grant No. 324672600.

References

[1] W. Knapp, “Tolerance and uncertainty,” in *Laser Metrology and Machine Performance V*, 2001, pp. 357–366.
[2] M. Smouni, *CZ 3D Metrology Solutions: The Golden Rule - Made by ZEISS.* [Online] Available: https://www.cz-3d-metrology.com/2017/12/14/the-golden-rule-made-by-zeiss/. Accessed on: Apr. 17 2019.
[3] G. Berndt, E. Hultzsch, and H. Weinhold, “Funktionstoleranz und Meßunsicherheit,” *Wissenschaftliche Zeitschrift der Technischen Universität Dresden*, vol. 17, no. 2, p. 465, 1968.
[4] *Computed tomography in dimensional measurement*, VDI/VDE 2630 Part 2.1, 2015.
[5] *Geometrical product specifications (GPS) -- Coordinate measuring machines (CMM): Technique for determining the uncertainty of measurement*, ISO 15530-3:2011, 2018.
[6] *International vocabulary of metrology: Basic and general concepts and associated terms (VIM) German-English version ISO/IEC-Guide 99:2007, Corrected version 2012,* 4th ed. Berlin: Beuth, 2012.
[7] A. M. Müller and T. Hausotte, “Comparison of different measures for the single point uncertainty in industrial X-ray computed tomography,” in *9th Conference on Industrial Computed Tomography*, 2019.
[8] A. M. Müller, F. Wohlgemuth, and T. Hausotte, “Simulation-based correction of systematic errors for CT measurements,” in *8th Conference on Industrial Computed Tomography*, 2018.
[9] A. M. Müller and T. Hausotte, “Utilization of single point uncertainties for geometry element regression analysis in dimensional X-ray computed tomography,” in *9th Conference on Industrial Computed Tomography*, 2019.
[10] A. M. Müller, T. Oberleiter, K. Willner, and T. Hausotte, “Implementation of parameterized work piece deviations and measurement uncertainties into performant meta-models for an improved tolerance specification,” *Proceedings of the International Conference on Engineering Design, ICED*, 2019.
[11] A. M. Müller, S. Metzner, and T. Hausotte, “Characterization of the single point precision of steel gear wheel measurements using tactile coordinate measurement machines in scanning mode: accepted contribution,” in *20. GMA/ITG-Fachtagung Sensoren und Messsysteme*, 2019.
[12] E. Baur, T. Osswald, and N. Rudolph, *Plastics handbook: The resource for plastics engineers,* 5th ed. Munich: Hanser, 2019.
[13] G. Menges, G. Wübken, and B. Horn, “Einfluß der Verarbeitungsbedingungen auf die Kristallinität und Gefügestruktur teilkristalliner Spritzgußteile,” *Colloid and Polymer Science*, vol. 254, no. 3, pp. 267–278, 1976.
[14] C. Raymond and S. Ronca, “Relation of Structure to Electrical and Optical Properties,” in *Brydson's Plastics Materials: Eighth Edition*, 2016, pp. 103–125.
[15] A. Müller, *Coloring of Plastics: Fundamentals - Colorants - Preparations*. München: Hanser, Carl, 2012.
[16] GOM GmbH, *ATOS Core – Optischer 3D-Scanner.* [Online] Available: https://www.gom.com/de/messsysteme/atos/atos-core.html. Accessed on: Apr. 17 2019.
[17] GOM GmbH, *ATOS Professional*, 2018.
[18] Carl Zeiss Industrielle Messtechnik, *3D Röntgenmessung in der Qualitätssicherung: ZEISS METROTOM.* [Online] Available: https://www.zeiss.de/messtechnik/produkte/systeme/computertomographie/metrotom.html. Accessed on: Apr. 17 2019.
[19] Volume Graphics GmbH, *VGSTUDIO MAX*, 2019.
[20] M. Hazewinkel, V. I. Bityutskov, R. V. Gamkrelidze, and J. V.'e. Prohorov, *Encyclopaedia of mathematics*. Dordrecht: Kluwer academic, 1988-1994.
[21] Joint Committee for Guides in Metrology (JCGM/WG 1), *Guide to the expression of uncertainty in measurement: JCGM 100:2008 - GUM 1995 with minor corrections*, 2008.

DOI 10.5162/sensoren2019/6.4.2

Entwicklung eines Systems für hochpräzise Mikroverzahnungsmessungen

Stephan Jantzen[1§], *David Metz*[2§], *Martin Stein*[1], *Karin Kniel*[1], *Andreas Dietzel*[2]

1 Physikalisch-Technische Bundesanstalt (**PTB**), Fachbereich Koordinatenmesstechnik, Bundesallee 100, 38116 Braunschweig, Deutschland
2 Technische Universität Braunschweig, Institut für Mikrotechnik (**IMT**), Alte Salzdahlumer Str. 203, 38124 Braunschweig, Deutschland
§ Beide Autoren haben zu gleichen Teilen zu der Arbeit beigetragen

Kurzfassung

Das von uns realisierte System besteht aus drei unabhängigen Entwicklungen: dem IMT-PTB-Mikrotaster, einem Mikroinnenverzahnungsnormal und einem Mikroenvironment. Der IMT-PTB-Mikrotaster ist ein auf Silizium -Feder-Parallelogrammen basierender Taster, der in kommerzielle Koordinatenmessgeräte integriert wird, um Mikromessungen zu ermöglichen. Bei dem Mikroinnenverzahnungsnormal handelt es sich um einen von der PTB, dem nationalen Metrologieinstitut Deutschlands, kalibrierten Prüfkörper mit Moduln zwischen 0,1 mm und 1 mm. Die dritte Systemkomponente, das Mikroenvironment, ist ein multifunktionales System, welches in Koordinatenmessgeräte integriert wird, um die Messung von Mikrostirnrädern zu verbessern und beschleunigen. Es besteht aus Aufspannung, Schutzumhausung, Reinigungs- und Sensorsystem sowie zwei Mikroskopkameras.
Der kombinierte Einsatz von Mikrotaster, -normal und -environment ermöglicht eine rückgeführte, optimierte Messung von Mikroverzahnungen auf kommerziellen Koordinatenmessgeräten. Das System ist eine kostengünstige Alternative zu Mikrokoordinatenmessgeräten.

Keywords: Fertigungsmesstechnik, Mikroverzahnung, membranbasierter isotroper Mikrotaster, Mikroenvironment, Mikroinnenverzahnungsnormal

Motivation

Die zuverlässige Messung von Mikroverzahnungen ist essentiell für deren Qualitätskontrolle in der Fertigung. Mögliche Abweichungen von der Sollgeometrie bei Verzahnungen können zu erhöhtem Verschleiß, Geräusch, Ausfall des Getriebes und verringerter Effizienz führen. Die taktile Messung von Verzahnungen ist weit verbreitet aufgrund der meist konkurrenzlosen Genauigkeit, den etablierten Technologien und der langen Entwicklungshistorie, was sich unter anderem in der umfangreichen und anerkannten Normung zeigt.
Messungen von Werkstücken mit Merkmalen im Sub-Millimeterbereich erfordern angepasste Messsysteme und -strategien. Gründe hierfür sind Schwierigkeiten bei der Handhabung, die benötigten kleineren Tastelemente und -kräfte und die hohen Anforderungen an die Kontaminationskontrolle. Aus diesem Grund haben die Physikalisch-Technische Bundesanstalt (PTB) und das Institut für Mikrotechnik der TU Braunschweig (IMT) in einem gemeinsamen Projekt ein System entwickelt, das die Herausforderungen in der Mikroverzahnungsmesstechnik überwindet.
Das entwickelte System besteht aus drei Komponenten:

- Isotroper IMT-PTB-Mikrotaster mit großen Auslenkungsbereichen
- Kalibriertes Mikroinnenverzahnungsnormal aus Hartmetall bzw. Titan
- Portables Mikroenvironment zur Prozessverbesserung und -beschleunigung

IMT-PTB-Mikrotaster

Der IMT-PTB-Mikrotaster beruht auf dem patentierten Prinzip von drei seriell angeordneten Silizium-Feder-Parallelogrammen (siehe Abbildung 1) [1]. Die Membrandicken sind in der Größenordnung von 25 µm. Anpassungen der einzelnen Parallelogrammgeometrien ermöglichen eine spezifische Optimierung von Isotropie und Steifigkeit in allen Raumrichtungen.
Der Parallelogrammaufbau ist die mikrotechnische Weiterentwicklung der in der Koordinatenmesstechnik etablierten Tastkopfaufbauten für Koordinatenmessgeräte (KMG) [2]. Die mikro-

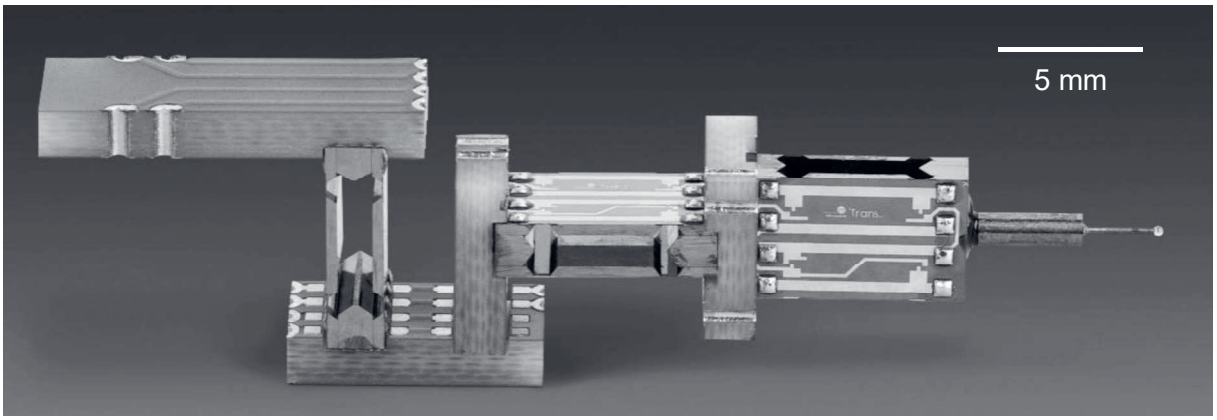


Abbildung 1: Aufbau des IMT-PTB-Mikrotasters (ohne Gehäuse). Die Steifigkeit des Tasters beträgt circa 0,3 mN·µm $^{-1}$. Die Verbindungsstücke zwischen den Parallelogrammen führen die elektrischen Signale weiter bis zur Kontaktierung am Gehäuse.

technische Adaption nutzt die Vorteile der Siliziumtechnik und integriert die Sensoren in die Festkörpergelenke (siehe Abbildung 2) [3]. Damit wird eine Miniaturisierung des Funktionsprinzips erreicht.
In die Siliziumfestkörpergelenke eindiffundierte Piezowiderstände messen die Auslenkung der Parallelogramme mithilfe von Wheatstone-Brücken. Mit dem neuartigen Prinzip sind Auslenkungen bis zu ± 400 µm möglich, was eine deutliche Verbesserung zu vergleichbaren membranbasierten Mikrotasterdesigns ist [4]. Der Taster kann mit Tastkugeldurchmessern zwischen 50 µm und 1 mm verwendet werden. Die exzellente Isotropie des Tasters aufgrund seines seriellen Parallelogrammaufbaus bietet messtechnische Vorteile, zum Beispiel wenn schräge Flächen angetastet werden.
Der IMT-PTB-Mikrotaster wurde in ein kommerzielles Verzahnungsmessgerät, die P40 der Firma Klingelnberg, erfolgreich integriert und experimentell verifiziert [5, 6, 7]. Die Integration läuft über einen in dem Messgerät vorhandenen A/D-Wandler. Die Messgerätsoftware ist angepasst, um die Auslenkungssignale des Mikrotasters anstelle der des Standardtastkopfs auszuwerten. Messungen laufen daher mit der Standardsoftware und den gleichen Messroutinen wie in der Standardkonfiguration, was eine sehr gute Bedienbarkeit ermöglicht.
Der IMT-PTB-Mikrotaster wurde anhand von Vergleichsmessungen an kalibrierten Prüfkörpern (Kugel, Würfel und dem Mikroinnenverzahnungsnormal (siehe folgenden Abschnitt)) verifiziert [8]. Die Messergebnisse zeigen Differenzen zu den Kalibrierwerten im Sub-Mikrometerbereich, was den Erfolg von Design und Realisierung bestätigt.

Mikroinnenverzahnungsnormal

Das in der PTB entwickelte Mikroinnenverzahnungsnormal [9] ermöglicht eine Erweiterung der Rückführung von Verzahnungsmessungen (zum Beispiel mit dem Substitutionsverfahren, siehe auch DIN EN ISO 15530-3).
Vier verschiedene Moduln sind an dem Prüfkörper realisiert: 1 mm, 0,5 mm, 0,2 mm und 0,1 mm (siehe Abbildung 3). Die Kalibrierung wurde auf dem Mikro-Koordinatenmessgerät Zeiss F25 mit einem Rubintastkugeldurchmesser von 120 µm durchgeführt. Die erweiterten Messunsicherheiten ($k = 2$) der taktilen Kalibrierung liegen zwischen 0,3 µm und 0,6 µm für die Winkel-, Form- und Gesamtabweichungen an den gemessenen Profilen und Flankenlinien.

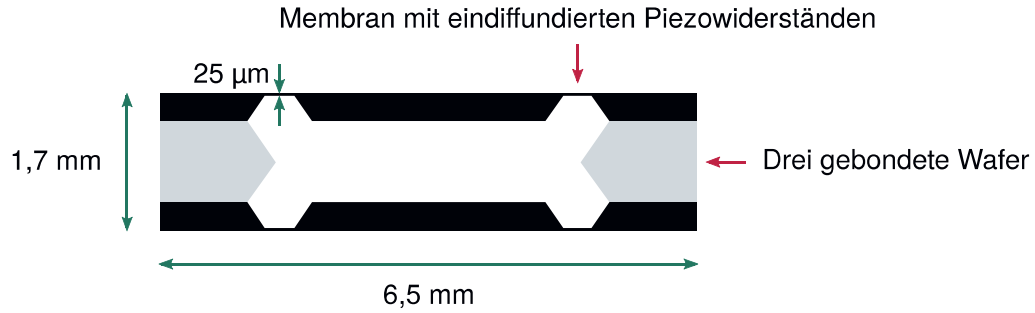


Abbildung 2: Querschnitt eines einzelnen Parallelogramms. Durch die Geometrie besitzen die Parallelogramme eine geringe Steifigkeit in einer Raumrichtung (Das Verhältnis der Steifigkeiten zwischen Vorzugsrichtung und den beiden anderen Achsen beträgt bis zu 1:300).

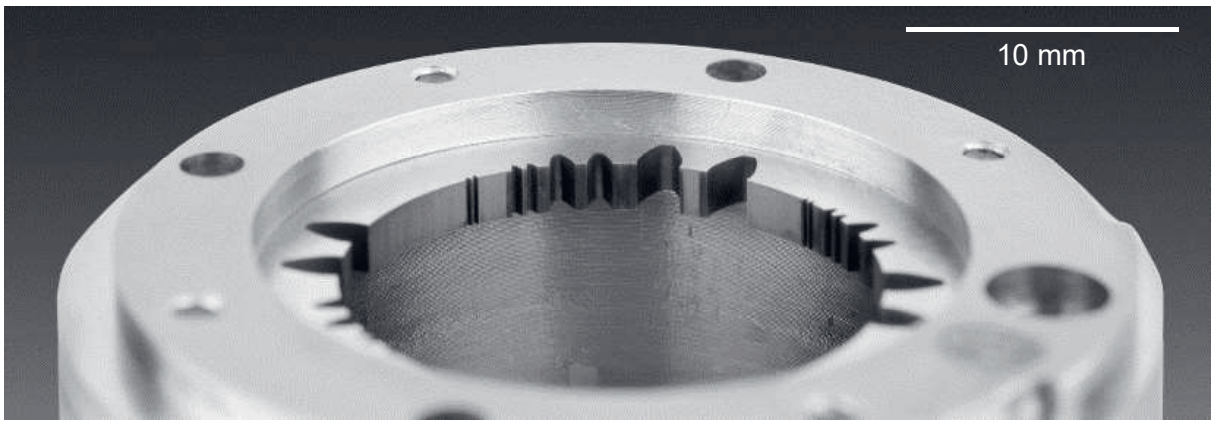


Abbildung 3: Detailaufnahme des Mikroinnenverzahnungsnormals.

Die Messunsicherheiten wurden GUM-konform ermittelt [10]. Die Eingangsparameter wurden mithilfe von Experimenten, Monte-Carlo-Simulationen und Expertenwissen bestimmt.
Der Prüfkörper besitzt einen Fußkreisdurchmesser von 20 mm und eine Zahnbreite von 2 mm. Die Zahnbreite orientiert sich an der Schaftlänge des Tasters. Ein möglicher Ansatz für beliebig große Zahnbreiten bei limitierten Mikrotaster-Schaftlängen ist das Stapeln und Verbinden von mehreren einzeln kalibrierten Prüfkörpern.
Zwei Versionen des Normals wurden in der PTB mit Drahterosion gefertigt. Eine Version besteht aus *Hartmetall* und ist aufgrund seiner Materialeigenschaften für taktile Messungen gedacht. Allerdings ist die Oberflächenhärte, die wir durch das Fertigungsverfahren zurzeit erreichen noch verbesserungswürdig. Die zweite Version besteht aus *Titan* und ist damit auch geeignet für Messungen mit Computertomografie. Beide Versionen wurden kalibriert.
Nach der taktilen Kalibrierung haben wir Vergleichsmessungen mit den Prüfkörpern mit unterschiedlichen Messprinzipien und -systemen durchgeführt. Hierbei ermöglichen die Computertomografie sowie die optischen Verfahren mittels Durchlicht und Fokusvariation auch die Messung des 0,1 mm Moduls.
Die Messung des Normals mit *Computertomografie* lieferte eine Voxelgröße von $(22{,}7\ \mu m)^3$ bedingt durch das große Verhältnis von Außendurchmesser (Ø 40 mm) zu Zahngrößen (Moduln von 1 mm bis 0,1 mm). Eine Messung bestehend aus 2500 Projektionen dauerte 167 Minuten. Die Abweichungen zu den Kalibrierwerten lagen im Bereich von 4 µm.
Bei der Messung mit *Fokusvariation* konnte nicht der ganze Prüfkörper erfasst werden (wie es bei Computertomografie möglich ist), sondern lediglich die Innenkontur und Teile der oberen Planfläche. Daher musste die Definition des Werkstückkoordinatensystems im Vergleich zu der taktilen Kalibrierung angepasst werden. An einzelnen Stellen konnte keine Oberflächeninformation erfasst werden, was zu Fehlstellen führte. Eine Messung dauerte 106 Minuten. Die Abweichungen zu den Kalibrierwerten lagen im Bereich von 4 µm und waren damit vergleichbar mit den Computertomografie-Ergebnissen.
Die *taktilen* Vergleichsmessungen wurden auf der Klingelnberg P40 durchgeführt. Hierbei verwendeten wir das Standardtastsystem als auch den IMT-PTB-Mikrotaster. Die Messstrategie wurde – soweit möglich – von der taktilen Kalibrierung übernommen. Die Messzeiten betrugen beispielsweise circa 20 Minuten sechs Lücken des 0,5 mm Moduls inklusive Prüfbunde. Die Abweichungen zu den Kalibrierwerten lagen im Sub-Mikrometer-Bereich.
Insgesamt bewerten wir die Vergleichsmessungen positiv. Die Ergebnisse und Erfahrungen verdeutlichen die Anforderungen, Vor- und Nachteile der jeweiligen Messverfahren [11].

Mikroenvironment

Das portable Mikroenvironment wurde entwickelt, um das Messsystem – bestehend aus IMT-PTB-Mikrotaster und Klingelnberg P40 – hinsichtlich der Praktikabilität für die Mikro-Koordinatenmesstechnik zu verbessern. Das Mikroenvironment besteht aus Lösungen für Aufspannung, Reinigung, Kontaminationskontrolle und Prozessüberwachung.
Die Aufspannung für Mikrostirnräder basiert auf einer elastischen Kunststoffschnur, deren Außendurchmesser dem Innendurchmesser der Durchgangsbohrung des Mikrostirnrades entsprechen sollte. Am Schnurende wird eine Kugel aufgeschmolzen (siehe Abbildung 4). Die Schnur fädelt man durch die Durchgangsboh-

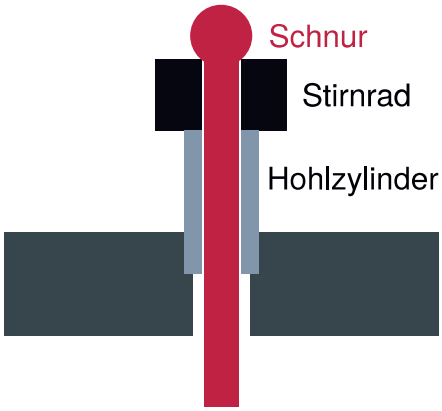


Abbildung 4: Prinzipskizze der Aufspannung für Außenverzahnungen mithilfe einer elastischen Schnur.

rung des Werkstücks und durch einen Hohlzylinder. Das Werkstück wird damit auf den Hohlzylinder zentriert (Präzision von wenigen Mikrometern) und fixiert, indem eine definierte und wiederholbare Spannkraft mithilfe von Gewichten erzeugt wird [12]. Durch die hohe Elastizität der Schnur (zum Beispiel aus Fluorcarbon) werden Spannungsspitzen, die dem Werkstück schaden können, im Normalfall vermieden (Ausnahmen können nachgiebige Werkstücke bei gleichzeitig hoher Spannkraft sein).

Die experimentell von uns erprobte Lösung für eine maschinennahe Reinigung von Werkstücken und Tastern ist das CO_2-Schneestrahlen [13]. Das Reinigungsverfahren ist schnell, effektiv und maschinennah einsetzbar. Nachteile sind die tiefe Temperatur des Strahls bestehend aus festem und gasförmigen CO_2 (–78 °C, was zu Kondensation an Oberflächen führen kann) und der hohe Strahldruck (vergleichbar mit einer Druckluftreinigung). Eine Beschädigung von Oberflächen ist aufgrund der geringen Härte der Trockeneispartikel ausgeschlossen, weshalb dieses Reinigungsverfahren auch bei Optiken angewendet wird [14].

Mögliche Optimierungen des Reinigungsverfahrens sind gepulste Gaszufuhr, Verwendung eines Schutzgases und Reinigung mit Dry-Box. Die Eignung des Verfahrens für die Reinigung von Mikrotastern haben wir ebenfalls untersucht. Die üblichen Schwachstellen von membranbasierten Mikrotastern sind die fragilen Membranen. Hier muss verhindert werden, dass der Krafteintrag durch den Gasstrom kritisch wird. Wir konnten dies durch Anpassung der Strahlrichtung, Pulsen und Klemmung des Tasterschaftes erreichen. Einen negativen Einfluss der niedrigen Temperaturen konnten wir nicht beobachten. Je nach Materialpaarung und für spezielle Aufbauten kann dieser Aspekt allerdings kritisch sein.

Eine Schutzumhausung mit integriertem Mikroskopkamera- und Sensorsystem verbindet die Abschirmung der direkten Messumgebung mit Überwachung der Umweltparameter (Temperatur, Feuchte und Partikeldichte) und erleichtert das Einrichten der Messung.

Die Schutzumhausung besteht aus einem Aluminiumprofilrahmen, der an drei Seiten mithilfe von Acrylglas geschlossen ist (s. Abbildung 5). Die verbleibenden offenen Seiten ermöglichen den Zugang zu Drehtisch, Tastkopf und Taster-wechselbank. Aufgrund der großen Flächen und der Nähe zum Werkstück und Mikrotaster sollten die Acrylglasscheiben antistatisch beschichtet sein (ESD-Material).

Die Schutzumhausung arbeitet passiv. Aktive Filterung, Temperierung und Luftströmung haben wir zur Reduzierung von Gewicht, Kosten und Entwicklungszeit verworfen.

Die Leistungsfähigkeit der passiven Lösung wurde in Experimenten mit Partikelfallen bestätigt: Die Partikeldichte sank um 32 % bei Verwendung der Schutzumhausung im Vergleich zur Standardkonfiguration der Maschine. Bei den Experimenten wurden jeweils zwei Partikelfallen aufgestellt. Eine Partikelfalle wurde maschinennah aufgestellt, um das Grundniveau der Partikelbelastung für jede Versuchsreihe zu ermitteln. Die zweite Partikelfalle platzierten wir neben dem Drehtisch der Maschine, sodass diese bei Verwendung der Schutzumhausung miteingeschlossen wird. Die Auswertung der

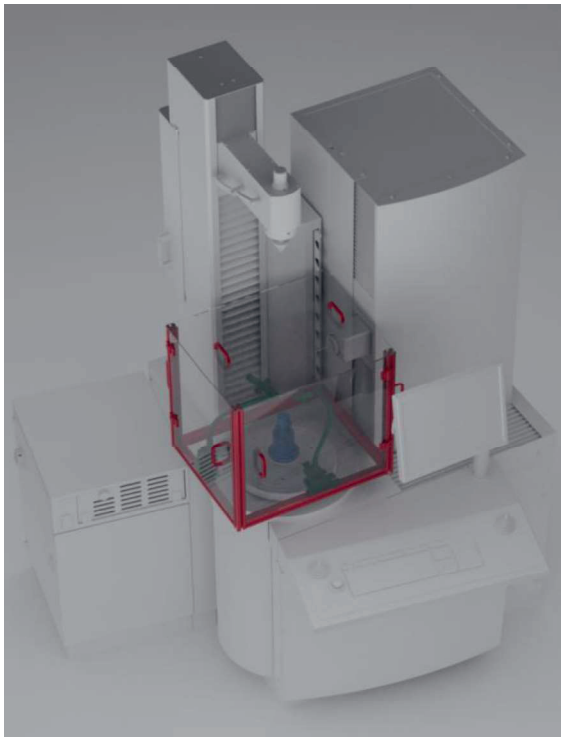


Abbildung 5: Digitales Modell des Mikroenvironments (in Farbe) integriert auf der Klingelnberg P40 (in grau)

Partikelfallen erfolgte über eine automatisierte Partikelerkennung und -analyse mithilfe eines Lichtmikroskops und dessen Software.

Diskussion

Die Messung von Mikrogeometrien ist ein wirtschaftlich wichtiger Prozess, was die Weiterentwicklung auf diesem Gebiet der Metrologie antreibt. Optische Messtechnik und dimensionelle Computertomografie haben in den letzten Jahren eine enorme Weiterentwicklung erfahren. Trotzdem ist die taktile Messtechnik oft weiterhin die beste Wahl, da sie zuverlässige, rückgeführte Resultate mit den meist geringsten Messunsicherheiten liefert. Multisensor-Koordinatenmessgeräte vereinen beispielsweise die Vorteile optischer und taktiler Technologien.
Optische Verfahren können bei Verwendung von Linien- oder Flächensensoren Vorteile bezüglich Geschwindigkeit bieten. Insbesondere die Computertomografie bietet durch die ganzheitliche Erfassung der Werkstücke (inklusive innenliegender Geometrien und Einschlüsse) großes Potenzial.
Hindernisse für die Mikromesstechnik im Allgemeinen sind die Notwendigkeit von speziellen Messverfahren und -systemen, das Fehlen von geeigneten Normalen und die hohen Anforderungen an Messunsicherheit und Zuverlässigkeit.
Neben der Weiterentwicklung von Hardware kommt der Software eine entscheidende Bedeutung zu. (Zukunfts-)Themen, die allerdings nicht nur beschränkt sind auf den Mikrobereich sind, sind Auswertealgorithmen, digitale Zwillinge, intelligente und lernende Systeme, Optimierung von Prüfplänen, Schnittstellen zum Datenaustausch, Kommunikation mit Fertigungsmaschinen oder anderer Software zur Steuerung von Unternehmensprozessen [15, 16, 17, 18, 19].

Zusammenfassung

Drei Systemkomponenten für die Mikrokoordinatenmesstechnik wurden entwickelt. Neben dem IMT-PTB-Mikrotaster haben wir ein Mikroinnenverzahnungsnormal und ein Mikroenvironment realisiert. Alle drei Komponenten zielen darauf ab, kommerzielle Koordinatenmessgeräte für die Mikromesstechnik zu erweitern.
Der IMT-PTB-Mikrotaster:

- ist isotrop
- ermöglicht Auslenkungen bis zu ± 400 µm in allen Achsen
- kann mit Tastkugeldurchmessern zwischen 50 µm und 1 mm verwendet werden
- hat Antastkräfte von circa 30 mN bei einer Auslenkung von 100 µm

Das Mikroinnenverzahnungsnormal ist ein neuartiger, rückgeführter Prüfkörper mit Messunsicherheiten im Sub-Mikrometer-Bereich. Es wurden eine Titan-Version für die Computertomografie und eine Hartmetall-Version für taktile und optische Messungen entwickelt.
Das Mikroenvironment vereint Komponenten zur Aufspannung von Mikrostirnrädern, CO_2-Schneestrahlen, Schutz vor Kontamination der direkten Messumgebung einschließlich Werkstück und Taster, Monitoring mithilfe von zwei Lichtmikroskopkameras und Erfassung von Umweltparametern (Temperatur, Feuchte, und Partikeldichte).
Die kombinierte Verwendung des Gesamtsystems soll eine hochpräzise, rückgeführte Mikroverzahnungsmessung mit kommerziellen Koordinatenmessgeräten ermöglichen.

Ausblick

Für die Zukunft planen wir die Integration des IMT-PTB-Mikrotasters in weitere Koordinatenmessgeräte wie zum Beispiel das in der PTB entwickelte Mikro-Koordinatenmessgerät [20]. Zudem streben wir eine Kommerzialisierung an. Weiterhin sind Messungen mit 50 µm Tastkugeln und weitere Tasterkonfigurationen in Stern-, T- oder L-Form in Vorbereitung. Diese Tasteranordnungen ermöglichen die optimierte Messung von komplexen Mikrogeometrien.
Zur Verbesserung der Oberflächenqualität des Mikroinnenverzahnungsnormals werden Fertigungsexperimente mit Schnellarbeitsstahl (HSS) und Superfeinkorn-Hartmetall durchgeführt. Damit erhoffen wir uns eine geringere Rauheit und eine größere Oberflächenhärte, die auch höchsten Flächenpressungen aufgrund der kleinen Antastkugeln standhält [21].

Danksagung

Wir danken Michael Neugebauer, Astrid Linkogel, Fabricio Borges de Oliveira und Dieter Schulz für die Durchführung der Vergleichsmessungen und Kalibrierungen.
Wir danken der DFG für die Projektförderung (KN 1181/1-1).
Wir danken Rudolf Meeß und Christian Wolpert für die Fertigung der Prüfkörper.

Literatur

[1] D. Metz, A. Dietzel, Mikrotaster und Verfahren zur Herstellung. Patent DE102017206145B3 https://patents.google.com/patent/-DE102017206145B3/de

[2] A. A. Weckenmann, T. Estler, G. Peggs, D. McMurtry, Probing systems in dimensional metrology CIRP Annals – Manufacturing Technology 53 (2) 657–684. DOI:10.1016/S0007-8506(07)60034-1.

[3] D. Metz, A. Dietzel, 3-d isotropic tactile microprobe based on a silicon parallelogram kinematic: From concept to fabrication, Journal of Microelectromechanical Systems (2018) 1–14. DOI:10.1109/jmems.2018.2877926.

[4] N. Ferreira, Entwicklung, Integration und Erprobung eines piezoresistiven 3D-Mikrotasters, Ph.D. thesis. Technische Universität Braunschweig. ISBN 978-3-8440-3546-9.

[5] D. Metz, S. Jantzen, D. Wessel, G. Mies, J. Lüdenbach, M. Stein, A. Dietzel, K. Kniel, Design of an isotropic microprobe and a microenvironment for high-precision coordinate metrology. Measurement Science and Technology, *submitted* (2019).

[6] D. Metz, N. Ferreira, J. Chaillot, M. Stein, K. Kniel, A. Dietzel, Integration of a piezoresistive microprobe into a commercial gear measuring instrument, Precision Engineering (2018). DOI:10.1016/j.precisioneng.2018.10.003.

[7] Klingelnberg, P 40 – Klingelnberg precision measuring center. https://www.klingelnberg.com/en/business-divisions/precision-measuring-centers/precision-measuring-centers/detail-page/product/p-40/

[8] S. Jantzen, Design of a high-precision microgear metrology system, Ph.D. thesis. Technische Universität Braunschweig. ISBN 978-3-95606-455-5

[9] S. Jantzen, M. Neugebauer, R. Meeß, C. Wolpert, A. Dietzel, M. Stein, K. Kniel, Novel measurement standard for internal involute microgears with modules down to 0.1 mm. Measurement Science and Technology 29 (12) (2018) 125012. DOI:10.1088/1361-6501/aae6f4.

[10] JCGM, Evaluation of measurement data – guide to the expression of uncertainty in measurement (GUM), Tech. rep., Joint Committee for Guides in Metrology (JCGM). http://-www.bipm.org/en/publications/guides/gum.html

[11] S. Jantzen, M. Stein, A. Dietzel, K. Kniel, Microgear measurement standards: comparing tactile, optical and computed tomography measurements, in: AGMA's 2018 Fall Technical Meeting, 18FTM16. https://members.agma.org/ItemDetail?iProductCode=18FTM16&Category=TECH%20PAPER

[12] S. Jantzen, R. Meeß, M. Stein, K. Kniel, A. Dietzel, Clamping of microgears with a compliant string, in: MacroScale 2017 – Recent developments in traceable dimensional measurements, 2017. DOI:10.7795/810.20180323B.

[13] S. Jantzen, T. Decarreaux, M. Stein, K. Kniel, A. Dietzel, CO_2 snow cleaning of miniaturized parts, Precision Engineering 52 (2018) 122–129. DOI:10.1016/j.precisioneng.2017.11.017.

[14] R. Kohli, K. L. Mittal (Eds.), Developments in Surface Contamination and Cleaning, Vol. 11, Elsevier Science, 2018.

[15] S. Boschert, R. Rosen, Digital Twin – The Simulation Aspect, Springer International Publishing, 2016, pp. 59–74. DOI:10.1007/978-3-319-32156-1_5.

[16] D. Heißelmann, M. Franke, T. Kistner, C. Schwehn, K. Wendt, Determination of measurement uncertainty by simulation, in: International Conference on Advanced Mathematical and Computational Tools in Metrology and Testing (AMTCM 2017), 2017. https://arxiv.org/-abs/1707.01091

[17] R. Schmitt, Das Internet of Production: Modellbasierte Qualitätssicherung in der digitalisierten Produktion, in: VDI-Berichte 2319, Messunsicherheit praxisgerecht bestimmen, 8. Fachtagung Erfurt, VDI Wissensforum GmbH, 2017.

[18] R. K. Leach, P. Bointon, X. Feng, S. Lawes, S. Piano, N. Senin, D. Sims-Waterhouse, P. Stavroulakis, R. Su, W. Syam, M. Thomas, Information-rich manufacturing metrology, in: Precision Assembly in the Digital Age, Springer International Publishing, 2018, pp. 145–157. DOI:10.1007/978-3-030-05931-6_14.

[19] D. Imkamp, R. Schmitt, J. Berthold, Blick in die Zukunft der Fertigungsmesstechnik. Technisches Messen 79 (10) 433–439. DOI:10.1524/teme.2012.0251.

[20] L. Yu, G. Molnar, S. Bütefisch, C. Werner, R. Meeß, H.-U. Danzebrink, J. Flügge, Micro coordinate measurement machine (uCMM) using voice coil actuator with interferometric position feedback, in: J. Tan, J. Lin (Eds.), Tenth International Symposium on Precision Engineering Measurements and Instrumentation, SPIE. DOI:10.1117/12.2511527.

[21] E. J. C. Bos, Aspects of tactile probing on the micro scale. Precision Engineering 35 (2) 228–240. DOI:10.1016/j.precisioneng.2010.09.010.

DOI 10.5162/sensoren2019/6.4.3

Modelle der Oberflächenform mechanisch bearbeiteter Werkstücke für die Messunsicherheitsbestimmung durch Simulation bei taktilen Koordinatenmessungen

Thomas Kistner[1], Dr. Dietrich Imkamp[1], Prof. Tino Hausotte[2]
[1]Carl Zeiss IMT GmbH, Carl-Zeiss-Straße 22, 73447 Oberkochen, Deutschland
[2]Lehrstuhl für Fertigungsmesstechnik, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Nägelsbachstraße 25, 91052 Erlangen

Zusammenfassung

Die Messunsicherheitsermittlung durch Simulation – bekannt unter dem Begriff „virtuelles Koordinatenmessgerät" – ist eine hoch entwickelte Methode im Bereich der Koordinatenmesstechnik. Durch Weiterentwicklung sind neue Funktionen zur Abschätzung von bisher nicht betrachteten Einflussgrößen hinzugekommen. Damit kann unter anderen der Einfluss der Messstrategie auf das Messergebnis in Abhängigkeit von der Formabweichung der betrachteten Werkstückoberfläche abgeschätzt werden. Durch Modellierung auf Basis von Informationen über die Bearbeitung werden die Werkstückoberflächen digitalisiert und können im Rahmen der Messunsicherheitsermittlung mit dem virtuellen Koordinatenmessgerät verwendet werden.

Keywords: Fertigungsmesstechnik, Messunsicherheit, Monte-Carlo-Simulation, Formmodellierung

Einleitung

Die Koordinatenmesstechnik dient zur dimensionellen Messung von Werkstücken vorwiegend im Rahmen der Qualitätsprüfung. Mit dem Begriff Koordinatenmesstechnik werden alle Messverfahren bezeichnet, welche das zu messende Objekt durch Koordinatenpunkte im Raum erfassen. Klassische Koordinatenmessgeräte sind kartesisch aufgebaut, daher beziehen sich die erfassten Koordinaten vorwiegend auf ein kartesisches Koordinatensystem. Bei taktilen Messungen werden die Punkte mit einem meist kugelförmigen Antastelement erfasst. Das Antastelement berührt das Werkstück und erkennt beim Erreichen einer vordefinierten Messkraft den Kontaktpunkt zwischen Werkstück und Antastelement als Messpunkt. Im Einzelpunktmodus werden die Messpunkte durch eine Antastung an einer definierten Stelle erfasst. Im Anschluss an die Punkerfassung löst das Tastelement den Kontakt mit der Werkstückoberfläche. Im Scanningmodus bleibt das Antastelement in Kontakt mit dem Werkstück, verfährt entlang einer definierten Messbahn und nimmt dabei kontinuierlich Messpunkte auf. Aus den erfassten Messpunkten wird durch Ausgleichsrechnung das assoziierte Element und damit das Ergebnis einer Messung berechnet. Dieses Ergebnis dient als Grundlage für die Bewertung der Qualität des gemessenen Objekts. Dabei wird in einem der Messwert mit dem Sollwert verglichen.

Der Messwert setzt sich aus dem wahren Wert sowie zufälligen und systematischen Messabweichungen zusammen. Zufällige Abweichungen sind im Gegensatz zu den systematischen Abweichungen nicht korrigierbar. Der Einfluss systematischer Komponenten auf das Messergebnis kann dagegen ausgeglichen werden. Dafür sind Kenntnis über den Messvorgang und seine Einflussparameter erforderlich. Aus diesen Informationen wird die Unsicherheit berechnet. Der Messwert wird damit um seine Unsicherheit erweitert.

In Forschung und Industrie spielt die Kenntnis der Messunsicherheit und ihre Größenordnung eine entscheidende Rolle Die Kenntnis über die eigenen Prozesse und deren Fähigkeit wird in Qualitätsrichtlinien gefordert. Die Messunsicherheit ist hierbei ein Indikator für deren Beurteilung. Die Kenntnis der Messunsicherheit kann neben dem Nachweis zur Eignung auch zur Verbesserung von Messprozessen genutzt werden.

Messunsicherheitsbestimmung mit dem virtuellen Koordinatenmessgerät

Die Bestimmung der Messunsicherheit durch Simulation basiert auf einer umfassenden Modellierung der Messaufgabe. Die Simulation löst das Problem der komplexen Überlagerung vieler Einflussgrößen unter Zuhilfenahme der

Wahrscheinlichkeitstheorie [2].
Mathematischer Hintergrund ist dabei die Monte-Carlo-Simulation. Die Umsetzung dieser Methodik für Koordinatenmessungen ist durch das virtuelle Koordinatenmessgerät, abgekürzt „VCMM“ (Virtual Coordinate Measuring Machine), realisiert [3]. Das VCMM befindet sich seit 2004 in Forschung und Industrie im Einsatz. Anwendung findet es hauptsächlich in Kalibrierlaboratorien. Durch die 2018 fertiggestellte Weiterentwicklung des virtuellen Koordinatenmessgeräts wurde dessen Funktionsumfang durch die Berücksichtigung von Scanning-Messungen und Tasterwechsel erweitert [4]. Die Simulation ist heute so in die Messsoftware der Hersteller integriert, dass bestehende Messprogramme im Anschluss an die Messung oder zu einem beliebigen späteren Zeitpunkt mit dem VCMM ausgewertet werden können. Die berechnete Messunsicherheit wird begleitend zum Messwert im Protokoll der Messung ausgegeben.

Die gerätebezogenen Einflussgrößen (z. B.: Restabweichungen der Führungen) werden bei der Installation des VCMM ermittelt und bei Änderungen am Gerät aktualisiert. Messaufgabenbezogene Größen (z. B.: Werkstücktemperatur) werden durch den Anwender direkt vor der Durchführung der Messaufgabe bestimmt. Zu den messaufgabenbezogenen Größen gehört auch die Oberflächenrauheit, die letztlich nur durch eine sehr hohe Anzahl von Messpunkten vollständig erfassbar ist. Der Abstand der Messpunkte ist so klein zu wählen, dass der Übergang von der Rauheit zu den Anteilen mit noch kürzeren Wellenlängen erfasst wird. Mit dieser Messstrategie wird eine ausreichende Strukturauflösung ($< \lambda s$ unter der Grenzwellenlänge nach [5]) erzielt.

Um für den praktischen Einsatz die Anzahl der Messpunkte zu begrenzen, und die Reproduzierbarkeit der Messung zu steigern, wird durch die punktgenaue Festlegung der Messstrategie als Teil der Messaufgabendefinition der Einfluss der Oberflächenform inklusive der Oberflächenwelligkeit (gesamtes Spektrum der Oberflächenformabweichungen gemäß Bild 8 in [6]) eliminiert. Dieses Vorgehen ist beim Kalibrieren von Normalen zweckmäßig und etabliert. Bei der Messung von Werkstücken, bei denen die unvollständige Erfassung der Oberfläche durch eine begrenzte Messpunktanzahl („Unterabtastung“) zur Messunsicherheit beiträgt, ist das allerdings nicht zielführend.

Beitrag der Messstrategie

Eine Werkstückoberfläche ist nur dann vollständig gemessen, wenn jede Oberflächencharakteristik bei der taktilen Messung erfasst wird. Bei einer unvollständigen Erfassung kann es vorkommen, dass Oberflächencharakteristika nicht erfasst werden und das Messergebnis somit unvollständig ist. Handelt es sich dabei um dominante Charakteristika, kann das Messergebnis dadurch verfälscht werden. Abweichungen hierbei betreffen sowohl Maß, Form und Lage des gemessenen Elements.

Durch die Analyse des Einflusses der Unterabtastung wird der Beitrag der Messstrategie (Messpunktverteilung) zur Unsicherheit abgeschätzt und im VCMM berücksichtigt [7]. Dazu werden Messungen mit verschiedenen Messstrategien miteinander verglichen. Das VCMM generiert für jedes Messmerkmal ein eigenes Formmodell. Dieses Formmodell wird entsprechend der zu erwarteten Formabweichung als sinusförmige Oberfläche modelliert und in Amplitude und Ortswellenlänge skaliert. Die Messmerkmale werden mit der vom Anwender als Standard für die Messung definierten Punktverteilung virtuell abgetastet.

Nach dieser virtuellen Messung findet die Messunsicherheitsermittlung durch das VCMM statt. Hierbei werden alle relevanten Einflüsse auf die Messung berücksichtigt. Im Anschluss an diese Standardmessung findet eine Simulation der Messung mit einer vielfachen Anzahl an Messpunkten statt. Diese Anzahl sollte so gewählt werden, dass die kürzeste Ortswellenlänge der Oberflächenform mit mindestens sieben Punkten [6] erfasst wird. Diese simulierte Messung dient zur vollständigen Erfassung der Formabweichung und wird als Referenz betrachtet. Der Unterschied der erfassten Formabweichungen bei den Messungen stellt den zusätzlichen Beitrag durch die Unterabtastung dar und fließt in die Messunsicherheit mit ein.

In Abbildung 1 werden solche simulierten Messungen mit den jeweils erfassten Formabweichungen dargestellt. Als Modell wird eine geschliffene Werkstückoberfläche (Werkzeug mit geometrisch unbestimmter Schneide) verwendet. Das erste (oberste) Bild der Abbildung zeigt die tatsächliche Oberfläche und die Berührpunkte mit 2 mm Taster-Durchmesser. Die Oberfläche wird mit insgesamt 20 000 Punkten simuliert und weist eine Formabweichung von insgesamt 7,4 µm auf. Im zweiten und dritten Bild der Abbildung wird das bisher im virtuellen Koordinatenmessgerät implementierte

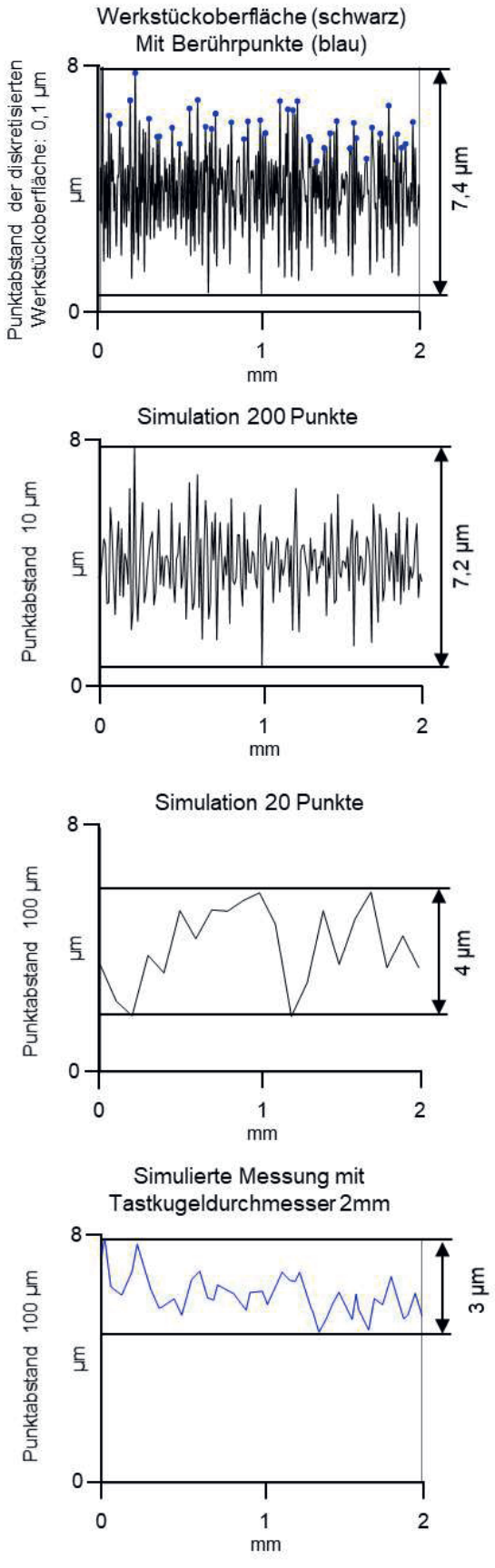


Abb. 1: Wirkung der Unterabtastung: Vergleich verschiedener Punktedichten und Herangehensweisen

Vorgehen gezeigt. Die Abtastung durch die Referenzmessung erfolgt mit einer vielfachen Punkteanzahl im Vergleich zur Standardmessung. Die sich ergebende Differenz der Formabweichung stellt den Beitrag durch Unterabtastung dar.

Im zweiten Bild der Abbildung werden Punkte für die Erfassung der Referenz mit dem Punktabstand von 10 µm aus der originalen Werkstückoberfläche extrahiert. Hierbei wird eine Formabweichung von 7,2 µm erfasst. Im dritten Bild der Abbildung werden die Punkte mit einem Punktabstand von 100 µm aus der originalen Werkstückoberfläche extrahiert. Dies stellt den vom Anwender festgelegten Messpunktabstand der beispielhaften Messaufgabe dar. Die erfasste Formabweichung dieser Standardmessung ist mit 4 µm deutlich geringer als die der Referenzmessung. Man erkennt, dass die Oberflächencharakteristik nicht mehr vollständig abgebildet wird. Der Beitrag durch Unterabtastung ergibt sich aus der Differenz der erfassten Formabweichungen von Standard und Referenz. Er beträgt hier 3,2 µm.

Im vierten Bild der Abbildung werden die Punkte nicht einfach aus der originalen Werkstückoberfläche extrahiert, sondern die mechanische Filterung des kreisförmigen Antastelements mit dem Durchmesser 2 mm wird mit einberechnet. Das erfasste Profil dieser simulierten Messung ergibt sich aus den Berührpunkten des Tastelements mit der Werkstückoberfläche. Durch die Geometrie des Antastelements werden die Täler der Werkstückoberfläche nicht mehr erfasst. Die erfasste Formabweichung beträgt nur noch 3 µm. Durch die Erfassung der Messpunkte an den Profilspitzen ist ein zudem ein Lageversatz der Messpunkte im Vergleich zum bisher implementierten Vorgehen erkennbar.

Weitere Untersuchungen werden zukünftig für Fertigungsverfahren mit geometrisch bestimmter und geometrisch unbestimmter Schneide durchgeführt. Bei den simulierten Messungen wird die Werkstückoberfläche durch das kreisförmig modellierte Tastelement mechanisch gefiltert. Bei kleinen Tastelementen findet eine detailliertere Erfassung der Werkstückoberfläche statt. Bei großen Tastelementen werden meist nur noch die Spitzen des Profils erfasst. Die erfassbare Formabweichung verringert sich somit mit der Vergrößerung des Tastelements. Die Unterschätzung der Form wird in der Gegenüberstellung der verschiedenen Messstrategien deutlich sichtbar. Findet eine ausreichende Erfassung der Werkstückoberfläche durch die Messstrategie

des Anwenders statt, sind die Unterschiede zwischen Referenz- und Standardmessstrategie zu vernachlässigen.

Formmodellierung

Durch die individuelle Anpassung des Formmodells an die Parameter des Fertigungsprozesses können aussagekräftige Angaben für diesen Beitrag zur Messunsicherheit gemacht werden. Bei der Modellierung wird zwischen Fertigungsverfahren mit geometrisch bestimmter und geometrisch unbestimmter Schneide unterschieden.

Erstere, zum Beispiel Fräsen und Drehen, werden anhand der Fertigungsparameter modelliert. Hierfür werden weitergehende Kenntnisse über die Eigenschaften der resultierenden Werkstückoberfläche aus den jeweiligen Fertigungsprozessen (z. B.: Vorschub und Schneidkantenradius [8]) genutzt. Diese Fertigungsverfahren sind im Hinblick auf die daraus entstehenden Oberflächenstrukturen bei einem idealen Prozess ausreichend bekannt.

Fertigungsverfahren mit geometrisch unbestimmter Schneide erzeugen eine zufällige Werkstückoberfläche. Beschreibungen der resultierenden Werkstückoberfläche bei perfekten Prozessparametern sind durch die Zufälligkeit des Fertigungsprozesses, resultierend aus der jeweiligen Schneidengeometrie der strukturgebenden Werkzeuge, nicht bekannt. Werkstückoberflächen mit diesen Eigenschaften werden daher durch ein Modell basierend auf sich überlagernden Sinuswellen erstellt. Die Informationen über die Parameter Amplitude, und Wellenlänge kommen aus Analysemessungen realer Bauteile mit Tastschnittgeräten mit einem Punktabstand < 0,1 µm. Gemäß dem Zufallscharakter der Oberfläche wird daraus ein Modell nicht statisch erstellt, sondern mit jedem Simulationsdurchlauf variiert. Ein einzelnes Modell besteht aus 82 verschiedenen Wellenlängen. Jeder Wellenlänge sind eine Amplitude mit oberen und unterm Grenzwert und eine zufällig generierte Phasenlage zugeordnet. Beide Modellierungsarten dienen der Simulation der aus dem Fertigungsprozess entstehenden Oberflächencharakteristik. Der Punktabstand der initial simulierten Werkstückoberflächen beträgt 0,1 µm. Dieser Abstand entspricht den Anforderungen an den Profilpunktabstand bei der Messung mit einem Tastschnittgerät [9]. Zudem ist der Abstand ausreichend gering, um die Oberfläche in Hinblick auf taktile Koordinatenmessungen mit kugelförmigen Antastelementen (Tastkugeldurchmesser > 0,2 mm) vollständig zu erfassen. Langwellige Abweichungen, welche z. B. aus Führungsabweichungen der Werkzeugmaschine stammen, werden in der Modellierung nicht berücksichtigt. Im Anschluss an die jeweilige Modellierung wird die mechanische Abtastung durch ein kugelförmiges Tastelement mittels morphologischer Filterung mit einem Kreis simuliert.

Verifikation des Modells

Die Verifikation der Modelle und der Vorgehensweise erfolgt durch den Vergleich mit Messungen an Oberflächennormalen (Oberflächenvergleichsmuster [10]). Neben Dreh- und Fräsprofilen als Vertreter der Fertigungsverfahren mit geometrisch bestimmter Schneide werden geschliffene Profile, Fertigungsverfahren mit geometrisch unbestimmter Schneide, betrachtet. Die Fertigungsparameter dieser Normale sind bekannt. Die Messung der Normale erfolgt mit verschiedengroßen, kugelförmigen Antastelementen, um die Bandbreite möglicher Messaufgaben, die mit unterschiedlichen Antastelementen ausgeführt werden, abzudecken. Die Messstrecke beträgt je nach Profil bis zu 30 mm. Die folgende Darstellung beschränkt sich auf den Schleifprozess.

Abbildung 2 zeigt ein Beispiel für den Vergleich von simulierten Messungen an einem Oberflächenmodell eines Schleifprozesses mit Messungen an einer realen Oberfläche, die ähnliche Fertigungsparameter wie das Modell aufweist. Die Profile werden dafür jeweils in einem Ausschnitt von 2 mm Länge dargestellt. Der Punktabstand der in Abbildung 3 gezeigten Beispiele beträgt jeweils 10 µm. Das gemessene Oberflächenprofil wurde mit einem Gauß-Hochpassfilter gefiltert, um langwellige Formabweichungen aus dem Fertigungsprozess zu eliminieren, und die Vergleichbarkeit zwischen beiden Profilen herzustellen. Die Messung wurde mit einem Antastelement mit dem Durchmesser 2 mm simuliert und an einem Koordinatenmessgerät durchgeführt. Aufgrund langwelliger Verformungen des gemessenen Oberflächennormals wurde das Profil mit einem Hochpass-Gaußfilter mit der Grenzwellenlänge von 2,5 mm gefiltert.

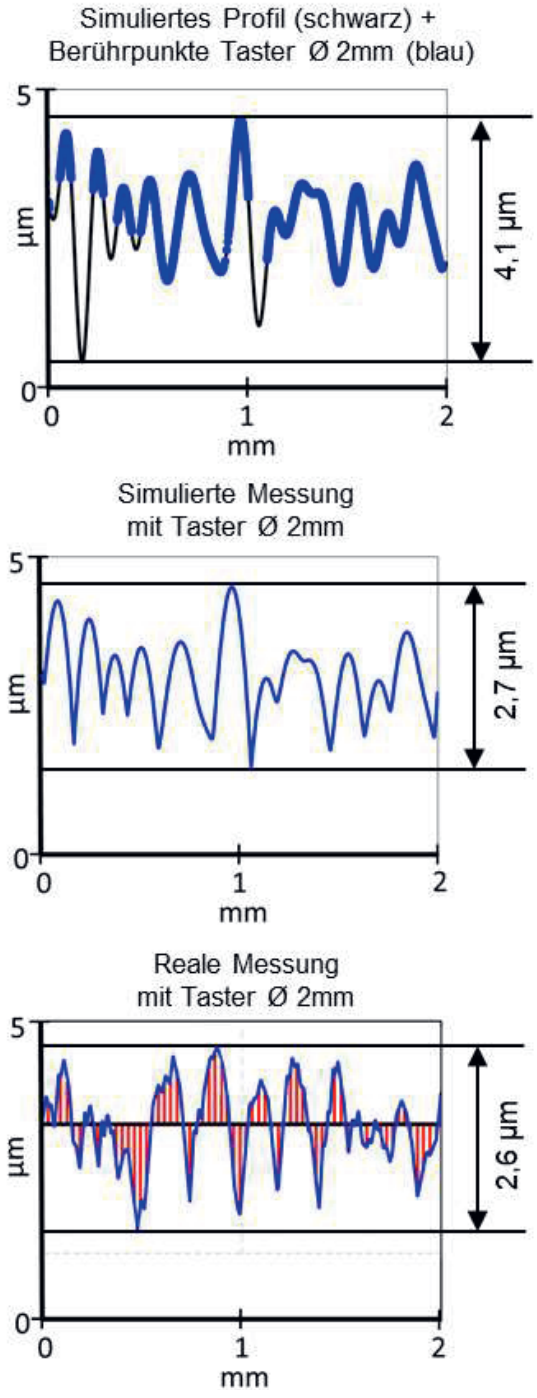


Abb. 2: Auszug einer beispielhaften Modellierung eines Schleifprofils: Simuliertes Ausgangs- und Messprofil im Vergleich zu einer real gemessener geschliffenen Werkstückoberfläche unter Berücksichtigung des mechanischen Filters durch das Tastelement (Durchmesser 2 mm)

Ein in Hinblick auf die mechanische Filterung relevanter Vergleichsparameter ist die mittlere Rillenbreite der Profilelemente (*Rsm*) und die mittlere Höhe der Profilelemente (*Rc*). Durch diese Parameter wird die Eintauchtiefe der Tastkugel in das gemessene Profil maßgeblich bestimmt. Die mittlere Rillenbreite *Rsm* beträgt 0,220 mm bei dem simulierten Schleifprofil und 0,226 mm bei dem real gemessenen Schleifprofil. Die mittlere Höhe der Profilelemente *Rc* beträgt 1,78 µm bei dem simulierten Schleifprofil und 1,791 µm bei dem real gemessenen Schleifprofil. Die Abweichungen zwischen den Profilen sind kleiner 3% für *Rsm* und kleiner 1% für *Rc*. Die Auswertungen beziehen sich jeweils auf die vollständigen Profile und nicht nur auf den in Abbildung 3 beispielhaft dargestellten Auszug.

Das gemessene Profil zeigt ein hochfrequentes Rauschen. Dieses Rauschen entsteht bei der realen Messung und wird bei der Simulation der Messung nicht erzeugt. Zukünftig ist es denkbar, auf die simulierten Modelle zusätzliche kurzwellige Störfrequenzen aufzubringen, um eine noch höhere Vergleichbarkeit mit taktilen Koordinatenmessungen herzustellen. Die Verifikation der in Simulation und Realität gemessenen Profile findet über die Parameter *Rsm* und *Rc* statt. Weitere Oberflächenparameter zur Verifikation sind Gegenstand der derzeitigen Untersuchung und können zukünftig mit hinzugezogen werden.

Ausblick

Durch die Modellierung der Werkstückoberfläche, die Simulation unterschiedlicher Messpunktverteilungen sowie der mechanischen Filterwirkung und der sich anschließenden Bewertung der Messstrategie auf Basis der Form der Profile können messaufgabenspezifisch Messstrategien hinsichtlich der Messunsicherheit in Verbindung mit dem VCMM optimiert werden. Weiterhin können die Modelle der Fertigungsverfahren zukünftig in das virtuelle Koordinatenmessgerät integriert werden. Dies ermöglicht dem Anwender die Abschätzung des Beitrags der Unterabtastung zur Messunsicherheit unter Verwendung realer Oberflächenformmodelle an seinen konkreten Messaufgaben. Die Integration der Messstrategie als Parameter zur Messunsicherheitsbetrachtung wird damit unterstützt.

Literaturnachweis

[1] Weckenmann, A. (Hrsg): *Koordinatenmesstechnik: Flexible Messstrategien für Maß, Form und Lage*; 2. Aufl. München : Hanser Verlag, 2012.

[2] JGCM 101:2008: *Evaluation of measurement data – Supplement 1 to the "Guide to the expression of uncertainty in measurement" – Propagation of distributions using a Monte Carlo method,* JGCM (Joint Committee for Guides in Metrology), 2008. (Internet: http://www.bipm.org/en/publications/guides/gum.html).

[3] Kistner, T.; Gabbia, A.; Imkamp, D: *Messunsicherheit durch Simulation: Stand der Technik und neue Entwicklungen für die Koordinatenmesstechnik in der industriellen Produktion,* 18. GMA/ITG-Fachtagung Sensoren und Messsysteme 2016, Nürnberg, 2016.

[4] Franke, M.; Kistner, T.; Hausotte, T.; Heißelmann, D.; Schwehn, C.; Wendt, K.: *Bestimmung der Messunsicherheit für Koordinatenmesssysteme*. tm - Technisches Messen, 84(5), S. 325-335.

[5] DIN EN ISO 4287:2010-07-00 *Geometrische Produktspezifikationen (GPS) – Oberflächenbeschaffenheit: Tastschnittverfahren – Benennungen, Definitionen und Kenngrößen der Oberflächenbeschaffenheit (ISO 4287:1997 + Cor 1:1998 + Cor 2:2005 + Amd 1:2009);* Deutsche Fassung EN ISO 4287:1998 + AC;2008 + A1:2009.

[6] VDI/VDE-Richtlinie 2631 Blatt 3, Formprüfung, Eigenschaften und Auswahl von Filtern, September 2018.

[7] Heißelmann, D.; Franke, M.; Kistner, T.; Schwehn, C.: *Das virtuelle Koordinatenmessgerät VCMM,* Koordinatenmesstechnik 2018 – Technologien und Anwendung, DIN-Tagung, Braunschweig, 2018.

[8] Klocke, F.; König, W.: *Fertigungsverfahren 1: Drehen, Fräsen, Bohren*, 4. Aufl. Berlin-Heidelberg : Springer, 2007. – ISBN 978-3-540-23458-6.

[9] DIN EN ISO 3274:1997 *Geometrische Produktspezifikationen (GPS) - Oberflächenbeschaffenheit: Tastschnittverfahren - Nenneigenschaften von Tastschnittgeräten (ISO 3274:1996);* Deutsche Fassung EN ISO 3274:1997.

[10] ISO 2632-1:1985 *Roughness comparison specimens – Part 1: Turned, Ground, Bored, Milled, Shaped and Planed,* September 1985.

[11] DIN EN ISO 3274:1997 *Geometrische Produktspezifikationen (GPS) - Oberflächenbeschaffenheit: Tastschnittverfahren – Oberflächen mit plateauartigen funktionsrelevanten Eigenschaften – Teil 2: Beschreibung der Höhe mittels linearer Darstellung der Materialanteilkurve (ISO 1356-2:1996);* Deutsche Fassung EN ISO 13565-2:1997.

DOI 10.5162/sensoren2019/6.4.4

Investigations on the sine fitting algorithm in the Planck-Balance

Shan Lin[1], Christian Rothleitner[1], Norbert Rogge[2]
[1]Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
[2]Institute of Process Measurement and Sensor Technology, Technische Universität Ilmenau, 98684 Ilmenau, Germany
shan.lin@ptb.de, christian.rothleitner@ptb.de, norbert.rogge@tu-ilmenau.de

Abstract

The Planck-Balance is a table-top version of a Kibble balance. In contrast to many other Kibble balances, the Planck-Balance generates an ac (sinusoidal) rather than a dc signal in the dynamic mode. A linear sine fitting algorithm is applied to estimate the amplitudes of the induced voltage and the coil motion, which determine the force factor Bl of the voice coil of the electromagnetic force compensated balance. Different sine fitting algorithms are compared in terms of harmonic distortion, additive Gaussian noise and non-coherent sampling. The biases and uncertainties associated with the estimated Bl are also compared by numerical simulations.

Keywords: Planck-Balance, three-parameter sine fit, multiharmonic sine fit, Monte Carlo simulation.

Introduction

After the redefinition of the kilogram, the Kibble balance is one possible approach to calibrate mass standards in terms of a fixed value of the Planck constant [1]. The Kibble balance has two measuring modes: static mode and dynamic mode. In the static mode, the gravitational force on a mass m is balanced by an electromagnetic force that is generated by a voice coil, i.e. $mg = BlI$, where g is the local gravitational acceleration, I the electrical current, B the magnetic flux density and l the length of the coil-wire in the B-field. The product Bl is commonly called the force factor. In the dynamic mode, Bl is obtained by moving the coil with a velocity v through the magnetic field, while measuring the induced voltage u across the coil ends synchronously, i.e., $u = Blv$. The Bl obtained in the dynamic mode is then used in the static mode to evaluate the mass m as a function of the coil current I.

The Planck-Balance (PB) is a table-top sized Kibble balance and is currently under development by the Physikalisch-Technische Bundesanstalt (PTB) and the Technische Universität Ilmenau (TUIL) [2]. In the dynamic mode of the PB, the coil is moved sinusoidally – in contrast to most other Kibble balance experiments, where the velocity is kept constant – through the B-field. If the coil motion is assumed to be perfectly sinusoidal, the velocity is simply the time derivative of the sinusoid. Therefore, the amplitude V of the velocity becomes $V = \omega S = 2\pi f_{\mathrm{sig}} S$, where S is the amplitude of the coil motion, ω the angular frequency and f_{sig} the oscillation frequency. Then, Bl is calculated by dividing the amplitude of the induced voltage U by the amplitude of the velocity V, i.e., $Bl = U/(2\pi f_{\mathrm{sig}} S)$. However, this is strictly valid only if the coil motion and the induced voltage are perfectly sinusoidal. Therefore, the accuracy of Bl is influenced by the estimation of amplitudes, when perturbations, such as noise or distortions caused by non-linearities, are present in the signal.

In this paper, a linear sine fitting algorithm is applied with perturbations like, e.g., higher order harmonics, additive Gaussian white noise and non-coherent sampling. The resulting bias is investigated in numerical simulations. The standard uncertainty associated with the force factor Bl is evaluated by the GUM uncertainty framework and Monte Carlo simulation, respectively.

Methodology

In dynamic mode of the PB, the coil sinusoidally oscillates through the B-field generated by the magnet system, and the coil position is measured by a laser interferometer. The movement of the coil induces an ac voltage across the coil ends, which is digitized by means of a multimeter. In the real measurement data of the induced voltage and the coil motion, higher order harmonics are present in the

signals after using Fourier analysis. Therefore, the signal of the induced voltage $u(t)$ with time t can be described by using a single-component sine wave with amplitude U_1 superimposed by higher order harmonics:

$$u(t) = U_0 + \sum_{n=1}^{N} U_n \sin(n\omega t + \varphi_{un}), \quad (1)$$

where U_n and φ_{un} are the amplitude and the initial phase of the nth-order harmonic of the induce voltage, respectively. U_0 denotes a dc offset, and N is the number of higher order harmonics, with $N \in \mathbb{N}$.

The coil motion $s(t)$ can also be described by a sum of N harmonic terms:

$$s(t) = S_0 + \sum_{n=1}^{N} S_n \sin(n\omega t + \varphi_{sn}), \quad (2)$$

where S_0 is a dc offset, S_n the amplitude of the nth-order harmonic, φ_{sn} the phase of nth-order harmonic.

The force factor Bl is calculated by using the signal frequency f_{sig} and the fundamental amplitudes of $u(t)$ and $s(t)$ as

$$Bl = \frac{U_1}{2\pi f_{sig} S_1}. \quad (3)$$

In the PB system, the signal frequency f_{sig} can be accurately achieved. Only the fundamental amplitudes U_1 and S_1 need to be estimated by the linear sine fitting algorithm. The following algorithms are applied to estimate the amplitudes:

- Three-parameter sine fit;
- Multiharmonic sine fit;
- Improved three-parameter sine fit.

Due to harmonic distortion in the measurement data, the accuracy of Bl is partly dependent on the fitting algorithm.

Three-parameter sine fit

An ideal single-component sine wave can be described with an amplitude Y and an initial phase φ as:

$$y(t) = Y_0 + Y \sin(\omega t + \varphi) \quad (4)$$

where Y_0 denotes a dc offset. Equivalently, the signal in Eq. (4) can be interpreted as a superposition of two shifted sine waves:

$$y(t) = Y_0 + A \sin(\omega t) + B \cos(\omega t) \quad (5)$$

where A and B are the amplitudes of in-phase and in-quadrature components, respectively. The signal frequency f_{sig} is assumed to be known in Eq. (5), and thus only three parameters Y_0, A and B are required to be estimated. The estimated parameter vector $\mathbf{p}$ can be represented as $\mathbf{p} = [A, B, Y_0]^{\mathrm{T}}$.

When a set of M samples $\{y_i\}_{i=1}^{M}$ from a sine wave is sampled at the time instants $\{t_i\}_{i=1}^{M}$, a linear least squares method can be used to determine the parameter vector $\mathbf{p}$ by minimizing the sum of the squares of the following errors:

$$\min_{A,B,Y_0} \sum_{i=1}^{M} (y_i - Y_0 - A \sin(\omega t_i) - B \cos(\omega t_i))^2. \quad (6)$$

The estimated parameters of the sine wave can be calculated in a matrix form:

$$\mathbf{p} = (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{y} \quad (7)$$

with $\mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_M]^{\mathrm{T}}$ and

$$\mathbf{D} = \begin{bmatrix} \sin(\omega t_1) & \cos(\omega t_1) & 1 \\ \sin(\omega t_2) & \cos(\omega t_2) & 1 \\ \vdots & \vdots & \vdots \\ \sin(\omega t_M) & \cos(\omega t_M) & 1 \end{bmatrix}. \quad (8)$$

The amplitude Y of $y(t)$ can be calculated as:

$$Y = \sqrt{A^2 + B^2}. \quad (9)$$

Multiharmonic sine fit

In the real measurement data of the induced voltage and the coil motion, higher order harmonics are present. These higher order harmonics may be caused, e.g., by the current source or the nonlinear B-field. However, the three-parameter sine fitting algorithm is not robust against the harmonic distortion with non-coherent sampling [3]. In order to improve the accuracy of the estimated amplitude, the multiharmonic sine fit has been proposed in [4]. The multiharmonic sine fit is the extension of the three- and four-parameter sine fit with the known and unknown signal frequency, respectively. In this paper, the linear optimization procedure based on the three-parameter sine fit is adopted.

When the signal contains higher order harmonic components, the signal is modeled by

$$y(t) = Y_0 + \sum_{n=1}^{N} Y_n \sin(n\omega t + \varphi_n), \quad (10)$$

where Y_n and φ_n are the amplitude and initial phase of the nth-order harmonic, respectively. The fitting model for the linear least squares method can be written as follows:

$$y(t) = Y_0 + \sum_{n=1}^{N} [A_n \sin(n\omega t) + B_n \cos(n\omega t)], \quad (11)$$

where A_n and B_n are the nth-order harmonic amplitudes of in-phase and in-quadrature components, respectively. By using the least squares method, the sum of the squares of the errors is minimized. The estimated parameters can be calculated to extend Eq. (7) as follows:

$$\mathbf{p}_{\mathrm{m}} = (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}\mathbf{D}^{\mathrm{T}}\mathbf{y} \quad (12)$$

where $\mathbf{p}_{\mathrm{m}}$ is the parameter vector, $\mathbf{p}_{\mathrm{m}} = [A_1, B_1, \ldots, A_N, B_N, Y_0]^{\mathrm{T}}$. Here,

$$\mathbf{D} = [\mathbf{S}_1 \quad \mathbf{C}_1 \quad \cdots \quad \mathbf{S}_N \quad \mathbf{C}_N \quad \mathbf{1}], \quad (13)$$

with

$\mathbf{S}_n = [\sin(n\omega t_1) \quad \sin(n\omega t_2) \quad \cdots \quad \sin(n\omega t_M)]^{\mathrm{T}}$,

$\mathbf{C}_n = [\cos(n\omega t_1) \quad \cos(n\omega t_2) \quad \cdots \quad \cos(n\omega t_M)]^{\mathrm{T}}$,

$n = 1, \ldots, N$.

The amplitude Y_n of the nth-order harmonic of $y(t)$ can be determined as:

$$Y_n = \sqrt{A_n^2 + B_n^2}. \tag{14}$$

Improved three-parameter sine fit

Due to harmonic distortion, the conventional three-parameter sine fitting algorithm will provide a bias of the estimated amplitude for the non-coherent sampling. Multiharmonic sine fit is useful to estimate the amplitude of each order harmonic. The bias of amplitude is relatively low, when the number of higher order harmonics included in the fitting model approaches to that included in the measurement data. However, it is not easy to select the number for the fitting model.

An alternative method is to use a coherent sampling instead of the non-coherent sampling by removing some points at the beginning and/or at the end of the data, which makes the remaining points cover an integer number of periods of the sine wave.

Firstly, the first zero-crossing sampling point is detected, which satisfies the following equation: $y(q_1 - 1) \cdot y(q_1) < 0$. Here $y(q_1) = y(t_{q_1})$ and q_1 denotes the first index of the detected sample. The same routine is utilized to detect the other zero-crossing sampling points $\{y(q_i)\}_{i=1}^{M_q}$, where M_q is the total number of the detected samples. Then, some ending points are removed if they are not inside the following range:

$$y_i \notin \begin{cases} \left[y(q_1), y\left(q_{M_q} - 1\right)\right], & \text{if } M_q \text{ is odd} \\ \left[y(q_1), y\left(q_{M_q - 1} - 1\right)\right], & \text{otherwise} \end{cases}. \tag{15}$$

Finally, the remaining points are extracted and used to estimate the amplitude by the three-parameter sine fit.

In practice, however, it is difficult to extract sampling points, which cover an exactly integer number of periods. Moreover, reduction of some samples can increase the bias and the standard deviation of the estimated amplitude [5]. If the number of samples is relatively large, the influence of reducing samples may be neglected in comparison to that due to harmonic distortion with non-coherent sampling. The simulation results will be shown in the following section.

Uncertainty evaluation

All the sine fitting algorithms used in this paper are linear least squares models, and thus the standard deviations of the estimated parameters can be evaluated by using the covariance matrix $\mathbf{V}$. The expression is given in [6] as follows:

$$\mathbf{V} = \sigma_{\mathrm{n}}^2 (\mathbf{D}^{\mathrm{T}}\mathbf{D})^{-1}. \tag{16}$$

Here, $\mathbf{D}$ is given by Eqs. (8) and (13) for the three-parameter sine fit and the multiharmonic sine fit, respectively. σ_{n} denotes the noise level of measurement. The result of $\mathbf{V}$ can be presented in the following format:

$$\mathbf{V} = \begin{bmatrix} u(A_1, A_1) & u(A_1, B_1) & \cdots & u(A_1, Y_0) \\ u(A_1, B_1) & u(B_1, B_1) & \cdots & u(B_1, Y_0) \\ \vdots & \vdots & \ddots & \vdots \\ u(A_1, Y_0) & u(B_1, Y_0) & \cdots & u(Y_0, Y_0) \end{bmatrix}, \tag{17}$$

where $u(p_i, p_j)$ is the covariance associated with estimates p_i and p_j.

The fundamental amplitude Y of the sine wave is calculated by using A_1 and B_1, and thus only the covariances associated with A_1 and B_1 are used in the following steps. The submatrix $\mathbf{V}_{\mathrm{x}}$ extracted from the covariance matrix $\mathbf{V}$ is

$$\mathbf{V}_{\mathrm{x}} = \begin{bmatrix} u(A_1, A_1) & u(A_1, B_1) \\ u(A_1, B_1) & u(B_1, B_1) \end{bmatrix}. \tag{18}$$

According to the GUM uncertainty framework (GUF) [7], the standard uncertainty σ_{y} associated with the amplitude Y is given by:

$$\sigma_{\mathrm{y}}^2 = \mathbf{C}_{\mathrm{x}} \mathbf{V}_{\mathrm{x}} \mathbf{C}_{\mathrm{x}}^{\mathrm{T}}, \tag{19}$$

where $\mathbf{C}_{\mathrm{x}}$ is the sensitivity matrix and evaluated as follows:

$$\mathbf{C}_{\mathbf{x}} = \begin{bmatrix} \frac{\partial Y}{\partial A_1} & \frac{\partial Y}{\partial B_1} \end{bmatrix} = \begin{bmatrix} \frac{A_1}{\sqrt{A_1^2 + B_1^2}} & \frac{B_1}{\sqrt{A_1^2 + B_1^2}} \end{bmatrix}. \tag{20}$$

According to Eq. (19), the standard uncertainties σ_{u} and σ_{s} associated with the amplitudes U and S of the induced voltage and the coil motion can be deduced, respectively.

As mentioned before, the force factor Bl is calculated by using Eq. (3). The standard uncertainty σ_{Bl} associated with Bl is calculated by:

$$\sigma_{\mathrm{Bl}}^2 = \mathbf{C}_{\mathrm{Bl}} \mathbf{V}_{\mathrm{Bl}} \mathbf{C}_{\mathrm{Bl}}^{\mathrm{T}}, \tag{21}$$

where $\mathbf{C}_{\mathrm{Bl}}$ and $\mathbf{V}_{\mathrm{Bl}}$ are the sensitivity matrix and covariance matrix for σ_{Bl} estimation, respectively. Here,

$$\mathbf{C}_{\mathrm{Bl}} = \begin{bmatrix} \frac{\partial Bl}{\partial U_1} & \frac{\partial Bl}{\partial S_1} & \frac{\partial Bl}{\partial f_{\mathrm{sig}}} \end{bmatrix},$$

$$\mathbf{V}_{\mathrm{Bl}} = \begin{bmatrix} u(U_1, U_1) & u(U_1, S_1) & u(U_1, f_{\mathrm{sig}}) \\ u(U_1, S_1) & u(S_1, S_1) & u(S_1, f_{\mathrm{sig}}) \\ u(U_1, f_{\mathrm{sig}}) & u(S_1, f_{\mathrm{sig}}) & u(f_{\mathrm{sig}}, f_{\mathrm{sig}}) \end{bmatrix},$$

where $u(U_1, U_1) = \sigma_{\mathrm{u}}^2$ and $u(S_1, S_1) = \sigma_{\mathrm{s}}^2$.

As the signal frequency can be accurately measured by a frequency counter (< 10^{-8}), the term in Eq. (21) representing the signal frequency contribution to the uncertainty is neglected in this case. After a series of deduction, the relative standard deviation associated with Bl is:

$$\sigma_{\mathrm{Bl_{rel}}} = \begin{cases} \sqrt{\frac{\sigma_u^2}{U_1^2} + \frac{\sigma_S^2}{S_1^2}}, & \text{if } u(U_1, S_1) = 0 \\ \sqrt{\frac{\sigma_u^2}{U_1^2} + \frac{\sigma_S^2}{S_1^2} - \frac{2u(U_1,S_1)}{U_1 S_1}}, & \text{otherwise} \end{cases} . \quad (22)$$

Numerical simulation of the relative bias of the amplitude

In order to investigate the influence of harmonic distortion on the Bl determination, the induced voltage and the coil motion are required to be simulated.

Firstly, the induced voltage is simulated by using a fundamental sine wave with the amplitude U_{ref} = 0.3051 V and superimposed by higher order harmonics up to 5th-order. The amplitudes of 2nd-order up to 5th-order harmonic are 0.45 mV, 62.9 µV, 2.40 µV and 1.92 µV, respectively. The total harmonic distortion is 0.15 %. These parameters are obtained according to real measurement data [8]. In order to simulate the non-coherent sampling, the signal frequency f_{sig} is changed from 3.82 Hz to 4.02 Hz with the same number of samples M = 10^4, i.e., the sampling frequency f_{s} = 1 kHz and the sampling time T = 10 s. For each f_{sig}, a data set of induced voltage $\{u_i\}_{i=1}^{10^4}$ is generated and taken as the nominal points. Three algorithms are applied to estimate the fundamental amplitude $\widehat{U}_1$. For the multiharmonic sine fit, 1st-order up to Nth-order harmonics (N = 1, 2, 3 and 4) are included in the model, respectively. When only 1st-order harmonic is included in the multiharmonic sine fit (i.e., N = 1), the algorithm is the same as the three-parameter sine fit. For error comparison of amplitude, the value of U_{ref} is taken as the reference value. The relative bias of the amplitude ε_{u} is calculated, as $\varepsilon_{\mathrm{u}} = \left|\widehat{U}_1 - U_{\mathrm{ref}}\right|/U_{\mathrm{ref}}$. The results are shown in figure 1.

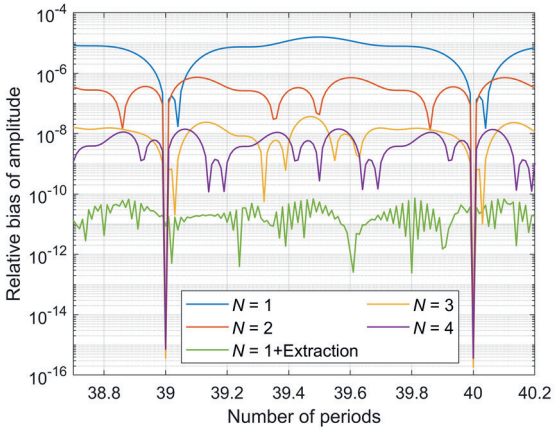


Figure 1: Relative bias of the amplitude as a function of the number of periods.

When the number of periods is an integer, i.e. coherent sampling, the amplitudes obtained from the three algorithms are comparable. Due to the numerical accuracy of the simulation data, the relative bias of the amplitude is in the order of 10^{-15}. Therefore, the three algorithms are robust against harmonic distortion with coherent sampling. However, it can be seen that the relative bias of the amplitude provided by the three-parameter sine fit is in the order of 10^{-6} with the non-coherent sampling. When the multiharmonic sine fit is applied, the higher the number of included harmonics is, the lower the bias of the estimated amplitude becomes. If higher order harmonics up to 5th-order (N = 5) are included in the fitting model, there is no bias. The bias can greatly be reduced for the non-coherent sampling if the number of samples is reduced to obtain an integer number of periods (green line in figure 1).

Moreover, the coil motion is also simulated by using a fundamental sine wave with the amplitude S_{ref} = 39.92 µm and superimposed by higher order harmonics up to 5th-order. The amplitudes of 2nd-order up to 5th-order harmonic are 32.29 nm, 2.93 nm, 0.14 nm and 0.13 nm, respectively. The total harmonic distortion is 0.08 %. Similar to the induced voltage, the data set $\{s_i\}_{i=1}^{10^4}$ of the coil motion is generated for each f_{sig}. Three algorithms are used to estimate the fundamental amplitude $\hat{S}_1$ of the coil motion. Then, Bl is determined from the estimated amplitudes, i.e., $\widehat{Bl} = \widehat{U}_1/\left(2\pi f_{\mathrm{sig}}\hat{S}_1\right)$. For error comparison, the fundamental amplitudes U_{ref} and S_{ref} in the simulated signals are used for the definition of the reference value of Bl, i.e., $Bl_{\mathrm{ref}} = U_{\mathrm{ref}}/\left(2\pi f_{\mathrm{sig}} S_{\mathrm{ref}}\right)$. Finally, the relative deviation of the calculated $\widehat{Bl}$ with respect to the reference Bl_{ref} is calculated, as $\varepsilon_{\mathrm{Bl}} = \left|\widehat{Bl} - Bl_{\mathrm{ref}}\right|/Bl_{\mathrm{ref}}$. The simulation results are presented in figure 2.

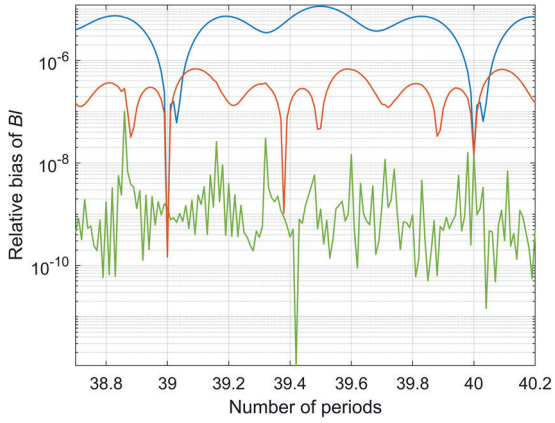


Figure 2: Relative bias of Bl as a function of the number of periods. The amplitudes are estimated by three-parameter sine fit (blue line), multiharmonic sine fit (N = 2, red line) and improved three-parameter sine fit (green line), respectively.

Similar results can be seen in figure 2: harmonic distortion with coherent sampling has almost no influence on Bl determination. However, when non-coherent sampling occurs, the three-parameter sine fit show the highest bias of Bl. Compared to the three-parameter sine fit, the multiharmonic sine fit (N = 2) can reduce the bias of Bl. If an integer number of periods is extracted from the non-integer number of periods by deleting some ending points, the bias can be greatly reduced.

Numerical simulation of standard deviation of amplitude

In this section, the GUM uncertainty framework (GUF) and the Monte Carlo method (MCM) are used to estimate the standard deviation of Bl. For the signal simulation, the nominal points $\{u_i\}_{i=1}^{10^4}$ of the induced voltage and $\{s_i\}_{i=1}^{10^4}$ of the coil motion are the same as those in the former section.

Firstly, the GUF is applied to estimate the relative standard deviation of Bl according to Eq. (22). According to one set of real measurement data, the standard deviations of noise are σ_{nu} = 0.1036 mV for the induced voltage and σ_{ns} = 1.8768 nm for the coil motion. The induced voltage and coil motion are measured by independent instruments, and here it is assumed that they are uncorrelated, i.e. $u(U_1, S_1) = 0$. The estimated relative standard deviation of Bl is shown in figure 3.

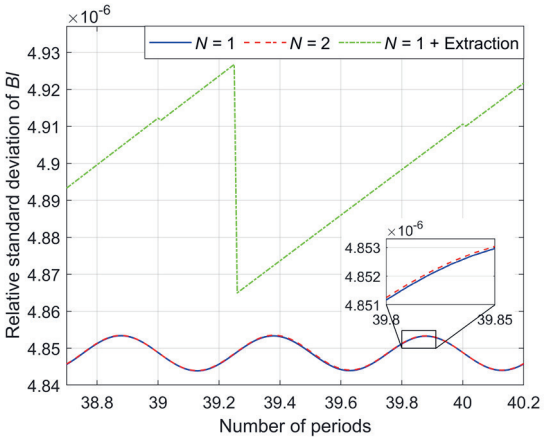


Figure 3: Relative standard deviation of Bl as a function of the number of periods. The amplitudes are estimated by three-parameter sine fit (blue line), multiharmonic sine fit (N = 2, red line) and improved three-parameter sine fit (green line), respectively.

In figure 3, all the relative standard deviations are in the same order of magnitude (i.e., 10^{-6}). The improved three-parameter sine fit achieves the highest standard deviation. That is because the reduction of some samples increases the standard deviation of estimated amplitude. Therefore, the relative standard deviation of Bl is increased in the order of 10^{-8}. If type A evaluation of standard uncertainty is applied, the relative uncertainty (k = 1) is increased in the order of 10^{-10}, which is much smaller than the accuracy aimed for the PB (about 10^{-8}). For the other two algorithms, the blue and red lines are almost indistinguishable, and the difference between them is in the order of 10^{-10}. Moreover, when the non-coherent sampling occurs, the change of relative standard deviation is in the order of 10^{-9}, which is also much smaller than the aimed accuracy of the PB. Therefore, the influence of non-coherent sampling on the standard deviation of Bl can be neglected.

In practice, the amplitudes U_1 and S_1 are correlated according to the data analyzed in a real measurement, and the correlation coefficient is about -0.06. Compared to the uncorrelated amplitudes U_1 and S_1, the acquired relative standard deviation is in the order of 10^{-8} higher. This difference is very low and can be neglected.

Then, the MCM is applied to evaluate the relative standard deviation of Bl. It is assumed that the amplitudes U_1 and S_1 are uncorrelated. For each f_{sig}, Gaussian white noise with the standard deviation σ_{nu} = 0.1036 mV and σ_{ns} = 1.8768 nm are superimposed on the nominal points of $\{u_i\}_i^{10^4}$ and $\{s_i\}_i^{10^4}$, respectively. Three algorithms are implemented to estimate the

amplitudes $\hat{U}_1$ and $\hat{S}_1$, and then $\widehat{Bl}$ is determined. The process is repeated 10^5 times. The bias and the standard deviation of $\widehat{Bl}$ are calculated and presented in figure 4.

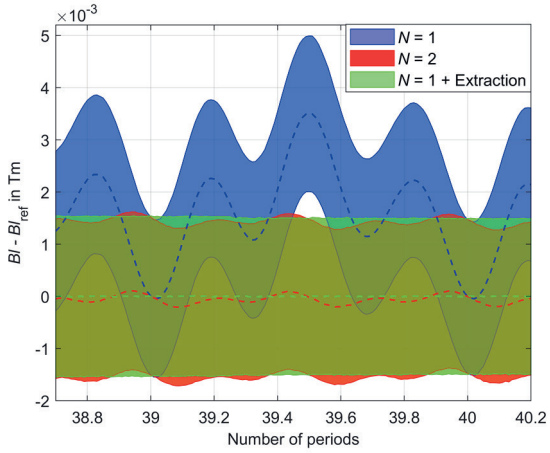


Figure 4: Bias and standard deviation of Bl estimated by the MCM. The dotted line represents the mean value of bias of Bl and the solid lines represent the confidence interval (k = 1). The amplitudes are estimated by three-parameter sine fit (blue), multiharmonic sine fit (N = 2, red) and improved three-parameter sine fit (green).

As shown in figure 4, the three-parameter sine fit provides a bias of the Bl with non-coherent sampling, which is identical to the blue line in figure 2. Compared to the bias provided by the three-parameter sine fit, the influence of non-coherent sampling on the standard deviation can be neglected, even though the improved three-parameter sine fit acquires the highest standard deviation in figure 3. The results of the other two algorithms are almost identical. The multiharmonic sine fit with N = 2 is comparable to the improved three-parameter sine fit in its stability.

Moreover, the relative standard deviation $\sigma_{\mathrm{Bl_{rel}}}$ determined by the GUF is compared with that provided by the MCM. The $\sigma_{\mathrm{Bl_{rel}}}$ is evaluated by Eq. (22), and it is assumed that the amplitudes U_1 and S_1 are uncorrelated. When using the MCM, the relative standard deviation $\sigma_{\widehat{\mathrm{Bl}}_{\mathrm{rel}}}$ is calculated by dividing the standard deviation $\sigma_{\widehat{\mathrm{Bl}}}$ associated with $\widehat{Bl}$ by the reference value Bl_{ref}. As shown in figure 3, the $\sigma_{\mathrm{Bl_{rel}}}$ provided by the three-parameter and multiharmonic sine fit are indistinguishable, and thus only the latter one is presented in figure 5.

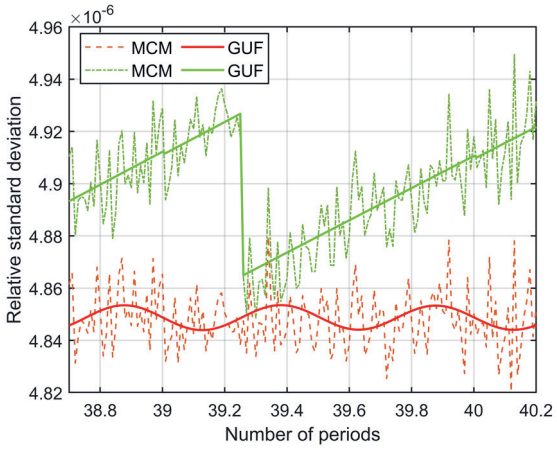


Figure 5: Relative standard deviation of Bl evaluated by the GUF and MCM, respectively. The amplitudes are estimated by multiharmonic sine fit (N = 2, red line) and improved three-parameter sine fit (green line).

In figure 5, the differences between $\sigma_{\mathrm{Bl_{rel}}}$ from the GUF and $\sigma_{\widehat{\mathrm{Bl}}_{\mathrm{rel}}}$ from the MCM are in the order of 10^{-8} for each algorithm. The difference is two orders of magnitude smaller than the values of $\sigma_{\mathrm{Bl_{rel}}}$ and $\sigma_{\widehat{\mathrm{Bl}}_{\mathrm{rel}}}$. Therefore, the results from GUF and MCM are comparable.

Conclusion

In order to obtain a high accuracy of the force factor Bl, the relative bias and standard deviation of Bl are investigated in numerical simulations. When the sampled sine wave includes higher order harmonics with non-coherent sampling, the extraction of integer periods of the sine wave or multiharmonic sine fit can reduce the bias of the estimated amplitude, and further improve the accuracy of Bl. As for the standard uncertainty of Bl, the GUF and MCM are applied and the results of all the algorithms are comparable. In the future, other influences on the amplitude estimation, e.g. time jitter, quantization error, and numerical approximation, will also be investigated.

Acknowledgement

The research of this project is funded via the program “Validierung des technologischen und gesellschaftlichen Innovationspotenzials – VIP+”, a program of the German Federal Ministry of Education and Research (BMBF), and is managed by VDI/VDE Innovation + Technik GmbH. The authors are grateful to members of Planck-Balance (PTB+TUIL) team for useful discussions.

References

[1] I. A. Robinson, S. Schlamminger, *Metrologia*, 53(5), A46-A74 (2016); doi: 10.1088/0026-1394/53/5/A46

[2] C. Rothleitner, J. Schleichert, N. Rogge, L. Günther, S. Vasilyan, F. Hilbrunner, D. Knopf, T. Fröhlich, F. Härtig, *Measurement Science and Technology,* 29(7), 074003(2018); doi: 10.1088/1361-6501/aabc9e

[3] J. Deyst, T. Sounders, O. Solomon, *IEEE Transactions on Instrumentation and Measurement*, 44(3), 637-642 (1995); doi: 10.1109/19.387298

[4] P. M. Ramos, M. F. d. Silva, R. C. Martins, A. M. C. Serra, *IEEE Transactions on Instrumentation and Measurement*, 55(2), 646-651 (2006); doi: 10.1109/TIM.2006.864260

[5] F. C. Alegria, *Measurement,* 42(5), 748-756 (2009); doi: 10.1016/j.measurement.2008.12.006

[6] A. B. Forbes, in *Data Modeling for Metrology and Testing in Measurement Science*, Birkhäuser Boston, 147-176 (2008); doi: 10.1007/978-0-8176-4804-6_5

[7] BIPM, Evaluation of measurement data — Supplement 2 to the “Guide to the expression of uncertainty in measurement” — Extension to any number of output quantities, (2011).

[8] C. Rothleitner, J. Schleichert, S. Vasilyan, N. Rogge, L. Günther, F. Hilbrunner, I. Rahneberg, D. Knopf, T. Fröhlich, F. Härtig, *Proc. of 2018 Conference on Precision Electromagnetic Measurements,* Paris, France, July 8-13 (2018); doi: 10.1109/CPEM.2018.8500904

DOI 10.5162/sensoren2019/6.4.5

Model based correction of motion deviations in the Planck-Balance

Huitian Bai [1], Norbert Rogge [2], Christian Rothleitner[3], Thomas Fröhlich[2]
[1] School of Precision Instruments and Opto-Electronics Engineering, Tianjin University, 92 Weijin Road, Nankai District, Tianjin, P.R.China
[2] Institute of Process Measurement and Sensor Technology, Technische Universität Ilmenau, 98684 Ilmenau, Germany
[3] Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

Tel.-Nr.: +86 13102030957, +493677691745, +495315921131
E-Mail: baihuitian@tju.edu.cn, norbert.rogge@tu-ilmenau.de, christian.rothleitner@ptb.de

Abstract

The Planck-Balance, a device planned to be used as a direct realization of the definition of the kilogram, utilizes the working principle of a Kibble balance [1]. Therefore, the motion of the coil under test has to be measured precisely. In the Planck-Balance, the coil is guided by the load carrier of a commercially available EMFC load cell. In order to reduce the measurement uncertainty of the velocity measurement, the tilt motions of this load carrier needs to be well known at the specified frequencies. The contribution presents different approaches to model the tilt of the load carrier during its sinusoidal excitation and the correction of tilt induced errors of the measurement.

Keywords: EMFC balance, mechanical simulation, tilt measurement, error correction, Kibble balance.

Introduction

In the Planck-Balance, a commercially available electromagnetic force compensated load cell (EMFC) is used to guide the moving coil of a voice coil actuator under test that is attached to the load carrier of the balance (see Figure 1). While the internal coil of the balance acts as a motor, which drives the system, the parallel lever system of the EMFC load cell provides a linear trajectory of the coil. This first mode of the experiment is utilized to determine the force factor Bl of the actuator under test (i.e. external voice coil), which is used to determine the mass of a weight in a second mode. The displacement of the coil and the induced voltage are measured as a function of time in order to calculate the force factor Bl as a ratio of the voltage and coil velocity. In the second mode the gravitational force produced by the weight is compensated by the force that is produced by the actuator when an electric current flows through the external coil.
Additional to the linear motion of the coil, the system also exhibits several tilt motions that cause errors in the velocity measurement if the displacement measurement system is not placed collinearly to the coil centre. These tilt motions should preferably be reduced as good as possible and the remaining tilt should be measured. In the experimental set up, three interferometer beams are used to measure the displacement of the load carrier and calculate two tilt angles from differences of the interferometer signals and the known distance between the laser spots. The induced voltage is measured with a digital voltmeter and the control algorithm is implemented on a commercially available digital signal processing system (DSP).

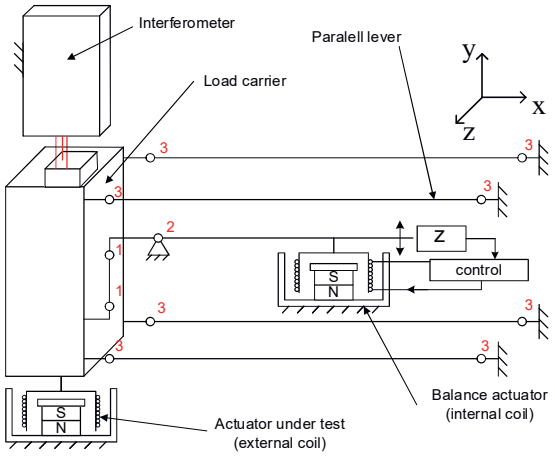


Figure 1: Schematic of balance and measurement set up.

Preceding investigations [2] indicated that the load carrier exhibits tilt oscillations when it is moved sinusoidal along the y-axis. The amplitude of these oscillations increases linear with the displacement amplitude and in a low frequency range of up to 50 Hz also with the excitation frequency.

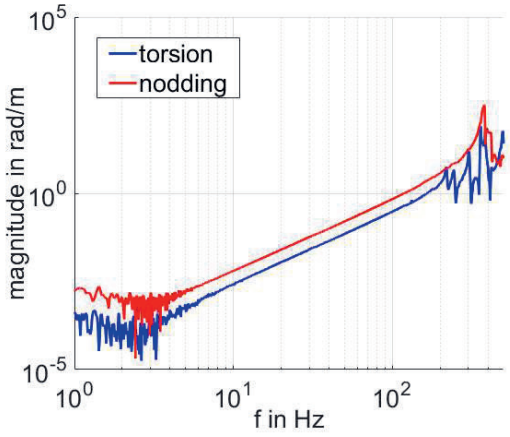


Figure 2: Torsion (blue) and nodding (red) tilt of the unloaded balance.

Furthermore, the mass distribution on the load carrier influences the amplitude of the oscillations themselves as well as the frequency range in which the increase is linear versus frequency. The frequency response of the system is obtained by exciting the system at the controller input with a chirp signal and applying the Fourier transform on the tilt signal and the displacement. Since the position sensor of this control system yields nonlinear characteristics, the applied chirp with frequencies of up to 30 Hz also produces higher harmonics, which can be used to identify also the system response for higher frequencies. The displacement acts as system input of the transfer system to make experiments with different excitation amplitudes comparable. An example of the occurring tilt amplitudes is shown in Figure 2, where “torsion” refers to the tilt about the x-axis, while the term “nodding” is used for rotations about the z-axis.

Model of the balance

In order to analyze the causes of the tilt oscillations, a model of the utilized balance is set up to investigate the sensitivity of the balances’ behavior versus the selected parameters. The model (shown in Figure 3) is mainly made up of four sections: mechanical system, position sensor, voice coil and U/I converter.

1. Mechanical subsystem:

The mechanical system is based on the CAD model and mechanical drawings of the device.

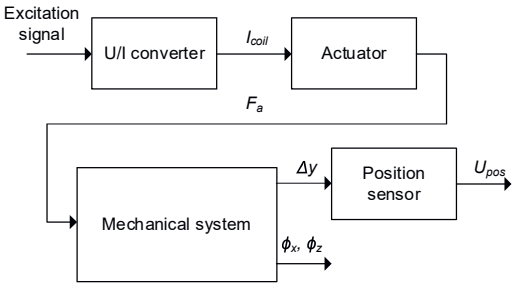


Figure 3: Main components of the balance model.

As proposed in [3] the mechanism is dissected into rigid bodies that are connected to each other by flexure hinges, which are simplified in Figure 1 as joints. The model contains two holders for putting on a mass piece, which are also available in the experimental set up of the tilt investigations and placed in a distance of 5 cm from the symmetry plane of the system. In the model, the position where the mass is located can be adjusted by setting the coordinates of the holders. The weight of the mass is defined by a constant and the unit of the constant is gram.

The joints in the model are assumed to possess different degrees of freedom, depending on their geometry and function within the system. The joints marked with number one in Figure 1 are able to rotate about the x-axis and z-axis, while the element between these two joints allows a deformation about the y-axis by adding another joint that is not shown in the schematic. The joints representing the lever bearing (number 2 in Figure 1) permit a rotation about the x-axis and z-axis, whereas the joints of the parallel lever system allow a rotational deformation about all axes and a translation along the x-axis.

The spring constants of the flexure hinges were calculated according to [4], where t, b and R represent geometrical parameters, α_z, Δy and Δx the deformation and M_z, F_y and F_x torque and applied forces, respectively, while E represents Young’s modulus:

Bending:
$$\frac{\alpha_z}{M_z} = \frac{9\pi R^{\frac{1}{2}}}{2Ebt^{\frac{5}{2}}} \quad (1)$$

Compression:
$$\frac{\Delta x}{F_x} = \frac{9\pi}{2Eb}\left(\frac{R}{t}\right)^{\frac{5}{2}} \quad (2)$$

Shear:
$$\frac{\Delta z}{F_z} = \frac{1}{Eb}\left[\pi\left(\frac{R}{t}\right)^{\frac{1}{2}} - 2.57\right]. \quad (3)$$

The additional rotational spring constants that cannot be calculated with the equations (1) to (3) were obtained by adjustment of their values

in order to fit to the measured behavior of the system.

2. Actuator subsystem:

The actuator subsystem of the model contains the electrical properties of the coil like its resistance and includes an interface between the mechanical and the electrical domain of the model. This interface is represented by the Bl of the internal coil of the balance that is the proportional factor of the relation between force F and the current I through the coil, as well as between the induced voltage U_{ind} and the coil velocity v:

$$Bl = \frac{U_{\mathrm{ind}}}{v} \tag{4}$$

$$Bl = \frac{F}{I} \tag{5}$$

The Bl of the internal coil was determined as 50 N/A in preceding experiments. Another parameter that plays an important role during the motion of the coil is the translational damping coefficient. It is determined as 9 N/(m/s) by fitting the model behavior to the experimental data. The mechanical part of the actuator system also contains the mechanical limits of the lever displacement, which is also present in the real system.

3. Position sensor:

The position sensor of the balance consists of a small aperture at the end of the transmission lever. The aperture partially blocks the light that is emitted by an infrared LED and that is pointed towards two photodiodes. Depending on the position of the aperture, the photodiodes produce photo currents that are equal, when both diodes receive the same amount of light. The difference of the photo currents is measured and converted into a voltage U_{pos}, which acts as the input signal for the lever position controller of the balance.
In order to control the displacement of the load carrier in the unit of meter instead of the position voltage, the relationship between them is considered. The load carrier displacement can be measured by the interferometer within the Planck-Balance set up. When the transmission lever is moving in a small angle, it is reasonable to assume from the geometry that the end of the transmission lever is displaced five times as much as the load carrier displacement and the directions of them are opposite. The function describing the relationship between the displacement of the end of the transmission lever relative to the balanced position and the position of load carrier is easily to get by fitting a polynomial to both signals.

4. U/I converter:

The U/I converter circuit board compares the voltage drop across a 500 Ω resistor to its voltage input, where the output current flows through the resistor and the actuator coil. The model of the converter also contains the saturation of the output current that depends on the limited voltage supply and the resistance of the load.

Model validation

To check the quality of the model, the measurements are carried out with a linear frequency chirp signal from 1 Hz to 30 Hz as input of the position controller and compared with simulations of the same input.

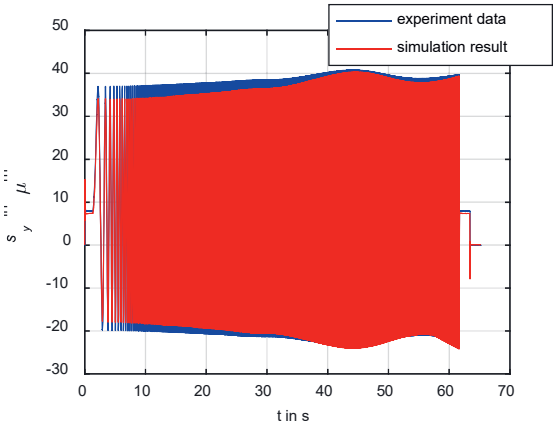


Figure 4: Displacement of the unloaded load carrier.

As depicted in Figure 4, the simulation of the displacement is in a quite good agreement with the measurement data, when the load carrier is unloaded. However, the torsion tilt oscillation does not appear in this case of simulation since the mechanical model is assumed perfectly symmetrical, which is in contrast to the real system.
The nodding oscillations do occur in the simulation with linear increase depending on the frequency (see Figure 5), but the amplitudes are smaller than the ones observed in the experiment.

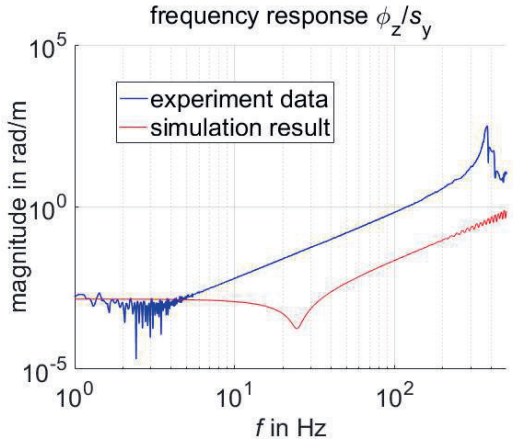


Figure 5: Nodding tilt amplitude relative to load carrier displacement, unloaded balance.

1. Mass distribution:

One parameter that is expected to be not symmetrical in the real system is the mass distribution on the load carrier. To quantify the influence of this parameter within the model, the model, as well as the balance, were loaded with a mass of 100 g in a distance of 5 cm relative to the symmetry plane. In this case, the simulation shows a linear increase of tilt amplitude not only for the nodding angle (Figure 6), but also for the torsion angle (Figure 7). In case of nodding, the simulated amplitudes are also significantly smaller. In addition to this, resonances occur around 50 Hz and 100 Hz, which are not present in the model due to the simplifications that were made to keep the focus on the linear tilt phenomena. For example, all levers in the model were assumed to be rigid bodies and do not possess elastic characteristics.

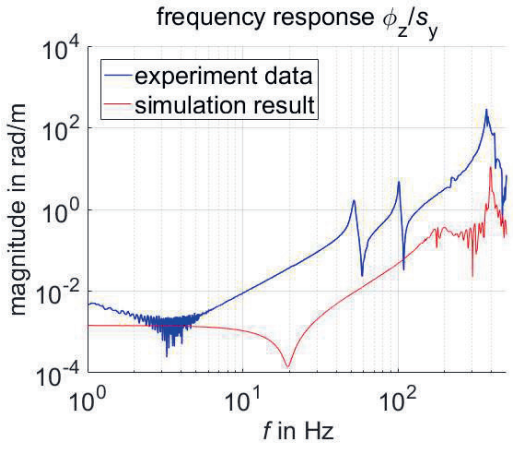


Figure 6: Nodding tilt amplitude relative to the load carrier displacement. The balance is loaded with 100 g on the right side.

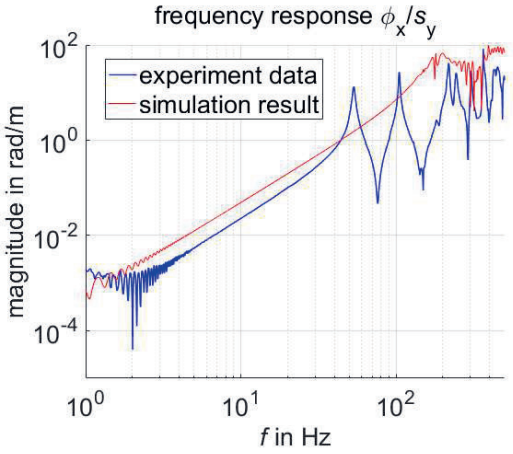


Figure 7: Torsion tilt amplitude relative to the load carrier displacement. The balance is loaded with 100 g on the right side.

In case of the torsion tilt, the frequency range with linear increase is more consistent with the measurement, but also only below the resonance effects.
In further simulations the mass of the unsymmetrically placed weight and its position in the xz-plane is varied to identify the linearity of the influence of these parameters on the behavior of the balance. Increasing the distance along the z-axis and the mass value, the amplitude linearly increases in the frequency range of 5 Hz to 30 Hz, while shifting the mass towards the actuator of the balance along x-axis decreases the amplitude. Furthermore, Figure 8 shows that the effect is not linear for a frequency of 30 Hz.

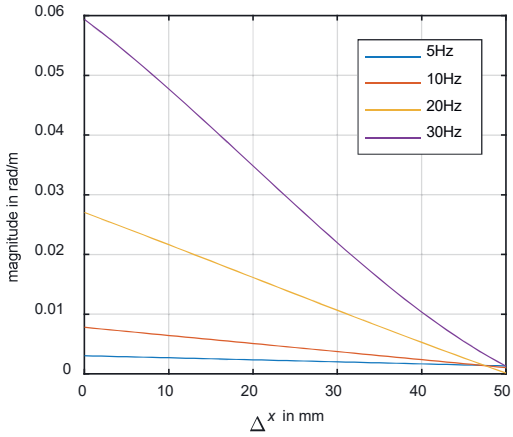


Figure 8: Relation between torsion tilt amplitude and weight placement along x-axis.

2. Joint stiffness:

The torsion tilt effects are not only caused by the unsymmetrical mass distribution, since they

also appear when the balance is unloaded and therefore the mass distribution can be assumed far more symmetrical than the state shown in Figure 7. Another possible cause for this type of tilt is the assumption, that not all joints possess the same spring constants. This influence is investigated by decreasing the rotational spring constants about the x- and z-axis and the translational spring constant along z-axis of one joint by a factor 10. As depicted in Figure 9, only the variation of the translational spring constant causes magnitudes close to the ones observed in the experiments.
The linearity of the influence of the translational spring constant along z-axis is checked in further simulations in which the value is varied in the range from 90 % to 99 % of the original value.

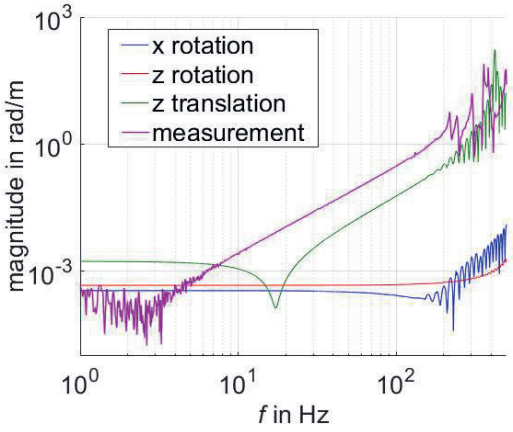


Figure 9: Torsion tilt amplitude response for decreased spring constant of one joint.

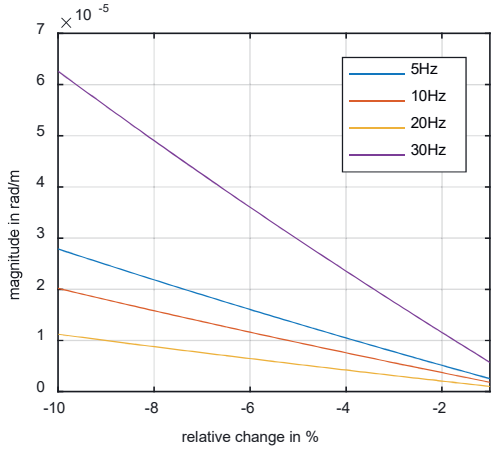


Figure 10: Torsion tilt amplitude for different translational spring constants of one joint.

The simulation results depicted in Figure 10 show a linear dependency for these small changes of the spring constant. The amplitude itself produced by this parameter a far too small to explain the results from the experiment and since the effect seems to be linear, it is not expected that other variations which are plausible according to the manufacturing tolerances can cause these effects.

Mounting deformation

An import factor that causes tilt oscillations is that the balance is mounted to the ground with three screws placed in the middle of the balance quiet close to each other. Since this fixation is expected to possess a finite stiffness, the mirror of the three beam interferometer was also placed onto the frame of the balance mechanism to examine the tilt of the mechanical reference frame to which the parallel lever system is connected to.
In this mode of the experiment, the measured tilt cannot be directly related to the load carrier displacement, since the interferometer only points to the frame of the balance. In case of a tilt that is caused by the chirp signal excitation with constant amplitude, a chirp signal is also expected for the tilt signal with an envelope corresponding to the transfer function of the tilt relative to the excitation amplitude. In order to obtain this transfer function, a chirp signal fit was applied to the tilt signals.

$$x(t) = A(t) \cdot \sin\left[2\pi \cdot \left(f_0 + \frac{k}{2} \cdot t\right) \cdot t + \varphi\right] \quad (6)$$

Equation (6) consists of the envelope function $A(t)$ multiplied with the chirp function, whose parameters are known since the signal is generated in the signal processing system. The envelope function is assumed as a second order polynomial and its parameters are estimated such that the chirp equation (6) fits to the measured chirp response within the tilt signal.
The transfer function of the tilt is then determined by dividing the envelope by the known excitation amplitude.
The results in Figure 11 show increasing tilt amplitudes for both types of tilt in the frequency range of up to 30 Hz. With the beforehand described method, it is not possible to obtain the frame tilt for higher frequencies, since it is not recommended to increase the frequency of the chirp due to the bandwidth of the utilized controller. In this range these amplitudes are smaller than the ones observed on the load carrier, but they represent a relevant portion of the load carrier tilt. This indicates that an improvement of the mounting design has the potential to reduce the tilt oscillations.

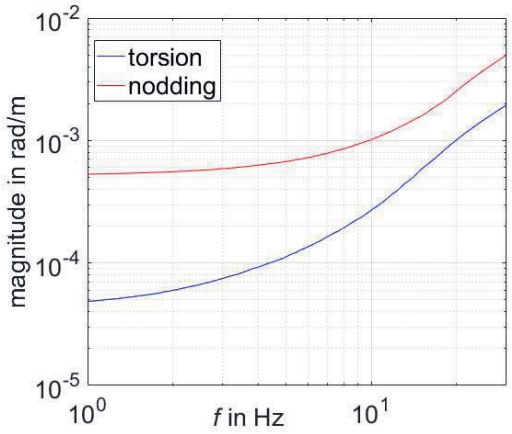


Figure 11: Amplitude transfer functions of the tilts measured on frame of the balance.

Bl determination

The preceding investigations can be utilized to improve further versions of the Planck-Balance set up. Nonetheless, also in the present set up the knowledge of the tilt motions is useful to correct the errors that are caused by them.
In the Planck-Balance set up, the motion that is performed to generate the induced voltage is sinusoidal in order to utilize the relatively small motion range of the load carrier of a commercially available EMFC balance. Assuming the sinusoidal nature of the displacement s and the induced voltage u, the Bl can be calculated from the displacement amplitude $\bar{S}$, the amplitude of the induced voltage $\bar{U}$ and the excitation frequency f_{sig}.

$$Bl = \frac{\bar{U}}{2\pi f_{\mathrm{sig}} \bar{S}} \tag{4}$$

The amplitudes are determined by fitting the parameters of a sinusoidal function with known frequency to the measured signals. The same fit can be used to obtain the amplitudes of the tilt signals that are derived from the signals of the three beam interferometer.
With the knowledge of the distances between the laser spots and the center of the actuator coil under test, the frequency dependency of the Bl determination can be reduced as depicted in Figure 12.

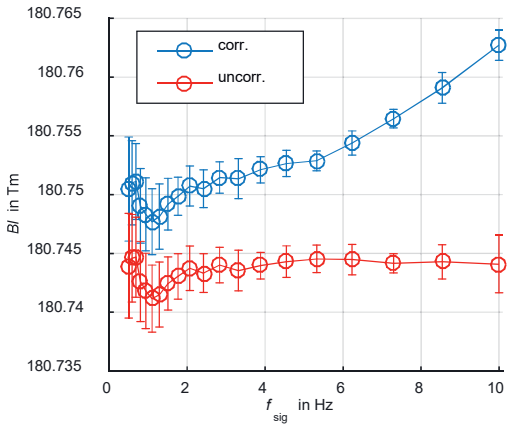


Figure 12: Relative change of the force factor Bl at different excitation frequencies.

Conclusion

This work presents investigations on two possible causes of the rotational motions that are observed in the velocity mode experiments of the Planck-Balance. This allows developing new approaches to decrease the amplitude of these motions. One option is to ensure a symmetrical mass distribution on the load carrier and shifting the center of gravity of the load carrier towards the middle of the balance. However, these measures should be done considering to keep the mass inertia small. Another reasonable method is to improve the mounting of the balance within its metrological frame. Nonetheless experiments show that a correction of the deviations caused by the tilt oscillations is feasible.

References

[1] C. Rothleitner, J. Schleichert, N. Rogge, L. Günther, S. Vasilyan, F. Hilbrunner, D. Knopf, T. Fröhlich and F. Härtig, "The Planck-Balance—using a fixed value of the Planck constant to calibrate E1/E2-weights," Measurement Science and Technology, vol. 29, no. 7, p. 074003, 2018.
[2] N. Rogge, J. Schleichert, C. Rothleitner, “Investigations on motion deviations of an EMFC balance”, Ilmenau Scientific Colloquium. Technische Universität Ilmenau; 59 (Ilmenau): 2017.09.11-15
[3] H. Weis, F. Hilbrunner, T. Fröhlich and G. Jäger, Parametric mechatronic model of a load cell with electromagnetic force compensation, IMEKO TC3 & TC5 & TC22 Int. Conf., pp 29–32, 2010.
[4] J. Paros, L. Weisbord. How to design flexure hinge. Mach Des 1965; 37:151–6.

Postersession

DOI 10.5162/sensoren2019/P1.1

Performanceanalyse von zwei optimierten Messverfahren in der kapazitiven Füllstandmessung

Jens Matheis, Andreas Uhl, Dr. Thomas Föhn, Furkan Dayi
EBE Elektro-Bau-Elemente GmbH, Sielmingerstr. 63, 70771 Leinfelden-Echterdingen

Zusammenfassung

Die kapazitive Füllstandmessung ist ein etabliertes Verfahren zur Erfassung von Füllständen von Flüssigkeiten sowie Schüttgütern. Eine Erweiterung der Messmethodik mittels eines breitbandigen Rauschsignals wird in diesen Artikel einer Performance Analyse unterzogen und dem klassischen Charge-Transfer-Messverfahren gegenübergestellt. Das Anwendungsszenario ist ein außen angebrachter kapazitiver Sensor an einer Gerätewand, hinter derer das eigentliche Mediengefäß wechselbar positioniert ist. Die Präsenzerkennung des Mediengefäßes sowie dessen Füllstands ist Ziel des Versuchsaufbaus auf Basis diskreter Bauelemente. Im Anschluss erfolgte eine Modellierung der Sensormessstrecke mit einem RLC-Ersatzschaltbild im direkten Vergleich der gemessenen Werte. Hier zeigte sich, dass die Leitfähigkeit σ_M des Mediums einen wesentlichen Einfluss bereits bei sehr dünnen Filmen generiert. Um eine sichere Präsenzerkennung des Mediengefäßes trotz Filmbildung zu gewährleisten, wurden elektrostatische Feldsimulationen durchgeführt, um die Kapazitätsänderung ΔC mit Permittivitäten von $\varepsilon_{Gefäß} << \varepsilon_{Medium}$ feststellen zu können. Es konnte gezeigt werden, dass das klassische Messverfahren auf Basis der Charge-Transfer-Technik für wässrige, leitfähige Medien die Anforderungen erfüllt. Bei viskosen, zähflüssigen Medien wie z.B. Öl oder Milch mit Neigung zur starken adhäsiven Filmbildung ist dagegen nur eine zuverlässigere Erfassung der Füllstände und der Mediengefäßpräsenz mit dem vorgestellten, rauschbasierten Verfahren möglich.

Keywords: kapazitive Füllstandmessung, kapazitive Füllstandschalter, adhäsive Filmbildung

Einleitung

Die Erfassung von Füllständen von Flüssigkeiten und Schüttgüter in der Industrie sind wohl verstanden und in den Märkten etabliert, z.B. als Schwellwertschalter. Zunehmend erhöht sich die Relevanz intelligenter Sensoren zur Füllstandmessung auch in dem Marktsegment der Medizintechnik sowie in den Produkten aus der Consumer-Electronics-Industrie.

Letztere Branche, welche sehr kostengerieben arbeitet, stellt erhöhte Anforderungen an Lösungen ohne den Einsatz von teuren Spezialbauelemente. Des Weiteren wird zunehmend gefordert, dass Sensoren für die Füllstanderkennung von wässrigen Lösungen auch die Präsenz von entnehmbaren Mediengefäßen erfassen und detektieren können mittels des gleichen Messprinzips.

Ein kapazitiv messender Sensor für wässrige, leitfähige Medien kann unsichtbar hinter einer oder mehreren dielektrischen Wänden verbaut werden (siehe Abbildung 1). Gegenüber konventionellen Reed-Schaltern bietet der kapazitive Sensor den Vorteil einer einfacheren Integration in das Gesamtsystem, ohne z.B. die Notwendigkeit einen Schwimmer zu verbauen, wodurch ein direkter Medienkontakt vermieden wird und Ausfällen durch Verkalkung und Verschmutzung und daraus folgendem Festsitzen der Schwimmer ausgeschlossen werden.

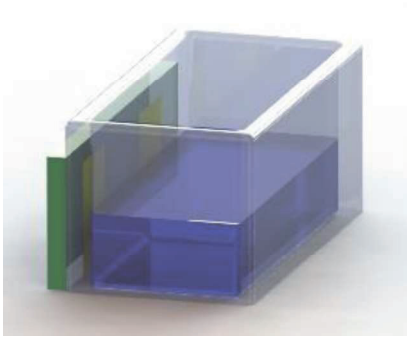


Abb. 1: Typische Aufgabenstellung einer kapazitiven Füllstandmessung inkl. Präsenzerkennung des Mediengefäßes.

Eine weitere Verbesserung ist die Möglichkeit zur kontinuierlichen Erfassung des tatsächlichen Füllstandes anstatt eines reinen Schwellwertschalters.

Dabei stellen insbesondere elektrisch leitfähige Flüssigkeiten besondere Herausforderungen an die kapazitive Messtechnik, da bereits dünne Flüssigkeitsfilme an den Gefäßwänden die Messungen stark beeinflussen können. Sind die Medien zudem noch adhesiv mit hoher Viskosität, so ergeben sich weitere Herausforderungen zur korrekten Detektion des Füllstandes. In dieser Arbeit betrachten wir zwei unterschiedliche Technologien, welche eine Mediengefäßerkennung erlauben und die Filmbildung der Medien bei der Füllstands-messung berücksichtigen.

Technologisch sind zwei Messmethoden für die kapazitive Füllstandmessung weit verbreitet. Im als Charge-Transfer-Messverfahren bekannten Verfahren wird die Messkapazität, welche stets durch eine geeignete Anordnung von Elektroden zur Anregung des den relevanten Erfassungsbereich abdeckenden elektrischen Feldes besteht, zyklisch Umgeladen [1],[2]. Die dabei erfasste Ladungsmenge Q wird dabei meist auf einer Referenzkapazität integriert und nach einer angemessenen Anzahl von Umladungen erfasst. Durch die Anzahl der Umladungen sowie durch die Größe der Referenzkapazität können die Messparameter des Sensors eingestellt werden. Des Weiteren spielt die Wahl einer geeigneten Umladezeit eine entscheidende Rolle, da die verschiedenen, zu erfassenden Szeanrien mit unterschiedlichen Frequnzgängen aufwarten. Diese Art der Messtechnik wird im Folgenden das erste System für den Performance Vergleich sein.

Ein weiteres, weit verbreitetes Verfahren ist die Einbindung der zu messenden Kapazität in einen resonanten LC-Schwingkreis, in welchem eine Verschiebung der Resonanz-frequenz, je nach Kapazitäts-änderung, getriggert durch z.B. eine Füllstandänderung, stattfindet. Es wird die Verschiebung der Schwingfrequenz messtechnisch erfasst [2]. Dieses Verfahren wird im Folgenden nicht näher betrachtet.

Ein weiterer Ansatz ist durch das innerhlab diesen Papiers dargelegten breitbandigen Messverfahren, welches im Folgenden näher betrachtetwird, gegeben. Gegenüber den bisherigen Messverfahren erfolgt eine Messung mit einem spektral breiten Messsignal. Dies steht im Gegensatz zu den schmalbandigen Messverfahren der zuvor genannten Messmethoden. Dies erlaubt eine bessere Performance des Messsystemes für adhäsive, filmbildende hoch viskose Medien.

Breitbandiges Rauschmessverfahren

Der Aufbau von kapazitiven Sensoren, welche auf dem Charge-Transfer-Verfahren basieren, ist bekannt und ausreichend in der Literatur dokumentiert [1]. Im Folgenden erfolgt die Vorstellung des zweiten kapazitiven Mess-verfahren, welches im Weiteren für die viskosen Medien zum Einsatz kommen wird.

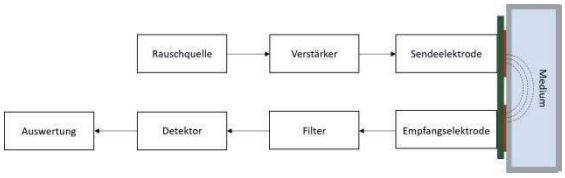


Abb. 2: schematischer Aufbau der breit-bandigen Rauschmessung

Der Sensoraufbau erfolgt mittels zwei PCB Elektroden, wobei eine als Sende- und die zweite als Empfangselektrode verwendet wird. Als Anregung wird eine Rauschquelle mit einem breitbandigen Spektrum eingesetzt, welche einen Frequenzbereich im VHF-Band abdeckt. Das erzeugte Rausch-signal wird moduliert, verstärkt und mit einem Anpassnetzwerk auf die Elektrode einge-koppelt. Das Signal an der Empfangselektrode wird nach Filterung auf einen Detektor gegeben und die Leistung gemessen. Im Anschluss erfolgt die Auswertung der Transmission der Messstrecke anhand der Empfangsleistung, wobei durch die eingesetzte Modulation eine hohe Störtoleranz und Genauigkeit erreicht wird.

Gegenüber der im vorherigen Abschnitt erläuterten Messverfahren erfolgt hier eine Anregung in einem breiten Frequenzspektrum anstatt einer Anregungsfrequenz von wenigen MHz bei dem Charge-Transfer-Verfahren oder einigen 10 MHz bei dem resonanten LC-Messverfahren, wodurch eine optimierte Signalantwort seitens des Mediensystems zu erwarten ist.

Präsenzerkennung des Mediengefäßes

Für eine Präsenzerkennung von Medien-gefäßen, wie z.B. bei einem Wassertank, müssen kleine Kapazitätsänderungen detektiert werden. Typische Wandstärken sind z.B. d_1 = 2 - 3 mm für die Gehäusewandung sowie d_2 = 2 - 3 mm für die Mediengefäßwand. Mögliche Materialien sind nicht leitfähige Kunststoffe wie beispielsweise ABS, PP oder PMMA, welche ein $\varepsilon_r \approx 1{,}8 - 2{,}4$ aufweisen.
Durch die sehr niedrige Permittivität erfolgt nur eine sehr geringe Kapazitätsänderung ΔC im

Gegensatz zu dem eigentlichen, zu messenden Medium mit $\varepsilon_{Gefäß} << \varepsilon_{Medium}$.
Um Aussagen über die zu erwartenden Kapazitätsänderungen geben zu können, wurden elektrostatische Feldsimulatonen durchgeführt.

In Abbildung 3 ist eine Feldlinienbild für eine beispielhafte Elektrodenanrodnung speziell zur Erfassung der Präsenz des Mediengefäßes gezeigt. Das gezeigte Szenario beinhaltet zunächst den Sensor und die Montagewand an der der Sensor befestigt wird. Diese Wandung ist in den üblichen Szenarien teil des Gehäuses des jeweiligen Geräts – bspw. einer Kaffeemaschine. Die Feldlinien verlaufen symmetrisch von den anregenden Elektroden hin zur Messelektrode und durchdirngen im wesentlichen die Montagewand, reichen jedoch kaum darüber hinaus.
Fügt man nun das Mediengefäß in die Simulation mit ein (Abbildung 4), so erkennt man eine geringfügige Verformung der Feldlinien einhergehend mit einer geringen messbaren Kapazitätsänderung durch das Mediengefäß.

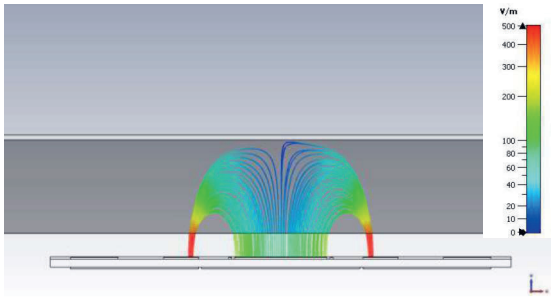


Abb. 3: Feldlinienbild des elektrischen Feldes einer Elektrodenanordnung ohne Mediengefäß, Schnitt durch Sensor-PCB (weiß, unten) und Montage-wandung (dunkelgrau, mitte) mit Luftbereichen (weiß) und Hintergrund-material (Luft, hellgrau, oben)

Betrachtet man die simulierten Kapazitäten bei vorhandenem sowie abwesendem Gefäß in Abbildung 3 und 4, so beträgt die Kapazitäts-änderung nur $\Delta C \approx 10$ fF. Diese sehr kleine Änderung erfordert weitere Maßnahmen zur Stabilisierung der Messtechnik hinsichtlich Temperatur- und Feuchtigkeitseinflüsse sowie mechanischen Toleranzen der Gesamt-anordnung. Eine Verbesserung der Stabilität der Präsenzerkennung des Gefäßes kann durch geeignete, weitere Elektroden und deren gemeinsame Auswertung erreicht werden.

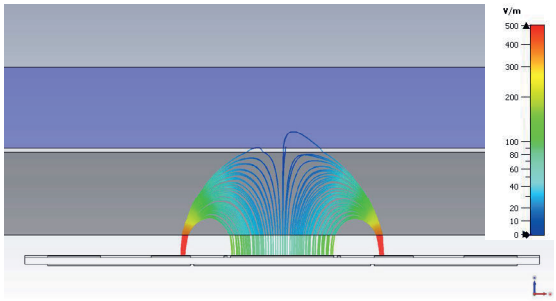


Abb. 4: Feldlinienbild des elektrischen Feldes einer Elektrodenanordnung ohne Mediengefäß, Schnitt durch Sensor-PCB (weiß, unten), Montagewandung (dunkelgrau, mitte), Gefäßwand (blau, oben) mit Luftbereichen (weiß) und Hintergrundmaterial (Luft, hellgrau, oben)

Die Kompensation der Temperatur-abhängigkeit kann weitestgehend elektronisch erfolgen. So ist die Änderung der gemessenen Kapazität im Bereich von +20°C bis +70°C kleiner als 1 fF. Der Einfluss der Feuchtigkeit ist dagegen um eine Zehnerpotenz größer als der Einfluss der Temperatur. Deswegen erfolgt eine aktive Korrektur der Kapazitätsmesswerte in der Software, wenn die Umgebungs-bedingung der Luftfeuchtigkeit im Bereich von 20–85 % rHd sich befindet.

Modellbildung der Applikationen

Für die Modellierung der Füllstandmessung wird der vereinfachte schematische Aufbau in Abbildung 5 mit einer Gefäßwand und zwei planaren Elektroden betrachtet.

Als Medium soll hier Wasser im Allgemeinen wie z.B. Leitungswasser betrachtet werden. Wasser weist eine hohe relative Dielektrizitätskonstante von $\varepsilon_r \approx 80$ auf. Aufgrund der darin vorhandenen Ionen besitzt Leitungswasser ebenfalls eine elektrische Leitfähigkeit, die wesentlich von den darin gelösten Salzen und den weiteren Verun-reinigungen bestimmt wird. Abhängig vom Messverfahren beeinflussen kapazitive und resistive Effekte die gemessenen Kapazitäts-werte auf unterschiedliche Weise.

In dem in Abbildung 5 dargestellten Ersatz-schaltbild modellieren die Kapazitäten C_0 die Einkopplung des Signals auf die lokalen Knoten im Sinne von elektrischen Knoten durch die Gefäßwand. Die variable Kapazität C_1 stellt den kapazitiven Einfluss der Signalstrecke durch das Medium Wasser dar.

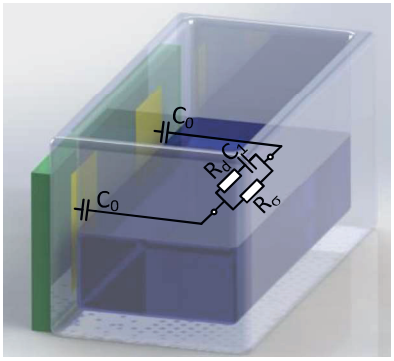


Abb. 5: Elektrotechnisches Ersatzschaltbild für die Modellierung der Applikation.

Dielektrische Verluste und die Leitfähigkeit des Wassers werden jeweils über die Widerstände R_d und R_σ modelliert.

Bei leitfähigen, ionenreichen Flüssigkeiten ist bereits bei einem sehr dünnen Flüssigkeitsfilm auf der Tankwand der Widerstand R_σ in Abhängigkeit der verwendten Messfrequenz vergleichsweise klein. In diesem Fall kann mit einer einfachen Messung nicht mehr unterschieden werden ob der Erfassungsbereich des Sensors großflächig mit einem Wandfilm überzogen ist oder der Füllstand tatsächlich ein hohes Level erreicht hat, da die Größe C_1, welche zu einem erheblichen Teil den Informationsgehalt über den tatsächlichen Füllstand trägt, nicht mehr gemessen werden kann. Um dieses Verhalten zu untersuchen werden Messungen an einem Referenzaufbau durchgeführt. Dabei werden leitfähige Filme unterschiedlicher Dicke durch in isotonische Kochsalzlöung (9g NaCl pro 1l Lösung) getränkte Papierstreifen nachgestellt.

Messung mit dem Charge-Transfer Verfahren

Die Messung erfolgt zunächst mithilfe eines Sensors der nach dem Charge-Transfer-Verfahren mit einstellbarer Umladezeit arbeitet. Eine 3 mm dicke PMMA-Platte bildet die Tankwand zwischen Sensor und Papierstreifen nach.

Durch Messungen mit unterschiedlichen Umladezeiten kann der Einfluss des Widerstand R_σ auf die Kapazitätsmessung anhand der RC-Zeitkonstante des Gesamtsystems untersucht werden. Hierbei ist es wichtig zu verstehen, dass die relevante Zeitkonstante nicht allein durch die Messkapazität nach Abbildung 5 bestimmt ist, sondern zusätzlich beschränkende Elemente in der Messchaltung mitwirken.

Die gemessene Abweichung der Kapazität durch die Präsenz von einem oder mehreren 100 µm starken Papierstreifen gegenüber dem Messwert ohne Papierstreifen ist in Abbildung 6 für Umladezeiten von 21ns bis 188ns dargestellt. Erkennbar ist die durch die Leitfähigkeit des durch das Papier gebundenen ionischen Films bestimmte Ladekurve der resultierenden RC-Schaltung, welche direkt die gemessene Ladungsmenge bestimmt. Bei dickeren Schichten ab etwa 4 Papierstreifen bestimmt zunehmend nicht mehr der Filmwiderstand, sondern die Komponenten der Schaltung die Zeitkonstanten.

Zeitkonstanten im Nanosekundenbereich stellen Sensoren, die nach dem Charge-Transfer-Verfahren arbeiten, vor Herausforderungen. Dies gilt insbesondere für Sensoranwendungen welche einen hohen Fokus auf die Herstellkosten aufweisen und auf diskrete Bauteile basieren. Die mimimale Umladedauer ist oft durch die Taktfrequenz des ansteuernden Mikrocontrollers begrenzt. Die verwendeten Analogschalter müssen entsprechend schnell schalten, ebenfalls muss das Timing für sämtliche relevanten Schaltungsteile über alle Betriebbedingungen sichergestellt werden. Eine weitere Limitierung besteht in der begrenzten Treiberfähigkeit der I/O-Pins gängiger Mikrocontroller in Verbindung mit großflächigen Messelektroden. Typischerweise müssen bei dieser Art Sensor Kapazitäten im Picofaradbereich umgeladen werden. Zusätzlich wirken parasitäre Kapazitäten im Bereich von einigen zehn Picofarad.

Anhand von Messungen mit mehreren Umladezeiten kann, für bekannte Flüssigkeiten, eine dünne Flüssigkeitsschicht erkannt, sowie die Dicke der Schicht abgeschätzt werden. Der betrachtete Sensor ist in der Lage die typischerweise verbleibenden Flüssigkeitsfilme bei der Füllstandmessung von Wasser zu erkennen sowie deren Einfluss zu kompensieren. Ebenfalls ist eine Präsenzdetektion des Tanks sowie eine kontinuierliche Füllstandmessung möglich. Mit zusätzlichen Techniken können bei dieser Art der Messung Einflüsse durch Berührung der Flüssigkeit sowie durch parasitäre Kapazitäten zur Sensorumgebung minimiert werden.

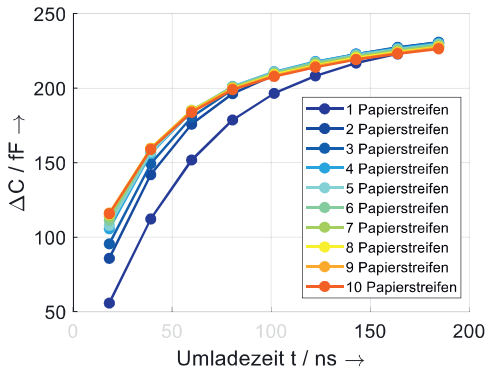


Abb. 6: Kapazitätsänderung von Filmen mit erhöhter Leitfähigkeit, hier NaCl.

Messung mit breitbandigem Rauschmessverfahren

Die Füllstandmessung von stark leitfähigen, zähflüssigen Medien mit dicken, verbleibenden Flüssigkeitsfilmen stellt die Messtechnik aufrgund der daraus resultierenden sehr kurzen Zeitkonstanten nach Abbildung 5 vor besondere Herausforderungen. In diesem Fall muss angestrebt werden trotz der hohen Leitfähigkeit R_σ tatsächlich die Kapazität C_1 auszuwerten. Der Füllstandsensor arbeitet deshalb mit Signalfrequenzen im Bereich von einigen bis zu mehreren hundert MHz. Dabei wird ein breitbandiges, weißes Rauschsignal moduliert und die von der Sendeelektrode auf die Empfangselektrode gekoppelte Leistung des Nutzsignals gemessen. Somit ist der Sensor in der Lage, Füllstände auch unter erschwerten Bedingungen, wie beispielsweise bei Einsatz von Waschmitteln oder Milch, zuverlässig zu bestimmen.

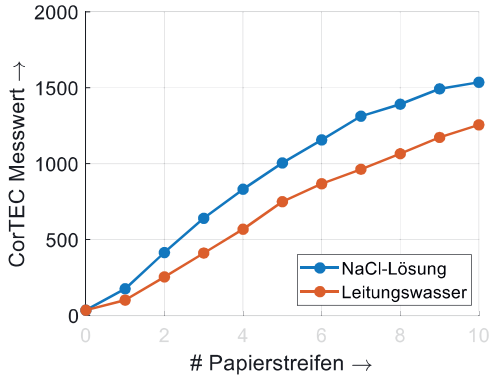


Abb. 7: Betragsänderung bei variierender, simulierter Filmdicke leitfähiger Medien.

Abbildung 7 zeigt den Verlauf des Messwertes des Sensors für ein bis zehn Lagen in Leitungswasser sowie in Kochsalzlösung getränkte Papierstreifen. Untersuchungen haben gezeigt, dass getränkte Papierstreifen ein repräsentatives Verfahren darstellt um Filme zu simulieren. Der Messaufbau entspricht dem des auf dem Charge-Transfer-Varfahren basierendem Vergleichssensor. Da bei dieser Messung im Wesentlichen der kapazitive Teil C_1 der Messstrecke ausgewertet wird, kann der Sensor selbst noch bspw. die neunte von der zehnten Papierschicht unterscheiden. Vergleichbare Ergebnisse liefert der Sensor ebenfalls für Papierstreifen, die sowohl in NaCl-Lösung als auch in destilliertem Wasser getränkt werden.

Ergebnis

Die Diskussion der Ergebnisse zeigt, dass das auf dem Charge - Transfer basierenden Messverfahren für wässrige Medien mit dünnen Wandfilmen sinnvoll einsetzbar ist. Für hoch viskose Medien jedoch erweist sich der Einsatz eines Messverfahrens mit höherer Messfrequenz, wie z.B. das in dieser Untersuchung eingesetzte modulierte breitbandige Rauschsignalverfahren, als vorteilhaft, da hier die hohe Leitfähigkeit der Wandfilme geringere Einflüsse auf das Messergebnis haben und die tatsächliche Kapazität sowie deren Änderung erfasst wird.

Beide Verfahren sind in der Lage, die geringe Kapazitätsänderung für die Präsenzerkennung des Mediengefäßes reproduzierbar zu erkennen und den Füllstand von wässrigen Medien sowie von adhäsiven, viskosen Medien trotz Filmbildung zu erfassen.

Literaturnachweis

[1] P. Hal, *Sensor Review* 19, Issue 2, 96-105 (1999)

[2] W. Du, *Series in Sensors*, *CRC Press,* (2015)

DOI 10.5162/sensoren2019/P1.5

Kostengünstige Nass/Nass-Differenzdruckmessung

Reinhard Werner, Henrik Schöpe
Analog Microelectronics GmbH, An der Fahrt 13, 55124 Mainz, Deutschland

Zusammenfassung

In dieser Veröffentlichung präsentieren wir ein Konzept für einen kostengünstigen, miniaturisierten Sensor zur Messung von Druckdifferenzen unter rauen / nassen Umgebungsbedingungen. Auf die in solchen Fällen oft eingesetzte Trennmembran mit Ölvorlage, die das Messelement vom jeweiligen Medium isoliert, wird hier verzichtet. Stattdessen werden zur Differenzdruckmessung zwei Siliziummesselemente eingesetzt und die Differenzdruckmessung durch zwei Druckmessungen mit relativem Bezug und anschließender Differenzbildung ersetzt. Konstruktionsbedingt kommen die sensiblen Messelemente nur mit ihrer unempfindlichen Rückseite mit dem jeweiligen Messmedium in Berührung, sodass auch elektrisch leitende Flüssigkeiten oder schwach korrosive Gase als Messmedium verwendet werden können.

Abstract

In this publication we present a concept for an inexpensive, miniaturized sensor to measure pressure differences in harsh or wet environments. The sensor has no separation membrane with oil buffer isolating the sensing element from the respective measuring medium. The differential measurement is done by performing two relative pressure measurements using two identical silicon sensing elements and subsequent subtraction of the sensing elements' output signals. Due to its design only the insensitive backside of the delicate sensing elements comes into contact with the respective measuring medium, so that even electroconductive fluids and weak corrosive gases can be used as measuring medium.

Keywords: Differenzdrucksensor, Medienresistenz, Nass / Nass, differential pressure sensor, harsh environment, wet / wet.

Einleitung

Auf dem Messen und Überwachen von Drücken, insbesondere von Druckdifferenzen, basiert die Funktion vieler Anlagen und Geräte. Beispielhaft sind folgende Einsatzgebiete zu nennen:

- Füllstandshöhe in Flüssigkeitstanks
- Überwachung von Filtern
- Durchfluss- Volumenstrommessung

Grundlagen und Stand der Technik

Die Umwandlung der mechanischen Größe „Druck" *p* in eine elektrische Größe erfolgt meist auf dem gleichen Weg: Der Druck *p*, definiert als Kraft *F* pro Fläche *A*

$p = F/A$ [Pa]

wirkt auf eine Membran und deformiert diese. Die Deformation / Auslenkung kann mittels diverser physikalischer Prinzipien gemessen werden, wobei hier nur die gängigen genannt werden sollen:

- induktives Messprinzip
- kapazitives Messprinzip
- resistives Messprinzip

In den meisten industriellen Drucksensoren kommt heute das resistive Prinzip zur Anwendung.

Die Methoden der modernen Halbleiterindustrie ermöglichen es, hochgenaue, stark miniaturisierte Druckmesselemente mit Membran, Messwiderständen und Leiterstrukturen in Silizium als MEMS (Micro-Electro-Mechanical-System) herzustellen. Abb. 1 zeigt den prinzipiellen Aufbau eines solchen Druckmesselements auf Siliziumbasis.

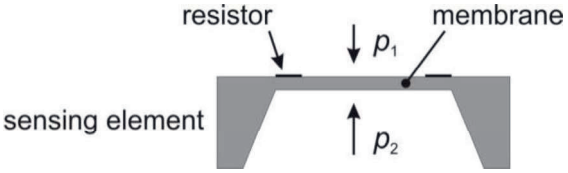


Abb. 1: Aufbau eines Druckmesselements auf Siliziumbasis.

Mit einem solchen MEMS Element lässt sich ein hoher Miniaturisierungsgrad, ein großes Nutzsignal auf Grund des piezoresistiven Effekts [1] und eine einfache Signalauswertung realisieren. Ferner ermöglicht das Halbleiterfertigungsverfahren große Stückzahlen und eine kostengünstige Fertigung.

Für die Druckmessung werden die mechanischen Spannungen in der Siliziummembran unter Ausnutzung des piezoresistiven Effekts gemessen. An geeigneten Stellen werden hierfür vier Messwiderstände implementiert und zu einer Wheatstonschen-Messbrücke verschaltet (siehe Abb. 2). Beim Anlegen einer Druckdifferenz $\Delta p = p_1 - p_2$ vergrößern sich die Werte der Widerstände R_2, R_3 und gleichzeitig reduzieren sich die Werte von R_1 und R_4.

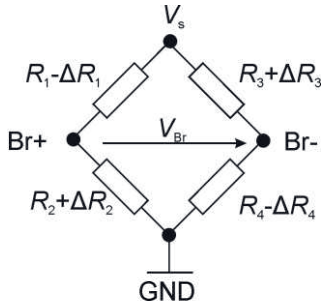


Abb. 2: Messbrücke

Für eine idealisierte Brücke mit vier gleichen Widerständen R und einer druckabhängigen Widerstandsänderung ΔR gilt für die Ausgangsspannung der Messbrücke V_{Br}:

$$V_{Br} = \frac{\Delta R(\Delta p)}{R} V_s$$

worin V_s die Versorgungsspannung bezeichnet.

Das resultierende Ausgangssignal kann mittels einer geeigneten Auswerteelektronik verstärkt, aufbereitet und als normiertes Signal ausgegeben werden. Hierfür bietet die Industrie passende Standard-lösungen als integrierte Schaltungen (ASIC) an [2][3].

Diese können zusätzlich Nichtlinearitäten und Temperaturabhängigkeiten des Brückensignals kompensieren.

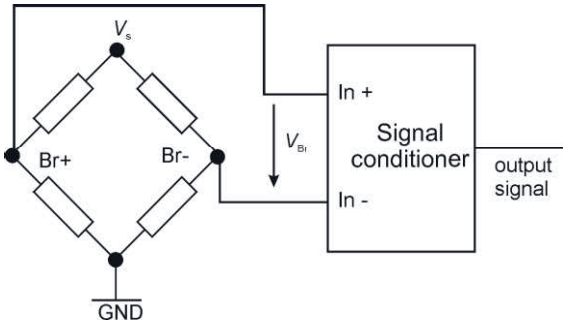


Abb. 3: Auswertung einer Sensorbrücke

In Abb. 4 ist der typischen Aufbau eines einfachen industriellen Differenzdrucksensors mit Siliziummesselement und Auswerteelektronik dargestellt [4]. Beide sind mittels Aufbau- und Verbindungstechnik zusammen auf einem Substrat montiert, das gleichzeitig der elektrischen Kontaktierung dient. Das empfindliche Innere wird durch eine druckdicht montierte Kappe geschützt. Die Druckzuführung erfolgt über die beiden Druckanschlüsse 1 und 2.

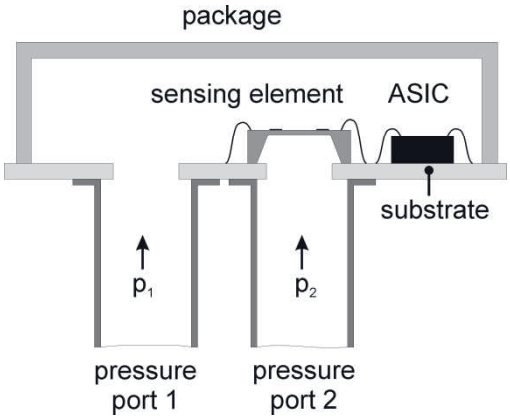


Abb. 4: Drucksensor mit einem Messelement

Diese einfachen Sensoren eignen sich nur zur Differenzdruckmessung von trockenen und nicht korrosiven Gasen. Für andere Medien sind sie nur sehr eingeschränkt verwendbar.

Um auch Differenzdruckmessungen an Flüssigkeiten und korrosiven Gasen durchführen zu können, muss das sensible Sensorinnere von dem Messmedium isoliert werden. Bei handelsüblichen Sensoren werden dazu verschiedene konstruktive Methoden [5][6] verwendet, auf die im Folgenden kurz eingegangen wird.

Die sicherste Möglichkeit, das Messelement und die dazu gehörige Auswerteelektronik vor dem Messmedium zu schützen, ist das Einsetzen einer elastischen Trennmembran aus Edelstahl und die Verwendung von Öl als Druckmittler (vgl. Abb. 5). Der konstruktive Aufwand führt zu einem entsprechenden Bauvolumen und in der Folge zu höheren Kosten.

Darüber hinaus verringert die Eigensteifigkeit der verwendeten Edelstahlmembran die Empfindlichkeit, so dass sich diese Methode nicht zur Messung von sehr kleinen Differenzdrücken eignet.

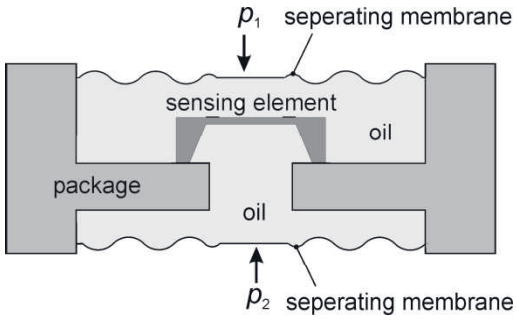


Abb. 5. Prinzipieller Aufbau eines Differenzdrucksensors mit Trennmembranen

Eine kostengünstige Alternative stellt das Abdecken / Vergießen des empfindlichen Sensorinneren dar. Dazu werden üblicherweise Silikongele oder Polymerschichten (Parylene) verwendet. Abb. 6 zeigt den Aufbau eines solchen Systems.

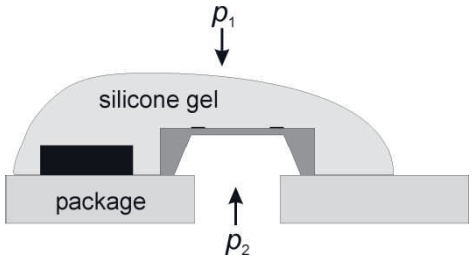


Abb. 6: Drucksensor mit Silikongelabdeckung

Dem Vorteil des geringen Platzbedarfs und der geringen Kosten steht der Nachteil einer nur eingeschränkten Medienkompatibilität gegenüber. Darüber hinaus bietet diese Methode nur einen eingeschränkten Langzeitschutz. Über einen längeren Zeitraum können Bestandteile des Mediums durch die Schutzschicht hindurch diffundieren oder sie unterkriechen.

Alternatives Lösungskonzept

In der Literatur [7][8] wird ein Differenzdrucksensor mit zwei Messelementen beschrieben, bei dem die Differenzdruckmessung durch zwei Relativdruckmessungen ersetzt wird. Dieser Ansatz, der von den Autoren auf dem aktuellen Stand der Technik umgesetzt wurde, verwendet zwei identische Silizium(Si)-Messelemente. Er ermöglicht eine kostengünstige Realisierung von medienkompatiblen Differenzdrucksensoren.

Abb. 7 zeigt den prinzipiellen Aufbau eines solchen Differenzdrucksensors. Die beiden Messelemente sind zusammen mit einer Signalauswerteelektronik auf einem Substrat montiert. Das empfindliche Sensorinnere ist durch eine montierte Kappe mechanisch geschützt. Die Druckzuführung erfolgt durch die beiden Anschlussstutzen 1 und 2. Eine zusätzliche Bohrung dient dem Druckausgleich im Sensorinneren.

Beide Si-Messelemente messen den an liegenden Druck p_1 bzw. p_2 relativ zum Umgebungsdruck. Dabei kommen nur die unempfindlichen Rückseiten der Siliziummesselemente mit dem Messmedium in Kontakt. Auf diese Weise kann eine ausgezeichnete Medienkompatibilität erreicht werden, die nur durch die eingesetzten Materialien beschränkt ist.

Der Systemdruck, der gleichzeitig an beiden Ports dieses Differenzdrucksensors anliegen darf, ist auf Grund der zweifachen Relativdruckmessung auf den Nenndruck der einzelnen Messelemente beschränkt.

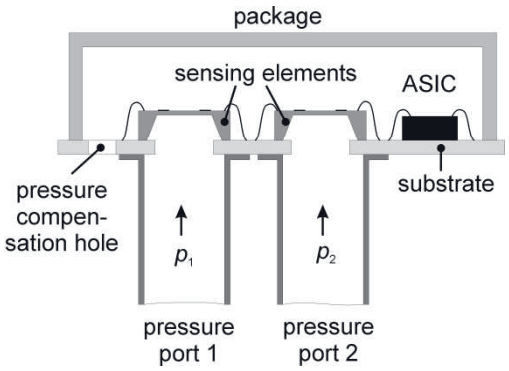


Abb. 7: Aufbau eines Differenzdrucksensors mit zwei Siliziummesselementen

Die verwendete Signalauswerteelektronik hat die Aufgabe, die Brückensignale der beiden Messzellen aufzubereiten und die Differenz zu bilden.

Die Signalaufbereitung kann rein analog oder nach analog / digital (A/D)-Wandlung digital erfolgen. Sie umfasst die Linearisierung, die Temperaturkompensation und den Abgleich auf ein Standardausgangssignal. Die Differenzbildung kann dabei vor der eigentlichen Signalaufbereitung oder erst am Schluss geschehen.

Um maximale Flexibilität und Genauigkeit zu erreichen, wurde bei dem realisierten Sensor die Signalauswertung gemäß Abb. 8 realisiert.

Die einzelnen Messbrückensignale werden hier getrennt voneinander verstärkt und A/D gewandelt. Anschließend erfolgt für jedes Digitalsignal eine Linearisierung, Temperaturkompensation und ein Abgleich auf ein Standardsignal mit Hilfe eines Mikro(µ)-Controllers.

Die Differenzbildung wird erst im letzten Schritt durchgeführt.

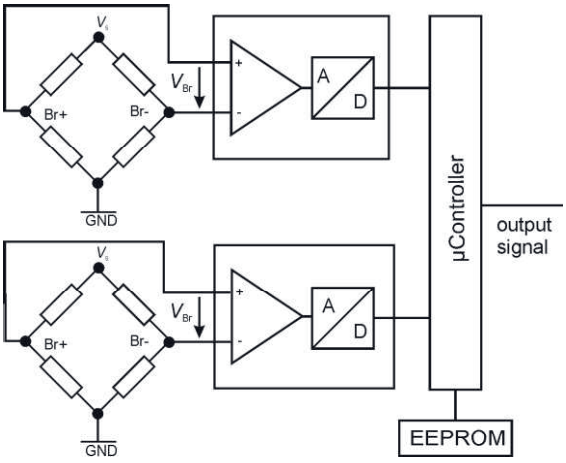


Abb. 8: Sensor mit Auswerteelektronik.

Realisierter Prototyp

Abb. 9 zeigt das Foto eines Prototyps des realisierten Differenzdrucksensors. Er wurden als miniaturisierte board mount Drucksensoren im DIL-Gehäuse auf Basis eines Keramiksubstrats (Abmessung: 15 x 15 mm^2) konzipiert.

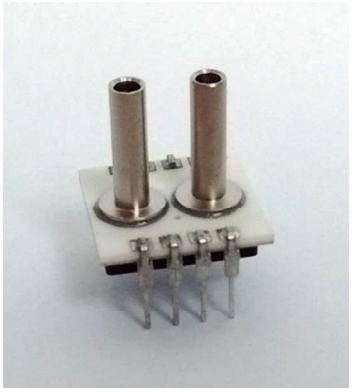


Abb. 9 Bild des Differenzdrucksensors

Kalibrierung

Zum Abgleich der Prototypen werden die beiden Messelemente des jeweiligen Sensors an drei verschiedenen Umgebungstemperaturen bei jeweils mehreren Drücken charakterisiert. Diese individuellen Messdaten werden mittels eines Algorithmus zur Berechnung von Korrekturkoeffizienten genutzt, die im EEPROM des jeweiligen Sensors gespeichert werden. Im späteren Betrieb errechnet der µController auf Basis der gespeicherten Korrekturkoeffizienten und der aktuellen, digitalisierten Brückensignale die standardisierten Relativdrucksignale und daraus als Differenz das standardisierte Differenzdrucksignal.

Messergebnisse

Die nachfolgenden Messergebnisse wurden mit Sensorprototypen ermittelt, die mit zwei identischen 1 bar Druckmesselementen aufgebaut wurden. Um möglichst gleiche Eigenschaften der Messelemente garantieren zu können, wurden jeweils benachbarte Messelemente von einem Wafer verwendet.

Die Sensorprototypen wurden nach der Kalibrierung in einem Druck- und Temperaturprüfstand untersucht.

In einem ersten Schritt wurden der Gleichlauf der beiden Relativdrucksensoren und die Systemdruckabhängigkeit des Ausgangssignals des Differenzdrucksensors bei Raumtemperatur überprüft. Dazu wurden die beiden Druckanschlüsse gemeinsam an einen Druckkalibrator angeschlossen und der Druck, beginnend bei 0 bar in 25 mbar Schritten bis 1 bar erhöht und anschließend in gleicher Weise wieder gesenkt. Nach dem Erreichen eines jeden Druckwerts wurden die standardisierten Relativdrucksignale und das Differenzdrucksignal aufgezeichnet. Insgesamt wurden 16 Druckzyklen durchlaufen. Die Temperatur wurde dabei konstant bei 25 °C gehalten.

Aus diesen Daten wurde der Fehler, bezogen auf das Nutzsignal (Full Span Output = FSO), bestimmt.

In Abb. 10 ist der Fehler der einzelnen Drucksignale als Funktion des anliegenden Systemdrucks für einen Sensor dargestellt. Der aufgetragene Fehler wurde aus der Differenz des jeweiligen Ausgangssignals zur idealen Kennlinie, normiert auf 1 bar, berechnet.

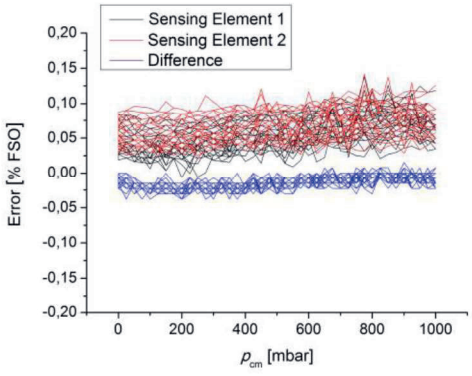


Abb. 10. Gleichlauf der zwei Si-Druckmesselemente bei Raumtemperatur

Im gesamten Druckbereich von 0 bis 1 bar weisen die beiden Relativdrucksensoren eine sehr gute Genauigkeit auf; der maximale Fehler ist kleiner als 0,15 %FSO. Das Differenzdrucksignal zeigt nahezu keine Abhängigkeit vom angelegten Systemdruck. Der maximale Fehler von 0,04 %FSO des Differenzdrucksignals und die geringe Streuung bestätigen den exzellenten Gleichlauf beider Relativdrucksensoren.

In einem zweiten Schritt wurden der Gleichlauf der Messelemente und die Systemdruckabhängigkeit der Prototypen unter Temperatureinfluss untersucht. Dazu wurde die Umgebungstemperatur im Intervall von -20 °C bis +80 °C in Schritten von 10 °C zyklisch variiert. Bei jedem Temperaturpunkt wurde der zuvor beschriebene Druckzyklus zwischen 0 und 1 bar einmal durchlaufen und die Relativdrucksignale sowie das Differenzdrucksignal aufgezeichnet.

Abb. 11 zeigt den Fehler der einzelnen Drucksignale als Funktion der Temperatur.

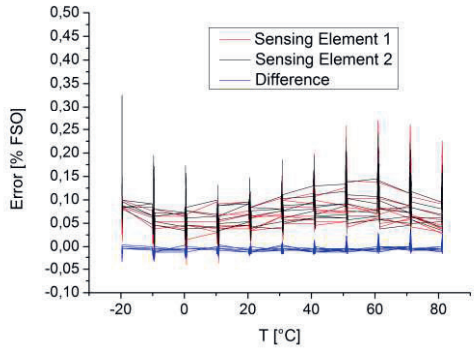


Abb. 11: Temperaturgang der zwei Si-Druckmesselemente und des Differenzsignals im Bereich von -20 °C bis 80 °C

Im untersuchten Temperatur- und Druckbereich weisen die beiden Relativdrucksensoren einen maximalen Gesamtfehler von kleiner als 0,35 %FSO auf; der Temperaturgang des Offsets ist kleiner als 0,15 %FSO.

Dies zeigt, dass die Temperaturkompensation bei beiden Messzellen gut funktioniert.

Das Differenzdruckausgangssignal weist im gesamten Temperatur- und Systemdruckintervall einen maximalen Fehler kleiner als 0,1 %FSO auf; die Temperaturabhängigkeit des Offsets ist kleiner 0,05 %FSO.

Die Messdaten zeigen, dass auch unter Temperatur ein exzellenter Gleichlauf der beiden Si-Messelemente erreicht werden konnte.

Für die weiteren Untersuchungen wurden Prototypen als Differenzdrucksensoren mit einem Druckbereich von 0 bis 100 mbar konfiguriert. Auf Grund der verwendeten 1 bar Si-Messelemente sind Systemdrücke von 0 bis 1000 mbar möglich.

Bei Raumtemperatur wurde an diesen Differenzdrucksensoren verschiedene Druckmessungen durchgeführt: Zum einen wurde die Differenzdruckkennlinie bei unterschiedlichen Systemdrücken zwischen 0 und 900 mbar aufgezeichnet. Zum anderen wurde die Systemdruckabhängigkeit des Ausgangssignals bei verschiedenen, konstant gehaltenen Druckdifferenzen untersucht.

In Abb. 12 ist der Fehler des Differenzdrucksignals eines Differenzdrucksensors bei Raumtemperatur als Funktion des anliegenden Differenzdrucks zwischen 0 und 100 mbar mit dem Systemdruck als Parameter dargestellt. Als Systemdruck wurden 0, 250, 500, 750 und 900 mbar gewählt. Zur Fehlerberechnung wurde die Abweichung des realen Messwertes von der idealen Kennlinie berechnet und auf 100 mbar FSO normiert. Im Rahmen der Fehlergrenzen konnte für den Differenzdrucksensor eine Messgenauigkeit von kleiner ±0,3 %FSO erreicht werden. Der Systemdruck zeigt dabei nur einen sehr geringen Einfluss. Die detaillierte Abhängigkeit des Ausgangssignales vom Systemdruck ist in Abb. 13 dargestellt. Aufgetragen ist der Fehler des Differenzdruckausgangssignales bei Raumtemperatur als Funktion des Systemdrucks für die Differenzdrücke von 0, 25, 50, 75 und 100 mbar. Nahezu unabhängig vom angelegten Differenzdruck kann eine leichte Systemdruckabhängigkeit beobachtet werden. Im Rahmen der Fehlergrenzen beträgt die maximale Abweichung von der idealen Kennlinie ±0,4 %FSO.

Beide Abbildungen zeigen, dass sich mit den beiden identischen 1 bar Messelementen 100 mbar Differenzdrucksensoren mit guter Messgenauigkeit realisieren lassen.

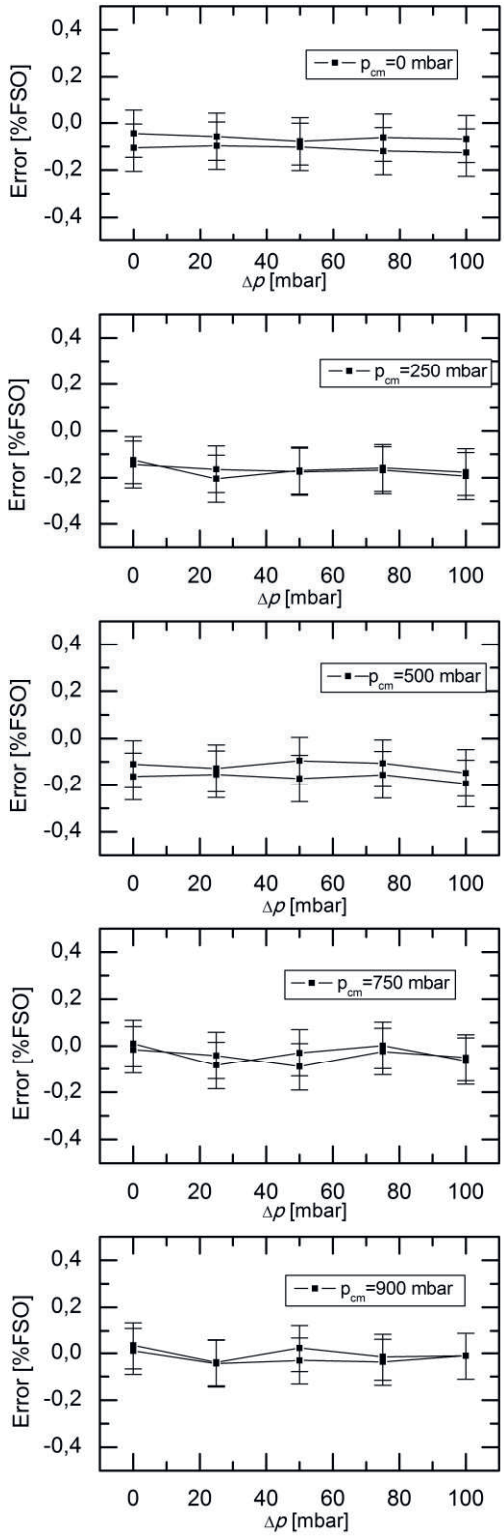


Abb. 12: Fehler des Differenzdrucksignals als Funktion des anliegenden Differenzdrucks bei unterschiedlichen Systemdrücken.

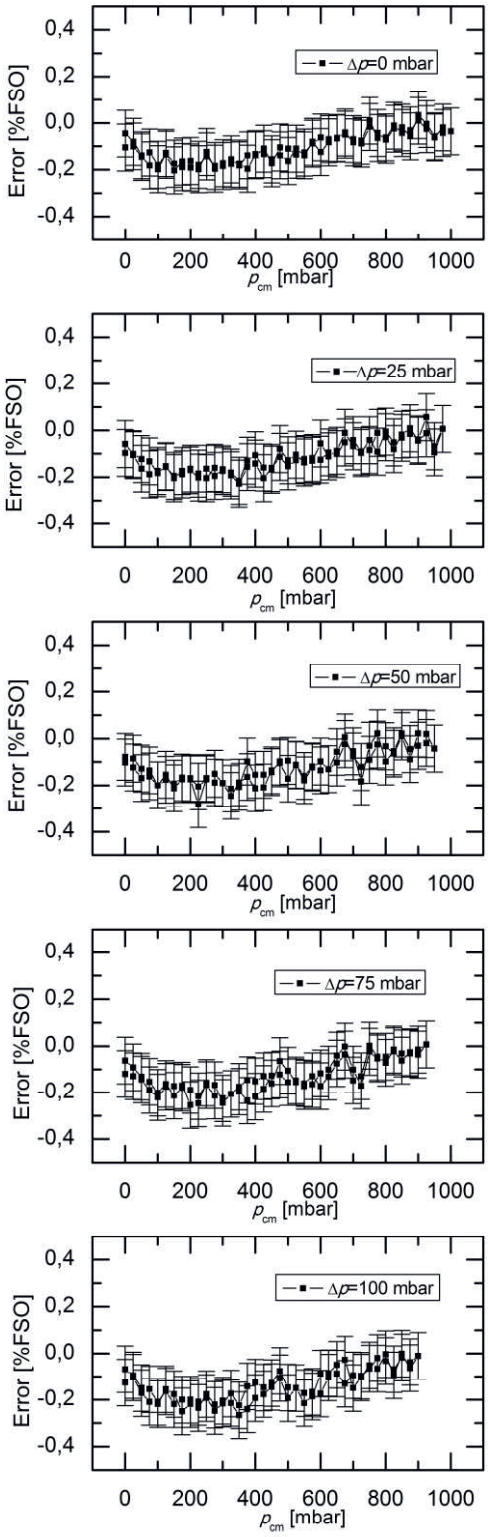


Abb. 13 Fehler des Differenzdrucksignals als Funktion des Systemdrucks für die Differenzdrücke von 0, 25, 50, 75 und 100 mbar.

Ausblick

Die Eigenschaften der mit 1 bar Messzellen aufgebauten 100 mbar Differenzdrucksensor-Prototypen haben ermutigende Ergebnisse geliefert. Daher soll das Konzept auf weitere Druckbereiche ausgeweitet werden.

Die vorliegenden Messungen sind an Prototypen mit Messelementen von demselben Wafer durchgeführt worden. Um die Serientauglichkeit des Konzeptes weiter zu untersuchen, müssen die Messungen an zusätzlichen Prototypen wiederholt werden, bei denen Messelemente von anderen Wafern / Losen zum Einsatz kommen. Erst dann kann ermittelt werden, mit welcher Streuung diese Differenzdrucksensoren produziert werden können.

Ferner ist noch zu klären, wie langzeitstabil der Gleichlauf zwischen den beiden Messelementen ist, bzw. wie stark sich deren Eigenschaften über die Zeit verändern.

Literaturnachweis

[1] G. Gerlach, R. Werthschützky, 50 Jahre Entdeckung des piezoresistiven Effekts-Geschichte und Entwicklungsstand piezoresistiver Sensoren, *Technisches Messen 72*, S. 53-76 (2005);

[2] ZSC31014 Resistive Sensor Signal Conditioner with Digital Output, Datenblatt, Integrated Device Technology, Inc., www.idt.com

[3] MAXQ7670A Microcontroller with 12-Bit ADC, PGA, 64KB Flash, and CAN Interface. Datenblatt, Maxim Integrated Products, Inc. www.maximintegration.com

[4] AMS 5915 - Drucksensor mit I2C-Ausgang für die PCB-Montage, Datenblatt. Analog Microelectronics GmbH. www.analog-micro.com

[5] Mark J. Clifford, Erfinder; Kavlico Corporation, Inhaber. Independent-Excitation crosscoupled differential-pressure transdurcer. United States patent US 6581468 B2. 2003 Jun. 24.

[6] D5100 Industrial Differential Pressure Transducer. Datenblatt. Measurement Specialties, Inc., a TE Connectivity company. www.TE.com/sensorsolutions D5100Industrial

[7] Walter Czarnocki, William, Erfinder; Motorola Inc., Inhaber. Gain-adjusting circuitry for combining two sensors to form a media isolated differencial pressure sensor. United States patent US 5471884. 1995 Dez. 5.

[8] Anthony D. Kurtz. Erfinder; Kulite Semiconductor Products, Inhaber. United States patent US 6272928 B1. 2001 Aug. 14.

DOI 10.5162/sensoren2019/P1.7

Elaboration of Reduced Graphene Oxide Electrodes Functionalized with rGO-ION Composite for Nitrite Detection in Water

A. Brahem[1]*, A. Al-Hamry[2*], A. Reddy[2], A. Bouhamed[2], S. Ben Nasrallah[1], L. G. Paterno[3], M. Ben Ali[1,4] and O. Kanoun[2]

[1]Electronic Department, University of Sousse, Higher Institute of Applied Science and Technology of Sousse, Sousse, Tunisia

[2]Technische Universität Chemnitz, Chair of Measurement and Sensor Technology, Chemnitz, Germany

[3]Laboratório de Pesquisa em Polímeros e Nanomateriais, Instituto de Química, Universidade de Brasília, Brasilia, Brazil

[4] NANOMISENE Lab, LR16CRMN01, Centre for Research on Microelectronics and Nanotechnology of Sousse, Technopole of Sousse B.P. 334, Sahloul 4034, Sousse, Tunisia

*email: ammar.al-hamry@etit.tu-chemnitz.de

Abstract

In this paper, we propose a novel nitrite sensor based on reduced graphene oxide electrodes elaborated by light scribing. . First, graphene oxide was deposited on flexible substrate, which was previously plasma treated and wetted by a Poly(diallyldimethylammonium chloride) polymer solution. Then, in order to tune and enhance the sensitivity, the working electrode has been modified by layer-by-layer deposition of iron oxide-reduced graphene oxide nanocomposite (ION-rGO). The working electrode has been characterized at different levels of modification by scanning electron microscopy (SEM) and dispersive X-ray spectroscopy (EDX) showing a successful deposition and functionalization. Furthermore, electrochemical characterizations have been carried out by impedance spectroscopy and cyclic voltammetry (CV) and quantitative measurements for nitrite detection by differential pulse voltammetry (DPV). The experimental results demonstrate a good electrochemical response for nitrite with a sensitivity of 0.314 µA/mM. Several successive scans by CV at 50 mV/s in 1 mM nitrite solution show that the electrode remains stable and the sensitivity slightly reduces due to surface contaminations. The proposed fabrication procedure has a significantly better performance in comparison with some reported state-of-the art electrodes [12].

Keywords-component; Electrochemical sensors, Nitrite detection; Light scribe Graphene; Layer-by-Layer Deposition, reduced Graphene Oxide (rGO); Iron-oxide nanoparticles

1 Introduction

Over the last few years, in Germany due to excessive use of fertilizers, ground water is contaminated by nitrite in several regions, so that it has been condemned by the European court of justice for breach of EU law in this respect. Exceeding the allowed level of nitrite in drinking water causes several health problems, such as the “blue baby syndrome”. Ingested nitrite can also react with amides and amines then forming potent carcinogens, such as nitrosamines [1-2]. The amount of nitrite in drinking water should therefore not exceed 3 mg/L referring to the World Health Organization (WHO) [3].

Hence, a reliable and accurate method for nitrite detection in water is important and recently received much attention. So far, many techniques have been adopted to detect nitrite, which involve tedious procedures and are time-consuming, such as chromatographic methods, ion spectrophotometry and chemiluminescence 4. Alternatively, electrochemical techniques are widely used because they provide highly selectivity, sensitivity and rapid nitrite determination compared with other traditional methods [5]. The electrochemical determination of nitrite is reported in literature either by oxidation or reduction. However, the oxidative nitrite determination is generally preferred because the nitrite reduction is basically known to suffer from interferences such as reduction of nitrate and molecular oxygen [6].

Recently, the applications of screen-printed electrodes (SPE) have significant advantages, such as simplicity, and low fabrication cost [7]. The whole electrode system, including working, reference and counter electrodes is printed on a substrate. Usually, for bare electrodes, the oxidation of nitrite occurs at high overpotential. Therefore, a modification of the screen printed electrodes is necessary in order to decrease the over potential and thereby to increase its sensitivity.

In this paper, we propose a novel nitrite sensors based on reduced graphene oxide electrodes produced by light scribing (LSG) [8-9]. To realize a high stability of the electrodes, graphene oxide was deposited on flexible substrate, which was plasma treated and wetted by a Poly(diallyldimethylammonium chloride) polymer solution. In order to tune and enhance the sensitivity, the working electrode has been modified by reduced graphene oxide- iron oxide nanoparticle (rGO-ION). The modification was made by layer-by-layer deposition of PDAC polymer and rGO-ION. The development of this modified electrode for nitrite detection is not expensive compared to conventional

electrodes. This novel fabrication technique realizes low-cost sensors able to be used on-site.

2 Material and methods

2.1 Chemicals

The graphene oxide (GO) is commercially available in aqueous solution with a concentration of 4 g/L, prepared by modified Hummers method and it is obtained from Graphenea S.A, Spain. Iron oxide-reduced graphene oxide nanocomposite (ION-rGO) was synthesized as following; 200 mL of GO 0.1 g/L with a pH of 11 was prepared. Under stirring, a solution of 0.5 g of $FeSO_4.7H_2O$ in 10 mL of deionized water was added to GO solution dropwise. The solution was stirred further for 30 m and the mixture was left decanting aided by a magnet. Cleaning steps were carried out by removing the supernatant and adding deionized water and decanting with the magnet. Finally, the particles were suspended NH3 solution (pH 9) using the horn sonicator and then centrifugated.

In this study, all reagents used were of analytical grade and distilled water was used throughout. Sodium nitrite was purchased from Aldrich. Freshly made 0.1 M phosphate buffer solution (PBS) in different pHs, consisting of sodium phosphate dibasic (Na2HPO4) and sodium phosphate monobasic (NaH2PO4) was used during experimental testes as the supporting electrolyte.

2.2 Preparation of the electrodes

The preparation of LSG was done by using a commercially available laser source (LightScribe DVD drive with 788nm) controlled by means of a computer software (Nero Cover Design) to reduce the graphene oxide (GO). For that, GO in an aqueous solution of 0.3 wt. % was casted directly onto polyethylene terephthalate (PET) substrate which was plasma treated. The substrate was already glued to a DVD disc by adhesive spray. GO takes 24 h to dry under ambient conditions. Then, the obtained film was reduced by light scribing method. More than 20 electrodes can be produced on a single DVD disc. The electrodes were prepared by pasting Ag/AgCl to the reference electrode (RE). Working electrodes (WE) were functionalized with the ION-rGO nanocomposite, while counter electrodes (CE) were kept unchanged i.e. only reduced graphene oxide (rGO). Finally, a passivation layer was applied to obtain final electrode as shown in Fig 1a.

rGO working electrodes were functionalized with Layer-by-layer technique using tow oppositely charged materials The first step was drop casting a layer of PDAC polymer on the working electrode WE and washed by deionized water in order to remove unbound material. The same procedure is repeated to deposit ION-rGO nanocomposite as shown in the figure 1b. This process forms one bilayer. Finally, the functionalized electrode was dried with nitrogen gas and it is ready for the electrochemical testes.

2.3 Characterization and Measurement Setup

The examination of the electrode morphology was carried out by a scanning electron microscopy (SEM) equipped with an energy dispersive X-ray spectroscopy (EDX). For the electrochemical experiments, PalmSens 4 (PalmSens BV) was used to perform cyclic voltammetry (CV) and differential pulse voltammetry (DPV) measurements.

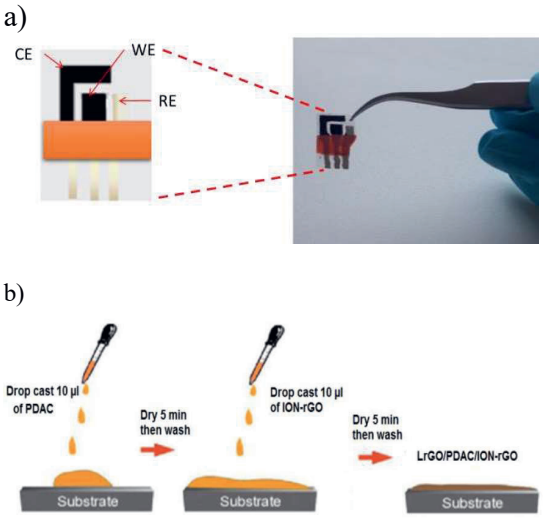


Fig1. a) LSG fabricated electrode configuration and b) Graphical illustration of the electrode functionalization

3 Results and Discussion

3.1 Surface Characterization electrodes

The morphology of electrodes after each preparation step can be clearly seen from the SEM images displayed in in Fig2. In Fig. 2a, the initial GO cast film shows a homogeneous film. After laser-scribing, when GO is photothermally reduced to rGO, the surface becomes rougher . This morphological transformation may be a consequence of a chemical bonds breaking during the reduction process. The lattice defects during the surface reconstruction after removing hydroxyl and epoxy groups may also contributed to change the morphology. The formation of rGO was further confirmed by EDX, which is performed using the same samples. EDX results show that carbon to oxygen relative concentration increase from GO to rGO samples, proving the reduction efficacy of the laser scribing process. . Finally, the last step of modification with PDAC/ION-rGO nanocomposite film is seen in Fig.2c. . It is possible to detect rGO sheets decorate with nearly spherical nanoparticles, which are the ION. EDX measurements detected the presence of iron and confirm the successful functionalization of the working electrode.

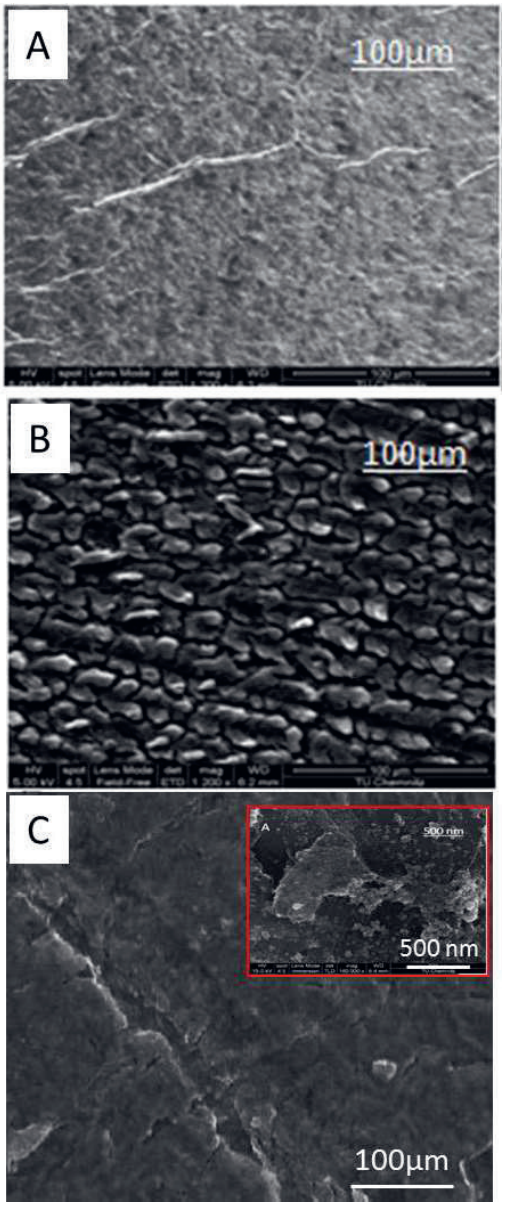


Fig 2. SEM images of (a) graphene oxide on top of the PET substrate, (b) rGO obtained after laser scribing, and (c) PDAC/ION-rGO nancomposite film on top of the laser scribed rGO electrode, Inset: image with higher magnification with noticeable ION nanparticles

3.2 Electrochemical behavior

First of all ,a test of the electrode without target ion with a solution composed of 0.1M of KCl and 10^{-3} M ferri/ferrocyanide [Fe (CN) 6]$^{3-/4-}$ was performed. For both bare and modified (PDAC/ION-rGO) electrodes we tested by CV, Fig.3. For modified surfaces, the [Fe (CN) 6]$^{3-/4-}$ redox couple provides a good electrochemical probe to study, because its oxidation and reduction proceed via a simple one-electron-transfer process. This chemical reaction (1) is the base of this redox-couple's electrochemistry:

$$[Fe(CN)_6]^{3-} + [\mathrm{Fe(CN)_6}]^{3-} e^- \leftrightarrow [Fe(CN)_6]^{4-} \quad (1)$$

An improvement between bare and modified electrodes in the electrochemical response can be seen in Fig.3 by the increase of the ferri/ferrocyanide oxidation and reduction peak currents after electrode modification with the PDAC/ION-rGO film. This improvement is ascribed to the increase of the surface area provided by the electrode modification.

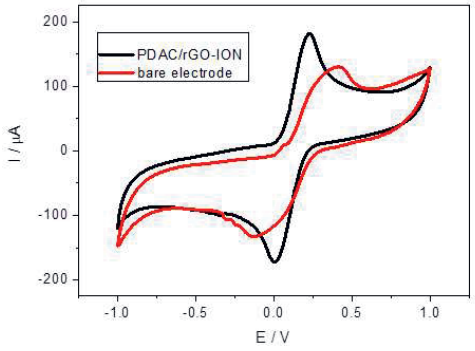


Fig.3.Cyclic voltammograms of Ferri-Ferrocyanide solution 10^{-3} M of bare and modified electrode in 25 mV/s

CV was also performed at different scan rates (25 mV/s, 50 mV/s, 100 mV/s, 150 mV/s, 200 mV/s. It was found a linear relationship between anodic/cathodic peak currents (i_p) and the square-root of the scan rate ($v^{1/2}$), shown in Fig.4.

The electro-active surface area of the electrodes was calculated by using the Randles-Sevcik equation. This equation predicts that the current peak should be proportional to the square root of the scan rate when voltammograms are taken at different scan rates. At 25 °C, the Randles-Sevcik equation can be reduced to the following equation (2) [11]:

$$i_p = 2{,}69\ 10^5\ n^{\frac{3}{2}} A D^{\frac{1}{2}} C v^{\frac{1}{2}} \quad (2)$$

in which n is the number of electrons appearing in half reaction for the redox couple (n=1 in our case), A is the active surface of the electrode (cm^2), C represents the concentration of the electroactive species (mol cm^{-3}), ν is the speed at which the potential is swept (V/s), and D is the diffusion coefficient of the analytes ($cm^2 s^{-1}$). From this equation, we can directly calculate the electroactive area for bare and modified electrodes.

Before the modification A is 0.2 cm^2, and it increases to 0.29 cm^2 after functionalization with the layer-by-layer assembled PDAC/ION-rGO film.

To choose the proper pH value at which the best detection of nitrite takes place, CV was performed by changing the pH value of the PBS solution: 3.5, 4.5, 5.7, 6.0, 6.5 and 7.4 as shown in in Fig.5a and b. The nitrite concentration is fixed at 1 mM as well as scan rate used is fixed at 50mV/s. It is found that oxidation current of nitrite ions for the modified electrodes has a better peak achieved at pH=5.7.

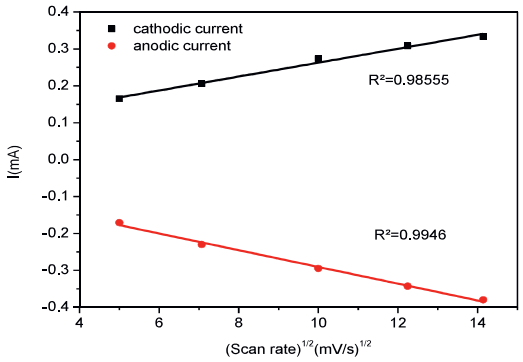


Fig. 4 Anodic and cathodic current of nitrite versus Square root of scan rate Vs

The sensor response to nitrite detection was tested by differential pulse voltammetry (DPV). The DPV curves of the modified electrode following the change in nitrite ion concentration were measured and shown in Fig. 6. The peak oxidation current increases with increasing the nitrite concentrations. This increase is almost linear.

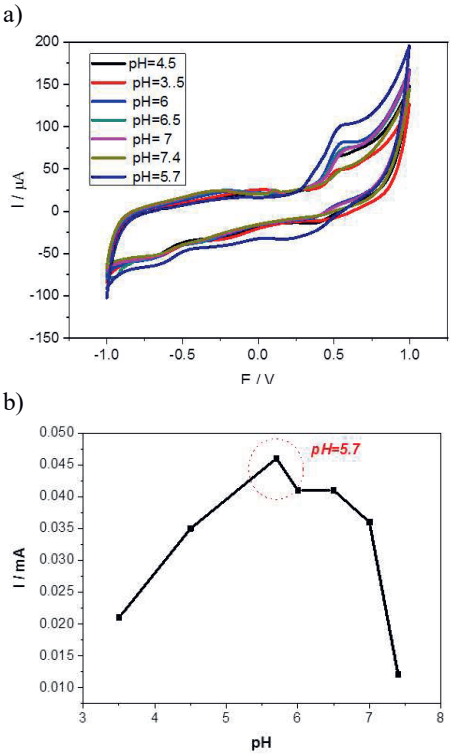


Fig. 5. (a) Cyclic voltammograms of detection of nitrite in different pHs and (b) Variation of nitrite oxidation peak current value with the pH. All measurements performed with the LSG PDAC/ION-rGO electrode.

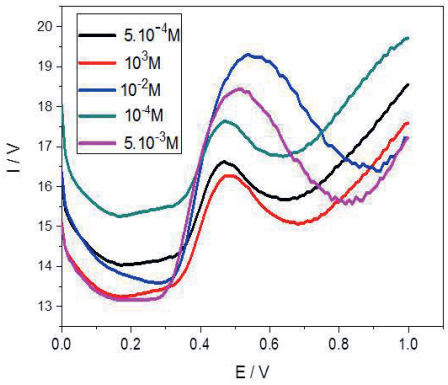


Fig. 6: DPV of modified electrode with RGO-IONs/PDAC for different nitrite concentrations

Fig. 7 illustrates the calibration curve describing the variation of currents as a function of increasing nitrite concentrations. We find that the method has a sensitivity of 0.314 µA/mA with (R^2=0.93).

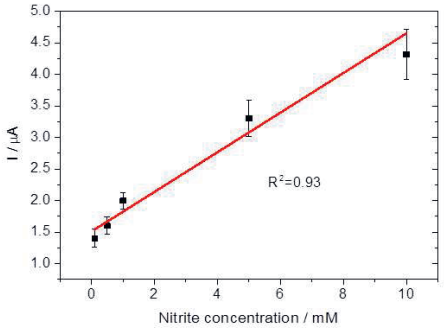


Fig.7 Nitrite calibration curve

Moreover, the reproducibility was studied as an important parameter of sensor characterization. The reproducibility was investigated from the current and potential response from CV using four electrodes independently prepared in the same way as seen in Fig. 8. Furthermore, the electrodes show stability of the oxidation peak current even after tens of cycles. Several successive scans by cyclic voltammetry of with a rate of 50 mV/s in 1 mM nitrite solution show that the electrode remains stable and using plasma and PDAC wetting make the electrode more stable.

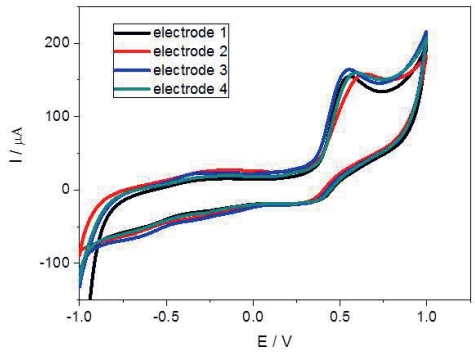


Fig. 8 Cyclic voltammograms of nitrite registered with four LSG PDAC/ION-rGO electrodes produced in different batches.

4 Conclusion

In this work, sensitive electrodes for the detection of nitrite in water were developed using PDAC/ION-rGO modified laser-scribed reduced graphene oxide electrode. The electrodes' modification was performed by layer-by-layer deposition technique. The sensor has a sensitivity of 0.314 µA/mM, which is better than state of the art sensors [12].

This study combines the use of nanomaterials and the realization of low-cost electrodes, which can be disposable, in the detection of nitrite in water. This can be an alternative for the detection and sensing of many other substances. The realized sensor has a fast response and is economically feasible.

Acknowledgement

This work was carried out with the support of the Nitramon project (project no. 100339427) funded by the Sächsische Aufbaubank (SAB) and the European Social Fund (ESF) and the project Electrochemical Sensors for Environmental Applications (ESEA) (project no. 249305) funded by DAAD. Authors acknowledge the support of Aleksandr Oshchepkov (within the project Nitramon) from Technische Universität Chemnitz, group of supramolecular chemistry, chemical department, for solutions preparation and discussions. A. A.-H. and O. K. acknowledge the financial support within the project InnoTeam-SimplySafe (contr. nr.100331073) funded by the Sächsische Aufbaubank (SAB) and the European Social Fund (ESF). L.G. P. and O. K. acknowledge the Alexander Von Humboldt Foundation.

Literature

1. Bryan, N.S., 2006. Nitrite in nitric oxide biology: Cause or consequence?: A systems-based review. Free Radical Biology and Medicine, 41(5), pp.691-701.
2. Chen, K.K. and Rose, C.L., 1952. Nitrite and thiosulfate therapy in cyanide poisoning. Journal of the American Medical Association, 149(2), pp.113-119.
3. Gorchev, H.G. and Ozolins, G., 1984. WHO guidelines for drinking-water quality. WHO chronicle.
4. Wang, Q.H., Yu, L.J., Liu, Y., Lin, L., Lu, R.G., Zhu, J.P., He, L. and Lu, Z.L., 2017. Methods for the detection and determination of nitrite and nitrate: A review. Talanta, 165, pp.709-720.
5. Kalimuthu, P. and John, S.A., 2009. Highly sensitive and selective amperometric determination of nitrite using electropolymerized film of functionalized thiadiazole modified glassy carbon electrode. Electrochemistry Communications, 11(5), pp.1065-1068.
6. Caro, C.A., Bedioui, F. and Zagal, J.H., 2002. Electrocatalytic oxidation of nitrite on a vitreous carbon electrode modified with cobalt phthalocyanine. Electrochimica Acta, 47(9), pp.1489-1494.
7. O'Halloran, M.P.; Pravda, M.; Guilbault, G.G. Prussian Blue bulk modified screen-printed electrodes for H2O2 detection and for biosensors. Talanta 2001, 55, 605–611.
8. Rodriguez, R. R. et al: "Reduced Graphene Oxide Nanostructures by Light: Going Beyond the Diffraction Limit" 2018 J. Phys.: Conf. Ser.1092 012124
9. Al-Hamry, A. et. al.: "Tuning the reduction and conductivity of solution-processed graphene oxide by intense pulsed light", Carbon, Vol. 102, June 2016, Pages 236-244
10. Cheng, K.H., Cheng, C.H., Chao, A.K.A. and Lo, D.K.C., 2015, April. LASER light scribe and microwave annealing reduction of graphene oxide for supercapacitor applications. In 10th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (pp. 380-383). IEEE.
11. M. E. G. Lyons and G. P. Keeley, "The Redox Behaviour of Randomly Dispersed Single walled carbon nanotubes both in the absence and in the presence of adsorbed glucose oxidase," Sensors, vol. 6, no. 12, pp. 1791–1826, 2006.
12. Amira Ghezal et. al.: "Electrochemical detection of nitrite using screen printed graphite electrode modified by ION/rGO", accepted for publication in the 16th International Conference on Systems, Signals and Devices 2019, 21-24 March 2019, Istanbul, Turkey

DOI 10.5162/sensoren2019/P1.8

Photo-Electrochemical Detection of Nitrite Ions on a SPE Functionalized with TiO_2 Nanoparticles for Environmental Monitoring

Mohamed Amin Elaguech[1], Ammar Al-Hamry[4, *], Frank Wendler[4], Ayda Bouhamed[4], Sami Ameur[3], Mounir Ben Ali[1,2], Olfa Kanoun[4]

[1]Electronic department , University of Sousse, Higher Institute of Applied Science and technologys of Sousse

[2]Laboratoire Nanomisene LR16CRMN01, Sousse, Tunisia

[3] Higher Institute of Agricultural Sciences of Chott Mariem, University of Sousse, BP N° 47, 4042 Sousse, Tunisia

[4]Professorship of Measurment and Sensor Technology, Technische Universität Chemnitz, Chemnitz, Germany

*email: ammar.al-hamry@etit.tu-chemnitz.de

Abstract

In this paper, a highly sensitive photo-electrochemical sensor based on a screen printed carbon electrode SPE was developed for nitrite detection. The sensor was functionalized with titanium oxide nanoparticles having excellent photocatalytic properties. The functionalization of the working electrode was investigated by performing scanning electron microscopy imaging and energy-dispersive-X-ray (EDX). Both methods confirmed the successful modification of the electrodes. In addition, an optimization of pH value of the phosphate buffered saline (PBS) solution was performed as it can influence the surface charge properties of the sensor. After the optimization, the sensor performance was characterized by means of cyclic voltammetry (CV) and square wave voltammetry (SWV) under illumination of an UV light source at a wavelength of 254 nm. The results show that in CV, the oxidation peaks have been enhanced around 20% by UV irradiation. This is the effect of the creation of electron/hole pairs in the structure of the TiO_2. The Square Wave Voltammetry measurements show two linear regions: The first in the range of 30 – 100 µmol/L and the second in the range of 100 – 1000 µmol/L and a limit of detection of 0.16 µmol/L.

Keywords- Electrochemical sensors; Nitrite detection; Photocatalytic effect; Titanium oxide nanoparticles; Screen printed electrodes (SPE)

1 Introduction

Recently due to excessive use of fertilizer, more nitrogen compounds are converted to nitrate and hence go to soil to be absorbed by plants or to tickle away to ground water. A high amount of nitrate in drinking and ground water can cause problems on humans and on the ecosystem in general. Under the influence of certain intestinal bacteria, nitrate is developed into nitrite in the body. Nitrite oxidizes the hemoglobin and prevents the oxygen supply to the cells, which can end up deadly especially for infants (blue baby syndrome) by internal suffocation [1]. For those reasons, the World Health Organisation (WHO) has fixed the values of nitrate and nitrite concentrations in drinking water respectively 50 mg/l and 3 mg/l [2]. Unfortunately, the concentration of nitrate exceeded this limit in several countries [3].

Therefore, there is an urgent need for highly sensitive and accurate quantitative determination of nitrite. Nitrite detection techniques can be divided into several categories, such as spectroscopic [**Fehler! Verweisquelle konnte nicht gefunden werden.**], optical [5] and catalytic methods [6]. These methods have shown good results, but they still hard to process and require heavy and expensive equipment. To avoid these limitations, electrochemical methods are preferred due to the high performance in terms of sensitivity, selectivity, fast response in addition to the possibility of in-situ detection of nitrite ions. Among electrochemical techniques, photo-electrochemical method presents a promising pathway for a sensitive and selective detection of nitrite, which combines the properties of the electrochemical and photocatalytic methods [7]. The most widely studied semiconductor photo catalyst for the photo-electrochemical detection is the titanium oxide TiO_2, which is well-known for its excellent photo-catalytic properties [8].

Therefore, in this work, a photo-electrochemical sensor based on a screen printed carbon electrodes (SPE) was used. The SPE is formed with three electrodes: working electrode (WE), counter electrode (CE) and reference electrode (RE). The WE of screen-printed electrodes was modified by drop casting of a TiO_2-sensitive layer over its surface. The functionalized layer was examined via scanning electron microscopy imaging and energy-dispersive-X-ray. Then, the functionalized electrode was tested to detect nitrite under UV irradiation of a UV lamp with a wavelength of 254 nm.

2 Material and methods

Standard nitrite solutions were prepared with sodium nitrite. Ferri/ferrocyanide redox probe solution was prepared with potassium ferrocyanide and potassium ferricyanide salts. Phosphate buffer solutions (PBS) were prepared with sodium phosphate (dibasic heptahydrate and monobasic monohydrate). All chemicals were purchased

from SigmaAldrich and used without additional purification.

A commercially available carbon screen printed electrodes (purchased from PalmSens BV, Houten, Netherlands) were used and functionalized with Titanium oxide P25 (TiO2 P25) purchased from Sigma Aldrich. In order to control the quality of the functionalized layer, scanning electron microscopy imaging and energy-dispersive-X-ray were done using Nova NanoSEM 200 microscope.

In this work, all the electrochemical measurements were carried out at room temperature using a potentiostat PalmSens 3. Cyclic voltammetry (CV) measurements were performed for the functionalized electrodes in different pH of PBS solutions prepared by mixing sodium phosphate monobasic and sodium phosphate dibasic, which were purchased from Sigma Aldrich. The measurements were carried out in solution containing 10^{-3} mol of nitrite under UV light source using an ultraviolet lamp Pen-Ray Standard UV with a wavelength of 254 nm.

3 Results and Discussion

3.1 Electrode surface characterization by SEM and EDX

The investigation of the SPE surface morphology of the bare and the functionalized electrode were carried out using the scanning electron microscope (SEM). As shown in figure 1, the SEM image of the bare carbon electrode surface reveals the presence of irregular and randomly oriented microscopic defects, which can increase the effective surface of the electrode and increase the number of possible active sites to capture the nitrite ions.

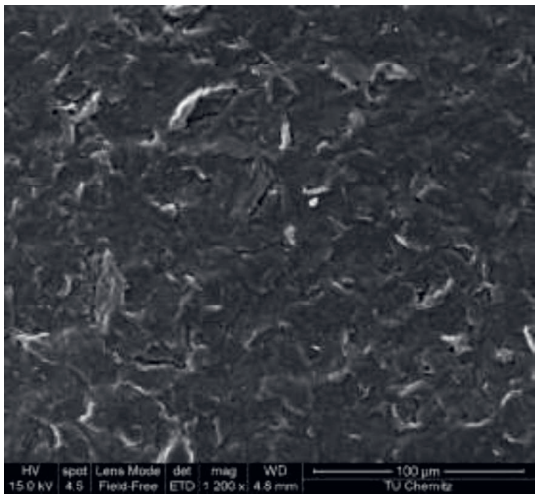


Figure 1: SEM image of the bare carbon electrode.

After functionalization with titanium oxide nanoparticles, a white layer has appeared on the surface of the working electrode as shown in figure 2.

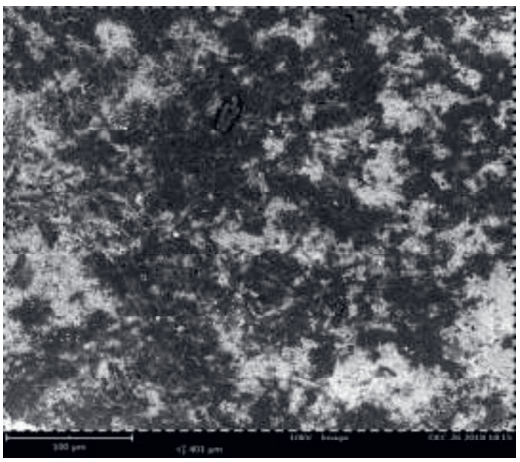


Figure 2:SEM image of the functionalized electrode

The evaluation of the chemical composition of the electrode surface was carried out using EDX technique. The EDX analysis of the bare electrode and TiO2 functionalied electrodes is presented in table 1.

Table 1: Atomic composition of the working electrode material in % concentration

Element	Bare electrode	TiO_2-functionlaized electrode
C	93.31	90.3
O	6.69	6.3
Ag	-	1.03
Ti	-	2.30

3.2 Electrochemical characterization of the electrodes

In this study, the electrochemical behaviour of the SPE and the kinetic transfer were investigated. The SPE was tested by cyclic voltammetry (CV) and electrochemical impedance spectroscopy (EIS) in presence of the redox pair ferri/ ferrocyanide $[Fe(CN)_6]^{3-/4}$ containing 0.1 M KCl as supporting electrolyte. The comparison between the cyclic voltammograms of the bare and the functionalized electrode shows a significant increase of the oxidation and reduction current peaks for the functionalized electrode as depicted in figure 3. This enhancement is related to the increases of the electro-active surface area of the electrode after the functionalization with the TiO_2 NPs.

In order to check the system reversibility, the difference between the oxidation and the reduction potential peaks ΔE have been calculated, $\Delta E = E_{pa} - E_{pc} = 250$ mV indicating that the system is quasi-reversible [9].

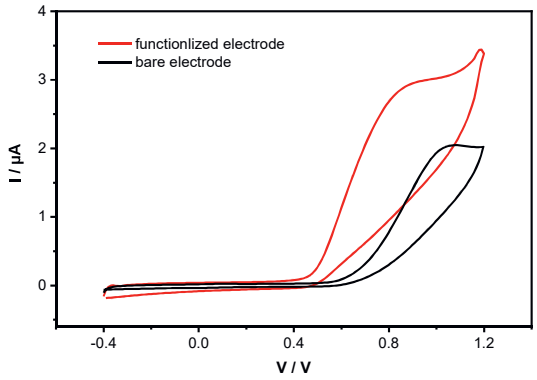


Figure 3: CVs of bare and functionlized electrodes in the presence of the redox pair ferri/ferrocyanide

The increase of the oxidation-reduction current peaks can be also referred to the reduction of the charge transfer resistance between the electrolyte and the electrode R_{ct} [10]. The reduction of R_{ct} value can be seen clearly through the EIS measurements in figure 4, where it is clear that the diameter of the semi-circle has been decreased after the functionalization with TiO_2 nanoparticles.

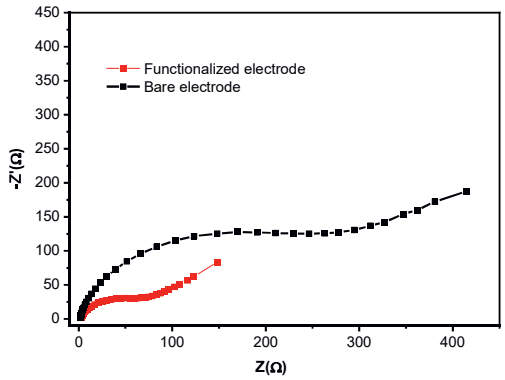


Figure 4: EIS of the bare and the functionalized electrode.

In order to get further information about the reasons behind the reduction of the charge transfer resistance and to estimate the electro-active surface of the electrode before and after functionalization, the Randles-Sevcik equation (equation 1) was applied on the bare and the functionalized electrode [11].

$$I_p = 268600\, n^{3/2}\, A.D^{1/2}\, C.v^{1/2} \qquad (1)$$

Where I_p is the oxidation current in Amperes, n is the number of electrons transferred during the redox event, D is the standard diffusion coefficient in cm²/s, C is the concentration of the redox pair in mol/cm^3, v is the scan rate in V/S and A is the electro-active surface area in cm².

The calculated electro-active surface areas show that before functionalization, the area was 0.0027 cm² and this number has been increased around two times after the functionalization and it becomes 0.0055cm².

3.3 Nitrite detection

a) Effect of the functionalization with TiO_2 NPs

In order to investigate the effect of the functionalization of the electrode with TiO_2 NPs, cyclic voltammetry measurements were run on the bare and the functionalized electrode in a PBS solution containing a nitrite concentration of 10^{-3} M. Figure 5 presents the two different CVs, it is clear from the voltammograms that the oxidation peak current has significantly increased after the functionalization of the electrode. We can refer this increase of the oxidation peak to the higher electro-active surface area of the functionalization electrode with TiO_2 NPs in comparison to the bare electrode [12].

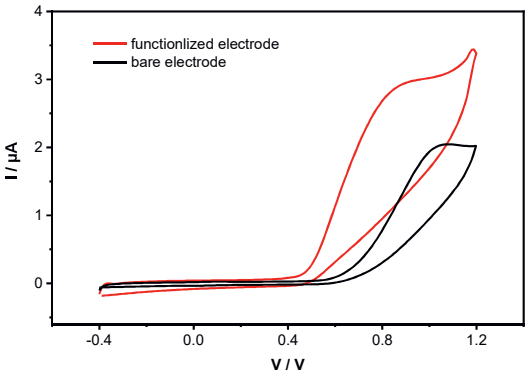


Figure 5: CVs of the bare and functionalized electrodes in 10^{-3} M of $NaNO_2$

b) Effect of pH and scan rate

The photo-catalytic reactions strongly depend on the pH value of the reaction solution because the surface charge properties of the semi-conductor can be influenced by the pH value. In acidic medium, TiO_2 surface becomes positively charged with the presence of H^+ ions ($TiOH_2^+$) which enhance the adsorption of nitrite ions (NO_2^-) to the electro-active sites of the photo-catalyst. Whereas in a basic medium the surface of the titanium oxide is negatively charged (TiO^-). In consequence, due to electrostatic repulsions, the adsorption of nitrite ions is unfavourable [13]. The effect of pH medium on the oxidation reaction was investigated by testing the sensor response in presence of 10^{-3} M of nitrite and at different pH (4.5, 5.7, 6.5 and 7.5). According to figure.6a, it is clear from the voltammograms that the suitable pH value that deliver the highest oxidation current is 5.7. To investigate the nature of the reaction, CV responses were recorded at different scan rates (0.3 – 0.9 V/S) on the functionalized electrode in a PBS solution pH =5.7 in presence of 10^{-3} M of nitrite. The CV responses in figure 6b show a proportional increase of the oxidation current in correlation to the increase of the scan rate. The insert in figure 7 presents the linear fitting of the oxidation current as a function of the square root of the scan rate, which shows a linear rela-

tionship between them, with a correlation factor R^2 = 0.997. As a conclusion, the electrode exhibited a surface controlled process [10].

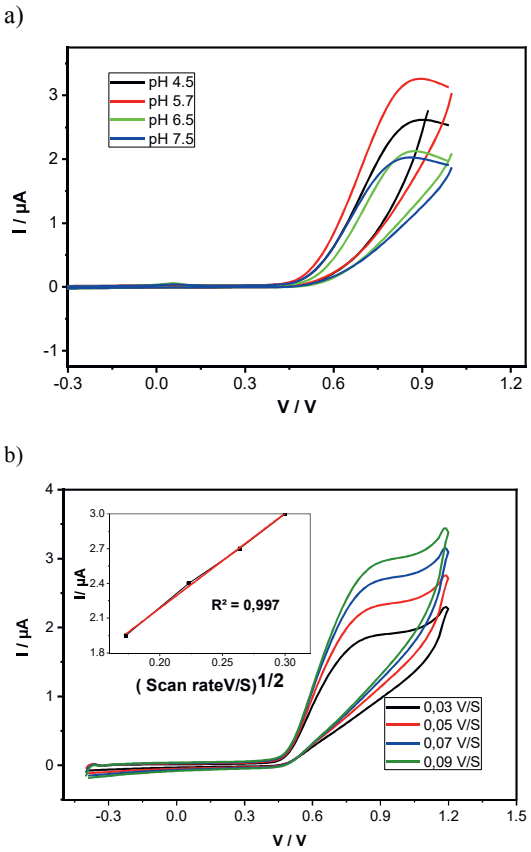


Figure 6 a) Response of the sensor in different pH values and b) CVs of 10-3 M of nitrite in different scan rate. Insert linear fitting of the oxidation current as a function of the square root of the scan rate.

c) Influence of UV irradiation on nitrite detection

In this part, the effect of the ultraviolet irradiation on the nitrite oxidation current peak was addressed. CV responses of the sensor were measured with three different nitrite concentrations (5.10^{-3}, 6.10^{-3}, 7.10^{-3}) in PBS solution pH=5.7 with and without UV-irradiation. The irradiation of the TiO_2 surface with a UV light source 254 nm brings a sufficient energy to excite an electron from the valance band VB to the conductive band CB of the TiO_2, which create a hole in the valance band and an electron in the conductive band of the TiO_2. Photo-catalytic reactions are initiated by the photo-generated electrons (e^-) and holes (h^+) when the TiO2 photo-catalyst is illuminated. The hole reacts with the adsorbed water or hydroxyl groups (OH^-) on the surface of the photo-catalyst to form hydroxyl radicals (OH•), which are, in general, accepted to be the reason for the initialization of the oxidation pathway [14]. As presented in figure 7, the UV irradiation has enhanced the oxidation peak around 20%.

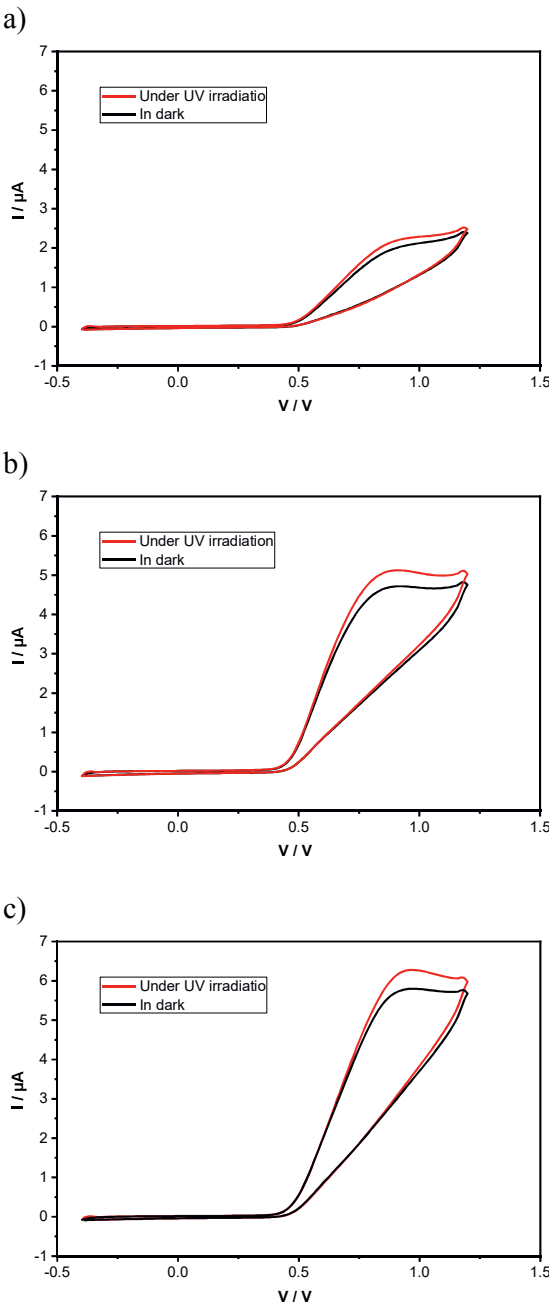


Figure 7 CVs response of different nitrite concentrations a) 5.10^{-3}, b) 6.10^{-3} and c) 7.10^{-3} mM

d) Nitrite detection using square wave voltammetry (SWV) technique

The response of the sensor upon different nitrite concentration was recorded using the square wave voltammetry technique in a PBS solution pH=5,7 under ultraviolet irradiation. Figure 8 presents the recorded SWV of different concentration of nitrite.

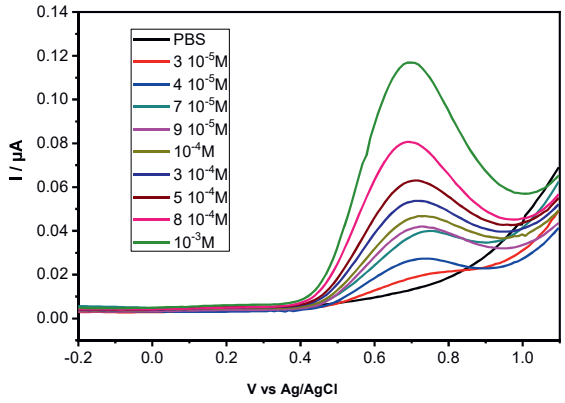


Figure 7: SWV recorded upon different concentration of nitrite

Figure 9 presents the linear fitting of the oxidation peak currents as a function of nitrite concentration, the linear fitting shows that the sensor present two linearity zones, the first for low concentrations 30 – 100 µmol with a correlation factor R^2= 0.99, and the second for the high concentrations 100 – 1000 µmol with a correlation factor R^2= 0.98. The limit of detection of the sensor was calculated using the formula in equation 2 [11]

$$LOD = 3 \times \delta/m \qquad (2)$$

Where δ is the stander deviation and m is the slop of the calibrations curve. This calculation gave a limit of detection LOD around 0.16 µmol.

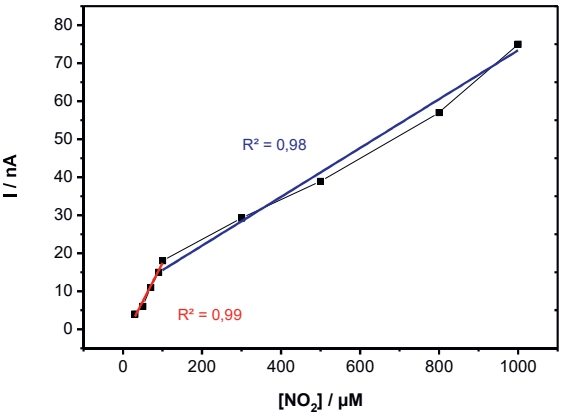


Figure 9 Linear fitting of the oxidation peak as a function of nitrite concentration.

4 Conclusion

In this paper, we have enhanced the performance of photo-electrochemical sensor via functionalization of SPE with TiO_2 NPs and selection of the suitable working conditions. The TiO_2/SPE sensor shows a good electrochemical behaviour in comparison with the bare electrode. The functionalization of the sensor with TiO_2 NPs increases the electro-active surface area and the oxidation current under the UV irradiation. The presented sensor exhibits a LOD 0.16 µmol, which satisfy the requirement of the WHO. Further investigations can be done in the future such as the interference test and a real test on drinking water.

Acknowledgement

This work was carried out with the support of the Nitramon project (project no. 100339427) funded by the Sächsische Aufbaubank (SAB) and the European Social Fund (ESF) and the project Electrochemical Sensors for Environmental Applications (ESEA) (project no. 249305) funded by DAAD. Authors acknowledge the support of Aleksandr Oshchepkov (within the project Nitramon) from Technische Universität Chemnitz, Chemical depart, group of Supramolecular Chemistry for solutions preparation and discussions. A.A.-H. and O. K. acknowledge the financial support within the project InnoTeam-SimplySafe (contr. nr.100331073) funded by the Sächsische Aufbaubank (SAB) and the European Social Fund (ESF).

Literature

1. Y. M. P. V. Yu. Titov, ‘Proposed Mechanism of Nitrite-Induced Methemoglobinemia’, Biochemistry, vol. 70, pp. 473–483, 2005.
2. E. C. 1991, ‘“Council Directive 91/676/EEC Concerning the Protection of Waters against Pollution Caused by Nitrates from Agricultural Sources’, Off. J. Eur. Communities, pp. 1–8, 1991.
3. M. Bejaoui, N. Khalloufi, M. Rihani, M. El, and W. A. Amar, ‘Impact de l’intensification de l’activité agricole et des rejets des eaux usées non traitées sur la qualité physico-chimique et bactériologique des eaux de la nappe " sidi ismail " a beja’, no. March, pp. 1–2, 2016.
4. Q. H. Wang *et al.*, ‘Methods for the detection and determination of nitrite and nitrate: A review’, *Talanta*, vol. 165, pp. 709–720, 2017.
5. X. Li *et al.*, ‘Nitrite optical sensing from core-shell structured nanosensors functionalized with rhodamine molecules’, *J. Lumin.*, vol. 200, pp. 225–232, Aug. 2018.
6. G. E. Dima, A. C. A. De Vooys, and M. T. M. Koper, ‘Electrocatalytic reduction of nitrate at low concentration on coinage and transition-metal electrodes in acid solutions’, *J. Electroanal. Chem.*, vol. 554–555, no. 1, pp. 15–23, 2003.
7. M. Kan, D. Yue, J. Jia, and Y. Zhao, ‘Photoelectrochemical reduction of nitrates with visible light by nanoporous Si photoelectrode’, *Electrochim. Acta*, vol. 177, pp. 366–369, 2015.
8. J. Bai and B. Zhou, ‘Titanium dioxide nanomaterials for sensor applications’, *Chem. Rev.*, vol. 114, no. 19, pp. 10131–10176, 2014.

9. D. A. C. Brownson and C. E. Banks, *The Handbook of Graphene Electrochemistry*. Springer, 2014.
10. M. A. Ali *et al.*, ‘Microfluidic impedimetric sensor for soil nitrate detection using graphene oxide and conductive nanofibers enabled sensing interface’, *Sensors Actuators, B Chem.*, vol. 239, pp. 1289–1299, 2017.
11. N. S. Neghmouche and T. Lanez, ‘Calculation of electrochemical parameters starting from the polarization curves of ferrocene at glassy carbon electrode’, *Int. Lett. Chem. Phys. Astron.*, vol. 4, pp. 37–45, 2013.
12. Y. T. Shi *et al.*, ‘The Enhanced Photo-Electrochemical Detection of Uric Acid on Au Nanoparticles Modified Glassy Carbon Electrode’, *Nanoscale Res. Lett.*, vol. 12, p. 7, 2017.
13. Y. A. Shaban, A. A. El, R. Kh, and A. Farawati, ‘Photocatalytic reduction of nitrate in seawater using C / TiO 2 nanoparticles’, *"Journal Photochem. Photobiol. A Chem.*, vol. 328, pp. 114–121, 2016.
14. T. Tan, D. Beydoun, and R. Amal, ‘Effects of organic hole scavengers on the photocatalytic reduction of selenium anions’, pp. 273–280, 2003.

DOI 10.5162/sensoren2019/P1.9

Sensitivity of Layer-by-Layer Deposited GO/PDAC to Volatile Organic Compounds

Ammar Al-Hamry[1,*], Tatiana Zubkova[2], Leonardo G. Paterno[3], Abdelhamid Errachid[4], Reinhard Baumann[2, 5], Olfa Kanoun[1]
[1]Measurement and Sensor Technology, Technische Universität Chemnitz, Chemnitz, Germany
[2]Print and Media Technology, Technische Universität Chemnitz, Chemnitz, Germany
[3]Instituto de Química, Universidade de Brasília, Brasilia, Brazil
[4]Université de Lyon, Institute of Analytical Sciences, Université de Lyon 1, Lyon, France
[5]Department Printed Functionalities, Fraunhofer ENAS, Chemnitz, Germany
*Tel.-Nr. +49 371 531 39541, E-Mail: ammar.al-hamry@etit.tu-chemnitz.de

Abstract

Films prepared by layer-by-layer deposition of graphene oxide/Poly(diallyldimethylammonium chloride) (GO/PDAC) have been investigated for detection of volatile organic components (VOC), such as acetone, ethanol, isopropanol and methanol. The films were deposited on top of silver interdigitated electrodes fabricated by inkjet-printing on a polyethylene naphthalate (PEN) substrate. The sensitive films were prepared with different bi-layers of GO/PDAC having different conductivities after thermal annealing converting GO to reduced GO. All fabricated sensor elements show a high selectivity towards Methanol compared to other detected VOCs. The resistance changes measured in existence of 2000 ppm concentration of VOCs relative to the resistance measured in ambient air at room conditions is 50% for methanol, 10% for acetone, 9% for ethanol, and 1% for isopropanol. This confirms the sensitivity of GO/PDAC to VOCs, which enables it to be considered as sensitive material to methanol. By further functionalization, the sensitivity to different VOCs can be adjusted, so that selective multi-sensor systems can be realized.

Keywords-component; Chemical sensors, Volatile Organic Compound (VOC), Graphene Oxide (GO), Reduced Graphene Oxide (rGO), Layer-by-Layer deposition, Methanol, Acetone, Ethanol, Isopropanol

1 Introduction

Detection of organic volatile compounds (VOCs) has become a very hot topic as they are a major contributor to bad indoor air quality. They are considered as dangerous environmental pollutants because they can cause health problems. For example, methanol is used as a reagent for chemicals such as cholesterol, in spray paints, in glass cleaners and as a fuel for picnic stoves and soldering torches [1]. Due to its toxicity, it was banned by the European Union from use in windscreen washing and defrosting liquids [2]. For the diagnosis of human diseases, the analysis of VOCs produced by the body, e.g. by breathe monitoring, is an invasive and economic method for early detection of diseases. And it is also an alternative approach for blood analysis [3-4].

VOCs are detected usually by gas chromatography/mass spectrometry techniques (GC-MS). However, due to its time, cost and complexity, rapid, easy and portable approaches are preferred [5]. Several sensor materials were developed to detect VOCs, *e.g.*, Si-doped epsilon-WO_3 nanostructures [6], SnO2-reduced graphene oxide [7] and nanostructured Pt-doped SnO_2 sensors [8] were used to detect acetone for diabetes patients.

Graphene oxide (GO) is used for different sensors, e.g. force, humidity, chemical sensing [9]. The inclusion in both conductive or insulating polymer and hybridization with nanoparticles can be performed in a manifold manner to increase its sensitivity and selectivity [10-13]. Graphene oxide multilayers reduced by hydrazine were already used to detect different VOCs where number of layers was controlled to tune the sensitivity [10]. Hybrid GO: ZrO2 and GO:CeO2 and GO:PEDOT were utilized as an electronic nose to distinguish different VOCs [5].

In this paper, we propose an easy processable and low cost GO based sensors for the detection of VOCs. GO is layered chemically derived graphene containing oxygen functional groups (epoxy, hydroxyl, carboxyl and carbonyl). These functional groups make it an insulator and having a resistance of hundreds of Mega Ohms [9]. The hydrophilic nature and surface properties of GO are altered by the change of its oxygen quantity. Thus the functional groups contained in GO give the sensitivity properties of GO and its derivative, the so-called reduced graphene oxide (rGO). The fabricated sensor is based layer-by-layer deposition and optimal thermal annealing. This process is followed by the characterization of optical and electrical properties. The sensor sensitivity to different VOCs i.e. methanol, acetone, isopropanol and ethanol is investigated and show a basis for detection of VOC especially to methanol based on piezoresistive change of the sensor resistance.

2 Experimental Investigations

2.1 Materials

GO utilized in this work was purchased from Graphenea Inc. with the initial concentration of 4 mg/mL. The GO was diluted in deionized water to 1 mg/mL. Slight sonication is required to prevent agglomeration after dilution. PDAC (20 wt.%) was purchased from Sigma Aldrich (CAS-No. 6062-79-32). All solvents for producing the vapor concentration are acetone, isopropanol, ethanol and methanol and are analytical grad from Sigma Aldrich.

2.2 Sensor Preparation

Silver interdigitated electrodes were initially inkjet printed on polyethylene naphthalate (PEN) substrate and annealed at 150 °C. The line width and finger spacing of the interdigitated electrodes are 250 µm. The total area of the sensor is 16 mm2. Next, prior to layer-by-layer deposition, PEN substrates are plasma cleaned and immediately dipped in a beaker of diluted PDAC in distilled water to 1%. The bi-layer of PDAC/GO is deposited by alternating dip coating the silver electrode/PEN for 5 m in GO (diluted to 1% in distilled water.) and PDAC. Cleaning with HCl diluted solution and drying with nitrogen, between alternating depositions was maintained. Finally, the film was dried with nitrogen. In this work we have opted to form 1, 2, 4, 6, 8 and 10 bilayers of PDAC/GO. After the film deposition the electrode with PDAC/GO bilayers is left overnight. The GO in a bi-layer deposited onto the electrodes was later reduced by thermal treatment. This was done by placing the electrodes on a hotplate in air atmosphere.

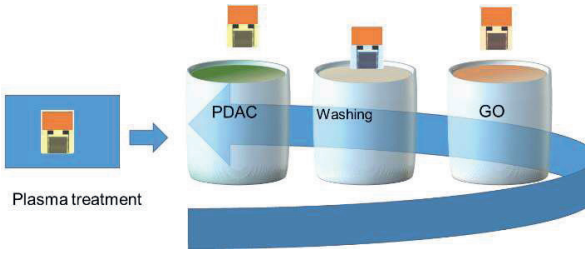


Fig 1 Layer-by-layer deposition of GO-PDAC

2.3 Characterization and Measurement Setup

Raman spectra, excited with diode-pumped solid state high brightness laser (excitation wavelength 532 nm), were collected on a DXR Raman microscope (Thermo Scientific, USA) equipped with an Olympus optical microscope and a CCD detector. The laser beam was focused on the sample using an objective (magnification 10×). The scattered light was analyzed by the spectrograph with a 900 lines mm^{-1} grating. Laser power on the sample was kept at 1 mW to prevent thermal degradation of the samples.

For the measurement of VOCs, experiments were performed at room temperature, and resistance measurement was taken by Keithly 2602 sourcemeter. The sensors were exposed to different gas concentrations. To generate certain concentration of VOCs, amount was injected in a sealed chamber i.e. glass bottle of 2.5 L using Hamilton syringe. The concentration of VOC was calculated according to the flowing equation (1) [12]:

$$V_{inj} = \frac{MPV_S}{R\rho} \cdot \frac{C}{T} \quad (1)$$

Where, M (g/mol) is the molecular weight of the liquid and it is 46, 48, 32.04 and 60.1 g/mol for ethanol, acetone, methanol and isopropanol, respectively. ρ (g/ml) is the density and it is 0.816, 0.788, 0.79 and 0.786 g/mL for ethanol, acetone, methanol and isopropanol, respectively. T (K) is the temperature of the chamber, V_s is the test chamber volume (i.e. 2.5 L), P is atmospheric pressure (101325 Pa), R is ideal gas constant (8.31441 J/(mol·K)) and C is resulted concentration of desired gas by injection of V_{inj} of liquid VOC. The relative response of the resistance was calculated by equation (2):

$$\frac{\Delta R}{R}\% = \frac{R_{gas} - R_{air,initial}}{R_{air,initial}} \quad (2)$$

Where $R_{air,\ initial}$ is the resistance of the sensor measured in air initially before injection of any concentration of VOC and R_{gas} is the resistance measured upon injection of certain gas concentration.

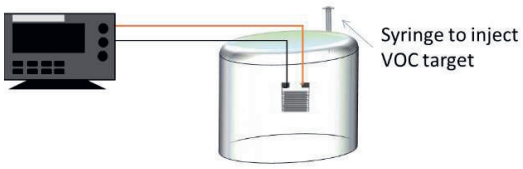


Fig 2 Schematic of gas sensing setup

3 Results and Discussion

The Raman spectra of PDAC/GO layers reduced at different temperatures show typical features of a highly disordered material [14-17] with D (disorder) and G (graphitic) bands, located around 1350 and 1580 cm−1, respectively, being the most pronounced features of the spectra (Figure 2). Also, the overtone bands, i.e. 2D region, are clearly visible. These bands are found at around 2670 cm−1 and 2930 cm−1. Following Ref. [18] the first band is denoted as 2D (G') while the second one is denoted either as D+D' or D+G. These bands are considered to be defect-activated and dependent on the number of layers in graphite/graphene. Moreover, their appearance is also associated with wrinkled graphene layers, which contain significant amount of defects [18]. There are variations of relative intensities of observed bands (certain increase of G-to-D band intensity and decrease of 2D region with increasing temperature), but the overall conclusion is that thermal treatment did not change the disorder in rGO significantly. However, we expect that conductivity of rGO is changed with respect to that of parental GO (which is very poor conductor). This conclusion is in line with previous results

considering conductivity changes of electrochemically reduced GO, showing that significant increase of conductivity is expected upon very mild reduction [19] and color change upon thermal treatment. We note that Raman spectroscopy showed homogeneous distribution of PDAC/rGO over silver electrodes and substrate.

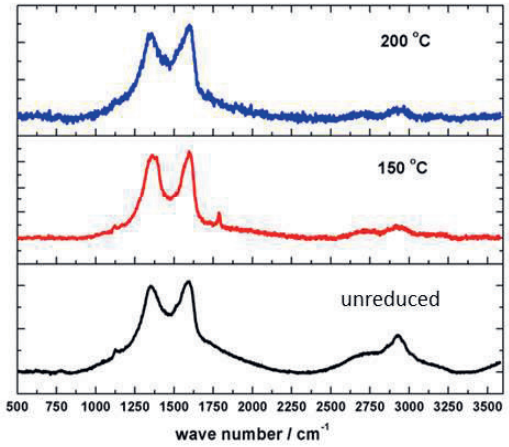


Fig. 3. Raman spectra of PDAC/rGO for unreduced and reduced at 150 and 200 °C

Due to the poor conductivity of PDAC/GO and partially reduced PDAC/rGO treated at low temperature, temperature of 200 °C was used to thermally reduce GO in PDAC/GO films. Therefore the resistance can be measured. In Fig. 4, the relation between electrical resistance and no. of layer is shown which concludes that a successful homogenous layer-by-layer deposition. The resistance was changed by two orders of magnitude to decrease from about 500 kΩ to several kΩ. After 8 layers, it seems that there is no significant change in the resistance in relation with number of layers.

Fig. 5 shows the relative response of resistance of the PDAC/rGO films exposed to various vapors i.e. ethanol, acetone, isopropanol and methanol in a concentration of 2000 ppm. These VOCs can be classified as aprotic polar vapors (acetone) and protic polar vapors (isopropanol, ethanol and methanol). In both cases both has a hydrogen atom bounded to an oxygen atom. Among them only isopropanol shows a negative response which has the lowest polar bond. The sensors show the highest sensitivity to methanol which could be explained that methanol has the highest polar bond. In addition, its smallest molecular size can easily diffuse and interact with the conductive layer of PDAC/GO [20].

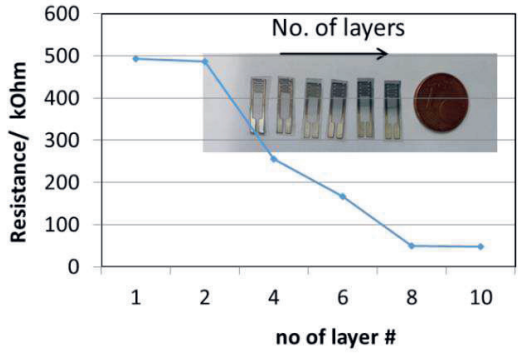


Fig. 4 Resistance change by no. of bi-layers reduced at 200 °C, Inset: photograph of the sensors

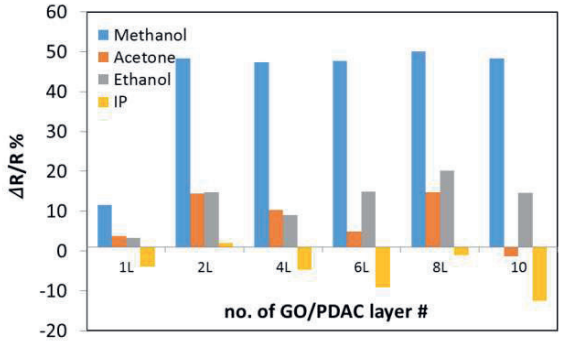


Fig. 5 Measured sensitivity of methanol, acetone, ethanol and isopropanol at 2000 ppm

For the measurements of response to methanol, Fig. 6 shows the relative change of resistance to different concentrations. It is noticed that resistance increases when more concentrations of methanol is added to the chamber. The sensitivity decreases by increasing the multilayer PDAC/rGO. However, for thicker films, the linearity is improved.

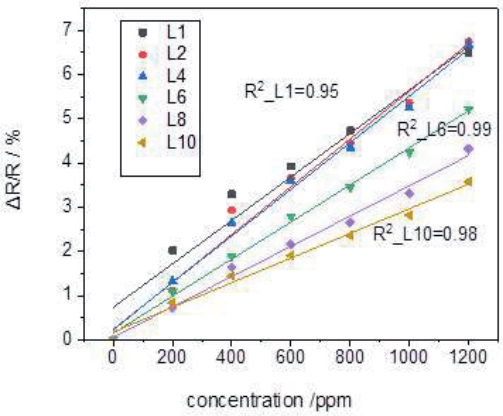


Fig. 6 Relative responses of resistance of the PDAC/rGO at various methanol concentrations

4 Conclusion

A sensor based on PDAC polymer and reduced graphene oxide was fabricated by layer-by-layer technique accompanied by relatively low temperature reduction. Different VOC vapors were used to test the sensing properties of these fabricated sensors. The relative resistance changes were 50% for methanol, 10% for acetone, 9% for ethanol, and 1% for isopropanol. This confirms the sufficient sensitivity of GO/PDAC to VOCs and that functionalization using layer by layer deposition is suitable as it improves response and presents a basis for detection of different VOCs specially methanol by measuring the change in electrical resistance. By further functionalization, the sensitivity to different VOCs can be adjusted, so that selective multi-sensor systems can be realized.

Acknowledgement

Presented results are obtained during the implementation of the projects InnoTeam-SimplySafe (contr. nr.100331073) funded by the Sächsische Aufbaubank (SAB) and "SMART-BIOTECH" (bilateral project Germany-France funded by DAAD). The authors acknowledge prof. Igor A. Pašti and D. B. Bogdanović (Belgrade University) for with Raman measurements. O.K and L.G.P. acknowledge the support of the Humboldt foundation for the support of prof. L.G.P. stay in Chemnitz.

Literature

[1] National Center for Biotechnology Information. PubChem Compound Database; CID=887, https://pubchem.ncbi.nlm.nih.gov/compound/887 (accessed Feb. 13, 2019).

[2] EUR-Lex - 32018R0589 - EN. [Online] Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32018R0589. Accessed on: Feb. 21 2019.

[3] H.-C. Lau, J.-B. Yu, H.-W. Lee, J.-S. Huh, and J.-O. Lim, "Investigation of Exhaled Breath Samples from Patients with Alzheimer's Disease Using Gas Chromatography-Mass Spectrometry and an Exhaled Breath Sensor System", Sensors, vol. 17, no. 8, 2017.

[4] ID. Kim, SJ. Choi, SJ. Kim, JS. Jang , (2015) "Exhaled Breath Sensors". In: Kyung CM. (eds) Smart Sensors for Health and Environment Monitoring. KAIST Research Series. Springer, Dordrecht

[5] M. R. Cavallari et al., "A Hybrid Electronic Nose and Tongue for the Detection of Ketones: Improved Sensor Orthogonality Using Graphene Oxide-Based Detectors," IEEE Sensors J., vol. 17, no. 7, pp. 1971–1980, 2017.

[6] M. Righettoni et al., "Breath acetone monitoring by portable Si:WO3 gas sensors", Analytica chimica acta, vol. 738, pp. 69–75, 2012.

[7] D. Zhang, A. Liu, H. Chang, and B. Xia, "Room-temperature high-performance acetone gas sensor based on hydrothermal synthesized SnO2-reduced graphene oxide hybrid composite", RSC Adv., vol. 5, no. 4, pp. 3016–3022, 2015.

[8] J. van den Broek, A. T. Güntner, and S. E. Pratsinis, "Highly Selective and Rapid Breath Isoprene Sensing Enabled by Activated Alumina Filter", ACS sensors, vol. 3, no. 3, pp. 677–683, 2018.

[9] A. Al-Hamry et al., "Tuning the reduction and conductivity of solution-processed graphene oxide by intense pulsed light," Carbon, vol. 102, pp. 236–244, 2016.

[10] M. A. Gross et al., "Reduced graphene oxide multilayers for gas and liquid phases chemical sensing," RSC Adv., vol. 4, no. 34, p. 17917, 2014.

[11] S. Konwer, A. Begum, S. Bordoloi, and R. Boruah, "Expanded graphene-oxide encapsulated polyaniline composites as sensing material for volatile organic compounds," J Polym Res, vol. 24, no. 3, p. 457, 2017.

[12] X. Li et al., "Reduced graphene oxide/MoS2 hybrid films for room-temperature formaldehyde detection," Materials Letters, vol. 189, pp. 42–45, 2017

[13] X. Li, Y. Chang, and Y. Long, "Influence of Sn doping on ZnO sensing properties for ethanol and acetone," Materials Science and Engineering: C, vol. 32, no. 4, pp. 817–821, 2012.

[14] T. K. Ega et al. " Detection of Dimethoate Pesticide using Layer by Layer Deposition of PDAC/GO on Ag electrode", International Multi-Conference on Systems, Signals and Devices, 16th IEEE International Multi-Conference on Systems, Signals and Devices

[15] A. T. Lawal, Talanta, vol. 131, pp. 424–443, 2015

[16] R. Beams, L. G. Cançado and L. Novotny, J. Phys.: Condens. Matter, 2015, 27, 083002

[17] M. M. Lucchese, F. Stavale, E. H. M. Ferreira, C. Vilani, M. V. O. Moutinho, R. B. Capaz, C. A. Achete and A. Jorio, Carbon, 2010, 48, 1592–1597

[18] A. Kaniyoor and S. Ramaprabhu, AIP Adv., 2012, 2, 032183

[19] S. J. Gutić et al., Physical chemistry chemical physics: PCCP, vol. 20, no. 35, pp. 22698–22709, 2018.

[20] K. Phasuksom, W. Prissanaroon-Ouajai, and A. Sirivat, "Electrical conductivity response of methanol sensor based on conductive polyindole," Sensors and Actuators B: Chemical, vol. 262, pp. 1013–1023, 2018.

DOI 10.5162/sensoren2019/P1.10

HTCC-Multilayerbasierte Ionensensitive Sensoren für die Wasseranalyse

Dr. Claudia Feller, Dr. Uwe Partsch, Dr. Stefan Körner
Fraunhofer-Institut für Keramische Technologien und Systeme IKTS
Abteilung Hybride Mikrosysteme
Winterbergstraße 28, 01277 Dresden, Deutschland

Zusammenfassung

Auf Basis eines mikrosystemtechnischen Ansatzes wurden elektrochemische Sensoren für die Wasseranalyse entwickelt. Zur Erfassung der verschiedenen Messgrößen sind spezielle Dickschichtpasten entwickelt worden, die für die Erzeugung sensitiver Funktionsschichten auf einem HTCC-Multilayer siebgedruckt werden. Je nach Anwendungsgebiet sind die sensitiven Schichten auf dem Sensorarray beliebig austauschbar. Für die Erfassung der Messgrößen pH-Wert, Cu^{2+}-Schwermetallion sowie der Kationen Ca^{2+}, NH_4^+ und K^+ in wässrigen Lösungen erfolgte eine Material- und Prozessentwicklung, um ionensensitive Dickschichtmembranen auf der keramischen Sensorplattform abzuscheiden. Alle ISEs besitzen im untersuchten Messbereich ein sehr gutes Ansprechverhalten und hohe Empfindlichkeiten. Die keramische Mehrlagentechnologie stellt damit ein geeignetes Verfahren da, um robuste und miniaturisierte Multisensoren für die Wasseranalyse herzustellen.

Keywords: Ionensensitive Elektroden, keramische Multilayertechnologie, Dickschichttechnik

Einleitung

Gewässer sind als Trinkwasserquelle und natürlicher Lebensraum, als Einleitungsort verschiedener Abfallprodukte von Mensch und Industrie in besonderem Maße schutzbedürftig. Die Überwachung ihrer Wasserqualität ist für den Schutz von Umwelt und Gesundheit unerlässlich. In den Bereichen Trinkwasser- und Abwasserüberwachung, intelligente Landwirtschaft, Lebensmittelindustrie usw. sind die Bestimmung unterschiedlicher Ionenkonzentrationen wie z. B. pH-Wert, Schwermetallionen und Pflanzennährstoffe NH_4^+, Ca^{2+}, K^+ sowie Leitfähigkeitsmessungen erforderlich.
Wissenschaft und Industrie stellen zunehmend höhere Anforderungen an die Sensoren bezüglich ihrer Messeigenschaften, Robustheit, Miniaturisierung, Geometrievariabilität und Herstellungskosten. Dabei zielen die Anforderungen von Anwendern an die Sensoren insbesondere auf Funktionsintegration und smarte, billige, miniaturisierte Multisensoren für Massenanwendungen. Nach dem Stand der Technik werden meistens potentiometrische Sensoren für Ionenkonzentrationsmessungen, d. h. sogenannte ionenselektive Elektroden (ISE), verwendet. Diese ISEs basieren auf verschiedenen, meist feinmechanischen Fertigungstechnologien. Ihre Herstellung ist daher recht teuer und das Miniaturisierungspotenzial begrenzt. Ziel war es daher, diese Einschränkungen zu überwinden und Technologien der keramischen Mehrlagentechnik einzusetzen, um kostengünstige und miniaturisierte Multisensoren für die Wasseranalyse herzustellen. Komponenten aus keramischer Mehrlagentechnologie, beheimatet in der Elektrotechnik, haben bereits breite Anwendungsgebiete in verschiedenen Bereichen der Elektronik und Mikrosystemtechnik erobert. Unter einem keramischen Multilayer versteht man ein intelligentes keramisches Packaging, welches durch 3D-Strukturierung und Materialintegration hergestellt wird. Durch die ausgezeichneten dreidimensionalen Strukturierungs- und Funktionalisierungsmöglichkeiten der Keramik sind robuste Sensorlösungen für die 3D-Aufbau- und Verbindungstechnik sensorischer Mikrosysteme realisierbar. Mit der Sensorminiaturisierung bietet die keramische Mehrlagentechnologie ausserdem einen wirtschaftichen Aspekt der Sensorfertigung. Das Ziel der Entwicklungsarbeit war es daher, unter Anwendung der **H**igh **T**emperature **C**ofired **C**eramics (HTCC)-Multilayertechnolgie eine Keramik basierte Integrationsplattform für verschiedene ionensensitive Sensoren zu realisieren, die in zahlreichen analytischen und industriellen Verfahren zur Wasseranalyse einsetzbar ist.

Materialien und Methoden

HTCC-Multilayertechnologie

Die Herstellung keramischer Multilayer basiert auf einer charakteristischen Technologieabfolge. Am Beginn der Prozesskette steht die Herstellung der keramischen Folien durch das Foliengießen. Die noch ungebrannten Folien werden anschließend mit Mikrobearbeitungsverfahren wie Stanzen, Lasern oder Prägen strukturiert. Nach dem Einbringen und Füllen der Durchkontaktierungen erfolgt der Siebdruck von zum Beispiel Leiterbahnen, Widerständen, kapazitiven und induktiven Schichten und von sensorischen Strukturen. Im Anschluss werden die einzelnen Lagen gestapelt und bei erhöhter Temperatur unter Druck laminiert. Während des folgenden Sinterprozesses werden die organischen Bestandteile ausgebrannt und der Grünkörper bei Temperaturen über 1000 °C zu einer porenfreien monolithischen Keramik verdichtet. Zuletzt können im Postprozess weitere Funktionselemente in Dickschichttechnik aufgedruckt oder in SMT-Technologie bestückt werden. Da die Fertigung im Nutzen erfolgt, müssen die einzelnen Elemente durch Lasern oder Wafersägen vereinzelt werden. Für die Herstellung der keramischen HTCC-Multilayerbasierten Sensorplattform (Abb. 1) wird ein bleifreies Hochtemperatur-PSZ-Tape der Firma ESL Electro-Science verwendet.

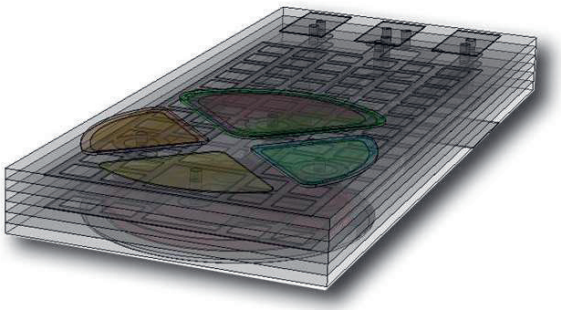


Abb. 1: HTCC-Multilayerbasierte Sensorplattform

Die inneren Metallisierungsschichten und Durchkontaktierungen werden mit Platinpasten der Firma ESL Electro-Science siebgedruckt. Im Postprozess werden die Platinmetallisierungen im Fall der ISEs auf dem Toplayer vergoldet bzw. im Fall der Referenzelektrode versilbert. Für die Kontaktpads auf den äußeren Layern wird eine AgPd-Paste ebenfalls der Firma ESL Electro-Science verwendet. Die Funktionsschichten der All-solid-state-Sensoren werden nachfolgend auf die Toplayer des Multilayers mittels Siebdruck abgeschieden und anschließend gesintert bzw. gehärtet.

Sensormaterialien

Für die pH-sensitive Glasschicht wird ein SiO_2-Na_2O-MgO-MnO_2-Glas mit Sinterende bei 640 °C verwendet. Das Glaspulver wird in einem organischen Binder, bestehend aus einem Lösungsmittel und einem Polymer, dispergiert. Die Glaspaste wird nachfolgend auf einem Dreiwalzwerk weiterverarbeitet, um eine homogene, siebdruckfähige Glassuspension zu erhalten.

Kupferhaltige Chalkogenidgläser dienen als sensitives Elektrodenmaterial für Cu^{2+}-ISE. Basierend auf Untersuchungen der Glasbildung im ternären System Cu-As-Se sowie im quarternären System Cu-Ag-As-Se wurden kupferhaltige Chalkogenidgläser CuAsSe und CuAgAsSe aus den Elementsubstanzen synthetisiert. Dabei wird das Pulvergemenge der Elementsubstanzen in geschlossenen, evakuierten Quarzglasampullen bei 600 °C geschmolzen. Untersuchungen zum Sinterverhalten der Chalkogenidgläser zeigen, dass die Gläser bei Temperaturen jenseits von 200 °C ohne das Auftreten von Sinterblockaden dicht sintern. Das Bulkmaterial wird zu feinem Glaspulver gemahlen, in einen Binder dispergiert und auf dem Dreiwalzwerk zu einer druckfähigen Suspen-sion verarbeitet.

Die ionensensitiven Membranen der Ca^{2+}-, K^+- und $NH4^+$-ISEs bestehen aus einer polymeren Stützmatrix (Silikonkautschuk), in die Weichmacher (DOS), Ionophor (Ca^{2+}: ETH129, K^+: Valinomycin, $NH4^+$: Nonactin) und Additiv (KTpClPB) eingearbeitet werden. Die Suspensionen werden mittels Handsiebdruck abgeschieden und bei Raumtemperatur gehärtet. Zwischen Goldableitung und ionensensitiver Membran wird ein elektrisch leitfähiges Polymer (PEDOT/PSS) als zweite Ableitung siebgedruckt.

Die Referenzelektrode ist eine Elektrode 2. Art und basiert auf dem System Ag/AgCl/NaCl, wobei das Chloridreservoir aus einem NaCl-Glas-Gemenge besteht. Dafür wurde ein SiO_2-B_2O_3-Glas mit einem Sinterende von 410 °C ausgewählt. Unter dem Chloridreservoir wird eine AgCl-Schicht siebgedruckt und bei 400 °C gesintert. Das Chloridreservoir wird zum Schutz gegen Auslaugung mit einer Polymerschicht abgedeckt, welche zur Realisierung des Kontakts mit der Messlösung mit einer Öffnung versehen wird. Die Herstellung der NaCl-Glas- bzw. AgCl-Paste erfolgt ebenfalls nach der für die pH-Glaspaste beschriebenen Verfahrensweise.

Elektrochemische Charakterisierung

Die ISEs sind potentiometrisch arbeitende Indikatorelektroden, die auf die Aktivität der zu bestimmenden Ionensorte ansprechen. Dabei

nutzt die Potentiometrie als ein elektroanalytisches Verfahren der quantitativen Analyse die Konzentrationsabhängigkeit des elektrochemischen Potentials. Die potentiometrische Messung zur Charakterisierung der Messeigenschaften des pH-Sensors wurde in den Standardpufferlösungen 1,68 pH, 4,01 pH, 6,86 pH und 9,18 pH durchgeführt. Die Bestimmung des Ansprechverhaltens und der Empfindlichkeit der ISEs erfolgte in Kalibrierlösungen im Messbereich von 10^{-1} M bis 10^{-4} M Cu^{2+}, Ca^{2+}, K^{+} und $NH4^{+}$. Die Potentiale wurden mit einem Elektrometer der Firma Keithley bei 25°C gemessen.

Ergebnisse und Diskussion

pH-Sensor

Vor den potentiometrischen Messungen wird der pH-Sensor (Layout 1, Abb. 2) mindestens zwei Stunden gewässert.

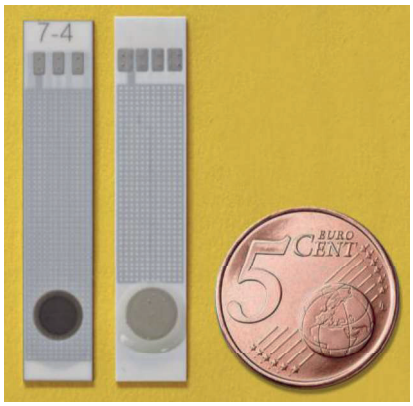


Abb. 2: Halbzellen pH-Elektrode (links, top) bzw. Referenzelektrode (rechts, buttom) des pH-Sensors (Layout 1)

Der pH-Sensor besitzt ein sehr gutes NERNST-Verhalten im pH-Bereich von 1,69 bis 9,18. Abbildungen 2 und 3 zeigen das Ansprechverhalten und die Kennlinie. Die Empfindlichkeit beträgt mindestens 90 % des Nernst-Wertes 59,2 mV/pH.

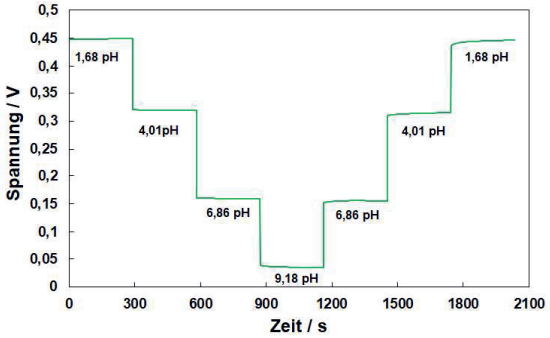


Abb. 3: Ansprechverhalten des pH-Sensors

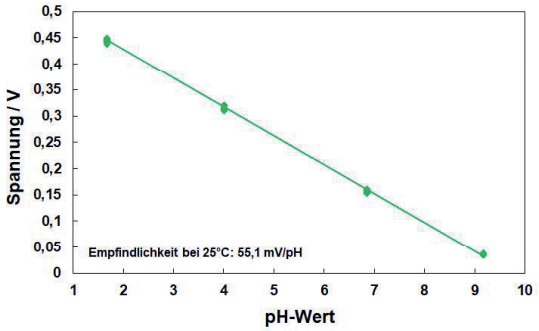


Abb.4: Kennlinie des pH-Sensors

Desweiteren wurde im Zeitraum von 12 Wochen die Potentialstabilität in den Standardpufferlösungen 1,68 pH, 4,01 pH, 6,86 pH und 9,18 pH untersucht (Abb. 5). Nach 15 Tagen ,,Voralterung" weist der Sensor in allen Lösungen relativ stabile Potentiale auf.

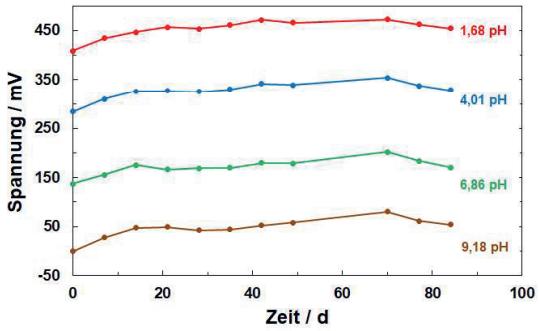


Abb.5: Potentialstabilität in Standardpufferlösungen 1,68 pH, 4,01 pH, 6,86 pH und 9,18 pH

Ausgehend vom Layout 1 wurde der Sensor nochmals miniaturisiert (Layout 2, Abb. 6).

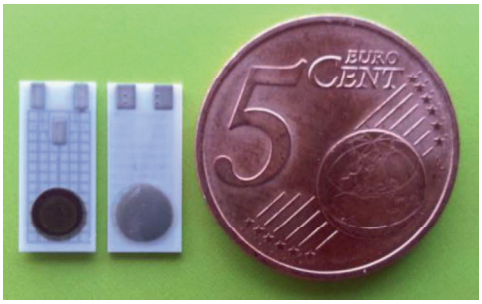


Abb. 6: Halbzellen pH-Elektrode (links, top) bzw. Referenzelektrode (rechts, buttom) des Mini-pH-Sensors (Layout 2)

Auch der Mini-pH-Sensor besitzt mit > 90% des Nernst-Faktors eine sehr gute pH-Empfindlichkeit.

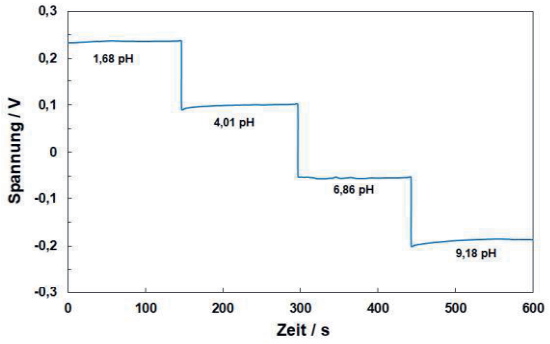


Abb. 7: Ansprechverhalten des Mini-pH-Sensors

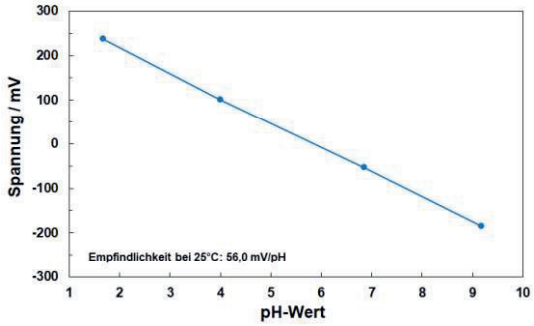


Abb.8: Kennlinie des Mini-pH-Sensors

Cu^{2+}-ISE

Entsprechend dem thermochemischen, Kristallisations- und Sinterverhalten der Kupferchalkogenidgläser wurden die aus den Gläsern hergestellten Dickschichtpasten so prozessiert, dass dünne Glasmembranen auf dem HTCC-Multilayer abgeschieden worden sind, welche kupferionensensitiv sind. Abbildungen 9 und 10 zeigen die Messeigenschaften von auf CuAsSe- bzw. CuAgAsSe-Glas basierten Cu^{2+}-ISE im Konzentrationsbereich 10^{-1} M bis 10^{-4} M Cu^{2+}. Unabhängig von der Zusammensetzung des Chalkogenidglases besitzen beide Glasvarianten eine sehr gute Messperformance.

Im Konzentrationsbereich 10^{-1} M bis 10^{-8} M Cu^{2+} liegt die Empfindlichkeit einer CuAgAsSe-Dickschichtelektrode im Messbereich von 10^{-2} M bis 10^{-6} M $Cu(NO_3)_2$-Lösung bei 27,3 mV/pCu^{2+}. Dieser Wert entspricht dem Nernstschen Verhalten für zweiwertige Ionen und korreliert sehr gut mit der Sensitivität des Bulkmaterials (Abb. 11).

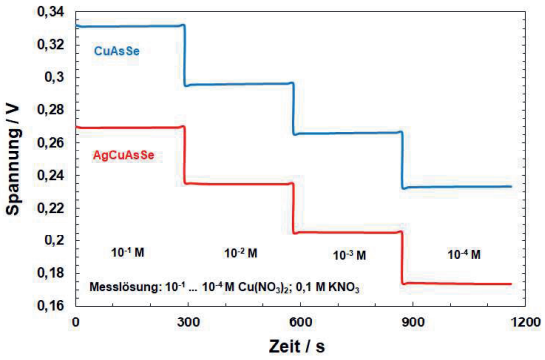


Abb. 9: Ansprechverhalten der Cu^{2+}-ISE

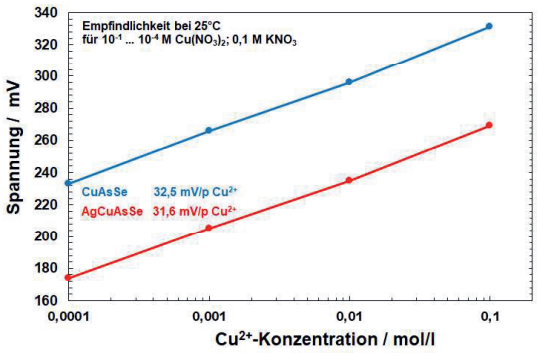


Abb.10: Kennlinien der Cu^{2+}-ISE

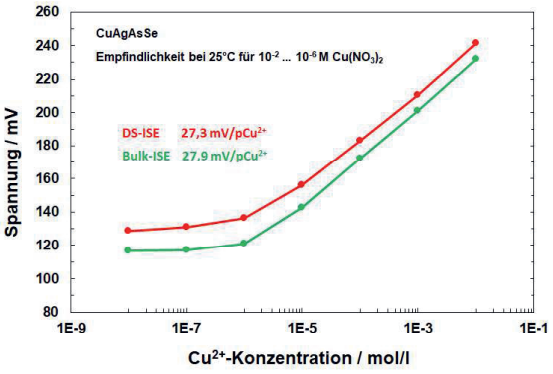


Abb. 11: Kennlinie der Cu^{2+}-ISE (CuAgAsSe) im Konzentrationsbereich 10^{-2}…10^{-8} M $Cu(NO_3)_2$

Ca^{2+}-ISE

Die Ca^{2+}-ISE wurde elektrochemisch auf ihre Funktionsfähigkeit in Ca^{2+}-haltigen Lösungen unterschiedlicher Konzentrationen bei 25 °C überprüft. Die Elektrode zeigt gegenüber Calciumionen ein sensitives Verhalten.

In den Abbildungen 12 und 13 sind das Ansprechverhalten und die Elektrodenkennlinie dargestellt. Als Referenz sind die Messwerte einer konventionellen Ca^{2+}-ISE-Stabelektrode vergleichend abgebildet. Die Empfindlichkeit der Ca^{2+}-ISE-Dickschichtelektrode liegt im

Messbereich von 10^{-1} M bis 10^{-4} M $Ca(NO_3)_2$-Lösung bei 30,0 mV/pCa^{2+}. Dieser Wert entspricht dem Nernstschen Verhalten für zweiwertige Ionen.

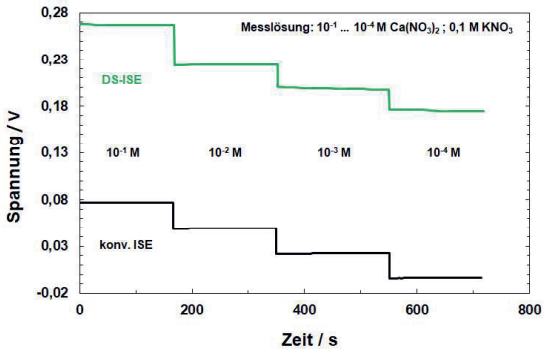


Abb. 12: Ansprechverhalten der Ca^{2+}-ISE

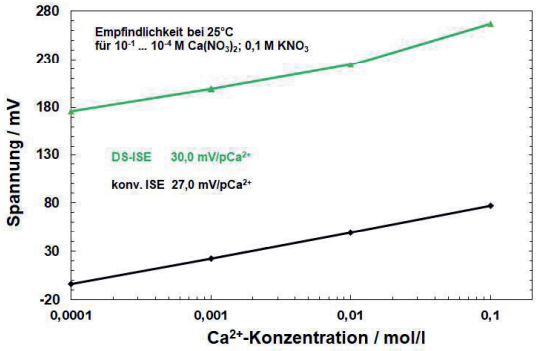


Abb.13: Kennlinie der Ca^{2+}-ISE

K^+-ISE

Die Abbildungen 14 und 15 zeigen das Ansprechverhalten und die Elektrodenkennlinie bei 25°C der K^+-ISE im Vergleich zu einer konventionellen K^+-ISE-Stabelektrode im Messbereich 10^{-1} M bis 10^{-4} M K^+-haltiger Lösungen. Die Empfindlichkeit der K^+-ISE-Dickschichtelektrode liegt im Messbereich von 10^{-1} M bis 10^{-4} M KNO_3-Lösung bei 53,5 mV/pK^+ und entspricht damit dem Nernstschen Verhalten für einwertige Ionen.

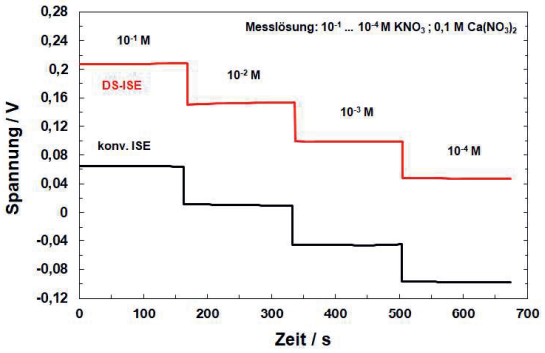


Abb. 14: Ansprechverhalten der K^+-ISE

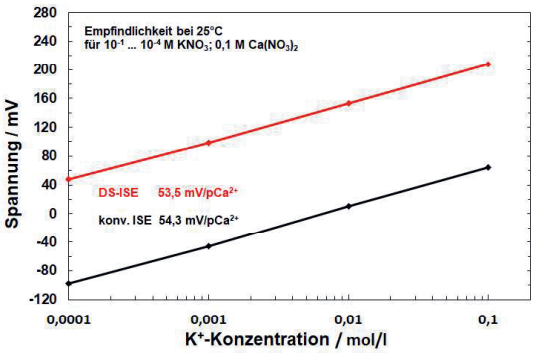


Abb.15: Kennlinie der K^+-ISE

NH_4^+-ISE

Die NH_4-Ionensensitive Elektrode wurde elektrochemisch auf ihre Funktionsfähigkeit in NH_4^+-haltigen Lösungen im Konzentrationsbereich 10^{-1} M bis 10^{-4} M bei 25°C überprüft. In Abbildung 16 bzw. 17 ist das Ansprechverhalten bzw. die Elektrodenkennlinie dargestellt. Zum Vergleich sind die Messwerte einer konventionellen NH_4^+-ISE-Stabelektrode abgebildet. Die Empfindlichkeit der NH_4^+-ISE-Dickschichtelektrode liegt im untersuchten Messbereich mit 57,0 mV/pNH_4^+ bei 96% Nernst.

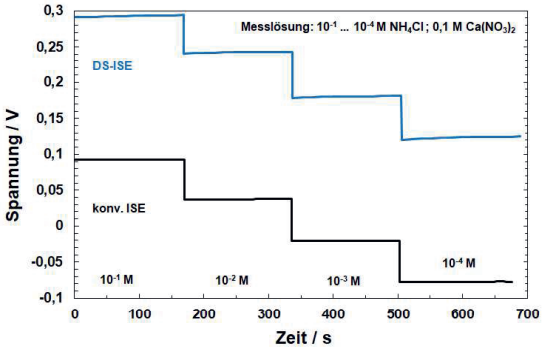


Abb. 16: Ansprechverhalten der NH_4^+-ISE

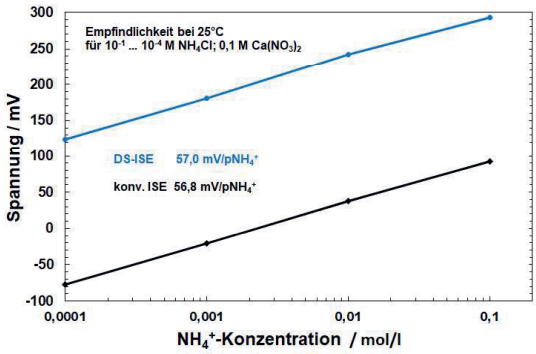


Abb.17: Kennlinie der NH_4^+-ISE

Zusammenfassung

Unter Anwendung der keramischen HTCC-Folien- und Multilayertechnolgie ist eine Plattform für die Integration elektrochemischer Sensoren zur Wasseranalyse entwickelt worden. Je nach Ausführung (Abbildung 18) können Sensorarrays unterschiedlichen Layouts gefertigt werden. Die einzelnen sensitiven Schichten werden durch das Abscheiden spezieller Funktionsschichten mittels Siebdruck auf der Keramik erhalten. Dabei sind die sensitiven Schichten je nach Anwendungsgebiet des Sensorarrays beliebig austauschbar.

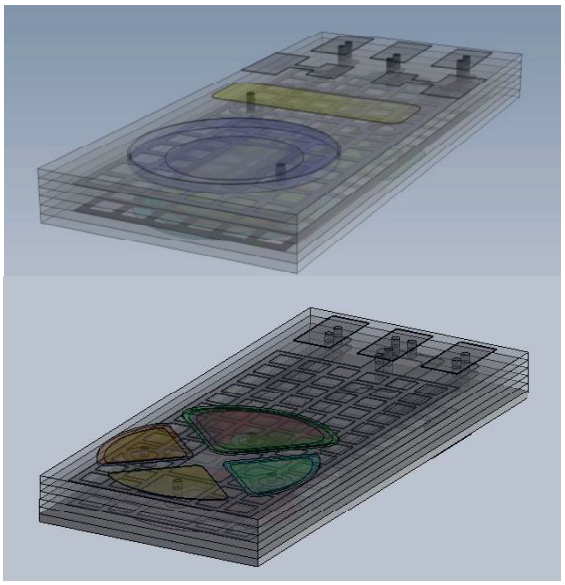


Abb.18: HTCC-Sensorplattform mit unterschiedlichen Layouts

Im vorgestellten Beispiel wurde der Funktionsnachweis für die Erfassung der Messgrößen pH-Wert, Cu^{2+}-Schwermetallion sowie der Kationen Ca^{2+}, NH_4^+ und K^+ in wässrigen Lösungen erbracht. Die Herausforderung bestand insbesondere in der Material- und Prozessentwicklung für die Abscheidung ionensensitiver Dickschichtmembranen auf der keramischen Sensorplattform. Alle ISEs besitzen im untersuchten Messbereich ein sehr gutes Ansprechverhalten und hohe Empfindlichkeiten. Damit stellt die keramische Mehrlagentechnologie ein geeignetes Verfahren da, um kostengünstige und miniaturesierte Multisensoren für die Wasseranalyse herzustellen.

DOI 10.5162/sensoren2019/P1.11

LHyCon (Low Hydrogen Concentration measurement sensor)

Arvind Balakrishnan[1], Saner Akan[1], Marko Völkel[1], Sebastian Geiger[2], Peter Rabenecker[2],
[1] LAMTEC Meß- und Regeltechnik für Feuerungen GmbH & Co. KG, Wiesenstraße 6, Germany
[2] Fraunhofer-Institut für Chemische Technologie ICT, Joseph-von-Fraunhofer-Str. 7, Germany

Abstract

This paper presents a newly developed Low Hydrogen Concentration measurement sensor for hydrogen leak detection. This project further developed a solid electrolyte sensor based on mixed potential theory. The main goals of this project were to develop a hydrogen sensor that has an increased sensitivity for 0-1000 ppb of Hydrogen in air, nearly linear sensor signal for 0-1000 ppb Hydrogen in air, quick response (t_{60} < 1 sec), low-cost production and reduced power consumption and portability through miniaturization. The senor type LH 44 which was fabricated within the framework of this project was found to be the most efficient variation which fulfils the defined goals. The final goal of the project, to develop a low concentration hydrogen sensor for mobile application through miniaturization, was demonstrated.

Keywords: Low Hydrogen Concentration measurement sensor, solid-state mixed potential electrochemical sensors, miniaturized hydrogen sensor and mobile application

Introduction

Components tightness testing is one of the industrial applications where helium is most commonly used as tracer gas [1]. But its drawbacks are high cost and requiring vacuum environment as well as a sophisticated measurement set up. Hydrogen is the lightest element followed by helium. Further, it is cheaper compared to Helium and can be detected easily in ambient air. Hence, its presence has increased in the last decades in this field of application. This has also increased the demand for a hydrogen sensor that is sensitive for low hydrogen concentration with a quick response time and a possible low cost production. Solid-state mixed potential hydrogen sensors became increasingly popular in the last three decades because they offer a simple means to measure hydrogen with high sensitivity and quick response time [2] [3]. A typical sensor of this sort consists of two different electrodes with a solid-state electrolyte in between [4]. The working principle, known as non-Nernst-behaviour [3], is based on a voltage generated due to the different electrochemical behaviour of the two electrodes [2].

The CarboSen sensor of LAMTEC GmbH is one of such solid-state mixed potential electrochemical sensors that has a wide field of application [5]. Depending on the application CarboSen is also referred as HydroSen [6]. The cross-section schematic view of CarboSen sensor is shown in *Figure* 1. A stacked assembly of working electrode – solid-state electrolyte – reference electrode is realized on an aluminium oxide (Al_2O_3) carrier substrate. Reference electrode is a platinum and the working electrode is a platinum-gold based cermet. The solid-state electrolyte used here is yttria stabilized zirconia (YSZ) and has a porous structure. The porous structure enables both electrodes to react with the measured gas simultaneously. YSZ which has to be operated in a temperature range between 400 °C and 900 °C in order to utilize oxygen ion conductivity [3]. Therefore, a heater is placed on the back side of the carrier substrate.

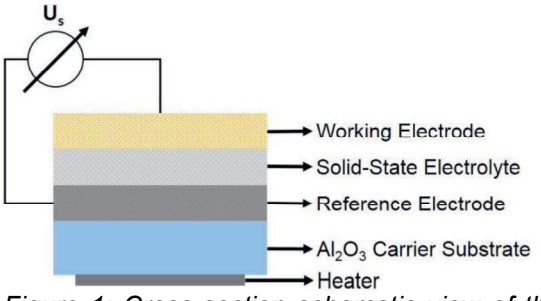


Figure 1: Cross-section schematic view of the CarboSen sensor construction

The working electrode and the reference electrode react to hydrogen in different ways. On the reference electrode, only catalytic reactions occur but on the working electrodes, both catalytic and electrochemical reactions

occur. This builds a measureable potential difference between the electrodes, which yields the sensor signal. As the hydrogen concentration changes, the mixed potentials of each electrode also change due to the change in the rate of the catalytic and electrochemical reactions.
One of the limitations of CarboSen is the unsatisfactory sensitivity for low Hydrogen concentrations at ambient air. A satisfactory measured sensor voltage is delivered by CarboSen only from the range of 10-1000 ppm of Hydrogen in ambient air.

Basing this as the motivation the focus of the LHyCon project was to develop a sensor system which reaches the following goals

- Increased sensitivity for 0-1000 ppb of Hydrogen in air.
- Nearly linear sensor signal for 0-1000 ppb Hydrogen in air (optional).
- Response time (t_{60} < 1 sec).
- Low-cost production.
- Reduced power consumption and portability through miniaturization.

Approach

The variation of surface morphology and the optimization of the functional materials corresponding to the electrode 1 and 2 were identified to be a suitable approach to reach the above-mentioned goals.
A design of experiments was used to generate a test matrix with the below-mentioned aspects as main focus. The test matrix was classified into 6 major sections, each section corresponds to one of the below-mentioned aspects.

- Variation of the functional electrode material thickness of electrode 2.
- Variation of the available three phase boundary on the functional electrode 2.
- Variation of the different functional material-I, on electrode 2.
- Variation of the different functional material-II, on electrode 2.
- Variation of the surface morphology of electrode 2 by fabrication process.
- Variation the electrode design of electrode 1.

Within each of the sections, there were certain variations. Details of each of these sensor variations will not be revealed in this paper due to the intellectual property rights of LAMTEC GmbH.
The nomenclature used for each test section was LH XX. LH stands for LHyCon, the name of the project and the XX corresponds to the test number. In total 75 different sensor types were tested on the available state-of-the-art CarboSen sensor platform.

Sensor fabrication

The sensors, fabricated in this project, are a planar type solid-state mixed potential electrochemical sensors. Planar sensors were fabricated with screen printing technology and industrially available Al_2O_3 was chosen as the carrier substrate by maintaining flexibility, reproducibility and the possibility for low-cost mass production. On one side of the carrier substrate two electrodes with contact pads and glass protective coating, on the other side the heater element with contact pads and glass protective coating were screen printed.

In the state-of-the-art CarboSen sensor, the solid-state electrolyte is YSZ, reference electrode is fabricated using an industrially available high purity platinum paste and the working electrode is a platinum-gold alloy cermet material and is fabricated using a hybrid fabrication process. The heating element is also screen printed using a platinum paste. Each layer of screen printing was followed by a high temperature sintering process. The Schematic view of the state-of-the-art CarboSen is shown in *Figure* 2.

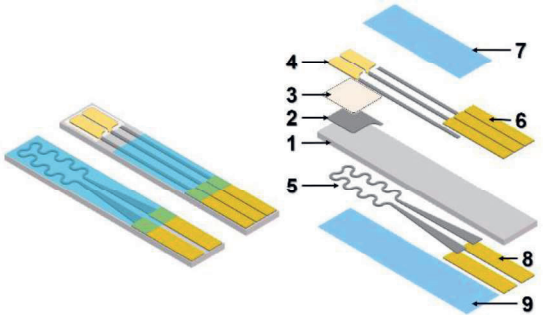


Figure 2: Schematic view of the state-of-the-art CarboSen sensor design.
1: carrier substrate, 2: reference electrode, 3: YSZ, 4: Working electrodes, 5: heating element, 6: electrode contact pads, 7: glass layer, 8: heating element contact pads and 9: glass layer.

The protective glass layer is used on the front side to protect the contact lines that extend from the electrodes to the gold contact pads. Similarly, on the back side the complete surface of the heating element is covered with the protective glass layer. The sensors fabricated in the LHyCon project were based on the same sensor design and construction as in the CarboSen, but the electrodes were varied according to the test matrix. One sensor type variation from the test matrix was fabricated on

one a 4" * 4" carrier substrate which consisted of 75 sensor elements.

Characterisation

A picture of a fully fabricated LHyCon sensor is presented in Figure 3.



Figure 3: Fully fabricated LHyCon sensor.

In Figure 3, the left-hand side image is the front side of the sensor and the right-hand side image is the backside. Figure 4 shows the cross-section SEM image of the fabricated sensor element

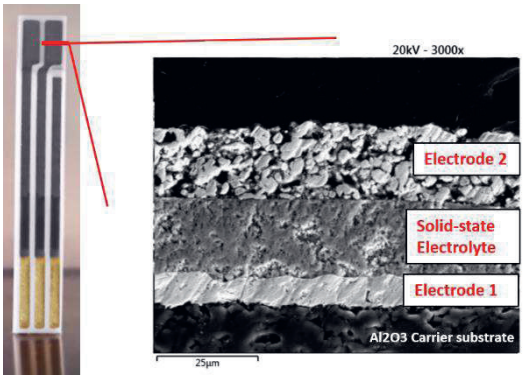


Figure 4: Cross-section SEM image of a specific location on the fabricated sensor.

Performance

The sensor elements were tested with good statistics (10 sensor elements simultaneously from every type) in various low concentrations of hydrogen in air and different sensor operating temperatures by yielding temperature characteristics as well as test gas characteristics of the fabricated sensor elements. Connected to a 10-fold sensor electronic rack, in a 10-fold measurement chamber the sensor elements were exposed to test gas concentrations that were well-defined through commissioning premixed test gas cylinders and high precision pc controlled digital flow regulators. A schematic view of the measurement setup is presented in Figure 5.

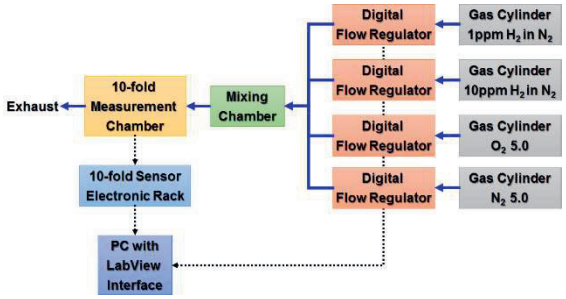


Figure 5: Test gas measurement setup

The 10-fold measurement chamber was an assembly, which consisted of two rectangular PEEK plates pressed on each other by a pneumatic actuator forming 10 individual slots as measurement chambers for each sensor element.
One of the rectangular PEEK plates was stationary on one end of the pneumatic actuator with sensor sockets and cables attached. The other one was the moving counter plate. The pneumatic actuator was operated on one horizontal axis through a control unit with an input of 4 bar air pressure. A picture of the 10-fold measurement chamber is presented in Figure 6.

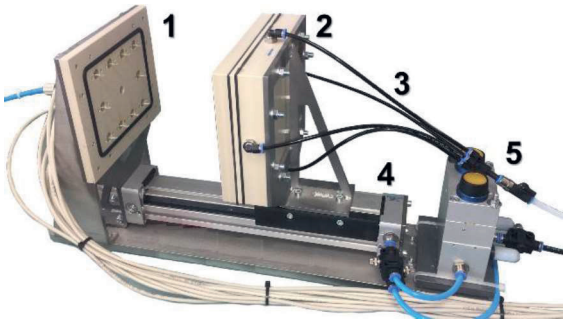


Figure 6: 10-fold measurement chamber,
1: plate with sensor connection, 2: counter plate with gas divider, 3: test gas connections, 4: Pneumatic actuator and 5: Pneumatic control with pressurized air.

All tests were carried out with three concentration ranges:

- 0-1000 ppb of hydrogen in air
- 1-7 ppm of hydrogen in air
- 7-100 ppm of hydrogen in air

The test gas cylinders with low concentrations of hydrogen and the gas mixing assembly were cross-verified using a mass-spectrometer. The mass spectrometer was able to detect concentrations lower than 1 ppm hydrogen in synthetic air (down to 50 ppb) that were mixed and supplied by the LHyCon gas mixing setup. Through a PC with LabView software interface, at each measurement point test gas mass flow

rate was fixed at 1600 sccm as well as the dwell time at 3 minutes.
The sensors were connected to the 10-fold sensor electronic Figure 7 where the sensor temperatures were manually adjusted and measurement data was recorded through a PC with LabView interface.



Figure 7: 10-fold Sensor Electronic Rack.
1: potentiometers for adjusting sensor temperature and 2: sensor connection module

Due to the nature of solid-state electrolytes, in the same hydrogen gas concentration the sensors perform differently at different operating temperatures. Therefore, temperature characteristics were obtained through various tests. The sensor temperature was controlled by adjusting the heater power and further control the heater resistance at a constant value. For the tests, three main heater powers values were specified: 2,1 W, 2,5 W and 2,9 W corresponding approximately to 450°C, 550°C and 600°C respectively. The optimum heater power for low concentrations of hydrogen in air was found at 2,1 W. The following measurement procedure was followed uniformly for all the sensors measured in the project:

- Pre-heating with 3 W heating power for a defined time to initialize the sensor.
- Sensor heating set to 2,9 W and the sensor signal measurement when exposed to 0-100 ppm of hydrogen concentration in air.
- Sensor signal measurement when exposed to 0-7 ppm of hydrogen concentration in air with heating powers 2,9 W, 2,5 W and 2,1 W.
- Sensor heating set to 2,1 W and the sensor signal measurement when exposed to 1-1000 ppb of hydrogen concentration in air.

Results and discussion

All fabricated sensors were measured using the three sets of different hydrogen concentration ranges in air. The measured results were analysed and concluded that the sensor type LH 44 best fulfilled the defined goals of the LHyCon project. The average measured sensor voltage of 10 LH 44 sensors when subjected to hydrogen concentrations of 0 ppm to 7 ppm is presented in Figure **8**, where the heating power is 2,1 W.

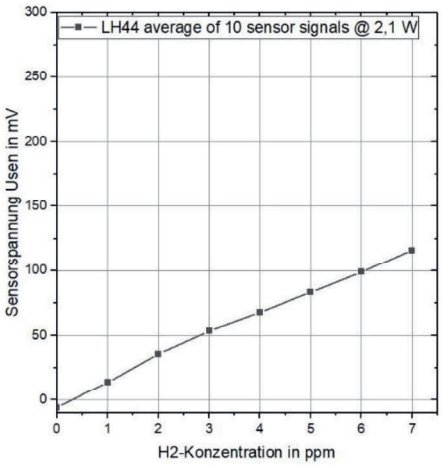


Figure 8: Sensor signal characteristic of LH 44 sensor type when exposed with 0-7 ppm hydrogen concentrations in air

Consecutively, the average measured sensor signal of 10 sensors of LH 44 sensor type when subjected to hydrogen concentration of 0 ppb to 1000 ppb is presented in Figure 9. The figure represents the behaviour of the sensor signal when the sensors were operated with heating power of 2,1 W.

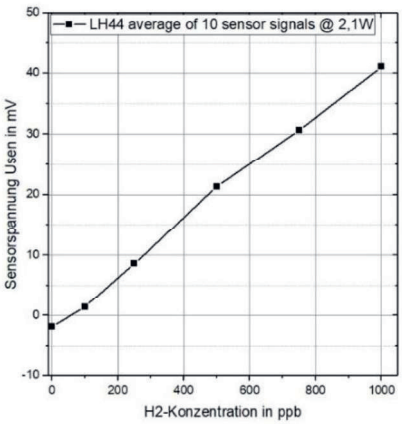


Figure 9: Sensor signal characteristic of LH 44 sensor type when exposed with 0-1000 ppb hydrogen concentrations in air

In order to the compare the amplitude of the average measured sensor signals of the LH 44 sensor type in 0 ppb to 1000 ppb range, 10 state-of-the-art CarboSen sensors were tested exactly as mentioned previously. The comparison of the average measured sensor

signals is presented in Figure 10. This figure reveals the increase in the sensor signal for the LH 44 sensor type in comparison to the CarboSen. This increase in the sensor signal is attributed to an increased electrochemical reactions dominated and reduced catalytic reactions of the LH 44 sensor type electrode surface.

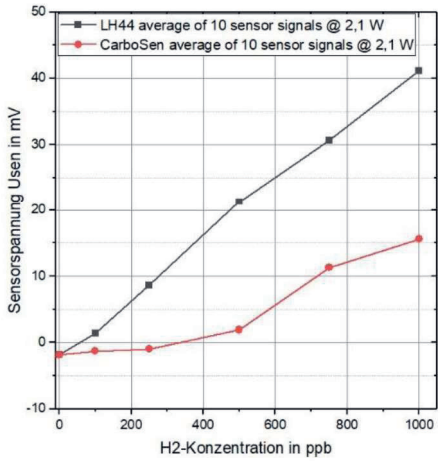


Figure 10: Comparison of sensor signals of LH 44 sensor type and the CarboSen when exposed to the hydrogen concentration range 0-1000 ppb.

Based on this inference, the final goal of the project was carried out, to achieve a low concentration hydrogen sensor for mobile application through miniaturization of the sensor that has reduced power consumption and better portability.

The miniaturized version of the LH44 sensor was designed with 3 mm*3 mm substrate dimensions. Therefore, the new sensor type is named as LH 44X3. The exploded schematic view of the miniaturized LH 44X3 sensors is shown in Figure 11.

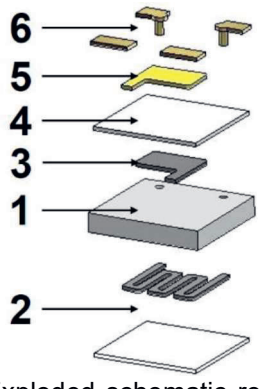


Figure 11: Exploded schematic representation of miniaturised LH 44X3 sensors.

1: carrier substrate, 2: heating element, 3: electrode 1, 4: solid-state electrolyte YSZ, 5: electrode 2 and 6: gold contact pads and the via hole connections.

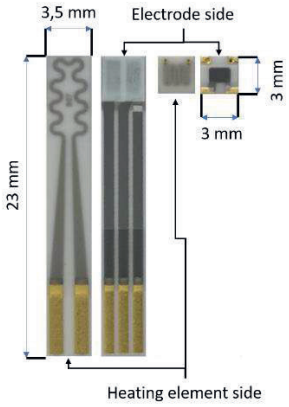


Figure 12: Picture comparing the fully fabricated LH 44 sensor in the standard size sensor and the miniaturized sensor.

A fully fabricated LH 44X3 was assembled in a Transistor Outline (TO) housing. A picture of the LH 44X3 in TO housing is shown in Figure 13.

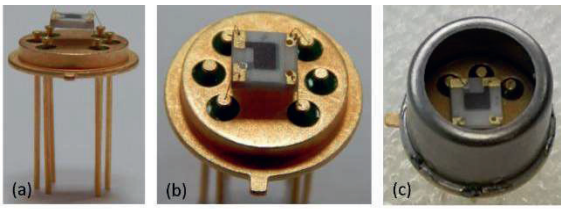


Figure 13: A fully fabricated LH 44X3 and assembled in a TO housing. (a) side view, (b) top view and (c) top view with full TO housing of the LH 44X3 sensors.

To achieve efficiency the new heater coil of LH 44X3 was designed thoroughly through elaborate scientific work.
The design simulation of the heating element was done within the framework of the project and it was optimized such that the minimum heating required to achieve the optimum sensor temperature was in the range of 0,8-1 W. The heating element was modelled in SolidWorks 2018 (Dessault Systémes). The Joule heating effects, voltage and temperatures were simulated and calculated using SolidWorks-Flow Simulation and Electronic Cooling modules for a current range from 0,55 A to 0,90 A.
A picture of the optimal heating element design obtained from the simulation is shown in Figure 14.

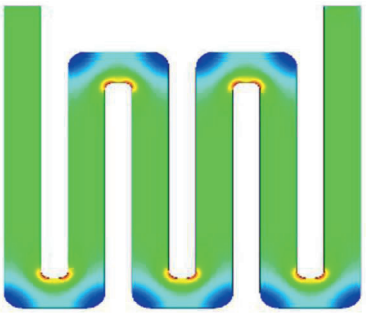


Figure 14: Picture of the optimal heating element design with Solidworks Flow Simulation module.

Figure 15, shows the calculated optimal operation point for LH 44X3 as required by specification at 0,7 A and 0,97 V with peak temperature of 454 °C in the centre reaching down to 440 °C at the outer edge of the 3 mm*3 mm substrate resulting in a calculated power draw of 0,68 W.
To verify the simulation a LH 44X3 sensor was fabricated and tested with the optimized heating element design. A picture of the fully fabricated heating element based on the obtained simulation results is shown in Figure 16.

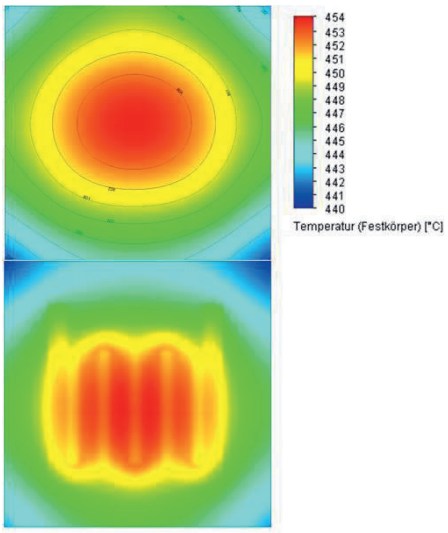


Figure 15: Simulation of the LH 44X3 sensor heating element, shown are resulting temperatures at the optimal working point of 0.7 A.

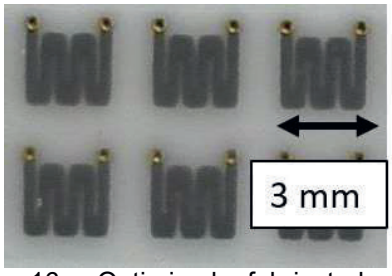


Figure 16: Optimized fabricated heating element of the LH 44X3.

The initial testing of the LH 44X3 was successful with the optimized heating element. The fully fabricated and assembled LH 44X3 sensors were heated using a high precision power source, Keithley 2601B. The power source was set in current controlled mode of operation. The current was also measured using an Agilent 34401A digital multimeter. Experiments were conducted in order to achieve the sensor operating temperature from 400°C to 650°C. The measured temperature was correlated with calculated power consumption values obtained from the simulation. Thermal imaging of the heated LH 44X3 sensor was performed using InfraTec VarioCam HD. A typical measured thermal image is presented in Figure 17. In the figure, the peak temperature in the middle of the sensor is about 450 °C. The thermal image was done with an emissivity of 0,95. The heating power was 1,18 W. The difference between the simulated and the measured heating power is attributed to the thermal losses that occur due to the bonding wires and the TO-housing.

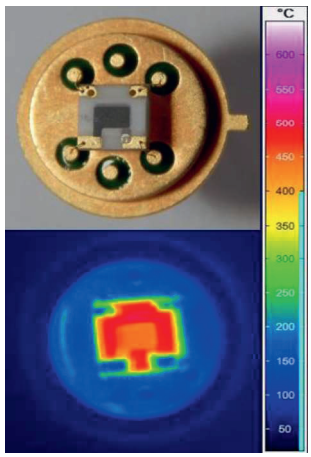


Figure 17: Thermal imaging of the LH 44X3 sensor with the optimized heating element.

Outlook

The following aspects are currently work in progress:

- The further optimization of the heater power for the newly developed LH 44X3.
- The testing of the fully fabricated LH 44X3 with various hydrogen concentrations in air.
- The development of a simple electronic control unit for LH 44X3 is in progress.

Future work includes a system design for a mobile hydrogen leak detection unit that is driven by industrially available batteries or using a standard USB port.

Acknowledgement

We would like to thank the Karlsruhe Institute of Technology, PTKA-PFT and the Bundensministereim für Bildung und Foschung for the financial and administrative support for the project LHyCon (02P16K090).

Reference

[1] INFICON, "INFICON Holding AG," 29 04 2019. [Online]. Available: https://products.inficon.com/en-us/nav-products/category/productgroup/leak-detectors/.

[2] M. Norio, S. Tomoaki, A. SriAyu, I. Hiroshi und Z. Serge, „A review of mixed-potential type zirconia-based gas sensors," *Ionics,* Bd. 20, pp. 901-925, 2014.

[3] M. Ralf, S. Kathy, F. Maximilian, G. Ulrich, B. Nicolae und W. Udo, „Solid State Gas Sensor Research in Germany – a Status Report," *Sensors,* Bd. 9, pp. 4323-4365, 2009.

[4] G. Fernando H., B. Eric L. und M. Rangachary, „THIN FILM MIXED POTENTIAL SENSORS". USA Patent US7264700B1, 04 09 2007.

[5] F. Hammer, „Vorrichtung und Verfahren zur Bestimmung der Konzentration eines Stoffes in einem Gas". Germany Patent DE102006054505A1, 29 05 2008.

[6] LAMTEC, „LAMTEC Sensors and Systems for Combustion technology," [Online]. Available: https://www.lamtec.de/en/products/sensorik/carbosen-coe.html. [Zugriff am 29 04 2019].

DOI 10.5162/sensoren2019/P1.12

Pulspolarisation: Einfluss der Polarisationsspannung auf die NO_x-Detektion mit dem System Pt|YSZ

N. Donker[1], A. Ruchets[2], D. Schönauer-Kamin[1], J. Zosel[2], U. Guth[3] und R. Moos[1]
[1] Lehrstuhl für Funktionsmaterialien, Universität Bayreuth, 95440 Bayreuth, Deutschland
[2] Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, 04736 Waldheim, Deutschland
[3] Fakultät Chemie und Lebensmittelchemie, TU Dresden, 01062 Dresden, Deutschland

Zusammenfassung

Mittels der Spannungs-Puls-Methode werden Stickoxide mit einem Pt | YSZ-Sensor gemessen. Bei dieser Methode wird der Sensor eine definierte Zeit lang polarisiert und der Spannungsverlauf bei der anschließenden Selbstentladung des Sensors gemessen. Hierbei führen bereits geringe NO-Konzentrationen zu einer deutlich beschleunigten Selbstentladung des Sensors. Diese beschleunigte Entladung wird dann als Sensorsignal genutzt. Hierbei wird der Einfluss der Höhe der Polarisationsspannung in Bezug auf die Empfindlichkeit des Sensorsignals untersucht. Es zeigt sich, dass eine Erhöhung der Polarisationsspannung nicht automatisch eine Steigerung des Sensorsignals zur Folge hat. Dies kann möglicherweise mit dem Einsetzen von Nebenreaktionen wie der Elektrolyse von vorhandenen Wasserdampf erklärt werden.

Keywords: Spannungs-Puls-Methode, Pt | YSZ, NO_x-Sensor, Gassensor, dynamisches Sensorprinzip

Einleitung

Kohlenmonoxid (CO), Kohlenwasserstoffe (HC) sowie Stickoxide (NO_x) gehören zu den Hauptemissionen bei Verbrennungsprozessen. Dabei sind insbesondere letztere in jüngster Zeit in den Fokus der Öffentlichkeit gerückt. Stickoxide bewirken neben Atemwegsreizungen auch die Bildung von saurem Regen sowie den Abbau der Ozonschicht.

Der Hauptverursacher solcher NO_x-Emissionen in Deutschland ist der Verkehr [1]. Um diese Emissionen zu reduzieren und auch um den immer strenger werdenden Abgasnormen zu genügen, ist es notwendig, den Verbrennungsprozess sowie die anschließende Abgasreinigung weiter zu optimieren. Hierfür ist eine kontinuierliche und selektive Messung der Stickoxidkonzentration nötig. Eine neuartige Möglichkeit hierfür ist die dynamische Spannungs-Puls-Methode an Pt|YSZ-basierten Sensoren.

Spannungs-Puls-Methode

Die Sensoren, welche für die Spannungs-Puls-Methode verwendet werden, sind ähnlich wie Mischpotentialsensoren aufgebaut. Bei diesem Sensortyp werden zwei ungleiche, unterschiedlich elektrokatalytisch aktive Elektroden auf einem Festelektrolyten verwendet. Dies führt dazu, dass sich an der Elektrode mit geringerer Aktivität kein thermodynamisches Gleichgewicht einstellen kann, wohingegen dieses an der Elektrode mit höherer Aktivität zumindest teilweise erreicht wird. Dadurch bildet sich an der Elektrode mit geringerer Aktivität ein Mischpotential, welches sich in Form einer Spannung zwischen den beiden Elektroden messen lässt.

Abweichend davon werden bei der Spannungs-Puls-Methode zwei identische Elektroden (hier Platin) auf einem Festelektrolyten (hier 8YSZ) verwendet (Abb. 1a). Sind beide Elektroden bei gleicher Temperatur der gleichen Gasatmosphäre ausgesetzt, führt dies dazu, dass die Spannungsdifferenz zwischen ihnen 0 V beträgt. Um ein Signal zu erhalten, wird der Sensor eine Zeitspanne t_{pol} lang mit der Spannung U_{Pol} polarisiert (siehe Abb. 1b). Anschließend entlädt sich der Sensor nach dem Entfernen der Spannungsquelle von selbst wieder. Es konnte gezeigt werden, dass der Verlauf dieser Selbstentladung stark von der NO_x-Konzentration in der umgebenden Atmosphäre abhängt und dass die Selbstentladung mit steigender NO_x-Konzentration beschleunigt abläuft [2,3]. Nach Ablauf der Entladezeit $t_{Entladung}$ wird der Sensor wieder für die Dauer t_{Pol} lang polarisiert, allerdings mit entgegengesetzter Polarität $-U_{pol}$. Ein solcher Zyklus mit alternierender Polarisation und zwischenzeitlicher Entladung ist schematisch in Abb. 1b gezeigt. Die gesamte Zyklusdauer t_{Zyklus} ergibt sich dabei aus der Summe der Polarisations- und Entladezeiten

($2t_{Pol} + 2t_{Entladung}$). Anhand der Spannungswerte zu einem definierten Zeitpunkt im Zyklus lässt sich dann ein Sensorsignal generieren. Diese Zeitpunkte sind in Abb. 1b schematisch als $U_{Signal1}$ und $U_{Signal2}$ gekennzeichnet. Im Falle eines perfekt symmetrischen Sensoraufbaus sind die Beträge dieser Spannungen im Rahmen der Messgenauigkeit gleich.

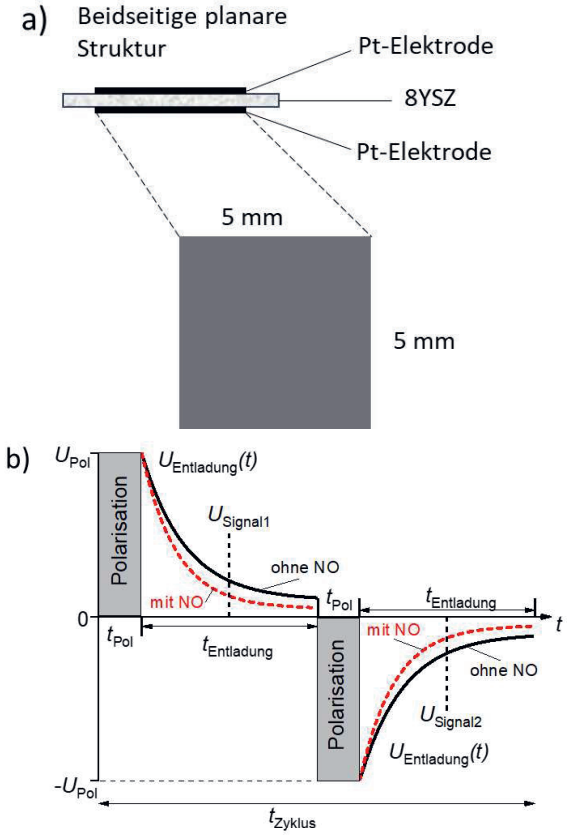


Abb. 1 a) Struktur des Sensors und b) schematischer Ablauf eines Spannungs-Puls-Zyklus

Experimentelles

Als Sensor dienten 5 mm x 5 mm große Platin-Elektroden, welche auf 300 µm dicke 8YSZ-Substrate siebgedruckt und anschließend bei 950 °C gebrannt wurden. Anschließend wurden die Sensoren mit Hilfe von 100 µm dicken Platindrähten kontaktiert.

Die darauffolgenden Messungen wurden bei T = 400 °C in einem gasspülbaren Rohrofen durchgeführt. Die Temperatur wurde hierzu mit Hilfe eines Thermoelements, welches direkt beim Sensor angebracht ist, geregelt.

Für die Untersuchungen der Gassensor-eigenschaften wurde ein Gemisch aus 10 % Sauerstoff (O_2) in Stickstoff (N_2) mit ca. 2 % Wasser (H_2O) als Grundgas verwendet. Nachdem der Sensor jeweils eine Stunde lang dem Grundgas ausgesetzt war, wurden zusätzlich schrittweise 5, 10, 25 und 50 ppm NO jeweils 15 min lang zum Grundgas hinzudosiert. Anschließend floss 15 min lang nur Grundgas. Dieses Gasprofil ist schematisch in Abb. 2 gezeigt.

Als Parameter für die Pulspolarisation wurde eine Polarisationsdauer t_{Pol} von 1 s verwendet. Die anschließende Selbstentladung $t_{Entladung}$ dauerte 10 s. Dies ergibt eine gesamte Zyklusdauer t_{Zyklus} von 22 s ($2t_{Pol} + 2t_{Entladung}$). Die Polarisationsspannung U_{Pol} wird in dem Versuch schrittweise von 50 mV auf 2000 mV erhöht. Dabei wird bei jedem Spannungsschritt das in Abb. 2 gezeigte Gasprofil durchlaufen und anschließend die Polarisationsspannung auf die nächst höhere Stufe erhöht.

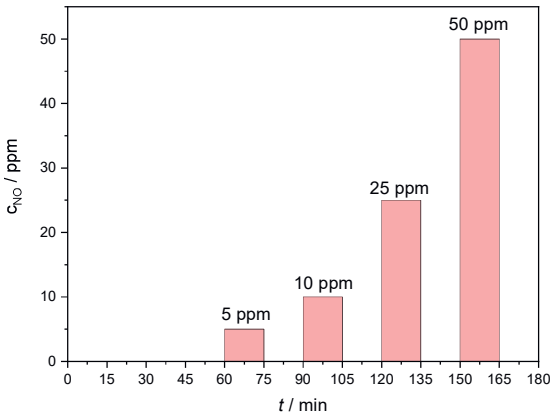


Abb. 2: Bei den Messungen verwendetes Gasprofil

Um die durch die Polarisation verursachten Veränderungen am Sensor zu beobachten, wurde zusätzlich vor und nach den Spannungs-Puls-Messungen ein Impedanzspektrum (f=10^7 - 10^{-3} Hz, U_{Eff}=10 mV) am Sensor aufgenommen.

Messergebnisse

Bei den Messungen mit der Spannungs-Puls-Methode konnte bei allen Polarisationsspannungen U_{Pol} eine deutlich beschleunigte Selbstentladung bei Vorhandensein von NO in der Gasatmosphäre beobachtet werden. Stellvertretend für die anderen Polarisationsspannungen ist ein vollständiger Zyklus mit U_{Pol} = 1000 mV exemplarisch in Abb. 3 gezeigt. Während die Selbstentladung des Sensors im Grundgas nur sehr langsam abläuft, ist die durch NO beschleunigte Entladung deutlich zu kennen. So beträgt die Spannung am Sensor nach positiver Polarisation und anschließender 10-sekündiger Selbstentladung im Grundgas noch ~360 mV, wohingegen bereits 5 ppm NO in der umgebenen Gasatmosphäre diese Restspannung auf ~180 mV halbieren. Höhere

Konzentrationen führen zu einer noch stärker beschleunigten Selbstentladung und damit zu noch niedrigeren Spannungswerten zu diesem Zeitpunkt.

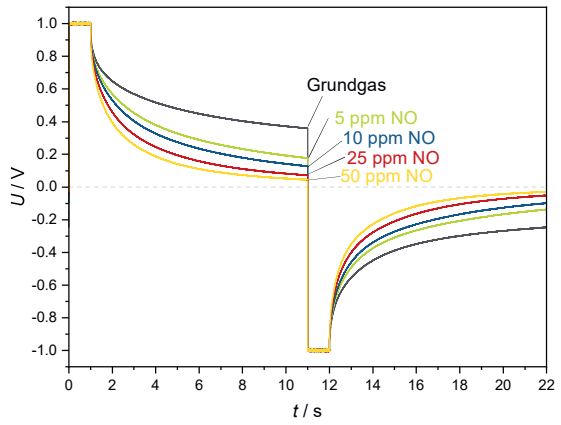


Abb. 3: Vollständer Spannungs-Puls-Zyklus im Grundgas (10% O_2 + H_2O in N_2) und zusätzlichen 5 bis 50 ppm NO mit U_{Pol}=1000 mV, t_{Pol}=1 s, $t_{Entladung}$=10 s

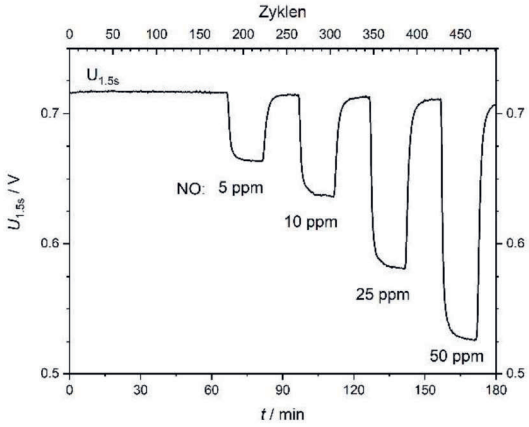


Abb. 4: Sensorspannung jedes Zyklus zum Zeitpunkt 1,5 s nach Zyklusbeginn im Grundgas (10% O_2 + H_2O in N_2) und zusätzlichen 5 bis 50 ppm NO mit U_{Pol}=1000 mV, t_{Pol}=1 s, $t_{Entladung}$=10 s.

Wie in Abb. 1b angedeutet, lässt sich auch aus den beschleunigten Selbstentladungen ein Sensorsignal generieren, indem die Spannungswerte zu einem festen Zeitpunkt im Zyklus ausgewertet werden [4]. Dies ist in Abb. 4 für eine Auswertezeit von 1,5 Sekunden nach Zyklusbeginn ($U_{1.5s}$) gezeigt. Bei dieser Art der Auswertung entspricht jeder Punkt dem Spannungswert zu einem definierten Zeitpunkt im jeweiligen Zyklus und ist somit auch ein Maß für die Stabilität des Sensors und des Signals. Die durch die beschleunigte Selbstentladung niedrigeren Spannungen sind für die einzelnen NO-Stufen in Abb. 4 deutlich zu erkennen. Der gerade Verlauf zeigt die hohe Stabilität, trotz über 400 durchlaufener Zyklen.

Zur Ermittlung der idealen Auswertezeitpunkte wurde die Spannungsdifferenz zwischen einer Entladung im Grundgas und einer Entladung mit zusätzlichen 50 ppm NO gebildet. Es zeigt sich, dass die Spannungsdifferenz zwischen einer Entladung im Grundgas $U_{Grundgas}$ und einer Entladung mit zusätzlichen 50 ppm NO $U_{50\ ppm\ NO}$ mit steigender Polarisationsspannung bis zu 1000 mV ansteigt. Auch zeigen alle Kurven einen ähnlichen Verlauf. Wird die Polarisationsspannung dagegen weiter erhöht (1000 mV bis 2000 mV), so sinkt die Differenz wieder ab. Zusätzlich verschieben sich die maximalen Spannungsdifferenzen zu späteren Zeitpunkten. Durch die so unterschiedlichen Kurvenverläufe zeigt sich, dass es keinen perfekten Auswertezeitraum gibt und das dieser je nach Polarisationsspannung gewählt werden muss.

Den gleichen Trend konnte auch bei geringeren NO-Konzentrationen beobachtet werden. Auch bei nur 5 ppm NO steigt die Differenzspannung bis zu einer Polarisationsspannung von 1000 mV an. Allerdings wird dieser Anstieg ab ca. 500 mV Polarisationsspannung deutlich langsamer und es kommt in dem Polarisationsspannungs-bereich von 900 bis 1000 mV zu einer Art Sättigung des Differenzsignals. Ab einer Polarisationsspannung von 1000 mV nimmt dann auch diese Spannungsdifferenz ab, wie bereits oben für 50 ppm NO beschrieben.

Zur Ermittlung der NO-Empfindlichkeit wurden die Entladespannungen des Sensors 5 Sekunden nach Zyklusbeginn U_{5s} über der zugegebenen NO-Konzentration aufgetragen. Der Zeitpunkt nach 5 Sekunden wurde gewählt, da hier alle Differenzkurven eine geringe Steigung aufweisen und somit die Auswirkungen des Auswertezeitpunktes gering sind. Es zeigt sich, dass für alle Polarisationsspannungen eine logarithmische Abhängigkeit zwischen der Entladespannung U_{5s} zu diesem Zeitpunkt und der NO-Konzentration beobachtet werden kann. Hierbei entspricht die Steigung der Geraden der Empfindlichkeit des Sensors.

Es ist ein kontinuierlicher Anstieg der Empfindlichkeit bis zu einer Polarisationsspannung von 1050 mV zu erkennen. Bei einer weiteren Erhöhung der Polarisationsspannung nimmt diese wieder ab. Allerdings zeigt sich bei U_{Pol} = 1700 mV eine kleine Erhöhung der Empfindlichkeit, welche

allerdings nicht an die Spitzenwerte bei einer Polarisierung mit 1050 mV heranreicht.
Dieses so ermittelte Maximum der Empfindlichkeit lässt sich bei einer Betrachtung der Sensorsignale bestätigen. Diese sind in Abb. 5 gezeigt. Aus Gründen der Übersichtlichkeit sind allerdings nicht alle Signale dargestellt. Auch hier ist die größte NO-Empfindlichkeit, welche der an der stärksten beschleunigten Entladung entspricht, bei einer Polarisation mit 1000 mV zu beobachten. Zusätzlich ist zu beachten, dass die Messungen nacheinander mit steigender Polarisationsspannung durchgeführt wurden. Dies bedeutet, dass der Sensor vor der Messung mit einer Polarisation von 2000 mV bereits über 19000 Zyklen (39 · 490 Zyklen) bei niedrigeren Polarisationsspannungen durchlaufen hat und dabei weiterhin ein stabiles Signal aufweist, was auf stabile Zyklen hindeutet. Weiterhin ist zu erkennen, dass es zwischen einer Polarisation mit 1500 mV oder mit 2000 mV im Grundgas 4 Sekunden nach der Polarisation (U_{5s} = 1 s Polarisation und 4 s Entladung) keine Spannungsunterschiede gibt. Dies deutet darauf hin, dass eine Sättigung auftritt, bei der eine weitere Erhöhung der Polarisationsspannung zu keiner weiteren Erhöhung im Entladeverhalten führt.

Weiterhin wurden vor und nach den Spannungs-Puls-Messungen Impedanzspektroskopieuntersuchungen durchgeführt. Diese zeigen eine deutliche Erhöhung des Widerstands durch die Spannungs-Puls Messung. Diese Veränderung am Sensor spiegelt sich allerdings nicht in einem unstabilen Sensorverhalten wider.

Diskussion

Der verwendete Sensor zeigt bei allen Polarisationsspannungen eine gute Zyklenstabilität, repräsentiert durch ein stabiles Sensorsignal. Außerdem konnte bei allen Messungen eine durch NO klar beschleunigte Entladung festgestellt werden. Diese Beschleunigung tritt auch schon bei der vergleichsweise geringen Polarisation mit 50 mV auf. Da bei diesen Spannungen vermutlich kein oder sehr wenig Platinoxid an der Grenzfläche gebildet werden kann, hat auch die Platinoxid-Bildung voraussichtlich keinen dominierenden Effekt auf das Sensorsignal [5]. Allerdings zeigt das Impedanzspektrum eine deutliche Erhöhung des Widerstandes, was wiederrum auf eine solche Bildung bei höheren Spannungen hindeutet. Doch auch falls dies der Fall sein sollte, so zeigt sich diese Bildung nicht in einer Verschlechterung des Sensorverhaltens.

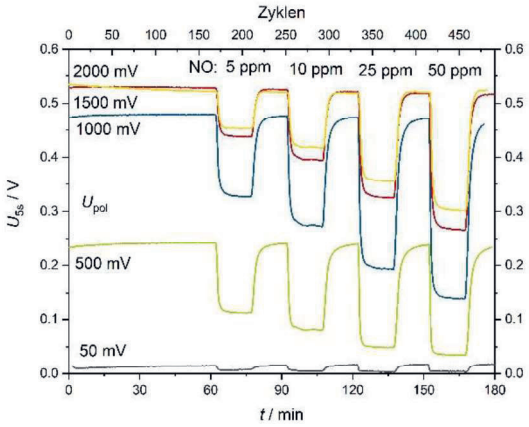


Abb. 5: Aus den Entladekurven ermittelte Sensorsignale nach vorheriger Polarisation mit 50 bis 2000 mV

Die Steigerung der Empfindlichkeit lässt nach derzeitigem Stand mit der sich erhöhenden Spannung erklären. So kann sich der Sensor auch durch das NO maximal bis zu einer Spannungsdifferenz in Höhe von 0 V entladen. Selbst für den theoretischen Fall, dass sich der Sensor im Grundgas überhaupt nicht entlädt, wäre die maximale Spannungsdifferenz so auf die Polarisationsspannung begrenzt. Eine Erhöhung der Polarisationsspannung lässt so auch eine Erhöhung der Spannungsunterschiede für die einzelnen NO-Konzentrationen zu.

Der allgemeine Effekt der beschleunigten Entladung könnte auf eine Verschiebung des NO/NO_2-Gleichgewichts zurückzuführen sein. So wird angenommen, dass aufgrund der hohen katalytischen Aktivität von Platin alle Gase an der Elektrode im thermodynamischen Gleichgewicht vorliegen. Im Falle von NO ist der Zerfall zu O_2 und N_2 jedoch kinetisch gehemmt, weswegen diese Reaktion hier ausgeschlossen wird [6]. Weitere thermodynamische Berechnungen führen zu dem Schluss, dass NO und NO_2 bei der betrachteten Temperatur von 400°C etwa 1:1 im Gleichgewicht liegen. Wird der Sensor nun polarisiert, so verschiebt sich dieses Gleichgewicht an der positiven Elektrode in Richtung NO_2 und an der negativen Elektrode in Richtung NO. Nach der Polarisierung streben beide Gase wieder einen Gleichgewichtszustand an, was zu der beschleunigten Entladung durch Stickoxide führt.

Die Senkung der Empfindlichkeit ab einer Polarisationsspannung von ca. 1050 mV hängt möglicherweise mit der einsetzenden Elektrolyse des vorhandenen Wasserdampfs

zusammen [7]. Diese Nebenreaktion reduziert den Effekt der Polarisierung auf die Stickoxide, was sich wiederrum durch eine Reduzierung der Empfindlichkeit auswirken kann. Hierzu sind aber noch weiterführende Untersuchungen notwendig und auch geplant.

Zusammenfassung und Ausblick

Es konnte bei den Spannungs-Puls-Messungen eine maximale NO-Empfindlichkeit nach der Polarisierung mit 1000 mV beobachtet werden. Dies wurde mit einer bei höheren Spannungen möglicherweise einsetzenden Elektrolyse des im Gas vorhandenen Wasserdampfes begründet. Um diese Hypothese zu bestätigen, wird Abhängigkeit der Empfindlichkeit von der Polarisationsspannung unter trockenen Bedingungen wiederholt. Es konnte außerdem gezeigt werden, dass der Sensoreffekt schon bei kleinen Polarisationsspannungen zu beobachten ist. Da sich Platinoxide vermutlich erst bei höheren Polarisationsspannungen bilden, scheinen diese für den Sensoreffekt eher eine untergeordnete Rolle zu spielen.

References

[1] Umweltbundesamt, Stickstoffoxid-Emissionen, 2018, https://www.umweltbundesamt.de/daten/luft/luftschadstoff-emissionen-in-deutschland/stickstoffoxid-emissionen, accessed 18 April 2019.

[2] S. Fischer, R. Pohle, B. Farber, R. Proch, J. Kaniuk, M. Fleischer, R. Moos, Method for detection of NO_x in exhaust gases by pulsed discharge measurements using standard zirconia-based lambda sensors, Sens. Actuators B 147 (2010) 780–785.

[3] S. Fischer, R. Pohle, E. Magori, M. Fleischer, R. Moos, Detection of NO by pulsed polarization of Pt I YSZ, Solid State Ionics 262 (2014) 288–291.

[4] S. Fischer, R. Pohle, M. Fleischer, R. Moos, Method for reliable detection of different exhaust gas components by pulsed discharge measurements using standard zirconia based sensors, Procedia Chemistry 1 (2009) 585–588.

[5] H. Pöpke, E. Mutoro, B. Luerßen, J. Janek, Oxidation of Platinum in the Epitaxial Model System Pt(111)/YSZ(111): Quantitative Analysis of an Electrochemically Driven PtO_x Formation, J. Phys. Chem. C 116 (2012) 1912–1920.

[6] M. Crocoll, Modellierung und Simulation der Pt-katalysierten NO-Oxidation in sauerstoffreichen Abgasen. Dissertation, Karlsruhe, 2003.

[7] F.M. Sapountzi, J.M. Gracia, C.J. Weststrate, H.O.A. Fredriksson, J.W. Niemantsverdriet, Electrocatalysts for the generation of hydrogen, oxygen and synthesis gas, Progress in Energy and Combustion Science 58 (2017) 1–35.

DOI 10.5162/sensoren2019/P1.13

Innovatives Sensorsystem und Auswerteverfahren zum Monitoring bei der Lebensmittelüberwachung

Rolf Seifert[1], *Hubert B. Keller*[1], *Heinz Kohler*[2]
[1]*Institut für Automation und Angewandte Informatik (IAI), Karlsruher Institut für Technologie (KIT), Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen*
[2]*Institut für Sensor- und Informationssysteme (ISIS), Hochschule Karlsruhe Technik und Wirtschaft (HsKA), Moltkestr. 30, 76133 Karlsruhe*

Zusammenfassung

Der Bedarf an innovativen online Analyse Systemen insbesondere beim Qualitätsmonitoring bei der Lebensmittelerzeugung bzw. –verarbeitung und -lagerung wächst ständig. In diesem Bericht wird ein Gassensorsystem für diesen Bereich vorgestellt, das durch seine zyklische Arbeitsweise gasspezifische Zeitwertleitprofile erzeugt, die zur Analyse der gemessenen Stoffprobe herangezogen werden können und so eine Stoffidentifikation und Konzentration der Stoffkomponenten ermöglichen. Für diese Stoffanalyse sind innovative Kalibrier- und Auswerteverfahren notwendig. Zwei leistungsstarke Verfahren werden in dieser Arbeit vorgestellt.

Keywords: Gassensorik, Lebensmittelüberwachung, Sensorsystem, thermozyklische Betriebsweise, Auswerteverfahren

Einleitung

Der Bedarf an innovativen online Analyse Systemen zur Überwachung von toxischen Gasen, zum Monitoring von flüchtigen Gasen in chemischen und biochemischen Prozessen oder zur Früherkennung von Schwelbränden wächst ständig. Ein wichtiges, weiteres Anwendungsgebiet ist das Qualitätsmonitoring zur Überwachung der Prozesskette, bestehend aus Lebensmittelerzeugung, –verarbeitung, -lagerung und -transport.
So muss zum Beispiel bei der Bierherstellung die Qualität der Rohstoffe wie Getreide, Hopfen und Hefe untersucht werden, um eine Kontamination mit 2,4,6-Trichloranisol (TCA) bereits vor der Verarbeitung zu erkennen [1]. TCA ist ein chlorhaltiger, aromatischer Kohlenwasserstoff, der durch seinen intensiv schimmlig-muffigem Geruch gekennzeichnet ist und demzufolge zu einer immensen Beeinträchtigung der Qualität des Produktes führen würde. Dabei ist die TCA Problematik nicht nur auf den Bereich der Bierherstellung beschränkt, sondern führt auch bei der Produktion von zahlreichen anderen Lebensmittelerzeugnissen, wie z. B. bei Wein durch kontaminierte Korken, zu erheblichen Qualitätseinbußen.
In allen diesen Fällen werden flüchtige organische Verbindungen (VOCs) freigesetzt, die durch innovative Sensor Systeme analysiert werden können und somit zur frühzeitigen Erkennung der beschriebenen Probleme dienen bzw. zur Qualitätssicherung der Prozess- und Lagerungsbedingungen herangezogen werden können.
In diesem Beitrag wird ein solches innovatives Sensorsystem mit den für die Gasanalyse ausgelegten mathematischen Kalibrier- und Auswerteverfahren vorgestellt.

Thermozyklische Betriebsweise

Der Einsatz von Metalloxid (MOX) Gassensoren zur Analyse von Gasen erscheint zunächst höchst attraktiv, da die Sensorelemente preiswert, hoch empfindlich, in einfachen Monogas-Anwendungen als Grenzwertgeber bewährt und langzeitstabil sind. Allerdings ist die Analysefähigkeit von Gasgemischen mittels MOX-Wirkschichten bei isothermem Betrieb in Folge von Querempfindlichkeiten unzureichend. Folglich ist dieser Sensortyp bei dieser Betriebsweise selbst zur Identifikation von Gaskomponenten und einfachen Gasgemischen nicht geeignet, da die spezifischen Oberflächenprozesse mit den Zielgasen selbst bei Modifikation der Wirkschichteigenschaften mit katalytischen Zusätzen bei isothermem Betrieb nicht selektiv genug sind, um spezifische Gaskomponenten bei ausreichender Diskriminierung anderer Komponenten zu analysieren [2]. Erst bei thermozyklischem Betrieb [3,4] können simultan aussagekräftige, gasspezifische Leitwert-Zeit-Profile (LZP) gemes-

sen werden, die zur numerischen Analyse der Gaszusammensetzung herangezogen werden kann. Abbildung 1 zeigt beispielhaft die Stofftypischen LZPs verschiedener Gase.

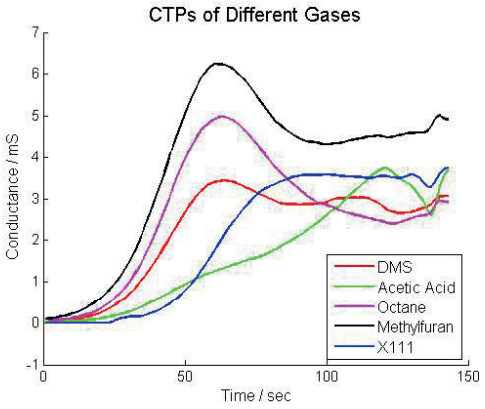


Abb. 1: Charakteristische LZP für verschiedene Stoffe.

Sensorsystem

Der am Institut für Sensor- und Informationssysteme (ISIS) der Hochschule Karlsruhe – Technik und Wirtschaft entwickelte Sensorchip (Abb. 2) ist ein 4-fach Sensorarray auf einem eigenbeheizten Aluminiumoxid Chip (4x4 mm^2). Auf der Vorderseite sind vier mikrostrukturierte Dünnschicht-Interdigitalelektroden (IDE) mit verschiedenen SnO_2/Additiv gassensitive Schichten in Dickschichttechnik und einem metallischen Widerstands-Temperatursensor integriert (Abb. 3) und auf der Rückseite ist eine Pt-Heizung mikrostrukturiert, die mit einer Glasschicht abgedeckt ist [5,6,7].

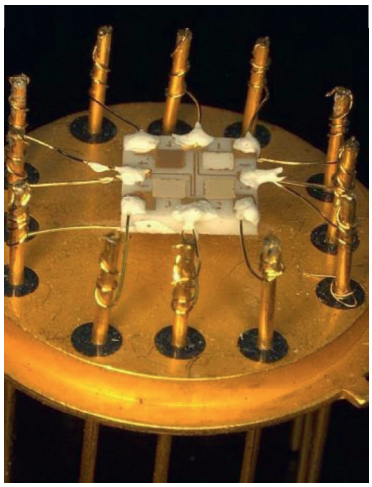


Abb. 2: In TO8-Gehäuse montiertes 4-fach Sensorarray mit verschiedenen SnO_2/Additiv- Schichten.

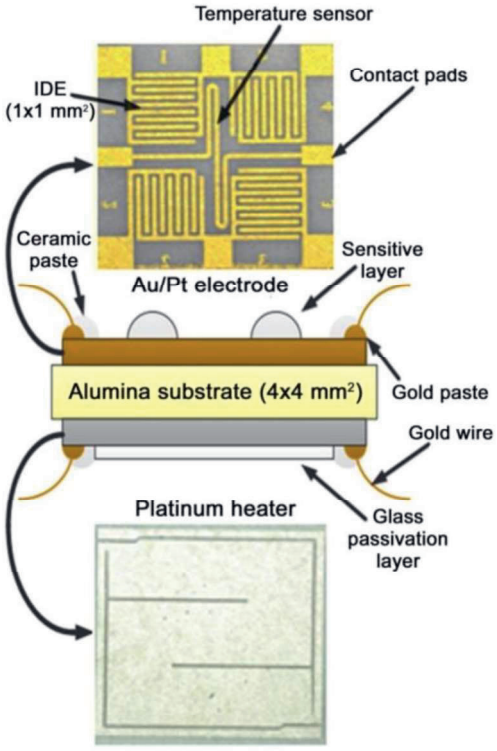


Abb. 3: Schematischer Aufbau des Sensorchip mit vier (unbeschichteten) IDE und dem Widerstands-Temperatursensor auf der Vorderseite und dem Pt-Heizer auf der Rückseite.

Durch thermozyklische Temperaturregelung und simultane Messung von vier spezifisch für die Messaufgabe ausgewählten gassensitiven SnO_2/Additiv-Wirkschichten wird eine hohe Identifikations- und Alarmierungssicherheit bei Einsatz von nur einem Sensor-Chip erreicht.
Bestandteil des Sensorsystems sind sensitive Schichten, welche entwickelt werden, um die bei thermozyklischem Betrieb relevanten Gaskomponenten im Umfeld variabler Umgebungsbedingungen zu detektieren. Die untere Sensitivitätsgrenze kann bei diesem Betriebsverfahren, abhängig von der Schichtauswahl (SnO2/Additiv), dessen Morphologie und der Gasart, bei wenigen ppm liegen.

Das Kalibrier- und Auswerteverfahren ProSens

Am Karlsruher Institut für Technologie (KIT) wurde speziell für die Analyse von LZP ein mathematisches Kalibrier und Auswerteverfahren zur Stoffidentifikation und Einzelkonzentrationsbestimmung der Komponenten eines Stoffgemisches auch bei variablen Umgebungsbedingungen namens ProSens [8] entwickelt. Das Verfahren besteht aus einem Kalibrier- und einem Auswerteteil. Im Kalibrier-

teil wird auf Basis von Kalibriermessungen ein parametrisches mathematisches Kalibriermodell erstellt und dem Auswerteteil in Form von Modellparametern zur Verfügung gestellt. Wie schon in zahlreichen Applikationen nachgewiesen werden konnte [6,9,10], ist das Verfahren in der Lage, sowohl eine sehr gute Stoffidentifikation als auch eine sehr genaue Bestimmung der Konzentrationen der Stoffkomponenten auch auf Basis von nur wenigen Kalibriermessungen durchzuführen.

Das Verfahren SimSens

In vielen Anwendungsgebieten können mehrere Zielstoffe bzw. Zielstoffgemische gleichzeitig auftreten. Das oben beschriebene Verfahren ProSens ist aber nur für die Analyse eines einzigen Stoffgemisches ausgelegt. Für die simultane Analyse mehrerer Zielstoffgemische wurde im KIT das Verfahren SimSens [11] entwickelt. Analog zu dem Verfahren ProSens besteht SimSens aus einem Kalibrier- und einem Auswerteteil. Im Kalibrierteil werden bei SimSens im Gegensatz zu ProSens nicht nur ein Kalibriermodell berechnet, sondern für jedes Zielstoffgemisch ein spezielles Kalibriermodell. Diese sind wiederum parametrische Modelle und die Modellparameter werden an den Auswerteteil übertragen.
Im Auswerteteil wird das gemessene LZP einer aktuellen Probe mit den theoretischen LZP vergleichen, die auf der Basis der verschiedenen Kalibriermodelle berechnet wurden. Hat eine dieser theoretischen Kalibrierkurven nur einen „geringen“ Abstand zu der gemessenen Kalibrierkurve, so identifiziert SimPlus die aktuell gemessene Probe mit dem entsprechenden Zielstoff und berechnet die Konzentration der Einzelkomponenten der Probe. In realen Anwendungen konnte die gute Performance des Verfahrens SimSens bereits nachgewiesen werden [12,13].

Fazit

Das oben beschriebene Sensorarray in Verbindung mit den vorgestellten Kalibrier- und Auswerteverfahren hat in mehrfachen Tests in anwendungsnahen Experimenten sein hohes Potential zur Lösung vielfältiger Gasanalyseprobleme bestätigt und kann deshalb als ein vielversprechendes Werkzeug im weiten Bereich der Lebensmittelüberwachung zur Anwendung kommen.

Acknowledgement

Das Sensornetzwerk HybridSensorNet (www.hybridsensornet.org) wird den Technologietransfer dieses vielversprechenden Sensorsystems unterstützen.

Literaturnachweis

[1] M. Platthaus, „Korkgeschmack analytisch betrachtet“, Zeitschrift Laborpraxis, 20.1.2011
[2] M. Ulrich, L. Ejsing, R. Graf, „Brandfrühwarnsysteme mit Halbleiter-Gassensoren“, atp edition, Automatisierungstechnische Praxis, 2009
[3] H. Kohler, J. Rober, N. Llink, and I. Bouzid, “New applications of tin oxide gas sensors – I. Molecular recognition by cyclic variation of the working temperature and reasonable numerical analysis”, *Sensors & Actuators* B: 163-169 (1999)
[4] H. Kohler, M. Bauer, and I. Bouzid, “Gas Recognition Using SnO2 Gas Sensors: Analysis of volatile organic substances dissolved in water”; *Proceedings Sensor Congress Nürnberg*, Vol. I AMA Service, Wunstorf. Germany: 125-128 (1999).
[5] K. Frank, V. Magapu, V. Schindler,Y. Arnanthigo,H. Kohler,H.B. Keller, R. Seifert, “Improving the analysis capability of tin oxide gas sensors by dynamic operation, appropriate additives and an advanced evaluation procedure”, *Sensor 2007*: 13th Internat.Conf., Nürnberg, May, 2007, Proc. Vol. 1, AMA Service GmbH, Wunstorf, 139–144 (2007).
[6] K. Frank, V. Magapu, V. Schindler, H. Kohler∗, H. B. Keller, and R. Seifert., “Chemical Analysis with Tin Oxide Gas Sensors: Choice of Additives, Method of Operation and Analysis of Numerical Signal”, *Sensors Letters* 6, 908-911 (2008).
[7] N. Illyaskutty, J. Knoblauch, M. Schwotzer, and H. Kohler,”Thermally modulated multi sensor arrays of SnO2/additive/electrode combinations for gas identification*”, Sensors and Actuators B* 217, 2-12 (2015).
[8] R. Seifert, H. B. Keller, K. Frank, and H. Kohler, *Sensor Letters*, Vol. 9/1, 7-10 (2011); doi.org/10.1063/1.4975280
[9] A. Jerger, H. Kohler, F. Becker, H. B. Keller, and R. Seifert, “New applications of tin oxide gas sensors II. Intelligent sensor system for reliable monitoring of ammonia leakages”, *Sensors and Actuators B*: Chemical, 81, 2-3, . 301-307 (2002).
[10] K. Frank, A. Hetznecker, H. Kohler, V. Schindler, U. Schönauer, Y. Arnanthigo, D. Falk, “Metal oxide gas sensors for field analysis: Novel SnO2/La2O3 sensor element for analysis of dissolved toluene/ethanol binary mixtures”, *Sensor 2005*: 12th Internat. Conf., Nürnberg, May, 2005, Proc. Vol. 2, AMA Service GmbH, Wunstorf, 207–209 (2005)
[11] H. B. Keller, R. Seifert, H. Kohler „SimSens – a New Mathematical Procedure for Simultaneous Analysis of Gases with Resistive Gas Sensors”, Sensors & Actuators: B. Chemical (2015), pp. 203-207 DOI information: 10.1016/j.snb.2014.10.133
[12] R. Seifert, H. B. Keller, N. Illyaskutty, J. Knoblauch, H. Kohler, „Numerical Signal

Analysis of Thermo-Cyclically Operated MOG Gas Sensor Arrays for Early Identification of Emissions from Overloaded Electric Cables" Sensors & Transducers journal (ISSN: 2306-8515, e-ISSN 1726-5479), Vol. 193, Issue 10, September 2015, pp.74-79

[13] R. Seifert, T. Conrad, J. Peter, H. B. Keller. „Breath Control in Respiratory Air and Simultaneous Analysis of Sensor Data" Vortrag 12th Asian Conference on Chemical Sensors (ACCS 2017), November 12 -15, Hanoi, Vietnam

DOI 10.5162/sensoren2019/P1.14

Fasergekoppelter konfokaler Sensor zur exakten Abstandsregelung für maskenlose Lithografieanwendungen

Johannes Kirchner[1], Rostyslav Mastylo[1], Uwe Gerhardt[1], Taras Sasiuk[1], Laura Weidenfeller[1], Martin Hofmann[2], Michael Kühnel[1], Stefan Sinzinger[3], Eberhard Manske[1]
[1]Institut für Prozessmess- und Sensortechnik, Gustav-Kirchhoff-Str. 1, 98693 Ilmenau, De
[2]Fachgebiet Mikro- und nanoelektronische Systeme, Gustav-Kirchhoff-Str. 1, 98693 Ilmenau, De
[3]Fachgebiet Technische Optik, Helmholtzring 1, 98693 Ilmenau, De

Zusammenfassung

Maskenlose Fotolithografie findet auch abseits der Halbleitertechnik hohe und immer weiterwachsende Bedeutsamkeit [1, 2]. Als dauerhaftes und wesentliches Entwicklungsziel hat sich die Vergrößerung der Packungsdichte der Mikro- bzw. Nanostrukturen herausgestellt. Dazu ist es nötig die angewendeten Lithografieprozesse zu optimieren und die Strukturen mit sehr hoher Präzision und einer sehr guten Reproduzierbarkeit zu fertigen. Eine Form der maskenlosen Lithografie wird durch Direct-Laser-Writing (DLW) beschrieben [3]. Bei diesem Prozess ist die Strukturbreite von der Fokussierbarkeit und der axialen Ausrichtung des verwendeten Fabrikationslasers abhängig. Um die Präzision dieser Lithografieanwendung und der zu entwickelnden Strukturgeometrie zu erhöhen, wird eine Nanopositionier- und Nanomessmaschine (NPMM) zur relativen Ausrichtung der beschichteten Probe zum Laserstrahl verwendet [4]. Ein neuentwickelter chromatisch-konfokaler Sensor scannt dafür die Topographie der Probenoberfläche und regelt den Maschinenhub so nach, dass der Polymerisationsprozess stets in der Brennebene des Fabrikationslasers stattfindet.

Keywords: Nano-CMMs, konfokale Sensorik, maskenlose Lithografie, Direct-Laser-Writing

1 Einleitung

Ein zentraler Forschungsbereich der Mikro- und Nanofabrikationstechnik beschäftigt sich mit der Auslegung und Umsetzung adäquater Positionier- und Messwerkzeuge. Durch die Realisierung und Entwicklung von Nano-Koordinatenmessmaschinen (Nano-CMMs) wurden die quantifizierbaren Größen Auflösung und Wiederholbarkeit in den letzten 20 Jahren kontinuierlich gesteigert [5]. Die Messunsicherheit solcher Maschinen sank dabei im Laufe der Entwicklung immer weiter [6]. Die präzise Positioniertechnik von Nano-CMMs im Bereich weniger Atombreiten bildet die Grundlage vieler neuer Anwendungsbereiche. Dazu gehören neben hochpräzisen Oberflächen-, Kontur- und Abstandsmessungen nun auch die Nanofabrikationstechnik. Dabei werden klassische und neue Fertigungsverfahren so optimiert, dass Strukturgrößen im Bereich von Mikro- bzw. Nanometern realisierbar werden. Zu den üblichen Fabrikationsmethoden für solche Abmessungen gehören u.a. die Field-Emission Scanning Probe Lithography (FE-SPL) [7], Nanoimprint Lithography (NIL) [8] und Soft-NIL [9], AFM-Scratching [10] sowie Direct-Laser-Writing (DLW) [11]. Beim DLW werden Fotolacke (Photoresiste) auf die Probe aufgetragen. Wird dieser Resist mit einer bestimmten Wellenlänge belichtet, polymerisiert er an dieser Stelle. Durch die anschließende Entwicklung des Lacks, nimmt die geschriebene Struktur eine reale Geometrie an und wird sichtbar [3]. DLW zeichnet sich vor allem als maskenlose, berührungslose, vergleichsweise schnelle und flexible Strukturierungsmethode aus. Während die FE-SPL Schreibgeschwindigkeiten von nur einigen Mikrometern pro Sekunde erreicht, können beim DLW problemlos mehrere Millimeter pro Sekunde strukturiert werden [12, 13]. Die klassische Halbleiterlithografie erreicht aufgrund simultaner Flächenbelichtung zwar noch höhere Produktionsgeschwindigkeiten - bedingt durch die Verwendung von Masken ist dieser Prozess in seiner Flexibilität und Anwendbarkeit jedoch sehr eingeschränkt und ausschließlich für hohe Stückzahlen ausgelegt. DLW-Lithografie ist sowohl für zwei- als auch dreidimensionale Strukturen realisierbar. Die dreidimensionale Methodik wird dabei durch Zwei- bzw. Mehrphotonen-Polymerisationsprozesse umgesetzt [14]. Diese Anwendung gehört zu den additiven Fertigungsmethoden und kann mit einem 3D-Druck auf mikroskopischer Ebene verglichen werden [15]. Zahlreiche Veröffentlichungen weisen auf eine hohe Anwendung dieser Technik vor allem in der Biomedizintechnik hin [1, 2]. Die Vergrößerung der Packungsdichte

steht bei Nanofabrikationsprozessen oft im Vordergrund, um immer mehr Strukturen auf immer kleineren Raum zu platzieren. Die Packungsdichte ist vom Abstand (Pitch) zweier Strukturen zueinander und somit von der erzeugten Strukturbreite abhängig. Um möglichst geringe Strukturbreiten zu erreichen, ist bei DLW-Anwendungen die korrekte Ausrichtung der Laserfokusebene zur Bearbeitungsfläche (2D) bzw. dem Bearbeitungsvolumen (3D) ein zentraler Untersuchungsgegenstand.

2 Nanopositionier- und Nanomessmaschine

Mit der Entwicklung der Nanopositionier- und Nanomessmaschinen (NPMM), stehen dem Institut für Prozessmess- und Sensortechnik der Technischen Universität Ilmenau Nano-CMMs höchster Präzision zur Verfügung. Diese dienen als Grundlage für künftige Nanofabrikationsarbeiten. Mit einer Auflösung von mehr als 1 nm, einer Positionierwiederholbarkeit von besser als 2 nm und einer 3D-Messunsicherheit von unter 10 nm über dem gesamten Arbeitsvolumen von (25 x 25 x 5) mm³ erreichen NPMM die nötige Präzision und einen ausreichenden Arbeitsbereich für DLW-Anwendungen [16, 17]. Um das Abbesche Komparatorprinzip einzuhalten, arbeiten diese Präzisionsmaschinen im Sample Scanning Mode (vgl. Abb. 1) [4, 18]. Dafür werden die Interferometer (1, 2, 3), sowie der Oberflächensensor (4), fest am metrologischen Rahmen (5) der Koordinatenmessmaschine angebracht. Die optischen Interferometerachsen, sowie die (mechanische bzw. optische) Sensorspitze, schneiden sich virtuell im Abbepunkt (6). Das Messobjekt wird auf dem beweglichen Probentisch (7) platziert. Der Probentisch wird in lateraler Richtung vom Bediener gesteuert. Die axiale Regelung übernimmt die Maschinensteuerung, so dass die Oberfläche der Probe ebenfalls kontinuierlich im Abbepunkt nachgeführt wird. Dafür wertet die Maschinensteuerung die Signale des Sensors aus und übermittelt eine Korrekturbewegung an den z-Antrieb des Tisches. Die Korrekturbewegung des Probentisches wird durch die Interferometer erfasst. Die gemessene Tischbewegung repräsentiert folglich die Oberfläche des Prüfkörpers. Um eine solch hochgenaue Probenplatzierung zu gewährleisten, wurden speziell entwickelte Direktantriebe verbaut. Das dynamische Verhalten dieser Antriebe ist nahezu linear und ermöglicht auch bei hohen Geschwindigkeiten eine sehr präzise Positionierung [19, 20]. Mit einer maximalen Verfahrgeschwindigkeit von über 5 mm/s ermöglichen NPMM im Hinblick auf Nanofabrikationsanwendungen einen hohen Durchsatz [6].

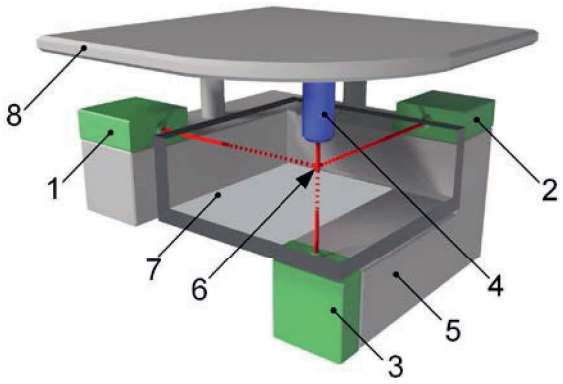


Abb. 1: Prinzipdarstellung der NPMM: 1, 2, 3 - Interferometer; 4 - Oberflächensensor; 5 - metrologischer Rahmen; 6 - Abbepunkt; 7 - beweglicher Probentisch; 8 - Sensorplatte

3 Entwicklung des Oberflächensensors

3.1 Sensoranforderungen

Für den beschriebenen Anwendungsbereich der NPMM ist ein neues Sensorsystem erforderlich. Optische Verfahren haben sich für diese Zwecke bewährt. Methodisch werden dafür spezielle Eigenschaften des Lichts ausgenutzt. So sind in der Literatur häufig astigmatische und interferenzoptische Lösungen oder das Foucaultsche Schneideverfahren zu finden [21–23]. Eine weitere optische Eigenschaft, die chromatische Aberration, wird bei chromatisch-konfokaler Sensorik bzw. Mikroskopie ausgenutzt [24]. Durch die Entwicklung des Fokussensors ist zwar ein adäquates optisches System zum präzisen Messen vorhanden, jedoch ist der Sensor auf eine numerische Apertur von NA = 0,6 ausgelegt und wird somit in seiner lateralen Auflösung begrenzt [6, 25]. Die Breite der Strukturen hängt jedoch umgekehrt proportional von der numerischen Apertur der verwendeten Optik ab (vgl. Gl. 1). Eine möglichst hohe NA führt also zu geringeren Strukturbreiten [26]. Das neue Sensorsystem muss somit auch bei Verwendung von Objektiven höherer numerischer Apertur zuverlässig arbeiten. Zur Reduzierung thermischer Einflüsse wird ein fasergekoppeltes Prinzip verfolgt. Durch Verwendung dieser Lichtwellenleiter (LWL) können wärmeeinbringende Lichtquellen außerhalb des Messaufbaus positioniert werden, ohne dabei das messtechnische Prinzip zu beeinflussen. Die Integrierbarkeit einer weiteren Laserdiode zur Auslösung des Polymerisationsprozesses im Photoresist bestimmt den konstruktiven Rahmen der Koppelstelle zwischen Sensorkopf und Objektiv und wird mittels Strahlteilerwürfel umgesetzt.

3.2 Konzept des neuen Sensors

Ein entsprechender Messaufbau wird in [27] beschrieben. Das Konzept basiert auf dem chromatisch-konfokalen Prinzip. Die Funktionsweise wird in Abbildung 2 deutlich. Das Licht zweier Laserdioden unterschiedlicher Wellenlänge (λ_{M1}, λ_{M2}) wird mittels Lichtwellenleiter-Koppler (1) in einer Faser kombiniert. Im Unterschied zu anderen chromatisch-konfokalen Sensoren fungiert hier das Faserende (2) selbst als konfokale Blende. Hinter der Faser breitet sich das Licht divergent aus und wird durch die Linse (3) kollimiert. In der Fokuslinse (4) kommt es zur Dispersion. Durch die chromatische Aberration entstehen die Fokuspunkte der beiden Laserwellenlängen in unterschiedlichen Ebenen entlang der optischen Achse [28]. Der relative axiale Abstand beider Fokuspunkte wird Schnittweitendifferenz genannt und wird, bei genauer Kenntnis der optischen Komponenten, durch die Sellmeier-Gleichung beschrieben [29]. Das Licht wird vom Messobjekt (5) reflektiert und auf dem Rückweg wieder in den Lichtwellenleiter eingekoppelt. Die rückgekoppelte Lichtintensität ist abhängig von der Position des Messobjekts. Steht das Messobjekt exakt in der Fokusebene einer respektiven Lichtwellenlänge, wird der messbare Rückreflex des entsprechenden Lasers maximiert. Am anderen Ende der Faserkombination trennen Farbfilter (6) die beiden Wellenlängen wieder, sodass zwei separate Intensitätssignale mit Hilfe der Photodioden (7) aufgezeichnet werden können. Werden die Intensitäten über den axialen Verstellweg z aufgezeichnet, erhält man nach Differenzbildung beider Signale einen annähernd linearen Bereich, der als Arbeitsbereich (8) zur Maschinenregelung genutzt wird.

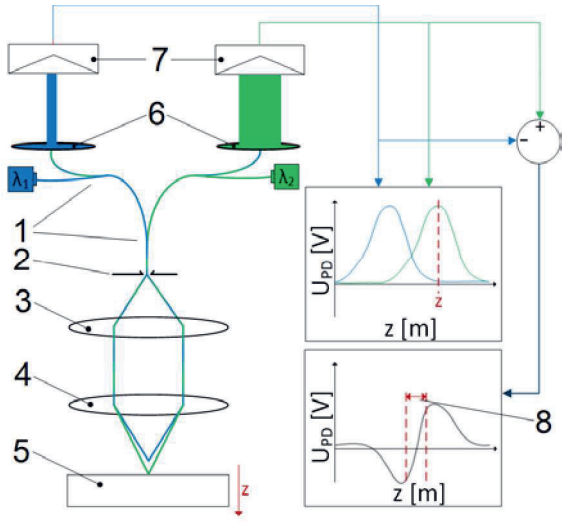


Abb. 2: Konzept des chromatisch-konfokalen Sensors: 1 - Lichtwellenleiter-Koppler (LWL-Koppler); 2 - Faserende (konfokale Blende); 3 - Kollimatorlinse; 4 - Fokuslinse; 5 - Messobjekt; 6 - Farbfilter; 7 - Photodioden; 8 - Arbeitsbereich zur Regelung der NPMM

3.3 Komponentenwahl und Konzeptprüfung

Die Auswahl der Laserdioden bestimmt das gesamte optische Design des Systems. Da der Sensor für DLW-Anwendungen genutzt werden soll, müssen die Messlaser in einem anderen Wellenlängenbereich als der Bearbeitungslaser liegen, da ansonsten schon während des Messvorgangs eine Resistpolymerisation ausgelöst wird. Das Rayleigh-Kriterium ist ein Maß für die Spotgröße von Licht in der Fokusebene und beschreibt somit die spätere theoretische Strukturbreite [26]. Diese hängt von der numerischen Apertur der verwendeten Optik, sowie von der Wellenlänge λ des Lasers nach Gl. (1) ab und soll für den Bearbeitungslaser minimiert werden:

$$R = \frac{1{,}22\lambda}{\mathrm{NA}} \qquad (1)$$

In [26] werden erzeugte Strukturbreiten von b = R < 300 nm beschrieben. Diese werden bei einer Wellenlänge von λ_P = 405 nm und einer numerischen Apertur von NA = 0,9 erzeugt und liegen somit sogar unter dem theoretisch berechneten Limit von R = 549 nm. Für die Strukturierungsarbeiten mit NPMM werden für die Bearbeitungslaser ebenfalls möglichst kurze Wellenlängen, im Farbbereich UV bis Blau, vorgesehen. Die Messwellenlängen müssen daher in einem anderen spektralen Bereich liegen.

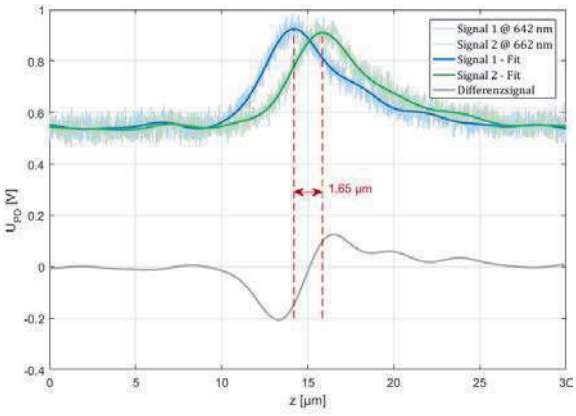


Abb. 3: Konzeptprüfung: Gemessen wird die Photospannnung U_{PD} entlang der optischen Achse z. Die Signale bilden sich wie erwartet aus (vgl. Abb. 2). Im Differenzsignal entsteht zwischen den Peaks ein nahezu linearer Bereich von 1,65 µm.

Für den Versuchsaufbau werden rote Laserdioden verwendet. Die Wellenlängen dieser Laser betragen λ_{M1} = 642 nm und λ_{M2} = 662 nm. Der Aufbau erfolgt nach dem beschriebenen Konzept (vgl. Abb. 2). Die aufgenommenen Spannungen der Photodioden werden elektronisch verstärkt und unter Nutzung eines 8-Bit AD-Wandlers digitalisiert. Ein Planspiegel dient als Messobjekt. Die Höhenänderung zur Fokuspunktdurchfahrt wird mit der NPMM realisiert. In Abbildung 3 werden die aufgenommen Photospannungssignale in Abhängigkeit der axialen Spiegelposition dargestellt. Das Signal-Rausch-Verhältnis (SNR) der Rohsignale ist noch relativ schlecht. Nur durch entsprechendes Fitten der Signale entsteht ein linearer Bereich im Differenzsignal. Und dennoch belegen die Signale die Funktionalität des Messprinzips. Es ist gut zu erkennen, dass aufgrund der Schnittweitendifferenz ein messbarer Abstand zwischen den Fokuspunktebenen der zwei Laser ausgebildet wird und ein Arbeitsbereich > 1 µm entsteht, der nach Signaloptimierung zur Regelung der Hochpräzisionsmaschinen genutzt werden kann.

3.4 Aufbau des Sensorsystems

Zur Verbesserung des SNR werden konzeptionelle und konstruktive Maßnahmen getroffen (vgl. Abb. 4 und Abb. 5). Ein Gehäuse schirmt die Verstärkerelektronik (1), sowie andere elektrische Bauteile und Leitungen ab. Im Gehäuse befinden sich zwei parallel aufgebaute optische Kanäle. Die verbauten Temperaturcontroller (2) verhindern ein thermisches Driften der Wellenlängen der Laserdioden (3). Die Faserenden, an denen die rückgekoppelten Photospannungen gemessen werden, werden mit Faserkollimatoren (4) ausgestattet. Pro Kanal dient ein Justierelement (5) der lateralen Ausrichtung der Photodioden (6). Mit diesem kann die Signalstärke der einzelnen Kanäle unabhängig voneinander verändert werden, um die Signalpeaks zueinander anzupassen. Der Lichtwellenleiter-Koppler (11), der beide Wellenlängen zusammenführt, wird an den optischen Gehäusedurchführungen (7) angebracht. Über die serielle Schnittstelle (8) werden die vorverstärkten Spannungssignale an den A/D-Wandler der NPMM übergeben. Der Sensorkopf (9) ist mit einem einstellbaren Kollimator (10) verbunden. Die Sensorkopfaufnahme an der Sensorplatte (12) ist mit einem Objektiv (13) hoher numerischer Apertur ausgestattet, um den Lichtfleck auf der Probe (14) möglichst klein zu halten und somit die laterale Auflösung zu erhöhen. Eine Kamera (15) dient der Beobachtung der zu scannenden Oberfläche. An Stelle der Kamera können seitlich auch andere Funktionsgruppen montiert werden. So wird für Strukturierungsversuche von dieser Position der Fabrikationslaser in den Strahlengang eingekoppelt. Zur Messung der Oberflächenneigung ist dort eine Positionsdiode (PSD) zur Verfolgung der Strahlauswanderung vorgesehen. Die Markerplatte (16) dient der Ausrichtung Laserstrahlen.

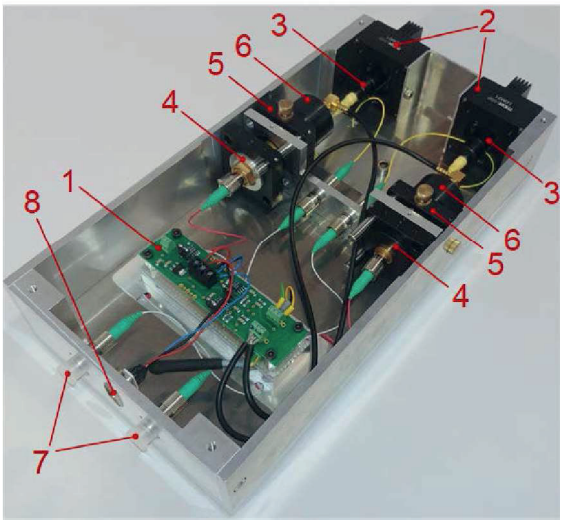


Abb. 4: Aufbau des Sensors: 1 - Verstärkerelektronik; 2 - Temperaturcontroller; 3 - Laserdioden; 4 - Faserkollimatoren; 5 - Justierelemente; 6 - Photodioden; 7 - optische Schnittstelle; 8 - serielle Schnittstelle

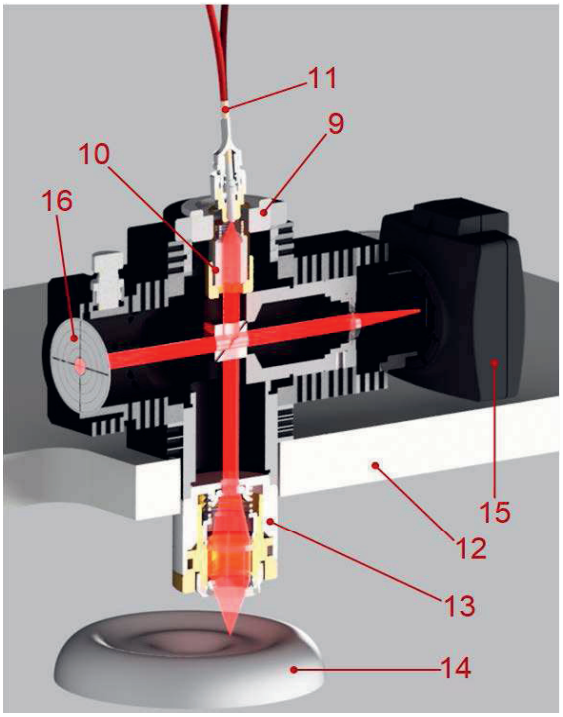


Abb. 5: Sensorkopfaufnahme: 9 - Sensorkopf; 10 - einstellbarer Faserkollimator; 11 - LWL-Koppler; 12 - Sensorplatte; 13 - Objektiv; 14 - Probe; 15 - Kamera; 16 - Markerplatte

Die Funktionsprüfung des Aufbaus erfolgt kongruent zur Konzeptprüfung (Abb. 3) anhand der rückgekoppelten Signalstärke. Abbildung 6 zeigt die beiden Signale. Verglichen mit der Konzeptprüfung ist eine deutlich geringes SNR zu sehen. Der Signalpeak ist von ungefähr 380 mV auf etwa 1610 mV angestiegen (nach Abzug des Offsets) und hat sich somit mehr als vervierfacht. Das Signalrauschen hat sich von circa 100 mV auf etwa 20 mV verringert. Das Aufstellen eines adäquaten Arbeitsbereichs zur Regelung der NPMM ist mit einer solchen Signalqualität problemlos möglich.

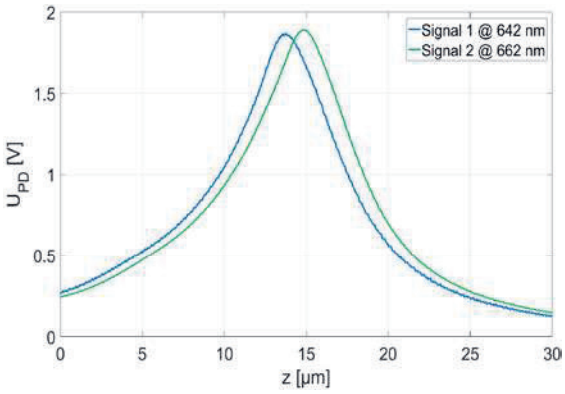


Abb. 6: Sensorsignale nach Überarbeitung des Aufbaus.

3.5 Verschiebung des Arbeitsbereiches zur Optimierung des DLW-Prozesses

Für DLW-Anwendungen wird noch eine weitere Wellenlänge λ_P in den optischen Strahlengang eingekoppelt. Aufgrund der chromatischen Aberration im Objektiv kommt es zu unerwünschten Effekten hinsichtlich der Strukturbreiten [28]. Abbildung 7a) verdeutlicht die Problematik. Die Schnittweiten der Messwellenlängen λ_{M1} und λ_{M2} (roter Farbbereich) spannen den Arbeitsbereich des Sensorsystems auf. Die Oberfläche des Photoresists wird von der NPMM auf die Mitte dieser beiden Schnittweiten positioniert und geregelt. Die Fokusebene des Fabrikationslasers (blauer bzw. ultravioletter Farbbereich) bildet sich auf der optischen Achse zwischen diesem Arbeitsbereich und dem Objektiv aus. Hinter der Fokusebene verlaufen die Strahlen divergent und belichten auf der Resistschicht einen relativ großflächigen Bereich, was sich negativ auf die Breite der Strukturen auswirkt. Durch eine Schnittweitenkorrektur ist es möglich den Arbeitsbereich entlang der optischen Achse so zu verschieben, dass die Fokusebene des Bearbeitungslasers exakt mittig zwischen den Fokusebenen der Messlaser liegt und somit direkt auf der Resistoberfläche. Der DLW-Prozess findet somit direkt im Fokus des Lasers statt, was zu einer extremen Verringerung der Strukturbreite führt. Umgesetzt wird diese Schnittweitenkorrektur durch Änderung des Kollimationszustandes des Messstrahls vor Objektiveintritt. Durch gezieltes konvergieren bzw. divergieren des Strahls, können die Schnittweiten der Messlaser und somit der Arbeitsbereich definiert verschoben werden (vgl. Abb. 7b)).

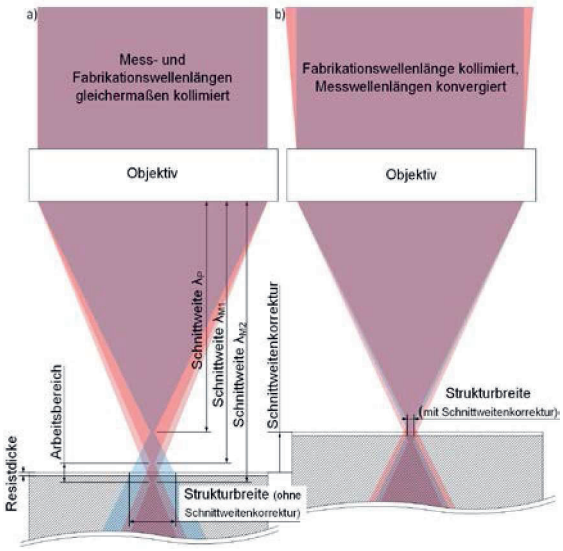


Abb. 7: Verschiebung des Arbeitsbereiches durch Dekollimation des Messstrahls: a) Mess- und Fabrikationslaser sind vor Objektiveintritt gleichermaßen kollimiert. Es entsteht aufgrund des Farblängsfehlers ein Schnittweitenversatz.
b) Der Messlaser wird vor dem Objektiveintritt konvergiert. Das führt zu einem Ausgleich des Schnittweitenversatzes, so dass die Brennebene des Fabrikationslasers im Arbeitsbereich des Sensors liegt.

4 Fazit und Ausblick

Die Untersuchungen zum Aufbau des chromatisch-konfokalen Sensors haben gezeigt, dass sich die in die Faser rückgekoppelten Intensitätssignale wie erwartet ausbilden und dass das Faserende des Lichtwellenleiters somit als konfokale Blende nutzbar ist. Zwischen den Fokusebenen der zwei Messwellenlängen stellt sich ein axialer Abstand ein, der als Arbeitsbereich zur Regelung von Nanopositionier- und Nanomessmaschinen verwendet werden kann. Der Aufbau des Sensors sowie die Funktionsprüfung sind dargestellt und eine Strategie zur Verschiebung des Arbeitsbereiches erklärt. Da die chromatische Aberration ausschließlich von der verwendeten Optik abhängt, werden zunächst unterschiedliche Sensor-Objektiv-

Paarungen und ihre Auswirkung auf die Kennlinien und die Größe des Arbeitsbereiches untersucht. Je besser die Farbkorrektur einer Optik, desto kleiner wird sich der axiale Arbeitsbereich ausbilden. Die Distanz zwischen den Fokusebenen gibt letztlich Aufschluss über die mögliche Auflösung des Sensors. Darauf aufbauend wird in weiterführenden Untersuchungen die Langzeitstabilität des Aufbaus geprüft. Reproduzierbarkeitsmessungen und definierte Stufenhöhenmessungen werden folgen und eine Abschätzung der Messunsicherheit zulassen. Zur Anwendung des Messsystems für Direct-Laser-Writing muss zunächst die Schnittweite der Bearbeitungswellenlänge bestimmt werden. Ein Laser der Wellenlänge $\lambda_P = 405\,\text{nm}$ steht dafür zur Verfügung. In [30] wird eine einfache Methodik zur Ermittlung der Fokusebene, unter Verwendung einer Photodiode, vorgestellt. Nachdem die Schnittweite bestimmt wurde, wird der Strahlengang der Messlaser so manipuliert, dass sich der Messbereich um die Fabrikationsbrennebene legt. Bereits ermittelte Parameter werden genutzt, um die Strukturbreiten der DLW-Lithografie zu verkleinern [13].

Förderung

Die Autoren bedanken sich für die Unterstützung durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen des Graduiertenkollegs „Tip- and laser-based 3D-Nanofabrication in extended macroscopic working“ (GRK 2182) an der Technischen Universität Ilmenau.

Literatur

[1] C. M. B. Ho, S. H. Ng, K. H. H. Li, Y.-J. Yoon, *Lab on a chip* 18, 3627–3637, (2015); doi: 10.1039/C5LC00685F.

[2] X. Li, R. Cui, L. Sun, K. E. Aifantis, Y. Fan, Q. Feng, F. Cui, F. Watari, *International Journal of Polymer Science* 24, 1–13, (2014); doi: 10.1155/2014/829145.

[3] A. Selimis, V. Mironov, M. Farsari, *Microelectronic Engineering*, 83–89, (2015); doi: 10.1016/j.mee.2014.10.001.

[4] G. Jäger, T. Hausotte, E. Manske, H.-J. Büchner, R. Mastylo, N. Dorozhovets, N. Hofmann, *Measurement* 9, 1099–1105, (2010); doi: 10.1016/j.measurement.2010.04.008.

[5] K. Takamasu, S. Ozawa, T. Asano, A. Suzuki, R. Furutani, S. Ozono. 1996. Basic concepts of nano-CMM (coordinate measuring machine with nanometer resolution). In.

[6] E. Manske, T. Hausotte, R. Mastylo, T. Machleidt, K.-H. Franke, G. Jäger, *Meas. Sci. Technol.* 2, 520–527, (2007); doi: 10.1088/0957-0233/18/2/S27.

[7] M. Hofmann, C. Lenk, T. Ivanov, I. W. Rangelow, A. Reum, A. Ahmad, M. Holz, E. Manske, *Journal of Vacuum Science & Technology* B 36, 06JL02-1 - 06JL02-6, (2018); doi: 10.1116/1.5048193.

[8] S. Y. Chou, P. R. Krauss, P. J. Renstrom, *Journal of Vacuum Science & Technology* 6, 4129–4133, (1996); doi: 10.1116/1.588605.

[9] S. Supreeti, J. Kirchner, M. Hofmann, R. Mastylo, I.W. Rangelow, E. Manske, M. Hoffmann, S. Sinzinger. 2019. Integrated soft UV-nanoimprint lithography in a nanopositioning and nanomeasuring machine for accurate positioning of stamp to substrate. In *Novel Patterning Technologies for Semiconductors, MEMS/NEMS, and MOEMS 2019*. SPIE, 44); doi: 10.1117/12.2514832.

[10] R. R. Marangoni, J. Schleichert, I. Rahneberg, R. Mastylo, E. Manske, T. Fröhlich, *tm - Technisches Messen* 9, 587–592, (2017); doi: 10.1515/teme-2016-0066.

[11] L. Weidenfeller, R. Schienbein, J. Kirchner, C. Reinhardt, E. Manske. Development of laser positioning system of high accuracy in the nanometer range. In *Proc. SPIE*, 105440E-1 - 105440E-7); doi: 10.1117/12.2312704.

[12] I. W. Rangelow, M. Kaestner, T. Ivanov, A. Ahmad, S. Lenk, C. Lenk, E. Guliyev, A. Reum, M. Hofmann, C. Reuter, M. Holz, *Journal of Vacuum Science & Technology B* 6, 06J102-1 - 06J102-9, (2018); doi: 10.1116/1.5048524.

[13] J. Kirchner, L. Weidenfeller, U. Gerhardt, R. Mastylo, M. Kühnel, S. Sinzinger, E. Manske. A combined laser scanning and DLW tool for measuring and fabrication tasks with NPMM.

[14] S. Maruo, O. Nakamura, S. Kawata, *Opt. Lett.* 2, 132, (1997); doi: 10.1364/OL.22.000132.

[15] C. M. B. Ho, S. H. Ng, Y.-J. Yoon, *Int. J. Precis. Eng. Manuf.* 5, 1035–1046, (2015); doi: 10.1007/s12541-015-0134-x.

[16] G. Jäger, E. Manske, T. Hausotte, A. Müller, F. Balzer, *Surf. Topogr.: Metrol. Prop.* 3, 34004, (2016); doi: 10.1088/2051-672X/4/3/034004.

[17] G. Jäger, E. Manske, T. Hausotte, H.-J. Büchner, *tm - Technisches Messen* 7-8/2000, (2000); doi: 10.1524/teme.2000.67.7-8.319.

[18] E. Abbe. Messaparate für Physiker. In *Zeitschrift für Instrumentenkunde - Organ für Mittheilungen aus dem gesammten Gebiete der wissenschaftlichen Technik*, 446–448.

[19] S. Hesse, C. Schäffel, H.-U. Mohr, M. Katzschmann, H.-J. Büchner, *Meas. Sci. Technol.* 7, 74011, (2012); doi: 10.1088/0957-0233/23/7/074011.

[20] A. Amthor, S. Zschack, C. Ament. 2008. Position control on nanometer scale based on an adaptive friction compensation scheme. In *2008 34th Annual Conference of IEEE Industrial Electronics*. IEEE, 2568–2573); doi: 10.1109/IECON.2008.4758361.

[21] C. Bricot, J. Lehureau, C. Puech, F. Carvennec, *IEEE Trans. Consumer Electron.* 4, 304–308, (1976); doi: 10.1109/TCE.1976.266827.

[22] K. Naganuma, K. Mogi, H. Yamada, *Opt. Lett.* 7, 393, (1990); doi: 10.1364/OL.15.000393.

[23] F. Zamkotsian and K. Dohlen, *Appl. Opt.* 31, 6532, (1999); doi: 10.1364/AO.38.006532.

[24] G. Molesini, G. Pedrini, P. Poggi, F. Quercioli, *Optics Communications* 4, 229–233, (1984); doi: 10.1016/0030-4018(84)90179-2.
[25] R. Mastylo, D. Dontsov, E. Manske, G. Jager. 2005. A focus sensor for an application in a nanopositioning and nanomeasuring machine. In . SPIE Proceedings. SPIE, 238); doi: 10.1117/12.612887.
[26] F. Zhu and J.-Y. Ma, *Chinese Phys. Lett.* 4, 48102, (2014); doi: 10.1088/0256-307X/31/4/048102.
[27] A. K. Ruprecht, C. Pruss, H. J. Tiziani, W. Osten, P. Lucke, A. Last, J. Mohr, P. Lehmann. 2005. Confocal micro-optical distance sensor. Principle and design. In . SPIE Proceedings. SPIE, 128); doi: 10.1117/12.612008.
[28] R. E. Bedford and G. Wyszecki, *J. Opt. Soc. Am.* 6, 564_1, (1957); doi: 10.1364/JOSA.47.0564_1.
[29] W. Sellmeier. Zur Erklärung der abnormen Farbenfolge im Spektrum einiger Substanzen. In *Annalen der Physik*, 272–282.
[30] Y. Shimizu, T. Maruyama, S. Nakagawa, Y.-L. Chen, H. Matsukuma, W. Gao, *Meas. Sci. Technol.* 7, 74006, (2018); doi: 10.1088/1361-6501/aac0a6.

DOI 10.5162/sensoren2019/P1.15

Mittels aerosolbasierter Kaltabscheidung bei Raumtemperatur hergestellte schichtförmige NTC-Thermistorbauelemente

Michaela Schubert[1], Christian Reichl[1], Christian Münch[2], Jaroslaw Kita[1], Ralf Moos[1]
[1]Lehrstuhl für Funktionsmaterialien, Universität Bayreuth, 95440 Bayreuth, Germany
[2]Vishay Electronic GmbH, Dr.-Felix-Zandman-Platz 1, 95100 Selb, Germany
funktionsmaterialien@uni-bayreuth.de; Tel.: 49 (0)921 55 7401

Zusammenfassung

In der vorliegenden Arbeit konnte gezeigt werden, dass eine aerosolbasierte Kaltabscheidung eines produktionsnahen NTCR-Keramikpulvers zu mechanisch stabilen Schichten möglich ist. Die Schichten zeichnen sich durch eine dichte nanokristalline Schichtstruktur und eine sehr gute Substratanbindung aus. Im ersten Versuchsteil konnte sowohl auf Edelstahl- als auch auf Al_2O_3-Substraten ein linearer Zusammenhang zwischen der erzielten Schichtdicke der aerosolkaltabgeschiedenen NTCR-Schicht und der Anzahl an Überfahrtsschleifen nachgewiesen werden. Basierend auf diesen Erkenntnissen wurden im zweiten Versuchsteil chipbasierte NTCR-Bauteile der Baugröße 1206 mit 10 und 20 Überfahrtsschleifen hergestellt. Die aerosolbasierte Kaltabscheidung des produktionsnahen Keramikpulvers erfolgte hierzu auf ein mit Sollbruchkanten und siebgedruckten Au-Interdigitalelektrodenstrukturen versehenes Al_2O_3-Substrat. Die Bauteile zeigen reproduzierbar die typische Charakteristik von konventionell gesinterten NTCR-Volumenkeramiken. Auch konnte der Nachweis erbracht werden, dass durch die Variation der Anzahl an Überfahrtsschleifen die Schichtdicke der aerosolkaltabgeschiedenen NTCR-Schicht und damit deren elektrische Widerstand gezielt eingestellt werden kann. Die charakteristische *B*-Konstante ändert sich dabei nicht.

Keywords: Aerosol Deposition Method (ADM), NTC Thermistor, Temperatursensor, Dickschicht

Einleitung

Bei NTC-Thermistoren (Negative Temperature Coefficient Thermally Sensitive Resistor [1]) handelt es sich größtenteils um halbleitende Keramiken auf Basis von Übergangsmetalloxiden, meist Übergangsmetallmanganaten mit Spinellstruktur $MeMn_2O_4$ (Me = Ni, Co, Fe, Cu, etc.) [2]. Diese zeichnen sich durch einen hohen negativen Temperaturkoeffizienten, geringe Kosten und eine hohe Zuverlässigkeit aus [3,4]. Angewendet werden NTC-Thermistoren vor allem zur präzisen Temperaturmessung und -kontrolle. Die große Temperaturabhängigkeit des elektrischen Widerstandes *R* der spinellbasierten NTC-Thermistoren basiert auf einem phononengestützten Hüpfmechanismus von Ladungsträgern entlang lokalisierter Zustände, dem sog. Small Polaronen Hopping [5]. Das *R*-*T*-Verhalten kann vereinfacht über eine Arrhenius-Beziehung (Gl. (1)) beschrieben werden.

$$R(T) = R_0 \cdot \exp(E_A/(k_B \cdot T)) \quad (1)$$

R_0 stellt dabei den elektrischen Widerstand bei unendlich hoher Temperatur, E_A die Aktivierungsenergie für den Hüpfprozess, k_B die Boltzmann-Konstante und *T* die Temperatur dar [6,7]. Aus dem Verhältnis E_A/k_B ergibt sich die Thermistorkonstante *B*. Die *B*-Konstante und der elektrische Widerstand bei 25 °C sind charakteristische Kenngrößen für die Beschreibung und den Vergleich von NTC-Thermistoren. Derzeit erfolgt die Herstellung der keramischen NTC-Thermistoren über traditionelle, sinterbasierte Verfahren bei Temperaturen oberhalb von 1000 °C [3]. Diese Verfahren sind jedoch hinsichtlich Flexibilität, Integrationsfähigkeit in elektronische Schaltungen und Miniaturisierbarkeit begrenzt [8].

Eine vielversprechende Alternative stellt das Verfahren der aerosolbasierten Kaltabscheidung (engl. Aerosol Deposition, AD) dar. Bei diesem neuartigen Beschichtungsverfahren können bei Raumtemperatur dichte keramische Schichten direkt aus einem keramischen Ausgangspulver auf unterschiedlichste Substratmaterialien (Glas, Keramik, Polymer, Me-

tall) aufgebracht werden [9,10]. Allgemein zeichnen sich die aerosolkaltabgeschiedenen Schichten durch eine hervorragende Substratanbindung, eine hohe Dichtigkeit und Funktionseigenschaften aus, die mit denen der klassischen Volumenkeramiken vergleichbar sind. Ein schematischer Aufbau einer Anlage für die aerosolbasierte Kaltabscheidung ist in der nachfolgenden Abb. 1 dargestellt.

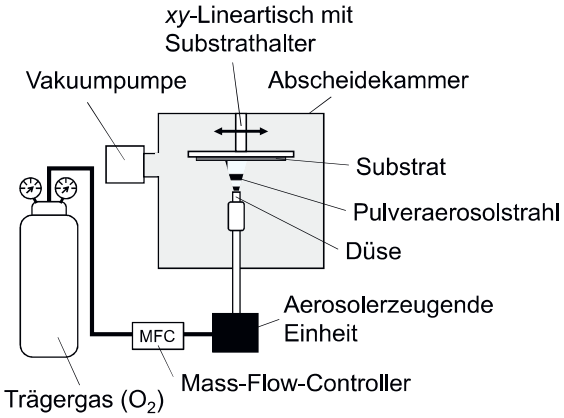


Abb. 1: Prinzipskizze einer Anlage für die aerosolbasierte Kaltabscheidung

Eine Anlage besteht grundsätzlich aus drei Hauptkomponenten - einer aerosolerzeugenden Einheit, einer Vakuumpumpe und einer Abscheidekammer. Während des Prozesses wird in der Abscheidekammer ein Vakuum erzeugt. In der aerosolerzeugenden Einheit wird das Ausgangspulver mit einem Trägergas (hier Sauerstoff) vermischt und ein Partikelaerosol erzeugt. Das generierte Partikelaerosol wird durch eine sich einstellende Druckdifferenz in die Abscheidekammer transportiert. Dabei passiert das Partikelaerosol eine Düse (hier Schlitzdüse 10 mm x 0,5 mm) und wird auf mehrere hundert Meter pro Sekunde beschleunigt. Die Partikel aus dem Aerosol treffen in der Abscheidekammer auf ein Substrat und bilden nach dem RTIC-Mechanismus (Room Temperature Impact Consolidation [10]) eine festanhaftende keramische Schicht aus.

Bisherige Untersuchungen an einem $NiMn_2O_4$-Modellpulver konnten bereits zeigen, dass eine Abscheidung von dichten keramischen Schichten mit typischer NTCR-Charakteristik möglich ist [11-13]. Für eine industrielle Umsetzbarkeit des Verfahrens ist es jedoch erforderlich, dass auch industrienahe Pulver abgeschieden werden können. Dadurch könnte das Verfahren ohne große Änderungen in den Produktionsprozess integriert werden.

Im Folgenden wird daher untersucht, ob ein produktionsnahes Pulver mittels der aerosolbasierten Kaltabscheidung zu mechanisch stabilen Schichten mit typischer NTCR-Charakteristik abgeschieden werden kann. Auch wird geklärt, ob eine gezielte Einstellung der Schichtdicke und damit des elektrischen Widerstandes möglich ist.

Experimentelles

Für die Untersuchung kam ein produktionsnahes Ni-Mn-Oxidpulver zum Einsatz. Um das Abscheideverhalten zu verbessern wurde das keramische Pulver durch Mahlen, Trocknen und Sieben AD-typisch präpariert. Das 30-minütige Pulvermahlen erfolgte bei 400 min^{-1} mit Hilfe einer Kugelmühle (Pulverisette 5, Fritsch). Für die Mahlung kamen Mahltiegel und -kugeln (Ø 10 mm) aus stabilisierten Zirkoniumdioxid und die Mahlflüssigkeit Cyclohexan (Sigma-Aldrich) zum Einsatz. Das Pulver wurde bei 200 °C in einem elektrisch betriebenen Umluftofen (UF110, Memmert) getrocknet. Für die Siebung wurde ein Sieb mit einer Maschenweite von 90 µm eingesetzt.

Im Anschluss erfolgte die aerosolbasierte Kaltabscheidung des präparierten produktionsnahen Keramikpulvers auf Edelstahl- und Al_2O_3-Substraten. Um den Einfluss der Anzahl an Überfahrten des Substrathalters auf die Schichtdicke *z* zu analysieren, wurde bei der aerosolbasierten Kaltabscheidung die Anzahl der Überfahrtsschleifen $N_{Schleife}$ (zwei Überfahrten bilden eine Überfahrtsschleife) im Bereich von 1 bis 20 variiert. Die Charakterisierung der hergestellten Schichten fand hinsichtlich der mechanischen Schichtstabilität, der Schichtmorphologie und der Schichtdicke mittels Tape-Test, REM (Leo 1450 VP, Zeiss) und Tastschnittverfahren (Perthometer, PGK, S2, Mahr) statt.

Um zu untersuchen, ob die aerosolkaltabgeschiedenen Schichten die typische NTCR-Charakteristik aufweisen und ob über eine Variation der Überfahrtsschleifen $N_{Schleife}$ der elektrische Widerstand bei 25 °C R_{25} gezielt eingestellt werden kann, wurden im zweiten Teil der Studie chipbasierte NTCR-Bauteile der Größe 1206 (1,6 mm x 3,2 mm) mit 10 und 20 Überfahrtsschleifen hergestellt. Die aerosolbasierte Kaltabscheidung des präparierten produktionsnahen Keramikpulvers erfolgte hierzu auf mit Sollbruchkanten und siebgedruckten Au-Interdigitalelektrodenstrukturen versehenen Al_2O_3-Substraten. Die Kontakte wurden während des Abscheidens mittels einer papierbasierten Einwegmaske abgedeckt. Nach dem Abscheiden fanden eine manuelle Vereinzelung der Bauteile sowie eine elektrische Charakterisierung im Temperaturbereich von 25 °C

bis 85 °C statt. Die elektrische Messung erfolgte in einem Thermostatbad (SL-12, Julabo) mit einem niederviskosen Silikonöl (200 Fluid, 5 cst, Dow Corning) mittels Vierleitermesstechnik und unter Einsatz eines Digitalmultimeters (2700, Keithley). Aus den ermittelten *R-T*-Daten wurden anschließend die charakteristischen NTCR-Kenngrößen, der elektrische Widerstand bei 25 °C R_{25} und die Thermistorkonstante *B* (Gl. (1)) bestimmt.

Ergebnisse

Die Untersuchung konnte zeigen, dass die aerosolbasierte Kaltabscheidung eines produktionsnahen Keramikpulvers zu mechanisch stabilen Schichten möglich ist (keine Ablösung durch Tape-Test). Die Schichten zeichnen sich, wie in Abb. 2 ersichtlich, durch eine dichte, nanokristalline Schichtstruktur und eine hervorragende Substratanbindung aus.

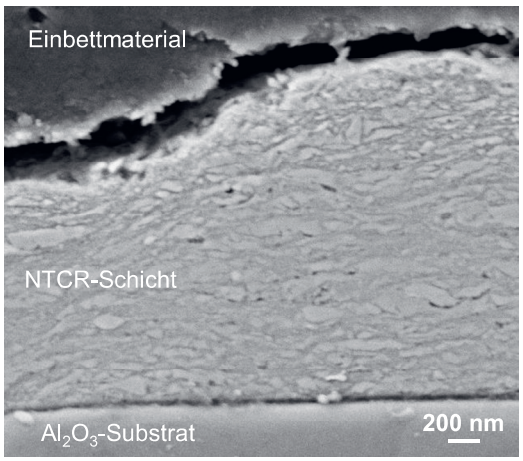


Abb. 2: Aerosolkaltabgeschiedene NTCR-Schicht auf einem Al_2O_3-Substrat im Schliff

Auch konnte sowohl auf Edelstahl- als auch auf Al_2O_3-Substraten ein linearer Zusammenhang zwischen der Anzahl an Überfahrtsschleifen $N_{Schleife}$ und der Schichtdicke *z* nachgewiesen werden (Abb. 3). Darüber hinaus konnte gezeigt werden, dass auf weicheren Substraten höhere Schichtdicken *z* erzielt werden können. So liegen die Schichtdicken der aerosolkaltabgeschiedenen NTCR-Schichten beim Edelstahlsubstrat bei 10 Überfahrtsschleifen im Bereich von 25 µm und bei 20 Überfahrtsschleifen sogar im Bereich von 50 µm. Im Gegensatz dazu beträgt die Schichtdicke der aerosolkaltabgeschiedenen NTCR-Schicht beim deutlich härteren Al_2O_3-Substrat bei 10 Überfahrtsschleifen nur etwa 2,5 µm und bei 20 Überfahrtsschleifen nur etwa 6 µm.

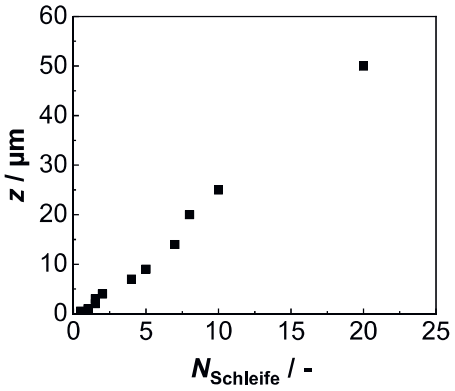


Abb. 3: Schichtdicke z in Abhängigkeit der Anzahl an Überfahrtsschleifen $N_{Schleife}$ auf Edelstahlsubstraten

Basierend auf diesen Erkenntnissen, wurden im zweiten Teil der Arbeit chipbasierte NTCR-Bauteile mit 10 und 20 Überfahrtsschleifen hergestellt. Hierbei konnte nachgewiesen werden, dass mittels der aerosolbasierten Kaltabscheidung des produktionsnahen Keramikpulvers erfolgreich Chipbauteile der Baugröße 1206 hergestellt werden können. Wie in Abb. 4 zu erkennen, sind die abgeschiedenen NTCR-Schichten auch nach der Demaskierung und der manuellen Vereinzelung der Bauteile stabil und zeigen keinerlei Risse oder Abplatzungen. Auch liegen die Kontaktpads vollständig unbeschichtet vor, was eine erfolgreiche Maskierung während der aerosolbasierten Kaltabscheidung bestätigt. Die elektrische Charakterisierung der Bauteile zeigt, dass diese reproduzierbar das typische NTCR-Verhalten eines annähernd exponentiellen Abfalls des elektrischen Widerstandes mit steigender Temperatur aufweisen.

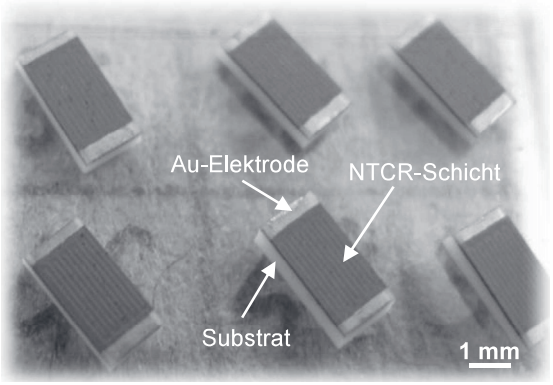


Abb. 4: Chipbasierte NTCR-Bauteile der Baugröße 1206 nach der manuellen Vereinzelung

Ein Vergleich der NTCR-Kenngrößen R_{25} und *B* von Bauteilen, die mit 10 und 20 Überfahrtsschleifen $N_{Schleife}$ hergestellt wurden, ist in Abb.

5 dargestellt. Es zeigt sich, dass der *B*-Wert in Abb. 5 a) erwartungsgemäß unabhängig von der Anzahl an Überfahrtsschleifen $N_{Schleife}$ ist. Mit 3865 K liegt der *B*-Wert zudem reproduzierbar im Bereich derer von konventionell gesinterten Bauteilen. Im Gegensatz dazu zeigt sich beim R_{25} in Abb. 5 b) ein deutlicher Einfluss der Anzahl an Überfahrtsschleifen $N_{Schleife}$. So verringert sich der R_{25} reproduzierbar von etwa 35 kΩ bei 10 Überfahrtsschleifen auf 14 kΩ bei 20 Überfahrtsschleifen. Die Abnahme des elektrischen Widerstandes *R* mit einer steigenden Anzahl an Überfahrtsschleifen ist gemäß der Beziehung $R \sim 1/z$ auf eine Zunahme der Schichtdicke *z* der aerosolkaltabgeschiedenen NTCR-Schicht zurückzuführen. So weisen die Bauteile, die auf Basis von 10 Überfahrtsschleifen hergestellt wurden, eine NTCR-Schichtdicke im Bereich von 2,5 µm und die Bauteile auf Basis von 20 Überfahrtsschleifen eine NTCR-Schichtdicke im Bereich von 6 µm auf.

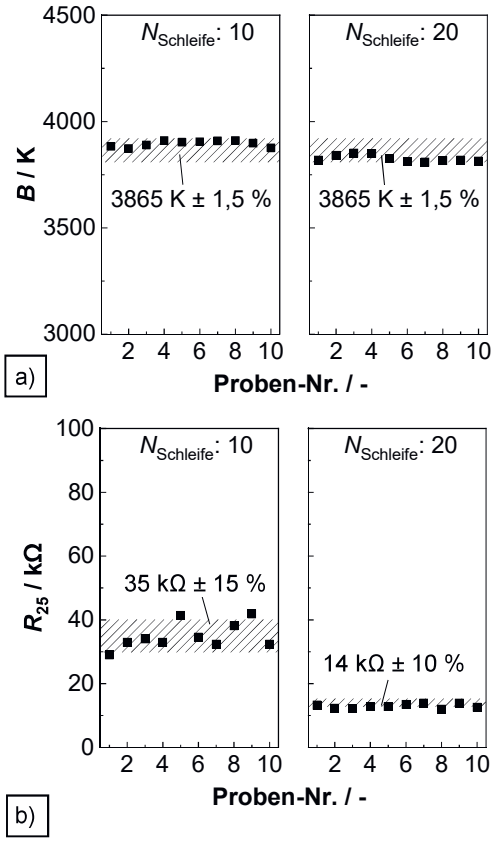


Abb. 5: a) Ermittelter B-Werte und b) R_{25}-Werte von chipbasierten Bauteilen in Abhängigkeit der Anzahl an Überfahrtsschleifen $N_{Schleife}$

Hervorzuheben ist hierbei, dass die erzielten Schichtdicken der aerosolkaltabgeschiedenen Schicht sehr gut mit den Schichtdicken der aerosolkaltabgeschiedenen NTCR-Schichten auf den Al_2O_3-Substraten aus Versuchsteil eins übereinstimmen. Folglich ist es möglich durch eine Variation der Anzahl an Überfahrtsschleifen gezielt die Schichtdicke der aerosolkaltabgeschiedenen NTCR-Schicht und damit deren Widerstand einzustellen.

Zusammenfassung

In der vorliegenden Arbeit konnte nachgewiesen werden, dass eine aerosolbasierte Kaltabscheidung eines produktionsnahen Keramikpulvers zu mechanisch stabilen Schichten mit typischer NTCR-Charakteristik möglich ist. Sowohl auf unstrukturierten Edelstahl- und Al_2O_3-Substraten als auch auf strukturierten Al_2O_3-Substraten konnte die Schichtdicke der aerosolkaltabgeschiedenen Schicht und damit der elektrische Widerstand gezielt über die Anzahl an Überfahrtsschleifen variiert und damit eingestellt werden.

Literaturnachweis

[1] M. Parlak, T. Hashemi, M. J. Hogan, A. W. Brinkman, *Thin Solid Films* 345, 307-311 (1999); doi: 10.1016/S0040-6090(98)01426-6
[2] K. Park, D. Y. Bang, *J. Mater. Sci.: Mater. Electron.* 14, 81-87 (2003); doi: 10.1023/A:1021900618988
[3] A. Feteira, *J. Am. Ceram. Soc.* 92, 967-983 (2009); doi: 10.1111/j.1551-2916.2009.02990.x
[4] G. D. C. Csete de Györgyfalva, I. M. Reaney, *J. Eur. Ceram. Soc.*, 21, 2145-2148 (2001), doi: 10.1016/S0955-2219(01)00190-X
[5] R. Schmidt, A. Basu, A. W. Brinkman, *Phys. Rev. B*, 72, 115101 (2005), doi: 10.1103/PhysRevB.72.115101
[6] P. Fau, J. P. Bonino, J. J. Demai, A. Rousset, *Appl. Surf. Sci.* 65/66, 319-324 (1993); doi: 10.1016/0169-4332(93)90679-6
[7] A. Rousset, R. Legros, A. Lagrange, *J. Eur. Ceram. Soc.* 13, 185-195 (1994); doi: 10.1016/0955-2219(94)90027-2
[8] A. Darmawansyah*, Prosiding SENTIA 2009,* 12.03.2009, Malang, Indonesia
[9] D. Hanft, J. Exner, M. Schubert, T Stöcker, P. Fuierer, R. Moos, *J. Ceram. Sci. Technol.* 2015, *6*, 147-182; doi: 10.4416/JCST2015-00018
[10] J. Akedo, *J. Therm. Spray Technol.* 2008, *17*, 181-198; doi: 10.1007/s11666-008-9163-7
[11] M. Schubert, C. Münch, S. Schuurman, V. Poulain, J. Kita, R. Moo, *J. Eur. Ceram. Soc.* 38, 613-619 (2018); doi: 10.1016/j.jeurceramsoc.2017.09.005
[12] M. Schubert, C. Münch, S. Schuurman, V. Poulain, J. Kita, R. Moos, *Sensors* 18, 3982 (2018); doi: 10.3390/s18113982
[13] M. Schubert, J. Kita, C. Münch, R. Moos, *Funct. Mater. Lett.* 10, 1750073 (2017); doi: 10.1142/S1793604717500734

DOI 10.5162/sensoren2019/P1.16

Vergleichende Betrachtung verschiedener Methoden zur Bestimmung der Inhomogenität von Thermoelementen

Philipp Germanow[1], Patrick Mehring[1], Herbert Neumann[1],
Silke Augustin[2], Thomas Fröhlich[2], Helge Mammen[2]
[1]Reckmann GmbH, Werkzeugstraße 19-23, D-58093 Hagen
[2]TU Ilmenau, Gustav-Kirchhoff-Str. 1, D-98693 Ilmenau

Zusammenfassung

Die Homogenität des Thermomaterials nimmt Einfluss auf die Kalibrierfähigkeit und die im Betrieb erreichbare Genauigkeit eines Thermoelementes. Eine Unstetigkeit im Thermomaterial, nachfolgend als Inhomogenität bezeichnet, wird jedoch schon während des regulären Einsatzes des Thermoelementes durch mechanische und thermische Beanspruchungen ausgebildet. Um nun Aussagen über den dadurch entstandenen Messfehler zu treffen, muss die Inhomogenität erfasst werden. Dies wird im Allgemeinen durch zwei unterschiedliche Messverfahren möglich. Im Rahmen der Einbindung einer neuen Messmethode zur Bestimmung der Inhomogenität in einem bestehenden Kalibrierlabor wurden Untersuchungen beider Verfahren vorgenommen. Die daraus entstandenen Versuchsanlagen beider Funktionsprinzipe und die dort aufgenommenen Inhomogenitätsprofile werden vergleichend gegenübergestellt. Abschließend erfolgt eine Einordnung der Messverfahren anhand ihrer Zweckmäßigkeit im bestehenden Kalibrierbetrieb.

Keywords: Thermoelement, Inhomogenität, Ein-Gradienten-Verfahren, Zwei-Gradienten-Verfahren, Messunsicherheitsbeitrag

Entstehung von Inhomogenitäten

Die Drift eines Thermoelementes, also die Veränderung der erzeugten Thermospannung im Laufe der Betriebszeit, geht stets von einer Änderung des Seebeck-Koeffizienten aus. Letztere sorgt dabei für die Ausprägung von Inhomogenitäten entlang des Thermoelementes. Eine Inhomogenität entsteht beispielsweise durch die Veränderung der chemischen Zusammensetzung, eine Veränderung des metallurgischen Zustandes oder das Einwirken äußerer mechanischer Spannungen [1].

Auswirkung von Inhomogenitäten

Im Gegensatz zu dem Missverständnis der alten Thermoelement-Theorie resultiert die erzeugte Thermospannung nicht aus der Temperaturdifferenz zwischen Mess- und Vergleichsstelle. Sie entsteht vielmehr durch den Temperaturgradienten und den lokalen Seebeck-Koeffizienten $S(T,x)$ entlang des Thermoelementes [1]:

$$\mathrm{U} = \int_0^L S(T,x)\,\frac{\partial T}{\partial x}\,dx$$

Aufgrund dieses Sachverhaltes wirkt sich die Inhomogenität deutlich auf die, in Abhängigkeit von der jeweiligen Einbausituation, erfasste Thermospannung des Sensors aus. Bei einem inhomogenen Thermoelement ist daher die erzeugte Thermospannung verschiedener Einbausituation, beispielsweise während der Kalibrierung und bei Inbetriebnahme, nicht direkt vergleichbar. Der Messfehler, ausgelöst durch die Inhomogenität des Sensors, muss daher zwingend in der Messunsicherheit der Kalibrierung des Sensors berücksichtigt werden.

Erfassung von Inhomogenitäten

Die Erfassung von Inhomogenitätsprofilen basiert grundsätzlich auf einem bekannten Temperaturgradienten, welcher über das Thermoelement bewegt wird. Der Temperaturgradient resultiert dabei aus dem axialen Temperaturverlauf entlang des zu untersuchenden Sensors. Auf Basis der lokal unterschiedlichen Seebeck-Koeffizienten entlang des Thermoelementes sorgt eine Verschiebung des Temperaturgradienten für eine Veränderung der erzeugten Thermospannung. Wenn die Thermospannung nun an diskreten Punkten entlang des Sensors aufgenommen wird, folgt daraus das

sogenannte Inhomogenitätsprofil des Sensors. Dabei ist die erfasste Thermospannung ein Maß für die Variation des Seebeck-Koeffizienten des Sensors.
Anhand der ausgebildeten Temperaturgradienten kann dabei in Ein-Gradienten- und Zwei-Gradienten-Messverfahren zur Erfassung von Inhomogenitätsprofilen unterschieden werden [2].

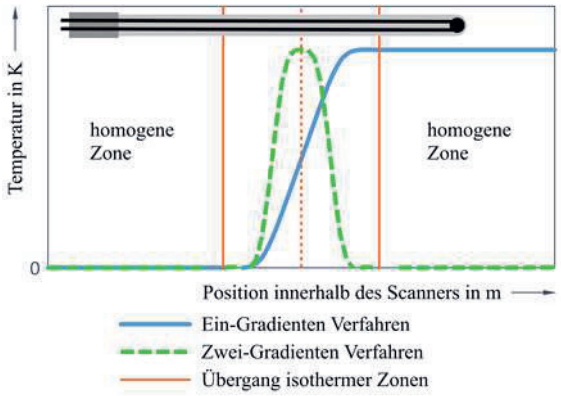


Abb. 1: Schematischer Temperaturverlauf des Ein- und Zwei-Gradienten Verfahrens [3]

Beim Ein-Gradienten-Verfahren wird ein Thermoelement von einer kalten homogenen Zone in eine heiße homogene Zone eingetaucht, beziehungsweise aus ihr herausgezogen. Mit dem Übergang zwischen den homogenen Zonen ergibt sich der in Abb. 1 dargestellte Temperaturverlauf entlang des Thermoelementes.
Im Gegensatz dazu wird beim Zwei-Gradienten-Verfahren nur eine lokal begrenzte Hitzequelle über das Thermoelement hinweg bewegt. Dabei bleiben die Mess- und die Vergleichsstelle des Sensors auf derselben Temperatur (vgl. Abb. 1).

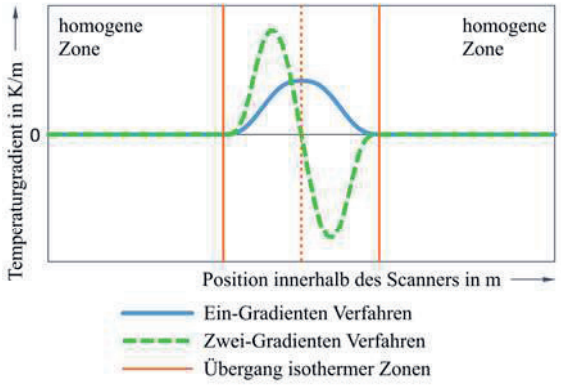


Abb. 2: Schematischer Temperaturgradient des Ein- und Zwei-Gradienten-Verfahrens [3]

Der Übergang zwischen den isothermen Zonen erfolgt dabei nicht sprunghaft. An den Grenzen der jeweiligen isothermen Zonen beider Messverfahren resultieren stets, die in Abb. 2 dargestellten, Temperaturgradienten. Der Temperaturgradient ist dabei ein Maß dafür, wie schnell die Temperatur entlang des zu untersuchenden Thermoelementes ansteigt. Hier ist jedoch stets darauf zu achten, dass der tatsächlich wirkende Temperaturgradient, auch bei idealer Abgrenzung der homogenen Zonen in der Versuchseinrichtung, stark von der Wärmeleitung entlang des Thermoelementes bestimmt wird [4].

Messungen im Ein-Gradienten-Verfahren

Die im Ein-Gradienten-Verfahren erfasste Thermospannung resultiert einzig aus dem einfachen Temperaturgradienten zwischen beiden homogenen Zonen. Dementsprechend beschreibt die erfasste Thermospannung hier direkt den innerhalb des Gradienten erfassten Seebeck-Koeffizienten (vgl. Abb. 3).

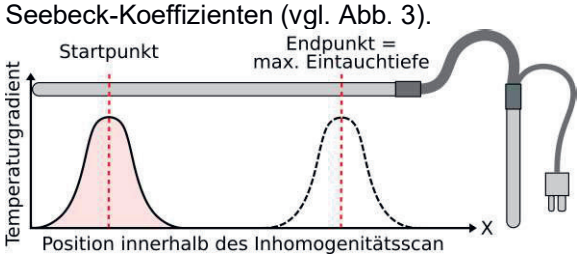


Abb. 3: Verschiebung des Temperaturgradienten innerhalb des Ein-Gradienten-Verfahrens

Um die Veränderung des lokalen Seebeck-Koeffizienten entlang des Thermoelementes dabei möglichst hoch aufzulösen, muss ein schmaler Temperaturgradient entlang des Sensors verschoben werden. Die Breite des Temperaturgradienten beschreibt dabei den Bereich, in dem die Variation des lokalen Seebeck-Koeffizienten erfasst werden kann. Je schmaler der Gradient ist, desto mehr Thermospannung resultiert aus einem kurzen Teilbereich des Thermoelementes. Lokal stark begrenzte Inhomogenitäten werden in breiten Temperaturgradienten nur in Verbindung mit homogenen Teilbereichen des Thermomaterials erfasst. Wird der Temperaturgradient daraufhin geschärft, verbessert sich die räumliche Auflösung des Scanners soweit bis der inhomogene Teilabschnitt des Thermoelementes vollständig erfasst wird.
Um eine geeignete Versuchsanlage zur Bestimmung der Inhomogenität zu entwickeln, muss der Temperaturgradient messtechnisch erfasst und dessen Breite rechnerisch

bestimmt werden. Nach der von Webster et al. [4] beschriebenen Vorgehensweise kann die maximale Auflösung für die nachfolgend vermessenen Typ S Thermoelemente mit 6 mm Keramikschutzrohr anhand des Temperaturverlaufs in der Versuchsanlage bestimmt werden. Dabei wird der Temperaturverlauf im Inhomogenitätsscanner unter Berücksichtigung der Wärmeleitung erfasst und in den wirkenden Gradienten umgewandelt. Der Temperaturgradient kann mathematisch durch die Ableitung des Temperaturverlaufs entlang des Thermoelementes bestimmt werden.

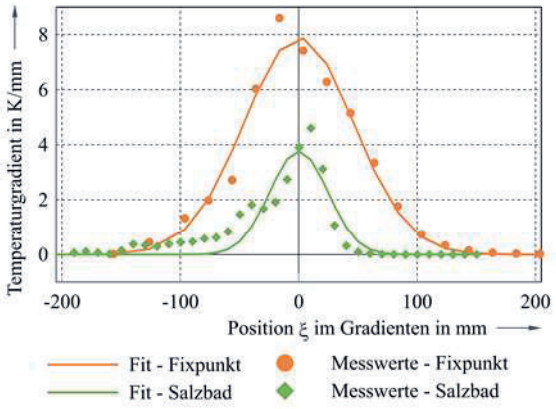


Abb. 4: Vergleich der Temperaturgradienten realer Ein-Gradienten-Scanner aufgenommen mit Typ S Thermoelement in 6 mm Schutzkeramik; Approximation über Gaußfunktion

In Abbildung 4 sind die vorherrschenden Temperaturgradienten in zwei typischen Versuchsanlagen zur Ein-Gradienten-Inhomogenitätsmessung gegenübergestellt.
Die Halbwertsbreite der Gauß-Fit-Funktion beschreibt dabei nach Hill et al. [5] die maximale räumliche Auflösung des Scanners. Diese lag bei den verwendeten Prüflingen bei 25 mm im Salz-Kalibrierbad und 46 mm im Silber-Fixpunkt. Aufgrund der besseren räumlichen Auflösung kann das Kalibrierbad, im Vergleich zu dem verwendeten Fixpunkt, auch deutlich kleinere Inhomogenitäten erfassen.

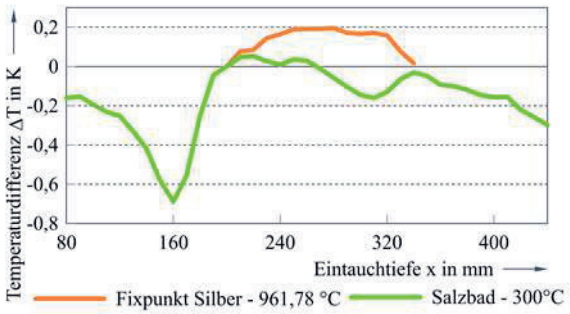


Abb. 5: Vergleich der Inhomogenitätsprofile eines Typ S Thermoelementes in 6 mm Schutzkeramik; Aufgenommen in verschiedenen Versuchsanlagen

Aus den Vergleichen verschiedener Kalibriereinrichtungen (in Abb. 4 und 5) geht hervor, dass sich vor allem Salz- und Öl-Kalibrierbäder zum Vermessen der Sensoren im Ein-Gradienten-Verfahren eignen.
Die direkten Auswirkungen der Gradientenbreite auf die erfassten Inhomogenitäten lassen sich in Abb. 5 erkennen. Die Untersuchung desselben Thermoelementes in mehreren Versuchsanlagen weist dabei deutliche Unterschiede im erfassten Inhomogenitätsprofil auf. Die lokal ausgeprägten Inhomogenitäten des vermessenen Thermoelementes werden hier nur vom Gradienten des Kalibrierbades erfasst. In der Silber-Fixpunktzelle werden diese lokal begrenzten Inhomogenitäten aufgrund des breiten Temperaturgradienten und des begrenzten Bauraums nicht erfasst (vgl. Abb. 5).



Abb. 6: Konstruktive Umsetzung des Ein-Gradienten-Inhomogenitätsscans am Salz-Kalibrierbad

Für die Untersuchungen der Homogenität von Thermoelementen bietet sich im Rahmen der verfügbaren Kalibrieranlagen ein Salz-Kalibrierbad mit großer Eintauchtiefe an (vgl. Abb. 6). Dieses wird zur automatisierten Umsetzung der Inhomogenitätsscans mit einer Linearachse und einer erweiterten Auswertesoftware ausgestattet.

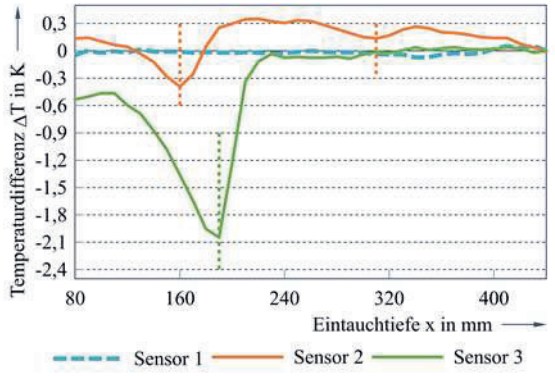


Abb. 7: Erfasste Inhomogenitäten diverser Typ S Thermoelemente mit 6mm Schutzkeramik im Salz-Kalibrierbad bei 300 °C

Über den vorgestellten Ein-Gradienten-Scanner wird die Aufnahme der Thermoelemente bis zu einer Länge von 440 mm möglich.
In Abbildung 7 sind die untersuchten Typ S Thermoelemente gegenübergestellt. Sensor 2 und 3 sind dabei ein exemplarisches Beispiel, welche Messfehler Inhomogenitäten im realen Betrieb erzeugen können. Die auf den Endwert bei maximaler Eintauchtiefe normierten Inhomogenitätsprofile zeigen Abweichungen von bis zu 2,07 K.

Messungen im Zwei-Gradienten-Verfahren

Beim Zwei-Gradienten-Verfahren wird das Thermoelement mit einer lokal begrenzten Hitzequelle überstrichen, während die Mess- und Vergleichsstelle auf einer konstanten Temperatur gehalten werden. Dadurch entstehen zwei gegenläufige Gradienten auf dem Thermoelement (vgl. Abb. 8) [3].

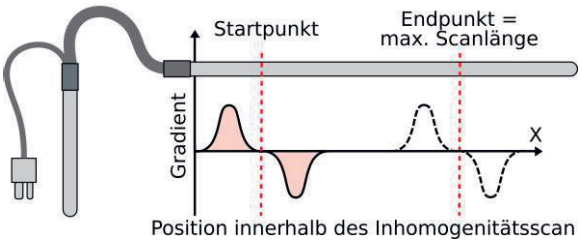


Abb. 8: Verschiebung der Temperatur-gradienten innerhalb des Zwei-Gradienten-Verfahrens

In den beiden gegenläufigen Gradienten entsteht jeweils eine Teil-Thermospannung. Bei einem homogenen Thermoelement sind die Thermospannungen in beiden Gradienten gleich groß, die gemessene Thermospannung beim Überfahren des Sensors ist gleich null. Während der Messung eines inhomogenen Thermoelementes sorgt der Unterschied zwischen den lokalen Seebeck-Koeffizienten in beiden Gradienten für eine von null abweichende Thermospannung. Diese Thermospannung ist dabei ein Maß für den Unterschied der beiden lokalen Seebeck-Koeffizienten innerhalb der gegenläufigen Gradienten. Dementsprechend kann die erfasste Thermospannung auch nicht direkt in eine äquivalente Temperatur umgerechnet werden.

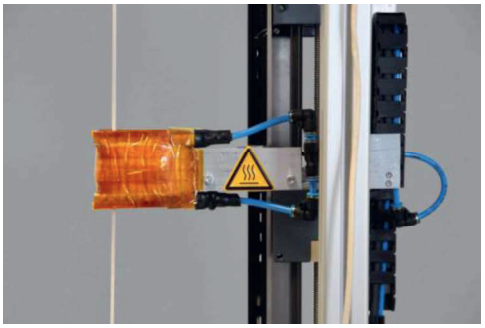


Abb. 9: Nahaufnahme des Zwei-Gradienten-Scanners

Um einen Zwei-Gradienten Scan durchführen zu können, wird eine räumlich stark begrenzte Hitzequelle mit einem klar definierten Temperaturfeld benötigt. Dieses kann beispielsweise durch einen nach außen isolierten Widerstandsheizer mit zusätzlicher Luftkühlung zur Schärfung der Gradienten erzeugt werden. Um automatisierte Inhomogenitätsprofile analog zum Ein-Gradienten-Verfahren aufnehmen zu können, wird dieser an einer geeigneten Lineareinheit angebracht und über das zu untersuchende Thermoelement gefahren (vgl. Abb. 9). Der untersuchte Sensor wird hingegen in der Messeinrichtung fixiert.

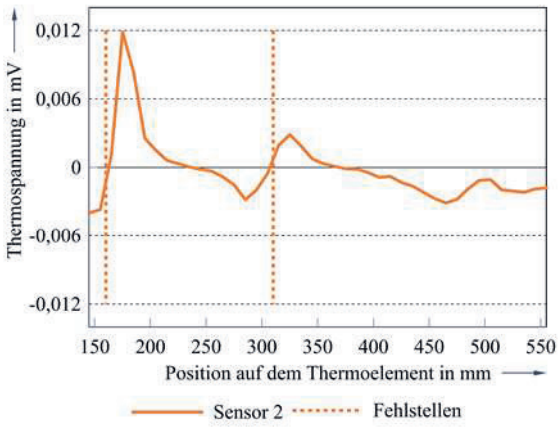


Abb. 10: Erfasstes Inhomogenitätsprofil des Sensors 2 (Typ S mit 6 mm Schutzkeramik) im Zwei-Gradienten-Verfahren bei 300 °C

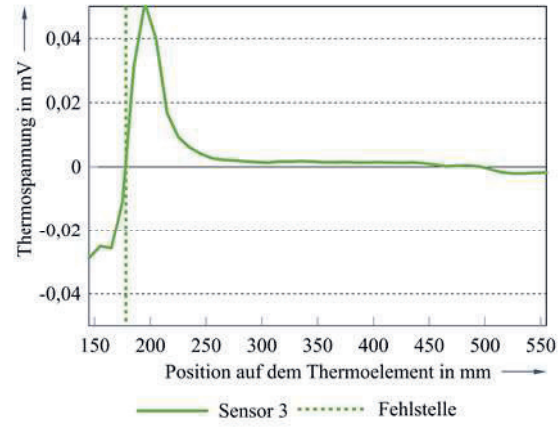


Abb. 11: Erfasstes Inhomogenitätsprofil des Sensors 3 (Typ S mit 6 mm Schutzkeramik) im Zwei-Gradienten-Verfahren bei 300 °C

Aufgrund der zwei gegenläufigen Gradienten folgt bei diesem Versuchsaufbau ein vollkommen unterschiedlicher Verlauf zu der im Ein-Gradienten-Verfahren erfassten Thermospannung. Einzelne Fehlstellen äußern sich im Zwei-Gradienten-Scan durch zwei charakteristische Peaks und einen Nulldurchgang im Inhomogenitätsprofil. (vgl. Abb.10 und 11).
Dieser charakteristische Verlauf resultiert einzig aus den gegenläufigen Temperaturgradienten. Wenn eine Inhomogenität im Zwei-Gradienten-Verfahren erfasst wird, so ändert sich erst die Teil-Thermospannung im ersten Gradienten. Anschließend erreicht die Inhomogenität die Mitte der Hitzequelle und trägt nicht zur erfassten Thermospannung bei. Erst beim Verlassen des Heizers wird die Inhomogenität wieder erfasst und erzeugt im gegenlaufenden Gradienten einen zweiten, umgekehrten Peak im Inhomogenitätsprofil. Die anhand der ausgeprägten Verläufe im Inhomogenitätsprofil erfassten Fehlstellen der Sensoren 2 und 3 korrelieren dabei mit den Messungen im Ein-Gradienten-Verfahren (vgl. Abb. 7, 10 und 11).
Die weitere Auswertung der erfassten Inhomogenitätsprofile ist hingegen schwierig, da die hier erfasste Thermospannung keine direkte Aussage über die tatsächlich ausgelösten Messfehler liefert. Um den realen Anwendungsfall eines Thermoelementes (näherungsweise in einem Gradienten) nachzuvollziehen, muss der hier erfasste Unterschied zwischen den Seebeck-Koeffizienten beider Gradienten umgerechnet werden. Eine direkte Umrechnung der Ein- und Zwei-Gradienten-Verfahren nach bekannten Methoden konnte, trotz vergleichbarer Temperaturgradienten und Absolut-temperaturen in beiden Versuchsanlagen, bei den bisher untersuchten Thermoelementen nicht erreicht werden.

Vergleich der verwendeten Versuchsanlagen und Messverfahren

Aus den entwickelten Verfahren und den Untersuchungen der Homogenität verschiedener Edelmetall-Thermoelemente gehen die jeweiligen Anwendungsbereiche der beiden Funktionsprinzipe hervor.
Dabei eignet sich der Ein-Gradienten-Inhomogenitätsscan aufgrund seines Temperaturverlaufs entlang des Thermo-elementes für direkte Aussagen über den lokalen Seebeck-Koeffizienten des Thermo-materials im Temperaturgradienten. In Abhängigkeit von der gewählten Einrichtung zur Erzeugung der homogenen Zone wird die zu untersuchende Länge des Thermo-elementes jedoch stark beschränkt. Des Weiteren wird in der Regel eine zusätzliche Kalibriereinrichtung mit hohen Anforderungen an die räumliche Homogenität der heißen Zone nötig.
Das Zwei-Gradienten-Verfahren zeichnet sich vor allem durch die nahezu unbegrenzte Länge des zu untersuchenden Sensors aus. Der Einsatzbereich des Zwei-Gradienten-Scanners wird dabei jedoch durch die Entstehung der Thermospannung in den gegenläufigen Gradienten begrenzt. Hier ist ohne weitere Anstrengungen keine direkte Aussage über die Ausprägung des lokalen Seebeck-Koeffizienten anhand der erfassten Thermospannung möglich.

Die Vergleichbarkeit der ermittelten Inhomogenitätsprofile ist jedoch unter gewissen Umständen gegeben. Wenn anstatt dem absoluten Betrag der Inhomogenität die Existenz und die Lage der Fehlstellen nachgewiesen werden soll, liefern Ein- und Zwei-Gradienten-Verfahren vergleichbare Ergebnisse. Die vollständige Beschreibung der ausgeprägten Inhomogenitäten wird aktuell jedoch nur im Ein-Gradienten-Verfahren möglich.

Literaturnachweis

[1] F. Bernhard, *Handbuch der Technischen Temperaturmessung*, 873 (2014); doi: 10.1007/978-3-642-24506-0

[2] Calibration Guide EURAMET cg-8, *Calibration of thermocouples*, Version 2.1, EURAMET e.V. Braunschweig, Oktober 2011

[3] M. Holmsten, J. Ivarson, R. Falk, M. Lisbeck, L.-E. Josefson, *Inhomogeneity Measurements of Long Thermocouples using a Short Movable Heating Zone*, Int. Journal of Thermophysics, (2008); doi: 10.1007/s10765-008-0418-9

[4] E. Webster, D. R. White, *Thermocouple homogeneity scanning*, Metrologia 52 (2015); doi: 10.1088/0026-1394/52/1/130

[5] K. D. Hill, D. J. Gee, *Quantifying the Calibration Uncertainity Attributable to Thermocouple Inhomogeneity*, AIP Conference Proceedings 1552, 520 (2013); doi: 10.1063/1.4819595

DOI 10.5162/sensoren2019/P1.18

Objektrekonstruktion in der Mikrowellen-Defektoskopie

Thomas Gagelmann[1], Maxim Nesterov[1], Sebastian Wöckel[1],
Cuihua Li[2], Sebastian Hantscher[2], Jörg Auge[2]
[1]Institut für Automation und Kommunikation, Werner-Heisenberg-Str. 1,
39106 Magdeburg, Deutschland
[2]Hochschule Magdeburg-Stendal, Breitscheidstr. 2, 39114 Magdeburg, Deutschland

Zusammenfassung

Dieser Beitrag stellt die Ergebnisse eines geförderten Projekts vor, dessen Ziel es war, die erreichbare Auflösung bei der Mikrowellen-Defektoskopie mit niedrigen Prüffrequenzen zu verbessern. Bei niedrigen Prüffrequenzen, wie die im Projekt verwendeten 24 GHz, wird eine größere Eindringtiefe der Mikrowellen ins Material erreicht. Es treten aber auch vermehrt Beugungseffekte auf, die die erreichbare Ortsauflösung verschlechtern und die Klassifizierung von Defekten erschweren.
Der im Artikel vorgestellte Ansatz verwendet eine der Messung nachgelagerte Analysetechnik, bei der die Messdaten mit einer zuvor ermittelten Transferfunktion entfaltet werden, um so die Beugungseffekte zu unterdrücken. Es wird gezeigt, dass mit dieser Technik für bestimmte Anwendungsfälle die Beugungsgrenze bei der Abbildung von kleinen, eng bei einander liegenden Defekten überwunden werden kann. Außerdem werden Defektkanten schärfer dargestellt, was eine genauere Klassifizierung der Defekte ermöglicht.

Keywords: zerstörungsfreie Werkstoffprüfung, Mikrowellen-Defektoskopie, Rekonstruktion, Punktspreizfunktion, Entfaltung

Einleitung

Der Markt für glasfaserverstärkte Kunststoffe wächst in Deutschland überdurchschnittlich stark. Deswegen werden kostengünstige und robuste zerstörungsfreie Prüfmethoden für diese Art dielektrischer Verbundmaterialien immer wichtiger um eine gleichbleibend gute Qualität und effiziente Produktion sicherzustellen. Die aktuell kommerziell verfügbaren zerstörungsfreien Prüfmethoden sind für glasfaserverstärkte Kunststoffe allerdings oft nur eingeschränkt geeignet. Zum Beispiel ist die Eindringtiefe bei der Thermografie auf wenige Millimeter begrenzt. Bei akustischen Prüfverfahren wird das Signal stark gedämpft und es treten Streueffekte an dem hohen Faseranteil im Material auf [1, 2]. Radiologische Prüfverfahren sind wiederum teuer, es werden zusätzliche Strahlenschutzvorkehrungen benötigt und die maximale Größe der Prüfobjekte ist oft begrenzt [3, 4]. Eine andere Methode zur Prüfung glasfaserverstärkter Kunststoffe, ist die Mikrowellen-Defektoskopie. Mit Ihr können zerstörungsfrei Defekte im Probekörper bildhaft dargestellt und klassifiziert werden. Die Wahl der Prüffrequenz hat dabei einen entscheidenden Einfluss auf die Eindringtiefe und die erreichbare Auflösung. Bei niedrigen Prüffrequenzen treten außerdem verstärkt Beugungserscheinungen an Inhomogenitäten im Material (Ecken, Kanten, Defekte) auf. In der Praxis stellen diese Beugungserscheinungen ein Problem dar, da sie insbesondere Defekte nahe von Ecken und Kanten maskieren können und die Klassifizierung der Defekte bezüglich Ihrer Größe Form und Tiefenlage erschweren [5-9]. Das bedeutet für die Qualitätskontrolle eine deutlich verringerte Zuverlässigkeit der Detektion von randnahen gegenüber randfernen Fehlern.
In der THz-Technik werden diese Probleme durch die Verwendung vergleichsweise hoher Prüffrequenzen umgangen. Bei dieser Technik treten Beugungsmuster von Ecken und Kanten im Vergleich zu niedrigen Frequenzen nur in abgeschwächter Form auf, dafür ist sie aufgrund des erhöhten technischen Aufwands allerdings auch immer noch recht teuer und die Eindringtiefe der hochfrequenten Mikrowellen in das Material ist aufgrund der starken Dämpfung begrenzt [10-13].
In diesem Zusammenhang bietet sich für die Detektion tiefliegender Defekte die Verwendung von niedrigeren Prüffrequenzen an. Im Projekt wurde eine Prüffrequenz von 24 GHz verwendet. Mit dieser Prüffrequenz können mit relativ geringem technischen Aufwand und einer geringen Sendeleistung größere Eindringtiefen in das dielektrische Material erzielt werden.
Der im Projekt verwendete Messaufbau ist in Abbildung 1 dargestellt. Bei einer Messung

bewegt sich die Hohlleiterantenne im gleichbleibenden Abstand von 4 mm zeilenweise über die Oberfläche des Prüfobjekts, bestrahlt es mit Mikrowellen einer konstanten Frequenz von 24 GHz (CW) und zeichnet den komplexen Reflexionsparameter S_{11} auf. Das Messsignal enthält das Interferenzmuster der Wellen, die an den Inhomogenitäten innerhalb des Probekörpers gestreut bzw. gebeugt wurden.

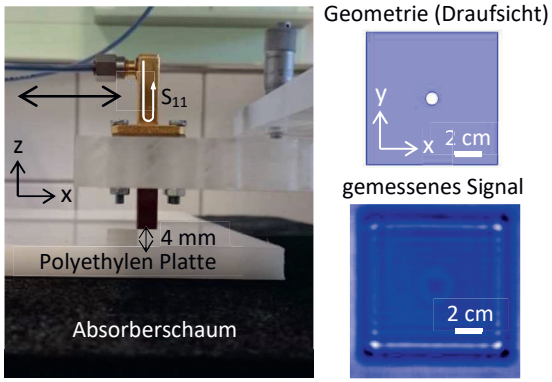


Abb. 1: Experimentelle Messanordnung mit der Hohlleiterantenne über dem Probekörper (links), ein Schema eines Probekörpers mit einer einzelnen Lochbohrung (oben rechts), und dem gemessenen Reflexionssignal $|S_{11}|$ (unten rechts).

In Abbildung 2 ist das Problem der durch die Beugungsmuster maskierten Defekte anhand einer Messung dargestellt, bei der vor allem der Defekt im rot markierten Bereich nur schwer erkennbar ist. Das Ziel des hier vorgestellten Projekts war es, eine Methode zu entwickeln, die trotz vorhandener Beugungserscheinungen eine klare Identifikation und Klassifikation von Defekten ermöglicht, deren Abmessungen klein gegenüber der Wellenlänge sind. Dafür müssen die Beugungseffekte möglichst unterdrückt werden, um so die Kantenschärfe der dargestellten Objekte zu erhöhen und sie in der korrekten Größe abzubilden. Dazu bedarf es einer Verfeinerung der Aufzeichnungstechnik, der Modellierung der Antenne sowie einer Bildverarbeitung bei der eine der Messung nachgelagerte Objektrekonstruktion durchgeführt wird um die Beugungsmuster zu reduzieren. Bei der Rekonstruktion wird davon ausgegangen, dass das gemessene Reflexionsprofil mit den Beugungsmustern h mathematisch der mit Gl. (1) beschriebenen Faltung des geometrischen Profils des Probekörpers f mit einer zunächst unbekannten Punktspreizfunktion (PSF) g entspricht, wobei mit $*$ der Faltungsoperator bezeichnet wird.

$$h = f * g \qquad \text{Gl. (1)}$$

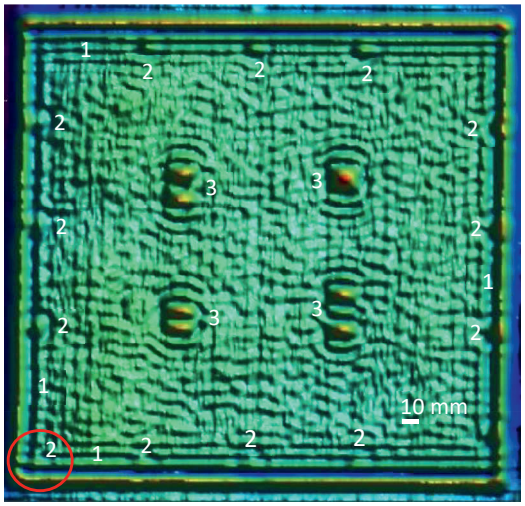


Abb. 2: Dargestellt ist der Realteil des Reflexionsparameters S_{11} vom 2D-Scan eines Probekörpers mit den an Geometriesprüngen entstehenden Beugungserscheinungen. Im Bild sind die verschiedenen Beugungserscheinungen mit Zahlen markiert. 1) Beugungseffekte an den Kanten des Probekörpers aus Polyethylen, 2) Beugungseffekte an mit Luft gefüllten Bohrungen von 3 mm Durchmesser 3) Beugungseffekte an mit Zylindern aus Aluminium gefüllten Bohrungen von 3 mm Durchmesser.

Zur Bestimmung der Objektkontur aus den Messdaten ist eine Entfaltung notwendig. Diese Entfaltungsmethode wird in der Optik, zum Beispiel in der Mikroskopie, bereits erfolgreich eingesetzt, um die Auflösung und Bildqualität zu verbessern [14]. In der Optik hängen die Eigenschaften und Qualität eines Bildes maßgeblich vom Fokus und dem Linsensystem des optischen Messsystems selbst und weniger vom abgebildeten Objekt ab. In der Mikrowellen-Defektoskopie hat aber neben dem Messsystem auch die Beschaffenheit des Prüflings einen entscheidenden Einfluss auf die abgebildeten Beugungsmuster, da sich der Prüfling im Nahfeld der Antenne befindet um eine hohe Auflösung zu erzielen. Deswegen wird im Projekt überprüft, inwiefern dieses Modell auch für die Mikrowellen-Defektoskopie anwendbar ist und welche Vor- und Nachteile sowie Limitationen bei der Entfaltungsmethode auftreten.

Die Punktspreizfunktion g ist definiert als die Antwort des Messsystems auf ein ideales infitesimal kleines punktförmiges Objekt. Die Faltung wird aus Gründen der einfacheren Berechnung im Frequenzbereich durch die Multiplikation der Fouriertransformierten von f und g berechnet. Die für die Entfaltung

benötigte Umkehroperation ist dann die in Gl. (2) dargestellte einfache Division, wobei die Fouriertransformierten der jeweiligen Funktionen in Großbuchstaben angegeben sind.

$$F = \frac{H}{G} \qquad \text{Gl. (2)}$$

Die Fouriertransformierte des rekonstruierten geometrischen Profils F wird nach der Entfaltung durch die Anwendung einer inversen Fourier-Transformation zurück in den Zeitbereich überführt. Sofern die Punktspreizfunktion g bekannt ist, kann das geometrische Profil des Probekörpers f aus dem gemessenen Signal h mithilfe der Entfaltungsoperation rekonstruiert werden. Auf diese Weise werden die abgebildeten Beugungseffekte an Ecken und Kanten von Objekten nachträglich unterdrückt. In Abbildung 3 wird das Ergebnis einer solchen Entfaltung anhand synthetischer Daten dargestellt.

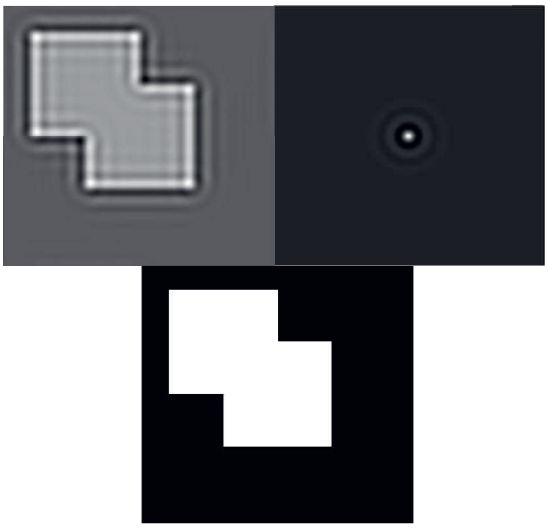


Abb. 3: Darstellung der Unterdrückung von Beugungserscheinungen anhand synthetischer Daten. Das Objekt mit Beugungserscheinungen h (links), die Transferfunktion bzw. PSF g (rechts) und das rekonstruierte Objekt f (unten).

Methodik

Es können zwei Hauptansätze verfolgt werden, um die Bildauflösung zu erhöhen, die Beugungseffekte zu reduzieren und die Defektklassifizierung in der Mikrowellen-Defektoskopie zu verbessern. Erstens kann die Messtechnik in Bezug auf ihre Emissionseigenschaften wie Strahlfokussierung, Rückflussdämpfung und Größe der Antennenfußfläche verbessert werden. Die Ergebnisse einer solchen Antennenoptimierungsstudie wurden in früheren Arbeiten vorgestellt [15].

Der aktuelle Beitrag zeigt die Ergebnisse eines zweiten Ansatzes, bei dem die Datenanalysetechniken in der Nachverarbeitung eingesetzt werden. Konzeptionell ist der Datenanalyseansatz in zwei Phasen unterteilt. Die erste Phase entspricht dem Training von Algorithmen, bei dem verschiedene Messszenarien numerisch simuliert und zur Bestimmung der spezifischen Punktspreizfunktionen verwendet werden. Die numerische PSF muss analytisch approximiert und eine Modellbibliothek erstellt werden, die die jeweiligen analytischen Punktspreizfunktionen für verschiedene Kombinationen von Antennen, Probenmaterial und Defekttypen enthält. In der zweiten Phase, der Anwendungsphase, werden die in der Modellbibliothek gespeicherten Punktspreizfunktionen unter realen Bedingungen verwendet, um das Beugungsmuster aus den gemessenen Bildern zu entfernen, wodurch klare und scharfe Objektprofile erzeugt werden. Nachdem die Bibliothek von Übertragungsfunktionen erstellt wurde, erfordert die Anwendungsphase keine großen Rechenressourcen mehr.
Der vorgeschlagene Arbeitsablauf kann mit den folgenden Schritten dargestellt werden.

Lernphase:
1. Numerische Simulation von PSFs
2. Analytische Näherung von PSFs
3. Erstellung der PSF-Bibliothek

Anwendungsphase:
4. Auswahl geeigneter PSF aus der Bibliothek
5. Entfaltung
6. Fehlerklassifizierung

Im ersten Schritt müssen für eine endliche Menge von Messszenarien PSF numerisch simuliert werden. Eine PSF ist dabei nur für ein bestimmtes, simuliertes Messszenario gültig. Um auch Punktspreizfunktionen für die große Menge an Messszenarien zu erhalten, die zwischen den simulierten Szenarien liegen (unterschiedliche Defektformen und -größen, Positionstiefe), werden die numerischen Punktspreizfunktionen im zweiten Schritt durch Anpassen geeigneter Funktionen analytisch approximiert. Jede Kombination von Antenne, Proben- und Defektmaterialien verfügt über eine eigene analytische PSF, die im Schritt 3 in einer Bibliothek zusammengefasst werden.
In der Anwendungsphase wird im Schritt 4 eine geeignete PSF aus der Bibliothek ausgewählt. Der Antennentyp und das Probenmaterial sind bei Messungen in der Regel bekannt. Der Defekttyp und das Defektmaterial müssen jedoch bestimmt werden. Dies kann durch Minimieren der Fehlerfunktion $\|h - f * g\|^2$ erreicht werden, wobei h ein gemessenes Reflexionsprofil, f ein unbekanntes Defektprofil und g eine analytische PSF aus der Bibliothek

darstellt. Die Minimierung wird für jede PSF aus der Bibliothek durchgeführt. Danach wird die PSF mit dem geringsten Minimierungsfehler für die Entfaltung ausgewählt. Die Entfaltung in Schritt 5 sollte dann das beugungsfreie Bild liefern, mit dessen Hilfe die Defekte in Schritt 6 klassifiziert werden können.
Im aktuellen Beitrag werden vor allem die Schritte 1 und 5 der vorgeschlagenen Datenanalyse basierend auf Simulationen, Messungen und empirischen Ergebnissen für ausgewählte Fehlerklassen erörtert.

Diskussion

Es wurde ein Satz von Probekörpern mit Defekten verschiedener Typen und Größen simuliert und gemessen. Danach wurde die auf Entfaltung basierende Rekonstruktionsmethode angewendet, um die Auflösung zu verbessern.
Für die Rekonstruktion wird eine zum Messszenario passende PSF benötigt. Die Generierung einer idealen Punktspreizfunktion aus Mess- oder Simulationsdaten ist allerdings schwierig, da die von einem infitesimal kleinem punktförmigem Objekt reflektierte Energie zu gering ist, um sie zu detektieren. Im Rahmen des Projekts wurden drei Ansätze getestet, mit denen zumindest eine gute Näherung der PSF aus dem Mess- bzw. Simulationsdaten generiert werden sollte.
Der erste Ansatz war die Anwendung der Entfaltung für ein Objekt, bei dem das ideale geometrische Profil bekannt war, so dass f synthetisch nachgebildet werden konnte. Durch das Umstellen der Gl. (2) sollte die Entfaltung von $G = H/F$ dann theoretisch die PSF g liefern. Allerdings stellte sich bei den praktischen Versuchen mit Messdaten heraus, dass mit dieser Methode für unterschiedliche Objektgrößen unterschiedliche PSF erzeugt wurden. Für im Vergleich zur Wellenlänge kleine Objekte wurden physikalisch sinnvolle lokalisierte PSF mit ähnlichem Funktionsverlauf erzeugt. Bei größeren Objekten lieferte die Entfaltung allerdings keine lokalisierten Funktionen mehr. Die erzeugten Funktionen glichen eher einem symmetrischen Rauschen und stellten eine reine mathematische Umkehrfunktion ohne physikalische Relevanz dar. Dieses Ergebnis ist ein erster Hinweis darauf, dass das aus der Optik übertragene Modell in der Mikrowellen-Defektoskopie nur eingeschränkt gültig ist.
Die zweite getestete Methode, eine PSF aus den Messdaten zu extrahieren, bestand darin, statt einem Punktdefekt eine scharfe Objektkante zu vermessen, um auf diese Weise die Kantenspreizfunktion (ESF) zu erhalten. Durch Ableitung der ESF orthogonal zum Kantenverlauf kann die Linenspreizfunktion (LSF) berechnet werden, die dem Querschnitt durch die dreidimensionale PSF entspricht [16]. Mithilfe dieser Methode konnten lokalisierte PSF erzeugt werden, die allerdings keine Spiegelsymmetrie entlang der y-Achse aufwiesen, wie sie für PSF eigentlich typisch ist. Bei der Entfaltung lieferten sie aber trotzdem ähnlich gute Ergebnisse, wie die mit der dritten Methode erzeugten PSF.
Die dritte Methode bestand darin, den kleinstmöglichen, noch messbaren Defekt zu vermessen und dessen Reflexionsprofil als PSF zu nutzen. So kann die für die Entfaltung benötigte PSF einfach und in guter Näherung erzeugt werden.
Aufgrund der einfacheren Umsetzung und dem symmetrischeren Funktionsverlauf wurde hauptsächlich auf diese dritte Methode zur Erzeugung der PSF zurückgegriffen. Der mit dem in Abbildung 1 dargestellten Messsystem kleinste messbare Defekt war ein zylindrischer Defekt aus Aluminium von 3 mm Durchmesser, der in eine 10 mm dicke Platte aus Polyethylen eingelassen war. Der Realteil des 2D-Scan dieses kleinen Referenzdefekts ist in Abbildung 5 oben links zu sehen.
Zur Vereinfachung der Bildanalyse wurden aus den jeweiligen 2D-Scans eindimensionale Schnittprofile, die mittig durch die Defekte verlaufen, extrahiert. Die Schnittebenen sind mit gestrichelten Linien angedeutet. Mit diesen Schnittprofilen wurde die Entfaltung durchgeführt.
In der Praxis sind die Messsignale in der Regel mit einem additiven Rauschen behaftet. Dieses Rauschen wird bei der Entfaltung mittels einer einfachen Division vor allem für die hohen Frequenzanteile im Messbild überproportional verstärkt. Um das zu vermeiden, wurde die Entfaltung mithilfe des Wiener-Filters umgesetzt [17, 18]. Die Aufgabe des Wiener-Filters ist es, die gestörten hohen Frequenzanteile aus den Messdaten herauszufiltern. Für den Frequenzbereich wird er mit Gl. (3) beschrieben.

$$F = \left[\frac{G^*}{|G|^2 + K}\right] \cdot H \qquad \text{Gl. (3)}$$

Dabei ist G^* die komplex Konjungierte von G und die Konstante K repräsentiert eine Abschätzung des Verhältnisses von Rausch- zur Signalleistung. Für niedrige Frequenzen entspricht der Wiener-Filter der in der Einleitung vorgestellten einfachen Division. Hohe Frequenzen werden jedoch herausgefiltert, wobei die Bandbreite des Filters durch die Wahl von K bestimmt wird.
Um die Frage zu beantworten, inwiefern die Qualität der Messungen durch die Rekonstruktionsmethode verbessert werden kann, soll

zunächst die Rekonstruktion eines im Vergleich zur Wellenlänge großen Defekts betrachtet werden. Abbildung 4 zeigt die Ergebnisse einer Messung eines großen, L-förmigen Defekts aus Aluminium, eingebettet in einer 10 mm dicken Platte aus Polyethylen.

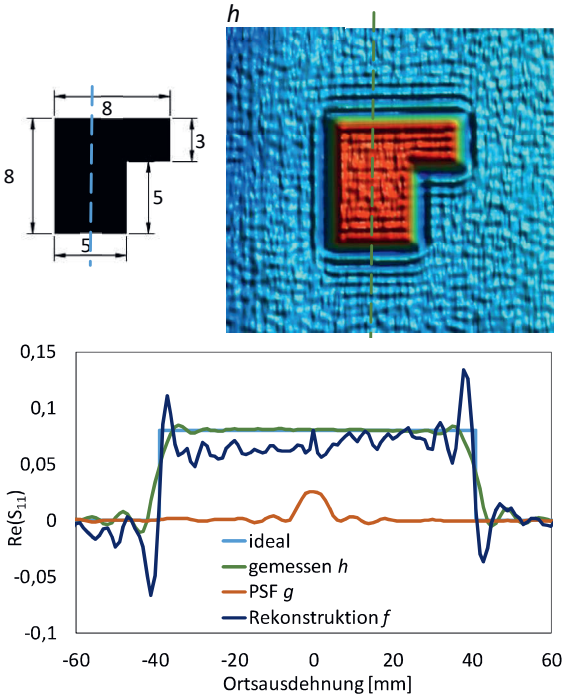


Abb. 4: Bildrekonstruktion für einen großen L-förmigen Defekt. Schematische Darstellung des Defekts (oben links), des Realteils des gemessenen Reflexionskoeffizienten S_{11} als 2D-Scan für den metallischen L-förmigen Defekt (oben rechts). Darstellung der 1D-Schnittprofile des gemessenen Defekts h, der PSF g, des rekonstruierten Defektprofils f und des idealen Defektprofils (unten).

Die Abmessungen des Defekts sind im Schema links oben in Millimetern angegeben. Das gemessene Signal wird als komplexer Reflexionskoeffizient S_{11} in einem 2D-Scan aufgezeichnet, dessen Realteil im Bild rechts oben dargestellt ist. In der unteren Grafik in Abbildung 4 sind die Schnittprofile der ursprünglichen Messung in grün und die Rekonstruktion des Defekts in dunkelblau dargestellt. Es ist erkennbar, dass der mit der Rekonstruktion einhergehende Informationsverlust Artefakte im Defektprofil erzeugt. Dadurch wird die Bildqualität insgesamt gesehen eher verschlechtert. Eine Verwendung des Rekonstruktionsalgorithmus zur Verbesserung der allgemeinen Bildqualität ist daher fraglich. Andererseits konnte die Kantenschärfe am Übergang vom Polyethylen zum Aluminium mit der Rekonstruktion deutlich verbessert werden.

Die Objektflanken folgen nach der Rekonstruktion dem idealen Profil sehr genau, was die Größenklassifizierung des Defekts vereinfacht.

Ein weiterer interessanter Anwendungsfall ist die Rekonstruktion von im Vergleich zur Wellenlänge kleinen Defekten. Die Wellenlänge bei 24 GHz in Luft beträgt ca. 12,5 mm. Vor allem kleine Defekte können leicht von Beugungsmustern überlagert werden, was ihre Erkennung erschwert. Es wurden zwei identische, zylindrische Defekte aus Aluminium mit 3 mm Durchmesser vermessen, die in einer 10 mm dicken Platte aus Polyethylen eingelassen waren. Die Defekte lagen so nahe beieinander, dass sich die von ihnen erzeugten Beugungsmuster überlagerten. In Abbildung 5 ist oben rechts das Schema der Defekte und der Realteil des 2D-Scans dargestellt. Wegen der Überlagerung der Beugungsmuster sind die einzelnen Defekte nicht unterscheidbar. Sie erscheinen im Bild wie ein einzelner, etwas größerer Defekt. Für die Rekonstruktion wurde wieder das Defektprofil eines einzelnen, zylindrischen Defekts aus Aluminium mit 3 mm Durchmesser als PSF verwendet, dessen 2D-Scan in Abbildung 5 oben links zu sehen ist.

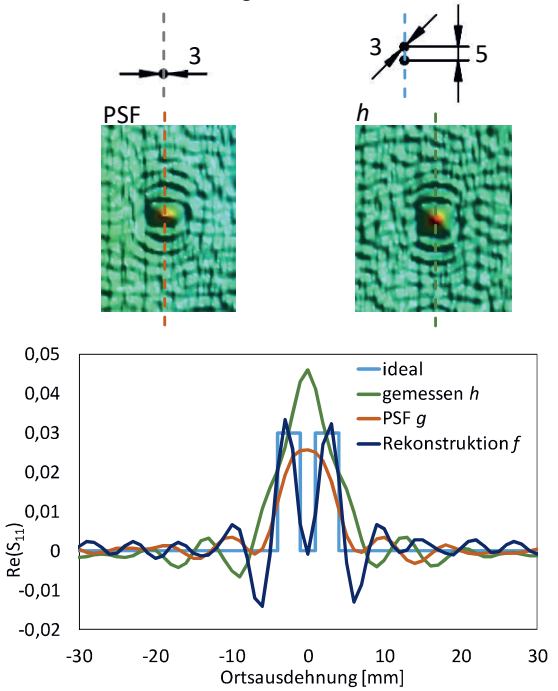


Abb. 5: Bildrekonstruktion für zwei kleine zylindrische Metalldefekte in unmittelbarer Nähe. Reflexions-koeffizient S_{11} als 2D-Scan für einen einzelnen Defekt (oben links) und für zwei Defekte (oben rechts). Das gemessene h, die PSF g, das rekonstruierte Profil f und das ideale geometrische Profil im Vergleich als 1D-Scan (unten).

In der unteren Grafik in Abbildung 5 sind die Schnittprofile durch die Defekte dargestellt. Während in der in grün dargestellten Originalmessung nur ein einzelner, zentraler Peak zu sehen ist, lassen sich nach der Rekonstruktion zwei klar getrennte Peaks ausmachen, die dem idealen Defektprofil sehr genau folgen. Mit Hilfe der Rekonstruktion konnten die Defekte also separiert werden.
Die Rekonstruktion von sich überlagernden Beugungsmustern funktioniert allerdings nur für ähnliche Defekte, für deren Entfaltung die gleiche PSF verwendet werden kann. Bei Defekten aus unterschiedlichen Materialien, zum Beispiel ein metallischer Einschluss neben einer mit Luft gefüllten Bohrung, funktioniert die Entfaltung nicht, da für die Rekonstruktion der verschiedenen Defekte unterschiedliche PSF verwendet werden müssten. Das ist der Hauptunterschied in der Nutzung des Entfaltungsmodells in der Mikrowellen-Defektoskopie und der Optik, wo die PSF nur vom Optischen System, nicht aber vom Prüfobjekt abhängig ist. In der Mikrowellen-Defektoskopie ist die Intensität und Form der Beugungsmuster um ein Objekt von dessen Reflexionsfläche, dem Abstand der Reflexionsfläche zur Antenne und des Unterschieds der Permittivität von Defekt- und dem Defekt umgebenden Material abhängig. Dadurch entstehen Beugungsmuster von unterschiedlicher Intensität, wobei Beugungsmuster starker Intensität die Beugungsmuster mit schwächerer Intensität überlagern können.

Zusammenfassung

In diesem Artikel wurde ein neuartiger zweistufiger Ansatz für die Mikrowellen-Defektoskopie vorgestellt, um Beugungseffekte, die an Inhomogenitäten im Prüfobjekt entstehen, nachträglich aus den Messdaten zu entfernen. Die vorgestellte Methodik basiert auf numerischen Simulationen zur Erzeugung von geeigneten Punktspreizfunktionen in der Lernphase und einem aus der Optik bekannten Entfaltungsansatz für die Objektrekonstruktion in der Anwendungsphase.
Es wurden außerdem drei Methoden getestet, um eine für die Entfaltung geeignete Punktspreizfunktion aus den Simulations- bzw. Messdaten zu extrahieren. Die im Projekt bevorzugt verwendete Methode bestand darin das Reflexionsprofil des kleinsten, messbaren Defekts in der Entfaltung zu verwenden. Die praktische Umsetzung dieser Methode ist einfach und das gemessene Reflexionsprofil stellt eine gute Näherung der idealen PSF dar.
Bei der Objektrekonstruktion wurden die gemessenen Daten mit der vorab generierten PSF unter Verwendung des Wiener-Filters entfaltet. Diese Rekonstruktionsmethode wurde durch die experimentelle Vermessung von einfachen synthetischen Defekten und durch 3D-Feldsimulationen verifiziert. Dabei konnte ein besseres Verständnis für das Potential, aber auch die Einschränkungen dieses Ansatzes in der Mikrowellen-Defektoskopie gewonnen werden. Es wurde gezeigt, dass mit der Entfaltungsmethode Beugungsmuster unterdrückt werden können, allerdings nur auf Kosten der allgemeinen Bildqualität, da durch den Informationsverlust bei der Entfaltung Artefakte entstehen. Die Defektgrenzen können mit gleichzeitiger Erzeugung von Artefakten, die eine Art Überschwingen an der Objektkontur darstellen, jedoch schärfer abgebildet werden. Wenn für die Defektklassifizierung die Darstellung der exakten Defektkonturen wichtiger ist, als die photographische Darstellung des Messbildes, kann das Rekonstruktionsverfahren hilfreich sein.
Die experimentellen Ergebnisse deuten außerdem an, dass zwei identische Defekte, die deutlich kleiner als die Wellenlänge des Messsignals sind und sehr nahe beieinander liegen, nach der Rekonstruktion einzeln aufgelöst werden können. Die klassische Beugungsgrenze des Messsystems kann in diesem speziellen Fall also überwunden werden. Auf den ersten Blick scheint die Beschränkung der Anwendbarkeit des Entfaltungssatzes bei der Überlagerung von Beugungsmustern auf identische Defekte eine starke Einschränkung zu sein. In der Praxis dominiert jedoch in einer großen Anzahl von Materialien meist eine Art von Defekt.
Die Experimente haben allerdings auch gezeigt, dass für jedes Messszenario in Abhängigkeit von Messsystem, Antenne, Material des Defekts und des den Defekt umgebenden Materials, der Größe der Reflexionsfläche und des Abstandes der Reflexionsfläche zur Antenne eine eigene PSF bestimmt werden muss. Das macht die Lernphase dieses Verfahrens sehr aufwändig und verhindert, dass es wie in der Optik als allgemeingültiges Werkzeug für ein bestimmtes Messsystem eingesetzt werden kann. Der im Artikel vorgeschlagene Ansatz für eine automatisierte Bildrekonstruktion und Defektklassifizierung, welche eine analytische Näherung einer geeigneten PSF verwendet, bietet Potential für einen begrenzten Satz von vorab definierten Messszenarien. Die große Menge von möglichen Messszenarien, verbunden mit dem hohen zeitlichen Aufwand der 2D-Simulationen bzw. Messungen zur Generierung geeigneter Punktspreizfunktionen beschränkt allerdings die praktische Anwendbarkeit des vorgestellten Ansatzes für allgemeinere Anwendungen.

Danksagung

Die finanzielle Unterstützung für die beschriebene Arbeit im Rahmen des IGF-Projekts 18827 BG durch die Forschungsgemeinschaft "Fördergemeinschaft für das SKZ" über AiF im Rahmen des Programms zur Förderung kooperativer industrieller Forschung und Entwicklung (IGF) des Bundesministeriums für Wirtschaft und Energie (BMWi), die auf einer Entschließung des Deutschen Bundestages basiert, wird dankend anerkannt. Darüber hinaus möchten wir uns beim Europäischen Fonds für regionale Entwicklung (EFRE) bedanken, der diese Forschung durch die Finanzierung der Radarsensoren ermöglichte.

Referenzen

[1] R. S. Panda, O. Karpenko, L. Udpa, M. Haq, P. Rajagopal, K. Balasubramaniam, *"Rapid non-contact inspection of composite ailerons using air-coupled ultrasound"*, AIP Conference Proceedings 1706, 080002 (2016)

[2] M. Webersen, S. Johannesmann, J. Düchting, L. Claes, B. Henning, *"Guided ultrasonic waves for determining effective orthotropic material parame-ters of continuous-fiber reinforced thermoplastic plates"*, Ultrasonics 84, S. 53–62, 2017.

[3] A. du Plessis, S. Roux, A. Guelpa, *"Case Studies in Nondestructive Testing and Evaluation"*, vol. 6, Part A, Nov. 2016, pp 17-25

[4] L. De Chiffre, S. Carmignato, R. Kruth, J.P. Schmitt, A. Weckenmann, *"Industrial applications of computed tomography"*, Man. Techn., 63 (2) (2014), pp. 655-677

[5] C. Ziehm, S. Hantscher, J. Hinken, C. Ziep, M. Richter, *„Near field focusing for nondestructive microwave testing at 24 GHz – Theory and experimental verification"*, Elsevier, Case Studies in Nondestructive Testing and Evaluation, vol. 6, pp. 70–78, Nov. 2016

[6] D. Beilken, J. Hinken, *"Microwave Based Nondestructive Testing Using Modified Eddy Current Systems"*, European Conference on Non-Destructive Testing, 2016

[7] S. Kharkovsky and R. Zoughi, *" Microwave and millimeter wave nondestructive testing and evaluation - Overview and recent advances"*, IEEE Instrumen-tation and Measurement Magazine, vol. 10, no. 2, pp. 26-38, Apr. 2007, doi: 10.1109/MIM.2007.364985

[8] C. Krebs, J. Rubart, A. Hommes, R. Brauns, D. Nüssler, *„High resolution measurements to determine the permittivity in artificial structures"*, International Conference on Millimeter Waves, Rome, Italy, 2010, doi: 10.1109/ICIMW.2010.5613004

[9] D. Nuessler, M. Schubert, S. Kose, N. Pohl, *"Swifter Security Scanning"*, IEEE Microwave Magazine, vol. 18, no. 6, Sept./Oct. 2017, doi: 10.1109/MMM.2017.2712018

[10] F. Ellrich, T. Weinland, J. Klier, J. Jonuscheit, and R. Beigang, *"Terahertz Waves – A New Spectral Band for Industrial Measurement Techniques",* Technisches Messen, vol. 77, no. 9, pp. 452-461, 2010

[11] G. von Freymann, J. Jonuscheit, *"THz technology - from hype to market"*, Photonik. Sonderheft LASER + PHOTONICS, pp. 102-105, 2016

[12] B. Baccouche, S. Mohammadzadeh, A. Keil, M. Kahl, P. Haring Bolivar, T. Loeffler, J. Jonuscheit, W. Sauer-Greff, R. Urbansky, F. Friederich, *"Illumination aspects of sparse line arrays for 3D terahertz imaging"*, 41st Int. Conference on Infrared, Millimeter, and Terahertz Waves, IRMMW-THz 2016, doi: 10.1109/IRMMW-THz.2016.7758825

[13] M. Nezadal, J. Schur, L. Schmidt, *"Non-destructive testing of glass fibre reinforced plastics with a synthetic aperture radar in the lower THz region"*, International Conference on Infrared, Millimeter and Te-rahertz waves, Kaiserslautern, 2012, doi: 10.1109/IRMMW-THz.2012.6380301

[14] R. Wegerhoff, O. Weidlich, M. Kässens. *„Basics of Light Microscopy & Imaging",* 18-19 (2011), GIT Verlag: Imaging & Microskopy

[15] Nesterov, M.; et.al.: *„Evaluation of Different Antenna and Data Analysis Concepts for Microwave Based Non-Destructive Testing"*, 12th European Conference on Non-Destructive Testing (ECNDT 2018), 11.-15.06.2018, Göteburg, Schweden, 2018.

[16] G. D. Boreman, *„Modulation Transfer Function in Optical and Electro-Optical Systems"*, Bellingham, Washington: SPIE-The International Society for Potical Engineering, 2001, S. 69-76, ISBN: 0-8194-4143-0

[17] Eva-Maria Brinkmann, „Dekonvolution in der Mikroskopie", Bachelorarbeit 2012, Westfälische Wilhelms-Universität Münster

[18] P. Sarder, A. Nehorai, *"Deconvolution methods for 3-d fluorescence microscopy images",* IEEE Signal Processing Magazine, S. 32–45, Mai 2006, doi: 10.1109/MSP.2006.1628876

Kontakt

Thomas Gagelmann: thomas.gagelmann@ifak.eu

DOI 10.5162/sensoren2019/P1.19

Dynamisches haptisches Feedback für Multi-Touch-Interaktionen

Falco Edner[1], Ulrike Steinmann[1]
[1]Otto-von-Guericke-Universität Magdeburg, Institut für Automatisierungstechnik, Lehrstuhl Messtechnik, Universitätsplatz 2, 39106 Magdeburg, Deutschland
falco.edner@ovgu.de

Zusammenfassung

In diesem Beitrag wird das Fokussierverhalten von elastischen Wellen in Platten durch akustische Zeitumkehr diskutiert, mit dem Ziel, mehrfache zeitlich und räumlich veränderliche taktile Anregungen auf einer Multi-Touch-Oberfläche (ebene Platte 720x920x3 mm, PMMA) zu erzeugen. Die grundlegenden Prinzipien und Vorgehensweisen zur Erzeugung von definierten taktilen Impulsen auf der Bedienoberfläche werden in dem Beitrag dargelegt. Es wird der jeweilige Fokuspunkt bezüglich seiner vertikalen Auslenkung und lateralen Auflösung bewertet. Zudem wird die Fokusqualität in Abhängigkeit von zeitlich gekürzten, re-emittierten Anregungssequenzen untersucht. Weiterhin werden Verfahren zur Amplitudenerhöhung einzelner Fokuspunkte aufgezeigt und im Hinblick auf die limitierte Anzahl der verwendeten Schallwandler ein Augenmerk auf das erreichbare Kontrastverhältnis gelegt.

Keywords: Haptik, Zeitumkehr-Fokussierung, Taktiles Display, haptische Interaktion

Einleitung

In diesem Artikel wird die Zeitumkehr-Technik (engl. time reversal, TR) genutzt, um auf einer Multi-Touch-Oberfläche eine taktile Rückmeldung an den Nutzer bereit zu stellen. Die Verwendung der TR-Methode bietet eine zusätzliche Interaktions- und Kommunikationsmöglichkeit für graphische Anzeigen oder Steuerungskonsolen, da auf der Oberfläche taktile Signale generiert werden können [1]. Ein weiterer Vorteil besteht bei der Zeitumkehr darin, dass die Wandler nicht im Interaktionsbereich des Nutzers angebracht werden müssen und somit die transparente Oberfläche für die Projektion frei bleibt. Zudem ermöglicht es das TR-Verfahren, ortsaufgelöst für mehrere Finger gleichzeitig, verschiedene Anregungen zu generieren. Die aktuellen Techniken zur taktilen Oberflächeninteraktion bieten diese Möglichkeit nicht, da sie die gesamte Struktur mittels Vibrationsmotoren stimulieren [2–4] oder durch Erzeugung von stehenden Körperschallwellen die Reibung zwischen Finger und der Oberfläche ändern [5–7]. Ebenso sind diese Verfahren für eine Interaktion mit mehreren Fingern ungeeignet. Die durch Fink [8] eingeführte TR-Technik der Fokussierung beliebiger Wellen in komplexen Medien überwindet zudem das dispersive Verhalten der dünnen Platte.
Ziel der Arbeit ist es, auf einer etwa 0,66 m², 3 mm starken Kunststoffplatte eine lokale und zeitlich variierbare taktile Rückmeldung zu generieren. Experimentell werden hierbei die entstehenden Biegewellen untersucht, da diese Verschiebungen aus der ebenen Oberfläche erzeugen und entscheidend für die für die taktile Wahrnehmung sind.

Prinzip

Die Zeitumkehrfokussierung von elastischen Wellen in Platten beruht üblicherweise auf zwei Abschnitten. Im ersten Schritt (Kalibrierung) wird in einem definierten Punkt A auf der Platte ein Impuls erzeugt und mittels einer festgelegten Anzahl von Wandlern, die sich bspw. am Rand der Platte befinden, die Verschiebungen aus der Ebene durch Ausbreitung und Nachhall der Plattenschwingungen innerhalb einer Zeitspanne T gemessen. Das dadurch aufgenommene Signal eines jeden Wandlers *i* (W_i) enthält die Impulsantwort $h_{AWi}(t)$ zwischen der Quelle im Punkt A und den einzelnen Wandlern W_i. Im zweiten Schritt wird die aufgenommene Impulsantwort eines jeden Wandlers zeitumgekehrt $h_{AWi}(T-t)$ und als Spannungssignal wieder an den jeweiligen Wandler zurückgespeist. In Folge der akustischen Emission erfolgt eine Re-Fokussierung der elastischen Wellen im Punkt A zu der Zeit $t = T$.
Diese Vorgehensweise ist auf den ersten Blick recht einfach, jedoch erweist sie sich in der Praxis als anspruchsvoll im Hinblick auf die i-

fach wiederholte Erzeugung des mechanischen Impulses während der Kalibrierung des Fokuspunktes. Eine alternative Methode verwendet daher die Reziprozität der genutzten Signale. Dieses Prinzip beruht darauf, dass die Umkehrung des ausgesendeten und aufgenommenen Signals das gleiche Impulsantwortsignal liefert und somit $h_{AWi}(t) = h_{WiA}(t)$ ist. Daher kann der Aufnahmeschritt auch dahingehend erfolgen, dass ein elektrischer Impuls $v_{Wi}(t)$ an den jeweiligen Wandler W_i gesendet wird und gleichzeitig die Verschiebung $u_{Ai}(t)$ am gewünschten Fokuspunkt A gemessen wird. Diese Verschiebungsmessung kann zum Beispiel mittels Laserdopplervibrometer LDV oder Beschleunigungssensor gemessen werden. Um die eingebrachte Energie in das System und damit auch den Signal-Rauschabstand SNR zu erhöhen, wird der elektrische Impuls mit einem breitbandigen Signal wie dem weißen Rauschen oder einem Sinus-Chirp ersetzt. In Abb. 1 ist exemplarisch dieser Vorwärtsschritt abgebildet, indem ein Sinus-Chirp-Signal v_{W1} auf den Aktor 1 gegeben und gleichzeitig mittels $LSDV$ die Verschiebung u_{A1} am Punkt A aufgenommen wird.

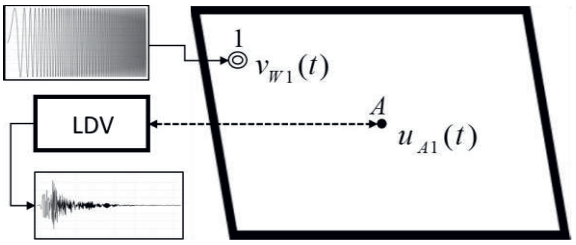


Abb. 1: Vorwärtsschritt mittels Chirp und gleichzeitiger Aufnahme der Verschiebung am späteren Fokuspunkt

Ein weiterer Schritt bei der Signalerzeugung nutzt die Tatsache aus, dass die Impulsantwort $h_{AWi}(t)$ die inverse Fourier-Transformierte der Übertragungsfunktion $H_{AWi}(\omega)$ ist. Das Verhältnis aus der gemessenen Verschiebung $U_{Ai}(\omega)$ und dem Spannungssignal am Aktor $V_{Wi}(\omega)$ bildet im Frequenzbereich die Übertragungsfunktion $H_{AWi}(\omega) = \frac{U_{Ai}(\omega)}{V_{Wi}(\omega)}$. Mittels der inversen Fourier Transformation F^{-1} ergibt sich für die Impulsantwort $h_{AWi}(t) = F^{-1}\left(H_{AWi}(\omega)\right)$. Im Folgenden wird die Impulsantwort $h_{AWi}(t)$ des jeweiligen Wandlers W_i zeitumgekehrt $h_{AWi}(T-t)$, wobei die Impulsantwort für jeden Wandler und Kalibrierungspunkt individuell zu bestimmen ist. Daraufhin wird, wie in Abb. 2 dargestellt, von allen Aktoren gleichzeitig die jeweils bestimmte Impulsantwort ausgesendet, woraufhin eine Re-Fokussierung $u_{FPA}(t)$ am kalibrierten Punkt A erfolgt.

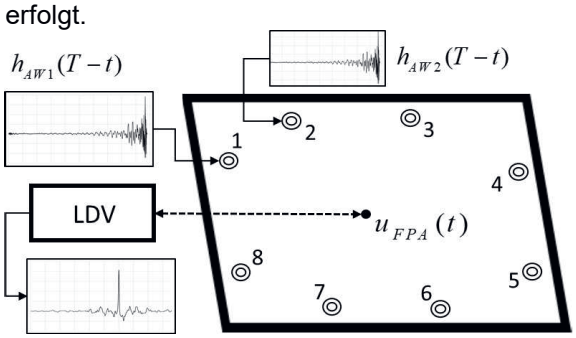


Abb. 2: Fokussierung mit simultaner Aussendung der Signale (hier exemplarisch für 2 der 8 Wandler)

Taktile Rückmeldung durch Vibration

Bei der Nutzerinteraktion mit der Eingabeoberfläche muss eine Empfindungsschwelle überwunden werden, um eine haptische Berührung wahrzunehmen. Haptik beschreibt die Wahrnehmung von mechanischen (taktil und kinästhetisch), thermischen und nozizeptorischen (schmerzhaften) Sinneseindrücken über die sensorischen sowie motorischen Fähigkeiten in der Haut, Gelenken, Muskeln und Sehnen. Für die hier avisierte Rückmeldung via Multi-Touch-Oberfläche werden die taktilen Rezeptoren, bestehend aus Merkelzellen, Ruffini-Körperchen, Meissner-Körperchen und Vater-Pacini-Körperchen in den äußeren Hautschichten genutzt, die mechanische Spannungen und Dehnungen in der Haut registrieren [9]. Dies wird demnach im Folgenden als taktile Rückmeldung bezeichnet. Dafür spielen unterschiedliche Parameter eine entscheidende Rolle: Um eine spürbare Rückmeldung für die Finger auf der Oberfläche zu erzeugen, muss nicht nur eine minimale Amplitude erreicht werden, sondern auch deren Abhängigkeit von der Frequenz, bei der der Fokuspunkt eine Vibration vollzieht, berücksichtigt werden (frequenzabhängige Wahrnehmungsschwelle). Die hier relevanten Rezeptoren sind a) Meissner-Körperchen, die eine maximale Empfindlichkeit im Bereich von 10-30 Hz besitzen, wobei eine minimale Amplitude von 5 µm für die Wahrnehmung benötigt wird. Im Fall der b) Vater-Pacini-Körperchen, deren maximale Sensitivität im Bereich von 250-300Hz liegt, sinkt die

Empfindungsschwelle auf einen theoretischen Wert von 0,1 µm [9–11].

Experimenteller Aufbau

Die Messung der Auslenkungen der Platte werden mittels Laserdopplervibrometer PSV-500-HV in Verbindung mit dem PSV-I-550 Scanning Head Xtra der Firma Polytec durchgeführt. Hierzu zählen die Signalsequenzen während Kalibrierung, die Amplituden und lateralen Auflösungen der Fokuspunkte sowie Vermessung der gesamten Plattenoberfläche während des Re-Emissionsprozesses. Die Kalibrierung erfolgt durch sequentielles Anlegen von Sinus-Chirp-Signalen an alle *i* Wandler in einem Bereich von 1 Hz bis 20 kHz mittels des integrierten Signalgenerators der PSV-E-530 Junction Box des LDV. Die Amplitude des Chirp-Signals (0,5 V_{pp}) wird zuvor durch einen analogen Verstärker (OPA552FA) auf 30 V_{pp} verstärkt. Sowohl für die Kalibrierung als auch alle experimentellen Untersuchungen wurden $i = 8$ Piezo-Aktoren (TDK PowerHap 2.5G), verwendet, die am Rand der Oberfläche verbaut wurden (Verteilung siehe Abb. 2). Sie sind zwischen der Platte (720x920x3 mm^3, PMMA) und deren punktuellen Halterungen/Auflagepunkten geklemmt. Die einzelnen zeitumgekehrten Impulsantworten werden simultan über eine PXI-6733 Messkarte der Firma National Instruments ausgegeben und wiederum mittels 8 analogen Verstärkern des Typs OPA522FA verstärkt. Limitierend für die verwendete Anzahl der Wandler ist die genutzte analoge Ausgabekarte mit ihren 8 zur Verfügung stehenden simultanen Ausgabe-kanälen.

Laterale Auflösung

Die laterale Auflösung L kann als Größe der räumlichen Ausbreitung bei der Hälfte der maximalen Amplitude des Fokuspunktes beschrieben werden und hängt im Fall einer endlichen Bandbreite und vernachlässigbarer Dispersion von der kleinsten Wellenlänge $\lambda_{\min}$ bzw. der größten Frequenz $f_{\max}$, die der Transducer anregt ab, sodass gilt [12],

$$L = \frac{\lambda_{\min}}{2} \tag{1}$$

Die Biegesteifigkeit B_{ST} einer dünnen Platte berechnet sich aus [13],

$$B_{ST} = \frac{h^3}{12}\frac{E}{(1-\nu^2)}, \tag{2}$$

wobei h die Dicke der Platte, ρ die Dichte, E der Elastizitätsmodul und ν die Querkontraktionszahl des Plattenmaterials ist. Die Plattengeschwindigkeit c_{BP} der Biegewelle definiert sich durch [13]

$$c_{BP} = \sqrt[4]{\frac{B_{ST}}{\rho h}\omega^2}, \tag{3}$$

wodurch sich für die minimale Wellenlänge $\lambda_{\min}$ folgende Beziehung ergibt:

$$\lambda_{\min}^2 = \frac{\pi h}{f_{\max}}\sqrt{\frac{E}{3\rho(1-\upsilon^2)}}. \tag{4}$$

Zusammengefasst ist so die mögliche laterale Auflösung, die zur fühlbaren Rückmeldung auf der Platte dient, abhängig von den mechanischen Eigenschaften der Platte und der höchsten Frequenz, die in das System eingebracht wird. Legt man die verwendete 3 mm PMMA-Platte zugrunde, erhält man den in Abb. 3 dargestellten Verlauf der räumlichen Auflösung in Abhängigkeit der maximalen Frequenz.

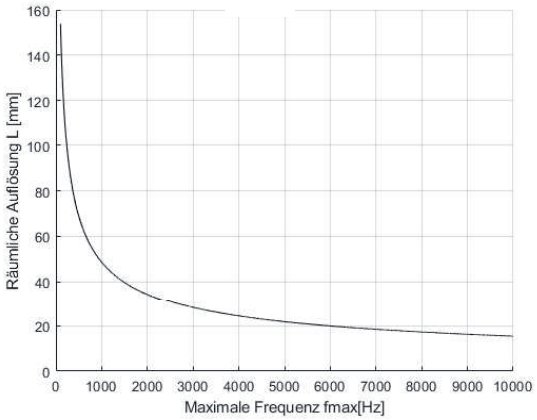


Abb. 3: *Räumliche Auflösung L der genutzten 3mm PMMA-Projektionsplatte*

Es ist ersichtlich, dass mit einer Steigerung der maximalen Frequenz im System die laterale Auflösung verbessert wird. Gerade im Bereich der unteren Frequenzen bis 3 kHz ist eine starke Verringerung und danach ein deutliches Abflachen der Kurve ersichtlich. Bei 10 kHz ergibt sich eine laterale Auflösung von 15 mm.

Re-Fokussierungszeit

Die Abhängigkeit der Amplitude und der lateralen Auflösung des Fokuspunktes von zeitlich gekürzten, re-emittierten Impulsantworten, soll im Folgenden untersucht werden. Die zeitumgekehrten Impulsantworten $h_{AWi}(T-t)$ werden in einem bestimmten Bereich abgeschnitten, was wiederum zu einer Verkürzung der Re-Fokussierungszeit (RFZ) führt. Ein besonderes Augenmerk wird bei den folgenden Untersuchungen auf 2 Grenzbereiche gelegt. Zum einen auf eine RFZ von 4 ms (Impulsfolge 250 Hz), um die höchste Sensitivität der Vater-Pacini-Körperchen zu nutzen. Zum anderen eine RFZ von 64 ms (Impulsfolge ca. 16 Hz), um den optimal fühlbaren Bereich der Meissner-Körperchen nutzen zu können. In Abb. 4 ist experimentell der Einfluss der Veränderung von RFZ auf die Amplitude des Fokuspunktes dargestellt.

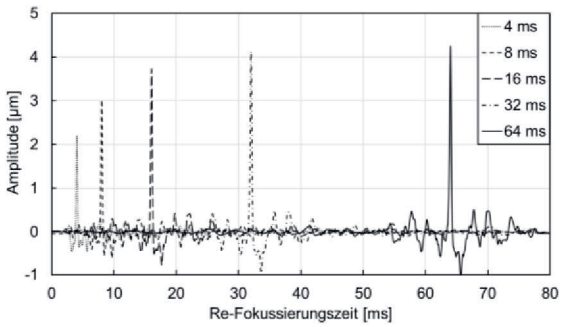


Abb. 4: Zeitlicher Verlauf am Fokuspunkt bei unterschiedlichen Re-Fokussierungszeiten

Es wird ersichtlich, dass die Amplitude von 4,25 µm bei 64 ms auf 2,19 µm bei 4 ms absinkt (Faktor 1,94).
Die laterale Auflösung des Fokuspunktes (Abb. 5) steigt mit Erhöhung der RFZ von 43 mm bei 4 ms auf 52 mm bei 64 ms und verschlechtert sich somit.

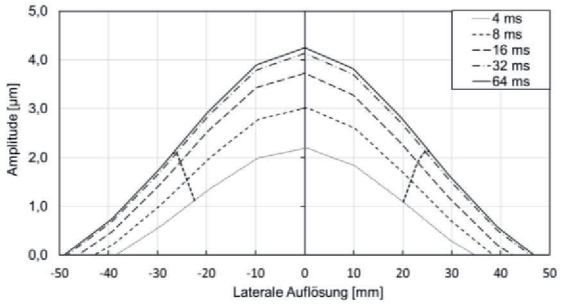


Abb. 5: Amplitude und laterale Auflösung des Fokuspunktes bei unterschiedlicher Re-Fokussierungszeit

Im höher-dynamischen Bereich von 4 ms wird die Amplitude der theoretischen Wahrnehmungsschwelle (0,1 µm) zwar überschritten, im weniger dynamischen Betriebsmodus von 64 ms Re-Fokussierdauer kann die Amplitude von 5 µm jedoch nicht erreicht werden. Nachfolgend werden daher Methoden zur Amplitudenerhöhung und deren Auswirkung auf die Fokusqualität weiter untersucht.

Amplitudenerhöhung durch 1-Bit-Verfahren

Durch das 1-Bit-Verfahren kann die Amplitude des Fokuspunktes erhöht werden. Derode et al. führten die Verwendung der 1-Bit-Methode in Zeitumkehr-Experimenten ein und konnten bei gleichbleibender Fokusqualität eine Steigerung der Amplitude bei gleicher Spitze zu Spitze Spannung V_{pp} erreichen [14].

Das 1-Bit-Verfahren umfasst folgende Signalverarbeitungsschritte (vgl. Abb. 6): a) Zeitumkehren der Impulsantwort, b) zeitliche Fensterung des Signals, c) das Signal nach Fensterung verwenden, um in d) eine 1-Bit Quantifizierung vorzunehmen, sodass das Signal am Ende über Gl. 5 in eine binäre Form (0/1) gebracht werden kann. Dabei kann über den Faktor V_{pp} die Höhe der Amplitude verändert werden.

$$v_W(t) = \frac{1}{2} V_{pp} \left[1 + \mathrm{sgn}\left(h_{AW}\left(T - t \right) \right) \right] \qquad (5)$$

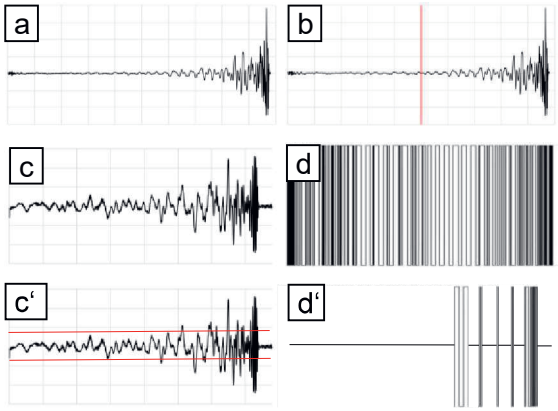


Abb. 6: (a) zeitumgekehrte Impulsantwort, (b) Fensterung des Signals, (c) Signal nach der Fensterung, (d) 1-Bit Quantifizierung, (c‘) Definition zweier Schwellwerte, (d‘) 1-Bit Quantifizierung mittels Schwellwerten

Die so veränderte Impulsantwort enthält daraufhin nur noch die Phaseninformation des ursprünglichen Signals. Um insbesondere die zu Beginn der Sequenz auftretenden Signalrauschanteile zu unterdrücken, können

zwei Schwellwerte innerhalb der Impulsantwort $h_{AW}(T-t)$ definiert werden (Abb.6c‘). Oberhalb der positiven Schwelle wird der Wert daraufhin auf 1, unterhalb der negativen Schwelle auf -1 und die Zwischenwerte auf 0 gesetzt (Abb.6d‘). Diese alternative Signalverarbeitung wurde in den Messungen jedoch nicht genutzt und soll lediglich Erwähnung finden. Der entscheidende Vorteil dieser 1-Bit-Methode ist, dass das binäre Signal mittels digitaler Technik ausgegeben und verstärkt werden kann.
Die Messergebnisse zeigen, dass sich wie erwartet und theoretisch beschrieben die Amplitude des Fokuspunktes gegenüber der normalen Zeitumkehr verbessert. In Anlehnung an die zuvor gewählte Darstellung (Abb. 4) wird eine Abhängigkeit der Amplitude zu der Re-Fokussierungszeit ersichtlich. Die Amplitude sinkt von einem Wert von 8,81 µm bei 64 ms auf einen Wert von 3,11 µm bei 4 ms (Faktor 2,83) (Abb. 7). Damit kann nunmehr für beide Ziel-Rezeptoren die jeweilige frequenzabhängige Empfindungsschwelle erreicht werden. Die laterale Auflösung des Fokuspunktes steigt wie schon beim Normal-Verfahren mit der RFZ an: von 44 mm bei 4 ms auf 59 mm bei 64 ms Zeitdauer.

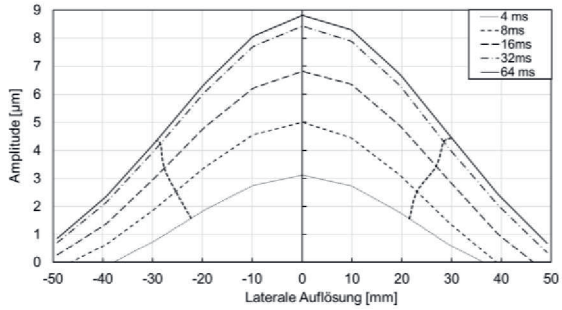


Abb. 7: Amplitude und laterale Auflösung des Fokuspunktes mittels 1-Bit-Verfahren zur Amplitudensteigerung

Amplitudenerhöhung durch Clipping

Dieses Verfahren zur Erhöhung der Amplitude des Fokuspunktes wurde von Heaton et al. als Verarbeitungsverfahren für die Zeitumkehr eingeführt und ist ähnlich dem 1-Bit-Verfahren [15]. Legt man alle Zeitumkehr-Methoden zu Grunde, zeigt das Clipping-Verfahren die höchste Amplitude [16] .
Wesentliche Signalverarbeitungsschritte bei diesem alternativen Verfahren sind (Abb. 8): (a) Zeitumkehren der Impulsantwort, b) zeitliche Fensterung des Signals, (c) Definition der minimalen und maximalen Schwellwerte, (d) das Signal an den Schwellwerten abschneiden und (e) Signal auf einen definierten Wert skalieren.

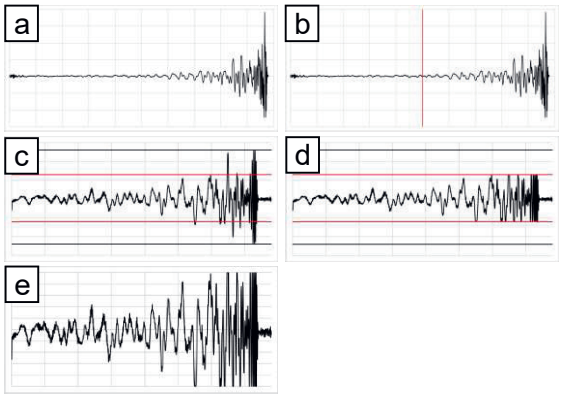


Abb. 8: (a) zeitumgekehrte Impulsantwort, (b) Fensterung des Signals, (c) Schwellwerte definieren, (d) Signal an Schwellwerten abschneiden (e) Signal auf einen bestimmten Wert skalieren

Der wesentliche technische Unterschied zum 1-Bit-Verfahren liegt in dem Beschneiden des Signals, da bei der Clipping-Methode das Signal zwischen den Schwellwerten unverändert bleibt. Entscheidend ist zudem bei dem Clipping-Verfahren die Schwellwert-optimierung.
Wie auch schon bei dem normalen und 1-Bit Verfahren gezeigt, ist auch hier ein Verringern der Amplitude in Abhängigkeit der Verkürzung der RFZ zu sehen. Diese sinkt von 12,73 µm bei 64 ms auf 3,44 µm bei 4 ms Zeitdauer (Faktor 3,7). Auch hier kann für die taktilen Rezeptoren eine ausreichend hohe Ansprechbarkeit bzgl. der frequenzabhängige Empfindungsschwelle erreicht werden.
Des Weiteren verschlechtert sich die laterale Auflösung von 49 mm bei 4 ms auf 61 mm bei 64 ms (Abb. 9).

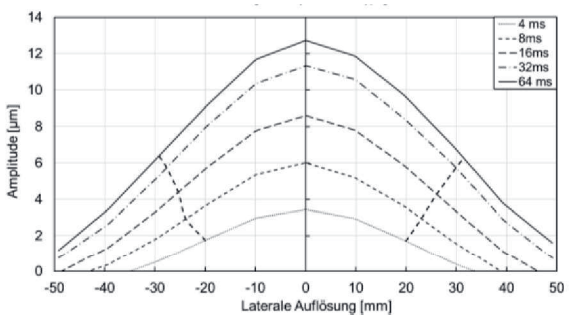


Abb. 9: Amplitude und laterale Auflösung am Fokuspunkt beim Clipping-Verfahren zur Amplitudensteigerung

Gegenüberstellung der 3 Verfahren

Abb. 10 zeigt eine Übersicht, wie sich die Amplitude des Fokuspunktes in Abhängigkeit

der genutzten Verfahren und der Re-Fokussierungszeiten unterscheidet.

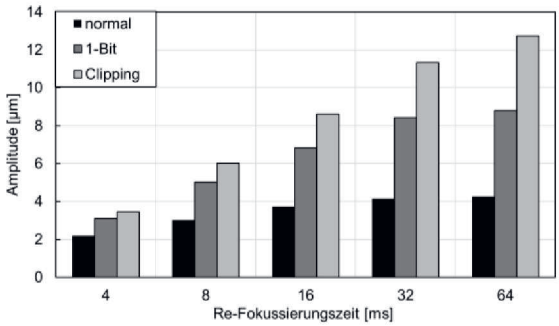


Abb. 10: Gegenüberstellung der maximalen Fokusamplituden zu der Re-Fokussierungszeit und zu den Verstärkungsverfahren

Im direkten Vergleich als auch über alle RFZ wird bei beiden Verstärkungsmethoden die Amplitude gegenüber der normalen Zeitumkehr gesteigert, wobei das Clipping-Verfahren den wesentlichsten Einfluss hat. Bei einer kurzen RFZ von 4 ms liegen die Amplituden aller Verfahren nah beieinander, wobei eine Steigerung der Amplitude bei dem 1-Bit-Verfahren um das 1,42-fache und bei dem Clipping-Verfahren um das 1,57-fache im Gegensatz zu der normalen Zeitumkehr beobachtet wird (Tab.1). Bei Zunahme der RFZ steigt auch der Unterschied zwischen den einzelnen Verfahren. Bei der normalen Zeitumkehr steigt der Wert von 4 ms gegenüber 64 ms um das 1,94-fache. Vergleicht man für die längste RFZ (64 ms) die Verfahren, so skalieren die Amplituden um das 2,07-fache (1-Bit) und das 3-fache (Clipping) relativ zur normalen Fokusamplitude (Tab.1).

Tab. 1: Normierte Verstärkungsfaktoren der Amplituden bei den verschiedenen Verfahren bezogen auf die normale Zeitumkehr TR

RFZ [ms]	Faktor Normale TR	Faktor 1-Bit	Faktor Clipping
4	1,00	1,42	1,57
8	1,38	1,66	1,99
16	1,70	1,83	2,31
32	1,88	2,04	2,74
64	1,94	2,07	3,00

Die laterale Auflösung des Fokuspunktes steigt bei allen drei Verfahren in Abhängigkeit der RFZ und somit der Fokusamplitude an. Die vorliegende Datenlage lässt allerdings keinen belastbaren Rückschluss darauf zu, inwiefern bei vorgegebener, für alle 3 Verfahren *identischer* Fokusamplitude eine Verbesserung der lateralen Auflösung erzielt werden kann. Es ist zu vermuten, dass das 1-Bit und Clipping Verfahren aufgrund der Signalmodifikationen einen Einfluss höherfrequenterer Anteile mit sich bringt und $\lambda_{\min}$ bzw. *L* folglich kleiner wird. Eine experimentell darauf ausgerichtete Untersuchung ist Gegenstand aktueller Arbeiten.

Kontrast des Fokuspunktes

Das Verhältnis der Amplitude am Fokuspunkt gegenüber dem quadratischen Mittelwert der Amplituden über die Platte zum Zeitpunkt des Impulses kann als Messwert für die Fokusqualität zu Grunde gelegt werden [17] .
Für die Beurteilung des Kontrastes wurde im Mittelpunkt der Multi-Touch-Platte mittels Zeitumkehr ein Impuls erzeugt und gleichzeitig die gesamte Oberfläche mit einem Raster von 20x20 mm vermessen (Abb. 11).

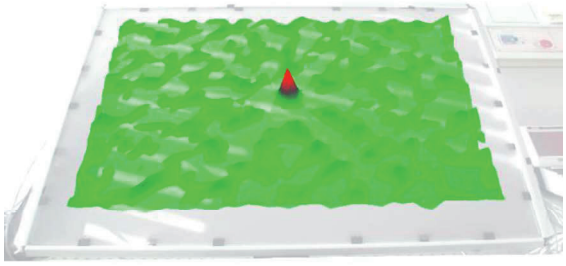


Abb. 11: Auslenkungen auf der Multi-Touch-Oberfläche zur Zeit der Impulserzeugung

Der quadratische Mittelwert aller Oberflächenamplituden (wobei bei der Berechnung der Bereich des Fokuspunktes nicht berücksichtigt wird) liegt gegenüber der normalen Zeitumkehr bei den Verstärkungsverfahren höher und führt zu einem verstärkten Grundrauschen der Platte (Abb. 12). Zudem steigt hier der Mittelwert wie auch im Vergleich zu den Ergebnissen aus Abb. 4 asymptotisch mit der RFZ an.

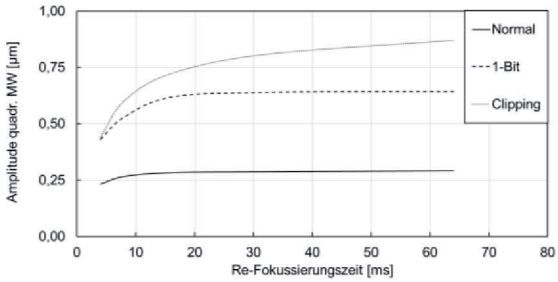


Abb. 12: Quadratischer Mittelwert der Oberflächenamplituden ohne Fokuspunkt bei veränderter Re-Fokussierungszeit und Verstärkungsverfahren

Wird der quadratische Mittelwert ins Verhältnis zu der Amplitude des Fokuspunktes gesetzt

(Abb. 13), kann eine Aussage über das Signal-Rausch-Verhältnis SNR der Verfahren bzw. deren Fokusqualität in Abhängigkeit von der RFZ getroffen werden. Es wird ersichtlich, dass durch die jeweiligen Verstärkungsverfahren (1-Bit, Clipping) die Fokusamplitude zwar erhöht wird, das aber mit einem ebenso verstärkten Grundrauschen einhergeht. In Bezug auf diese Betrachtungsweise (SNR) zeigen die drei Verfahren ein relativ vergleichbares Kontrastverhältnis und ähnliche Fokusamplitude. Letztere steigt mit zunehmender Zeitdauer bei allen drei Verfahren an und erreicht ab ca. 30 ms einen konstanten Wert.

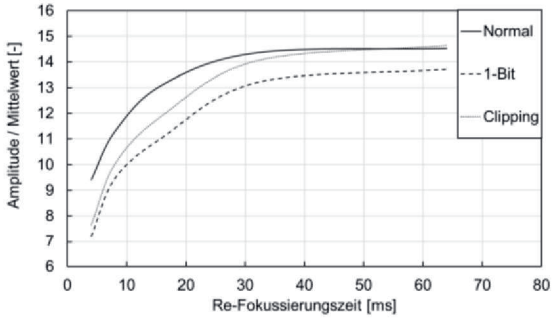


Abb. 13: Verhältnis der Amplitude des Fokuspunktes zum quadratischen Mittelwert der Oberflächenamplituden ohne Fokuspunkt bei veränderter Re-Fokussierungszeit und in Abhängigkeit des Verstärkungsverfahrens

Zusammenfassung

In diesem Beitrag konnte gezeigt werden, dass mittels einer begrenzen Anzahl von 8 Wandlern überwiegend spürbare Amplituden auf der verwendeten Multi-Touch-Oberfläche erzeugt werden konnten und dass das Verfahren der Zeitumkehr-Fokussierung für die Multi-Touch-Interaktion geeignet ist. Mit zunehmender Re-Fokussierungszeit steigt die Amplitude des Impulses asymptotisch bis zu einem gleichbleibenden Wert an und mittels des 1-Bit und Clipping-Verfahrens konnte die Amplitude bei 64 ms Re-Fokussierungszeit von 4,2 µm bei normaler TR bei dem 1-Bit-Verfahren um Faktor 2,07 auf 8,7 µm und bei dem Clipping-Verfahren um Faktor 3,0 auf 12,6 µm gesteigert werden. Ein hinreichend hoher Kontrast des Fokuspunktes zur gesamten Oberfläche konnte beobachtet werden, wobei alle drei verwendeten Verfahren relativ betrachtet, gleiche Kontrastverhältnisse liefern. Es konnte die Möglichkeit einer dynamischen Anregung von Punkten mittels Impulsfolgen bis zu 250 Hz mit ausreichend hoher Amplitude nachgewiesen werden. Zusammenfassend konnte belegt werden, dass das verwendete TR-Verfahren für den Anwendungsfall geeignet ist.

Literaturverzeichnis

[1] C. Hudin, J. Lozada, and V. Hayward, “Localized Tactile Feedback on a Transparent Surface through Time-Reversal Wave Focusing,” (eng), *IEEE transactions on haptics*, vol. 8, no. 2, pp. 188–198, 201510.1109/TOH.2015.2411267.

[2] J. Pasquero *et al.,* “Haptically Enabled Handheld Information Display With Distributed Tactile Transducer,” *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 746–753, 200710.1109/TMM.2007.895672.

[3] M. E. Altinsoy and S. Merchel, “Electrotactile Feedback for Handheld Devices with Touch Screen and Simulation of Roughness,” (eng), *IEEE transactions on haptics*, vol. 5, no. 1, pp. 6–13, 201210.1109/TOH.2011.56.

[4] J. Reisinger and J. Wild, “Haptische Bedienschnittstellen,” in *Infotainmentsysteme im Kraftfahrzeug*, Wiesbaden: Vieweg, 2008, pp. 115–145.

[5] T. Nara *et al.,* “Surface acoustic wave tactile display,” *IEEE Comput. Grap. Appl.*, vol. 21, no. 6, pp. 56–63, 200110.1109/38.963461.

[6] E. C. Chubb, J. E. Colgate, and M. A. Peshkin, “ShiverPaD: A Glass Haptic Surface That Produces Shear Force on a Bare Finger,” (eng), *IEEE transactions on haptics*, vol. 3, no. 3, pp. 189–198, 201010.1109/TOH.2010.7.

[7] M. Wiertlewski, D. Leonardis, D. J. Meyer, M. A. Peshkin, and J. E. Colgate, “A High-Fidelity Surface-Haptic Device for Texture Rendering on Bare Finger,” in *Lecture Notes in Computer Science, Haptics: Neuroscience, Devices, Modeling, and Applications*, M. Auvray and C. Duriez, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, pp. 241–248.

[8] M. Fink, “Time reversal of ultrasonic fields. I. Basic principles,” (eng), *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 39, no. 5, pp. 555–566, 199210.1109/58.156174.

[9] R. T. Verrillo and S. J. Bolanowski, “Tactile Responses to Vibration,” *Handbook of Signal Processing*, vol. 4, pp. 1185–1213, 200810.1007/978-0-387-30441-0_62.

[10] L. A. Jones and N. B. Sarter, “Tactile displays: guidance for their design and application,” (eng), *Human factors*, vol. 50, no. 1, pp. 90–111, 200810.1518/001872008X250638.

[11] T. A. Kern, *Entwicklung Haptischer Geräte*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[12] Didier Cassereau and Mathias Fink, “Time-Reversal of Ultrasonic Fields-Part III: Theory of the Closed Time-Reversal Cavity,” *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 39, no. 5, pp. 579–592, 1992.

[13] R. Lerch, G. Sessler, and D. Wolf, *Technische Akustik*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.

[14] A. Derode, A. Tourin, and M. Fink, “Ultrasonic pulse compression with one-bit time reversal through multiple scattering,” *Journal of Applied Physics*, vol. 85, no. 9, pp. 6343–6352, 199910.1063/1.370136.
[15] C. Heaton, B. E. Anderson, and S. M. Young, “Time reversal focusing of elastic waves in plates for an educational demonstration,” (eng), *The Journal of the Acoustical Society of America*, vol. 141, no. 2, p. 1084, 201710.1121/1.4976070.
[16] M. L. Willardson, B. E. Anderson, S. M. Young, M. H. Denison, and B. D. Patchett, “Time reversal focusing of high amplitude sound in a reverberation chamber,” (eng), *The Journal of the Acoustical Society of America*, vol. 143, no. 2, p. 696, 201810.1121/1.5023351.
[17] C. Hudin, J. Lozada, and V. Hayward, “Spatial, temporal, and thermal contributions to focusing contrast by time reversal in a cavity,” *Journal of Sound and Vibration*, vol. 333, no. 6, pp. 1818–1832, 201410.1016/j.jsv.2013.11.016.

DOI 10.5162/sensoren2019/P1.20

Design und Modellierung eines 3D-Nanotasters

Jan Thiesler[1], Gaoliang Dai[1], Rainer Tutsch[2]
[1]Physikalisch-Technische Bundesanstalt, Bundesallee 100, Deutschland
[2]Technische Universität Braunschweig, IPROM, Schleinitzstraße 20, Deutschland

Zusammenfassung

In der Nanometrologie werden Strukturen immer kompakter und komplexer, dabei basiert deren Funktionalität auf genau bekannten dreidimensionalen geometrischen Abmessungen und definierter Rauigkeit der Oberfläche. So ist eine Miniaturisierung in der Halbleiterindustrie vielfach durch Wandrauigkeiten (LER = line-edge roughness) limitiert und eine genaue Kenntnis des Fertigungsprozesses ist notwendig. Das vorgestellte Design eines 3D-Nanotasters basiert auf dem Prinzip eines Biegebalkens eines CD-AFM Cantilevers. Dieser strukturierte Biegebalken beinhaltet zwei Festkörpergelenke, ist hinsichtlich isotroper Steifigkeit (<1 N/m) und 3D-Sensitivität optimiert, besitzt eine erste Eigenfrequenz von 59,3 kHz und ist als Taster für Spitzendurchmesser unter 100 nm konzipiert.

Keywords: Nanotaster, CD-Cantilever, 3D-AFM, Critical Dimension, µKMG

Einleitung

Seit der Erfindung des Rasterkraftmikroskopes durch G. Binnig et al. 1985/86 [1] hat sich das Design der mit der Probe wechselwirkenden Sonde als Biegebalken, zur Detektion von Oberflächentopographien im Nanometer-bereich für Rastersondenmikroskope, durchgesetzt. Dieser Biegebalken oder Cantilever verformt sich elastisch durch die Kräfte der wechselwirkenden Oberflächenpotentiale von der Spitze der Sonde und Probe. Diese Verformung kann interferometrisch, kapazitiv, piezoresistiv oder über ein Lichtzeigerprinzip ausgelesen werden. Dabei ist der klassische Cantilever (**Abb. 1**) ein 1D-Sensor, der lediglich die Kräfte senkrecht zu seiner Längsachse bei glatter Probenoberfläche als eine Biegung zeigt.

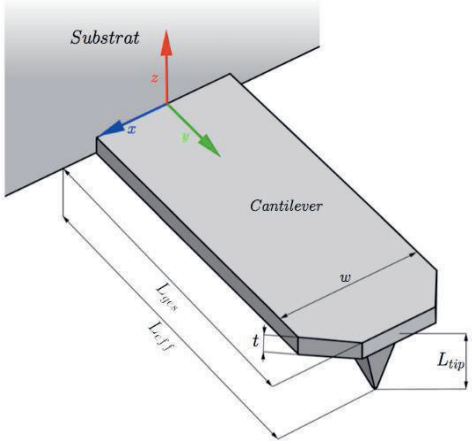


Abb. 1: Skizze eines Cantilevers mit Variablen (Breite = w, Dicke = t, effektive Länge = L_{eff} und Spitzenlänge = L_{tip}) und Koordinatensystem

Spezielle Cantilever mit am Ende verbreiteter Spitze für CD-AFMs, im Folgenden als CD-Cantilever (CD = critical dimension) (**Abb. 2**) bezeichnet, sind mit dem entsprechenden Tastkopf des Rasterkraftmikroskops, in der Lage als 2D-Sensor zu arbeiten. Es wird die Biegung und Torsion des Cantilevers detektiert [2, 3].

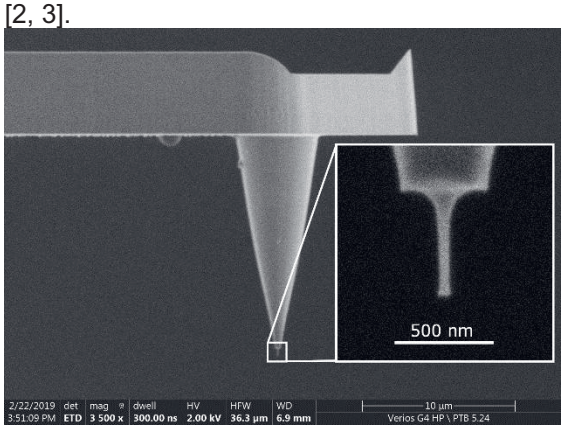


Abb. 2: Seitliche REM-Aufnahme des Spitzenbereiches eines CD-AFM Cantilevers

Die CD-Cantilever weisen weiterhin eine am Ende ausgestellte Spitze mit Durchmessern von 850 nm bis 15 nm auf [4, 5], die es ermöglicht Hinterschneidungen, an den sonst nicht zugänglichen Seitenwänden der Probenoberfläche, abzubilden. Jedoch besitzen diese CD-Cantilever eine um Größenordnungen geringere Sensitivität, durch Kombination der inversen Steifigkeiten von Spitze und Cantilever, bei seitlichem Antasten der Spitze und sind anisotrop. Diese von CD-AFM oder auch 3D-AFM benutzten Cantilever, sind in der Halbleiterindustrie weit verbreitet

und werden zum Messen von Masken für die Lithographie, Rauheitsmessung und Prozesskontrolle, sowie für die geometrische Überprüfung komplexer Nanostrukturen verwendet. Dabei ist zu beachten, dass das bezeichnete 3D-AFM keine echt 3D-messenden Eigenschaften besitzt [2, 6]. Der verwendete Cantilever kann lediglich zwei orthogonale Richtungen auflösen, durch Konturschnitte wird ein dreidimensionaler Eindruck der Oberfläche erzeugt [6]. Da die fertigungsbedingt möglichen Geometrien von Mikro- und Nanostrukturen immer komplexer werden ist ein Messgerät gefordert, dass ähnlich einer Mikro-Koordinaten-Messmaschine (µKMG) dreidimensionale Strukturen auch dreidimensional antasten kann. Jedoch besitzen µKMGs derzeit Sonden mit einigen Mikrometern Spitzendurchmesser [7-15], der ein Antasten kleiner Radien und Strukturbreiten im unteren Mikrometerbereich verhindert und ein Messen von beispielsweise Mikroverzahnungen unmöglich macht. Um diesen Umständen zu begegnen wird derzeit an echt 3D-messenden Nano-Positioniermaschinen gearbeitet beispielsweise die NPMM-200 [16, 17]. Bedauerlicherweise sind Geräte bezüglich der taktilen Messung mit Spitzendurchmessern im Nanometerbereich noch in der Entwicklung.

Es gilt somit einen 3D-Nanotaster zu entwickeln, der Eigenschaften wie quasi-isotrope Steifigkeit des µKMG-Taster beinhaltet und gleichzeitig die Vorteile eines 3D-AFMs mit den geringen Spitzendurchmessern und hoher Eigenfrequenz der CD-Cantilever vereint.

Im Interesse der Zugänglichkeit für komplex geformte Proben, wird auf ein Cantilever-Design für die Sonde, der 3D-Nanotaster, nicht verzichtet. Weiterhin gilt es, ein ausgewogenes Verhältnis für die Sensitivität und Selektivität der Detektionsrichtungen zu erhalten und auf ein prototypengerechtes Design zu achten, das mit einer Ionenfeinstrahlanlage hergestellt werden kann.

Design des Nanotasters

Ausgehend von einem klassischen rechteckförmigem Cantilever mit rechteckigem Querschnitt, zeigt dieser eine Anisotropie bezüglich seiner Komponenten c_{ii} mit (i = x,y,z) der inversen Steifigkeitsmatrix für kleine Auslenkungen, $\underline{\boldsymbol{C}}_c$ siehe **Gl. (1)**. Das Verhältnis dieser Anisotropie liegt für (x/y/z) in der Größenordnung von (20/20/1). Weiterhin besitzt die symmetrische Matrix $\underline{\boldsymbol{C}}_c$ Elemente auf den Nebendiagonalen die ungleich Null sind. Diese Elemente haben ein Übersprechen der bisher nicht detektierbaren y-Richtung auf die z-Richtung (und umgekehrt) zur Folge. Dies führt zur Fehlinterpretation bei dem Bestimmen geometrischer Abmessungen und ist für den 3D-Nanotaster zu minimieren.

Die Ableitung der inversen Steifigkeitsmatrix ist die Matrix für die Auslenkwinkel des Cantilevers $\underline{\boldsymbol{A}}_c$ siehe **Gl. (2)**. Diese zeigt in Abhängigkeit der wirkenden Kräfte oder Verschiebung der Spitze eine Winkeländerung des vorderen freien Endes des Cantilevers. Die Winkeländerung wird über einen Lichtzeiger detektiert.

$$\underline{\boldsymbol{C}}_c = \begin{bmatrix} \frac{2L_{tip}^2}{L_{eff}^2} + \frac{t^2}{w^2} & 0 & 0 \\ 0 & \frac{3L_{tip}^2}{L_{eff}^2} & \frac{3L_{tip}}{2L_{eff}} \\ 0 & \frac{3L_{tip}}{2L_{eff}} & 1 \end{bmatrix} c \;;\; c = \frac{4L_{eff}^3}{Ewt^3} \qquad (1)$$

$$\underline{\boldsymbol{A}}_c = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{bmatrix} 0 & \frac{3L_{tip}}{L_{eff}^2} & \frac{3}{2L_{eff}} \\ \frac{2L_{tip}}{L_{eff}^2} & 0 & 0 \\ \frac{3t^2}{2w^2 L_{eff}} & 0 & 0 \end{bmatrix} \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} c \qquad (2)$$

Durch Kombination mit der inversen Steifigkeit einer biegeweichen Spitze $\underline{\boldsymbol{C}}_s$ (siehe **Gl. (2)**) kann die Isotropie hergestellt werden, jedoch skaliert die Größenordnung der Spitze mit ihrer Steifigkeit. Mit abnehmendem Durchmesser, bei gleichem Aspektverhältnis (üblicherweise 6:1), nimmt ebenfalls die Steifigkeit ab. Dies hat zur Folge das mit kleineren Spitzendurchmessern die Sensitivität abnimmt, beziehungsweise die minimal detektierbare Verschiebung zunimmt, welches die Unsicherheit vergrößert, da die Verschiebung der Spitze nicht detektiert wird, siehe **Gl. (4)**.

$$\underline{\boldsymbol{C}}_s = \begin{bmatrix} \frac{64L_s^3}{3Ed^4\pi} & 0 & 0 \\ 0 & \frac{64L_s^3}{3Ed^4\pi} & 0 \\ 0 & 0 & \frac{4L_s}{Ed^2\pi} \end{bmatrix} \qquad (3)$$

$$\underline{\boldsymbol{A}}_s = 0 \qquad (4)$$

Somit wird ein Design (**Abb. 3**) vorgestellt, bei dem die Nebendiagonalen nach einer Optimierung Null sind und die kombinierte inverse Steifigkeit für die Diagonalelemente $c_{xx} \approx c_{yy} \approx c_{zz}$ ist, ein Cantilever mit isotroper Steifigkeit. Dadurch wird ein unkontrolliertes Rutschen der Tastspitze des 3D-Nanotasters für ein Antasten einer Oberfläche in Normalenrichtung vermieden. Es wird ein an der yz-Ebene symmetrisches Design gewählt, da hier unter Berücksichtigung der zy- und yz-Komponenten die Nebendiagonalen bereits in guter Näherung Null sind. Das vorgestellte Design weist zwei Festkörpergelenke auf, die die Steifigkeit der Struktur beeinflussen. Der zweite Hals besitzt ein negatives Vorzeichen

bezüglich zy- und yz-Komponenten und kann somit zu deren Minimierung beitragen. Weiterhin lässt sich der nicht detektierbare, aber für ein von Übersprechen von α und β verantwortlicher Winkel γ reduzieren (**Gl. 2**). Die 3-Selektivität wird durch ein mit dem Reduzieren der Nebendiagonalen einhergehenden Minimieren der Auslenkung $\alpha(F_z)$ erreicht.

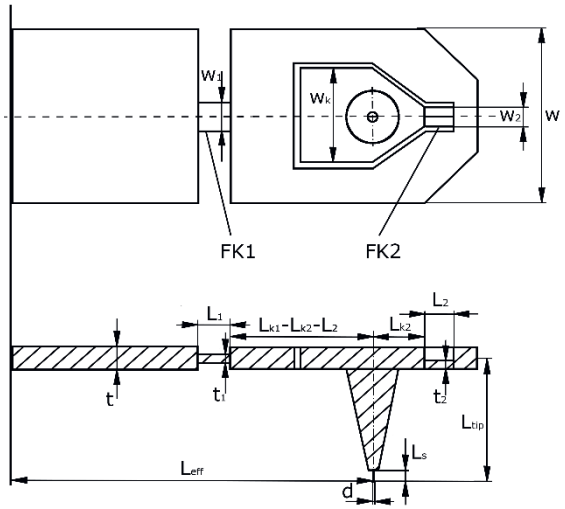


Abb. 3: Skizze des Designs des 3D-Nanotasters mit Variablen für die Berechnung (FK = Festkörpergelenk).

Der Kopfbereich des 3D-Nanotasters zeigt für eine Verschiebung in z-Richtung ein Heben und Senken (Auslesemöglichkeit: Interferometer) und für Verschiebungen bzw. Kräfte in x-Richtung oder y-Richtung eine Winkeländerung (Auslesemöglichkeit: Quadranten-Photodiode). Durch die Miniaturisierung sind Eigenfrequenzen von einigen 10 kHz bei Steifigkeiten von 1 N/m zu erwarten.

Modellbildung

Das oben in **Abb. 3** vorgestellte Design zeigt den 3D-Nanotaster mit zwei Festkörpergelenken. Dabei besitzt das Gelenk 1 einen rechteckigen Querschnitt und Gelenk 2 aus fertigungstechnischen Gründen der Ionenfeinstrahlanlage einen dreieckigen Querschnitt. Die Modellierung erfolgt mit Hilfe der Euler-Bernoulli Balkentheorie, da in den interessanten Bereichen die Querkräfte im Vergleich zu den anliegenden Momenten klein sind und elastische Auslenkungen im Bereich einiger zehn µrad zu erwarten sind. Es wird das Superpositionsprinzip für die Matrizen der inversen Steifigkeiten und der Auslenkwinkel angewendet (im Folgenden nicht weiter betrachtet). Dabei wird von einer festen Einspannung $\underline{\underline{C}}_{Basis} = 0$ ausgegangen und die Komponenten addiert (**Gl. (5)**).

$$\underline{\underline{C}}_{ges} = \underline{\underline{C}}_{Basis} + \underline{\underline{C}}_{c} + \underline{\underline{C}}_{Gelenk1} + \underline{\underline{C}}_{Kopf1} + \underline{\underline{C}}_{Gelenk2} + \underline{\underline{C}}_{Kopf2} + \underline{\underline{C}}_{s} \quad (5)$$

Somit ergibt sich die Inverse Steifigkeitsmatrix $\underline{\underline{C}}_{ges}$ des 3D-Nanotasters mit

$$\underline{\underline{C}}_{ges} = \begin{bmatrix} c_{xx} & c_{xy} & c_{xz} \\ c_{yx} & c_{yy} & c_{yz} \\ c_{zx} & c_{zy} & c_{zz} \end{bmatrix} \quad (6)$$

Exemplarisch wird die nominelle Steifigkeit c_{zz} des 3D-Nanotasters angegeben, wobei E den E-Modul des Materials beschreibt und $L_3 = L_2+L_{k2}$ ist.

$$c_{zz} = \frac{4L_{eff}^3}{Ewt^3} - \frac{4(L_{k1}+L_1)^3}{Ewt^3} + \frac{4(L_{k1}+L_1)^3}{Ew_1t_1^3} - \frac{4L_{k1}^3}{Ew_1t_1^3} + \frac{4L_{k1}^3}{Ewt^3} - \frac{4(L_k+L_2)^3}{Ewt^3} + \frac{4(L_k+L_2)^3}{E(w-w_k)t^3} - \frac{4L_{eff}^3}{E(w-w_k)t^3} + \frac{2L_2^3}{Ew_3t^3} - \frac{4L_{eff}^2 L_2}{Ewt^3} - \frac{6(L_{k1}+L_1)^2L_1}{Ewt^3} + \frac{6(L_{k1}+L_1)^2L_1}{Ew_1t_1^3} - \frac{6L_{k1}^2L_3}{Ew_1t_1^3} + \frac{6L_{k1}^2L_3}{Ewt^3} - \frac{6(L_k+L_2)^2L_3}{Ewt^3} + \frac{6(L_k+L_2)^2L_3}{E(w-w_k)t^3} - \frac{4L_2^2L_3}{E(w-w_k)t^3} + \frac{4L_2^2L_3}{Ew_3t^3} + \frac{12L_3^3}{Ew_2t_2^3} - \frac{12L_{k2}^3}{Ew_2t_2^3} + \frac{4L_{k2}^3}{Ew_kt^3} - \frac{6L_{eff}^2L_3}{Ewt^3} - \frac{6(L_{k1}+L_1)^2L_3}{Ewt^3} + \frac{6(L_{k1}+L_1)^2L_3}{Ew_1t_1^3} - \frac{6L_{k1}^2L_3}{Ew_1t_1^3} + \frac{6L_{k1}^2L_3}{Ewt^3} - \frac{6(L_k+L_2)^2L_3}{Ewt^3} + \frac{6(L_k+L_2)^2L_3}{E(w-w_k)t^3} - \frac{6L_2^2L_3}{E(w-w_k)t^3} + \frac{3L_2^2L_3}{Ew_3t^3} + \frac{12L_3^2L_{eff}}{Ewt^3} - \frac{12(L_k+L_2)L_3^2}{Ewt^3} + \frac{12(L_k+L_2)L_3^2}{Ew_1t_1^3} - \frac{12L_3^2L_{k1}}{Ew_1t_1^3} + \frac{12L_3^2L_{k1}}{Ewt^3} - \frac{12(L_k+L_2)L_3^2}{Ewt^3} + \frac{12(L_k+L_2)L_3^2}{E(w-w_k)t^3} - \frac{12L_2L_3^2}{E(w-w_k)t^3} + \frac{6L_2L_3^2}{Ew_3t^3} + \frac{4L_s}{Ed^2\pi} \quad (7)$$

Dieses Modell wird mit einer FEM-Simulation (*Ansys 18.0*) überprüft.

Optimierung und Vergleich mit Finite Elemente Simulation

Ziel der Optimierung ist, wie bereits beschrieben, eine isotrope Steifigkeit des 3D-Nanotasters, bei einer ersten Eigenfrequenz von >50 kHz (**Abb. 4)**, mit ausgewogener möglichst hoher Sensitivität.

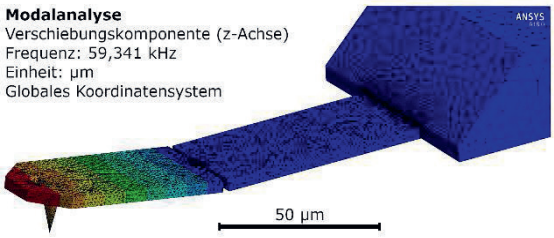


Abb. 4: Grundmode des 3D-Nanotasters

Dementsprechend werden die inversen Steifigkeiten $c_{xx} \approx c_{yy} \approx c_{zz}$ und $c_{yz} = c_{zy} = 0$ (**Abb. 5**), sowie die Auslenkwinkel $\alpha(F_z) \approx 0$ optimiert und $\gamma(F_z)$ minimiert (**Abb. 6**).

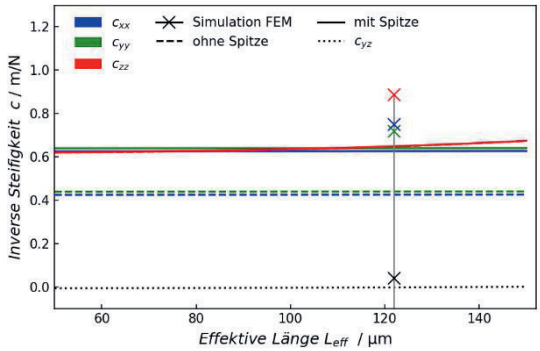


Abb. 5: Inverse Steifigkeiten des Modells und der Simulation über der Effektiven Länge des Cantilevers, mit und ohne der CD-Spitze.

Die **Abb. 5** zeigt den Einfluss der sonst die Steifigkeit bestimmenden effektiven Länge des Cantilevers für den 3D-Nanotaster. Der Kurvenverlauf zeigt, dass die mechanischen Eigenschaften der Struktur durch die Festkörpergelenke dominiert wird. Für eine zuvor durch Optimierung festgelegte Struktur, weist die FEM-Simulation eine Abweichung bezüglich der Inversen Steifigkeiten auf. Die inverse Steifigkeit ist insbesondere für die nominelle inverse Steifigkeit erhöht, welches auf die nicht im Modell berücksichtigten elastischen Einspannungen der Festkörpergelenke zurückzuführen ist.

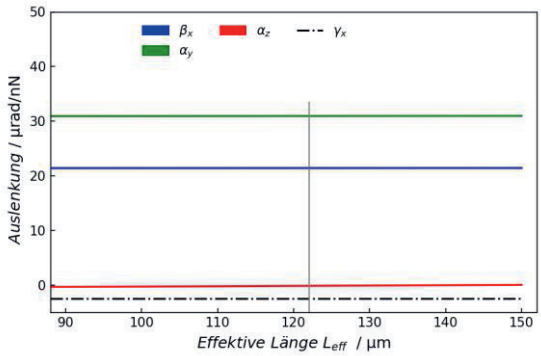


Abb. 6: Winkelauslenkung des Cantilevers für die an der Spitze wirkenden Kräfte bei einem Nanonewton über der effektiven Länge.

Aus **Abb. 6** wird ersichtlich das mit dem 3D-Nanotaster Antastkräfte von weniger als 1 nN detektierbar sein sollten. Die Abbildung zeigt eine ähnliche Auslenkung für an der Spitze wirkende Kräfte, für die x- und y-Richtungen, und das zuvor geforderte minimierte Verkippen des Kopfbereiches für eine Kraft in z-Richtung. Die Verbesserung der Eigenschaften ist in **Tab. 1** dargestellt.

Tab. 1: FEM-Simulierte und Modellbasierte Eigenschaften des 3D-Nanotasters im Vergleich zum CD-Cantilever mit d = 64,5 nm*

Eigenschaft (x/y/z)	CD-Cantilever	3D-Nanotaster
Sensitivität in µrad/nm	(0,28/0,50/12,0)	(34,2/48,4/0,21)
Isotropie Steifigkeits-verhältnis	(7,09/7,13/1,00)* (6,68/6,71/1,00)	(0,84/0,81/1,00)* (0,97/0,99/1,00)

Es wird gezeigt, dass die Isotropie um den Faktor sieben verbessert wurde bei gelichzeitiger 95-fachen Erhöhung der Sensitivität.
Die Grundmode von 59,3 kHz (Material: anisotropes Silizium) wurde mit einer Modalanalyse simuliert.

Diskussion

Das vorgestellte Design des 3D-Nanotasters zeigt, dass durch modellbasierte Optimierung ein quasi isotroper Taster erzeugbar ist, bei dem die Spitze dezentral angeordnet ist. Jedoch tritt bei dem Vergleich der Simulation zum Modell eine Abweichung auf. Diese Abweichung ist auf den Vergleich eines linearen Modells mit einer nichtlinearen Simulation zurückzuführen, die die Einspannbedingungen der Festkörpergelenke abbildet, wobei der Schubeinfluss vernachlässigbar klein ist. Diese nicht berücksichtigten Einspannbedingungen sind maßgeblich für die Abweichungen in den inversen Steifigkeiten. Es existieren Festkörpergelenke die in ihrer Geometrie einen Übergangsbereich bezüglich des eigentlichen Verformungsbereichs des Gelenkes zeigen und die Relevanz der Einspannung verringern, Fertigungsbedingt wurde jedoch eine rechteckige Form der Festkörpergelenke gewählt. Weiterhin sind aufgrund der Fertigungsabweichung, der einige 100 nm dicken Strukturen, Abweichungen in dieser Größenordnung von 20% zu erwarten, so dass dieses Modell für eine Optimierung verwendet werden kann. Die Idee wurde in einem Patent (Nr.: 10 2018 105 756 B3) festgehalten.

Zusammenfassung

Das Modell des isotropen Cantilevers auf Basis eines CD-AFM Cantilevers (CD = Critical Dimension), ist eine Möglichkeit einen 3D-Nanotaster zu erzeugen. Dabei zeigt der quasi isotrope 3D-Nanotaster eine 95-fach erhöhte Sensitivität im Vergleich zu bisherigen Sonden mit Spitzendurchmessern von nominell 70 nm (CD-AFM Cantilever). Der 3D-Nanotaster wurde so optimiert, dass sein freies Ende (Kopfbereich) für ein Anzeigesystem benutzt

werden kann. Eine Kraft senkrecht zu seiner Längsachse und Probenebene in z-Richtung führt zu einer Verschiebung des Kopfbereiches, orthogonal orientierte Kräfte dazu in x- und y-Richtung zu einer Verkippung des Kopfbereiches. Somit wäre ein Interferometer mit gekoppelter Quadranten-Photodiode eine Ausleseoption [16]. Die Grundmode des Designs für das Material Silizium liegt bei 59,3 kHz. Das Design findet sich im Patent Nr.: 10 2018 105 756 B3.

Danksagung

Spezieller Dank gilt Jannick Langfahl-Klabes für die Aufnahmen im Rasterelektronenmikroskop, sowie der DFG finanzierten Research Training Group 1952 Metrology for Complex Nanosystems, *NanoMet* und International Graduate School of Metrology *B-IGSM*.

Literaturnachweis

[1] G. Binning, C.F. Quate, Ch. Gerber, *Physical Review Letters* 56, 6 (1986); doi: 10.1103/PhysRevLett.56.930
[2] Y. Matrin, K Wickramasinghe, *Appl. Phys. Lett*. 64, 2498 (1994); doi: 10.1063/1.111578
[3] G. Dai, M. Heidelmann, C. Kübel, R. Prang, J. Fluegge, H. Bosse, *Meas. Sci. Technol*., 24, 085001 (2013); doi: 10.1088/0957-0233/24/8/085001
[4] Team Nanotec GmbH, Wilhelm-Schickardt-Str. 10, 78052 Villingen-Schwenningen, Germany https://www.team-nanotec.de/index.cfm
[5] nanotools GmbH, Reichenbachstraße 33 D - 80469 München https://www.nanotools.com/products/blue-line/cdr-ebd/
[6] R. Dixson, N. Orji, I. Misumi, G. Dai, *Ultramicroscopy. 194*, (2018); doi: 10.1016/j.ultramic.2018.08.011.
[7] V. Nesterov, P. Pornnoppadol, U. Brand, (2003). *Metrological characterization of a new three-axis piezoresistive micro probe*, Proc. of Euspen 2002 s. 737-740
[8] G. Dai, S. Bütefisch, F. Pohlenz, H.-U. Danzebrink, *Measurement Science and Technology,* 20. 084001, (2009); doi: 10.1088/0957-0233/20/8/084001
[9] J. Claverley, R. Leach, (2011). *Verifying the capability of a three axis vibrating micro-scale CMM probe to operate in a non-contact mode*. Proceedings - ASPE 2011 Annual Meeting. 52.
[10] J. Cui, J. Li, K. Feng, J. Tan, *Optics Express*. 23. 26386-26398. (2015); doi: 10.1364/OE.23.026386.
[11] K. Alblalaihid, et. al., *Sensors*. 16. 492. (2016); doi: 10.3390/s16040492
[12] L. Zou, H. Ni, P. Zhang, X. Ding, *Sensors (Basel, Switzerland)*, *17*(11), 2652 (2017); doi:10.3390/s17112652
[13] R.-L. Li, P. Xu, P.-Y. Wang, K.-C. Fan, R.-J. Cheng, (2018). *Sensors.* 18. 4229. (2018); doi: 10.3390/s18124229
[14] BS Precision Engineering Deutschland GmbH, Leitzstraße 45, 70469 Stuttgart, Germany http://www.ibspe.com/category/triskelion-touch-probe.htm
[15] Werth Messtechnik GmbH, D-35394 Gießen https://www.werth.de/fileadmin/media/pdf/Fachartikel/FA-Messen_von_Mikrogeometrien_mit_dem_Fasertaster-Mikroproduktion_02.2016.pdf
[16] Felix G Balzer *et al*, *Meas. Sci. Technol.* 22 094018, (2011); doi:10.1088/0957-0233/22/9/094018
[17] G. Jäger, E. Manske, T. Hausotte, A.Müller, F. Balzer, *Surf. Topogr.: Metrol.* Prop.4 034004 (2016); doi: 10.1088/2051-672X/4/3/034004

DOI 10.5162/sensoren2019/P1.21

Simulation von thermischen Einflüssen auf MEMS-Mikrofone

Hanna Ebbinghaus[1], Gregor Feiertag[1], Sebastian Walser[2]
[1]Hochschule München, Lothstraße 64, 80335 München
[2]TDK Electronics AG, Rosenheimer Straße 116b, 81669 München
Kontakt: hanna.ebbinghaus@hm.edu

Zusammenfassung

Um die Genauigkeit von Flip-Chip MEMS-Mikrofonen zu verbessern, wurden Untersuchungen zum Temperaturverhalten der Empfindlichkeit durchgeführt. Es wurde das Verhalten auf einen Temperaturzyklus gemessen. Die Messungen zeigten, dass die Umgebungstemperaturen Einfluss auf die Empfindlichkeit des Mikrofons haben und eine Temperaturhysterese verursachen. Anschließend wurde das Verhalten auf den gleichen Temperaturzyklus simulativ berechnet. Eine Übereinstimmung zwischen Messung und Simulation liegt vor. Dies ermöglicht es, weitere Untersuchungen simulativ durchführen zu können.

Keywords: MEMS- & Nanosensoren, Modellbildung, Simulation und Validierung

1 Einleitung

Aufgrund ihrer sehr kleinen Bauform und ihren guten elektroakustischen Eigenschaften sind MEMS-Mikrofone schon seit längerem Stand der Technik in der Unterhaltungselektronik. Im Bereich der Mobilkommunikation werden inzwischen fast ausschließlich MEMS-Mikrofone verwendet. Für die Störgeräuschunterdrückung ist es wichtig, dass sich die Empfindlichkeiten der in einem Gerät verwendeten Mikrofone möglichst wenig unterscheiden [1]. Streuungen kleiner ± 1 dBV/Pa bei Empfindlichkeiten von - 38 dBV/Pa zählen derzeit zum Stand der Technik [2]. Zu Empfindlichkeitsstreuungen kann es aufgrund von mechanischen oder thermomechanischen Spannungen kommen. In diesem Beitrag wird die Temperaturhysterese eines MEMS-Mikrofons mittels Simulation berechnet. Die Simulationen werden mit experimentellen Untersuchungen verglichen

1.1 Flip-Chip MEMS-Mikrofon

Bei dem in diesem Beitrag untersuchten Mikrofon handelt es sich um ein Mikrofon, das in Flip-Chip-Montage aufgebaut wird (siehe Abb.1). [3]
Durch die Flip-Chip Montage der Chips können kleinere Bauteilgrößen erreicht werden. In Bild 2 ist zu erkennen, dass das Mikrofon aus einem differentiellen Sensor-Chip und einem ASIC-Chip (Application-specific integrated circuit chip) besteht. Der Schalleintritt erfolgt durch Löcher im HTCC Substrat (High Temperature Cofired Ceramic). Das sehr kleine Vorvolumen wird durch eine Polymerfolie vom Rückvolumen getrennt. Das Rückvolumen wird mit Hilfe einer Metallkappe verschlossen. [3]

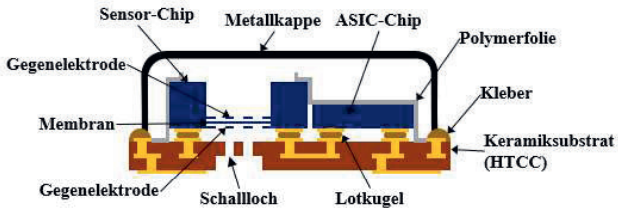


Abb. 1 Schematischer Querschnitt eines Flip-Chip MEMS-Mikrofons [3]

1.2 Differentieller Sensor-Chip

MEMS-Mikrofone bestehen aus einem elektroakustischen Wandler (Sensor-Chip) und einem ASIC-Chip. Der Sensor-Chip ist ein kapazitiver Wandler mit einer beweglichen Membran und zwei perforierten Gegenelektroden, siehe Abb. 2. [3]

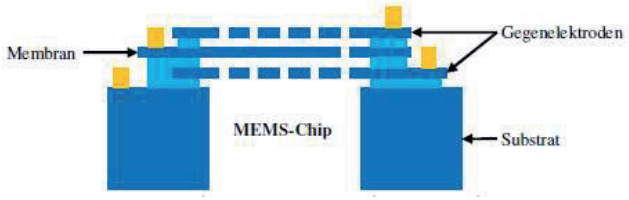


Abb. 2 Aufbau eines differentiellen Sensor-Chips mit Membran und zwei Gegenelektroden [2]

Die etwa 500 nm dünne Silizium-Membran folgt den Druckschwankungen des Schallfeldes. Über eine, zwischen Membran und den beiden Gegenelektroden angelegte, elektrische Spannung wird die Schwingung anhand der auftretenden Kapazitätsänderung in ein differentielles elektrisches Signal umgewandelt. [3]

2. Thermisches Verhalten von MEMS-Mikrofonen

Erste Untersuchungen zum thermischen Verhalten von MEMS-Mikrofonen zeigten, dass sich die Empfindlichkeit der Mikrofone durch eine Temperatureinwirkung ändert. Bei Temperaturen unter oder über der Raumtemperatur entstehen aufgrund der unterschiedlichen thermischen Ausdehnungskoeffizienten mechanische Spannungen. Wird das Mikrofon wieder auf Raumtemperatur gebracht, entsteht so ein anderer Spannungszustand als vor der Temperaturbelastung. Dies entspricht einer Temperaturhysterese der Empfindlichkeit des Mikrofons. Aufgrund der Temperaturbelastung ändert sich die Zugspannung der Membran. Dadurch kommt es zu Veränderungen in der Empfindlichkeit. Das thermische Verhalten der MEMS-Mikrofone wird zum einen experimentell und zum anderen simulativ mit der Finiten Elemente Methode (FEM) bestimmt. [4]

2.1 experimentelle Untersuchungen

Zur experimentellen Bestimmung des Temperaturverhaltens wurde untersucht, wie sich MEMS-Mikrofone verhalten, wenn sie groben Änderungen der Umgebungstemperatur ausgesetzt werden [5]. Dafür wurden die Mikrofone einem Temperaturzyklus ausgesetzt. Dies wurde für 20 Mikrofone durchgeführt. Zuerst wurde die Empfindlichkeit bei Raumtemperatur bestimmt, um im späteren Verlauf die Empfindlichkeitsänderung zu bestimmen. Anschließend wurde der Temperaturzyklus gestartet. Dafür werden die Mikrofone zunächst für 30 Minuten bei 85 °C gelagert. Nach einer Wartezeit von fünf Minuten bei Raumtemperatur wird die Empfindlichkeit erneut bestimmt. Nach weiteren fünf Minuten bei Raumtemperatur werden die Mikrofone für eine Dauer von 30 Minuten bei – 40 °C gelagert. Auch hier wird nach einer Wartezeit von fünf Minuten bei Raumtemperatur die Empfindlichkeit erneut bestimmt. Dies wird insgesamt dreimal durchgeführt. Nach der letzten Wiederholung wird noch eine Messung nach einer Lagerung nach 85 °C durchgeführt. Abschließend werden die Mikrofone noch für eine halbe Stunde bei Raumtemperatur gelagert und erneut die Empfindlichkeit bestimmt. Abbildung 3 zeigt den Messplan.

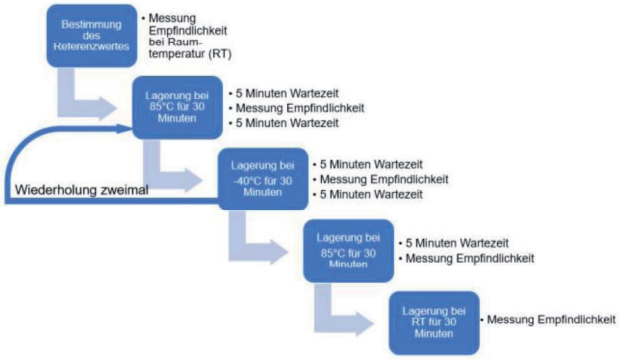


Abb. 3: Schematische Darstellung des Messplans

Die Empfindlichkeit der Mikrofone wurde in einer Messvorrichtung mit einer Druckkammer bestimmt. Abb. 4 zeigt den schematischen Messaufbau.

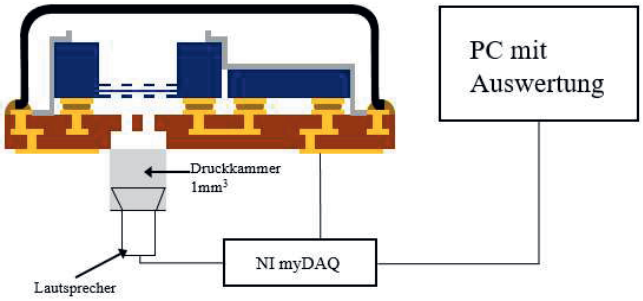


Abb. 4: Schematische Darstellung des Messaufbaus

Das Mikrofon wird zur Bestimmung der Empfindlichkeit an die Druckkammer angeschlossen. Das Mikrofon wird mit einer Schallwelle mit einer Frequenz von 1 kHz bei einem Schalldruck von 1 Pa beschallt [4].

Für die Empfindlichkeitsänderung wurde der Mittelwert der Messungen bestimmt. Es wurde die Differenz zwischen dem Anfangswert und den Werten nach der Temperaturbelastung bestimmt. In Abb. 5 ist das Ergebnis der experimentellen Bestimmung dargestellt. Auf der Ordinate ist die Empfindlichkeitsänderung dargestellt und auf der Abszisse, die Temperatur, bei der die Mikrofone für 30 Minuten gelagert wurden. Die Messung fand bei Raumtemperatur statt.

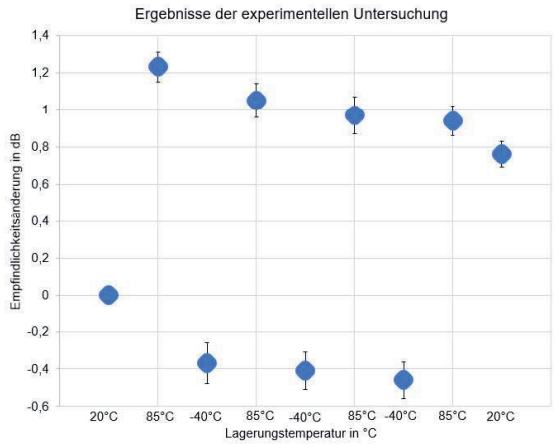


Abb.5: gemessene Änderung der Empfindlichkeit relativ zum Startwert. Messung bei RT nach Lagerung bei den auf der Ordinate angegebenen Temperaturen

In Abb. 5 ist erkennbar, dass sich die Empfindlichkeit des MEMS-Mikrofons durch Temperatureinwirkung ändert. Dies liegt an den thermisch induzierten Spannungen. Diese thermisch induzierten Spannungen führen dazu, dass die Empfindlichkeit bei den späteren Messungen nach den Lagerungen bei 85 ° C und – 40 ° C nicht auf den gleichen Wert zurück gehen. Dass die Empfindlichkeit nach der letzten Lagerung bei 20 ° C nicht auf null zurück geht, liegt an dem Kriechverhalten der Lotstellen.

2.2 Simulation

Für die FEM-Berechnungen wird die Kapazitätsänderung bestimmt, da sich die Kapazitätsänderung proportional zur Empfindlichkeitsänderung verhält. Die Simulationen werden mit der Open-Source-Software Code-Aster durchgeführt [6]. Dass sich diese Software für Simulationsberechnungen eignet, wurde schon in [7] gezeigt. In [7] wurde gezeigt, dass sich die Membran bei einem einwirkenden Druck kugelförmig verformt (siehe Abb. 6).

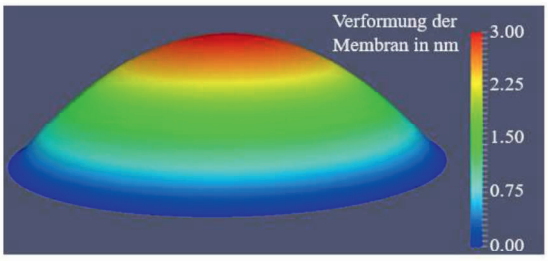


Abb. 6: Verformung der Membran eines MEMS-Mikrofons [7]

Die Kapazität eines Kondensator, bei dem eine der Platten kugelförmig verformt ist, berechnet sich nach Gl. (1) [8].

$$C = \frac{\varepsilon_0 * A}{g_0} * \left(1 + \frac{1}{3}\frac{g'}{g_0}\right) \qquad \text{Gl. (1)}$$

Dabei entspricht ε_0 der elektrischen Feldkonstante, A der Fläche, g_0 der Anfangsabstand zwischen der Membran und der Gegenelektrode und g' der Auslenkung in der Mitte der Membran bei einem einwirkenden Druck. Nach Gl. (1) wurde die Kapazität in der Simulation bestimmt.

Für die Simulation wurde das geometrische Modell des Mikrofons erstellt und vernetzt. ES wurde das folgende Modell dafür verwendet (Abb. 7). Bei den Lotstellen handelt es sich um Zylinder, die mit einem elastisch-plastischen Materialmodell simuliert worden sind.

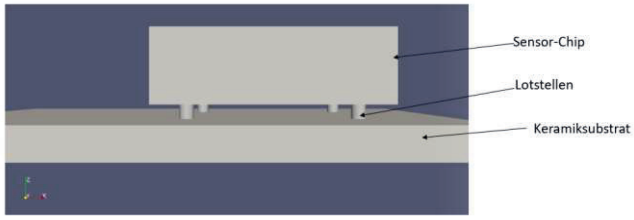


Abb. 7: Modell des simulierten Mirkofons

In der Simulation wurde der gleiche Temperaturzyklus (siehe Abb. 3) berechnet. Es wurde für jede Temperatur eine Simulation bei + 1 Pa und – 1 Pa durchgefürt, da dies einer Schallwelle mit einer Frequenz 1 kHz und 1 Pa Schalldruck entspricht. Für jede Temperatur wurde die Differenz bestimmt. Anschließend wurde das logarithmische Verhältnis zum ersten simulierten Wert (=Ausgangswert) bestimmt. Abb. 8 zeigt das Ergebnis der Simulation.

In Abb. 8 geht die Empfindlichkeit nach einer simulierten Lagerung bei den Temperaturen 85 ° C und – 40 ° C nicht auf den Startwert zurück. Dies liegt an den thermisch induzierten Spannungen im MEMS-Mikrofon. Dass der letzte Wert nach einer simulierten Lagerung der Mikrofone bei 20 ° C nicht weiter abnimmt, liegt daran, dass in der Simulation ein elastisch-plastisches Materialverhalten der Lotstellen angenommen wurde und das Kriechverhalten nicht berücksichtigt wurde.

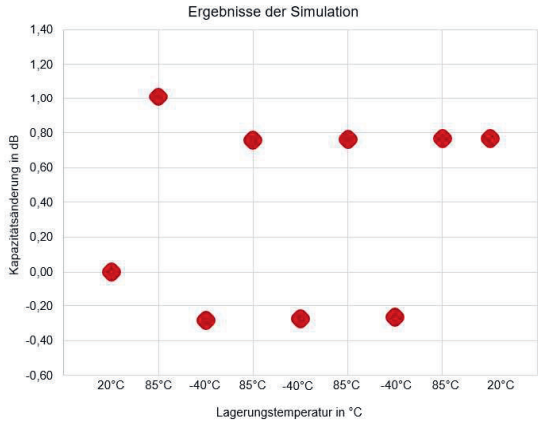


Abb. 8: simulaiv bestimmte Änderung der Kapazität relativ zum Startwert. Simulation bei RT nach Lagerung bei den auf der Ordinate angegebenen Temperaturen

3 Vergleich von Experiment und Simulation

Abb. 9 zeigt einen Vergleich von den Messergebnissen und den FEM-Berechnungen.

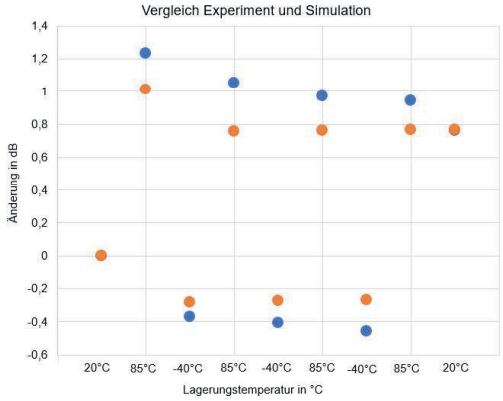


Abb. 9: Vergleich Messung und Simulation

Sowohl die Messungen als auch die Simulationen zeigen, dass die Empfindlichkeit bei Temperaturen über der Raumtemperatur zunimmt. Bei Temperaturen unter der Raumtemperatur nimmt die Empfindlichkeit ab.. Dies zeigt, dass die Temperatur einen Einfluss auf die Empfindlichkeit des Mikrofons hat.

Es ist zu erkennen, dass die Tendenz der Ergebnisse bei der Messung und bei der Simulation gleich ist. Allerdings zeigt Abb. 9 auch, dass die Ergebnisse der Simulation im Schnitt um ca. 20 % abweichen. Des Weiteren lässt sich in Abb. 9 erkennen, dass die Empfindlichkeit nicht mehr auf den Referenzwert zurückgeht. Bei den Messungen nimmt die Empfindlichkeit im späteren zeitlichen Verlauf dennoch weiter ab. Dies liegt an dem Kriechverhalten der Lotstellen. Das Kriechen wurde in der Simulation nicht berücksichtigt.

4 Zusammenfassung und Ausblick

In unserem Beitrag wurde das thermische Verhalten von MEMS-Mikrofonen experimentell und simulativ bestimmt. Es hat sich gezeigt, dass die Ergebnisse gut übereinstimmen. Die Simulation liefert glaubwürdige Ergebnisse.

Durch die Simulation wird es möglich sein, neue Mikrofonaufbauten auf ihr Temperaturverhalten zu prüfen. Außerdem soll der Einfluss des Kriechverhaltens der Lotstellen auf die Empfindlichkeit untersucht werden.

Literatur

[1] Z.-H. Fu, F. Fan, J.-D. Huang; *Dual-microphone noise reduction for mobile phone application*; in Acoustic, Speech and Signal Processing Conference, Vancouver, Canada, May 26-31, 2013, pp. 7239-7243

[2] S. Walser, G. Feiertag, M. Loibl, W. Pahl, M. Winter, C. Siegel, A. Leidl; *MEMS-Mikrofone mit hohem Signal zu Rausch Abstand und geringen Sensitivitätsstreuungen;* in Tagungsband Mikrosystemtechnik Kongress 2015, Karlsruhe, Deutschland, 26. – 28. Oktober 2015, S. 171-174

[3] S. Walser, C. Siegel, M. Winter, G. Feiertag, M. Loibl, A. Leidl; *MEMS microphones with narrow sensitivity distribution;* Sensors and Actuators A; vol. 247; pp. 663-670, 2016

[4] H. Ebbinghaus, A P. Chew, G. Feiertag, S. Walser; *Messung und Berechnung der Temperaturhysterese von Flip-Chip MEMS-Mikrofonen*, Tagungsband Mikrosystetechnikkongress 2017, München, Deutschland, 23. – 25. Oktober 2017 München

[5] C. Philipp: *Analyse temperaturbedingter Relaxationsprozesse bei MEMS-Mikrofonen*, Masterarbeit 2016, HS München, Deutschland

[6] www.code-aster.org

[7] H. Ebbinghaus, G. Feiertag, S. Walser: *Calculation of the Deformation of a Circular Fixed Silicon Membrane with Prestress by Finite Element Analysis,* Sensoren und Messsysteme 2018, Nürnberg, Deutschland, 26. – 27. Juni 2018

[8] D. T. Martin, *Design, Fabrication, and characterization of a MEMS dual-backplate capacitive microphone*, Dissertation 2007, University of Florida, USA

DOI 10.5162/sensoren2019/P1.22

Dynamic measurement methods for solid electrolyte gas sensors

Daniela Schönauer-Kamin[1], Nils Donker[1], Anastasiya Ruchets[2], Jens Zosel[2], Ulrich Guth[2] und Ralf Moos[1]
[1]Lehrstuhl für Funktionsmaterialien, Universität Bayreuth, 95440 Bayreuth, Deutschland
[2]Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg, 04736 Waldheim, Deutschland

Summary

Dynamic measurement methods are an appropriate tool for the realization of highly selective solid-electrolyte based gas sensors. Three dynamic methods are used in this work to investigate Pt|YSZ-based sensors. The polarization of Pt|YSZ sensors with alternating voltage pulses and the measurement of the resulting self-discharge behavior leads to a very sensitive and selective NO_x sensor. By using the dynamic cyclic voltammetry method, it is possible to detect H_2 and O_2 simultaneously and selectively in low concentration ranges under certain measuring conditions. The third method, the coulometric measurement method, is used for the determination of the amount of exchanged oxygen during polarization of the system Pt|YSZ. For 400 °C, an uptake of O_2 of the polarized Pt|YSZ sensor was found.

Keywords: dynamic sensor systems, solid-electrolyte, Pt-YSZ, gas sensor

Novelty

Dynamic methods provide a great potential to realize sensitive and highly selective electro-chemical gas sensors. This study shows examples for applications of dynamic methods. The pulsed polarization method of Pt-YSZ sensors to detect NO_x and dynamic cyclic voltammetry, a tool for selective measurement of H_2 and O_2 in non-equilibrated gas mixtures by commercial solid electrolyte gas sensors based on Pt-YSZ are presented.

Introduction

The detection of various gas concentration, e.g. NO_x, is still a huge task considering exhaust-aftertreatment systems and air quality monitoring applications. Typical gas sensing devices measure during stationary equilibria meaning the sensor is in equilibrium with the surrounding atmosphere and a static electrical property like resistance or potential difference is measured in dependence on the analyte concentration. A new approach utilizes dynamic interactions between gas phase and polarized electrodes of solid electrolyte based sensors, here in the system Pt|YSZ. The dynamic behavior provides on the one side a possibility to improve the sensing characteristics like sensitivity and selectivity and on the other side enables investigation of processes occurring at the three-phase boundary Pt|YSZ|gas, at the electrode and at the electrolyte. Here, the dynamic pulsed polarization method [1], the cyclic voltammetry (CV) analysis of solid electrolyte sensors [2] and the coulometric measurement of oxygen exchange [3] are combined for investigation of the dominating processes.

Experimental

The dynamic behavior of the system Pt|YSZ is investigated by pulsed polarization technique and by dynamic cyclic voltammetry.

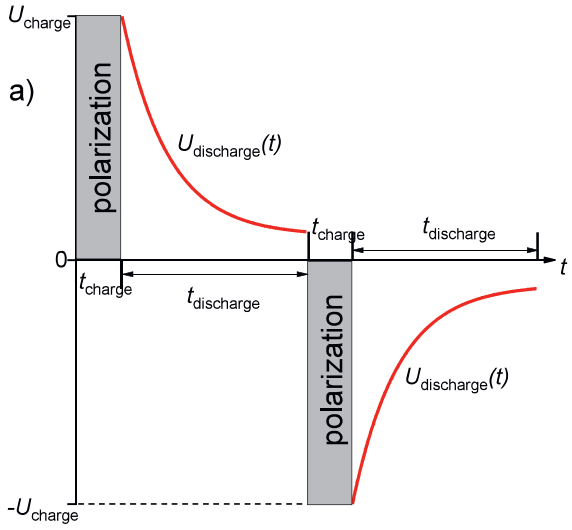


Fig. 1: Scheme of pulsed polarization method with constant voltage polarization, adapted from [1].

Fig. 1 shows a scheme of the pulsed polarization technique. A sensor consisting of two Pt electrodes screen-printed on both sides of an 8YSZ disk (Pt|YSZ|Pt) is heated in a tube furnace to 400 °C. The base gas consists of 10 % O_2, 2 % H_2O and N_2 as balance. The sample is polarized by alternating voltage pulses of 1 V for 1 s and the self-discharging (depolarization voltage) of the sensor is measured during 10 s of depolarization [1]. The dynamic response is analyzed for varying O_2 concentrations in dependence of the NO concentration.

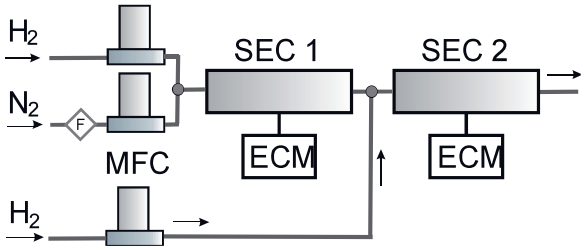


Fig. 2: Experimental setup for dynamic CV measurements, SEC = solid electrolyte cell, ECM = potentiostat, MFC = massflow controller

The experimental setup for dynamic CV measurements (Fig. 2) shows two commercial solid electrolyte Pt|YSZ|Pt cells (SEC). Main added gases are N_2 and H_2. SEC 1 is polarized by a defined pumping voltage to adjust the O_2 concentration. SEC 2 is the measuring cell, measuring by CV [2] in dependence of O_2 and H_2 concentrations. Another setup for the coulometric determination of the oxygen exchange of a heated Pt|YSZ sensor presents Fig. 3 [3]. Here, defined O_2 concentrations are adjusted by cell 1 (SEC) (e.g. 1 ppm O_2). A voltage pulse of 1 V is applied for 1 s to the Pt|YSZ|Pt sensor (at 400 °C) to activate the O_2 exchange reactions and the resulting O_2 concentration (sum of dosed c_{O2} and O_2-uptake or release of the polarized sensor) is measured downstream by cell 2.

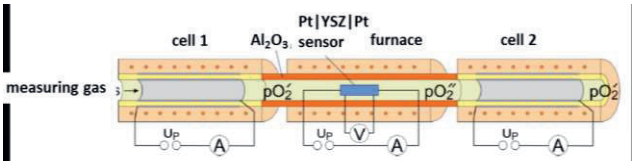


Fig. 3: Setup for coulometric measurement of oxygen exchange of heated Pt|YSZ|Pt sensor, adapted from [3].

Results and Discussion

Fig. 4 shows the discharge voltages after positive polarization pulse of the Pt|YSZ|Pt sensor for different NO concentrations and for different signal evaluation times (1, 2, 4, and 8 s) at 400 °C. The discharge voltage decreases with increasing NO concentration. That means that NO accelerates the self-discharge behavior resulting in a faster discharge. Additionally, the O_2 concentration shows a comparable effect, the discharging is faster with increasing O_2 concentration. Various effects contribute to the sensing mechanism, which is not fully understood yet.

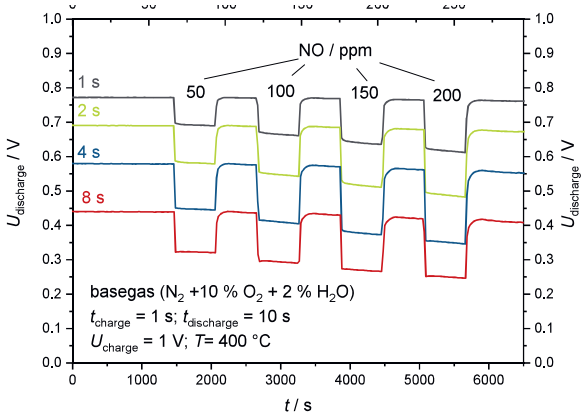


Fig. 4: Discharge voltage curves for varying evaluation times of discharge curves and different NO concentrations of Pt|YSZ|Pt sensor at 400 °C.

The results in Fig. 5 show the CV analysis of SEC 2 of the setup shown in Fig. 2 at 750 °C. A hydrogen related peak appears in anodic scan direction at low scan rates and an oxygen related peak occurs at higher scan rates in cathodic direction. This clear difference between O_2- and H_2-related peaks allows determining the concentration of the both gases simultaneously with high selectivity. Additionally, reaction steps on electrode processes can be analyzed.

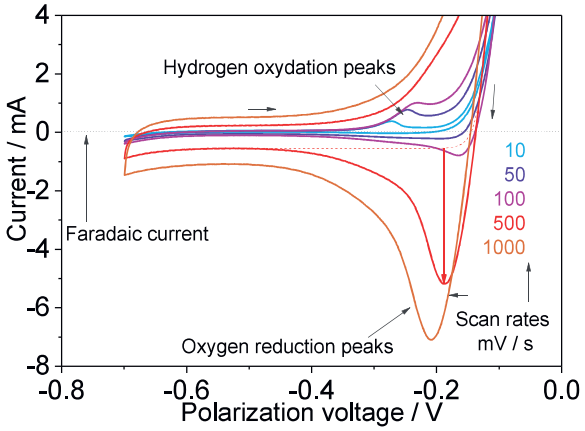


Fig. 5: Cyclic voltammograms CV for different scan rates for SEC 2 at 750 °C for 50 ppm H_2 and 10 ppm O_2 in N_2.

Fig. 6 shows the measured current of cell 2 of the oxygen exchange measurement (setup see Fig. 3) of a Pt|YSZ|Pt sensor at 400 °C with 10 ppm O_2. The polarization of the sensor with 1 V seems to result in an uptake of oxygen by the sensor device. One explanation could be the formation of PtO_x on the Pt electrodes. These effects need further investigation.

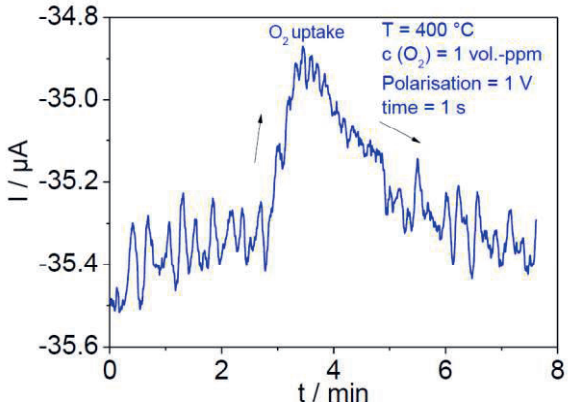


Fig. 6: Measured current at cell 2 for oxygen exchange measurement of Pt|YSZ|Pt sensor at 400 °C and 10 ppm O_2.

References

[1] S. Fischer, R. Pohle, B. Farber, R. Proch, J. Kaniuk, M. Fleischer, R. Moos, Method for detection of NO_x in exhaust gased by pulsed discharge measurements using standard zirconia-based lambda sensors, *Sens. Act. B: Chem.* **147** (2010) 780 – 785, doi: 10.1016/j.snb.2010.03.092.

[2] M. Schelter et al.: Electrolyte sensor for the analysis of gaseous mixtures, *J. Sens. Sens. Syst.* **5** (2016) 319 – 324, doi: 10.5194/jsss-5-319-2016

[3] J. Zosel, A. Herms, J. Yao, M. Schelter, V. Vashook, U. Guth, M. Mertig, Oxygen solid electrolyte coulometry: Chances, challenges and new approaches, Proceedings of *International Meeting on Chemical Sensors* IMCS July 2018, Vienna, Austria.

DOI 10.5162/sensoren2019/P1.23

CH_4-Sensitivity of Thermoelectric Gas Sensors

Julia Wohlrab[1], Gunter Hagen[1], Heinz Kohler[2], Ralf Moos[1]
[1] Department of Functional Materials, University of Bayreuth, D-95440 Bayreuth, Germany
[2] Institute for Sensor and Information Systems (ISIS), D-76133 Karlsruhe, Germany

Abstract

Recently, thermoelectric gas sensors were used for in-situ determination of the CO/HC-content during wood burning. The sensors measure a thermoelectric voltage between a catalytically activated and an inert area. Since the reaction of hydrocarbons takes place exothermically at the catalytic layer, a temperature gradient is generated between the catalyst coating and the inert layer. At an operating temperature of 600°C, a linear relationship can be established between the hydrocarbon concentration and the sensor signal. With methane, however, a huge scattering occurs. Measurements have shown that the oxidation of the highly stable methane molecules at this catalyst just starts at 600°C. The sensitivity of the sensor to methane increases with increasing operating temperature up to 800°C. This information can also be used for a new operating strategy of the sensor.

Keywords: thermoelectric gas sensor, exhaust gas sensor, methane sensitivity, hydrocarbon sensor, thermopile

Introduction

This paper describes the methane-sensitivity of thermoelectric gas sensors. The thermovoltage between a catalytically activated and an inert area of a special designed thermopile structure is measured. Such sensors are ideal candidates to be used in harsh environments like automotive exhaust or the flue gas of firewood combustion processes [1]. As methane (CH_4) is a climate-damaging pollutant, the need for reliable gas sensing is obvious.

State of the art

Recently, thermoelectric gas sensors were successfully applied in the flue gas of firewood combustion. Their signals showed impressive correlation with data from FTIR-analysis during the whole burning process [2]. Such sensors might be applied in future for in-situ determination of the CO/HC-content during wood burning. So, a control algorithm to regulate air streams automatically will help to decrease emissions significantly.

Description of the sensor setup

Sensor elements are built up in thick-film technology on alumina substrates (Rubalit 708, CeramTec, Germany). The sensing structure on the front side consists of 15 serial connected thermopiles (made of Pt and PtRh, both from DuPont), meandering between two areas: One area is catalytically coated (we used a porous alumina thick-film with 1 wt.% Pt-loading), the second area is covered inert (QM42, DuPont). A measurable voltage occurs when reducing gases exothermically react at the catalyst and generate a temperature gradient between the catalyst coating and the inert area. To activate the catalyst, the absolute sensor temperature must be controlled to a certain value, which is high enough to exceed the light-off-temperature of the catalyst for most gases. For that purpose, a thick-film heater (made of Pt, LPA-88, Heraeus) was integrated on the reverse side of the substrate (Figure 1). For more details on the sensor setup see Ref. [3].

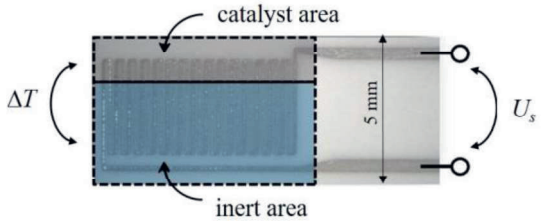


Figure. 1: Top view of the functional part of the sensing element

Experimental

To investigate the sensor behavior with respect to methane sensitivity, we conducted lab measurements in synthetic gas atmospheres. In a first experiment, the absolute sensor temperature was kept constant at 600 °C. Methane was admixed in increasing

concentration steps (up to 7000 ppm, base gas: 10 % O_2 in N_2) to provide the sensor characteristic curve. As second experiment, the temperature dependency of the sensor behavior was investigated (base gas: 10 % O_2, 3 % CO_2 and 3 % H_2O in N_2, total flow: 1000 ml/min). With and without test gas (2100 ppm CH_4), the absolute sensor temperature was varied from 150 °C to 800 °C (heating curve) in steps of 25 °C and analogously backwards (cooling curve).
The sensor response was calculated as the difference of the sensor signal with and without test gas (Equation 1) :

$$\Delta U_S = U_S(base\ gas + CH_4) - U_S(base\ gas)$$

Results and Discussion

As characteristic curve one expects a linear correlation between test gas concentration and signal change with high reproducibility. We found that for several test gases in former investigations [3]. In the here presented data for CH_4 (Figure 2) a huge scattering is visible.

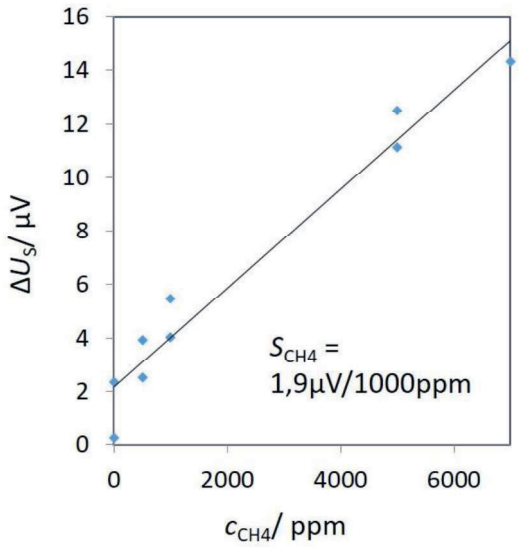


Figure 2: Characteristic curve at 600 °C absolute sensor temperature (signal change for several CH_4 concentrations)

This behavior can be explained by the second experiment (Figure 3). Here, it becomes clear that the sensor starts measuring firstly at 600 °C and then its sensitivity strongly increases up to 800 °C. That means, the oxidation of the highly stable methane molecule needs at least 600 °C to be activated at this catalyst. The higher the temperature, the more of the methane reaching the sensors surface is converted and contributes to the sensor response. Slight changes in the catalyst temperature will strongly influence the response.

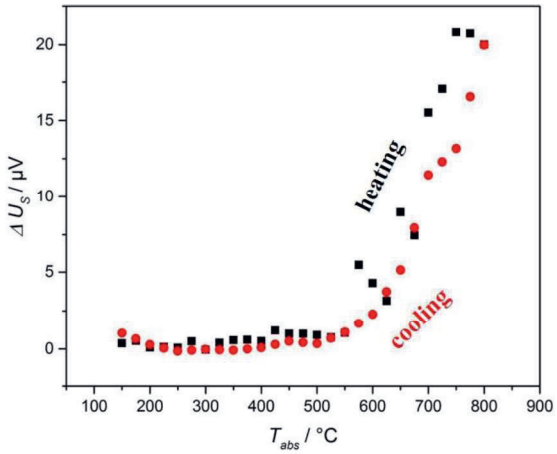


Figure 3: Temperature dependent sensor behaviour (signal change upon 2100 ppm CH_4)

One could take these findings as an advantage to discriminate between CH_4 and other gases. Therefore, the operation strategy of the sensor should include measurements at two different temperatures above and below 600 °C and the difference in the response will be attributed to methane in the test atmosphere.

References

[1] P. Butschbach, F. Hammer, H. Kohler, A. Potreck, Th. Trautmann; Extensive reduction of toxic gas emissions of firewood-fueled low power fireplaces by improved in situ gas sensorics and catalytic treatment of the exhaust gas, Sensors and Actuators B137(1) (2009) 32-41

[2] B. Ojha, G. Hagen, H. Kohler, R. Moos, „Exhaust Gas Analysis of Firewood Combustion Processes: Application of a Robust Thermoelectric Gas Sensor", Proceedings 1 (4), 457, doi: 10.3390/proceedings1040457

[3] S. Wiegärtner, G. Hagen, J. Kita, W. Reitmeier, M. Hien, P. Grass, R. Moos, "Thermoelectric hydrocarbon sensor in thick-film technology for on-board-diagnostics of a diesel oxidation catalyst", Sensors and Actuators B: Chemical 21, 234–240, doi: 10.1016/j.snb.2015.02.083

DOI 10.5162/sensoren2019/P2.1

Messgenauigkeit und Reproduzierbarkeit bei der Messung der Brillouin-Frequenzverschiebung für die faseroptische Bestimmung der Temperaturverteilung

Aleksander Wosniok
Bundesanstalt für Materialforschung und –prüfung (BAM), Unter den Eichen 87, 12205 Berlin

Zusammenfassung

Auf der Grundlage der in DIN EN 61757-2-2:2017 beschriebenen Standardisierungsmethoden wurden in dieser Arbeit wichtige Leistungsparameter von vier kommerziellen Messsystemen zur ortsaufgelösten faseroptischen Erfassung der temperatur- und dehnungsabhängigen Brillouin-Frequenzverschiebung untersucht. Gemäß der oben genannten Norm konzentrierte sich hier die Ermittlungsprüfung auf messtechnische Aspekte reiner Temperaturmessungen. Dabei diente die Ermittlung von einem festgelegten Satz dreier Betriebsverhaltenskenngrößen in Form von quantitativen Angaben des Messfehlers, der Wiederholgenauigkeit sowie der räumlichen Unsicherheit einem Vergleich der getesteten Messsysteme.
Alle faseroptischen Messungen wurden mit einer Faserschleife ausgeführt, d. h. als Sensorkonfiguration der Messfaser wurde eine Schleifenanordnung gewählt. Die Messverfahren der getesteten Messysteme basierten somit auf der Brillouin-optischen Zeit- bzw. Frequenzbereichsanalyse. Die Betriebsverhaltenskenngrößen wurden ferner bei unterschiedlichen Werten zweier messtechnischer Parameter, des Frequenzschrittes und der Mittelungsrate, für Standardeinstellungen der Laserleistungswerte der in die Sensorfaser einzukoppelnden Laserlichtsignale ermittelt. Auf diesem Weg lassen sich relevante Aussagen zur Optimierung der Messgenauigkeiten in Bezug auf Messdauer treffen.

Keywords: Faseroptischer Sensor, ortsaufgelöste Brillouin-Sensorik, verteilte faseroptische Temperaturmessung, Brillouin-Frequenzverschiebung, Betriebsverhaltenskenngrößen.

Einleitung

Vor dem Hintergrund der weltweit steigenden wirtschaftspolitischen Anforderungen an Zuverlässigkeit und Betriebssicherheit von kritischen Anlagen, wie Kraftwerke, Rohleitungen, Deiche sowie Straßen- und Bahnnetz, bieten örtlich auflösende faseroptische Temperaturmessverfahren ein ausgezeichnetes Tool für lückenlose Überwachung über kilometerlange Strecken. Aufgrund der Materialeigenschaften der faseroptischen Sensoren, ihrer geringen Abmessungen, Unempfindlichkeit gegenüber elektromagnetischen Störungen und Einsetzbarkeit unter hohem Druck sowie in einer explosionsgefährderten Umgebung, kann der Einsatzbereich dieser Sensoren auf weitere industrielle Bereiche erweitert werden. Dazu zählt beispielhaft die Bohrlochüberwachung sowie Monitoring von Starkstromnetzen.
Die in den letzten Jahren vorangetriebenen Arbeiten zur Festlegung von Rahmenbedingungen zwecks Gewährleistung der Vergleichbarkeit zwischen unterschiedlichen faseroptischen Temperaturmessystemen mündeten in einer Vereinheitlichung der Standards im Rahmen der im Jahr 2016 veröffentlichten Norm IEC 61757-2-2 und ihrer deutschen Fassung DIN EN 61757-2-2:2017 [1]. Der Gültigkeitsbereich dieser Norm erstreckt sich über alle ortsaufgelösten faseroptischen Temperatursysteme (engl.: distributed temperature sensing - DTS) auf Basis von temperaturabhängigen Lichtstreueffekten in Lichtwellenleitern (LWL), die als ortsauflösende Sensoren benutzt werden. DTS basiert somit auf der messtechnischen Nutzung sowohl der elastischen Rayleigh-Streuung als auch der unelastischen Lichtstreuarten in Form der Raman- und der Brillouin-Streuung.
Die vorliegende Arbeit liefert einen quantitativen Vergleich zwischen vier kommerziellen faseroptischen Brillouin-basierten Messsystemen unterschiedlicher Hersteller. Die dabei durchgeführten Temperaturtests dienten der Ermittlung der drei relevanten in der Norm festgelegten

Betriebsverhaltenskenngrößen. Dazu zählt der Messfehler sowie die Wiederholbarkeit und die räumliche Unsicherheit der Temperaturmessung. Diese Leistungsparameter sind ferner in einer Schleifenanordnung der faseroptischen Sensoren untersucht worden. Die möglichen Sensorkonfigurationen der Brillouin-Messtechnik sind im nächsten Abschnitt dargelegt.

Grundlagen der Brillouin-Messtechnik

Die Brillouin-basierten Messverfahren gehören zu den weit verbreiteten faseroptischen Methoden für die ortsaufgelöste Dehnungs- und Temperaturmessung über kilometerlange Messstrecken. Die im Handel kommerziell erhältlichen Geräte verwenden die bekannten Ansätze zur Bestimmung der sogenannten Brillouin-Frequenzverschiebung ϑ_B sowohl im Zeit- als auch im Frequenzbereich. Die messtechnischen Eigenschaften der beiden Messprinzipien sind in [2] und [3] detailliert vorgestellt. Im Allgemeinen können die Brillouin-Systeme entweder als Reflektometer betrieben oder als eine faseroptische Brillouin-Analyse realisiert werden. Die Messungen auf Basis der Brillouin-Reflektometrie werden ferner mit einer einseitig offenen Sensorfaser ausgeführt. Dies wiederum bedeutet, dass die Laserlichtsignale (Pump-Signale) zur Erzeugung der Brillouin-Streuung an einem Faserende eingekoppelt werden. Durch eine zusätzliche Stimulierung der Brillouin-Streuung, bei einer Einkopplung von weiteren frequenzverschobenen Laserlichtsignalen (Probe-Signalen) an dem zweiten Faserende in die entgegengesetzte Richtung, bietet die faseroptische Brillouin-Analyse dagegen höhere Messgenauigkeiten. Die hier erforderliche Schleifenanordnung der Sensorfaser ist jedoch bedingt durch die Lichteinkopplung an den beiden Faserenden nicht für jede Anwendung realisierbar.

Bei den für die hier vorgestellten Untersuchungen relevanten Analyse-Verfahren, BOTDA (engl.: Brillouin Optical Time Domain Analysis) und BOFDA (engl.: Brillouin Optical Frequency Domain Analysis), wird die Brillouin-Frequenzverschiebung ϑ_B aus den aufgenommenen lorentzförmigen Brillouin-Verstärkungsspektren ermittelt. ϑ_B entspricht ferner dem Frequenzwert beim jeweiligen Maximum der Brillouin-Spektren, d. h. der Frequenzverschiebung, die an jedem Ort entlang der Messfaser zu einer maximalen Brillouin-Streuung führt. Die Erfassung der ortsverteilten Brillouin-Spektren entlang der Sensorfaser erfolgt dabei durch Abscannen des Frequenzbereiches um den Wert ϑ_B. Das Abscannen selbst wird durch eine schrittweise Änderung der Frequenzverschiebung zwischen den Pump- und Probe-Signalen mit einem konstanten Frequenzschritt realisiert. Abb. 1 zeigt beispielhaft einen Verlauf der Brillouin-Verstärkungsspektren entlang der hier untersuchten 4,5 km langen Messstrecke. Die Verschiebung der Brillouin-Spektren in Bereichen S1 und S2 (siehe Abb. 2) zu höheren Frequenzwerten ist temperaturbedingt, dies verursacht durch Temperaturanstieg bis auf 60 °C. Im Gegenteil zu BOFDA, erfordert der BOTDA-Ansatz eine zusätzliche Mittelung vieler Einzelmessungen zwecks Verringerung des Signal-Rausch-Verhältnisses (engl.: signal-to-noise ratio - SNR).

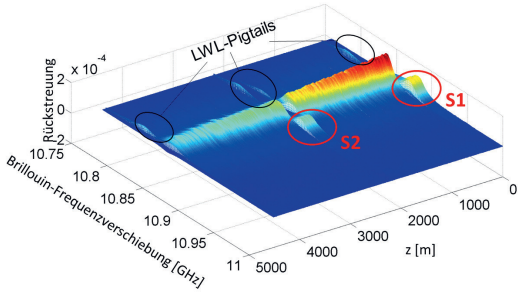


Abb. 1: Messung der Brillouin-Verstärkungsspektren entlang der getesteten Sensorstrecke im BOFDA-Verfahren.

Unabhängig von der Sensorkonfiguration und dem gewählten Messverfahren lässt sich die Temperatur- bzw. Dehnungsverteilung mittels folgender Gleichung bestimmen:

$$\vartheta_B = \vartheta_{B0} + C_T \cdot (T - T_0) + C_\varepsilon \cdot \varepsilon \qquad (1)$$

wobei C_T und C_ε die Fasermaterial- und wellenlängenspezifischen Temperatur- bzw. Dehnungskoeffizienten darstellen. Der Bezugswert ϑ_{B0} in Gl. (1) ist als eine stark von der Faserkerndotierung abhängige charakteristische Brillouin-Frequenzverschiebung einer ungedehnten Faser bei einer Referenztemperatur T_0 zu verstehen. Bleibt die Sensorfaser mechanisch entlastet (ungedehnt), so lassen sich mit einem bekannten Wert von ϑ_{B0} aus jedem gemessenen Wert ϑ_B direkte Rückschlüsse über die Temperaturverteilung entlang dieser Messfaser ziehen. Der in Abb. 1 präsentierte Frequenzversatz zwischen den gemessenen ϑ_B-Werten der Sensorfaser und den an die Sensorfaser angespleißten Verbindungsfasern (LWL-Pigtails) weist bedingt durch eine konstante Umgebungstemperatur sowie die vorhandene mechanische Entkopplung auf ihre unterschiedlichen ϑ_{B0}-Werte hin. Dies wiederum belegt somit geringe Unterschiede in der Faserkerndotierung der Sensorfaser und der LWL-Pigtails.

Versuchsanordnung

Die untersuchten Betriebsverhaltenskenngrößen wurden im Prüfaufbau gemäß Abb. 2 bei drei konstanten Temperaturwerten (-10°C, 20 °C und 60°C) separat für die sich in einer kalibrierten Temperaturkammer VTL 4003 befindenden Faserabschnitte S1 und S2 (202 m bzw. 204 m lang) bestimmt. Als Sensorfaser wurde eine Singlemode-Standardfaser SMF28e eingesetzt.

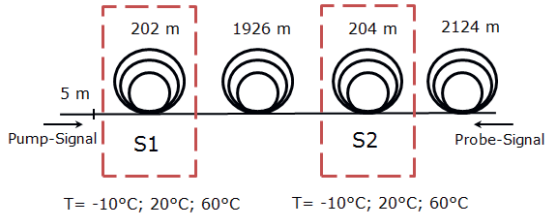


Abb. 2: Sensorkonfiguration der Testfaser - Schleifenanordnung.

Der Zustand der Sensorfaser lässt sich durch die optischen Verluste entlang der gesamten Länge der Faser quantifizieren. Abb. 3 zeigt den Verlauf der optischen Dämpfung entlang der Teststrecke gem. Sensorkonfiguration in Abb. 2. Der Dämpfungsverlauf in Abb. 3 stellt realitätsnahe Dämpfungswerte dar, die bei Sensorapplikation bzw. -integration für die Aufgaben des Zustandsmonitorings zur Überwachung von ausgedehnten Objekten und Anlagen realisierbar sind.

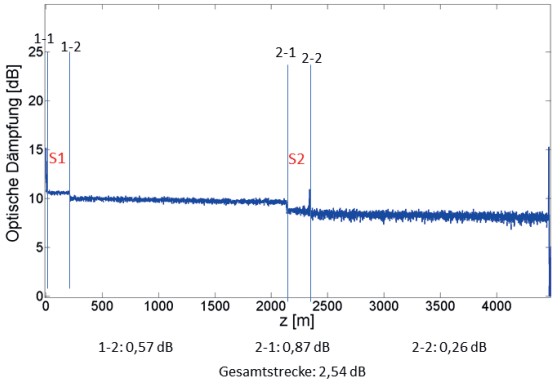


Abb. 3: Optische Dämpfung entlang der Testfaser.

Drei von den vier getesteten Geräten nutzten das BOTDA-Verfahren. Bei einem weiteren Gerät handelte es sich um ein BOFDA-Messsystem.

Die Pump- und Probeleistungen der an den beiden Enden der Sensorfaserstrecke einzukoppelnden Laserlichtsignale wurden stets nach den Standardvorgaben des jeweiligen Herstellers eingestellt.

Das Abscannen des Frequenzbereiches zwecks Ermittlung der temperaturabhängigen Brillouin-Frequenzverschiebung ϑ_B erfolgte in vorgebbaren Frequenzschritten von 1 MHz, 2 MHz bzw. 4 MHz.

Die gewählte Mittelungsrate im BOTDA-Verfahren variierte zwischen 2^{10}, 2^{13} und 2^{15}. Die konstante Ortsauflösung lag bei 1 m und der Messpunktabstand im Bereich um 20 cm.

Prüfverfahren zur Leistungsbewertung

Die Messungen zur Ermittlung der Betriebsverhaltenskenngrößen aller vier Messsysteme sind in einem klimatisierten Labor bei einer stabilisierten Temperatur (23,5 ± 0,5) °C durchgeführt worden. Die Gewährleistung einer möglichst konstanten Betriebstemperatur ist eine wichtige Voraussetzung für die Vergleichbarkeit der Messergebnisse. Darüber hinaus sind vor der Ausführung der Tests die Aufwärmzeiten der jeweiligen Geräte einzuhalten.

Die Leistungsbewertung erfolgte zwecks erweiterter Vergleichbarkeit bei unterschiedlich voreingestellten Messparametern [Anzahl der Mittelungen (gilt nur für BOTDA), Frequenzschritt]. Die Messungen wurden stets zwanzig Mal für jeden Messparametersatz und einen konstanten Temperaturwert der zu untersuchenden Bereiche S1 sowie S2 (siehe Abb. 2) wiederholt. Die auf diese Weise aufgenommenen Temperaturspuren dienten somit zur Berechnung aller drei Betriebsverhaltenskenngrößen. Die detaillierte Vorgehensweise zur Kennwertberechnung des Temperaturmessfehlers, der Temperaturwiederholgenauigkeit sowie der räumlichen Temperaturunsicherheit ist in [1] nachzuschlagen.

Im Allgemeinen wird die Messgenauigkeit der Messsysteme mit dem Temperaturmessfehler wiedergegeben, während die beiden anderen Kenngrößen die Reproduzierbarkeit der ortsverteilten Messung beschreiben. Der Messfehler wird somit als Höchstwert der Differenz zwischen einem zentrierten und gleitenden Mittelwert der gemessenen Temperatur und einer Referenztemperatur (dem in S1 sowie S2 konstanten wahren Wert der Temperatur) in den beiden temperierten Bereichen S1 und S2 gedeutet. Die Glättung der ermittelten Temperaturmittelwerte erfolgt dabei zentriert über 51 Messpunkte.

Die Temperaturwiederholbarkeit soll als ein geglätteter Verlauf der erweiterten Messunsicherheit der Temperaturmessdaten an jedem Messort gegeben durch eine entsprechende zweifache Standardabweichung des Temperaturmittelwertes verstanden werden. Wird für jede einzelne Temperaturabtastspur eine zweifache Standardabweichung über 51 örtlich

aufeinander folgende Temperaturwerte gebildet, so entspricht die räumliche Temperaturunsicherheit den jeweiligen aus zwanzig Messungen ermittelten Mittelwerten dieser zweifachen Standardabweichung. Die räumliche Temperaturunsicherheit gibt somit die örtliche Reproduzierbarkeit der Temperaturwerte an.

Ergebnisse

Die Norm [1] schreibt vor, dass die zu untersuchenden Betriebsverhaltenskenngrößen für drei unterschiedliche Messzeiten zu bestimmen sind. Ferner handelt es sich um die kürzeste durch das DTS-Messsystem zur Verfügung gestellte Zeit, eine empfohlene Zeit und die längstmögliche Zeit. Aufgrund der Differenzen in den Messzeiten aller vier getesteten Messgeräte werden die geforderten unterschiedlichen Messzeiten einheitlich zwecks bestmöglicher Vergleichbarkeit über den Parametersatz [Anzahl der Mittelungen (gilt nur für BOTDA), Frequenzschritt] wie folgt definiert:

- [2^{13}, 4 MHz] – kürzeste Zeit
- [2^{13}, 2 MHz] – empfohlene Zeit
- [2^{13}, 1 MHz] – längste Zeit

Tab. 1 zeigt alle ermittelten Messfehlerwerte in Abhängigkeit von der Temperatur für die drei oben angegebenen Messzeiten.
Da das System B messtechnisch einen Einsatz einer Nachlauffaser erfordert, konnten für dieses Gerät keine Kenngrößen im Abschnitt S2 bestimmt werden. Die in diesem Fall erforderliche Verlängerung der Messstrecke hätte die in Abb. 2 gezeigte Sensorkonfiguration verändert und somit die Vergleichbarkeit der Messwerte beeinträchtigt. Zwecks Sicherstellung der Vergleichbarkeit wurde hier auf die Verlängerung der Messstrecke verzichtet.
Tab. 1 zeigt, dass sich mit den getesteten Brillouin-basierten Systemen Messgenauigkeiten von 0,3 °C bis 0,4 °C erreichen lassen. Die Messfehlerwerte zeigen jedoch allgemein Temperatur- und Ortsabhängigkeiten, je nachdem für welchen Abschnitt (S1 oder S2) die Werte ermittelt worden sind. Dieses Verhalten ist durch die in [4] thematisierte Dämpfung der Brillouin-Verstärkungsspektren in Abhängigkeit sowohl von der gewählten Pump- und Probeleistung der Laserquellen als auch von der Fasertemperatur selbst zu erklären. Werden die in Abb. 1 gezeigten Spektren bei unterschiedlichen Werten der Pump- und Probeleistung aufgenommen, so verändert sich die Höhe und die Breite der Spektren entlang der Faserstrecke, was den Messfehler und die weiteren Leistungsparameter ortsabhängig beeinflussen kann. Darüber hinaus sind die Höhe und die Breite der Brillouin-Spektren temperaturabhängig. Wie in [4] gezeigt, führen die Unterschiede in der Fasertemperatur auch zur Dämpfungsänderung der erfassten Brillouin-Spektren.

Tab. 1: Ermittelte Messfehlerwerte in °C - Anzahl der Mittelungen: 2^{13} (nur im Zeitbereichsverfahren)

Temperatur [°C]	Frequenzschritt [MHz]		
	1	2	4
S1 – Messsystem A			
-10	0,36	0,46	0,43
20	0,42	0,40	0,42
60	**0,29**	0,31	0,33
S2 – Messsystem A			
-10	0,46	0,42	0,37
20	0,72	0,65	0,71
60	0,43	0,43	0,42
S1 – Messsystem B			
-10	**0,33**	0,34	0,35
20	0,35	0,34	0,34
60	0,42	0,42	0,40
S1 – Messsystem C			
-10	0,56	0,44	0,40
20	0,41	0,47	0,41
60	**0,33**	0,34	0,35
S2 – Messsystem C			
-10	0,50	0,38	0,35
20	0,47	0,45	0,49
60	0,61	0,66	0,81
S1 – Messsystem D			
-10	0,61	0,65	0,62
20	0,56	0,58	0,61
60	0,58	0,57	0,59
S2 – Messsystem D			
-10	0,54	0,54	0,54
20	0,49	0,50	0,49
60	**0,46**	0,47	0,48

Entgegen der allgemeinen Erwartung, dass sich am Anfang der Faser im Abschnitt S1 die höchste Messgenauigkeit erreichen lässt, wurden für das Messystem D gerade im Abschnitt S2 kleinere Messfehler ermittelt. Dies kann ein Indiz für eine nicht optimale Anpassung der Pump- und Probeleistung sein. Eine weitere Optimierung der Laserleistung könnte hier auch in der Verbesserung der Messfehlerwerte und der Reproduzierbarkeit der Messung resultieren.
Im Allgemeinen lässt sich die Messzeit sowohl durch die Herabsetzung der Mittelungsrate als auch durch die Erhöhung des Frequenz-

schrittes reduzieren. Vergleichbar kurze Messzeiten wie bei [2^{13}, 4 MHz] lassen sich somit auch bei [2^{10}, 1 MHz] erreichen. Tab. 2 zeigt die quantifizierten Leistungsunterschiede für die beiden Messparametersätze aller drei BOTDA-Systeme. Tab. 2 belegt auch, dass die durch die Erhöhung der Mittelungsrate erwartete Verbesserung der Messgenauigkeit und Reproduzierbarkeit im Fall der mit der Mittelungsrate von 2^{15} durchgeführten Messungen zweier Messsysteme nicht bestätigt werden kann. Ferner resultiert die deutliche Erhöhung der Mittelungsrate in den signifikant längeren Messzeiten. Dabei wird die Sensorfaser über längere Zeiten leichten Temperaturschwankungen der Temperaturkammer ausgesetzt, was wiederum die ermittelten Messfehlerwerte beeinträchtigt. Die Verkleinerung des Frequenzschrittes im BOFDA-System bis auf 0,5 MHz führte zum gleichen Effekt.

Tab. 2: Normierter Vergleich der Betriebsverhaltenskenngrößen in Abhängigkeit von dem Frequenzschritt und der Anzahl der Mittelungen – BOTDA-Messsysteme

Frequenzschritt [MHz] / Mittelungsrate	Messsystem		
	1.	2.	3.
Messfehler			
1 / 2^{10}	0,99	0,82	0,98
4 / 2^{13}	1	1	1
1 / 2^{15}	---	0,95	0,98
Höchstwert der Temperaturwiederholbarkeit			
1 / 2^{10}	1,01	1,04	1,00
4 / 2^{13}	1	1	1
1 / 2^{15}	---	1,15	0,63
Höchstwert der räumlichen Temperaturunsicherheit			
1 / 2^{10}	1,20	1,04	1,13
4 / 2^{13}	1	1	1
1 / 2^{15}	---	1,16	1,02

Die relevanten Messergebnisse zu der untersuchten Temperaturwiederholbarkeit sind in Abb. 4 – Abb. 6 dargestellt. Die ersten zwei Diagramme belegen die Temperaturabhängigkeit dieses Leistungsparameters. Abb. 6 zeigt dagegen eine messbare Verschlechterung der Temperaturwiederholbarkeit bei Erhöhung des Frequenzschrittes von 2 MHz auf 4 MHz.

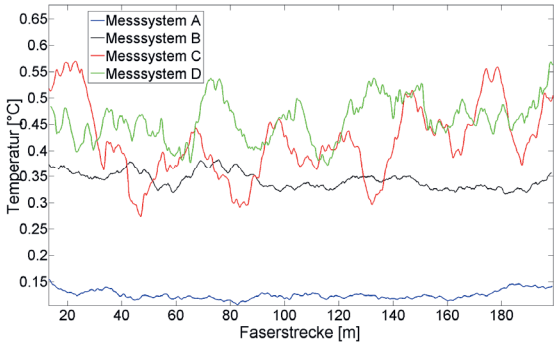


Abb. 4: Temperaturwiederholbarkeit für S1; T = -10 °C, Frequenzschritt: 2 MHz, Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

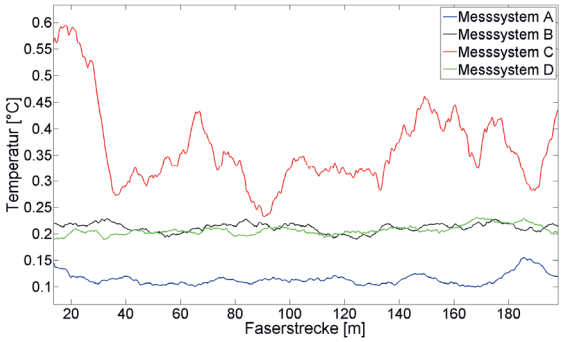


Abb. 5: Temperaturwiederholbarkeit für S1; T = 60 °C, Frequenzschritt: 2 MHz, Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

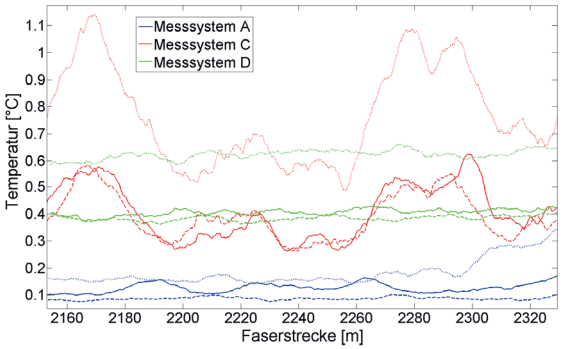


Abb. 6: Temperaturwiederholbarkeit für S2; T = 60 °C, Frequenzschritt: 1 MHz (gestrichelte Linie) / 2 MHz (durchgezogene Linie) / 4 MHz (gepunktete Linie), Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

Die in Tab. 3 temperaturübergreifend zusammengefassten genausten Höchstwerte der Temperaturwiederholbarkeit zeigen, dass schon mit einem Frequenzschritt von 2 MHz optimierte Leistungsparameterwerte realisierbar sind.

Tab. 3: Höchstwerte der Temperatur-wiederholbarkeit in °C - Anzahl der Mittelungen: 2^{13} (nur im Zeitbereichs-verfahren)

Abschnitt	Frequenzschritt [MHz]		
	1	2	4
Messsystem A			
S1	0,12	0,16	0,22
S2	0,10	0,27	0,29
Messsystem B			
S1	0,28	0,24	0,39
Messsystem C			
S1	0,56	0,57	0,64
S2	0,58	0,62	0,60
Messsystem D			
S1	0,27	0,23	0,34
S2	0,41	0,39	0,58

Wie in Abb. 7 - Abb. 9 gezeigt, sind die erwähnten Temperatureffekte auch in den Verläufen der räumlichen Temperatur-unsicherheit zu erkennen.

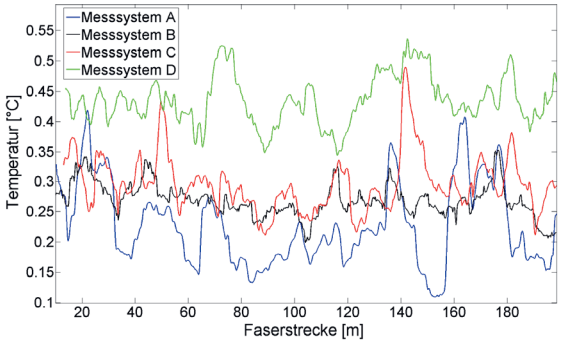


Abb. 7: Räumliche Temperaturunsicherheit für S1; T = -10 °C, Frequenzschritt: 2 MHz, Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

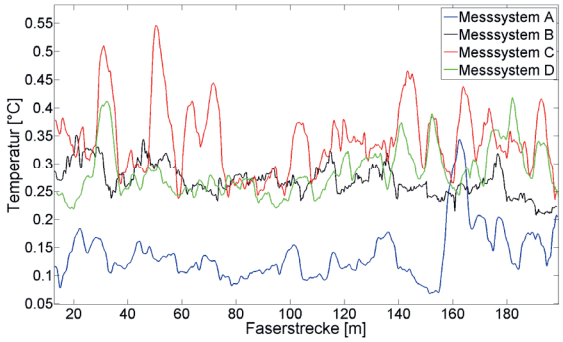


Abb. 8: Räumliche Temperaturunsicherheit für S1; T = 20 °C, Frequenzschritt: 2 MHz, Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

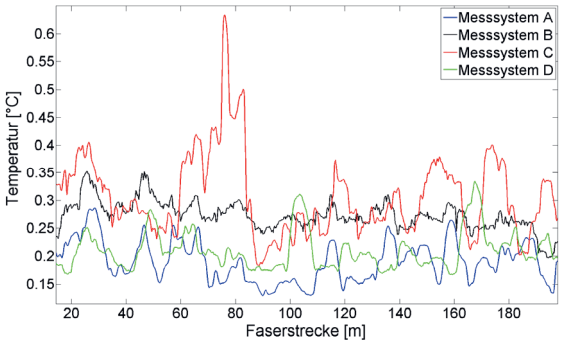


Abb. 9: Räumliche Temperaturunsicherheit für S1; T = 60 °C, Frequenzschritt: 2 MHz, Anzahl der Mittellungen: 2^{13} (nur im Zeitbereichsverfahren)

Anhand der in Abb. 10 präsentierten Verläufe der Temperaturunsicherheit und der in Tab. 4 angegebenen Höchstwerte lässt sich generell feststellen, dass die Erhöhung des Frequenz-schrittes auf 4 MHz auch zur Verschlechterung dieser Kenngröße führen kann.

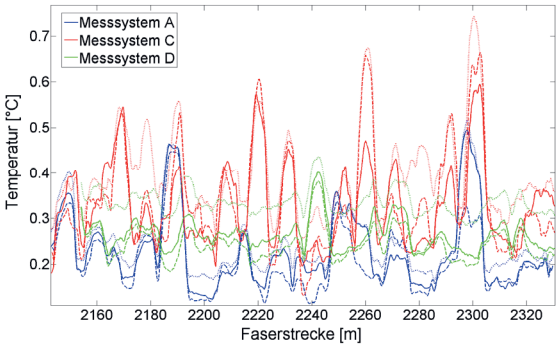


Abb. 10: Räumliche Temperaturunsicherheit - für S2; T = 60 °C, Frequenzschritt: 1 MHz (gestrichelte Linie) / 2 MHz (durchgezogene Linie) / 4 MHz (gepunktete Linie), Anzahl der Mittellungen: 2^{13} (nur im BOTDA)

Tab. 4: Höchstwerte der räumlichen Temperaturunsicherheit in °C

Abschnitt	Frequenzschritt [MHz]		
	1	2	4
Messsystem A			
S1	0,29	0,29	0,37
S2	0,45	0,49	0,52
Messsystem B			
S1	0,33	0,35	0,35
Messsystem C			
S1	0,51	0,49	0,44
S2	0,46	0,41	0,47
Messsystem D			
S1	0,39	0,33	0,40
S2	0,37	0,38	0,41

Schlussfolgerungen

Die ermittelten Betriebsverhaltenskenngrößen weisen Temperatur- und ortsabhängige Differenzen auf. Der dabei erreichte geringste Messfehler liegt bei 0,29 °C für das Messsystem A, 0,33 °C für die Messsysteme B und C, sowie 0,46 °C für das Messsystem D.
Die Messgenauigkeit und die Reproduzierbarkeit der auf der Bestimmung der Brillouin-Frequenzverschiebung basierenden Temperaturmessung können durch die Wahl der Leistungswerte der Pump-Probe-Lasersignale beeinflusst werden. In Bezug auf die Optimierung der Messunsicherheiten führen die von den Herstellern voreingestellten Standardwerte somit nicht zwangsläufig zum erwarteten bzw. vorgegebenen Niveau der untersuchten Kenngrößen. Vielmehr sollen die Pump- und Probe-Leistungen vor der eigentlichen Messung anwendungsbezogen an die Länge der Messstrecke, die optischen Verluste und die gewählte Ortsauflösung angepasst werden.
In vielen Anwendungsfällen wird von der Messtechnik neben präzisen Messergebnissen die Einhaltung kurzer Messzeiten gefordert. Auf Basis der durchgeführten Prüfuntersuchungen liefern alle drei im Zeitbereichsanalyseverfahren getesteten Messsysteme bei der Mittelungsrate von 2^{13} und dem Frequenzschritt von 2 MHz relativ kurze Messzeiten, die zu keinem relevanten Genauigkeitsverlust führen. Eine vergleichbare Effizienz wurde im Frequenzbereichsanalyseverfahren beim Frequenzschritt im Bereich von 1 bis 2 MHz erreicht. Eine genauere Frequenzabtastung mit dem Schritt von 0,5 MHz erwies sich als nicht zielführend.

Literaturnachweis

[1] *DIN EN 61757-2-2:2017*; doi: 10.31030/2833739
[2] M. A. Soto, Distributed Brillouin Sensing: Time-Domain Techniques, *Handbook of Optical Fibers*, ,1-91 (2017); doi: 10.1007/978-981-10-1477-2_7-1
[3] A. Wosniok, Distributed Brillouin Sensing: Frequency-Domain Techniques, *Handbook of Optical Fibers*, ,1-25 (2017); doi: 10.1007/978-981-10-1477-2_8-1
[4] A. Wosniok, Untersuchungen zur Unterscheidung der Einflussgrößen Temperatur und Dehnung bei Anwendung der verteilten Brillouin-Sensorik in der Bauwerksüberwachung, *Dissertation TU Berlin*, (2017); doi: 10.14279/depositonce-3828

DOI 10.5162/sensoren2019/P2.2

Rotationsgeschwindigkeitsmessung mit Hilfe von “Self-Mixing”-Interferometrie

Yang Liu[1,2], Dr. -Ing. habil. Jigou Liu[1], Qiankun Song[1,3], Prof. Dr. -Ing. Dr. h.c. Ralph Kennel [2]
[1] ChenYang Technologies GmbH & Co. KG., Markt Schwabener Str. 8, 85464 Finsing, Deutschland
[2] Technische Universität München, Arcisstraße 21, 80333 München, Deutschland
[3]Hochschule München, Lothstraße 64, 80335 München, Deutschland
E-Mail: yang.liu@chenyang.de

Zusammenfassung

In diesem Beitrag wird die Optimierung eines Messsystems beschrieben, mit der über die Methodik der „Self-Mixing“-Interferometrie ein Signal (SMI-Signal), das für die präzise Rotationsgeschwindigkeitsmessung eines Servomotors in Echtzeit genutzt werden kann, generiert wird. Das Messsystem zeigt initial ein hohes Rauschverhalten, was das SMI-Signal in erheblichem Maße stört und die Rotationsgeschwindigkeitsmessung signifikant erschwert. Als Ursache ist der Servomotor selbst identifiziert worden, dessen Einfluss durch Abschirmungen und durch die Eliminierung von Systemschwachstellen minimiert worden ist. Dadurch konnte das Rauschverhalten des Messsystems stark verbessert werden. Nach der Beseitigung des Störsignals wird eine neuentwickelte Signalverarbeitungsmethode zur Ermittlung der Rotationsgeschwindigkeit über das SMI-Signal vorgestellt. Diese Methode nutzt zur Rauschunterdrückung die Autokorrelationsfunktion und selektiert mithilfe eines Algorithmus Signalabschnitte, die für eine präzise Messung geeignet sind. Weiterhin wird über eine Nullpunktdetektion die Frequenz des SMI-Signals, die für die Ermittlung der Rotationsgeschwindigkeit essentiell ist, ermittelt. Mithilfe dieser Methode konnten vielversprechende Ergebnisse erzielt werden.

Keywords: Drehzahlmessung, Drehzahlgeber, „Self-Mixing“-Interferometrie, Autokorrelation, Rauschunterdrückung

Einleitung

In Anwendungen der Automobilindustrie erweisen sich Geschwindigkeit und Position als bedeutsame Parameter. Insbesondere bei Servomotoren sind sie wichtige Themen, da hier eine dreifach kaskadierte Regelungsstruktur für Torsions-, Geschwindigkeits- und Positionsregelung verwendet wird [1]. Während Servomotoren mit analogen Regelungen für jede Regelungsdimension jeweils einen separaten Sensor benötigen, ist bei digitalen Regelungen nur ein Sensor vonnöten. Für Geschwindigkeitsmessungen wird daher in vielen Fällen ein optischer Encoder verwendet [2]. Dabei werden eine Lichtquelle und ein Fotodetektor auf verschiedene Seiten einer Scheibe mit transparenten und undurchsichtigen Bereichen positioniert. Das Licht der Lichtquelle passiert lediglich die transparenten Bereiche. Dadurch wird eine Pulssequenz generiert, mit der die Geschwindigkeit sowie die Position der Scheibe ermittelt werden kann. Da bei einem optischen Encoder 10000 bis 50000 Linien pro Umdrehung möglich sind, erweist sich ihre Auflösung als hoch. Für hohe Auflösungen müssen jedoch hohe Kosten in Kauf genommen werden. Darüber hinaus kommt beispielsweise in elektrischen Antrieben in der Elektromobilität der Geschwindigkeit im Gegensatz zu Industrieantrieben ein höheres Gewicht als der Position zu. In solchen Fällen wird eine direkte Geschwindigkeitsmessung angestrebt, die nicht über den Umweg der Position erfolgt. Deshalb konzentriert sich die Forschung auf die Laserinterferometrie.

Bei der Laserinterferometrie handelt es sich um eine etablierte Technologie, die in vielen Bereichen wie die mechanische Messtechnik, Profilometrie, Vibrometrie und Geschwindigkeitsanalyse Anwendung findet. Im Gegensatz zu optischen Encoder, deren Funktionsweise auf die Zählung der Striche am Zielobjekt beruht, erhalten Interferometrie-Sensoren ihre Messergebnisse durch die Interferenz ihrer intrinsischen Wellenlängen. Dadurch erhält man inhärent hohe Auflösungen und Genauigkeiten [3].

Die „Self-mixing“-Interferometrie (SMI) ist eine Subkategorie der Laserinterferometrie, die sich wachsender Beliebtheit in der Forschung erfreut [4] [5]. Dabei handelt es sich um eine Messtechnik, in welcher der Laserstrahl einer

Laserdiode von einem Objekt zurück zum Laser reflektiert wird. Gemäß dem schematischen Aufbau in Abbildung 1 kann diese Messtechnik für die Rotationsgeschwindigkeitsmessung genutzt werden.

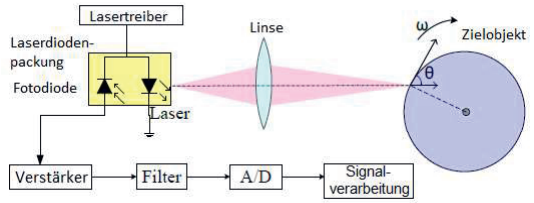


Abb.1: Schematische Aufbau einer SMI-Geschwindigkeitsmessung (in Anlehnung an [3])

Es kommt zu einer Interferenz zwischen dem reflektierten Licht und dem Lichtstrahl des Lasers, wodurch die Amplitude und Frequenz des Lasers infolge des Doppler-Effekts verändert wird. Die Frequenzverschiebung (f_D) moduliert die Ausgangsleistung des Lasers und ist gemäß Gl. (1) direkt proportional zur Rotationsgeschwindigkeit eines Objekts.

$$f_D = \frac{2v\cos\theta}{\lambda} \qquad (1)$$

Dabei repräsentiert λ die Wellenlänge der Laserdiode, θ den Winkel zwischen der Einfallsrichtung des Laserstrahls und der Bewegungsrichtung des Zielobjekts und *v* die lineare Geschwindigkeit des Zielobjekts. Indem die Ausgangsleistung eines Lasers mit einer in der Laserdiode integrierte Fotodiode beobachtet wird [6], kann ein Signal (SMI-Signal) extrahiert werden, mit dem die Rotationsgeschwindigkeit des Objekts über analoge und digitale Signalverarbeitungsmethoden ermittelt werden kann [3].

Im Folgenden wird das Messsystem, das für die SMI-Signal Extrahierung aufgebaut worden ist, sowie ein initiales Systemdefizit vorgestellt. Anschließend wird erläutert, wie das Defizit behoben werden konnte. Im zweiten Teil wird eine neuentwickelte Signalverarbeitungsmethode präsentiert, mit dem die Rotationsgeschwindigkeit des Objekts mithilfe des SMI-Signals präzise ermittelt werden soll.

Messsystemaufbau und Optimierung

Abbildung 2 präsentiert den Aufbau des Messsystems.

Über die Stromversorgung und der Treiberschaltung wird der Laserdiode im Sensorkopf ein konstanter Strom zugeführt. Die Laserdiode sendet ein Laser, der von den Linsen im Sensorkopf fokussiert wird und auf die Mantelfläche einer zylindrischen Scheibe trifft, die mit der Achse eines Servomotors verbunden ist. Der Servomotor wird dabei über einen Computer gesteuert und rotiert die Scheibe mit einer festgelegten Rotationsgeschwindigkeit um die eigene Achse. Es wird ein optisches Feedback im Laserstrahl generiert, der sich in der Ausgangsleistung des Lasers bemerkbar macht und von der Fotodiode, die in der Laserdiode integriert ist, erfasst wird. Die Fotodiode leitet die Ausgangsleistung an die Vorverarbeitungsschaltung weiter, in der das SMI-Signal extrahiert und verstärkt wird. Abschließend sendet die Schaltung das SMI-Signal an das Oszilloskop weiter, welches das SMI-Signal visualisiert. Aus dem Oszilloskop können Datensätze zum SMI-Signal entnommen werden, die mithilfe von verschiedenen Softwareprogrammen bearbeitet werden können.

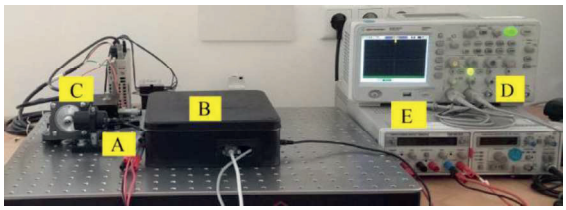


Abbildung 2: initialer Systemaufbau A: Sensorkopf; B: Treiberschaltung + Vorverarbeitungsschaltung; C: Servomotor; D: Oszilloskop; E: Stromversorgung

Das System hat jedoch Schwächen vorzuweisen. Eine davon ist das hohe Rauschverhalten, welchem das System ausgesetzt ist.

Abbildung 3 visualisiert das Rauschverhalten des Systems sowie das dadurch generierte SMI-Signal.

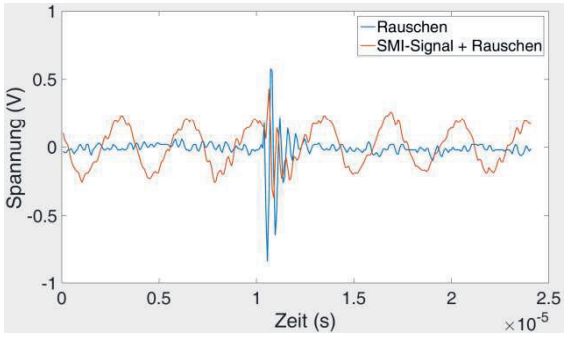


Abbildung 3: initiales Rauschverhalten des Messsystems und generiertes SMI-Signal

Das Rauschen überlagert sich mit dem SMI-Signal, wodurch das Signal in erheblichem Maße verzerrt wird. Auch sind in regelmäßigen

Abständen sehr hohe Spitzen des Rauschverhaltens zu finden. Somit wird das SMI-Signal signifikant gestört, was die Nutzung des Signals für die präzise Geschwindigkeitsmessung erschwert.

Aus diesem Grund sind Maßnahmen erforderlich, mit denen das Rauschverhalten des Messsystems auf ein akzeptables Maß reduziert wird. Es können dabei drei Vorgehensweisen verfolgt werden:

1. Eliminierung der Störquellen
2. Analoge Rauschunterdrückung in der Schaltung
3. Digitale Rauschunterdrückung durch Signalverarbeitungsmethoden

Dieser Teil des Beitrags befasst sich mit der Verfolgung der ersten Vorgehensweise. Auf Basis einer ausführlichen Ursachenanalyse sollen die Störquellen identifiziert und mithilfe geeigneter Maßnahmen eliminiert oder ihre Wirkungen zumindest auf ein Minimum reduziert werden.

Die Ursachenanalyse hat ergeben, dass das Rauschen insbesondere durch den Einfluss des verwendeten Servomotors entsteht. Im Betriebsmodus sendet der Motor elektromagnetische Wellen sowie Schallwellen aus, welche die Vorverarbeitungsschaltung im erheblichen Maße stören.

Um die Einfluss beider Wellen zu beseitigen sind zwei Ansätze verfolgt worden. Als ersten Schritt ist um den Servomotor ein Aluminiumgehäuse angebracht worden. So können beide Wellen um einen gewissen Maß abgedämpft werden. Abbildung 4 stellt die Wirkung dieser Maßnahme dar.

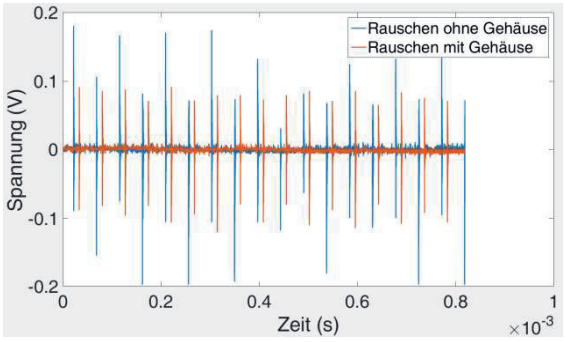


Abbildung 4: Rauschen des Messsystems mit und ohne Aluminiumgehäuse zur Dämpfung hochfrequenten Schalls

Es ist deutlich zu erkennen, dass die Spitzen kleiner geworden sind (siehe Abb. 4), weshalb die Maßnahme als Erfolg quittiert wird. Um die Dämpfung der Schallwellen zu erhöhen, muss entweder die Blechdicke des Gehäuses erhöht oder ein Material mit höherer Masse pro Rohdichte verwendet werden. Wird eine industrielle Anwendung der „Self-mixing"-Interferometrie angestrebt, müssten zur Verringerung der Rauschleistung durch vollständige Dämpfung der Schallwellen möglicherweise hohe Materialkosten und/oder ein hohes Gewicht des Sensors in Kauf genommen werden.

Um diese Probleme zu vermeiden, ist für den Umgang mit beiden Wellenarten ein weiterer Ansatz verfolgt worden. Anstatt die Schallwellen zu dämpfen, ist die Suche nach der Stelle des Messsystems initiiert worden, die am stärksten durch die Schallwellen gestört wird. Als Systemschwachstelle hat sich dabei der Verbindungskabel zwischen Fotodiode und Vorverarbeitungsschaltung herausgestellt. Über diesem Kabel sendet die Fotodiode die Ausgangsleistung des Lasers als Stromfluss an die Vorverarbeitungsschaltung, wo er als ersten Schritt über einen Transimpedanzverstärker in eine Spannung umgewandelt wird. Der Stromfluss erweist sich als sehr störanfällig, weshalb die Strecke, die der Ausgang der Fotodiode als Stromfluss zurücklegt, minimiert werden muss. Aus diesem Grund ist der Verbindungskabel entfernt und die Vorverarbeitungsschaltung ausschließlich durch eine Steckverbindung mit der Laserdiode angeschlossen worden.

Abbildung 5 visualisiert das Resultat dieser Maßnahme.

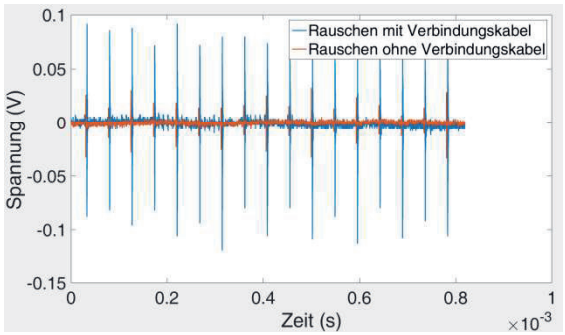


Abbildung 5: Rauschen des Messsystems mit und ohne Verbindungskabel zwischen Fotodiode und Verarbeitungsschaltung

Es ist eine deutliche Reduzierung des Rauschens festzustellen. Die Verbesserung erweist sich als höher als bei der Dämpfung der Schallwellen durch das Aluminiumgehäuse.

Abbildung 6 zeigt das Resultat beider Maßnahmen.

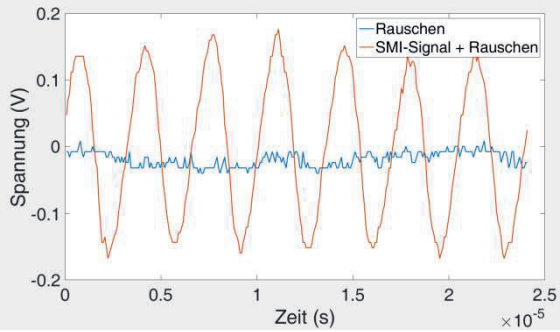


Abbildung 6: Rauschverhalten des Messsystems und generiertes SMI-Signal nach den Maßnahmen

Insgesamt kann eine erhebliche Reduzierung der Rauschleistung festgestellt werden. Das SMI wird durch Rauschen in geringerem Maße gestört und weist eine sauberere Form auf.

Signalverarbeitungsmethode

In den letzten Jahren sind Anstrengungen unternommen worden, Signalverarbeitungsmethoden für eine industrielle Nutzung des SMI-Signals für Rotationsgeschwindigkeitsmessungen zu nutzen. [7] verwendet die analoge Phasenregelschleife-Methode sowie eine angemessene Abtastfrequenz, um im Messbereich von 10-470mm eine Genauigkeit von 1.3% zu erreichen. In [3] wird die Fast-Fourier-Transformation zur Abschätzung der Rotationsgeschwindigkeit mit einer Genauigkeit von 1% genutzt, während sich [8] [9] mit dem Einfluss verschiedener Parameter auf den SMI-Signal und die Genauigkeit von Geschwindigkeitsmessungen befassen. Eine industrielle Anwendung konnten mit den Methoden aus [3] [7] aber noch nicht umgesetzt werden.

In diesem Teil des Beitrags wird deswegen eine alternative Signalverarbeitungsmethode, die für die präzise Geschwindigkeitsmessung entwickelt worden ist, vorgestellt. Mit ihr soll es möglich sein, die Frequenz des SMI-Signals sowohl in Echtzeit als auch mit hoher Genauigkeit zu ermitteln. Wie bereits erläutert worden ist, wird das Verhältnis zwischen Frequenz des SMI-Signals und der Rotationsgeschwindigkeit des Zielobjekts durch Gl. (1) beschrieben. Es ist hierbei eine lineare Abhängigkeit festzustellen. In der Regel sind die anderen Parameter der Gleichung wie beispielsweise Einfallswinkel des Lasers und Laserwellenlänge nach einer festen Positionierung des Sensors Konstanten, wodurch die Rotationsgeschwindigkeit die einzige Variable der Gleichung darstellt. Wenn die Frequenz bekannt ist, kann nach einer Umstellung der Formel die Rotationsgeschwindigkeit mit Leichtigkeit ermittelt werden. Somit stellt die Frequenzabschätzung der wichtigste Teil der Rotationsgeschwindigkeitsmessung dar.

Abbildung 7 visualisiert ein typisches SMI-Signal, das aus dem System generiert wird.

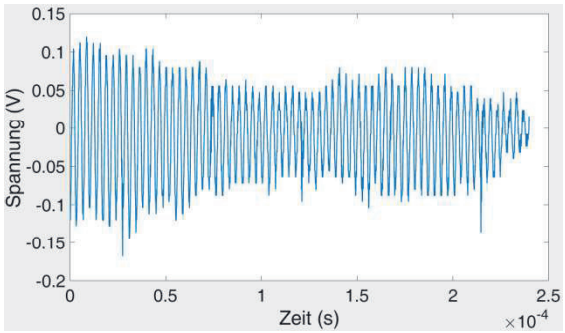


Abbildung 7: typisches SMI-Signal

Es handelt sich um eine sinusförmige Welle mit variierender Amplitude, die durch Rauschen gestört wird. Der Signalverarbeitungsalgorithmus muss in der Lage sein, den Einfluss des Rauschens zu minimieren, ohne die Frequenz des Nutzsignals zu verändern.

Dementsprechend werden im entworfenen Signalverarbeitungsalgorithmus folgende Schritte durchgeführt.

1. Autokorrelation
2. Signalselektion
3. Nullpunktdetektion und Frequenzbestimmung

Im Folgenden werden die einzelnen Schritte näher erläutert.

Autokorrelation

Die Autokorrelation ist ein Begriff der Signalverarbeitung, mit dem die Ähnlichkeit bzw. Korrelation eines Signals mit sich selbst bei unterschiedlichen Zeitverschiebungen τ zwischen den betrachteten Funktionswerten ermittelt wird. Allgemein ist sie wie folgt definiert:

$$\phi_{xx}(\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t) * x(t-\tau) dt \qquad (2)$$

Folgende Eigenschaften macht die Autokorrelationsfunktion für die Rauschunterdrückung interessant [10]:

- Die Autokorrelationsfunktion eines periodischen Signals ist ebenfalls

periodisch und besitzt dieselbe Frequenz.
- Unabhängig von seiner Phasenlage ist für jede harmonische Schwingung die Autokorrelationsfunktion eine Kosinusfunktion.
- die Autokorrelationsfunktion eines stochastischen offsetfreien Rauschsignals stellt einen Deltaimpuls dar, die mit wachsender Verzögerungszeit τ gegen null verläuft.

Mit der Autokorrelationsfunktion ist es somit möglich, das Rauschen eines Signals zu unterdrücken, ohne Informationen über die Signalfrequenz zu verlieren.

Aus diesen Gründen wird im ersten Schritt der Signalverarbeitungsmethode die Autokorrelationsfunktion des SMI-Signals berechnet und fortlaufend aktualisiert. Dabei wird die diskrete Version der Autokorrelation angewandt [11].

$$\phi_{xx}(n) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} x(k) * x(k-n) \qquad (3)$$

Für die Ermittlung der Signalfrequenz f_D muss n mindestens die Anzahl an Punkte für dreiviertel einer Signalperiode betragen. Für bessere Messgenauigkeiten hat sich jedoch während unserer Forschungsarbeiten eine Betrachtung von zwei ganzen Perioden bewährt, weshalb eine Punktezahl für mindestens das 2.25-fache einer Signalperiode empfohlen wird. Diese Anzahl wird vom Algorithmus zuvor ermittelt.

N kann prinzipiell beliebig groß gewählt werden. Eine vollständige Rauschunterdrückung ist bei einer unendlich großen Menge an Abtastpunkte ($N \to \infty$) möglich. Dies ist zwar in der Praxis nicht realisierbar, doch auch mit einer begrenzten Zahl an Abtastpunkte sind Rauschunterdrückungen bis zu einem gewissen Maß umsetzbar. Dabei gilt, dass gegenüber der geometrischen Standardabweichung $\overline{\sigma_{xx}}(n)$ des Produkts der Abtastwerte in einer Periode die Standardabweichung der Autokorrelationsfunktion $\sigma_{\phi xx}(n)$ entsprechend dem Faktor $\sqrt{N}$ verkleinert wird [11].

$$\sigma_{\phi xx}(n) \approx \frac{\sigma_{xx}(n)}{\sqrt{N}} \qquad (4)$$

Abbildung 8a und b visualisieren das Ergebnis der Autokorrelation.

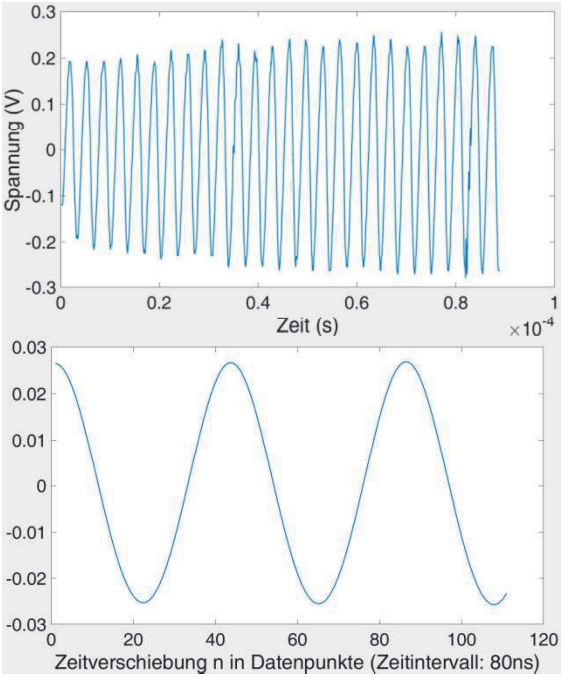


Abbildung 8: a) betrachteter SMI-Signalabschnitt; b) Autokorrelationsfunktion des Signalabschnitts

Aus dem betrachteten Signalabschnitt in Abbildung 8a resultiert die Autokorrelationsfunktion in Abbildung 8b. Für die Parameter τ und N sind die Werte 110 und 1000 gewählt worden. Es entsteht eine saubere Kosinusfunktion, an der keine Störungen zu erkennen sind.

Somit gestaltet sich die Rauschunterdrückung als stärker, je größer die N gewählt wird. Auf diese Weise kann die Messgenauigkeit erhöht werden. Es besteht jedoch die Möglichkeit, dass große Werte für N die Reaktionsfähigkeit der Signalverarbeitungsmethode auf Geschwindigkeitsänderungen reduziert. Diese Vermutung muss in spätere Forschungsvorhaben geprüft werden.

Signalselektion

Wie Abbildung 7 verdeutlicht, erweist sich das SMI-Signal als eine sinusförmige Welle mit variierender Amplitude. Dadurch sind bestimmte Teile des SMI-Signals mehr für die Dopplerfrequenzbestimmung geeignet als andere. Dies liegt im Signal-Rausch Verhältnis (=SNR) begründet. Sie ist wie folgt definiert [10].

$$SNR = 10 \lg \left(\frac{P_{Signal}}{P_{Rauschen}} \right) dB \qquad (5)$$

Es kann davon ausgegangen werden, dass sich bei gleichbleibenden Umgebungsbedingungen und Systemeigenschaften die

durchschnittliche Rauschleistung über die Zeit nicht oder nur in geringfügige Maße verändert. Dementsprechend erhöht sich mit steigender Signalamplitude auch die SNR und umgekehrt. Für eine präzise Messung muss der Einfluss des Rauschens so gering wie möglich gehalten werden, weshalb ein hoher SNR angestrebt wird. Im Signalverarbeitungsalgorithmus wird deshalb eine Signalselektion implementiert, in der Signalabschnitte mit zu geringem SNR von den nachfolgenden Verarbeitungsschritten ausgeschlossen werden. Als Entscheidungskriterium wird dabei der Effektivwert der Autokorrelation eines Signalabschnitts gewählt. Erst wenn sie einen bestimmten Schwellenwert überschritten hat, wird eine Weiterverarbeitung des betrachteten Signalabschnitts initiiert.

Nullpunktdetektion und Frequenzbestimmung

An der Autokorrelationsfunktion der Signalabschnitte, die der Signalselektionsalgorithmus als geeignet für die Weiterverarbeitung befunden hat, wird anschließend eine Nullpunktdetektion durchgeführt. Die Signalverarbeitungsmethode detektiert die Stellen, an denen sich die Vorzeichen der Signalwerte ändern. An diesen Stellen wird über eine lineare Regression mit Datenpunkten um die Stellen herum eine Ausgleichsgerade generiert, mit der die Nullstelle ermittelt wird (siehe Abbildung 9).

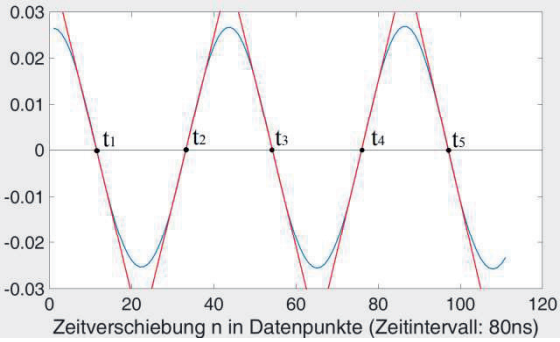


Abbildung 9: Ausgleichsgeraden für die Nullpunktbestimmung

Sind die Positionen von alle erforderlichen Nullstellen in der Autokorrelationsfunktion ermittelt worden, so wird die Frequenz des SMI-Signals über folgende Formel ermittelt.

$$f_D = \frac{1}{0.5*(t_5-t_1)*f_s} \quad (6)$$

Dabei stehen t_5 und t_1 für die Position der Nullstellen, während f_s die Abtastfrequenz repräsentiert. Abschließend kann die Rotationsgeschwindigkeit mithilfe von Gl. 1 ermittelt werden.

In unseren Forschungsanstrengungen sind mehrere Möglichkeiten für die Ermittlung der Dopplerfrequenz mithilfe von maximal fünf Nullstellen überprüft worden. Zum Zeitpunkt dieses Beitrages ist Gl. 6 die Berechnungsmethode, bei denen die Nichtlinearitäten am geringsten gewesen sind. Weitere Forschungsarbeiten für die ideale Bestimmung der Dopplerfrequenz über Nullstellen sind geplant.

Messergebnisse

Die Wirksamkeit des Signalverarbeitungsmethode ist wie folgt geprüft worden:

Im ersten Schritt ist die Linearität zwischen der Dopplerfrequenz, die über die Signalverarbeitungsmethodik ermittelt worden ist, und der Rotationsgeschwindigkeit geprüft worden. Dazu sind unter folgenden Parametern SMI-Signale extrahiert worden.

Tabelle 1: Parametereinstellungen des Messsystems

Parameter	Wert
Rotationsgeschwindigkeit	5, 10, 20, 30, 40, 50, 75, 100, 125, 150, 175, 200 RPM
Objektdurchmesser	40 mm
Laserwellenlänge	785 nm

Anschließend ist die Signalverarbeitungsmethode an Ihnen angewandt worden. Dabei sind für n 400 und für N 3000 ausgewählt worden. Aus 15000 Messwerten pro Rotationsgeschwindigkeit sind für jede Geschwindigkeit ist Mittelwerte erzeugt worden. Tabelle 2 visualisiert das Resultat.

Tabelle 2: Ergebnisse des Messexperiments

Rotationsgeschwindigkeit (in RPM)	Ergebnis (in Hz)
5	19709.384
10	39452.750
20	78844.825
30	118441.050
40	157385.265
50	197586.159
75	295866.171
100	392123.353
125	492757.194
150	591602.127
175	688991.484
200	787732.170

Mithilfe dieser Datenpunkte wird anschließend über die lineare Regression eine Ausgleichsgerade erstellt, um die Nichtlinearität $NL(w)$ der Daten zu ermitteln. Die Nichtlinearität bei einer Geschwindigkeit wird in diesem Beitrag wie folgt definiert:

$$NL(w) = \frac{f_o(w) - f_{ag}(w)}{f_{ag}(w)} * 100\% \quad (7)$$

f_o repräsentiert dabei die Dopplerfrequenz der Signalverarbeitungsmethodik, während f_{ag} die Dopplerfrequenz der Ausgleichsgerade darstellt. Das Resultat ist in Abbildung 10 und Tabelle 3 zu sehen.

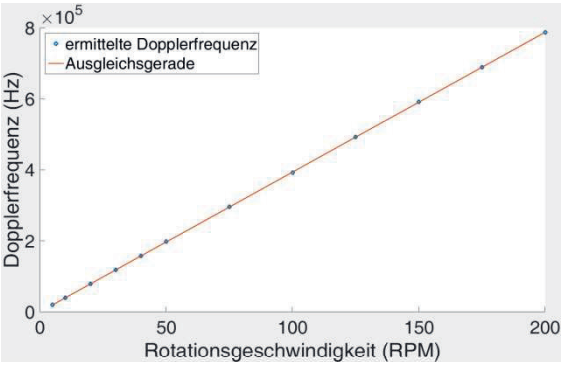


Abbildung 10: Proportionalität zwischen Rotationsgeschwindigkeit

Tabelle 3: Nichtlinearität der ermittelten Dopplerfrequenzen

Rotationsgeschwindigkeit (in RPM)	Dopplerfrequenz der Ausgleichsgerade (in Hz)	Nichtlinearität (in %)
5	19783.872	0.38
10	39474.372	0.05
20	78855.372	0.01
30	118236.372	-0.17
40	157617.372	0.15
50	196998.372	-0.30
75	295450.872	-0.14
100	393903.372	0.45
125	492355.872	-0.08
150	590808.372	-0.13
175	689260.872	0.04
200	787713.372	0.00

Es ist eine maximale Nichtlinearität von 0.45 % festzustellen. Somit besteht grundsätzlich eine Proportionalität zwischen Messergebnisse der Verarbeitungsmethode und Rotationsgeschwindigkeit, was eine Ermittlung der Geschwindigkeit mithilfe dieser Methode möglich macht. Es besteht die Möglichkeit dass die Abweichungen in Abbildung 10 reduziert oder sogar eliminiert werden, falls eine größere Menge an Werte für die Durchschnittsbildung angewandt wird. Dieser Vermutung wird in weitergehenden Forschungsanstrengungen nachgegangen.

Tabelle 4 stellt weitere Informationen bezüglich der Messwerte der Methode zur Verfügung. Es sind die maximale Abweichungen sowie die Standardabweichungen von 15000 Messwerten pro Geschwindigkeit zum entsprechenden Wert der Ausgleichsgerade ermittelt und in Prozentzahlen dokumentiert worden. Auf diese Weise werden erste Aussagen über die Genauigkeit der Methode gemacht.

Tabelle 4: Abweichungen der Messergebnisse vom Mittelwert

Rotationsgeschwindigkeit (in RPM)	Max. Abweichung (in %)	Standardabweichung (in %)
5	1.50	0.60
10	1.11	0.53
20	0.86	0.46
30	1.03	0.57
40	1.48	0.64
50	1.72	0.73
75	1.26	0.62
100	1.31	0.37
125	1.13	0.57
150	0.77	0.33
175	0.98	0.33
200	0.57	0.28

Die Ergebnisse in Tabelle 4 erweisen sich als sehr vielversprechend. Die maximale Abweichung beträgt höchstens 1.72%, während die Standardabweichung mit 0.73% als höchsten Wert stets unter 1% bleibt. Bei 200 RPM beträgt die Standardabweichung zum Mittelwert sogar 0.278 %, was das große Potenzial dieser Methode verdeutlicht, auch industriell genutzt werden zu können.

Zusammenfassung und Ausblick

In Rahmen eines Forschungsprojektes wird ein Messsystem genutzt, der mithilfe „Self-Mixing"-Interferometrie ein Signal (SMI-Signal), das für die präzise Messung von Rotationsgeschwindigkeiten eines Servomotors genutzt werden kann, erzeugt. Dieses Messsystem weist hohes Rauschverhalten auf, das sich mit dem SMI-Signal überlagert und Messergebnisse signifikant beeinträchtigt. Nach einer Ursachenanalyse sind Schallwellen und elektromagnetische Wellen, die der

Servomotor aussendet und auf die Vorverarbeitungsschaltung des Messsystems einwirken, als die primären Störquellen identifiziert worden. Die Wirkung dieser Wellen ist minimiert worden, indem der Servomotor mit einem Aluminiumgehäuse abgeschirmt und beide Wellen abgedämpft worden sind. Außerdem ist mit dem Verbindungskabel zwischen Fotodiode und Vorverarbeitungsschaltung das Systemteil entfernt worden, das am empfindlichsten auf Störfaktoren reagiert. Dadurch konnte die Rauschleistung des Messsystems stark verringert und die Messergebnisse verbessert werden. Es ist darüber hinaus eine neue Signalverarbeitungsmethode zur Ermittlung der Rotationsgeschwindigkeit über das SMI-Signal entstanden. Diese Methode nutzt im ersten Schritt die Autokorrelationsfunktion, da sie in der Lage ist, das Rauschen eines Signals zu unterdrücken, ohne zu einem Verlust von Informationen bezüglich der Signalfrequenz zu führen. Mithilfe eines Selektionsalgorithmus werden anschließend Signalabschnitte, die zu schlechten Messergebnissen führen, von der Weiterverarbeitung ausgeschlossen. An den verbleibenden Abschnitten wird über eine Nullpunktdetektion die Frequenz des SMI-Signals, die für die Ermittlung der Rotationsgeschwindigkeit essentiell ist, ermittelt. Die Nullpunktdetektion erfolgt dabei über eine Ausgleichsgerade, die über eine lineare Regression mit Datenpunkten um die Nullstelle herum gebildet wird. Erste Messergebnisse erweisen sich als vielversprechend und zeigen das Potenzial der Methode, auch in industriellen Anwendungen genutzt werden zu können. Dennoch müssen weitere Untersuchungen bezüglich folgender Aspekte unternommen werden:

- Verbesserung der Signalverarbeitungsmethodik zur Erhöhung der Genauigkeit
- Reduzierung der Nichtlinearitäten der Signalverarbeitungsmethodik
- Analyse der Reaktionsfähigkeit der Methodik auf Geschwindigkeitsänderungen
- Prototypenherstellung

Literaturnachweis

[1] R. M. Kennel. IEEE Transactions on Industry Applications 43, 1572-1577 (2007); doi: 10.1109/TIA.2007.908161

[2] P. Drabarek, R. M. K. Fiet. *4th IET Conference on Power Electronics Machines and Drives*, 149-153 (2008); doi: 10.1049/cp:20080501

[3] H. Sun, J. G. Liu, Q. Zhang, R. Kennel. *Applied Optics* 55, 236-241 (2016). doi: 10.1364/AO.55.000236.

[4] S. Donati. *Laser & Photonics Reviews 6*, 393-417 (2012). doi: 10.1002/lpor.201100002

[5] T. Taimre, M. Nikolic, K. Bertling, Y. L. Lim, T. Bosch, A. D. Rakic. *Advances in Optics and Photonics 7*, 570-631 (2015). doi: 10.1364/AOP.7.000570

[6] G. Giuliani, M. Norgia, S. Donati, and T. Bosch, *Journal of Optics A Pure and Applied Optics 4*, 283 (2002). doi: 10.1088/1464-4258/4/6/371

[7] W. Huang, H. Q. Gui, L. Lu, J. P. Xie, H. Ming, D. Y. He, H. Q. Wang, T. P. Zhao. *Optics Communications 281*, 1662-1667 (2008). doi: 10.1016/j.optcom.2007.11.029

[8] H. Sun, J. G. Liu, R. Kennel. *Adaptive Optics: Analysis, Methods & Systems*, JT3A.4 (2016). doi:10.1364/3D.2016.JT3A.4

[9] H. Sun, J. G. Liu, R. Kennel. 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 236-241 (2017). doi: 10.1109/I2MTC.2017.7969922

[10] J. Detlefsen, U. Siart. *Grundlagen der Hochfrequenztechnik.* 2. Auflage, Oldenbourg, München (2006), ISBN: 3-486-57866-9

[11] J.G. Liu. *Eigenkalibrierende Meßverfahren und deren Anwendungen bei den Messungen elektrischer Größen.* Technische Universität Dresden, Dissertation (1999).

DOI 10.5162/sensoren2019/P2.3

Entwicklung eines akustischen Füllstandsensors für mobile Schüttgutsilos

Dipl.-Ing. Paaranan Sivasothy, B. Sc. Hüseyin Cavuldak, M. Sc. Vitali Rjasanzew, Prof. Dr.-Ing. Jörg Seewig
Lehrstuhl für Messtechnik und Sensorik, Technische Universität Kaiserslautern,
67663 Kaiserslautern, Gottlieb-Daimler-Str. 44, Deutschland

Zusammenfassung

Starre Silobehälter stehen oft unter Druck und sind mit Gefahrstoffen gefüllt. Der Füllstand kann in der Regel nicht per Sichtprüfung bestimmt werden. Daher wird beispielsweise der Füllstand von mobilen Schüttgutsilos, die häufig auf Baustellen eingesetzt werden, von einer Arbeitskraft überprüft, der einen Stein gegen das Silo wirft und den nicht sichtbaren Silohohlraum anhand des Klangs abschätzt. Allerdings ist diese Methode subjektiv und weist erfahrungsgemäß große Fehler auf. Das vorgeschlagene Verfahren „Silopic“ soll das Prinzip der „Steinwurfmethode“ maschinell umsetzen. Eine Sensoreinheit mit einem Schlagwerk gibt einen mechanischen Impuls auf die Siloaußenwand und der entstehende Klang wird mit einem Mikrofon aufgezeichnet. Aus der akustischen Impulsantwort soll ein Kennwert berechnet werden, der zur Abschätzung des Füllstands geeignet ist. Weiterhin wird in dieser Arbeit der Einfluss der eingebrachten Erregerenergie, der Sensorpositionierung auf der Mantelfläche des Silos und Mikrofonabstandes auf das Messverfahren untersucht.

Keywords: Füllstandmesstechnik, Schüttgut, Silo, Wavelet-Transformation, Akustik,

Motivation

Seit vielen Jahrzehnten sind verschiedenste Füllstandsensoren für Gase und Flüssigkeiten etabliert. Fluide können dabei als eine homogene Masse betrachtet werden. Füllstandmesssysteme für Schüttgüter sind jedoch seltener, komplexer und entsprechend teurer. Der Grund hierfür liegt in der komplexen Natur von Schüttgütern, die sich aus einer immensen Menge an heterogenen Festkörperteilchen zusammensetzen.
Gerade beim Einsatz von mobilen Schüttgutsilos, wie sie auf den meisten größeren Baustellen vorzufinden sind, wird daher auf jegliche Art von Füllstandsensorik verzichtet. Stattdessen wird der Füllstand manuell abgeschätzt. Hierzu kommen in der Regel zwei unterschiedliche Verfahren zum Einsatz:

Methode 1: Das Abklopfen mit einem Besenstiel
Eine gängige Form der Füllstandmessung bei Schüttgutsilos ist das Abklopfen mit einem Besenstiel. Eine Arbeitskraft der Baustelle führt hier mehrere Schläge auf die Siloaußenwand aus, wobei er dabei mit dem Schlag örtlich vertikal nach oben wandert. Wandelt sich die dabei resultierende akustische Impulsantwort subjektiv von einem tiefen, dumpfen Klang zu einem hellen, lauten Klang, so gilt die Grenzschicht zwischen Schüttgutmaterial und dem Hohlraum im inneren des Silos als gefunden. Auf Basis dieser Füllhöhe versucht die Arbeitskraft nun den Füllstand in Tonnen abzuschätzen und diese Information zum Zwecke der Nachbestellung an den Bauherrn weiterzugeben.

Methode 2: Das Steinwurf-Verfahren
Eine schnellere Methode, die gerade von erfahreneren Mitarbeitern bevorzugt wird, ist die Steinwurfmethode. Die Arbeitskraft wirft einen Stein auf einen bestimmten Bereich der Siloaußenwand. Auf Grundlage seines erfahrungsbasierten Wissens über den Zusammenhang zwischen Füllstand und akustischer Impulsantwort versucht die Arbeitskraft den tatsächlichen Füllstand abzuschätzen.

Bei der Verwendung dieser beiden manuellen Verfahren zur Füllstandabschätzung stellen sich gewisse Verluste ein. Bei der ersten vorgestellten Methode kann relativ gut die Füllstandhöhe bestimmt werden, sofern der Füllstand im für den Arbeiter zugänglichen

Bereich liegt. Äußerst kritisch ist jedoch die subjektive Abschätzung des Füllgewichts auf Basis der Füllstandhöhe. Vor allem aufgrund des Trichtergeometrie und der Einbauten im Siloinneren ist diese Zuordnung nicht vollständig linear und daher ohne eine detaillierte Berechnung kaum zuverlässig zu erbringen.
Die zweite vorgestellte Methode basiert vollkommen auf der Erfahrung der ausführenden Arbeitskraft. Die Qualität der Messung hängt somit stark von zeitlich veränderlichen Faktoren wie Alter, Gesundheit und Tagesform der einzelnen Arbeitskraft ab. Daher ist diese Methode als hochgradig unzuverlässig zu beurteilen, ungeachtet der zeitlichen Einsparung, die diese Methode grundsätzlich mit sich bringen kann. Zudem entstehen durch die Steinwurfmethode erhebliche Schäden auf der Mantelfläche des Silos, was regelmäßig kostenintensive Wartungsarbeiten mit sich zieht.
Bei manuellen Messungen wird die zuverlässige und regelmäßige Durchführung durch eine Arbeitskraft vorausgesetzt. Im alltäglichen Stress der wechselhaften Baustellenarbeiten ist dies jedoch relativ unrealistisch.
Durch das Leerlaufen eines Silos kann es zum Stillstand von Bauarbeiten kommen. Dieser Stillstand zieht eine Stagnation des Projektfortschritts mit sich, währenddessen die zeitvariablen Kosten wie Mitarbeiterlohn und Miete für Arbeitsmaschinen weiterhin entrichtet werden müssen. Der deutsche Bundesgerichtshof fällte im Oktober 2018 ein Urteil, wonach stillstandbedingte Mehrkosten durch den Auftragnehmer, also dem Bauunternehmen, übernommen werden müssen. Durch den Wunsch diesem Risiko vorzubeugen, steigt innerhalb der Baubranche die Nachfrage nach einer automatisierten kontinuierlichen Füllstandüberwachung erheblich an.

Stand der Füllstandmesstechnik

Die Einsatzbedingungen für die Füllstandmesstechnik in mobilen Schüttgutsilos gelten als recht rau [1]. Innerhalb des metallenen Behälters kommt es bei Befüll- und Entleervorgängen zu einer enormen Staubbildung. Am Sensor anhaftendes Material ist eine häufige Ursache für fehlerhafte Messungen. Darüber hinaus liegt aufgrund der Anzahl an laufenden Maschinen und Arbeiten eine enorme akustische Belastung vor. Weiterhin führt die mechanische Beanspruchung durch bewegende Feststoffpartikel und die damit einhergehenden abrasiven Prozesse zu einer kurzen Lebensdauer von Sensoren im Innenraum des Silos [1]. Die Montage heutiger Sensoren erfordert aufwändige Nachrüstarbeiten (Aufschneiden, Einsetzen, Schweißen, Schleifen, etc.), die mitsamt den Sensorkosten i.d.R. zwischen 5.000 – 10.000 € liegen. Wird der Sensor jedoch von außen an das Silo angebracht, so ist dieser unterschiedlichsten Umgebungseinflüssen ausgesetzt.
Aufgrund der rauen Einsatzbedingungen resultieren folgende Forderungen an ein Füllstandmesssystem für mobile Schüttgutsilos: der Sensor muss

- den Füllstand stufenlos ermitteln,
- robust gegen äußere Einflüsse sein,
- nicht invasiv an das Silo montierbar sein,
- energieeffizient genug für einen Batteriebetrieb über mehrere Jahre sein,
- wartungsfrei sein und
- kostengünstig sein (< 1000 €).

Heute gibt es eine Vielzahl von Messsystemen mit unterschiedlichen Messprinzipien, wobei keins die oben beschriebenen Anforderungen vollständig und zufriedenstellend erfüllt.
Kostengünstig und hinreichend zuverlässig sind vor allem die als Grenzwertschalter verwendeten Sensoren. Als Beispiel sei hier der Drehflügelschalter genannt [2]. Dabei wird eine kleine, rotierende Welle, die in der Mantelfläche gelagert wird, von außen permanent durch einen Elektromotor angetrieben. Am anderen Ende, welches in den Innenraum des Silos ragt, befindet sich ein Drehflügel. Kann die Welle frei rotieren, so ist der Füllstand des Silos niedriger als die Montageposition des Sensors. Wird die Rotation jedoch verhindert, so ist der Füllstand höher. Als großer Nachteil des Drehflügelschalters gelten seine große Abmessung, die Montage und der Verschleiß der bewegten Teile [2].
Ein anderes Beispiel für einen Grenzschalter ist der Vibrationsschalter. Hierbei ragt eine Stimmgabel, die von außen piezoelektrisch angeregt wird, ins Innere des Silos [3] [4]. Die Stimmgabel schwingt frei mit ihrer Eigenfrequenz. Taucht die Stimmgabel jedoch in das Schüttgut ein, so kann eine signifikante Dämpfung der Schwingungsamplitude gemessen werden [3]. Der Vibrationsschalter zeichnet sich vor allem durch seine Robustheit gegenüber Ablagerungen und Anhaftungen bei staubartigem und granuliertem Schüttgut aus [3].
So robust wie sie auch sein mögen, haben die Sensordaten von Grenzwertschalter eine geringe Aussagekraft. Für die zuverlässige

Planung der Bauarbeiten ist eine stufenlose Messung des Füllstandes erforderlich.
Die fünf gängigsten Messverfahren bei der stufenlosen Füllstandüberwachung in Schüttgutsilos sind

- Elektromagnetisches Lotsystem
- Kapazitive Sensorik
- Radar (freistrahlend, geführt)
- Radiometrie
- Ultraschall

Elektromagnetische Lotsysteme verfügen über ein Tastgewicht, welches über ein gewickeltes Messband oder –seil auf die Füllgutoberfläche herabgelassen wird [5]. Die abgewickelte Messbandlänge wird über ein Zählrad und einen dazugehörigen Induktivgeber überwacht [5]. Das Auftreffen des Tastgewichtes auf der Materialoberfläche wird durch das plötzlich abfallende Lastmoment sensorisch detektiert und anschließend der Füllstand berechnet. Elektromagnetische Lotsysteme sind vor allem für besonders große Silos mit einer Höhe von bis zu 70 m mit Schüttgütern wie Kalk, Steine oder Koks geeignet [5]. Die Vorzüge dieses einfachen Systems kommen vor allem bei der Entstehung von brennbaren Stäuben zum Ausdruck. Für mobile Silos ist das System u.a. wegen der Sensitivität gegenüber Drücken über 2 bar ungeeignet.
Kapazitive Messverfahren sehen ein Kondensatorgebilde aus Silowand und einer Messsonde vor [2] [5] und können bei einer Messfrequenz zwischen 20 kHz und 2 MHz sowohl zur Grenzwertmessung [4] als auch zur kontinuierlichen Messung [2] genutzt werden. Die Kondensatorkapazität, die direkt von der Dielektrizitätskonstante abhängig ist, variiert mit der Veränderung der Materialeigenschaften der Stoffe zwischen Silowand und der kapazitiven Messsonde [2] [4] [5]:

- Isolation der Sonde
- Silogeometrie
- Füllguttemperatur
- Körnung des Schüttgutes
- Abrasivität
- Chemische Agressivität
- Druck im Siloinneren
- Kondensatbildung

Die Kapazität wird entweder durch Resonanzschaltungen oder selbstabgleichende Wechselstrombrücken gemessen. Grundsätzlich gilt, dass mit steigendem Füllstand die Kapazität ebenfalls steigt [5].
Kapazitive Füllstandmessverfahren sind kostengünstig [6], hinreichend robust gegen Druckunterschiede [2] und ziehen gegenüber allen anderen gängigen Messsystemen den geringsten Aufwand mit sich [4]. Die Nachteile sind jedoch die Vielzahl an Einflussparametern auf die Kondensatorkapazität und der hohe Energieverbrauch.
Seit 2001 wird Radar zur Füllstandermittlung in Schüttgutsilos eingesetzt [7]. Dabei strahlt die trichterförmige Antenne (vgl. Abb. 1), die in einer kleinen Behälterstutze montiert ist [1], einen Funkwellenimpuls aus, welcher sich im Frequenzbereich zwischen 10-79 GHz mit Lichtgeschwindigkeit ausbreitet [5] [6] [8]. Dieses Signal wird auf der Materialoberfläche, wegen des Dielektrizitätszahlunterschieds zwischen Luft und Schüttgut, reflektiert [6]. Zur Auswertung der reflektierten Funkwelle können zwei Verfahren zum Einsatz kommen:

a. Puls-Radar-Verfahren
b. Frequenzmodulierte Dauerstrichmethode (z. engl. **F**requency **M**odulated **C**ontinuous **W**ave radar)

Beim Puls-Radar-Verfahren wird die Laufzeit vom Entsenden des Signals bis zum Empfangen des reflektierten Echosignals gemessen. Trotz des geringen Energiebedarfs und den unwesentlichen Betriebskosten [9], birgt dieses Verfahren gewisse Nachteile. Diese resultieren aus den bei Füllstandmessungen üblich kleinen Laufzeiten im Nano-Sekundenbereich (10^{-9} s) [5] und der Überlappung des Sende- und Empfangssignals. Die Auflösung und Filterung der hochwertigen Auswerteelektronik muss diesen Phänomenen genügen.
Das häufiger verwendete FMCW-Verfahren sieht eine geringfügige Variation der kontinuierlichen Trägerfrequenz vor [5]. Somit kann anhand der exakten Frequenz des reflektierten Signals die Laufzeit und somit der Füllstand kontinuierlich überwacht werden.
Der Einsatz von Radar zur Füllstandüberwachung bringt eine Robustheit gegenüber Staubbildung [1] [7], Anhaftung an der Antenne und Temperatur- und Druckschwankungen mit sich [1] [5] [7]. Darüber hinaus sind der große Messbereich von bis zu 100 m und die Einsetzbarkeit unter dynamischen Prozessbedingungen vorteilhaft. Rein qualitativ wird die Radarfüllstandmesstechnik für Schüttgutsilos oft als das „beste“ Füllstandmessverfahren gewertet [6]. Die Nachteile sind vor allem der hohe Energieverbrauch, die Anschaffungskosten und der enorme Aufwand bei der Erstinbetriebnahme [7]. Weitere Einschränkungen liegen vor, wenn die empfangenen Reflektionen des Sendesignals weniger direkt von der Schüttgutoberfläche, sondern vielmehr von Reflektionen auf Einbauten, Rührwerk, Schweißnähten, etc. herrühren (vgl. Abb. 1) [7] [8].

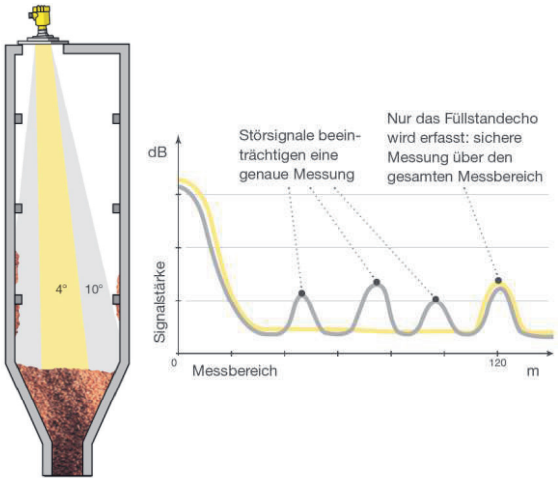


Abb. 1: Radar-Füllstandmesstechnik [8]

In solchen Fällen kann der Einsatz eines geführten Radars anstelle eines freistrahlenden Radars zielführend sein. Dabei wird die Funkwelle nicht in den freien Raum abgestrahlt, sondern in einen Metallstab, welcher im Material eingetaucht ist. Der Teil der Radarwellen, der an der Schüttgutoberfläche reflektiert wird, ermöglicht eine Laufzeitmessung [10].

In gewissen Anwendungen wird der Füllstand eines Schüttgutsilos mittels der Radiometrie ermittelt. Das Prinzip beruht dabei auf der Schwächung von Gammastrahlen durch die Masse des durchstrahlten Materials [5]. Dabei befindet sich an einer Stelle am äußeren Mantel des Silos ein Gammastrahler mit radioaktiven Isotopen und auf der anderen Seite ein Strahlenempfänger. Das Gamma-Strahlenbündel durchdringt das Silo bspw. diagonal durch die Mittelachse. Das Messverfahren auf Basis der Radiometrie weist in der Regel keine Einschränkungen durch Betriebsbedingungen oder Materialeigenschaften auf [5]. Jedoch ist der Einsatz von radioaktivem Material mit hohen Sicherheitsrisiken verbunden [5]. Beschaffung, Lagerung, Handhabung und Transport unterliegen strengen Auflagen. Daher kommt die kostenintensive radiometrische Messung lediglich dann zum Einsatz, wenn alle anderen Messverfahren versagen [5].

Das gängigste Messverfahren für die Füllstandüberwachung bei mobilen Schüttgutsilos ist Ultraschall. Im Gegensatz zum Radar werden hier steilflankige Schallwellenimpulse mit einer Dauer von 1 ms im Frequenzbereich zwischen 20 kHz und 2 MHz in der Regel mittels eines piezoelektrischen, früher noch oft magnetorestriktiven Wandlers erzeugt [2] [3] [5]. Die Schallwelle kann zusätzlich durch eine akustische Linse und einer gekrümmten Wandleroberfläche fokussiert werden [3]. Dieser Impuls breitet sich mit Schallgeschwindigkeit aus und wird an der Schüttgutoberfläche aufgrund des Dichteunterschiedes zwischen Luft und Schüttgut reflektiert [6]. Mit zunehmender Frequenz wird der Ausbreitungscharakter der Schallwelle gradlinig und sie verhält sich wie ein Lichtstrahl [2]. Daher gilt in dem Fall das Reflexions- und Brechungsgesetz aus der Optik. Wichtig ist erneut einen anwendungsspezifischen, zeitlichen Mindestabstand zwischen zwei Sendeimpulsen einzuhalten, um eine Überlagerung des Sendeimpulses mit dem Echosignal zu verhindern. Das Verfahren ist wirtschaftlich und birgt flexible Einbaumöglichkeiten [6]. Die Genauigkeit, die Wartungsfreiheit und die Robustheit gegenüber zeitabhängigen Einflussparametern begründen den dominanten Einsatz von Ultraschall [3][6]. Ultraschall versagt jedoch bei Staubbildung [7]. Es kommt zu Reflektionen an den Staubpartikeln, Einbauten und Schweißnähten und ist daher für dynamische Einsatzbedingungen kritisch. Die Temperaturabhängigkeit der Schallgeschwindigkeit bringt gerade bei Anwendungen im Außenbereich einen Laufzeitfehler mit sich [3] [6].

Konzept „Silopic“

Der in [11] vorgestellte "Silopic"-Ansatz zur Füllstandabschätzung in Schüttgutsilos beinhaltet unter anderem ein neues Sensorkonzept des Instituts für **M**ess**t**echnik und **S**ensorik (**MTS**) der Technische Universität Kaiserslautern in Kooperation mit WIKON Kommunikationstechnik GmbH (**WIKON**) und HUEBINET Informationsmanagement GmbH & Co. KG (**HUEBINET**) auf Grundlage der "Steinwurfmethode". Das System besteht aus einem Aktuator und einem Sensor, der von außen mit Magneten am Silomantel befestigt ist (vgl. Abb. 2). Der Aktuator ist ein elektromagnetisches Schlagwerk, das die Silowand durch einen mechanischen Impuls anregt. Dadurch wird das Silo zur Schwingung angeregt und die akustische Impulsantwort über den Sensor aufgezeichnet. Die Schwingungsdaten werden über das Mobilfunknetz an eine Cloud weitergeleitet. Dort befindet sich ein Software-Modul, das die Sounddaten auswertet. Ziel ist die Zuordnung vom Klang zu einem Füllstand. Hierzu werden die Erkenntnisse verwendet, die aus analytischen Berechnungen, numerischen Simulationen mit Kopplung der Finite-

Elemente-Methode und der Diskrete-Elemente-Methode, experimentellen Tests und Ansätzen des maschinellen Lernens resultieren. Die Füllstandinformation wird dem Benutzer über eine App zur Verfügung gestellt. Diese App beinhaltet ebenfalls eine automatisierte und optimierte Logistikplanung für den Schüttgutzulieferer.

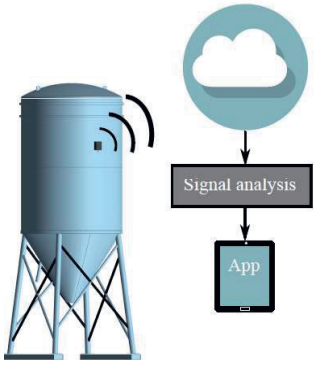


Abb. 2: Sensorkonzept „Silopic“ [11]

Im Gegensatz zu herkömmlichen Messsystemen hat das Silopic-System den Vorteil, dass die Messeinheit nicht invasiv montiert werden kann, sodass keine Änderungen am eigentlichen Silo vorgenommen werden müssen. Da Silos eine lange Lebensdauer von ca. 20 Jahren haben, ist die uneingeschränkte Montage des Sensors über Magnete an der Siloaußenwand eine komfortable Lösung. Dieser Aspekt spiegelt sich im Kaufpreis wider. Der Einsatz derselben Sensoreinheit für mehrere Silos ist denkbar. Da die Sensoreinheit auf der Außenseite des Silos montiert wird, ist das Gehäuse zwar den Umwelteinflüssen ausgesetzt, jedoch nicht der stark abrasiven, druckbehafteten Bestrahlung durch Feststoffpartikel. Voraussichtlich ist die Lebensdauer von Silopic länger als die von konventionellen Systemen. Die mobile funkbasierte Kommunikationsfähigkeit der Sensoreinheit ermöglicht einen hohen Automatisierungs- und Vernetzungsgrad bei der Bestellung und Bereitstellung wesentlicher Baustoffe. Dies kann die Effizienz und damit die Rentabilität von Bauprojekten erheblich steigern.

Analyse der akustischen Impulsantwort

Die akustische Impulsantwort ist eine konstant gedämpfte Schwingung und gemäß DIN 1311 [12] als transiente Schwingung zu kategorisieren. Diese entsteht bei einer sprung- oder impulsartigen Anregung und zeichnen sich durch ein chaotisches Schwingverhalten zu Anfang aus (vgl. Abb. 3). Anschließend stellen sich typischerweise geordnete Grund- und harmonische Oberschwingungen ein. Eine spezifische Schwingung einer gewissen Frequenz kann dabei am Erregungsort punktuell abklingen, während sie entlang der Mantelfläche wandert. Kommt es nun an einer Grenzfläche wie z. B. einer Schweißnaht oder Füllstandgrenzfläche zu einer Reflektion der Schallwelle, so kann am Erregungsort ein Aufklingen dieser spezifischen Schwingung beobachtet werden. Somit ist das Frequenzspektrum der akustischen Impulsantwort über die Zeit zu beobachten.

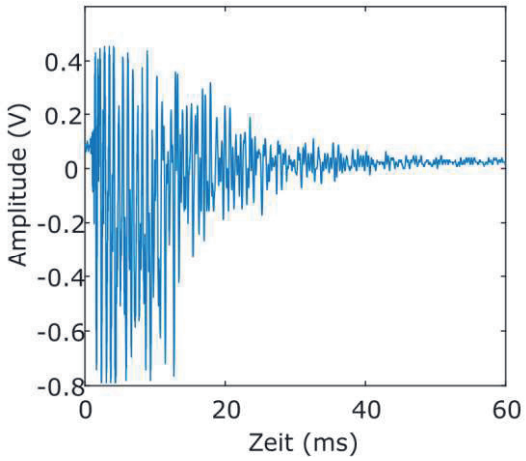


Abb. 3: Impulsantwort eines befüllten Silos

Hierzu werden drei Messverfahren in Erwägung gezogen:

- Beschleunigungssensor
- Mikrofon
- Akustische Kamera

Heutige Beschleunigungssensoren arbeiten in der Regel piezoelektrisch oder kapazitiv und können zur Messung von Körperschall eingesetzt werden (vgl. Abb. 4). Die piezoelektrischen Beschleunigungssensoren verfügen über eine seismische Masse, die an einem piezokeramischen Sensorplättchen befestigt ist [4]. Dieser Piezowiderstand wandelt dynamische Druckschwankungen in elektrische Signale um.

Kapazitive Beschleunigungssensoren sind in der Regel **M**ikro-**E**lektro-**M**echanische **S**ysteme (**MEMS**). Diese Feder-Masse-Systeme werden in der Regel aus Silizium hergestellt [4]. Dabei wird eine H-förmige Masse von vier fest verankerten Stegen in der Funktion als tragende Feder gehalten (vgl. Abb. 4). Durch die Bewegung der Masse findet eine relative Verschiebung zwischen dieser Masse und der unbeweglichen Bezugselektroden statt, welche durch die lineare Änderung der Kapazität erfasst werden kann [4].

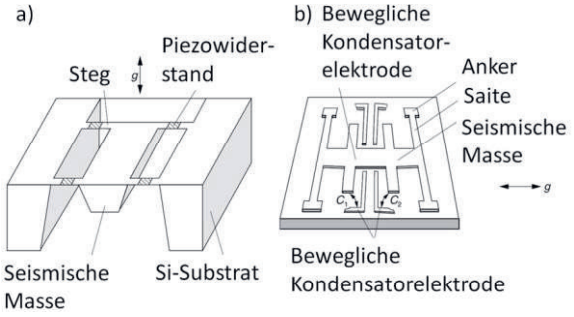


Abb. 4: a) Piezoelektrischer Beschleunigungssensor, b) MEMS [13]

Gängige Mikrofone lassen sich heutzutage weitestgehend in fünf Kategorien eingliedern [14]:

- Kontaktwandler
- Piezoelektrisches Mikrofon
- Elektrodynamisches Mikrofon
- Elektrostatisches Mikrofon
- Hitzedraht-Mikrofon

Allesamt eignen sich üblicherweise zur Messung von Luftschall im Frequenzbereich zwischen 20 Hz bis 20 kHz. Die unterschiedlichen Messprinzipien seien in diesem Rahmen nicht näher erläutert.
Eine Akustische Kamera besteht aus mehreren gezielt angeordneten Mikrofonen (daher auch bekannt als Mikrofonarray) [14]. Eine Schallwelle breitet sich kugelförmig von ihrem Quellort aus und trifft im homogenen akustischen Fernfeld auf jedes Mikrofon mit einer unterschiedlichen zeitlichen Verzögerung auf. Nun wird durch die Interferenzrekonstruktion der Quellort bestimmt [14] und mithilfe der bildgebenden Daten einer Kamera visualisiert. Eine akustische Kamera ist geeignet um stark abstrahlende Stellen zu detektieren.
Da ein mobiles Schüttgutsilo eine Vielzahl von inhomogenen, lokalen Versteifungen besitzt, konnte messtechnisch nachgewiesen werden, dass die akustische Impulsantwort eine starke Ortsgebundenheit aufweist. Silopic muss jedoch gewisse Toleranzen hinsichtlich der Montageposition kompensieren können. Daher wurde auf die Verwendung von Beschleunigungssensoren für den Prototyp verzichtet. Bei der Verwendung der akustischen Kamera war die Positionierung außerhalb des inhomogenen akustischen Nahfeldes nicht möglich, da die Störfaktoren aus der Umgebung die akustische Impulsantwort zu stark überlagerten. Daher konnte lediglich im Nahfeld festgestellt werden, dass die beiden vertikalen Enden des senkrechten Silos eine hohe Abstrahlungsenergie aufweisen (vgl. Abb. 5).



Abb. 5: Analyse des vertikalen Silos mit der akustischen Kamera

Für den Silopic-Prototyp eignet sich vor allem ein elektrostatisches Mikrofon. Diese sind preiswert, können die akustische Impulsantwort mit ausreichender Qualität messen und weisen trotz der Positionierung im Nahfeld eine gewisse Toleranz hinsichtlich der Montageposition auf.
Da die Messung an einem mobilen Schüttgutsilo im Realeinsatz eine gewisse Komplexität mit sich zieht, werden die Messungen im Rahmen dieser Arbeit an einem Prüfstand in Laborumgebung mit konstanter Temperatur und ohne weitere wesentliche Störeinflüsse untersucht (vgl. Abb. 6).



Abb. 6 Silo-Prüfstand

Die Frequenzverteilung oder das Spektrum eines Zeitsignals kann in den meisten Fällen ohne weiteres nicht beurteilt werden. Daher ist ein geeignetes Spektralanalyseverfahren erforderlich. Die bekannteste Spektralanalyse ist die **F**ourier**t**ransformation (**FT**), die sich wiederum in vier verschiedene Verfahren untergliedert [15].
Die FT approximiert im Grunde ein kontinuierliches Zeitsignal in ein kontinuierliches Spektrum, wobei das Signal als eine Summe von periodischen und stationären Sinus oder Kosinus unterschiedlicher Frequenzen und unterschiedlicher Amplituden rekonstruiert wird [16]. Da die Fouriertransformierte zwar eine

Lokalität im Frequenzbereich aber nicht im Zeitbereich hat, ist die Voraussetzung für die Approximation die stationäre Periodizität des Zeitsignals [15]. Da in diesem Anwendungsfall eine auf- und abklingende und damit eine instationäre Schwingung vorliegt, ist somit die Verwendung der FT unzulässig [12] [15] [16].
Für instationäre Signale kann die Kurzzeit-Fouriertransformation (zu engl. **S**hort-**T**ime **F**ourier **T**ranformation [**STFT**]) geeignet sein [15] [16]. Dabei wird nicht das gesamte Signal fouriertransformiert, sondern mittels eines Fensters in mehrere Teilsegmente unterteilt und anschließend jedes Teilsegment quasi unabhängig analysiert.
Dadurch wird zusätzlich zur Frequenzauflösung eine grobe Zeitauflösung gewonnen [16]. Durch eine geeignete Fensterung können sogenannte Leck-Effekte [16] kompensiert werden. Die STFT eignet sich für Signale, die über die Zeit nur allmähliche Veränderungen im Spektrum aufweisen. Bei einer Impulsanregung klingt jedoch die Schwingung aus dem Ruhezustand rapide auf und unmittelbar danach wieder ab. Signale mit einer derartigen Dynamik können mit einer STFT nur sehr eingeschränkt analysiert werden, da i. d. R. weiterhin eine stationäre, periodische Funktion als Basisfunktion verwendet wird. Als Ausnahme sei hier die Gabortransformation genannt, eine Sonderform der STFT, die starke Analogien zur **W**avelet**t**ransformation (**WT**) aufweist [15].
Die WT ist eine lineare Zeit-Frequenz-Transformation, die oft als eine „Verbesserung der STFT“ gesehen wird [15]. Obwohl die Grundidee bereits 1910 von Haar beschrieben wurde, nahm sie erst seit 1990, nach Entdeckung von kompakten, stetigen und orthogonalen Wavelets durch Daubechies 1988 und die Entwicklung der schnellen WT 1989, Einzug in die breite Anwendung [16]. Eine Wavelet ist eine im Grunde beliebige Funktion, die eine wellenartige Oszillation aufweist und zu Anfang und Ende gegen null läuft (vgl. Abb. 7) [16]. Die meisten Wavelets haben keine geschlossene Form. Bei der WT wird die gewählte Wavelet, die sich voraussichtlich am besten zur Approximation eines spezifischen Rohsignals eignet, als Basisfunktion verwendet. Für die hier vorliegende rasch auf- und abklingende Schwingung wird die Morlet-Waveletfunktion gewählt.

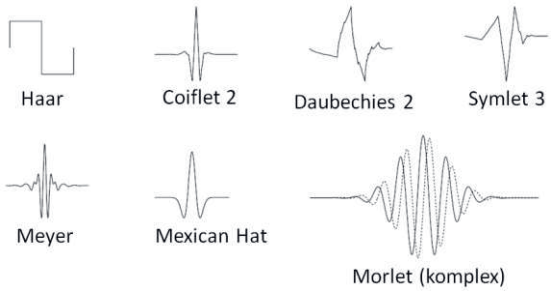


Abb. 7: Bekannte Waveletfunktionen [16]

Kennwertdefinition

Primäres Ziel der experimentellen Untersuchungen war es, den Einfluss des Schüttgutfüllstandes auf die Impulsantwort in Erfahrung zu bringen. Anschließend ist ein Kennwert zu definieren, der ausreichend mit dem Füllstand korreliert. Hierzu wurden zwei Messreihen durchgeführt, wobei jeweils das volle Silo auf dem Prüfstand gleichmäßig entleert wurde. Zwischen den beiden Messreihen wurde der Sensor um 3 cm verschoben, um eine Aussage hinsichtlich der Robustheit gegenüber Montagetoleranzen treffen zu können.
Abb. 8 zeigt den Vergleich zwischen der Impulsantwort eines a) vollen und b) leeren Silos. Über die Abszisse ist die Zeit in ms und über die Ordinate die Frequenzverteilung bei 0-16 kHz dargestellt. Die Dämpfung durch das Schüttgut (vgl. Abb. 8a)), die mit zunehmender Frequenz stärker wird, ist offensichtlich. Die Impulsantwort klingt bei einem vollen Silo nach ca. 30 ms ab, wobei die Schwingung bei einem leeren Silo (vgl. Abb. 8a)) zum gleichen Zeitpunkt lediglich ein wenig gedämpft ist.
Um aus den transformierten Daten einen zulässigen Füllstands-Kennwert zu berechnen, werden für unterschiedliche Zeit- und Frequenzbereiche die sieben gängigsten statistischen Kenngrößen ermittelt (Mittelwert, Median, Minimum, Maximum, Modus, Streuung, Varianz) [17].
Dabei zeigt der Medianwert im Zeitbereich 11-14 ms nach dem Schlag und in einem Frequenzbereich von 2500 – 5000 Hz eine deutliche Sensitivität hinsichtlich des Füllstandes zwischen 100 % und 95 % und zwischen 40 % und 0 % (vgl. Abb. 9). Bei einem Füllstand zwischen 95 % und 40 % ist die Aussagekraft des Kennwertes aufgrund der großen Streuung gering.

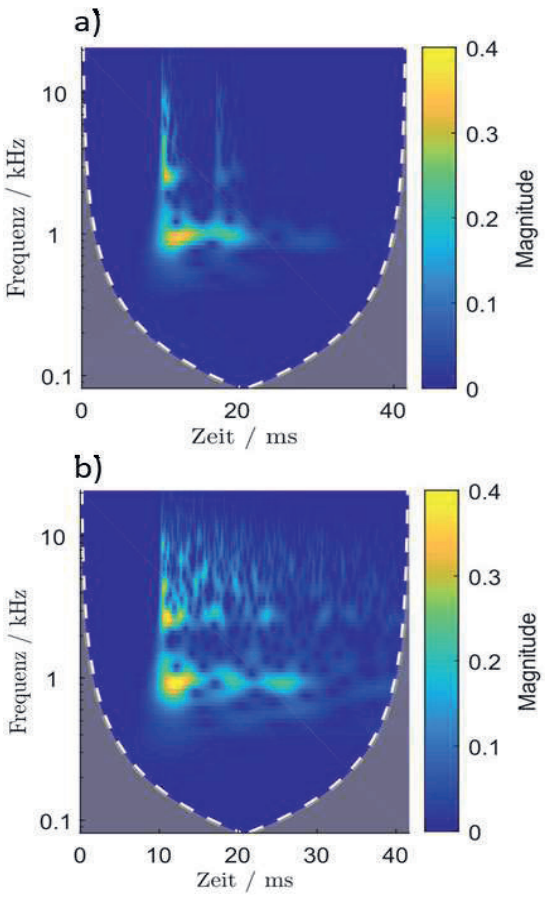


Abb. 8: Akustische Impulsantwort eines a) vollen Silos und b) eines leeren Silos

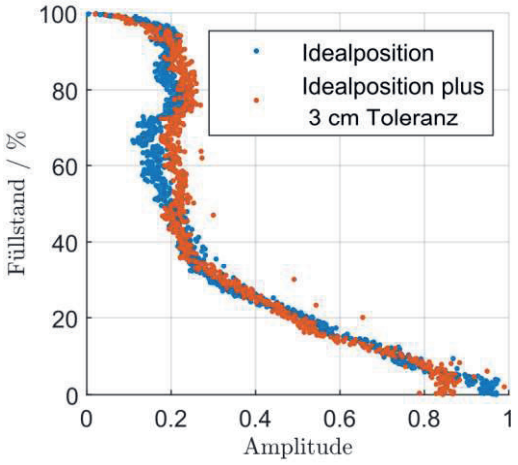


Abb. 9: Verlauf des Silopic-Kennwertes im Frequenzbereich von 2500 - 5000 Hz und Zeitbereich von 11 – 14 ms nach Impulsanregung

Zusammenfassung

Im Rahmen dieser Arbeit wurde der heutige Stand der Füllstandmesstechnik für Schüttgüter erläutert. Weiterhin wurde ein neues Konzept für einen akustischen Füllstandsensor zur Anwendung an mobilen Schüttgutsilo präsentiert. Die Eignung verschiedener Sensoren wie Beschleunigungssensoren, Mikrofone oder akustische Kameras wurden beurteilt. Zur Analyse der Impulsantwort wurde eine geeignete Methode der Signalverarbeitung gewählt. Erste experimentelle Untersuchungen haben den Einfluss des Füllstandes auf die akustische Impulsantwort eines Silos auf dem Prüfstand gezeigt. Auf Basis dieser Daten konnte erstmalig ein prototypischer Kennwert berechnet werden, der hinreichend mit dem Schüttgutfüllstand korreliert.

Literaturnachweis

[1] Der Joker im Silo*; PROCESS;* url: https://www.process.vogel.de/der-joker-im-silo-a-55399/*;* zugegriffen am 08.05.2019
[2] K. W. Bonfig, A. Liske; Fullstand-Meßtechnik; *Hans Holzmann Verlag KG*, (1973)
[3] E. Hering, H. Steinhart; Taschenbuch der Mechatronik; *Carl Hanser Verlag*, (2015); ISBN: 978-3-446-43857-6
[4] G. W. Schanz; Sensoren – Fühler der Meßtechnik; *Dr. Alfred Hüthig Verlag Heidelberg*, (1988)
[5] G. Strohrmann; Meßtechnik im Chemiebetrieb; *R. Oldenburg Verlag München Wien*, (1995); ISBN: 3-486-22999-0
[6] M. Rohn; Füllstandmessung mit Ultraschall; *Endress+Hauser*, (2014)
[7] H. Bohlmann, S. Mühlenkamp; Wie das Radar Schüttgut lernte; *PROCESS*; url: https://www.process.vogel.de/wie-das-radar-schuettgut-lernte-a-421144/; zugegriffen am 08.05.2019
[8] Der Radarsensor für Schüttgüter VEGAPULS 69; *VEGA*, (2014)
[9] M. J. Heim, A. Mayr, *Endress+Hauser GmbH*, (2001)
[10] Geführtes Radar: VEGAFLEX Serie 80; *VEGA*, (2015)
[11] L. Anslinger, P. Sivasothy, F. Torner, J. Seewig; Mobile sensor concept for acoustic level estimation in bulk solids silos; *IEEE Sensors Applications Symposium (SAS)*, (2018); doi: 10.1109/SAS.2018.8336760
[12] Schwingungen und schwingungsfähige Systeme; *Deutsches Institut für Normung*; (2000); DIN 1311-1
[13] W. Schmidt; Sensorschaltungstechnik; *Vogel Buchverlag*, (2002); ISBN: 978-3-8343-3111-3
[14] M. Eichler; Breitbandige Beamforming-Algorithmen zur Erfassung von Audiosignalen mit kompakten Mikrofon-Arrays; *Dissertation*, (2009)
[15] A. Mertins; Signaltheorie, *Springer Verlag*, (2013); doi: 10.1007/978-3-8348-8109-0
[16] S. Tamm; Hochaufgelöste Zeit-Frequenz-Analysen ereigniskorrelierter EEGOszillationen mittels S-Transformation; *Dissertation*, (2005)
[17] L. Fahrmeir, R. Künstler, I. Pigeot, G. Tutz; Statistik; *Springer Verlag*, (2007); ISBN: 978-4-540-69713-8

DOI 10.5162/sensoren2019/P2.4

ExtruSens – Low-Cost-Spektroskopie zur Inline-Prozesskontrolle in der Kunststoffverarbeitung

Norbert Halmen[1], Christoph Kugler[1], Benjamin Baudrit[1], Thomas Hochrein[1], Martin Bastian[1], Hans-Peter Jungbauer[2], Harald Nitsche[2]
[1]SKZ – Das Kunststoff-Zentrum, Friedrich-Bergius-Ring 22, 97076 Würzburg, Deutschland
[2]J & M Analytik AG, Willy-Messerschmitt-Straße 8, 73457 Essingen, Deutschland

Zusammenfassung

Der Ruf nach einer lückenlosen Qualitätskontrolle wird auch in der Kunststoffaufbereitung und -verarbeitung lauter. Hierbei wird es immer wichtiger, bereits im laufenden Prozess entscheidende Materialparameter zu kennen. Allerdings sind Messsysteme auf Basis optischer Spektroskopie, welche eine Vielzahl von relevanten Kenngrößen ermitteln können, sehr teuer und oft zu komplex in der Anwendung.
Um die Anwendung der Spektroskopie in der Prozesskontrolle zu vereinfachen und eine deutlich preiswertere Lösung anbieten zu können, wurde die modulare Systemplattform ExtruSens entwickelt. Diese ermöglicht die gezielte Inline-Bestimmung verschiedener Prozesskenngrößen mittels kostengünstiger LED-Beleuchtung und Halbleiter-Sensorik. Durch den Einsatz druck- und temperaturstabiler Sonden kann direkt in der Kunststoffschmelze gemessen werden. Es wurden drei grundlegende Funktionalitäten realisiert: Bestimmung des Feuchtegehalts von Granulaten und Pulvern, Detektieren des hydrolytischen Abbaus von hygroskopischen Polymeren wie Polyamid (PA) und Polyethylenterephthalat (PET) sowie Überwachen der Farbe der Kunststoffschmelze für einen besseren Einblick in Farbwechsel- und Ausspülprozesse.
Im Vergleich mit den Referenzverfahren liefert ExtruSens gut reproduzierbare Ergebnisse und bietet zudem noch Potential für weitere Entwicklungsstufen.
Durch die spezialisierte Anwendung sind die einzelnen Module zu einem Fünftel des Preises klassischer spektroskopischer Lösungen verfügbar. Als weitere Einsatzfelder für die ExtruSens-Plattform sind die Lebensmittel- und Pharmaindustrie interessant.

Keywords: Spektroskopie, Kunststoffschmelze, Farbmessung, Feuchte, Polymerabbau

Motivation

Spektroskopische Untersuchungen im ultravioletten (UV), sichtbaren (VIS) und nahinfraroten (NIR) Bereich gehören in zahlreichen Einsatzbereichen, wie der Lebensmittel-, Pharma-, Petro- und chemischen Industrie, mittlerweile zu den Standardmethoden der Qualitätsüberwachung. Auch die Kunststoffaufbereitung und -verarbeitung stellt ein vielversprechendes Anwendungsfeld dar, da gerade dort der Ruf nach einer lückenlosen Qualitätskontrolle im laufenden Prozess immer lauter wird.
Speziell für den Einsatz in der Kunststoffschmelze gibt es jedoch bis jetzt keine kompakte inlinefähige Lösung. Bislang werden meist sehr umfangreiche und universelle Spektroskopie-Systeme angeboten. Mit diesen kann eine Vielzahl von Kenngrößen aus den Spektren ermittelt werden. Insbesondere die Detektion der Zusammensetzung bzw. die qualitative und quantitative Bestimmung von Additiven und Füllstoffen [1, 2] stellen die Kernanwendungen bei der Überwachung von Kunststoffaufbereitungsprozessen dar.
Dem gegenüber stehen jedoch erhebliche Nachteile. Die hohen Temperaturen bis zu 400 °C und enormen Drücke von einigen hundert bar beim Einsatz in Extrudern stellen hohe Ansprüche an die Robustheit möglicher Spektroskopiesonden. Weiterhin erfordern eine komplexe Bedienung sowie die aufwändige Kalibrierung oftmals teures Fachpersonal. Hinzu kommen hohe Kosten für Spektrometer und Wartung. Daher stellt die Nutzung derartiger Systeme in der Kunststoffindustrie nicht den Stand der Technik dar.

Grundlagen der optischen Spektroskopie

Bei der UV/VIS/NIR-Spektroskopie wird die Wechselwirkung von elektromagnetischer Strahlung (Licht) mit Materie ausgenutzt, um

z. B. Eigenschaften wie Bindungsverhältnisse und Struktur zu untersuchen.
Bei Bestrahlung mit Licht aus dem UV (200 – 380 nm) oder VIS (380 – 750 nm) werden Valenzelektronen angeregt.
Im NIR (750 – 2.500 nm) erfolgt die Absorption des Lichts durch Molekülschwingungen. Dies sind harmonische Oberschwingungen und kombinierte Schwingungen der Grundschwingungen aus dem mittleren IR.
Die Wellenlänge des absorbierten Lichts ergibt sich aus der für die jeweilige Anregung benötigten Energie:

$$E = h \cdot f = (h \cdot c)/\lambda \quad (1)$$

wobei *E* die Energie, *h* das Plancksche Wirkungsquantum, *f* die Frequenz, *c* die Lichtgeschwindigkeit und *λ* die Wellenlänge der elektromagnetischen Welle sind.
Die jeweiligen Valenzelektronen und verschiedenen funktionellen Gruppen, wie z. B. CH, CH_3, OH und NH besitzen charakteristische Frequenzen, bei denen sie eingestrahltes Licht absorbieren. Da der absorbierte Anteil im am Detektor ankommenden Licht fehlt, ergeben sich charakteristische Absorptionsbanden in den Spektren. Da diese Frequenzen bzw. Spektralbanden bekannt sind, können sie für die Substanzidentifizierung herangezogen werden [3, 4].
Bei der Messung an einer Probe wird ein Teil der eindringenden Lichtintensität in dieser absorbiert oder gestreut. Beide Mechanismen bewirken eine Abnahme der einfallenden Intensität. Diese Strahlungsabschwächung (Extinktion) wird durch das Lambert-Beersche Gesetz beschrieben:

$$E_x = -\log\left(I/I_0\right) = e^{-\alpha x} \quad (2)$$

mit der Intensität I_0 des eintretenden Lichtstrahls, der Intensität *I* des austretenden Lichtstrahls, dem wellenlängenabhängigen Absorptionskoeffizienten α sowie der Dicke *x* der Probe.
Die Messung kann sowohl in Transmission- als auch in Reflexionsanordnung erfolgen. Letztere wird beim Einsatz an Extrudern oftmals bevorzugt, da nur ein einzelner Zugang zum Schmelzekanal notwendig ist.

Charakteristische Wellenlängen für Feuchte und Polymerabbau

Der Feuchtegehalt von PA kann anhand charakteristischer Spektralbanden im NIR bestimmt werden. Dies wurde bereits an Folien aus PA 6 und spritzgegossenen Probekörpern aus PA 6.6 demonstriert [5, 6]. In der zuletzt genannten Arbeit wurde durch die Verwendung von geeigneten chemometrischen Modellen eine Nachweisgenauigkeit von 0,05 % erreicht.
Es existieren zudem einige Untersuchungen des Feuchtegehalts von Kunststoffschmelzen im NIR. Dabei konnte mittels Inline-Spektroskopie im NIR der Feuchtegehalt sowohl in PET-Schmelze beim Spritzguss [7] als auch in Polylactid (PLA)-Schmelze [8] detektiert werden. Im Falle von PLA betrug die Nachweisgrenze 0,04 %.
Der Einfluss von Feuchte und Temperatur auf den Abbau von PLA-Fasern wurde ebenfalls anhand charakteristischer Banden im NIR deutlich gemacht [9]. Zudem wurden charakteristische Veränderungen im VIS-Spektralbereich durch den thermischen Abbau von PLA-Schmelze nachgewiesen [10].

Anpassung an die konkrete Anwendung

Um die Anwendung der Spektroskopie in der Prozesskontrolle zu vereinfachen und zu deutlich geringeren Kosten zu ermöglichen, wurde im Rahmen eines Forschungsvorhabens der J & M Analytik AG, Essingen, und dem Kunststoff-Zentrum – SKZ, Würzburg, ein neuer Weg beschritten. Anstelle von teuren Universallösungen wurden spezifische Sensoren konzipiert, welche nur bei definierten Wellenlängen arbeiten und standardisierte Aufgaben in der Kunststoffverarbeitung abdecken. Prinzipiell können Messungen bei unterschiedlichen Wellenlängen auf zwei Arten erfolgen:

- Verwendung eines Detektors und mehrerer diskreter Wellenlängen zur Beleuchtung (LED, Laser)
- Nutzung einer breitbandigen Lichtquelle (Halogenlampe) und eines scannenden Systems zur Detektion

Die Verwendung mehrerer Detektoren z. B. eines Dioden-Arrays ist für eine kostengünstige Lösung zu teuer.
In Kombination mit einem modularen Plattformkonzept können solche Sensoren mit definierten Wellenlängen um eine Größenordnung preisgünstiger als universelle Mess- und Spektroskopie-Systeme angeboten werden. Möglich wird dies durch die in den letzten Jahren immer weiter sinkenden Kosten – bei gleichzeitiger Leistungssteigerung – von Halbleiterkomponenten sowie der Fortschritt der LED-Technologie.
Kommerziell sind bereits z. B. Inline-Systeme zur Feuchtebestimmung von Agrargütern oder für die Lebensmittelanalytik verfügbar [11]. Diese arbeiten bei wenigen definierten Wellenlängen im VIS- oder NIR-Bereich. Der Einsatz von LEDs oder Lasern als Lichtquelle zur Untersuchung von Flüssigkeiten und Gasen hat sich mittlerweile bewährt und bietet eine ausreichend hohe Messgüte [12, 13].

Es wurden Sensoren auf Basis von LED-Beleuchtung mit drei grundlegenden Funktionen entwickelt [14, 15]:

- Bestimmung des Feuchtegehalts von Granulaten und Kunststoffschmelze,
- Detektion des hydrolytischen Abbaus von hygroskopischen Polymeren in der Kunststoffschmelze,
- Überwachen der Farbe in der Kunststoffschmelze.

Bei der Ermittlung von Farbveränderungen erfolgt die Beleuchtung aufgrund der kostengünstigen Verfügbarkeit durch LEDs und entsprechende Fotodetektoren.
Zur Bestimmung des Feuchtegehalts und des Polymerabbaus sind theoretisch zwei Wellenlängen ausreichend: Die zur Anregung der Wasserbande bzw. OH-Bande und eine weitere Wellenlänge außerhalb dieser Wellenlängen als Referenz. Aufgrund der Vielzahl unterschiedlicher Kunststoffe ist aber eine größere Mindestanzahl an Wellenlängen für diese Anwendung erforderlich. Einerseits kann eine andere Referenzbande notwendig sein und zum anderen kann die Wasserbande verschoben sein. LEDs im NIR sind teurer als im VIS und nur für eingeschränkte Wellenlängen verfügbar, jedoch für diese Anwendung immer noch preiswerter als ein Universalspektrometer. Zudem bieten diese LEDs heute eine sehr gute Langzeitstabilität. Dadurch werden die Systeme für Anwender attraktiver.

Messsonden für Extruder

Für den Einsatz unter den extremen Umgebungsbedingungen, welche in der Kunststoffaufbereitung vorherrschen, wurden spezialisierte faseroptische Sonden konstruiert, welche bei Temperaturen bis 400 °C und Drücken bis zu 300 bar eingesetzt werden können (vgl. Abb. 1).

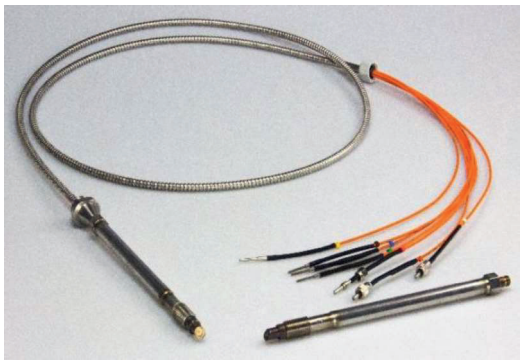


Abb. 1: ExtruSens-Extrudersonde mit sechs Glasfasern für Beleuchtung und einer Glasfaser für den Sensor/Detektor

Die Sonden sind für die Verwendung in Bohrungen mit dem in der Kunststoffindustrie etablierten ½“-20 UNF-Gewinde ausgelegt. Ein Saphirfenster schützt die Glasfasern vor direktem Kontakt mit der Kunststoffschmelze. Das Fenster ist für eine bessere Selbstreinigung im Schmelzestrom um 30° geneigt. Entsprechend des Konzepts zur Nutzung verschiedener Wellenlängen wurde die Sonde mit sieben Einzelfasern versehen – sechs für LEDs und eine für den Detektor. Die Fasern werden über den gängigen SMA 905-Faseranschluss an das Messsystem angeschlossen.

Modulares Messsystem ExtruSens

Das entwickelte Messsystem ExtruSens ist als modulare Systemplattform konzipiert. Die Sensoren für die verschiedenen Applikationen sind mit einem Gerät bzw. einer Software steuerbar. Die Sensoren sind zur Verwendung von faseroptischen Sonden optimiert. Sie bestehen aus einem Detektor (Si: UV-VIS, InGaAs: NIR) und bis zu sechs LEDs (380 – 2.150 nm). Der Betriebsstrom ist für jede LED separat einstellbar.
Im Zuge einer Messung werden der Dunkelstrom des Sensors und anschließend die Reflexionswerte für jede Wellenlänge erfasst. Um das Signal-zu-Rausch-Verhältnis zu verbessern, wird mit einer Frequenz von 1 kHz gemessen und innerhalb der Elektronik auf ca. 1 Hz gemittelt. Die Rohdaten werden nach einer Korrektur des Dunkelstroms auf der Elektronik zur Visualisierung und weiteren Berechnung auf einer Benutzeroberfläche zur Verfügung gestellt. Der Anwender erhält lediglich einen Kennwert. Um dies zu ermöglichen, erfolgt eine werkseitige Kalibrierung für die unterschiedlichen Anwendungsfälle mit gängigen Referenzverfahren.
Die aktuelle Entwicklungsstufe des Systems ist in Abb. 2 dargestellt. Eine weitere Miniaturisierung auf die Größenordnung des Farbumschlagsensors (oben aufliegende Box) ist als nächster Schritt geplant.

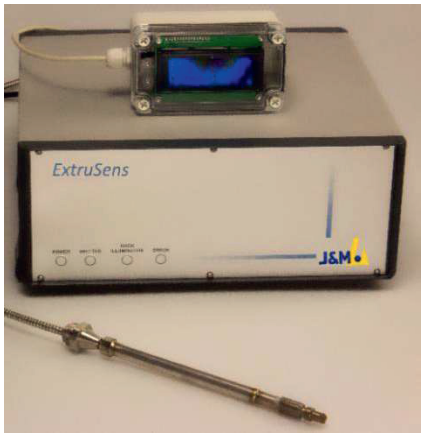


Abb. 2: Prototyp der ExtruSens-Plattform mit Farbumschlagsensor-Modul

Untersuchte Materialien

Für die Versuche zu Feuchte und Polymerabbau wurden beispielhaft zwei handelsübliche hygroskopische Materialien verwendet, ein PA 6 (Ultramid B27 E der BASF SE, Ludwigshafen) und ein PET (PET BC112 der Sabic Deutschland GmbH, Düsseldorf). In Tab. 1 sind die Sättigungswerte sowie die maximalen bzw. empfohlenen Feuchten für die beiden Materialien laut Hersteller enthalten.

Tab. 1: Angaben zu Materialfeuchte für eingesetzte Materialien laut Herstellerdatenblatt [16, 17]

Material	Feuchte /Gew.-%	
	Sättigung	Maximal bzw. empfohlen für Verarbeitung
Ultramid B27 E	2,2	0,1
PET BC112	< 0,350	0,003 – 0,004

Für die Versuche zum Farbumschlag wurde das Polypropylen (PP) Sabic PP 505P der Sabic Deutschland GmbH sowie verschiedene Farbmasterbatches auf PP-Basis verwendet.

Verwendete Referenzverfahren

Zum Referenzieren der Feuchte der Granulate wurde der HydroTracer FMX (aboni GmbH, Schwielowsee) verwendet. Das System arbeitet nach der Calciumhydrid-Methode (CHM) und ist selektiv auf Wasser. Durch eine herstellerseitige Kalibrierung erreicht dieses in der Kunststoffindustrie gängige Atline-Messsystem vergleichbare Messwerte wie die Karl-Fischer-Titration (KFT) – dem Standard für die Feuchtebestimmung im Labor [18].
Zur Quantifizierung des Polymerabbaus anhand der Kettenlängenverteilung kam die Gel-Permeations-Chromatographie (GPC) mit einer PSS SECcurity 1260 (PSS Polymer Standards Service GmbH, Mainz) zum Einsatz.
Für die Erkennung von Farbveränderungen sind keine absoluten Farbwerte notwendig. Dennoch wurden die zum Vergleich verwendeten Spektrometer auf Farbstandards kalibriert.

Prozessmessungen unter industrienahen Bedingungen

Für alle Versuche wurde ein gleichläufiger Doppelschneckenextruder ZE 25 A UTX der KraussMaffei Berstorff GmbH, Hannover, verwendet (Temperaturen: 200 °C (PP), 280 – 320 °C (PET); Drücke: 15 bar (PP), 3 bar (PET); Extruderdrehzahl: 300 U/min, Durchsatz: 10 kg/h). Die Farbversuche mit PP und den Farbmasterbatches wurden zusätzlich auf dem Doppelschnecken-Compounder ZSK 18 Mc (Coperion GmbH, Stuttgart) durchgeführt (Temperatur PP: 200 °C, Druck: 15 bar, Extruderdrehzahl: 400 U/min, Durchsatz: 5 kg/h). Die Messsonden wurden jeweils in Adaptern zwischen Extruderschnecke und Düse des Extruders eingesetzt.
Für den Polymerabbau während der Verarbeitung spielen Temperatur, Verweilzeit und Scherung eine wichtige Rolle. Die beiden ersten Größen wurden im Labormaßstab variiert. Dies erfolgte auf einem MiniLab HAAKE Rheomex CTW5 (Thermo Fisher Scientific GmbH, Schwerte). Die Bestimmung der Feuchte war hier ebenfalls wichtig, da Hydrolyse beim Polymerabbau ein wichtiger Faktor ist. Es wurden Granulate aus PA 6 und PET mit Feuchtewerten zwischen 0,30 Gew.-% und unter 0,01 Gew.-% mit Verweilzeiten zwischen 2 und 20 Minuten verarbeitet. Die Temperaturen betrugen 240 °C für PA 6 und 250 °C für PET. Mit dem System wurden kleine Plättchen für die spektroskopische Untersuchung hergestellt.

Bestimmung des Feuchtegehalts von Kunststoffgranulaten

Der Feuchtegehalt von Kunststoffgranulaten ist insbesondere bei der Verarbeitung von hygroskopischen Polymeren, wie PA, PET und PLA von hoher Relevanz für die spätere Bauteilqualität.
Zur Erzeugung verschiedener Feuchtewerte wurden die Granulate aus PA 6 und PET an Luft bis zur Sättigung mit Feuchte gelagert und danach in einem Granulattrockner Luxor BS 30 (motan-colortronic Gmbh, Friedrichsdorf) bei definierten Temperaturen (PA 6: 110 °C, PET: 140 °C) getrocknet. Zu ausgewählten Zeitpunkten wurden kleine Mengen der getrockneten Granulate entnommen und spektroskopisch sowie mittels CHM vermessen. Durch die Trocknung konnten je nach Material Feuchtewerte zwischen 0,30 Gew.-% und unter 0,01 Gew.-% erreicht werden.
Für die spektroskopischen Untersuchungen wurden ein NIR-Spektrometer (TIDAS NIR 1121, 1.100 – 2.100 nm) sowie ein IRIS 7-Reflextionsmessaufbau der J & M Analytik AG, verwendet. Die gemessenen Spektren bei verschiedenen Feuchtegehalten zeigen nach einer Standard Normal Variate (SNV)-Korrektur deutliche Unterschiede zwischen feuchten und trockenen Granulaten. Dies ist in Abb. 3 beispielhaft für PET gezeigt.

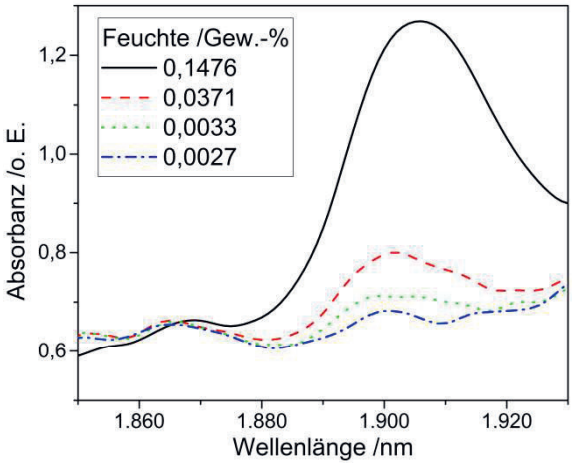


Abb. 3: Beispiel der Feuchtebanden des verwendeten PET für vier verschiedene Feuchtegehalte nach der SNV-Korrektur

Für die absolute Feuchtebestimmung erfolgte eine Kalibrierung der aufgenommenen Spektraldaten mit den gemessenen Feuchtegehalten. Hierfür wurde die Absorptionsbande von Wasser – in diesem Fall bei 1.910 nm – verwendet. Die Messungen zeigten eine gute Reproduzierbarkeit. Die Messgenauigkeit und Korrektur der Streuung an unterschiedlichen Granulatgrößen ist Inhalt weiterer Untersuchungen.
Der Auswertealgorithmus der finalen Anwendersoftware von ExtruSens ermittelt zunächst die basislinienkorrigierten Werte des Sensors. Die Feuchte wird danach anhand einer bereits werkseitig hinterlegten Kalibration berechnet – basierend auf Referenzmessungen wie CHM oder KFT. Der Endanwender hat auf Wunsch die Möglichkeit zur Eingabe von Referenzwerten für die Erstellung einer eigenen univariaten Kalibrierung. Alternativ zur absoluten Feuchtebestimmung kann der Feuchteverlauf auch mittels Grenzwerten für die Rohdaten des Sensors verfolgt werden.

Feuchtemessung in Kunststoffschmelze

Die Veränderung der Materialfeuchte während der Verarbeitung lässt sich nicht mittels CHM ermitteln. Daher wurden Granulate zu definierten Zeitpunkten nach der Trocknung sofort im Extruder zudosiert und die Feuchte zu diesem Zeitpunkt mittels CHM gemessen. Gleichzeitig wurde auf eine Entgasung verzichtet, um außer bei der Zudosierung einen weiteren Verlust von Feuchte im Prozess zu verhindern. Die spektroskopischen Messungen mit der Sonde erfolgten sowohl mit der entwickelten Elektronik als auch mit einem klassischen Spektrometer TIDAS NIR 1121.

Die Auswertung der Spektren der Extrudermessungen erforderte eine Basislinienkorrektur, um Temperatur- und Streueinflüsse zu kompensieren. Zusätzlich erfolgte eine gezielte Variation der Prozessparameter, Temperatur und Materialdurchsatz, um deren Einfluss auf die Spektraldaten zu untersuchen.
Der Wellenlängenbereich um die bei der Messung an Granulaten deutlich erkennbare Wasserbande war in den Spektren der Schmelzemessungen stark übersteuert (vgl. Abb. 5). Dies ist mit großer Wahrscheinlichkeit auf ein vermehrtes Vorkommen von hydrolytischen Abbauprodukten mit OH-Gruppen zurückzuführen, wie im nächsten Abschnitt beschrieben. Eine direkte Bestimmung der Feuchte war somit mittels Spektrometer nicht möglich.

Detektion von hydrolytischem Polymerabbau

Die im MiniLab hergestellten PA 6- und PET-Proben mit verschiedenen Abbaugraden lagen als dünne, transparente Plättchen vor. Diese sind beispielhaft für PA 6 in Abb. 4 dargestellt.

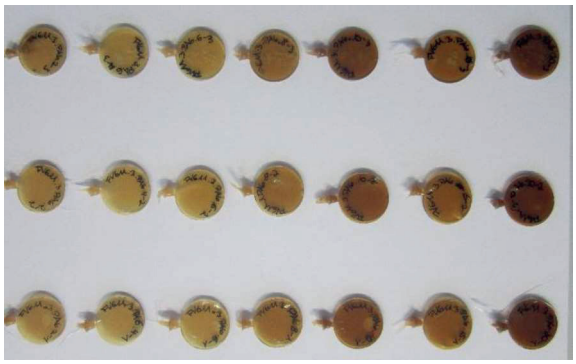


Abb. 4: Im MiniLab mit verschiedenen Feuchtegehalten und Verweilzeiten abgebaute PA 6-Proben

Aufgrund der Transparenz wurde in Transflexion gemessen. Hierfür wurden eine breitbandige Halogenlichtquelle (360 – 2.500 nm) in Verbindung mit Spektrometern im VIS- (300 – 1.100 nm) und NIR-Bereich (1.100 – 2.100 nm) sowie ein Weißstandard als Referenz eingesetzt. Dabei waren v. a. wachsenden Feuchtebanden im NIR (1.940 nm) und die zunehmende Gelb- bzw. Braunfärbung im VIS bei stärkerem Polymerabbau charakteristisch. Dies wurde auch durch kürzere mittlere Kettenlängen in der GPC bestätigt. Eine multivariate Kalibrierung mit den Ergebnissen der GPC war jedoch aufgrund der geringen Datenmenge nicht in ausreichender Qualität möglich.
Im MiniLab ist der Schereintrag aufgrund der verwendeten reinen Förderschnecken verhältnismäßig gering verglichen mit einem Doppelschneckenextruder, welcher in der Regel auch

über Knet- und Mischblöcke auf der Schnecke verfügt. Die Verweilzeit lässt sich am Extruder über Durchsatz und Drehzahl steuern. Zudem waren für die Messungen am Extruder die Parameter (Anfangs-)Feuchte und Temperatur maßgeblich. Die Untersuchungen erfolgten mit PET mit der entwickelten Sonde und dem TIDAS NIR 1121. In den Extruderversuchen zeigte sich eine starke Übersteuerung im Bereich der Wasserbande bzw. bei höheren Wellenlängen mit zunehmendem Polymerabbau. Mit zunehmender (Anfangs-)Feuchte schiebt sich die Flanke der gezeigten Bande immer weiter hin zu kleineren Wellenlängen. Der gleiche Effekt zeigt sich auch bei Erhöhung der Temperatur oder längeren Verweilzeiten während der Verarbeitung von Material mit einer definierten Anfangsfeuchte. Dies ist in Abb. 5 beispielhaft für drei Temperaturen gezeigt.

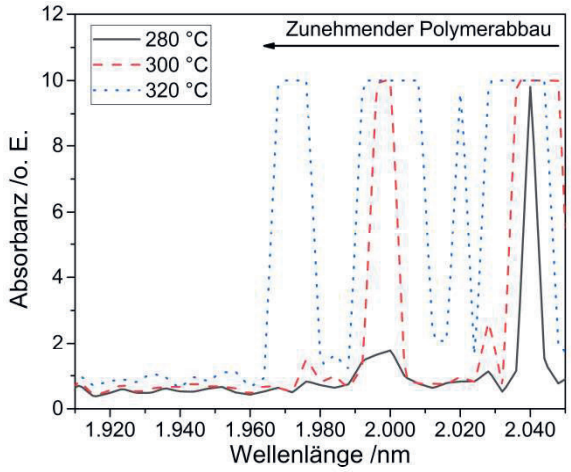


Abb. 5: Spektroskopische Messung der hydrolytischen Polymerabbaus von PET-Schmelze bei verschiedenen Temperaturen

Die verschiedenen Positionen der Flanken waren gut reproduzierbar. Die noch nicht erfolgte Quantifizierung des Abbaus, z. B. per GPC, und die Betrachtung weiterer Materialien ist im Rahmen weiterer Untersuchungen vorgesehen. Der derzeitige Algorithmus zur Detektion des Polymerabbaus ist eine einfache Schwellwertauswertung. Als Wellenlängen werden 1.200 nm (Referenz) sowie 1.940 nm und 1.960 nm verwendet. Somit ergibt sich eine Kombination des Signals aus Feuchte und hydrolytischem Polymerabbau. Für beide Applikationen kann derselbe Sensor zum Einsatz kommen.

Zusätzlich wurde zur Überwachung des Materialabbaus ein Inline-Farbmesssystem der ColVisTec AG, Berlin, eingesetzt. Mit diesem konnte die zunehmende Gelbfärbung gemessen werden. Der Einsatz eines Farbmesssystems zur Abbaudetektion ist allerdings nur bei nicht oder nur gering eingefärbten Kunststoffschmelzen möglich.

Messung von Farbveränderungen

Basierend auf vorhandenen Kenntnissen im Bereich der Farbmessung wurde ein Konzept für eine einfache Farbumschlagsmessung entwickelt. Dieses basiert auf der Nutzung charakteristischer Wellenlängen im VIS-Bereich (z. B. 370, 420, 475, 570, 650 und 810 nm). Eine Absolutmessung der Farbe ist somit nicht möglich, aber auch zur Bestimmung einer Farbänderung nicht erforderlich. Der Messalgorithmus berechnet aus den dunkelstromkorrigierten Messwerten den Mittelwert und die Standardabweichung aufeinanderfolgender Messungen. Anhand letzterer wird ein Wert für Farbkonstanz bzw. -umschlag ermittelt und ausgegeben. Dies ermöglicht eine kalibrierfreie Überwachung des Farbumschlags. Falls gewünscht kann auch eine Kalibrierung auf Basis der Rohdaten mittels einer univariaten Kalibrierung erfolgen. Damit können Farbwechsel- und Ausspülprozesse sowie die Prozesskonstanz überwacht werden.

Bei den Versuchen wurden Farbwechsel verschiedener Farben und Deckungsgrade (deckend, transluzent) sowie Ausspülvorgänge betrachtet. Als Basispolymer dienten PP und verschiedene Farbmasterbatches auf PP-Basis. Da keine absoluten Farbwerte ermittelt wurden, war auch die chemische Zusammensetzung der Farbmasterbatches (Pigmente oder Farbstoffe) nicht relevant. Ebenso wurden Temperatur, Durchsatz und Drehzahl am Extruder auf Standardwerte eingestellt, da deren Einfluss auf die Einfärbung für diese in den Versuchsreihen nicht Ziel der Untersuchung war. Das Inline-Farbmesssystem von ColVisTec wurde zum Vergleich der spektralen Daten im VIS eingesetzt. Augenmerk lag auf der Bestimmung der Nachweisgrenzen beim Wechsel zweier sehr ähnlicher Farben oder sehr geringen Unterschieden im Deckungsgrad. Die Einflüsse der Parameter Temperatur und Druck wurden ebenfalls punktuell untersucht. Abhängig von der Prozesstemperatur wurde eine andere Farbe gemessen (Thermochromie). Zur Steigerung der Empfindlichkeit des Farbumschlages wurden die LEDs zu einem Rot-Grün-Blau (RGB)-Verbund gruppiert. Der Farbsensor zeigte insgesamt sehr gute Ergebnisse und eine ausreichend hohe Empfindlichkeit. Er eignet sich prinzipiell auch für die Detektion des Polymerabbaus farbloser Schmelzen (Gelbwert). In Abb. 6 ist die Benutzeroberfläche der Software des Farbumschlagsensors mit den RGB-Verläufen während eines Farbwechsels dargestellt.

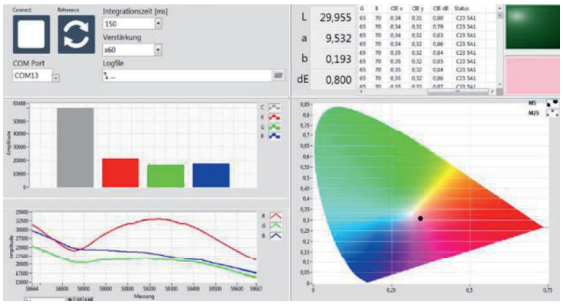


Abb. 6: Benutzeroberfläche der Software des Farbumschlagsensors während eines Farbwechsels

Die Einordnung in einen Farbraum ist dabei optional und erfordert die Berücksichtigung der Strahlcharakteristik der LEDs.

Korrelation der spektroskopischen Ergebnisse mit Labormessungen

Die Quantifizierung der Kenngröße Feuchte anhand der Laborversuche funktionierte gut und diente daher als Grundlage für die Kalibrierung. Eine Kalibrierung anhand der unterschiedlichen Feuchtegehalte im Extruder selbst ist allerdings nicht sinnvoll, da Extruder nicht perfekt dicht sind.

Die Versuche zeigten, dass mit denselben LEDs sowohl ein Feuchte- als auch ein Abbau-Sensor realisiert werden kann. Zum jetzigen Zeitpunkt lassen sich noch keine quantitativen Aussagen zum Polymerabbau treffen. Jedoch lässt sich ein Inline-Sensor für eine Gut/Schlecht-Aussage (Ampelsystem) realisieren.

Das gewählte Konzept für die Ermittlung des Farbumschlags erwies sich in den Extruderversuchen als praktikabel. Der Parameter Druck hatte keinen großen Einfluss auf die Messungen. Zudem wurde bei konstanten Temperaturen gemessen, da der Thermochromie-Effekt bei den Farben nicht einfach korrigiert werden kann.

Die jeweiligen Sensoren sind funktionsfähig, bedürfen aber noch weiterführender Tests und Anpassungen bis zur finalen Marktreife. Dabei ist zu prüfen, wie gängige Additive die gewählten für die Kalibrierung verwendeten Banden beeinflussen. Für die Funktionalität des Farbumschlagsensors ist der Additivgehalt nicht relevant.

Die angestrebte Umsetzung eines spezialisierten Sensorsystems zu einem günstigeren Preis als bei Vollspektrometer-Messsystemen war somit erfolgreich. Der Fokus der Systemplattform liegt derzeit auf der Extrusion und Compoundierung. Eine Erweiterung auf das Spritzgießen soll ebenfalls erfolgen.

Zusammenfassung

Zur Vereinfachung der Spektroskopie in der Prozesskontrolle wurde erfolgreich die modulare Systemplattform ExtruSens entwickelt. Diese ermöglicht die gezielte Bestimmung verschiedener Materialkenngrößen mittels kostengünstiger LED-Beleuchtung und Halbleiter-Sensorik direkt im Prozess. Hierzu wurden zwei Sensoren für grundlegende Funktionalitäten entwickelt: Ein Kombisensor für die Bestimmung des Feuchtegehaltes von Granulaten (oder Pulvern) und des hydrolytischen Polymerabbaus von Kunststoffschmelze sowie ein Sensor zur Überwachung von Farbveränderungen der Kunststoffschmelze. Die Sensoren liefern im Vergleich mit den Referenzverfahren gut reproduzierbare Ergebnisse und bieten noch Potential für weitere Entwicklungsstufen. Durch die spezialisierte Anwendung sind die Sensoren deutlich günstiger als klassische spektroskopische Lösungen verfügbar. Zudem ist auch die zukünftige Nutzung des Systems in weiteren Einsatzfeldern wie z. B. der Lebensmittel- und Pharmaindustrie denkbar.

Danksagung

Die Entwicklung des Prüfgeräts wurde über die Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF) im Rahmen des Programms zur Förderung des zentralen Innovationsprogramms Mittelstand (ZIM) vom Bundesministerium für Wirtschaft und Energie (BMWi) aufgrund eines Beschlusses des Deutschen Bundestages gefördert (ZF4026601RE5). Wir bedanken uns für die finanzielle Unterstützung.

Literaturnachweis

[1] I. Alig, D. Fischer, D. Lellinger, B. Steinhoff: Combination of NIR, Raman, Ultrasonic and Dielectric Spectroscopy for In-Line Monitoring of the Extrusion Process, *Macromol. Sy. 230*, 51-58 (2005); doi: 10.1002/masy.200551141

[2] T. Hochrein, I. Alig: Prozessmesstechnik in der Kunststoffaufbereitung, Würzburg: Vogel Buchverlag (2011)

[3] M. Hesse, H. Meier, B. Zeeh: Spektroskopische Methoden in der organischen Chemie, Stuttgart: Thieme-Verlag (2005)

[4] J. Workman, L. Weyer: Practical Guide to Interpretative Near Infrared Spectroscopy, New York: CRC Press (2007)

[5] M. Fukuda, M. Miyagawa, H. Kawai, N, Yagi, O. Kimura, T. Ohta: Fundamental Studies on the Interactions between Moisture and Textiles V. FT-IR Study on Moisture Sorption Isotherm of Nylon 6, *Polym. J. 19(7)*, 785-804 (1987); doi: 10.1295/polymj.19.785

[6] W. Camacho, A. Vallés-Lluch, A. Ribes-Greus, S. Karlsson: Determination of Moisture Content in Nylon 6,6 by Near-Infrared Spectroscopy and

Chemometrics, *J. Appl. Polym. Sci. 87*, 2165-2170 (2003); doi: 10.1002/app.11595
[7] O. R. Dumitrescu, D. C. Baker, G. M. Foster, K. E. Evans: Near Infrared Spectroscopy for In-line Monitoring during Injection Moulding, *Polym. Test. 24*, 367-375 (2005); doi: 10.1016/j.polymertesting.2004.10.003
[8] W. Becker, E. Inone-Kauffmann, W. Eckl, N. Eisenreich: Near Infrared Spectroscopy for In-Line Control of Biopolymer Processing, *AIP Conf. Proc.* (2007)
[9] H. Shinzawa, M. Nishida, T. Tanaka, W. Kanematsu: Accelerated Weathering-Induced Degradation of Poly(Lactic Acid) Fiber Studied by Near-Infrared (NIR) Hyperspectral Imaging, *Appl. Spectrosc. 66(4)*, 470-474 (2012); doi: 10.1366/11-06540
[10] Y. Wang, B. Steinhoff, C. Brinkmann, I. Alig: In-line Monitoring of the Thermal Degradation of Poly(L-lactic acid) during Melt Extrusion by UV-vis Spectroscopy, *Polymer 49*, 2008, 1257-1265 (2008); doi: 10.1016/j.polymer.2008.01.010
[11] K. A. Bakeev: Near-Infrared Spectroscopy as a Process Analytical Tool, Part II: At-line and On-line Applications and Implementation Strategies, *Spectroscopy 19(1)*, 39-42 (2004)
[12] Endress+Hauser GmbH+Co. KG: Prozessphotometer, Kompetenz in der Flüssigkeitsanalyse (2016)
[13] R. Szedlak, A. Harrer, M. Holzbauer, B. Schwarz, J. P. Waclawek, D. MacFarland, T. Zederbauer, H. Detz, A. M. Andrews, W. Schrenk, B. Lendl, G. Strasser: Remote Sensing with Commutable Monolithic Laser and Detector, *ACS Photonics 2016(3)*, 1794−1798 (2016); doi: 10.1021/acsphotonics.6b00603
[14] M. Bauer, D. Beilke, H. Nitsche, N. Halmen, C. Kugler, T. Hochrein, M. Bastian, A. Bogomolov: Spektroskopische Sensoren zur Überwachung von Extrusionsprozessen, Qualitätsgipfel Kunststoff 2017, Würzburg (2017)
[15] H. Nitsche, N. Halmen: Neues preiswertes und modulares Messsystem für die Qualitätssicherung in der Kunststoffindustrie, Extrusion, 07 (2018)
[16] BASF SE: Ultramid B27E, Materialdatenblatt (2008)
[17] SABIC: SABIC PET BC112, Materialdatenblatt (2015)
[18] Aboni GmbH: HydroTracer FMX/FLV, Wassergehaltsbestimmung in Feststoffen (2017)

DOI 10.5162/sensoren2019/P2.5

Detektion von Kunststoffen in Wasser mithilfe von Fluoreszenz

Maximilian Wohlschläger[1], *Martin Versen*[1]
[1]*Technische Hochschule Rosenheim, Hochschulstraße 1, 83024 Rosenheim, Deutschland*

Zusammenfassung

Die wohl größte aktuelle Umweltverschmutzung mit Kunststoffen findet in Ozeanen, Seen und Flüssen statt. Bis jetzt gibt es kein etabliertes Verfahren, mit dem Kunststoffe in Umwelt- bzw. Wassermatrizen detektiert werden können. Deshalb wird mit einem Verfahren, welches auf dem Prinzip der Fluoreszenz basiert, untersucht, ob unter Laborbedingungen Kunststoffe in Wasser detektiert und identifiziert werden können. Anhand eines mathematischen Modells, einer Simulation und mit Hilfe von optischen Experimenten wird gezeigt, dass der Nachweis und auch eine Identifikation von Kunststoffen in Abhängigkeit der Wassertiefe, mit diesem Verfahren möglich ist.

Keywords: Kunststoffe, Umwelt, Fluoreszenzmesstechnik, Fluoreszenz

1. Einleitung

Weltweit sind Kunststoffe mittlerweile in allen Bereichen der Industrie als Aufbewahrungs-, Schutz- und Verpackungswerkstoff unverzichtbar. Gleichzeitig steigt die Anzahl verschiedener Kunststoffsorten stetig an und die Verfahren zum Kunststoffrecycling werden immer weiter verbessert, um eine möglichst hohe Qualität bei der Wiederverwendung zu gewährleisten.

Kunststoffe können jedoch nur dann wiederverwendet werden, wenn diese ordnungsgemäß entsorgt werden. Leider ist dies aber aufgrund von z.B. mangelnder Aufklärung oder fehlenden Müllentsorgungssystemen nicht immer der Fall. Daher wächst die Umweltbelastung derzeit ins Unermessliche.

So zeigt beispielsweise eine Studie aus Hongkong, dass 27.909 Kunststoffpartikel in einem Wasservolumen von 100m³ in Küstenregionen vor Hongkong zu finden sind. Betrachtet man Studien aus Neuseeland in der Küstenregion von Queen Charlotte Sound, erscheinen die Werte aus Hongkong harmlos, denn dort wurden 763.000 Partikel in 100 m³ Wasser festgestellt. In Ostchina wurden wiederrum nur 17 Partikel pro 100 m³ und im Küstenraum von Portugal 4 Partikel pro 100 m³ Wasser nachgewiesen (siehe [1] und die darin enthaltenen Referenzen).

Die Verschmutzung der Umwelt ist allerdings nur ein Aspekt dieses Problems, denn Meereslebewesen wie Fische nehmen die im Wasser enthaltenen Partikel auf, wie Studien malaysischer und kanadischer Forscher zeigen [2, 3]. Durch den Verzehr der kontaminierten Fische kann dies auch zum Gesundheitsrisiko für den Menschen führen.

Zum Nachweis der Kunststoffe in Wasser und zur Analyse werden die Kunststoffe aus dem Wasser manuell ausgesiebt, ins Labor transportiert und dort mit spektrometrischen Verfahren wie zum Beispiel der Infrarot oder Raman Spektroskopie untersucht [4]. Es gibt keine Methode, welche den Kunststoff vor Ort in einer Wassermatrix detektieren und identifizieren kann. Um diese Lücke zu schließen, wird ein optisches Verfahren gesucht, dass Kunststoff in Wasser erkennt und identifiziert.

Derzeit gibt es zwei unter Laborbedingungen getestete Verfahren, die Kunststoffe detektieren und voneinander unterscheiden können.

Eines davon ist die Trennung der Kunststoffe aufgrund ihrer spezifischen Fluoreszenzabklingzeiten. Eine Probe wird dabei mit einem Laserlichtimpuls angeregt und es wird die Zeit gemessen, wie lange der Kunststoff nachleuchtet. Die sich daraus ergebende Fluoreszenzabklingzeit ist für jeden Kunststoff spezifisch und lässt eine Unterscheidung zu [5, 6]. Da die Abklingzeiten in der Größenordnung von Nanosekunden liegen, sind für die Realisierung eines Messaufbaus synchronisationsschnelle und folglich teure Geräte erforderlich.

Ein zweites Verfahren betrachtet die spezifische Fluoreszenz-Quanteneffizienz. Dabei wird der Kunststoff mithilfe einer

Lichtquelle angeregt und das Fluoreszenzsignal gemessen. Durch spektrale Rechenoperationen kann nach der Messung die Anzahl der Fluoreszenz-Photonen berechnet werden. Ist das Absorptionsspektrum des Kunststoffes bekannt, kann die Anzahl der absorbierten Photonen ermittelt werden. Die spezifische Quanteneffizienz wird dann berechnet aus dem Verhältnis der Anzahl an Fluoreszenz-Photonen zur Anzahl der absorbierten Photonen. Darüber hinaus kann bei diesem Verfahren eine kunststoffspezifische Apparatekonstante gebildet werden, die definiert ist als das Verhältnis aus der Anzahl der von der Detektionseinheit gezählten Photonen zur Anzahl der absorbierten Photonen [7, 8].
Bei beiden Verfahren werden Anregungswellenlängen im nahen UV/VIS-Bereich (ultravioletter visueller- Bereich) des elektromagnetischen Spektrums eingesetzt. Vorteilhaft ist, dass Licht in diesem Wellenlängenbereich eine hohe Eindringtiefe in Wasser hat, wie Abb. 1 zeigt (siehe Wellenlängenbereich von 380-400nm).

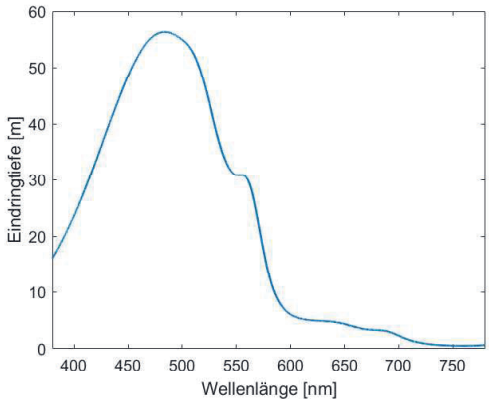


Abb. 1: Darstellung der Lichteidringtiefe in Wasser, in Abhängigkeit der Wellenlänge, berechnet mit dem Absorptionskoeffizienten aus [9].

Zur Berechnung wird der Absorptionskoeffizienten von reinem Wasser aus [9] verwendet, welcher keine Schwebepartikel, Verunreinigungen etc. berücksichtigt.
Im Folgenden wird untersucht, ob Kunststoffe aufgrund ihrer spezifischen Quanteneffizienz bzw. Apparatekonstante in Wasser detektiert und identifiziert werden können. Mit einer Simulation wird die Wassertiefe bestimmt, bis zu der ein Kunststoff auffindbar ist. Darüber hinaus soll überprüft werden, ob ein Zusammenhang aus der Theorie und Realität hergestellt werden kann.
Der Beitrag gliedert sich in die Weiterentwicklung und Modifikation des theoretischen und heuristischen Modells aus [7, 8], dem Simulationsmodell und den Simulationsergebnissen, sowie den experimentellen Untersuchungen und den daraus resultierenden Ergebnissen. Am Ende wird eine Zusammenfassung einen Überblick über die Ergebnisse und den Zusammenhang von Theorie und Realität geben.

2. Theoretisches Modell zur Detektion des Kunststoffes in Wasser

In diesem Kapitel wird in den Stand der Technik eingeführt und die Modifikation des theoretischen Modells erläutert.

2.1. Stand der Technik

Das Modell zur Detektion und Identifikation von Kunststoffen wurde bereits in [7, 8] vorgestellt (siehe modifiziertes Modell Abb. 2). Ein Lichtstrahl trifft mit einer definierten Anzahl an Photonen (2) auf die Kunststoffprobe (3). Die Absorption eines Photons hat zur Folge, dass ein Elektron von einem niedrigen Energieniveau auf ein höheres Energieniveau gehoben wird. Da das Elektron bestrebt ist, in seinen energetischen Grundzustand zurückzukehren, gibt es seine Energie einerseits als nicht strahlende Energie und andererseits als strahlende Energie, einem Fluoreszenz-Photon, ab. Durch den Energieverlust, ist das Fluoreszenzphoton niederenergetischer und hat eine höhere Wellenlänge. Diese Wellenlängenverschiebung zwischen Anregungs- und Fluoreszenzlicht wird als Stokes-Verschiebung bezeichnet.
Die anregende Lichtquelle (1) ist dabei eine UV-LED mit einer CW (Zentralen Wellenlänge) von 390 nm. Die exakte Strahlführung für (2), (4) und (5) wird mit einem Mikroskop gewährleistet. Als Detektionseinheit dient ein Minispektrometer. Mit diesem werden wellenlängenabhängige Aussagen getroffen, wie viele Photonen abhängig von der Wellenlänge detektiert werden.
Die Berechnung der Quanteneffizienz von den drei Kunststoffproben ist in [8] gezeigt und führt unter Berücksichtigung der Gaußschen Normalverteilung zu den in Tab. 1 eingetragenen Ergebnissen.

Tab. 1: Berechnete Werte der Quanteneffizienz QE und Apparatekonstante QE_A für die Kunststoffe PA (Polyamid), PE (Polyethylen) und PET (Polyethylenterephthalat) [8]

Polymer	*QE [%]*	QE_A
PA	2,3±0,1	$2{,}2\cdot10^{-8}\pm1\cdot10^{-9}$
PE	3,0±0,2	$2{,}9\cdot10^{-8}\pm1{,}9\cdot10^{-9}$
PET	3,4±0,2	$3{,}4\cdot10^{-8}\pm1{,}9\cdot10^{-9}$

2.2. Theoretisches Modell

Um zu untersuchen, ob eine Detektion und Identifikation im Wasser möglich ist, wird das aus [8] bekannte Modell modifiziert und erweitert (siehe Abb. 2). Ein entscheidender Faktor für die Detektions- bzw. Identifikationsgrenze ist die Wasserschicht (7) die sich auf der Kunststoffprobe (3) befindet.

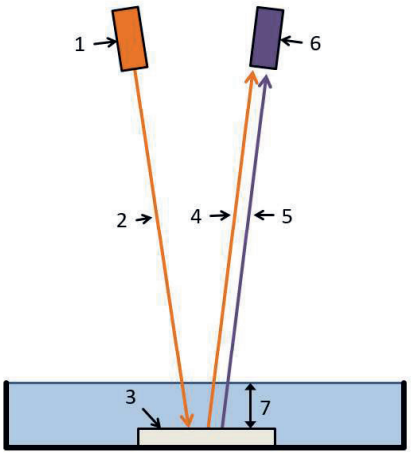


Abb. 2: Theoretisches Modell zur Detektion und Identifikation in Wasser mit Lichtquelle1, dem anregenden Lichtstrahl 2, der Kunststoffprobe 3, dem reflektierten Lichtanteil 4, dem fluoreszierenden Lichtanteil 5, der Detektionseinheit 6 und der Wasserhöhe über der Kunststoffprobe 7.

Um die exponentielle Absorption von Wasser zu berücksichtigen, wird der Absorptionskoeffizient von Waser nach [9] in das theoretische Modell aus [8] aufgenommen. Dabei ergibt sich die absorbierte Photonenanzahl N_A nach Gl. (1):

$$N_A = N_E \cdot (1 - e^{-\alpha_P \cdot d_P}) \cdot e^{-\alpha_W \cdot d_W} \quad (1)$$

wobei N_E die auf den Kunststoff treffenden Photonen (in Abb. 2 siehe Nummer 2), α_P der Absorptionskoeffizient, d_P die Dicke des Kunststoffes, α_W der Absorptionskoeffizient von Wasser und d_W die Höhe der Wasserschicht auf der Kunststoffprobe sind.

Mit der Anzahl der absorbierten Photonen N_A und der Quanteneffizienz QE kann dann die Anzahl der Fluoreszenz-Photonen N_F nach Gl. (2) berechnet werden.

$$N_F = N_A \cdot QE \quad (2)$$

Aus den Fluoreszenz-Photonen wird dann mit Gl. (3) die Anzahl der detektierten Photonen N_D unter Verwendung des Transmissionsgrades des Mikroskops T_M bestimmt.

$$N_D = N_F \cdot e^{-\alpha_W \cdot d_W} \cdot T_M \quad (3)$$

Die Anzahl der detektierten Photonen ist eine der beiden Größen, die Aufschluss über eine mögliche Detektierbarkeit und Identifizierbarkeit liefern soll. Als zweites kann mit ihr die kunststoffspezifische Apparatekonstante QE_A nach Gl. (4) berechnet werden.

$$QE_A = N_D / N_A \quad (4)$$

Beide Größen können in Abhängigkeit der Wasserschicht auf dem Kunststoff dargestellt werden, was nachfolgende Simulation zeigt.

3. Simulation der Identifikation in Wasser

Zur Simulation wird ein Programm in der Berechnungssoftware MATLAB verwendet. Im Folgenden werden die Funktionsweise des Programmes und die resultierenden Simulationsergebnisse beschrieben

3.1. Simulationsmodell

Das erstellte Simulationsprogramm, arbeitet nach dem in Abb. 3 gezeigtem Ablaufplan.

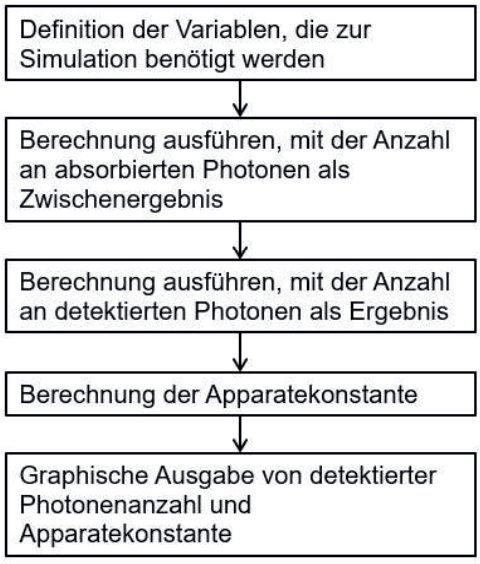


Abb. 3: Programmablaufplan des zur Simulation verwendeten Berechnungsprogramms.

Nachdem in einem ersten Schritt die für die Simulation benötigten Variablen definiert sind, wird die Berechnung der Anzahl der absorbierten Photonen ausgeführt und als Zwischenergebnis gespeichert. Mit dem Zwischenergebnis wird im vierten Schritt zusammen mit der Quanteneffizienz die Anzahl der Fluoreszenz-Photonen berechnet. Im dritten Schritt wird mit der Anzahl der Fluoreszenz-Photonen die Anzahl der detektierten Photonen

ermittelt. Zusätzlich wird aus der Anzahl der detektierten Photonen und der Anzahl der absorbierten Photonen die spezifische Apparatekonstante bestimmt.
Die berechneten Werte werden zugehörig zu ihrer Wasserschichtdicke in einer Matrix gespeichert, d.h. jeder Wasserhöhenwert hat einen zugehörigen Wert für die detektierte Photonenanzahl und Apparatekonstante.
Am Ende der Simulation werden zwei Diagramme ausgegeben, wobei eines davon die Anzahl der detektierten Photonen und das andere die Apparatekonstante als Funktion der Wasserhöhe auf dem Kunststoff zeigt.
Für den Graphen der Apparatekonstante wird in bestimmten Abständen der zugehörige Fehlerbalken eingetragen, der der Gauß'schen Standardabweichung der experimentell bestimmten Quanteneffizienz entspricht. Bei einer Überschneidung von zwei Fehlerbalken, ist es unwahrscheinlich, die Kunststoffe anhand ihrer spezifischen Apparatekonstante zuverlässig zu unterscheiden.

3.2. Definition der Variablen

Zur Simulation wurde die bereits in [7, 8] verwendeten Werte zur Berechnung der Quanteneffizienz und Apparatekonstante verwendet. Benötigt werden noch die in Tab. 1 eingetragenen Quanteneffizienzen der drei Kunststoffproben und die Wasserhöhe. Für diese wurde ein Wertebereich von 0mm bis 50000mm in 1,0mm Schritten definiert.

3.3. Simulationsergebnisse

Die Ergebnisse der Simulation sind in Abb. 4 und Abb. 5 dargestellt, wobei Abb. 4 die Anzahl der detektierten Photonen Abb. 5 die Apparatekonstante als Funktion der Wasserschicht auf dem Kunststoff zeigt.

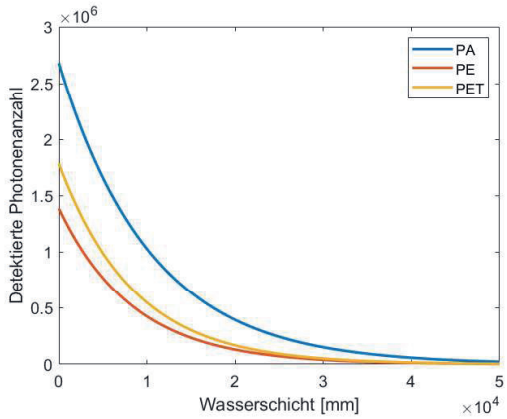


Abb. 4: Anzahl der detektierten Photonen in Abhängigkeit der Wasserhöhe, im Bereich von 0mm bis 50000mm für die Kunststoffe PA, PE und PET

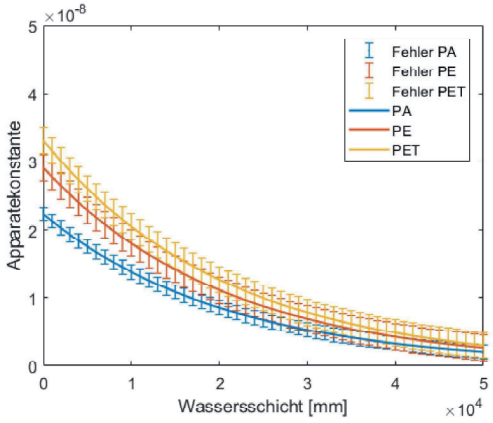


Abb. 5: Apparatekonstante mit Fehlerbalken in Abhängigkeit der Wasserhöhe, im Bereich von 0mm bis 50000mm für die Kunststoffe PA, PE und PET

Die Identifikation mit der Apparatekonstante ist allerdings bei 20m nicht mehr möglich. Um eine simulative Aussage über die Identifizierbarkeit zu treffen, wird deshalb der Bereich zwischen 0mm und 20000mm (20m) vergrößert.
Die Vergrößerung ist in Abb. 6 dargestellt. Die Fehlerbalken von PA und PE überlappen bis zu einer Wassertiefe von ca. 18m nicht. Eine Unterscheidung zwischen PA und PET ist sogar in tieferen Gewässern möglich. Folglich kann der Kunststoff PA mit hoher Wahrscheinlichkeit bis zur Wassertiefe von ungefähr 18m als Polyamid identifiziert werden ohne mit PE oder PET verwechselt zu werden.
Zwischen den Kunststoffen PE und PET ist keine zuverlässige Unterscheidung bei über 2m möglich, da dort eine Überschneidung der Fehlerbalken stattfindet. Folglich können diese zwei Kunststoffe nur bis zu einer Wassertiefe von kleiner 2m relativ sicher identifiziert und unterschieden werden.

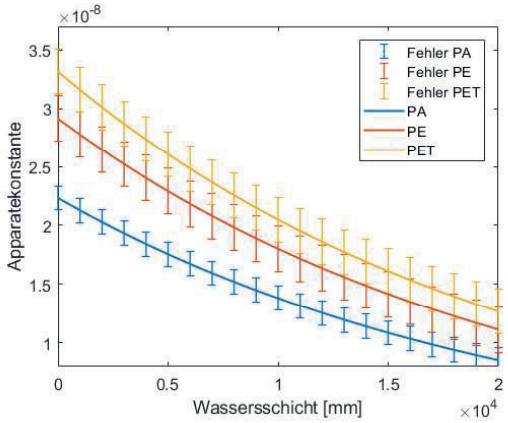


Abb. 6: Apparatekonstante mit Fehlerbalken in Abhängigkeit der Wasserhöhe, im Bereich von 0mm bis 20000mm für die Kunststoffe PA, PE und PET

4. Experimente zur Detektion und Identifikation in Wasser

Um das simulierte Ergebnis zu überprüfen, werden die Kunststoffe in Wasser präpariert und nach den in [7, 8] gezeigten Methoden ausgewertet. Im Folgenden wird die Probenvorbereitung, der für die Experimente verwendete Versuchsaufbau, das Vorgehen bei den Untersuchungen und die die daraus resultierenden Ergebnisse gezeigt.

4.1. Probenvorbereitung

Die in den Versuchen verwendeten Proben sind Folien der Kunststoffe PA, PE und PET, welche an der Technischen Hochschule Rosenheim hergestellt worden sind.
Die Folien werden in Stücken von 2cm x 2cm in einer Petrischale mithilfe von Klebestreifen befestigt (siehe Abb. 7). Dadurch wird verhindert, dass die Kunststoffe aufschwimmen, da alle eine geringere Dichte als Wasser aufweisen.

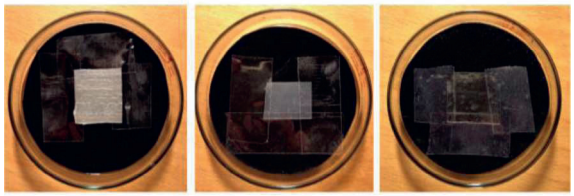


Abb. 7: Kunststoffproben PA (Links), PE (Mitte) und PET (Rechts) in einer Petrischale, mit Klebeband gegen Aufschwimmen befestigt.

4.2. Experimenteller Aufbau

Als experimenteller Aufbau wurde der in [7, 8] bereits erklärte Aufbau verwendet. Dieser besteht aus einer Lichtquelle einer Probe-Station und einem Minispektrometer.
Als Lichtquelle dient eine NVUS233B UV-LED von Nichia [10]. Wie bereits erwähnt, besitzt die LED eine zentrale Wellenlänge von 390 nm, was optimal für die Anregung der Kunststoff-proben ist.
An der Probe-Station EPS150FA von Cascade Microtech können die Komponenten, also Lichtquelle und Minispektrometer, befestigt werden. In der Probe-Station ist ein Mikroskop integriert, an welchem eine zwei, zehn oder zwanzigfache Vergrößerung eingestellt werden kann. Dies ist insbesondere vorteilhaft, um Photonen aus extrem kleinen Proben aufzuzeichnen [11].
Das Minispektrometer C10083CAH von Hamamatsu fungiert als spektrale Detektions-einheit. Es kann mit einer Auflösung von 1nm den Wellenlängenbereich von 300 bis 1080 nm aufzeichnen [12].

4.3. Vorgehen bei den Experimenten

Bei den Experimenten wurde nach dem in Abb. 8 dargestellten Ablaufplan vorgegangen. Anzumerken ist, dass die Petrischale eine maximale Füllhöhe von 8mm hat und daher keine Versuche in tieferem Wasser durchgeführt werden können.

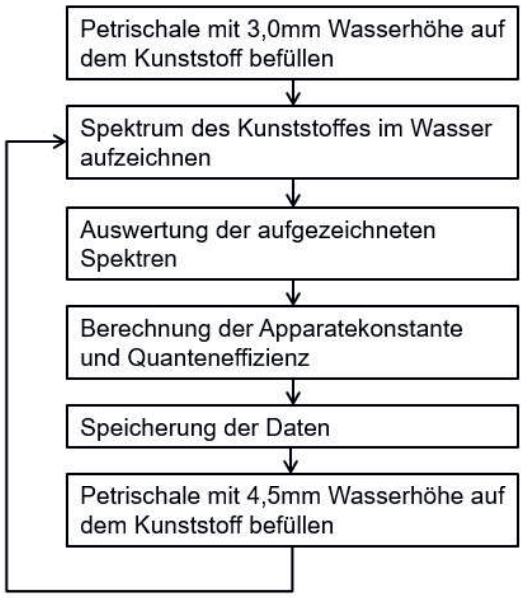


Abb. 8: Versuchsablaufplan zur Vermessung der Kunststoffproben in 3,0mm, 6,0mm und 8,0mm Wasser

Die Kunststoffproben sind im Trockenen und in den Wassertiefen von 3,0mm, 4,5mm, 6,0mm, 7,0 mm und 8,0mm vermessen worden.
Die Versuche im Trockenen werden wie in [8] beschrieben durchgeführt und zeigen im Rahmen des Fehlers ähnliche Werte für die Quanteneffizienzen und Apparatekonstante. Zurückgeführt werden diese Änderungen auf statistische Einflüsse wie beispielsweise eine nicht kontrollierte Raumtemperatur bei den Versuchen. Es ergeben sich die in Tab. 2 eingetragenen Werte für die Quanteneffizienz und Apparatekonstante.

Tab. 2: Experimentell bestimmte Werte der Quanteneffizienz QE und Apparatekonstante QE_A für die Kunststoffe PA (Polyamid), PE (Polyethylen) und PET (Polyethylenterephthalat)

	QE [%]	*QE_A*
PA	1,97±0,08	$1{,}91\cdot10^{-8}\pm8{,}16\cdot10^{-10}$
PE	3,65±0,03	$3{,}52\cdot10^{-8}\pm1{,}84\cdot10^{-9}$
PET	4,36±0,03	$4{,}21\cdot10^{-8}\pm2{,}99\cdot10^{-10}$

Nach der Bestimmung der Quanteneffizienz im Trockenen, wird die Petrischale mit einem Wasserstand von 3,0mm auf dem Kunststoff befüllt und unter das Mikroskop gestellt, um das Spektrum aufzunehmen. Die Kunststoffprobe wird an 20 unterschiedlichen Positionen vermessen. Die dabei aufgenommenen Spektren werden in Abhängigkeit der Wellenlänge in eine Tabelle übertragen und mit MATLAB ausgewertet. Anschließend werden die Mittelwerte und die Standardabweichungen unter Annahme einer Gauß'schen Normalverteilung berechnet.
Nach der Auswertung mit einer Wasserhöhe von 3,0mm auf dem Kunststoff, wird die Petrischale weiter bis auf 4,5mm mit Wasser befüllt. Anschließend wird der beschriebene Vorgang einschließlich der Auswertung der Daten wiederholt.
Wenn die Untersuchungen mit einer Wasserschicht von 4,5mm abgeschlossen sind, wird der Wasserstand weiter auf 6,0mm, 7,0mm und 8,0mm erhöht und die Messungen werden wiederholt.
Wenn die erste Kunststoffprobe in den fünf Wassertiefen vermessen ist, wird der Versuch mit den anderen zwei Kunststoffproben nach dem gleichen Schema durchgeführt.

4.4. Experimentelle Ergebnisse

Die Auswertung der Daten zeigt, dass sich sowohl die Apparatekonstante als auch die Quanteneffizienz mit steigender Wasserhöhe ändert.
In Abb. 9 ist die Quanteneffizienz als Funktion der Wasserschicht für Messungen im Trockenen, bei 3,0mm, bei 4,5mm, bei 6,0mm, bei 7,0mm und bei 8,0mm Wasserhöhe auf dem Kunststoff dargestellt. Die Quanteneffizienz nimmt mit zunehmendem Wasserstand exponentiell ab.

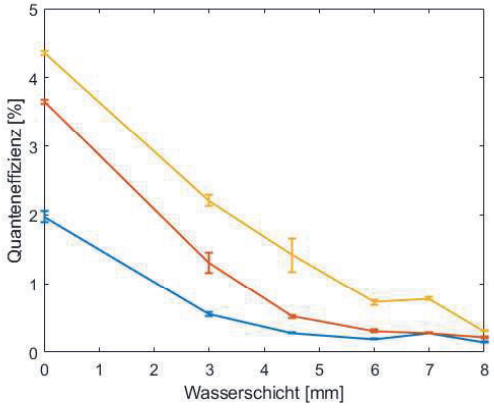


Abb. 9: Darstellung der Quanteneffizienz als Funktion der Wasserschicht für die drei verwendeten Kunststoffproben PA, PE und PET

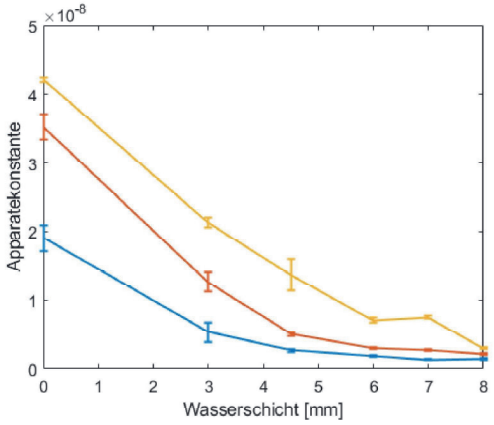


Abb. 10: Darstellung der Apparatekonstante in Abhängigkeit der Wasserschicht für die drei verwendeten Kunststoffproben PA, PE und PET

Für die Apparatekonstante in Abhängigkeit der Wasserschicht ergibt sich das in Abb. 10 gezeigte Diagramm.
Wie bei der Quanteneffizienz in Abb. 9 ist auch hier ein exponentieller Abfall zu sehen, welcher erwartet wird. Allerdings ist dieser deutlich höher, als erwartet, was im Folgenden näher betrachtet wird.

4.5. Auswertung der experimentellen Ergebnisse

Um eine Aussage über die experimentellen Ergebnisse zu treffen, werden Ausgleichsgraphen durch die gemessenen Punkte gelegt. Für die sinkenden Quanteneffizienzen ergeben sich dabei die in Abb. 11 dargestellten Graphen.
Die Graphen werden mit einem exponentiellem Fit der in Gl. (5) gezeigten Form modelliert.

$$QE = a * \exp(b * d_W) \tag{5}$$

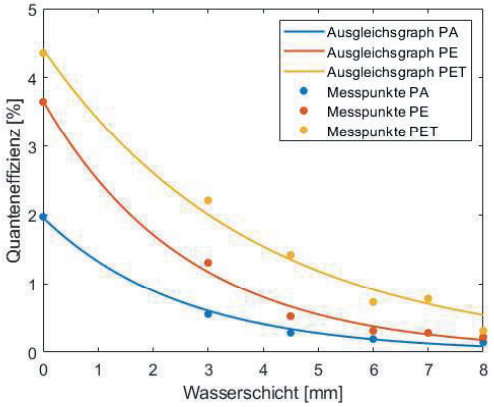


Abb. 11: Darstellung des Ausgleichsgraphen für die sinkende Quanteneffizienz der Kunststoffproben PA, PE und PET

Tab. 3: Ermittelte Koeffizienten a und b der exponentiellen Ausgleichgraphen für die Quanteneffizienz

	a [%]	*b [1/mm]*
PA	1,96	-0,39
PE	3,66	-0,38
PET	4,41	-0,27

Dabei ergeben sich für die Parameter *a* und *b* die in Tab. 3 eingetragenen Werte für die verschiedenen Kunststoffproben.
Der Koeffizient *a* gibt den kunststoffspezifischen Wert und *b* den vom Wasser verursachten Verlust an.
Die gleiche Auswertung wird für die Apparatekonstante durchgeführt (siehe Abb. 12).

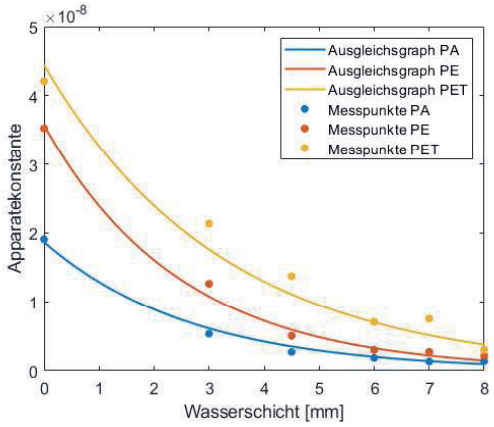


Abb. 12: Darstellung des Ausgleichsgraphen für die sinkende Apparatekonstante der Kunststoffproben PA, PE und PET

Die Koeffizienten für die Apparatekonstante sind in Tab. 4 eingetragen.
In beiden Fällen sind die Koeffizienten von *b* erwartungsgemäß nahezu gleich, da sie die Absorbtionseigenschaften des Wassers darstellen.
Ein Vergleich der Werte für die ermittelten Quanteneffizienzen nach den Tab. 2 und 3 zeigen ebenfalls erwartungsgemäß keine signifikante Veränderung, da die Quanteneffizienz ein kunststoffspezifischer Wert ist, welcher nicht von Apparaten oder der Wasserschicht etc. abhängt.

Tab. 4: Ermittelte Koeffizienten a und b der exponentiellen Ausgleichgraphen für die Apparatekonstante

	a [%]	*b [1/mm]*
PA	$1{,}87 \cdot 10^{-8}$	-0,37
PE	$3{,}57 \cdot 10^{-8}$	-0,40
PET	$4{,}45 \cdot 10^{-8}$	-0,31

Quantitativ ergibt sich eine Veränderung des experimentell ermittelten Absorptionskoeffizienten *b* im Vergleich zum Literaturwert aus [9]. Dieser ist um vier Größenordnungen niedriger als der Wert, der sich bei der experimentellen Auswertung zu ca. 0,35·1/mm ergibt.
Der Unterschied der beiden Werte ergibt sich daraus, dass der Absorptionskoeffizient des Referenzwertes aus [9] ohne den Einfluss von Streuung, Beugung und Brechung ermittelt worden ist. Im hier beschriebenen Aufbau sind diese Effekte enthalten und werden mit einem effektiven Absorptionskoeffizienten als Ersatzgröße berücksichtigt, die den Strahlungsverlust durch Absorption, Beugung, Streuung und Brechung von Wasser einschließt. Mit dieser Ersatzgröße gelingt die qualitative und quantitative Beschreibung der Experimente.

5. Zusammenfassung der Ergebnisse

Zusammenfassend betrachtet haben die simulativen und experimentellen Untersuchungen gezeigt, dass Kunststoffe in Wasser detektiert und identifiziert werden können.
Das aus [7, 8] bereits bekannte Verfahren wurde mit einem Absorptionskoeffizienten für Wasser so modifiziert, dass eine Erkennung und Charakterisierung des Kunststoffes in einer Wasserschicht ermöglicht wird.
Die Auswertung experimenteller Daten zeigt übereinstimmende Ergebnisse mit den in [7, 8] gewonnenen Quanteneffizienzen. Darüber hinaus wird ein exponentieller Abfall der Quanteneffizienz in Abhängigkeit der Wasserschichtdicke beobachtet. Der exponentielle Abfall wird qualitativ erwartet, jedoch ist der experimentell gewonnene Absorptionskoeffizient vier Größenordnungen größer ist als der Referenzwert aus [9]. Grund für die Abweichung ist, dass bei dem Referenz-Absorptionskoeffizienten Effekte wie Streuung, Brechung und Beugung unberücksichtigt geblieben sind.
Für den exponentiellen Abfall wird eine Ersatzgröße gebildet, die diese Einflüsse berücksichtigt und in künftigen Untersuchungen statt des Literaturwertes verwendet wird.
Die Experimente legen nahe, dass eine Identifikation nur in geringen Wassertiefen möglich ist. Eine Detektion ohne Unterscheidung der Kunststoffe in Wasser ist hingegen auch in größeren Tiefen möglich.
Zukünftige Untersuchungen sollen zeigen, ob das Detektionsverfahren auch zum Nachweis von Mikroplastik in Wasser eingesetzt werden kann.

Literaturnachweis

[1] Y. Tsang, C. Mak, C. Liebich, S. Lam, E. Sze, K. Chan, *Microplastic pollution in the marine waters and sediments of Hong Kong*, Marine Pollution Bulletin 115, 20-28 (2017): doi: 10.1016/j.marpolbul.2016.11.003.

[2] A. Karami, A. Golieskardi, Y. Ho, V. Larat, B. Salamatina, *Microplastics in eviscerated flesh and excised organs of dried fish*, Scientific Reports 7, 5473, 1-6 (2017); doi: 10.1038/s41598-017-05828-6

[3] S. Campbell, P. Williamson, B. Hall, *Microplastics in the gastrointestinal tracts of fish and the water from an urban prairie creek*, Facets Journal 2, 395-409 (2017); doi: 10.1139/facets-2017-0008

[4] A. Silva, A. Bastos, C. Justino, J. Costa, A. Duarte, T. Rocha-Santos, *Microplastics in the environment: Challenges in analytical chemistry - A review*, Analytica Chimica Acta 1017, 1-19 (2018); doi: 10.1016/j.aca.2018.02.043

[5] H. Langhals, D. Zgela, T. Schlücker, *High Performance Recycling of Polymers by Means of Their Fluorescence Lifetimes*, Green and Sustainable Chemistry 4, 144-150 (2014); doi: 10.4236/gsc.2014.43019

[6] H. Langhals, D. Zgela, T. Schlücker, *Improved High Perfomance Recycling of Polymers by Means of Bi-Exponential Analysis of Their Fluorescence Lifetimes*, Green and Sustainable Chemistry 5, 92-100 (2015); doi: 10.4236/gsc.2015.52012

[7] M. Wohlschläger, M. Versen, *Bestimmung einer Quanteneffizienz für das automatisierte Sortieren von Kunststoffen durch Fluoreszenz*, 16. AALE Konferenz, VDE Verlag Berlin, 179-187 2019 (2019); ISBN: 978-3-8007-4860-0

[8] M. Wohlschläger, M. Versen, H. Langhals: *An apparatus specific quantum efficiency approach for sorting of plastics*, IEEE Sensor Applications Symposium (SAS), Sophia Antipolis, 2019 (in print).

[9] G. Hale, R. Querry: *Optical Constants of Water in the 200-nm to 200-µm Wavelength Region*, Applied Optics12, 3, 555-563 (1973); doi: 10.1364/AO.12.000555

[10] Nichia*, NVSU233B*, Datasheet, 2018.

[11] Cascade Microtech, *EPS150FA, A dedicated 150mm manual probing solution for failure analysis*, Datasheet, 2018

[12] Hamamatsu, *Mini spectrometer C10083CAH*, Datasheet, 2018.

DOI 10.5162/sensoren2019/P2.6

Analysis of Photoelastic Properties of Monocrystalline Silicon

Markus Stoehr[1], Gerald Gerlach[2], Thomas Härtling[2], Stephan Schoenfelder[1]
[1]Leipzig University of Applied Sciences, Faculty of Engineering,
Karl-Liebknecht-Str. 134, 04277 Leipzig
[2]TU Dresden, Faculty of Electrical and Computer Engineering, Helmholtzstr. 18, 01069 Dresden

Abtract

Photoelasticity could be a useful measurement tool for non-destructive, contactless determination of mechanical stresses or strains in production of silicon wafers. It describes a change in indices of refraction of materials when stressed. Since silicon has a diamond lattice structure, the stress-dependent change in refractive indices varies with loading direction. In this work an anisotropic stress-optic law is derived and compared to existing models for photoelasticity in silicon from literature.

Keywords: Photoelasticity, Monocrystalline Silicon, Infrared, Non-destructive Testing

Introduction

Photoelasticity is a contactless, non-destructive method to optically measure mechanical stresses or strains. Therefore, it offers an interesting in-line measurement method to determine and characterise transparent materials. Currently, it is employed for the production of glasses, but it could also be useful in the production process of silicon wafers. Though, while glass has an isotropic structure, silicon has an inherent mechanical and photoelastical anisotropy because of its lattice structure. Therefore, a corresponding model is required to determine the photoelastical properties of silicon.

Photoelasticity is based on birefringence caused by mechanical stress (respectively strain). Birefringence describes the property of materials to split an incident electro-magnetic wave into two refracted waves instead of one. These two refracted waves of light show different coefficients of refraction and different states of polarisation [1], [2]. A mechanical stress applied to a material susceptible to birefringence results in a change in the difference of the two states of polarisation and, hence, a change in the indices of refraction n_1 and n_2:

$$\Delta n = n_1 - n_2. \tag{1}$$

This change can be measured with a polariscope by measuring the phase difference δ between the two refracted light waves that increases with the thickness t of the material:

$$\delta = t\,(n_1 - n_2). \tag{2}$$

For mechanically and photoelastically isotropic materials the phase difference is proportional to the difference in the first and second principal stress (σ_I, σ_{II}) for plane stress conditions:

$$\delta = \mathrm{C\,t}\,(\sigma_\mathrm{I} - \sigma_\mathrm{II}). \tag{3}$$

The stress-optic coefficient C depends in general on both the material and the wavelength. It relates the difference in the principal stresses σ_I and σ_{II} to the phase difference δ.

For mechanically and photoelastically anisotropic materials, e.g. silicon, Eq. (2) does not hold true, because the lattice structure of silicon leads to a direction-dependent behaviour. In literature several different models exist for the description of the photoelastic properties of silicon. However, they lead to vastly different results (compare [3]–[5]) even when based on the same approach as shown in Fig. 1.

Therefore, in this work a new model is presented to describe the photoelastic material properties of silicon. This is based on the phenomenological approach for birefringence and photoelasticity by Pockels and Neumann [6]–[8].

Theoretical Analysis: Photoelasticity in (100)-Silicon

The indices of refraction due to birefringence of an unstressed material can be described by Maxwell's equations. However, this approach leads to impractically long equations. In order to simplify this, an analogy is used consisting of an ellipsoid that expresses the material properties.

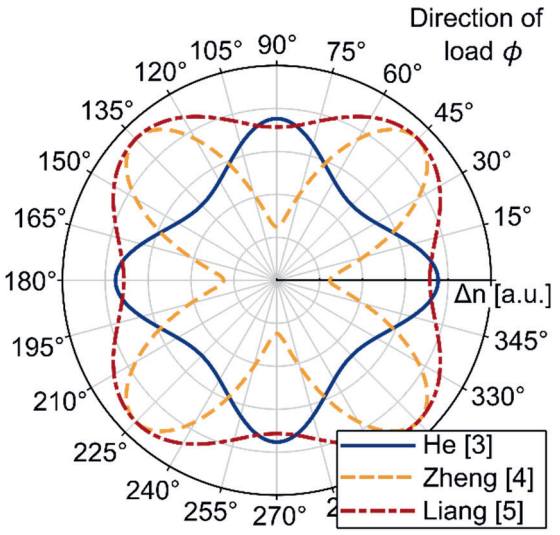


Fig. 1: Comparison of different models for the photoelasticity of silicon from literature expressed as the difference Δn between the coefficients of refraction

In literature this ellipsoid is usually called indicatrix [9]. It is based on an imaginary light ray falling into the centre of the indicatrix. Perpendicular to this ray, a plane is constructed and the intersection of the plane with the indicatrix generates an ellipse. The lengths of the two half axis of that ellipse correspond to the two indices of refraction. This analogy is shown in Fig. 2.

The indicatrix is a quadratic surface with in general six independent parameters:

$$\mathrm{B_{ij}} x_i x_j = 1 \text{ for } i = (1,2,3), \tag{4}$$

where B_{ij} is a 3×3 symmetric second-order tensor that depends on the material. It is called impermeability tensor. Vector x_i represents a Cartesian coordinate system which is in this work defined so that $x_1 \rightarrow x, x_2 \rightarrow y$ and $x_3 \rightarrow z$. By using the Einstein summation convention, Eq. (4) is expanded to:

$$\begin{aligned} &\mathrm{B_{11}x^2} + B_{22}y^2 + B_{33}z^2 \\ &+2\,B_{12}xy + 2\,B_{13}xz + 2B_{23}yz = 1. \end{aligned} \tag{5}$$

By rotating the indicatrix to align with the used (x, y, z)-coordinate system it can be expressed by only three independent values:

$$\mathrm{B_I x^2} + B_{II}y^2 + B_{III}z^2 = 1. \tag{6}$$

Here, B_I, B_{II} and B_{III} are the principal values of B_{ij}. By definition, the indicatrix can also be expressed as the reciprocal of the squared indices of refraction (n_I, n_{II}, n_{III}):

$$\frac{1}{n_I^2}\mathrm{x^2} + \frac{1}{n_{II}^2}y^2 + \frac{1}{n_{III}^2}z^2 = 1, \tag{7}$$

or in a more general form:

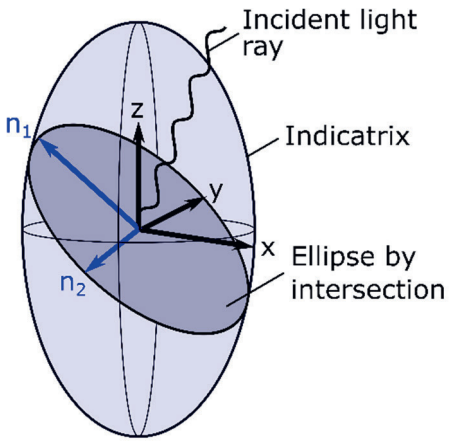


Fig. 2: Indicatrix as an analogy to describe the two indices of refraction for a certain direction of incident light in a birefringent material

$$B_{ij} = \frac{1}{n_{ij}}. \tag{8}$$

Pockels' and Neumann's phenomenological approach [6]–[8] to describe the effect of photoelasticity links the change of the impermeability ΔB_{ij} with the mechanical stress tensor σ_{ij} and the mechanical strain tensor ε_{ij}, respectively, by a fourth-order tensor. In terms of the analogy of an indicatrix, this means that mechanical stresses or strains deform the indicatrix:

$$\Delta B_{ij} = p_{ijkl}\sigma_{kl}, \tag{9a}$$

$$\Delta B_{ij} = q_{ijkl}\varepsilon_{kl}. \tag{9b}$$

Here, p_{ijkl} and q_{ijkl} are the stress-optical and the strain-optical tensors, respectively. Both can be expressed by each other under consideration of Hooke's law for linear elasticity [10]. Therefore, in this work only the stress-optical relationship is considered.

The change in impermeability can be expressed as the difference between the stressed and the unstressed indicatrix given by the impermeability B_{ij} and B_{ij}^o for the stressed and the undeformed indicatrices. Using Eq. (8) they can be expressed by the refraction indices n_{ij} and n_{ij}^o:

$$\Delta \mathrm{B_{ij}} = B_{ij} - B_{ij}^o, \tag{10a}$$

$$\Delta \mathrm{B_{ij}} = \frac{1}{\left(n_{ij}\right)^2} - \frac{1}{\left(n_{ij}^o\right)^2}. \tag{10b}$$

By converting to a common denominator this can be rewritten to:

$$\Delta \mathrm{B_{ij}} = \frac{(n_{ij}^o + n_{ij})(n_{ij}^o - n_{ij})}{\left(n_{ij}\right)^2\left(n_{ij}^o\right)^2}. \tag{11}$$

Since the change in impermeability can be considered small in comparison to the unstressed

impermeability (meaning: $n_{ij} \approx n_{ij}^o$), two simplifications can be made:

$$\left(n_{ij}^o\right)^2 + \left(n_{ij}\right)^2 \approx 2\, n_{ij}^o, \tag{12a}$$

$$\left(n_{ij}^o\right)^2 \left(n_{ij}\right)^2 \approx \left(n_{ij}^o\right)^4. \tag{12b}$$

With those simplifications, Eq. (11) can be expressed as:

$$\Delta \mathrm{B_{ij}} \approx \frac{2}{\left(n_{ij}^o\right)^3} \left(n_{ij}^o - n_{ij}\right). \tag{13}$$

Leading to an expression in which the change between the indices of refraction of the stressed and the unstressed material is linked with the change in impermeability and, therefore, with the mechanical stress by Eq. (9a). In the following, the approximation sign is omitted, though, it still is an approximation only valid for small changes of impermeability. Rearranging Eq. (13) yields:

$$\left(n_{ij} - n_{ij}^o\right) = -\frac{\left(n_{ij}^o\right)^3}{2} \Delta \mathrm{B_{ij}}. \tag{14}$$

For an indicatrix whose half axis are aligned to the coordinate system this leads to a set of three equations:

$$\left(n_I - n_I^o\right) = -\frac{\left(n_I^o\right)^3}{2} \Delta \mathrm{B_I}, \tag{15a}$$

$$\left(n_{II} - n_{II}^o\right) = -\frac{\left(n_{II}^o\right)^3}{2} \Delta \mathrm{B_{II}}, \tag{15b}$$

$$\left(n_{III} - n_{III}^o\right) = -\frac{\left(n_{III}^o\right)^3}{2} \Delta \mathrm{B_{III}}. \tag{15c}$$

With regard to Eq. (9), the change in indices of refraction can be calculated for a stress state σ_{ij} and a known stress optic tensor p_{ijkl}. Because the impermeability and the stress tensors are symmetric second-order tensors, the stress-optic tensor must show certain symmetries [10]. This allows it to be written as a 6×6 matrix in Voigt notation. In this form the indices ij and kl of the tensor p_{ijkl} are reduced to $11 \to 1$, $22 \to 2$, $33 \to 3$, $13 \to 4$, $23 \to 5$ and $12 \to 6$. In the following, Voigt notation will be indicated by $^{(\mathrm{V})}$.

Due to the diamond structure of monocrystalline silicon there are only three independent parameters of the stress-optic tensor [10]:

$$p^{(\mathrm{V})} = \begin{pmatrix} p_{11} & p_{12} & p_{12} & 0 & 0 & 0 \\ p_{12} & p_{11} & p_{12} & 0 & 0 & 0 \\ p_{12} & p_{12} & p_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{44} & p_{44} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{16}$$

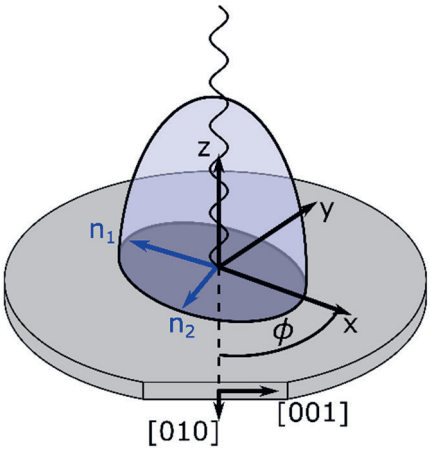


Fig. 3: Orientation of a (100)*-Si wafer and the indicatrix describing birefringence with reference to the global* (x, y, z) *coordinate system*

In order to account for the different crystalline orientations of the silicon, a rotation matrix R_{ij} is introduced. The rotated stress-optic tensor p'_{ijkl} accounts for different orientations of the silicon lattice structure by applying the rotation matrix R_{ij} to it:

$$p'_{ijkl} = R_{im} R_{jn} R_{ko} R_{lp} p_{mnop}. \tag{17}$$

For simplicity, in the following only mechanical stresses in plane with the (100) plane of silicon are discussed. The <100> direction is further assumed to be parallel with the z-axis of the coordinate system. A situation in which these simplifications hold true is represented by a (100)-silicon wafer as shown in Fig. 3. In that case the rotation matrix $R_{ij}(\phi)$ describing the rotation around the z-axis by an angle ϕ is:

$$R_{ij}(\phi) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{18}$$

Since the (100)-silicon wafer is considered sufficiently thin, only plane stresses are evaluated:

$$\sigma_{ij} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 \\ \sigma_{12} & \sigma_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{19}$$

Inserting the rotation matrix of Eq. (18) into Eq. (17) and applying both to Eq. (9a) yields the impermeability tensor for the plane stress with stress tensor:

$$\Delta B_{ij} = \begin{pmatrix} B'_{11}(\phi) & B'_{12}(\phi) & 0 \\ B'_{12}(\phi) & B'_{22}(\phi) & 0 \\ 0 & 0 & B'_{33}(\phi) \end{pmatrix}. \tag{20}$$

If the incident light ray falling into the indicatrix is parallel to the z-axis of the chosen coordinate system, Eq. (15c) can be neglected. Therefore, the change in the indices of refraction can be expressed by subtracting Eq. (15b) from (15a).

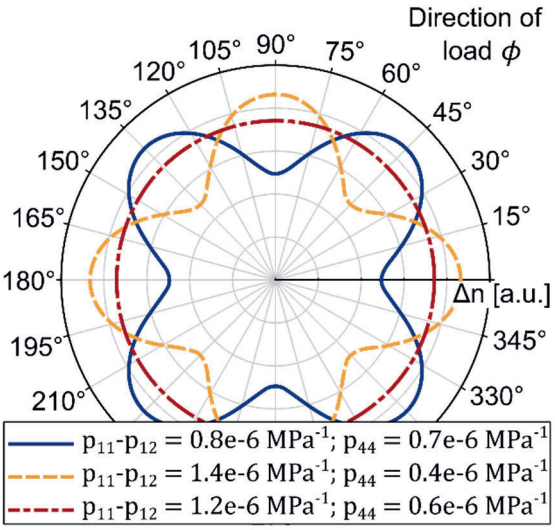


Fig. 4: Derived model for photoelasticity with arbitrarily chosen material parameters $p_{11} - p_{22}$ and p_{44} for a plane stress state without shear stresses ($\sigma_{12} = 0$)

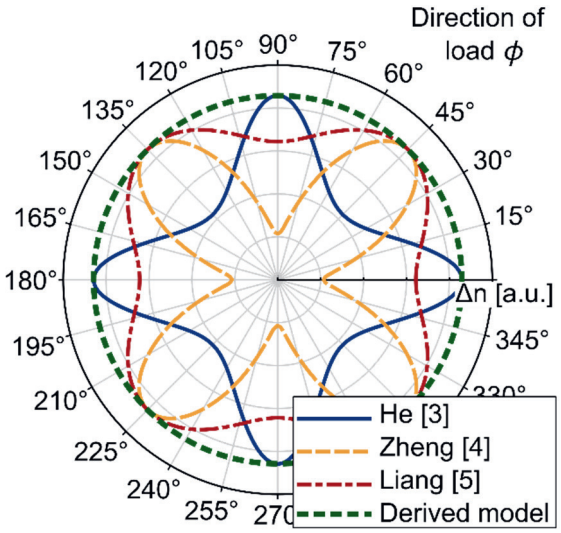


Fig. 5: Comparison of the derived model with models from literature for the birefringence under plane stress without shear stresses ($\sigma_{12} = 0$) for an isotropic material ($p_{44} = (p_{11} - p_{12})/2$)

Further, the unstressed birefringence is comparably small against the stress-induced birefringence $\left(n_{ij}^{o} = n^{o}\right)$ [11], yielding:

$$(n_I - n_{II}) = -\frac{(n^o)^3}{2}(\Delta \mathrm{B_I} - \Delta B_{II}). \tag{21}$$

To obtain the principal values ΔB_I and ΔB_{II} a simple eigenvalue analysis on the impermeability tensor of Eq. (20) can be performed:

$$\begin{aligned} \Delta \mathrm{B_I} &= \frac{1}{4}\left(a - \sqrt{b + c\,d}\,\right), \\ \Delta \mathrm{B_{II}} &= \frac{1}{4}\left(a + \sqrt{b + c\,d}\,\right), \end{aligned} \tag{22}$$

in which are:

$$\begin{aligned} a &= 2\,(\mathrm{p_{11}} + \mathrm{p_{12}})\,(\sigma_{11} + \sigma_{22}), \\ b &= 2\,((p_{11} - p_{12})^2 + 4p_{44}^2)\,(4\sigma_{12}^2 \\ &\quad + (\sigma_{11} - \sigma_{22})^2), \\ c &= 2\,((p_{11} - p_{12})^2 - 4p_{44}^2), \\ d &= (-4\sigma_{12}^2 + (\sigma_{11} - \sigma_{22})^2)\cos 4\phi \\ &\quad + 4\,\sigma_{12}(\sigma_{11} - \sigma_{22})\sin 4\phi. \end{aligned} \tag{23a}$$

Inserting the principal values of Eq. (22) into Eq. (21) yields the stress-optic law for a (100)-silicon wafer:

$$\Delta n = (n_I - n_{II}) = \frac{(n^o)^3}{4}\sqrt{b + c\,d}\,. \tag{24}$$

In this model there are only two independent variables since the difference between p_{11} and p_{12} cannot be separated. In Fig. 4 the derived model for photo-elasticity in silicon is shown for three arbitrarily chosen parameter combinations.

To describe a photoelastically isotropic material the stress-optic tensor of Eq. 16 can be replaced with the stress-optic tensor for an isotropic material [10]:

$$p_{iso}^{(\mathrm{V})} = \begin{pmatrix} p_{11} & p_{12} & p_{12} & 0 & 0 & 0 \\ p_{12} & p_{11} & p_{12} & 0 & 0 & 0 \\ p_{12} & p_{12} & p_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{p_{11} - p_{12}}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{p_{11} - p_{12}}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{p_{11} - p_{12}}{2} \end{pmatrix}. \tag{25}$$

Following the same derivation as for the anisotropic case, the stress-optic law reduces to:

$$\begin{aligned} \Delta n &= \frac{(n^o)^3}{2}(p_{11} - p_{12})\sqrt{(\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2} \\ &= \frac{(n^o)^3}{2}(p_{11} - p_{12})(\sigma_I - \sigma_{II}), \end{aligned} \tag{25}$$

which is the law for isotropic photo-elasticity as given in literature. The same can be achieved by replacing p_{44} with $(p_{11} - p_{12})/2$ in Eq. (24). This special case is shown in Fig. 4 as a red dash-dotted line.

Inserting the same replacements into the models from literature shows, that those are not able to produce the isotropic case as shown in Fig. 5 whereas the newly derived model shows the expected constant difference Δn in indices of refraction independent from loading direction ϕ.

Discussion and Conclusion

Using Pockels' and Neumann's phenomenological approach to describe photo-elasticity in anisotropic materials, a model was derived that

also yields the same results as derived in literature by the same approach depending on the material parameters $p_{11} - p_{12}$ and p_{44}. Additionally, the model is capable to describe photoelasticity in isotropic materials such as glass by choosing an appropriate combination of material parameters what is not possible with other models found in literature (compare [3]–[5]).

As a small caveat, the model is derived for a certain orientation of the crystalline diamond lattice structure in regard to the applied plane stress state and the direction of incident light. This simplification was made to allow an easy and concise determination of the principal values of the impermeability tensor B_{ij}. For application with different lattice orientations, the introduced rotation of the stress-optic tensor p_{ijkl} can be changed accordingly and with known material parameters carried out numerically.

Acknowledgement

This work has been financially supported by the European Union as part of the ESF-Program (Project No K-7531.20/462-9), which is gratefully acknowledged. Additionally, Christiane Schuster (Fraunhofer IKTS, Germany) and Ringo Köpge (Fraunhofer IMWS, Germany) deserve thank for their support.

Literature

[1] K. Ramesh, *Digital Photoelasticity - Advanced Technisques and Applications*, 1st ed. Springer-Verlag Berlin Heidelberg, 2000.

[2] W. Zinth and U. Zinth, *Optik*, 2nd ed. München: Oldenbourg Verlag München, 2009.

[3] S. He, T. Zheng, and S. Danyluk, “Analysis and determination of the stress-optic coefficients of thin single crystal silicon samples,” *J. Appl. Phys.*, vol. 96, no. 6, pp. 3103–3109, 2004.

[4] T. Zheng and S. Danyluk, “Study of stresses in thin silicon wafers with near-infrared phase stepping photoelasticity,” *J. Mater. Res.*, vol. 17, no. 1, pp. 36–42, 2002.

[5] H. Liang, Y. Pan, S. Zhao, G. Qin, and K. K. Chin, “Two-dimensional state of stress in a silicon wafer,” *J. Appl. Phys.*, 1992.

[6] F. Pockels, *Lehrbuch der Kristalloptik*, 1st ed. Leipzig: B. G. Teubner, 1906.

[7] F. Pockels, “Ueber den Einfluss elastischer Deformationen, speciell einseitigen Druckes, auf das optische Verhalten krystallinischer Körper,” *Ann. Phys.*, vol. 273, no. 6, pp. 269–305, 1889.

[8] K. E. Neumann, “Die Gesetze der Doppelbrechung des Lichts in comprimirten oder ungleuichförmig erwärmten unkrystallinischen Körpern,” *Ann. der Phys. und Chemie*, vol. 54, no. 12, pp. 449–476, 1841.

[9] E. Hecht, *Optik*, 7. de Gruyter, 2018.

[10] T. S. Narasimhamurty, *Photoelastic and Electro-Optic Properties of Crystals*, 1st ed. Springer US, 1981.

[11] C. Krüger, D. Heinert and A. Khalaidovski, “Birefringence measurements on crystalline silicon,” *Class. Quantum Gravity*, vol. 33, no. 1, 2015.

DOI 10.5162/sensoren2019/P2.7

Study of interdigital electrode structures for the detection of water spray

Jakob Döring, Lakshan Tharmakularajah, Karl-Ludwig Krieger
Institute of Electrodynamics and Microelectronics (ITEM.ae), University of Bremen
{doering, lakshan, krieger}@item.uni-bremen.de

Abstract

This paper presents a study of interdigital electrode structures suitable for the detection of road surface wetness. The structures are studied by means of the finite element method (FEM) with the aim of analyzing the effect of varying design parameters on the detection of water. A variation of the number of digits on a fixed electrode area shows a trade-off between dynamic range and penetration depth. Furthermore, increasing the ratio of electrode width and distance results in an enhanced signal-to-noise ratio. In addition, the effects of varying the substrate thickness and applying a shield electrode are evaluated. The study of placing a coating layer on the electrode structure shows the significance of the choice of its thickness and permittivity. Finally, simulations with different water drop sizes show the usability of detecting even small amounts of drops and verify the assumptions made before.

Keywords: Planar capacitive sensors, interdigital transducers, wetness detection, vehicle safety systems, finite element method.

I. Introduction

Even nowadays, in times of steadily increasing sensor functionality and related driver assistance functions, there are numerous traffic accidents involving personal injury every year. In 2017, there were 302,656 of such accidents in Germany in which a total of 393,492 people were injured [1]. In around 5 % of these cases, the cause of the accident is stated to be the road surface condition. Although today's vehicles are already equipped with extensive environment sensors, including radar, lidar, ultrasound and camera systems, these do not provide direct information about a road's current state. Currently, the driver of the vehicle estimates the current state intuitively and experience-based and adapts his driving accordingly. Misjudgments can have fatal consequences for the driver and other traffic participants.

With the increasing level of automation in today's motor vehicles, the detection of road surface conditions is continuing to increase in importance. Particularly with regard to highly or fully automated driving (SAE Level 4 and 5), knowledge of the current condition is of great relevance in order to be able to handle safety-critical driving situations due to weather conditions without the visual control of the driver. In addition to existing data from environment and motor vehicle sensors, innovative approaches are needed to reliably monitor road surface condition. Thus, the driver or the system can be warned in time of potentially dangerous situations that could lead to an unstable vehicle condition due to loss of traction.

A road surface condition that can lead to a significant loss of traction is road wetness. In our previous work [2], we presented a novel approach for detecting road surface wetness, which can prospectively contribute to increase the safety of vehicle passengers and traffic participants in times of highly or fully automated driving. Here, planar capacitive sensors are placed on a motor vehicle's wheel arch liner. If the road surface is wet, the tires cause water to whirl up, which impinges upon the wheel arch liner in the form of drops (water spray). The capacitive sensor can detect the water drops and thus indirectly infer a wet road surface. Since a greater water film height leads to more water being whirled up (to a certain extent) and thus more drops impinging on the sensor, a basic distinction between wetness levels is possible.

As the focus in [2] was on demonstrating the feasibility of the approach, a limited design space was used for the simulations. Furthermore, effects on varying design parameters were neglected. In this paper, we present an enhanced study of interdigital electrode structures for the detection of water spray. The structures are studied by means of the finite element method (FEM) in COMSOL Multiphysics. Fundamental characteristics of interdigital electrode structures, including electrode width and distance, are initially examined by assuming a static water layer. Furthermore, the transducer's substrate thickness and the effect of a shield

electrode are studied. Since the main application is road wetness detection, a coating layer is needed to mechanically protect the electrodes. For this reason, the influence of a coating layer is studied in the simulations. Finally, we simulate selected electrode structures with water drops of different size to show the assumptions previously made for a static water layer hold.

II. Planar capacitive sensors

Due to its wide range of features, a possible type of sensor suitable for the detection of road surface wetness is the capacitive sensor, as shown in [2]. Because of desirable features like robustness, low costs and flexibility in electrode design [3, 4, 5] they have already been used in a wide range of applications. For instance, they are used for liquid level sensing in containers [6], the detection of ice in liquid conveyor pipes [7] or the measurement of soil water content [8].

The measuring principle of capacitive sensors is based on the interaction of a monitored medium and an applied electric field between the sensor electrodes [3, 9]. Since the electric field penetrates through the medium, electric displacement is generated inside it, which alters the charge stored between the sensor electrodes. Thus, the sensor capacitance changes, allowing inferences to be drawn about the permittivity of the medium. Hence, the determination of system parameters that correlate with the permittivity of the medium (e.g. moisture) is possible.

The more conventional parallel-plate capacitive sensor is based on the same principle of operation as a planar capacitive sensor [10]. Between the two parallel plates, one representing the driving electrode and the other representing the sensing electrode, an electric field is distributed uniformly [3]. Due to the much larger dimensions of the electrode surface compared to the electrode thickness, the fringe field is negligible. Since the electrodes are opened up in case of a planar capacitive sensor, the fringe field becomes predominant, which is why this kind of sensor is often denoted as a fringe field sensor in literature [3, 10, 11]. Even in the case of the planar capacitor, the electric field lines penetrate into the monitored medium. Thus, the sensor enables single-sided access to the medium, which is one of the most important advantages of a planar capacitive sensor in addition to the simple and low-cost fabrication process [9].

In order to meet the requirements of an application, in this case the detection of water spray, the choice of sensor geometry is critical as it is a major determination factor for sensor performance [12]. A suitable design of a planar capacitive sensor is an interdigital electrode structure, since the contribution of the fringing field effect in sensor capacitance is maximized [8, 11].

An interdigital electrode structure consists of two electrodes, one being the sensing electrode and the other being the driving electrode. Both electrodes are generally built up by the same number of digits n, which are separated by a distance d. Another parameter commonly used in the context of interdigital electrode structures is the spatial wavelength λ [10]. It can be defined as the distance between the centerlines of adjacent digits belonging to the same electrode

$$\lambda = 2 \cdot (w + d). \tag{1}$$

In order to design an interdigital sensor, λ has to be considered since the penetration depth $\gamma_{97\%}$ of the electric field is approximately proportional to it [10]. There is no unique definition for the penetration depth in literature, but one possible way to evaluate it, is to study how deep the electric field penetrates into a monitored medium [9]. For this purpose, the thickness of the monitored medium is increased incrementally until the point is reached, where 97 % of the asymptotic capacitance is obtained

$$\frac{C(t_\mathrm{m} = \gamma_{97\%}) - C(t_\mathrm{m} = 0)}{C(t_\mathrm{m} = \infty) - C(t_\mathrm{m} = 0)} \cdot 100\ \% = 97\ \%. \tag{2}$$

$C(t_\mathrm{m} = \gamma_{97\%})$ is the capacitance, where the thickness of the monitored medium t_m equals the penetration depth $\gamma_{97\%}$, $C(t_\mathrm{m} = \infty)$ is the asymptotic capacitance for the monitored medium and $C(t_\mathrm{m} = 0)$ the capacitance in the absence of a medium.

In addition to the penetration depth, there are further parameters usually used to evaluate interdigital electrode structures including signal strength, dynamic range and measurement sensitivity [3, 12]. While the signal strength describes the output capacitance of a sensor for a monitored medium of a certain thickness $C(t_\mathrm{m})$, the dynamic range DR indicates the difference of the maximum capacitance for a monitored medium and the basic capacitance *BC* (absence of a medium)

$$DR = C(t_\mathrm{m} = \infty) - C(t_\mathrm{m} = 0). \tag{3}$$

By maximizing the dynamic range, and simultaneously minimizing the basic capacitance, a maximum signal-to-noise ratio can be achieved. The measurement sensitivity S can be defined as the ratio of the change in the sensor output to the change in the system variable. In the context of this paper, the thickness of the monitored medium represents the system variable. Thus, the sensitivity can be calculated from the ratio of changes in capacitance to changes of the medium's thickness

$$S = \frac{\mathrm{d}C(t_\mathrm{m})}{\mathrm{d}t_\mathrm{m}}. \tag{4}$$

III. FEM simulations

The analytical modeling of interdigital electrode structures is not trivial compared to classic electrode shapes such as parallel-plate electrodes in which the electric field is distributed homogeneously between the plates [8, 10]. For that reason, FEM is commonly used for sensor modeling, optimization and performance evaluation [12]. In order to study interdigital electrode structures and their design parameters with regard to the detection of water, we use COMSOL Multiphysics for two- and three-dimensional electrostatic FEM simulations.

Since the target application is the detection of water spray on a vehicle's wheel arch liner, we have limitations in design parameters, which are considered in the simulations, as for example the electrode area is fixed at 50 mm × 50 mm. Furthermore, we focus on printed circuit board (PCB) technology as it enables to fabricate small-scale units with high precision at low per-unit costs [3, 8, 11]. The relevant parameters and dimensions studied in the simulations are depicted in Fig. 1.

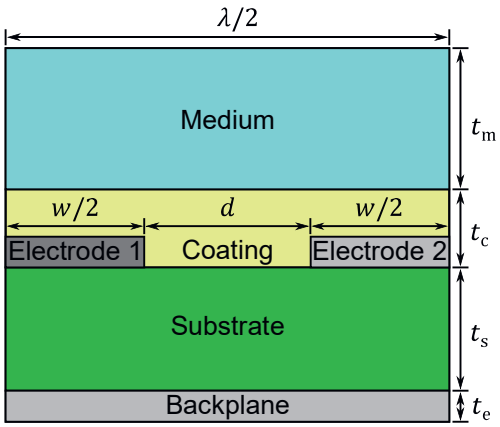


Fig. 1: Schematic cross-section view of an interdigital electrode structure with its relevant parameters and dimensions.

In general, we assume a static water layer thickness t_m above the electrode structure, which is varied from 0 to 10 mm. Additionally, a water layer thickness of 200 mm is simulated to determine the asymptotic capacitance for the evaluation of the penetration depth $\gamma_{97\%}$. The relative permittivity ε_r is assumed to equal 1 for air, 80 for water and 4.5 for FR4, which is a common material for PCBs. The remaining parameters are varied and therefore described in the corresponding subsections.

A. Number of digits

The first parameter to be varied is the number of digits n from 3 to 50, to evaluate its influence on the structure's characteristics. As the electrode area is fixed at 50 mm × 50 mm, a variation of n at the same time means a variation of the spatial wavelength λ (see Eq. (2)) from around 18 mm to 1 mm. For the study of n, the substrate thickness t_s equals 1.5 mm and the electrode thickness t_e equals 36 µm, which are common values for PCBs. Moreover, the backplane and the coating, depicted in Fig. 1, are neglected and evaluated later in this paper.

The result for the effect of varying n on the capacitance against water layer thickness is shown in Fig. 2. Since the simulations are carried out in two dimensions, the capacitance is given in pF/mm and would have to be multiplied with the length of the electrodes (50 mm) to get the approximated overall capacitance (three dimensions). However, for the evaluation of the effect, we restrict the simulations to two dimensions.

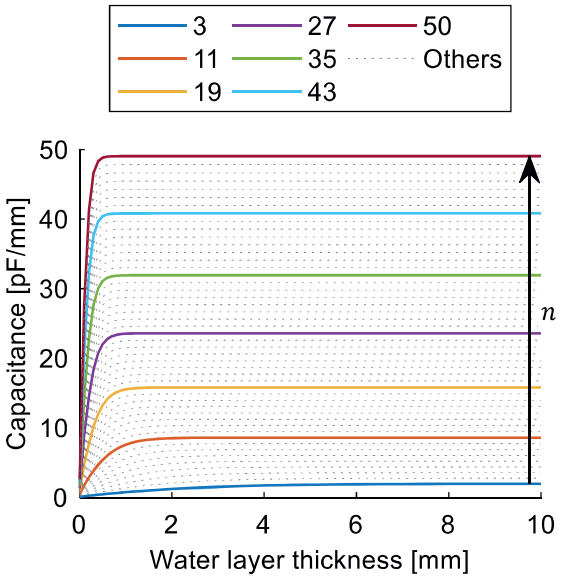


Fig. 2: Capacitance against water layer thickness for a variation of the number of digits n from 3 to 50.

As Fig. 2 shows, the signal strength increases with the number of digits n. While the capacitance of 3 digits (10 mm water layer thickness) equals around 1.96 pF/mm, the capacitance of 50 digits with around 49 pF/mm is 25 times greater. Besides the signal strength, also the measurement sensitivity (see Eq. (4)) increases, especially in close range, as the electrodes are moving together and the electric field gets more intensive in this range.

In Fig. 3 the penetration depth and the dynamic range are shown against the number of digits. By increasing n, the penetration depth decreases while the dynamic range increases. Since more digits lead to lower electrode gaps, the electric field lines pass a smaller distance and the electric field strength decreases with the height. Thus, the penetration depth declines from

around 6.8 mm to around 0.4 mm. In contrast to the penetration depth, the dynamic range increases with the number of digits from approximately 1.9 pF/mm to 46.4 pF/mm. However, at the same time the basic capacitance increases, too. Since the growth is low in comparison (~0.09 pF/mm to ~2.6 pF/mm), an improvement of the signal-to-noise ratio can be achieved by increasing the number of digits.

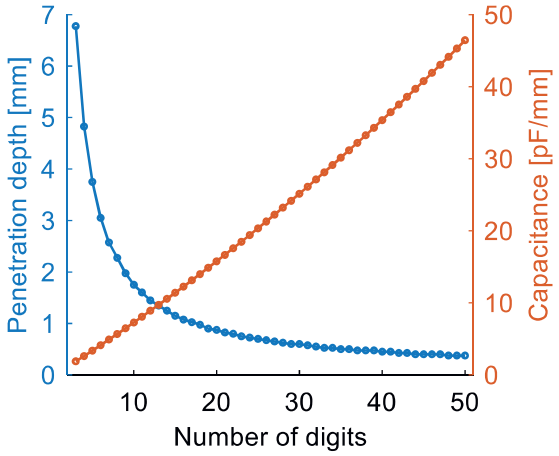


Fig. 3: Penetration depth and dynamic range against number of digits n.

By means of the electric field line and potential distribution the significant area around the interdigital electrode structure can be visualized. In Fig. 4 a part of the structure for 25 digits, a water layer thickness of 1 mm and a substrate thickness of 1.5 mm is exemplarily shown.

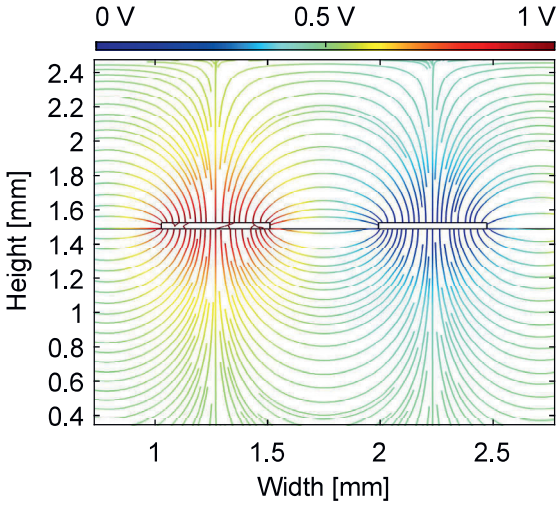


Fig. 4: Electric field (streamlines) and potential distribution (color) for 25 digits.

The electric potential distribution is described by color coding. While blue represents 0 V and red 1 V, green represents the part of the electric field which is not altered by the electrodes. As illustrated, the electric field penetrates into the medium and the substrate leading both to contribute to the capacitance. Furthermore, the figure enables the penetration depth, which equals 0.7 mm for 25 digits, to be displayed.
The results demonstrate, there is always a trade-off between penetration depth and dynamic range or measurement sensitivity when designing an interdigital electrode structure considering the number of digits, respectively the spatial wavelength. Thus, the design parameters have to be adjusted to the application's requirements.

B. Electrode width

The next parameter to be studied is the electrode width w and hence, the ratio of width w and distance d. Therefore, we introduce the dimensionless variable r, which is used to calculate the electrode width and the electrode distance depending on the electrode area's width a for the simulations

$$w = r \cdot \frac{a}{4n-1}, \; d = \frac{a-(2nw)}{2n-1}. \tag{5}$$

In order to evaluate the electrode width's effect on the electrode structure's characteristics we exemplarily present results for 25 digits and vary r from 0.3 to 1.7. In this case 0.3 yields to an electrode width of approximately 0.15 mm and 1.7 to a distance of around 0.15 mm, which are commonly the minimum track width and the minimum spacing at PCB manufacturers. As for the study of the number of digits, we assume a substrate thickness t_s of 1.5 mm, an electrode thickness t_e of 36 µm and neglect backplane and coating.

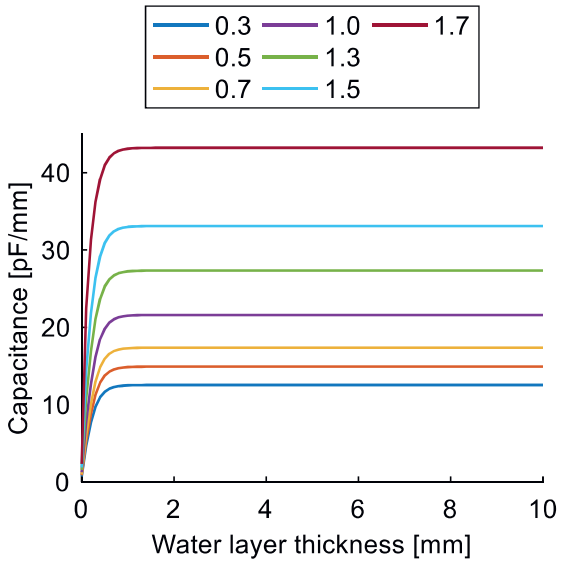


Fig. 5: Capacitance against water layer thickness for a variation of r from 0.3 to 1.7.

The resulting capacitance against water layer thickness curves for a selection of r are shown

in Fig. 5. All curves show a similar exponential behavior. Increasing r results in a larger signal strength, which ranges from approximately 12.5 pF/mm in saturation (0.3) to 43.2 pF/mm (1.7) due to the broader electrodes moving together. As additionally illustrated in Fig. 6, not only the capacitance in saturation increases, but also the dynamic range by more than three times from around 11.8 pF/mm to 40.9 pF/mm, since the basic capacitance has comparably small values. As a result, the signal-to-noise ratio can be extended.

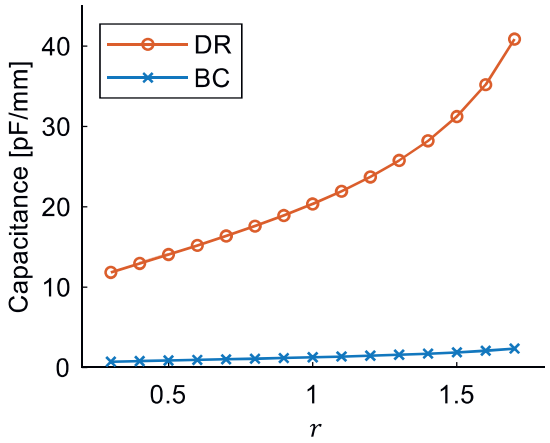


Fig. 6: Dynamic range and basic capacitance against r.

C. Substrate thickness

In this section, the effect of varying the substrate thickness t_s is studied. For this purpose, we exemplarily present results for an interdigital electrode structure with 25 digits and a varying substrate thickness from 0.1 mm to 1.5 mm. The electrode thickness t_e is assumed to be 36 µm and a coating is neglected. In the first part of the study, we also neglect the backplane depicted in Fig. 1.

An increase in substrate thickness results in a higher signal strength in saturation of about 0.5 pF/mm for the whole range. It is caused by the electric field that penetrates into a substrate with a permittivity of 4.5 instead of air with a permittivity of 1. A comparable effect would be caused by changing the substrate's material and thereby the permittivity of it. As Fig. 7 illustrates, the basic capacitance increases with the substrate thickness by a similar value as the signal strength. Thus, the dynamic range, and also the measurement sensitivity, is almost unaffected and the signal-to-noise ratio declines. The basic capacitance increases, but the maximum value is limited due to the penetration depth of the interdigital electrode structure, which is constant (in this instance at 0.7 mm) for all substrate thicknesses.

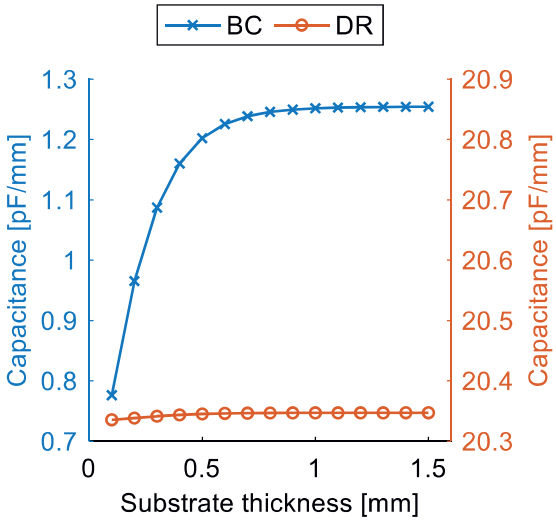


Fig. 7: Dynamic range and basic capacitance against substrate thickness.

In Fig. 8 the electric field line and potential distribution for a part of the structure is shown for a substrate thickness of 0.1 mm and a water layer thickness of 1 mm. The electric field penetrates through the substrate in the air on the backside. Therefore, not only the monitored medium on the frontside is in the focus of the measurement, but also the medium on the backside, which could lead to misinterpretations and in the context of road wetness detection to fatal errors. Thus, it is preferable to have a substrate thickness thicker than the electrode structure's penetration depth, even though the basic capacitance rises.

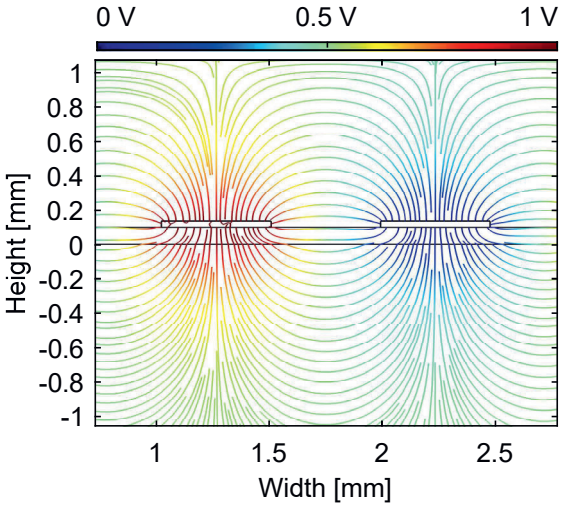


Fig. 8: Electric field (streamlines) and potential distribution (color) for 25 digits and a substrate thickness of 0.1 mm.

Another possibility to limit the measuring range to the frontside, is to use a backplane which shields the electrode structure from noise. In order to evaluate the effect of the backplane, we repeat the simulations with the same settings. In

contrast to the results without a backplane, a decrease in the substrate thickness leads to a higher signal strength in saturation. The increase in signal strength from a substrate thickness of 1.5 mm to 0.1 mm is approximately 6.7 pF/mm and is thus in a different order of magnitude compared to the simulations without a shield due to the capacitance built between the electrode structure and the backplane. For the same reason, the basic capacitance rises with a decreasing substrate thickness, which is shown in Fig. 9. Besides the basic capacitance, the dynamic range and the measurement sensitivity also increase with reducing the substrate thickness.

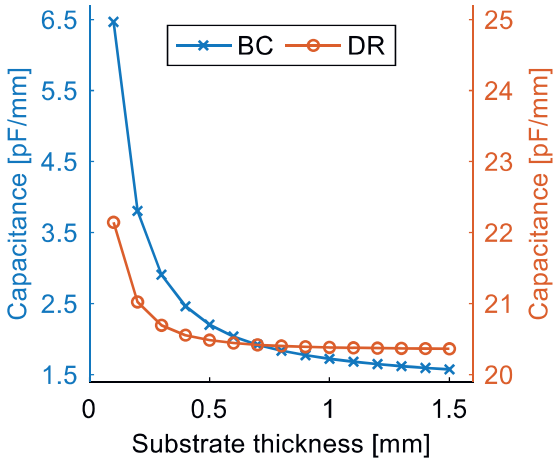


Fig. 9: Dynamic range and basic capacitance against substrate thickness with a backplane.

The electric field line and potential distribution for a part of the studied structure is shown for a substrate thickness of 0.1 mm and a water layer thickness of 1 mm in Fig. 10.

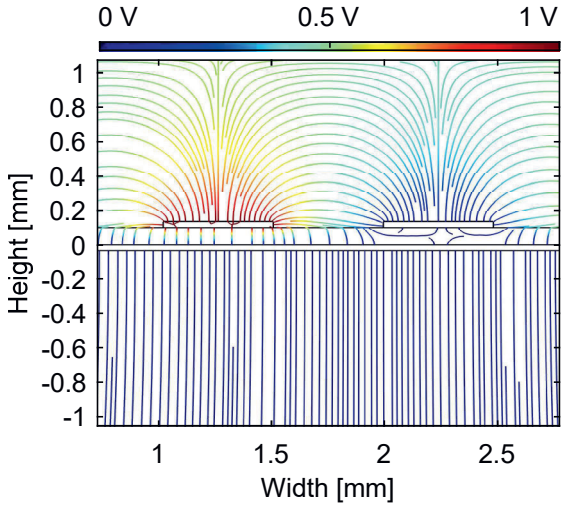


Fig. 10: Electric field (streamlines) and potential distribution (color) for 25 digits and a substrate thickness of 0.1 mm with a backplane.

In contrast to Fig. 8, the electric field lines cannot penetrate through the substrate due to the backplane. Therefore, only the electrode structure's frontside is relevant for the measurement. The positive electrode and the grounded backplane build a kind of parallel-plate capacitor which results in a high basic capacitance.

D. Coating

In this section, we evaluate the effect of the thickness and the permittivity of a coating on top of an interdigital electrode structure. For this purpose, the coating thickness t_c is varied from 0 to 1 mm. Additionally, we study the effect of the coating's permittivity by evaluating two materials. For the simulations, a substrate thickness t_s of 1.5 mm and an electrode thickness t_e of 36 µm are assumed. Additionally, the backplane is neglected. In the following we exemplarily present the results for 25 digits.

In Fig. 11, the simulated capacitance against water layer thickness for a variation of coating thickness and a relative permittivity of 3.4 (e.g. polyimide) is presented. As the figure shows, the capacitance is damped increasingly with the coating thickness. Up to a certain coating layer thickness, the course of the capacitance curves shows a similar exponential behavior as before. However, if the coating thickness approaches the structure's penetration depth, the curves get almost constant.

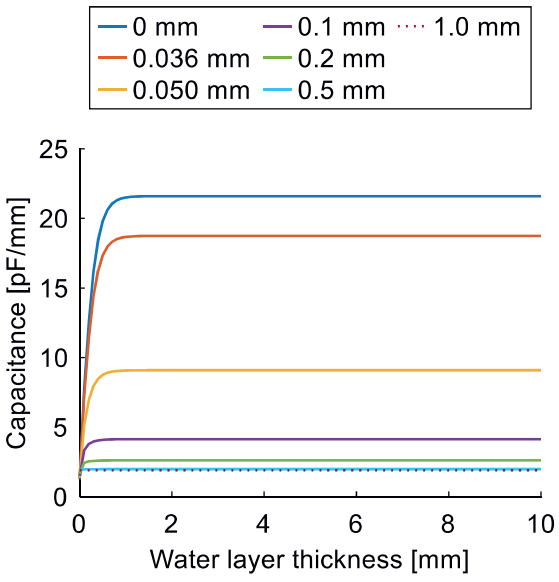


Fig. 11: Capacitance against water layer thickness for a variation of the coating thickness from 0 to 1 mm and a relative permittivity of 3.4.

Fig. 12 illustrates the electric field line and potential distribution for a coating thickness of 0.1 mm. In contrast to the behavior in Fig. 8, where no coating is applied, the electric field strength in

the coating layer is significantly higher than at the transition to the medium. Thus, the coating accounts for a large part of the capacitance, as more field lines pass through it.

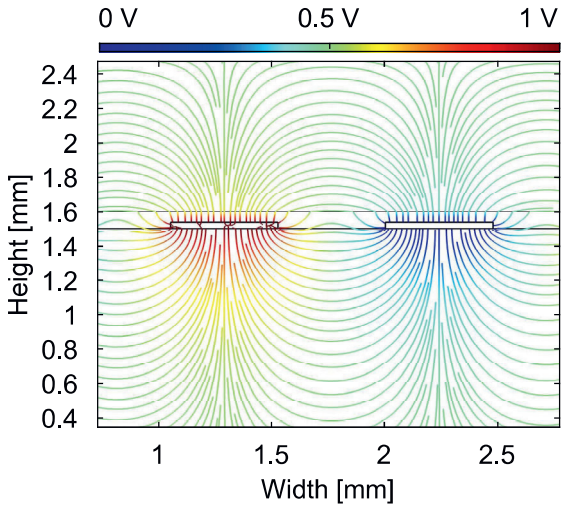


Fig. 12: Electric field (streamlines) and potential distribution (color) for 25 digits and a coating thickness of 0.1 mm and a relative permittivity of 3.4.

By increasing the permittivity of the coating layer to 11.7 (e.g. silicon), the behavior alters, as shown in Fig. 13. As the difference in permittivity between the coating and the medium reduces, more electric field lines pass through the coating into the medium resulting in a lower damping of the capacitance. Comparing Fig. 12 and Fig. 13, which represent the same coating thickness, the significant part of the electric field lines penetrating into the medium is noticeably larger.

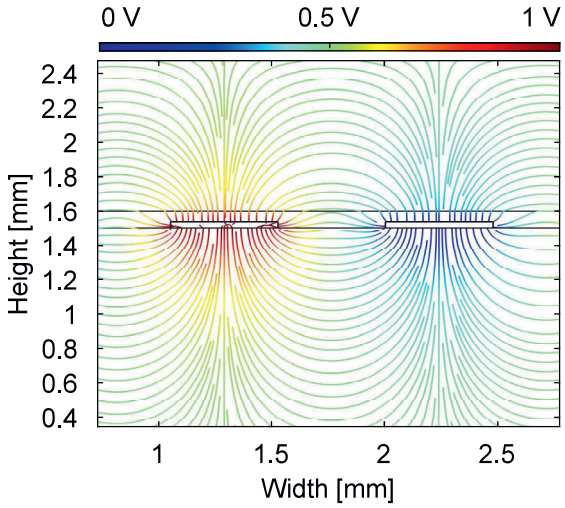


Fig. 13: Electric field (streamlines) and potential distribution (color) for 25 digits and a coating thickness of 0.1 mm and a relative permittivity of 11.7.

Fig. 14 shows the dynamic range and basic capacitance against coating thickness for two materials. Due to more electric field lines passing the coating layer, increasing the thickness results in a higher basic capacitance. Furthermore, increasing the permittivity shows the same effect. Similar to the results for the substrate thickness, the maximum basic capacitance is limited due to the penetration depth. By increasing the coating's thickness, the dynamic range decreases significantly in both cases. While there is no difference between the two materials for a thickness of 36 µm, as the electrodes are not covered on top, a significant variation is noticeable already with a thickness of 50 µm. Here, the percentage decrease in dynamic range for a relative permittivity of 3.4 is about 62 % and for 11.7 about 38 %. Already at 0.1 mm, there is only 13 % of the dynamic range left for a relative permittivity of 3.4, while there are 32 % left for 11.7. Thus, both curves for the dynamic range converge to zero at different rates showing advantages for a higher relative permittivity.

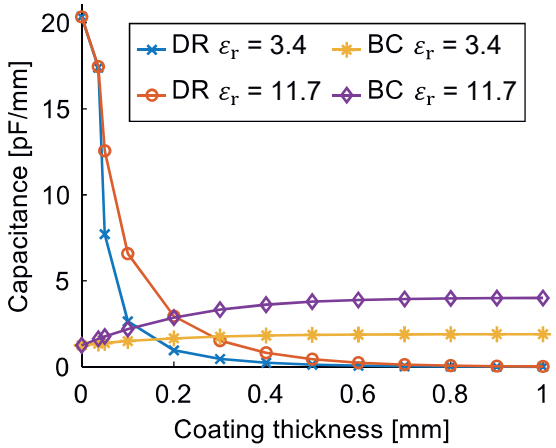


Fig. 14: Dynamic range and basic capacitance against coating thickness.

E. Water spray

In contrast to the previous simulations, where a static water layer is assumed to simplify the simulations and compare the structure's characteristics, in this section, the structure's behavior towards the presence of water drops is presented. For this purpose, two designs (7 and 25 digits) are exemplarily compared in three-dimensional electrostatic simulations. With regard to the results of the previous sections, both designs have an electrode distance of 0.15 mm, a substrate thickness of 1.5 mm and a backplane. The water drops are realized by hemispheres of three different radii (0.5 mm, 1 mm and 2.5 mm) and are placed as arrays of size 1 × 1 to 5 × 5 on top of the electrode structures.

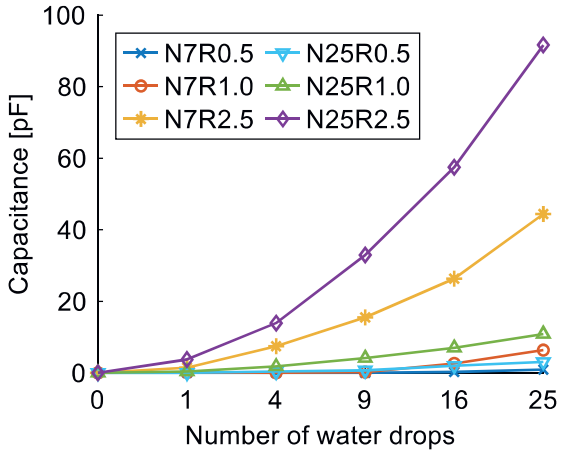


Fig. 15: Capacitance range against number of water drops for two designs.

In Fig. 15, the capacitance range against the number of water drops curves are shown for the two designs. As for the simulations with a static water layer, the increase in capacitance and the measurement sensitivity are higher for a greater number of digits. Especially for smaller drops and low amounts, the design with 25 digits is advantageous, as the increase is considerably greater. The results show the possibility of detecting even small amounts of water drops and confirm the assumptions made in the previous sections.

IV. Conclusions

This paper presented a study on interdigital electrode structures suitable for the detection of water spray. By means of FEM simulations in COMSOL Multiphysics the structures were analyzed with regard to the effect of varying their design parameters. A variation of the number of digits on a fixed electrode area shows a trade-off between dynamic range and penetration depth. Furthermore, an improvement for the signal-to-noise ratio is possible, as also for increasing the ratio of electrode width to distance. The choice of substrate thickness is directly related to the required penetration depth and a possible shield electrode. In addition, the effect of coating thickness and material were evaluated. Here, the capacitance is damped increasingly with the thickness. By increasing the coating's relative permittivity, the attenuation decreases. Finally, we presented results for the behavior to the presence of water drops and confirmed the assumptions made before.

References

[1] Statistisches Bundesamt, "Verkehrsunfälle 2017," *Verkehr,* 2018.

[2] J. Döring, L. Tharmakularajah, J. Happel and K.-L. Krieger, "A novel approach for road surface wetness detection with planar capacitive sensors," *Journal of Sensors and Sensor Systems,* vol. 8, no. 1, pp. 57--66, 2019, doi: 10.5194/jsss-8-57-2019.

[3] X. Hu and W. Yang, "Planar capacitive sensors - designs and applications," *Sensor Review Vol. 30 Issue: 1,* pp. 24-39, 2010, doi: 10.1108/02602281011010772.

[4] L. K. Baxter, Capacitive Sensors: Design and Applications, New-York: Wiley-IEEE Press, 1997, doi: 10.1109/9780470544228.

[5] G. Brasseur, "Design rules for robust capacitive sensors," *IEEE Transactions on Instrumentation and Measurement,* vol. 52, no. 4, pp. 1261-1265, 2003, doi: 10.1109/TIM.2003.816812.

[6] H. Canbolat, "A Novel Level Measurement Technique Using Three Capacitive Sensors for Liquids," *IEEE Transactions on Instrumentation and Measurement,* vol. 58, no. 10, pp. 3762-3768, 2009, doi: 10.1109/TIM.2009.2019715.

[7] J. Happel, J. Döring, K.-L. Krieger, J. Deitschun and D. Godlinski, "Printed Capacitive Sensors for Contactless Ice Detection in Automotive Liquid Conveyor Pipes," *AMA Conferences -- Proceedings Sensor,* pp. 127-132, 2017, doi: 10.5162/sensor2017/P3.2.

[8] J. Mizuguchi, J. C. Piai, J. A. d. França, M. B. d. M. França, K. Yamashita and L. C. Mathias, "Fringing Field Capacitive Sensor for Measuring Soil Water Content: Design, Manufacture, and Testing," *IEEE Transactions on Instrumentation and Measurement,* pp. 212-220, January 2015, doi: 10.1109/TIM.2014.2335911.

[9] M. J. da Silva, Impedance Sensors for Fast Multiphase Flow, Technical University of Dresden, 2008.

[10] A. V. Mamishev, K. Sundara-Rajan, F. Yang, Y. Du and M. Zahn, "Interdigital sensors and transducers," *Proceedings of the IEEE 92,* pp. 808-845, May 2004, doi: 10.1109/JPROC.2004.826603.

[11] R. N. Dean, A. K. Rane, M. E. Baginski, J. Richard, Z. Hartzog and D. J. Elton, "A Capacitive Fringing Field Sensor Design for Moisture Measurement Based on Printed Circuit Board Technology," *IEEE Transactions on Instrumentation and Measurement,* vol. 61, no. 4, pp. 1105-1112, April 2012, doi: 10.1109/TIM.2011.2173041.

[12] X. B. Li, S. D. Larson, A. S. Zyuzin and A. V. Mamishev, "Design principles for multichannel fringing electric field sensors," *IEEE Sensors Journal,* vol. 6, no. 2, pp. 434-440, 2006, doi: 10.1109/JSEN.2006.870161.

DOI 10.5162/sensoren2019/P2.8

Simultaneous measurement of resistance and temperature changes in bridge circuits

Klemens Gintner
University of Applied Sciences Karlsruhe, Faculty MMT, Moltkestrasse 30, 76133 Karlsruhe, Germany

Abstract

Wheatstone Bridge Circuits are commonly used in sensor applications to measure the resistance change of sensor elements, which represent typical metrics such as pressure, force or magnetic flux density. In many applications, for example, magnetoresistive or pressure-sensitive sensor elements are used in Wheatstone bridges. A temperature change of the Wheatstone bridge leads not only to a change of the resistance of the sensor elements, but also of the reference resistors. The reason for this is the (mostly positive) temperature coefficient of the resistors and also the dependence of the sensor effect on the temperature. In order to obtain a more accurate result, the temperature of the bridge circuit should therefore be known. By using several different voltages in a bridge circuit and knowing the temperature coefficients, it is possible to determine the exact resistance change of the sensor elements and the temperature change simultaneously without using an additional temperature sensor. A possible offset of the bridge circuit due to a mismatch of the four resistors is not relevant.

Keywords: Bridge-Circuit, Temperature, resistance change

1. Method for determining the resistances in a Wheatstone bridge under reference conditions

In this proposal it is essential to know the exact values of the four resistances in the Wheatstone bridge under reference conditions, i.e. the temperature is given; then the temperature change is zero (dT = 0, therefore the temperature coefficients a and b are not relevant) and there is no change in resistance as no measured quantity is applied (then: dR = 0).

To accurately measure each of the four resistors, a floating zero-resistance ammeter is suggested, as shown in Figure 1a and 1b. It is important to note that the bridge circuit with R1 to R4 has already been produced and is ready for use. The basic idea is to short-circuit one of the two resistors connected in series connection (in Figure 1a and 1b: node 1 and node 2 are short-circuited so that R4 is no longer relevant) and measure the current (here: I(R3)) using operational amplifiers so that, for example, the resistance of R3 can be determined by the voltage U_3 (then $U_3 = U_V$) divided by the measured current $I(R_8)$; here: $I(R_8) = I(R_3)$ with $I(R_3)$ as the current through the (reference) resistor R3 in the Wheatstone bridge; all this is resulting in: $R3 = U_3/I(R_3) = U_V/I(R_8)$. The current $I(R_8)$ (= $I(R_7)$) can be derived by using the difference of the voltages at the output of the two operational amplifiers (U_{A1} - U_{A2}) according to equation (1); it is important to know that ($R_5 \cdot R_7$) is equal to ($R_6 \cdot R_8$).
Equation (1) for $I(R_3)$ results in:

$$I(R_3) = \frac{U_{A1} - U_{A2}}{R_7 \cdot \left(1 + R_5 / R_6\right)} = I(R_8) = I(R_7) \quad (1)$$

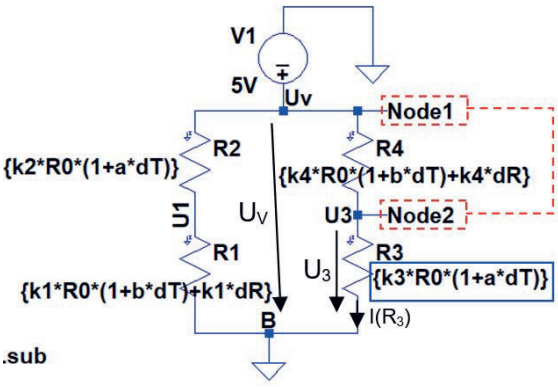


Fig. 1a: Floating zero-resistance ammeter for measuring the resistances at reference conditions (dR = 0 and dT = 0) (for further details see [1])

In this example k3*R_0 is the value of the resistance at dT = 0: $R3 = U_V/I(R3) = 1.05\ k\Omega$ which corresponds with the correction factor k3 = 1.05 multiplied with the nominal resistance R0 = 1 kΩ. All this can be realised at the end of the production process and possibly the data can be stored in the sensor (e.g. by means of laser fuses) and can be used in signal conditioning circuits. With a micro-

controller the change of the resistance dR and the change of the temperature dT can be evaluated as described in the following chapters.

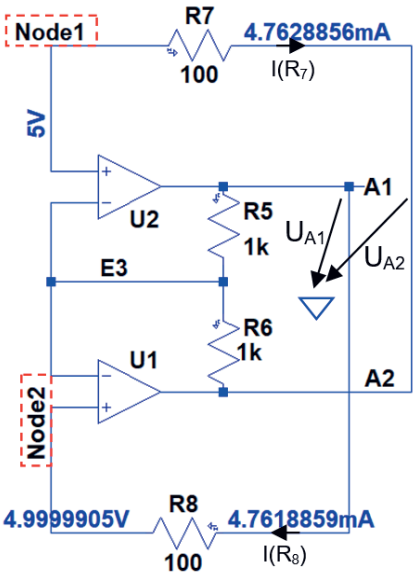


Fig. 1b: Floating zero-resistance ammeter for measuring the resistances at reference conditions (dR = 0 and dT = 0) (for further details see [1])

2. Method for simultaneously determining the change in resistances and temperature in Wheatstone bridges supplied with a voltage source

A new concept for the simultaneous determination of the change in resistance dR of sensor elements and the change in temperature dT in Wheatstone bridges is proposed. The basic idea is to consider the known temperature coefficients of all resistors used in the bridge circuit. Then the absolute change of the resistance dR and the temperature dT compared to the reference conditions (with dR = 0 and dT = 0) can be determined by using different voltages of the bridge circuit. The equations depend on the circuit, so a quarter bridge with one sensor element, a half bridge with two sensor elements or a full bridge circuit with four sensor elements have different equations. For each sensor element, the relative change of the resistance dR/R_0 is the same when the same measured variable is applied, where R_0 represents the nominal resistance at reference conditions.

2.1 Quarter bridge circuit with one sensor element and voltage source

In the quarter bridge depicted in Figure 2, the sensor element is represented by the resistor R1; R_0 is the nominal value (here: 1 kΩ) and k1 is the factor that determines the exact value of R1 at reference conditions – same is valid for k2 for R2, k3 for R3 and k4 for R4; b is the constant temperature coefficient of R1 and dR represents the absolute change of R1 due to the measured variable. The resistors R2, R3 and R4 are reference resistors with different constant temperature coefficients (TC) a and g with $a \neq g$. It is possible: a = b or g = b with b as TC of the sensor element; this means that either R3 or (R2 and R4) can be replaced by an insensitive sensor element which only reacts to temperature changes but not to changes in the measured variable. But always the temperature coefficients a and g of R3 and R4 need to be different: $a \neq g$.

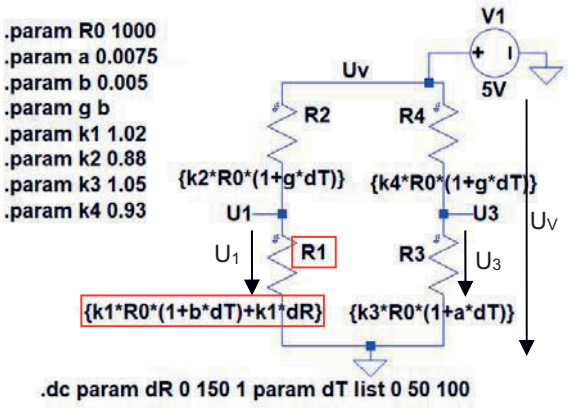


Fig. 2: Quarter bridge circuit with one sensor element (R1) and three reference resistors (R2, R3, R4)

Sensor resistance:

$$R_1 = k_1 \cdot R_0 \cdot (1 + b \cdot dT) + k_1 \cdot dR$$

Reference resistances:

$$R_2 = k_2 \cdot R_0 \cdot (1 + g \cdot dT)$$

$$R_3 = k_3 \cdot R_0 \cdot (1 + a \cdot dT)$$

$$R_4 = k_4 \cdot R_0 \cdot (1 + g \cdot dT)$$

In equation (2) the calculation of the temperature change dT is shown; equation (3) describes how the relative change of the resistance dR/R_0 of the sensor element can be derived with R_0 as nominal resistance.

Figure 3 shows the result of this calculation in a simulation with LTspice [3]. It can be seen that both the temperature change dT and the relative resistance change dR/R_0 can be derived simultaneously using equations (2) and (3).

For the change of the temperature dT equation (2) is used:

$$dT = \frac{k_3 \cdot U_V - U_3 \cdot (k_3 + k_4)}{U_3 \cdot (a \cdot k_3 + g \cdot k_4) - a \cdot k_3 \cdot U_V} \quad (2)$$

and equation (3) for the relative change dR/R_0 of the resistance of the sensor element R1:

$$\frac{dR}{R_0} = \frac{\left[k_3 \cdot U_V - U_3 \cdot (k_3 + k_4)\right]}{(U_V - U_1) \cdot k_1} \cdot \frac{\left[k_2 \cdot g \cdot U_1 - k_1 \cdot b \cdot (U_V - U_1)\right]}{\left[U_3 \cdot (a \cdot k_3 + g \cdot k_4) - a \cdot k_3 \cdot U_V\right]} - \frac{(k_1 \cdot (U_V - U_1) - k_2 \cdot U_1)}{(U_V - U_1) \cdot k_1} \cdot \frac{\left[U_3 \cdot (a \cdot k_3 + g \cdot k_4) - a \cdot k_3 \cdot U_V\right]}{\left[U_3 \cdot (a \cdot k_3 + g \cdot k_4) - a \cdot k_3 \cdot U_V\right]} \quad (3)$$

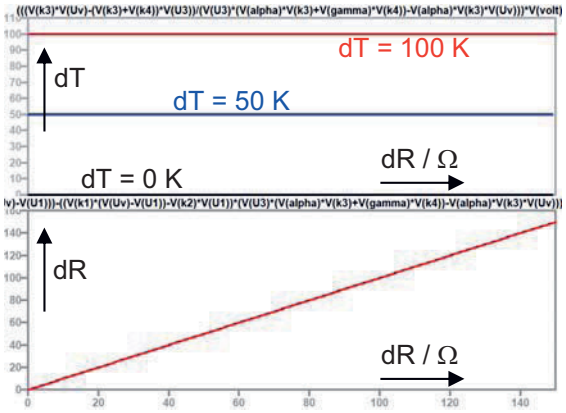


Fig. 3: Quarter bridge circuit (see Fig. 2) depending on change of R (dR) with temperature change dT as parameter: Top: Temperature change dT; Bottom: Resistance change dR

2.2 Half bridge circuit with two sensor elements and voltage source

In the half bridge shown in Figure 4 two sensor elements R1 and R4 are used. The resistors R2 and R3 represent the reference resistors with different temperature coefficients (TC) a and g: a ≠ g.

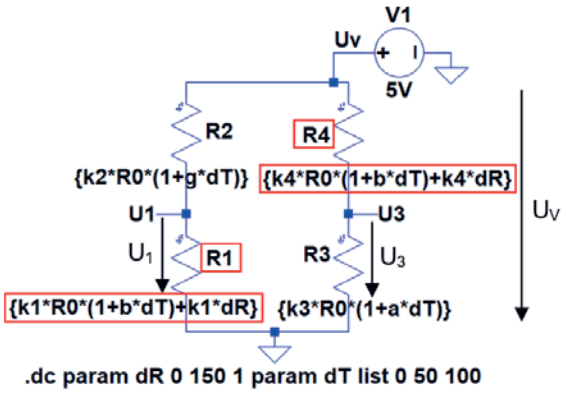


Fig. 4: Half bridge circuit with two sensor elements (R1, R4) and two reference resistors (R2, R3)

Similar to the previous chapter, dR (or dR/R₀) and dT can be calculated using equations (4) and (5). In this example the following also applies: g = b; so R2 can be realised as an insensitive sensor element. It is important that the TC of the reference resistors R2 and R3 are different: a ≠ g.

Equation (4):

$$dT = \frac{k_4 \cdot U_3 \cdot \left[k_2 \cdot U_1 - k_1 \cdot (U_V - U_1)\right] + \ldots}{k_1 \cdot (U_V - U_1) \cdot \left[k_3 \cdot a \cdot (U_V - U_3) - k_4 \cdot b \cdot U_3\right] + \ldots} \frac{-k_1 \cdot (U_V - U_1) \cdot \left[k_3 \cdot (U_V - U_3) - k_4 \cdot U_3\right]}{+k_4 \cdot U_3 \cdot \left[k_1 \cdot b \cdot (U_V - U_1) - k_2 \cdot g \cdot U_1\right]} \quad (4)$$

Equation (5):

$$\frac{dR}{R_0} = \frac{k_3 \cdot (U_V - U_3) \cdot (1 + a \cdot dT)}{k_4 \cdot U_3} - \frac{k_4 \cdot U_3 \cdot (1 + b \cdot dT)}{k_4 \cdot U_3} \quad (5)$$

The results of equations (4) and (5) are depicted in Figure 5; again, the change in temperature dT and the change in resistance dR of the sensor element can be calculated using equation (4); then the change in resistance dR can be derived using equation (5).

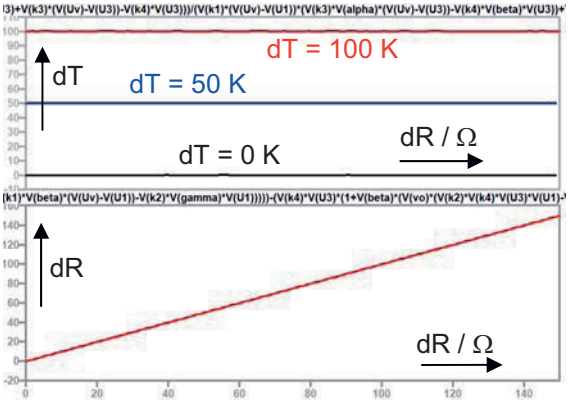


Fig. 5: Half bridge circuit (see Fig. 4) depending on change of R (dR) with temperature change dT as parameter: Top: Temperature change dT Bottom: Resistance change dR

3. Method for simultaneously determining the resistance and temperature change of Wheatstone bridges with additional resistor and voltage source

If there is an additional resistor R5 used that is connected in series to the bridge circuit (see Figure 6), it is also feasible to calculate the change of the temperature dT and the change of the resistance dR of the sensor elements. The idea is to get the information about the

total current flowing through the bridge circuit in order to calculate dT and dR. This is shown in the following chapters for quarter bridge, half bridge and full bridge Wheatstone circuit.

3.1 Quarter bridge circuit with one sensor element and additional series resistor

Figure 6 shows a quarter bridge circuit (with one sensor element R1 and three reference resistors R2, R3, R4) with an additional resistor R5. As in the previous chapters, it is possible to derive the temperature change dT and the resistance change dR (or relative change dR/R_0) of the sensor element by using voltages, temperature coefficients (TC) and resistances at reference conditions - see equations (6) und (7); the temperature change dT as a result of equation (6) is used in equation (7). Here the temperature coefficients a of the three reference resistors may be similar (without R5 they need to be different, see chapter 2.1) – the temperature coefficient a can be similar to the temperature coefficient b of the sensor element R1; therefore: a = b. If R5 is not constant, then the temperature coefficient f (with R5 = k5*R0*(1+f*dT), see Fig. 6: f ≠ a) and the temperature of R5 must be known; it is still possible: a = b. R5 may be located in close proximity to the signal conditioning circuit (e.g. microcontroller or electronic control unit); the temperature change and hence the change of R5 is then usually limited.

$$dT = \frac{\frac{k_3}{k_2}\cdot U_2 + U_3}{a\cdot\left(U_V - U_{BR}\right)\cdot\frac{k_3\cdot R_0}{R_5}} - \frac{1}{a} \quad (6)$$

, with R5 as constant resistor and equation (7):

$$\frac{dR}{R_0} = \frac{k_2}{k_1}\cdot\left(1 + a\cdot dT\right)\cdot\left(\frac{U_1}{U_{BR} - U_1}\right) - \left(1 + b\cdot dT\right)$$

In Figure 7 the result of the simulation with LTspice [3] is shown – again the change of the temperature dT and the change of the resistance dR of the sensor element can be calculated by using different voltages and temperature coefficients.

If the voltage source V1 and the resistor R5 are substituted by a current source I_5, equation (6) for changing the temperature dT is changed to equation (8); for dR equation (7) can be used.

$$dT = \frac{\frac{k_3}{k_2}\cdot U_2 + U_3}{a\cdot I_5\cdot k_3\cdot R_0} - \frac{1}{a} \quad (8)$$

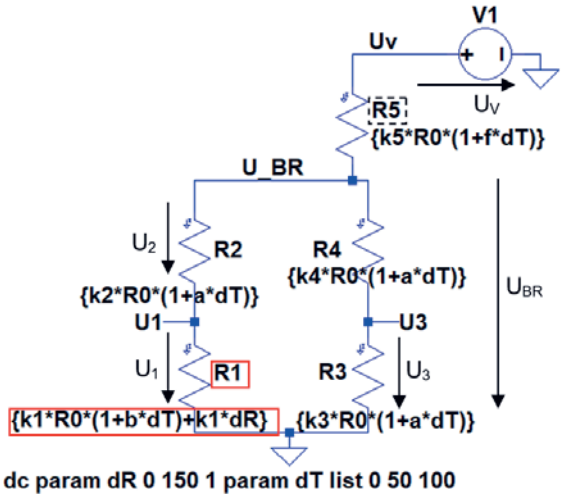


Fig. 6: Quarter bridge circuit with additional resistor R5 and a bridge circuit with one sensor element (R1) and three reference resistors (R2, R3, R4)

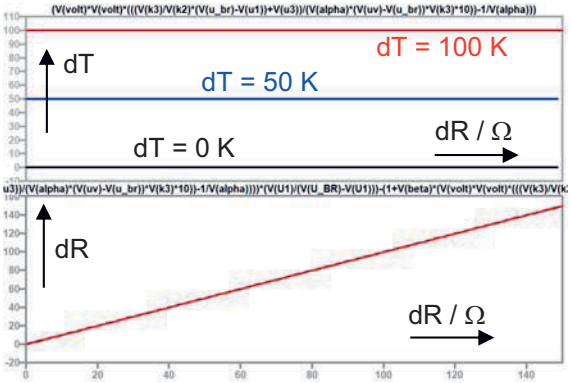


Fig. 7: Quarter bridge circuit (see Fig. 6) depending on change of R (dR) with temperature change dT as parameter: Top: Temperature change dT Bottom: Resistance change dR

3.2 Half bridge circuit with two sensor elements and additional series resistor

Figure 8 shows a half bridge circuit with two sensor elements (R1 and R4), two reference resistors (R2 and R3) and additionally the resistor R5 in series connection to the bridge circuit. The temperature coefficient f of the series resistor R5 must differ from the temperature coefficient a of R2 and R3 (hence: f ≠ a), but it is possible: a = b with b as temperature coefficient of R1 and R4.

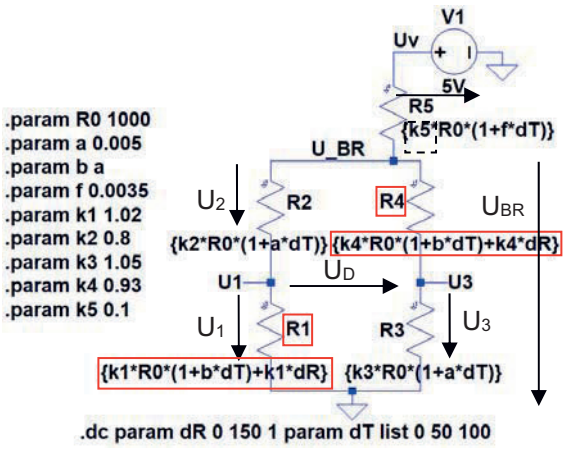


Fig. 8: Half bridge circuit with additional resistor R5 and a bridge circuit with two sensor elements (R1, R4) and two reference resistors (R2, R3)

Equation (9a) applies to the change in temperature dT under the condition that the value of R5 is constant (possible: a = b) and equation (9b) applies to dT if R5 (with f as constant temperature coefficient) is at the same temperature as the bridge circuit; dT can be used to calculate dR (or dR/R₀) in equation (10).

$$dT = \frac{\frac{k_3}{k_2}\cdot U_2 + U_3}{a\cdot\left(U_V - U_{BR}\right)\cdot\frac{k_3\cdot R_0}{R_5}} - \frac{1}{a} \qquad \text{(9a): with}$$

R5 as constant resistor
and equation (9b) with f as TC of R5:

$$dT = \frac{k_2\cdot k_3\cdot U_5 - k_2\cdot k_5\cdot\left(U_{BR}-U_D\right)}{\left[k_2\cdot k_5\cdot\left(U_{BR}-U_D\right)+k_5\cdot U_2\cdot\left(k_3-k_2\right)\right]\cdot f} \ldots \frac{-k_5\cdot\left(U_{BR}-U_1\right)\cdot\left(k_3-k_2\right)}{-k_2\cdot k_3\cdot U_5\cdot a} \qquad (9b)$$

$$\frac{dR}{R_0} = \frac{U_{BR}\cdot k_3 - U_3\cdot\left(k_3+k_4\right)}{U_3\cdot k_4} + \frac{dT\cdot\left[U_{BR}\cdot k_3\cdot a - U_3\cdot\left(a\cdot k_3 + b\cdot k_4\right)\right]}{U_3\cdot k_4} \qquad (10)$$

In Figure 9 the results of equation (9b) and (10) are depicted. The temperature change dT and the change of the sensor resistance dR can be determined again simultaneously. Similar to the previous chapter, the voltage source V1 and the resistor R5 can be replaced by a current source I5; then equation (8) is valid again.

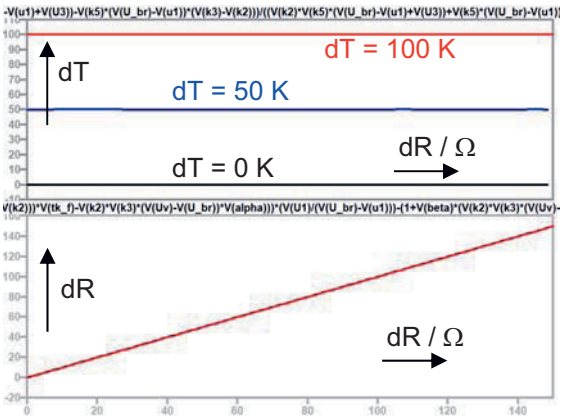


Fig. 9: Half bridge circuit (see Fig. 8) depending on change of R (dR) with temperature change dT as parameter: Top: Temperature change dT Bottom: Resistance change dR

3.3 Full bridge circuit with four sensor elements and additional series resistor

A full bridge circuit consists of four sensor elements as shown in Figure 10. Then the two sensor elements R1 and R4 must show reciprocal behaviour compared to the other two sensor elements R2 and R3, so that the resistance of the bridge circuit can be assumed to be constant - only depending on temperature change dT. Equation (11a) applies to the change in temperature dT provided that the value of R5 is constant; equation (11b) applies to dT when R5 is [R5 = k5*R₀*(1+f*dT)], where f is the constant temperature coefficient of R5 at the same temperature as the bridge circuit; f must be different to b as TC of the sensor elements (→ f ≠ b); equation (12) can be used to calculate dR or the relative change dR/R₀.

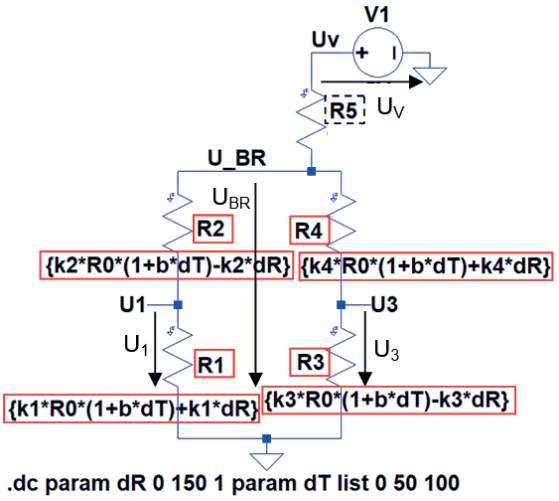


Fig. 10: Full bridge circuit with four sensor elements (R1, R2, R3, R4) and an additional resistor R5

Equation (11a) with R5 as constant results in

$$dT = \frac{U_{BR} \cdot (k_1 + k_2 + k_3 + k_4) \cdot R_5}{b \cdot (U_V - U_{BR}) \cdot R_0 \cdot (k_1 + k_2) \cdot (k_3 + k_4)} - \frac{1}{b}$$

and equation (11b) with f as TC of R5

$$dT = \frac{(U_V - U_{BR}) \cdot (k_1 + k_2) \cdot (k_3 + k_4) + \dots}{U_{BR} \cdot (k_1 + k_2 + k_3 + k_4) \cdot k_5 \cdot f + \dots}$$
$$\dots \frac{-U_{BR} \cdot (k_1 + k_2 + k_3 + k_4) \cdot k_5}{-b \cdot (U_V - U_{BR}) \cdot (k_1 + k_2) \cdot (k_3 + k_4)}$$

and equation (12):

$$\frac{dR}{R_0} = \frac{\left[U_{BR} \cdot k_1 - U_{1\cdot} \cdot (k_1 + k_2)\right] \cdot (1 + b \cdot dT)}{U_{1\cdot} \cdot (k_1 - k_2) - U_{BR} \cdot k_1}$$

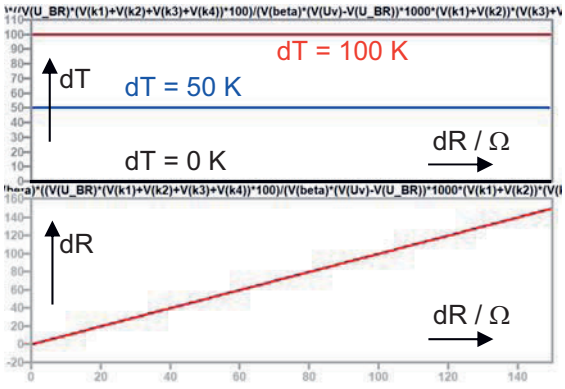


Fig. 11: Full bridge circuit (see Fig. 10) depending on change of R (dR) with temperature change dT as parameter: Top: Temperature change dT Bottom: Resistance change dR

Figure 11 shows the results for dT (equation (11a)) and for dR according to equation (12). Similar to the previous chapter, the voltage source V1 and the resistor R5 can be replaced by a current source I_5, then the equation (11c) applies instead of the equation (11a) or (11b). Equation (11c) with current source I5 instead of voltage source V1 results in:

$$dT = \frac{U_{BR} \cdot (k_1 + k_2 + k_3 + k_4)}{b \cdot I_5 \cdot R_0 \cdot (k_1 + k_2) \cdot (k_3 + k_4)} - \frac{1}{b} \quad (11c)$$

4. Method for simultaneously determining the change in resistances and temperature in a Wheatstone bridge using AMR sensor elements

This example shows the new concept with an AMR sensor element implemented in LTspice [3] as variable resistor. The model for the AMR sensor element is described in [2]. Here the temperature coefficient (TC) of the AMR sensor element is +0.003/K (relevant for R_0) and the TC of the AMR effect is -0.002/K (relevant for dR).

4.1 Quarter bridge circuit using one AMR sensor element

In Figure 12 a quarter bridge circuit with one AMR sensor element R1 is shown. The TC of R3 and R4 must be different (a ≠ g); a as TC of R3 may be equal to b as TC of the AMR sensor element (a = b), e.g. by using an insensitive AMR-sensor for R3. The change in temperature dT can be derived using equation (13) which is similar to equation (2). As depicted in Figure 13, the typical AMR characteristic (R depending on the magnetic field strength *H* in the direction of the hard axis) is influenced by the temperature change dT: The higher dT, the lower the change of the resistance dR due to the negative TC (-0.002/K) of the AMR effect.

$$dT = \frac{k_3 \cdot U_V - U_3 \cdot (k_3 + k_4)}{U_3 \cdot (a \cdot k_3 + g \cdot k_4) - a \cdot k_3 \cdot U_V} \quad (13)$$

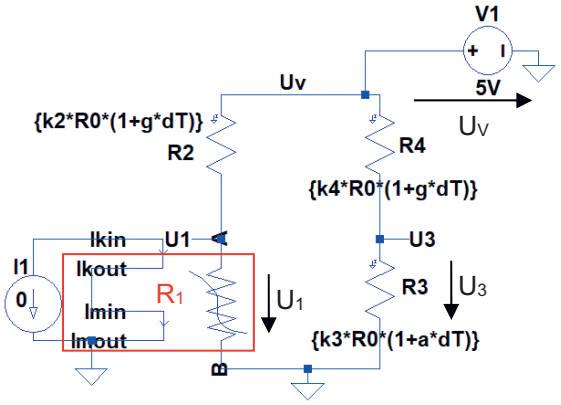


Fig. 12: Quarter bridge circuit with one sensor element (AMR-sensor R1) and three reference resistors

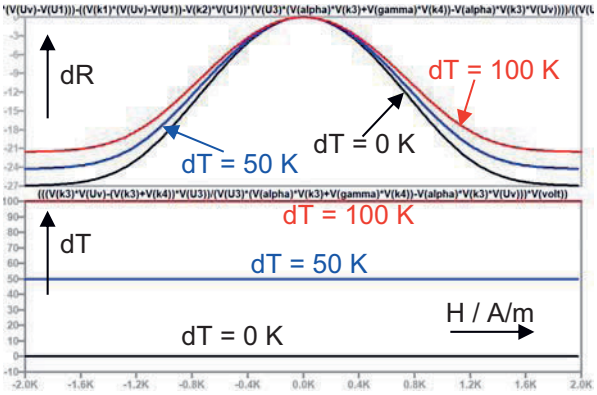


Fig. 13: Quarter bridge circuit (see Fig. 12) depending on applied magnetic field strength H with temperature change dT as parameter. Top: Resistance change dR; Bottom: Temperature change dT

4.2 Full bridge circuit using four AMR sensor elements

In Figure 14 a full bridge circuit with four AMR-sensor elements R1 to R4 is depicted. The resistors R1 and R4 must exhibit opposite behaviour to R2 and R3, which can be achieved by barber pole structures. In Figure 15 the results of the simulation is depicted: The temperature change dT (see equation (14)), the resistance change dR and also the output voltage U_D can be determined simultaneously using the corresponding equations.

Equation (14) is valid for the full bridge circuit under the condition that R5 is constant (see equation (11a)):

$$dT = \frac{U_{BR} \cdot (k_1 + k_2 + k_3 + k_4) \cdot R_5}{b \cdot (U_V - U_{BR}) \cdot R_0 \cdot (k_1 + k_2) \cdot (k_3 + k_4)} - \frac{1}{b}$$

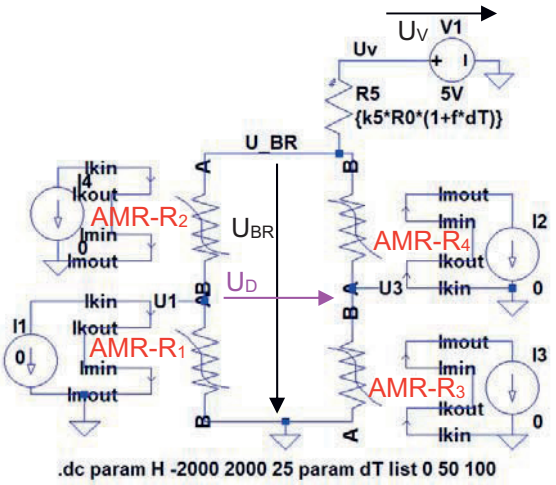


Fig. 14: Full bridge circuit with four sensor elements (AMR-sensors R1-R4)

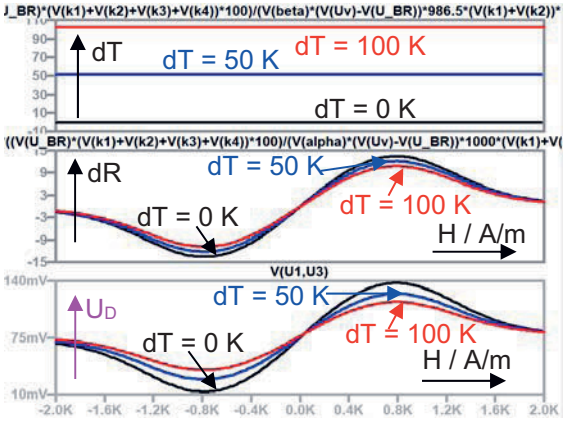


Fig. 15: Full bridge circuit (see Fig. 14) depending on applied magnetic field strength H with temperature change dT as parameter. Top: Resistance change dR; middle: Temperature change dT; bottom: Voltage U_D

5. Conclusion

A new concept is described which allows the simultaneous measurement of the change in resistance dR and the change in temperature dT. This is achieved by using the applied voltages and temperature coefficients of all resistors in the Wheatstone bridge. This concept is applicable to quarter, half and full bridge circuits with different numbers of sensor elements. It is shown how the relevant information (resistors R1 to R4) is determined in an already manufactured Wheatstone bridge circuit with an operational amplifier circuit. A possible offset of the bridge circuit due to a mismatch of the four resistors is not relevant. The additional information about the temperature can help to obtain more accurate results. The bridge circuit can not only be used for compensation of the temperature influence, but also as a temperature sensor for condition monitoring or diagnostic functions in the application. Thus, the bridge circuit can provide two different information simultaneously: resistance change dR and the temperature change dT. In this example, the temperature coefficient is assumed to be constant - if this is not the case, the calculation can be performed using a numerical approach. A patent application has already been filed.

References

[1] Tietze, U.; Schenk, Ch.: Halbleiterschaltungstechnik, 9. Auflage, Springer-Verlag, Berlin Heidelberg, 1990, S. 864f

[2] Gintner, K.: Ein Sensor auf Basis des anisotropen magnetoresistiven Effekts, Dissertation, Lehrstuhl für Technische Elektronik, Friedrich-Alexander-Universität Erlangen-Nürnberg, 1999

[3] Engelhardt, M.: LTspice XVII(x64), 2018, Linear Technology Corporation

DOI 10.5162/sensoren2019/P2.9

Comparison of deep feature extraction techniques for varying-length time series from an industrial piercing press

Christian Thiel[1,2], Carolin Steidl[2], Bernd Henning[2]
[1] BENTELER Steel/Tube GmbH, Residenzstraße 1, 33104 Paderborn, Germany
[2] Measurement Engineering Group, Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany

Abstract

The continuous refinement of sensor technologies enables the manufacturing industry to capture increasing amounts of data during the production process. As processes take time to complete, sensors register large amounts of time-series-like data for each product. In order to make this data usable, a feature extraction is mandatory. In this work, we discuss and evaluate different network architectures, input pre-processing and cost functions regarding, among other aspects, their suitability for time series of different lengths.

Keywords: Feature Extraction, Dynamic Time Warping, Masking, Neural Networks

1 Introduction

Companies in the manufacturing industry measure increasing amounts of product and process data in order to improve their processes. This data is used to stabilize production, identify factors on the product quality and improve process efficiency. As processes take time to complete, sensors register large amounts of time-series-like data for each product. In contrast to single-value features measured once for each product, temporally or locally resolved measurements provide much richer information about the underlying process. The high complexity and dimensionality of these time series of different lengths point towards the necessity of Feature Extraction (FE) methods.
Traditionally, domain experts would design handcrafted composite features for each time series. This process is prone to errors and unlikely to capture every detail. Recent advances in deep learning have allowed various Neural-Network(NN)-based auto encoding FE techniques to emerge, allowing FE in an objective unified way.
These new architectures include some, that can natively handle varying length inputs. Those architectures include Recurrent Neural Networks (RNNs), which we discuss in section 4. Besides RNNs, also Fully Convolutional Networks are suited for varying length inputs [7]. However, their hidden dimensionality is not constant, so that less features are produced for shorter inputs. As this simply shifts the problem of unequal dimensionality to the next step of analysis, we do not evaluate these methods further.
For common NN-based Autoencoders (AE) however, handling varying-length time series still poses major problems, as only fixed-dimensional inputs are natively supported. While the use of zeropadding or interpolation already enables the use of common AE, we show that naive application of these methods is not ideal as new patterns are introduced to the original time series. In this work we discuss and compare two possibilities to overcome this shortcommming:

1. Zeropadding introduces new patterns and discontinuities to time series. We propose the use of masking to avoid training on these artificially created patterns in subsection 3.1.
2. Interpolation stretches and compresses patterns in time series. This effect can be mitigated using elastic cost functions such as Dynamic Time Warping (DTW) as discussed in subsection 3.2.

The outline of the paper is as follows: In section 2 the process and data preprocessing is discussed. Our proposed modifications to common AE architectures are then motivated and presented in section 3. With RNNs a more complex network architecture is discussed in section 4. Finally our experiments are presented in section 5.

2 Process and Data Description

In the tube production process a piercing press is used to form a hole in a solid block of hot metal. The piercing step is crucial in production and strongly influences many critical quality factors such as the eccentricity of the final tube. Thus, data from the piercing process is of importance in many analysis tasks.
As the process duration of the piercing press depends on the size of the block as well as the chosen process speed, signals resulting from equidistant measurements in time inherently have different

lengths. In this work we evaluate all approaches on 100000 instances of pressing force time series from the tube production process. The sampling rate for all time series is 20 milliseconds. The resulting time series have lengths between 156 and 258 values.
Signals from industrial processes often contain outliers due to the rough environments. In order to be able to use non-robust estimators such as means or standard deviations, a Hampel-Filter is applied to each time series during preprocessing. All identified outliers are then replaced using interpolation.
The importance of time series normalization when measuring time series similarities has been shown multiple times in the literature [6]. When using DTW or the Mean Squared Error (MSE) it is evident that already a small offset between two time series leads to potentially large distances. This problem becomes more significant, the longer the time series gets. Usually, each time series is thus normalized to zero mean and unit variance. For the piercing press data used in this work we deviate from this approach slightly and use trimmed mean and standard deviation as more robust estimators for normalization. This mainly results in the normalization focusing on the center part of the signal without taking edge effects such as very low pressing forces at the start or end into account.

3 Adjustments for common architectures

Common AE based techniques, as for example Deep Convolutional Autoencoders (DCAE) and Fully Connected Autoencoders (FCAE), only support fixed-dimensional inputs [4]. This restriction originates from the matrix multiplication in Fully-Connected (FC) Layers.
In order to use DCAE and other architectures that require a fixed input dimensionality, simply equalizing the length of time-series prior to FE naturally comes to mind. In this work we consider zero-padding and interpolation as two means to equalize the lengths of time series. While this already enables the use of powerful techniques such as the FCAE and DCAE, the results might not be perfect. We provide two intuitive explanations that support this assumption:

I. Zeropadding introduces new patterns to the time series. This does not only relate to the perfectly straight line in the zeropadded part of the signal, but especially to the transition from signal end to zero. Even for standardized time series with zero mean and unit variance significant discontinuities can occur as seen in the curve on top of Figure 1. Even small inaccuracies of the discontinuities position in the AE's reconstruction yield high losses. Thus the AE spents a reasonable amount of information in the latent representation to exactly model the discontinuity while performance in meaningful areas suffers.

II. Interpolation causes characteristic local patterns in the original time series to be stretched or compressed depending on the difference of time series length to chosen input dimensionality. These alterations of the original pattern can not be matched as good as the original pattern by a single filter mask. Thus either more filters representing different alterations of the same original pattern are required or the overall performance can be expected to drop.

We propose the use of masking when zeropadding is used to counteract I in subsection 3.1. Also Dynamic Time Warping (DTW) based losses in combination with interpolation to counteract II are discussed in subsection 3.2.

3.1 Masking layer

Masking is already used to simplify training of RNNs with inputs of varying dimensionality. It is implemented for RNNs in popular frameworks such as TensorFlow (`tf.keras.layers.masking`).
We propose to use masking in combination with zeropadding also for common architectures such as DCAE. In our masking layer, we simply force all values in the padded part of the reconstructed signal to be zero before calculating the loss and performing back propagation. Errors in the reconstruction of the padded parts as well as the artificially introduced discontinuity no longer contribute to the cost function and thus do not affect training. The overall DCAE model including the masking layer is shown in Figure 1.

3.2 Dynamic Time Warping as loss

Dynamic Time Warping (DTW) is widely used as an elastic similarity measure in time series analysis. While DTW is already able to compare time series of unequal lengths, this feature has no relevance for common architectures as the output always has a fixed dimensionality so that unequal lengths never occur. However, as an elastic measure, DTW is very well suited to compensate for stretched or compressed patterns introduced by interpolation. Following, we firstly give a brief introduction to the classic DTW approach in subsubsection 3.2.1, as we have already shown its suitability for our data in [10]. As this classic DTW approach is not differentiable, it cannot be used for training using backpropagation. Secondly we introduce the recently published soft-DTW algorithm in subsubsection 3.2.2, which is a differentiable modification of DTW [2].

3.2.1 Classic Dynamic Time Warping

To calculate the DTW distance, firstly the signal difference matrix $\mathbf{S}$ containing all pairwise distances of two time series $\vec{x} = \{x_1, \ldots, x_M\}$ and $\vec{y} = \{y_1, \ldots, y_N\}$:

$$\mathbf{S}_{m,n} = d(x_m, y_n) \quad \forall\, m = 1, \ldots, M \quad n = 1, \ldots, N, \quad (1)$$

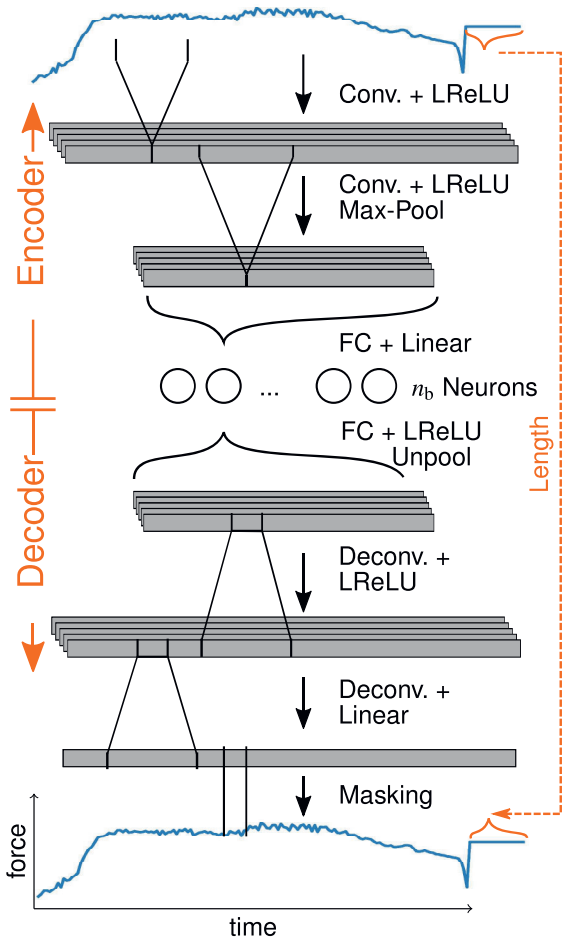


Figure 1: *Deep Convolutional Autoencoder (DCAE) with additional masking to avoid training on zero-padded parts of the input signal.*

where d is usually the Euclidean Distance (ED). For brevity we denote $d(x_m, y_n)$ as $d_{m,n}$. The DTW alignment of the series is defined by a warping function $\mathcal{W} = \{\vec{w}_1, ..., \vec{w}_K\}$ through $\mathbf{S}$, which results in the lowest cumulated (weighted) distances:

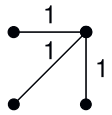
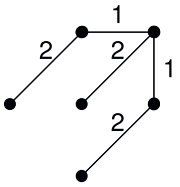
$$\mathcal{W}^* = \underset{\vec{w}_1,...,\vec{w}_K}{\operatorname{argmin}} \sum_{k=1}^{K} d(x_{w_{k,1}}, y_{w_{k,2}}), \quad \vec{w}_k = \{w_{k,1}, w_{k,2}\}, \quad (2)$$

where each warping function has to fullfill the well-known conditions [8]

1. Monotony: $w_{k,1} \geq w_{k-1,1}$ and $w_{k,2} \geq w_{k-1,2}$
2. Continuity: $w_{k,1} - w_{k-1,1} \leq 1$ and $w_{k,2} - w_{k-1,2} \leq 1$
3. Boundary conditions: $w_1 = (1,1)$ and $w_K = (M,N)$

These conditions allow a reformulation as a Dynamic Programming (DP) algorithm using Bellman's equation through a cumulated signal difference matrix $\tilde{\mathbf{S}}$

Table 1: *Constraining the slope of DTW warping functions using stepping patterns [8].*

Formula for $\tilde{\mathbf{S}}_{m,n} =$	Schema
Symmetric-P0 (SymP0)	
$\min \begin{bmatrix} \tilde{\mathbf{S}}_{m,n-1} + d_{m,n} \\ \tilde{\mathbf{S}}_{m-1,n-1} + d_{m,n} \\ \tilde{\mathbf{S}}_{m-1,n} + d_{m,n} \end{bmatrix}$	
Symmetric-P1 (SymP1)	
$\min \begin{bmatrix} \tilde{\mathbf{S}}_{m-1,n-2} + 2d_{m,n-1} + d_{m,n} \\ \tilde{\mathbf{S}}_{m-1,n-1} + 2d_{m,n} \\ \tilde{\mathbf{S}}_{m-2,n-1} + 2d_{m-1,n} + d_{m,n} \end{bmatrix}$	

calculated as follows:

$$\tilde{\mathbf{S}}_{0,0} = 0, \quad \tilde{\mathbf{S}}_{m,0} = \infty \; \forall \; m > 0, \quad \tilde{\mathbf{S}}_{0,n} = \infty \; \forall \; n > 0 \quad (3)$$

$$\tilde{\mathbf{S}}_{i,j} = \min \begin{bmatrix} \tilde{\mathbf{S}}_{m,n-1} + d_{m,n} \\ \tilde{\mathbf{S}}_{m-1,n-1} + d_{m,n} \\ \tilde{\mathbf{S}}_{m-1,n} + d_{m,n} \end{bmatrix}, \quad (4)$$

where Equation 4 is known as the stepping pattern. The commonly used unrestricted Symmetric-P0 stepping pattern can easily lead to unreasonable mappings for real world data, as it allows arbitrary many points of one time series to map to just one point of the other. The original authors of DTW proposed the use of more robust steppings such as Symmetric-P1, which restrict the slope of the warping function and thus prohibits degenerated alignments [8, 10]. A visualization of the Symmetric-P1 in comparison to the default Symmetric-P0 stepping is given in Table 1. Finally, the overall DTW cost is equivalent to the corner element $\tilde{\mathbf{S}}_{M,N}$ of the cumulated signal difference matrix. It is equivalent to the cumulated costs of the optimal warping function $\mathcal{W}^*$:

$$\mathbf{dtw}(\vec{x}, \vec{y}) = \tilde{\mathbf{S}}_{M,N} = \min_{\vec{w}_1,...,\vec{w}_K} \sum_{k=1}^{K} d(x_{w_{k,1}}, y_{w_{k,2}}) \quad (5)$$

3.2.2 Differentiable Dynamic Time Warping

DTW, as presented in subsection 3.2.1, is not continuously differentiable due to the minimum operator in Equation 2 and Equation 4 respectively. While for a fixed warping path a derivative with respect to one of the time series can easily be calculated, the minimum operator leads to discontinuities for values of $\vec{x}$, where small changes result in a different warping path. Following the idea of the Global Alignment Kernel (GAK) introduced in 2011 [1], γ-soft-DTW considers the set

of all possible alignments $\mathcal{W}_{M,N}$ instead of only the one alignment resulting in minimal costs $\mathcal{W}^*$ [2].
In order to incorporate all warping paths, the authors introduce a generalized soft minimum operator with smoothing parameter $\gamma \geq 0$ [2]:

$$\min^{\gamma}\{a_1,\ldots,a_n\} := \begin{cases} \min_{i\leq n} a_i, & \gamma = 0, \\ -\gamma \log \sum_{i=1}^{n} e^{-a_i/\gamma} & \gamma > 0. \end{cases}, \tag{6}$$

which for $\gamma = 0$ results in a regular minimum operator and for $\gamma > 0$ in a soft-minimum incorporating all values in the given set. With the generalized soft minimum defined, γ-soft-DTW follows as:

$$\mathbf{dtw}_{\gamma}(\vec{x},\vec{y}) := \min^{\gamma}\left\{\sum_{k=1}^{K} d(x_{w_{k,1}}, y_{w_{k,2}}), \{\vec{w}_1,\ldots\vec{w}_K\} \in \mathcal{W}_{M,N}\right\}, \tag{7}$$

where $\mathcal{W}_{M,N}$ denotes the set of all possible alignment paths for two signals of lengths M and N. For the reformulation as a DP algorithm, the minimum operator of Equation 4 is simply replaced by the soft minimum operator with the corresponding parameter γ.
Finally, for $\gamma > 0$, Equation 7 can be explicitly differentiated as shown in [2]. This enables its use as a loss in any gradient based setting, such as backpropagation in NN.

4 Recurrent Network Architectures

RNNs are inherently able to handle varying lengths inputs. In contrast to common architectures where the whole input sequence is fed to the network simultaneously, RNNs are fed sequentially. In a time series setting, one value of the time series is fed after the other. RNNs are obtained by using neurons outputs of one time step as inputs of the next. In bidirectional RNNs also the outputs from future time steps are used as inputs, at the cost of loosing causality [9]. Additionally, memory blocks are often introduced, to capture dependencies over time explicitly.
To train RNNs, gradients do not only have to be propagated through the network itself but also through all time steps. The algorithm used for training is referred to as Backpropagation-Through-Time (BPTT). Due to the unfolding in time, RNNs can be seen as very deep networks, although with strong parameter sharing. RNNs high depth leads to vanishing and exploding gradient problems. Multiple techniques have been introduced to counteract these problems. One architecture that gained significant attention are Long-Short Term Memory Networks (LSTMs) [5].
For FE, the encoder is realized as a Bidirectional LSTM (Bi-LSTM). To obtain a fixed dimensional representation for the varying length inputs, the last outputs of the deepest LSTM layers are interpreted as features. The concatenated outputs of the forward and backward models is the overall representation of the input time series in the feature space. For the decoder, this representation is repeated to match the length of the input time series, and then fed to the decoding forward and backwards models in each timestep. The final prediction for a timestep is obtained by collecting the outputs of the deepest layer of the forward and backward LSTM respectively, and feeding them into a single neuron with linear activation. The weights of the neuron is identical for every timestep. The number of steps for which the decoder is run is equivalent to the length of the input.

Table 2: *Model Architectures. Encoder architecture is abbreviated as EA. The decoder is build transposed to the encoder without weight sharing. Activations before output and feature representation are always linear. Convolutional Layers are denoted as Conv(Filters, Filtersize, Stride).*

Model	Architecture
Bi-LSTM	Activation: tanh Recurrent Activation: hard sigmoid EA: $2\text{x}50 \rightarrow 2\text{x}50 \rightarrow 2\text{x}5$ Feature Dimension: 10 Number of Parameters: 170,981
DCAE	Activation: Leaky-ReLU EA: $\text{Conv}(8,6,2) \rightarrow \text{Conv}(16,6,2) \rightarrow 10$ Feature Dimension: 10 Pooling: Not used. Number of Parameters: 23,515
FCAE	Activation: Leaky-ReLU EA: $200 \rightarrow 200 \rightarrow 10$ Feature Dimension: 10 Number of Parameters: 169,020

5 Experiments

In order to validate our proposed adjustments from section 3, experiments on zeropadded (Experiment 1) and interpolated (Experiment 2) signals are performed. Results for the RNN are obtained using the original (normalized) signals without zeropadding or interpolation.

5.1 Experiment Description

Besides experiment specific performance measures discussed below, also epochs until convergence and training time per epoch are taken into account. In this work we define convergence as the last epoch which lead to an improvement of more than one percent compared to the previously best result. If no such improvement occurs for more than 5 epochs, the training is stopped. All architectures used for the experiments are shown in Table 2. All models are trained using the Adamax optimizer with a batch size of 200.
Experiment 1: In order to asses the influence of masking for zeropadded signals, a comparable metric for models trained with and without masking is

required. We use the mean squared error (MSE) between the prediction and the true signal in the non-zeropadded part of the signal. We motivate this twofold: Firstly, the accuracy of the zeropadded part is simply stated irrelevant. Secondly, as discussed in subsection 3.1, the length of the signal can be seen as an additional feature and is thus known. Using this knowledge any error in the reconstruction of the zeropadded part can be rectified.

Experiment 2: Strict dynamic time warping using stepping restrictions and warping windows result in reasonable alignments for pressing force signals [10]. However, neither the score of classic DTW using other steppings than SymP0 nor the soft DTW score can be interpreted easily as some *average error of the prediction in the original domain of the data*, following the MSE notion. In order to obtain such interpretable errors, we use the resulting MSE of the *aligned* signals. The alignment is obtained using classic DTW with SymP1 steppings and a Sakoe-Chiba bands of width 30. We denote the resulting metric as $\mathrm{MSE}_{\mathrm{DTW}}$. During training we use the soft-DTW implementation by the original authors written as a C-Extension for Python.

5.2 Results

The results of both experiments and the reference Bi-LSTM are shown in Table 3.

Experiment 1: For both, the DCAE and FCAE masking leads to overall better results. The convergence time of both models is significantly lower and additionally also the overall MSE reduced. These results directly support assumption I from section 3. While training time increases when masking is used, the overall training is still extremely fast.

Experiment 2: Using soft-DTW in training leads to significantly lower DTW losses, even for strict steppings and warp path regulations which were not used during training. This shows that soft-DTW is a suitable replacement in training for non-differentiable strict DTW. Please note that as described in subsection 5.1 $\mathrm{MSE}_{\mathrm{DTW}}$ follows the notion of average error per point in the original data domain, and is thus directly comparable to the MSE. Supporting assumption II, $\mathrm{MSE}_{\mathrm{DTW}}$ is consistently lower when soft-DTW is used for training. The significant increase in training time results from the high complexity of soft-DTW and even supersedes the slow training of RNNs.

Table 3: *Results of the experiments. Columns: Proposed Method Applied (P), Epochs until Convergence (EC), training Time per Epoch (TE), Mean Squared Error (MSE), Mean Squared Error after constrained DTW alignment ($\mathrm{MSE}_{\mathrm{DTW}}$). The proposed method in Experiment 1 is masking, in Experiment 2 training with soft-DTW.*

Model	P	EC	TE	MSE	$\mathrm{MSE}_{\mathrm{DTW}}$
Experiment 1 (zeropadded data)					
DCAE	yes	19	2 s	**0.10**	
DCAE	no	30	1,7 s	0.12	
FCAE	yes	**12**	1,4 s	0.14	
FCAE	no	26	**1,1 s**	0.18	
Experiment 2 (interpolated data)					
DCAE	yes	29	25 min	0.19	**0.08**
DCAE	no	20	1,7 s	**0.13**	0.10
FCAE	yes	18	25 min	0.20	0.10
FCAE	no	**8**	**1,1 s**	0.15	0.12
RNN					
Bi-LSTM	no	8	16 min	0.14	0.12

6 Literature

[1] M. Cuturi: Fast global alignment kernels, *Proceedings of the 28th international conference on machine learning (ICML-11)*, 2011, pp. 929-936.

[2] M. Cuturi, M. Blondel: Soft-DTW: a Differentiable Loss Function for Time-Series, *Proceedings of the 34th International Conference on Machine Learning*, Vol. 70, Mar. 2017 .

[3] H. A. Dau, E. Keogh, K. Kamgar, C. Yeh, Y. Zhu, G. Shaghayegh, C. Ratanamahatana, A. Chotirat, Yanping and B. Hu: The UCR Time Series Classification Archive, Oct. 2018.

[4] J. Geng, J. Fan, H. Wang, X. Ma, B. Li and F. Chen: High-Resolution SAR Image Classification via Deep Convolutional Autoencoders, *Proceedings of the IEEE conference on Geoscience and Remote Sensing Letters*, 2015.

[5] S. Hochreiter and J. Schmidhuber: Long short-term memory, *Neural computation*, 9.8, 1997, pp. 1735-1780.

[6] E. Keogh, S. Kasettey: On the Need for Time Series Data Mining Benchmarks: A Survey and Empirical Demonstration, *Data Mining and Knowledge Discovery*, Vol. 7, Oct. 2003.

[7] J. Long, E. Shelhamer and T. Darrell: Fully Convolutional Networks for Semantic Segmentation, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.

[8] H. Sakoe and S. Chiba: Dynamic programming algorithm optimization for spoken word recognition, IEEE transactions on acoustics, speech, and signal processing, Vol. 26, No. 1, 1978, pp. 43-49

[9] M. Schuster and K. Paliwal, Bidirectional recurrent neural networks, *IEEE Transactions on Signal Processing*, Vol. 45, No. 11, 1997, pp. 2673–2681.

[10] C. Thiel, N. Feldmann and B. Henning: Extraction of Interpretable Features from Temporal Measurements using Approximate Prototypes, *Sensors and Measuring Systems*, 19th ITG/GMA-Symposium, Nuremberg, Germany, 2018, pp. 536-540.

DOI 10.5162/sensoren2019/P2.10

Entwicklung eines taktilen Mikrotaster-Messsystems für Hochgeschwindigkeitsmessung von Form, Rauheit und mechanischen Eigenschaften

Michael Fahrbach, Erwin Peiner
Institut für Halbleitertechnik (IHT) und Laboratory for Emerging Nanometrology (LENA), Technische Universität Braunschweig, Braunschweig, Deutschland
m.fahrbach@tu-braunschweig.de, e.peiner@tu-braunschweig.de

Zusammenfassung

Im Rahmen eines europäischen EMPIR-Projekts soll das Messprinzip der Kontaktresonanzspektroskopie auf 5 mm × 200 µm × 50 µm ($L \times B \times D$) große, piezoresistive Tastsensoren angewandt werden. Dieses Messverfahren ist im Bereich der Rasterkraftmikroskopie bereits etabliert. Da die hier verwendeten Sensoren jedoch deutlich größer sind als im Bereich der Rasterkraftmikroskopie üblich, sind einige der dort geltenden Vereinfachungen nicht mehr anwendbar.
In dieser Arbeit werden Verfahren vorgestellt, die die Auswertung von Messwerten der verwendeten Sensoren verbessern, indem Schwingungsmodelle systematisch angepasst werden. Dazu gehören die Berücksichtigung der Luftdämpfung, von frequenzabhängigem Verhalten der Anregung sowie Nichtlinearitäten der Messbrücke.

Keywords: Mikrotaster, Cantilever-Sensor, Kontaktresonanzspektroskopie, Modellierung, Messdatenauswertung

Einleitung

Im Zuge der fortwährenden Digitalisierung der industriellen Fertigung werden neuartige Sensoren und Messverfahren benötigt, die Hochgeschwindigkeitsmessungen von Form, Rauheit und mechanischen Eigenschaften an Werkstücken und Werkzeugmaschinen während der Produktion ermöglichen [1]. Taktile Mikrotaster zeigen hier viel Potential, da sie genutzt werden können, um Oberflächen mit Geschwindigkeiten bis zu 15 mm/s abzutasten [2]. Um solche Sensoren weiterzuentwickeln, wird ein europäisches EMPIR-Projekt finanziert [3]. Eines der Ziele dieses Projekts ist die Entwicklung eines Hochgeschwindigkeits-Messsystems zur Messung der oben genannten Größen.
Bei den hier verwendeten Sensoren handelt es sich um Silizium-Cantilever-Sensoren mit einer integrierten piezoresistiven Wheatstone'schen Messbrücke und einer Tastspitze nahe ihrem freien Ende der Firma CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt vom Typ CAN50-2-5 [4]. Die geometrischen Abmessungen dieser Sensoren sind in Tabelle 1 aufgeführt.

Tab. 1: Nominelle Geometrie des Mikrotasters.

Parameter	Symbol	Nomineller Wert
Länge	L	5 mm
Position der Spitze	L_1	4,9 mm
Breite	w	200 µm
Dicke	b	50 µm

Messaufbau und Theorie

Wie in Abbildung 1 dargestellt, werden die Sensoren auf sogenannten Mikrotasterträgern montiert.

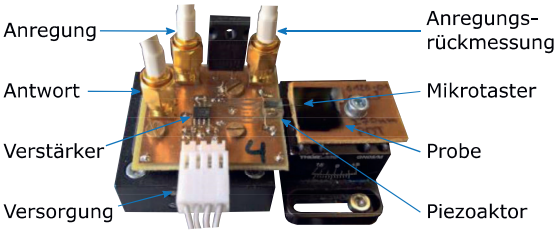


Abb. 1: Übersicht eines Mikrotasterträgers neben einer zu untersuchenden Probe.

Dabei handelt es sich um kleine Platinen, die einen Piezoaktor (Miniatur-Piezoaktor,

5 × 5 × 2 mm³, PL 055.30 PICMA®Chip Actuator, www.piceramic.de) zur dynamischen Anregung des Tasters, und einen Vorverstärker (Chip-Vorverstärker, AD8421, Analog Devices Inc.) zur Messung der Ausgangsspannung der Wheatstone'schen Messbrücke bereitstellen. Zur Anbindung an eine Messelektronik sind drei SMA-Buchsen vorgesehen, über die alle nötigen Signale abgegriffen werden können. Zusätzlich ist eine Pinleiste zur Spannungsversorgung der Bauteile verbaut.
Der Mikrotasterträger ist auf einem Positioniersystem, dargestellt in Abbildung 2, montiert. So kann er in der *xy*-Ebene über Positionierschrauben und in der *yz*-Ebene über Piezotische (PI P-518.ZCD und PI P-621.1CD, Physik Instrumente (PI) GmbH & Co.KG, www.physikinstrumente.de) verfahren werden. Neben dem Mikrotasterträger angeordnet kann eine Probe auf einem *xyz*-Positionierer angebracht werden.

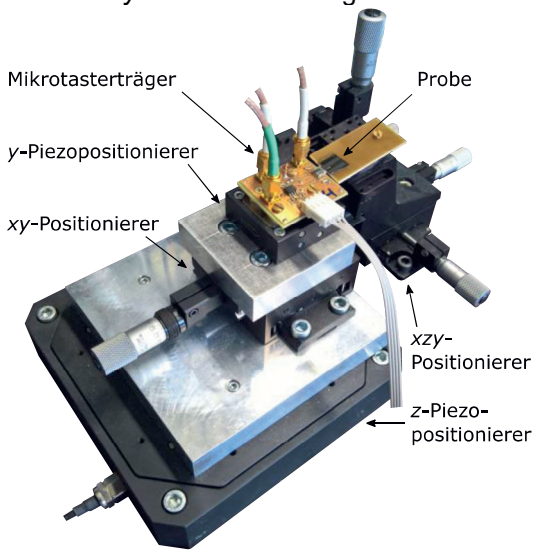


Abb. 2: Messtisch mit Mikrotasterträger und Probe. Die Spitze des Mikrotasters ist in Kontakt mit der Probe, die auf einem weiteren Positionierer befestigt ist.

Während einer Messung wird die Probe zunächst über die Positionierschrauben an den Taster herangeführt. Daraufhin wird der Taster mit dem *z*-Piezotisch an die Probe angenähert, bis Kontakt zwischen Tastspitze und Probenoberfläche hergestellt ist und der Cantilever beginnt sich zu biegen. Diese Bewegung wird so lange fortgeführt, bis die Kraft, die der Taster, entsprechend dem Produkt seiner Auslenkung und Steifigkeit auf die Probe ausübt, einen vorgegebenen Wert erreicht hat. Sobald dies geschehen ist, wird der Mikrotaster an seiner Basis durch den verbauten Piezoaktor in eine out-of-plane-Biegeeigenschwingung angeregt.
Wie in Abbildung 3 dargestellt interagiert der Taster mit der Probenoberfläche, wodurch die Resonanzfrequenzen und -amplituden, im Vergleich zum freischwingenden Sensor, zu höheren Werten verschoben bzw. verringert werden.

Modellierung des gemessenen Verhaltens

In [5, S. 55-58] stellt Rabe ein Modell auf, das das Verhalten von AFM-Sensoren in Kontakt mit Probenoberflächen unter Annahme verschiedener Randbedingungen und Vereinfachungen beschreibt. In ihm werden Länge, Breite, Dicke, Dichte und Elastizitätsmodul des Cantilevers, Position, Masse und Massenschwerpunkt der Tastspitze, Kontaktsteifigkeit und Dämpfung sowohl senkrecht als auch parallel zur Oberfläche sowie der Winkel zwischen Cantilever und Probenoberfläche berücksichtigt.

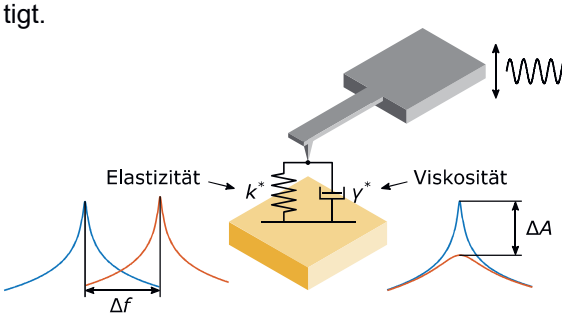


Abb. 3: Messprinzip der Kontaktresonanzspektroskopie. Der Sensor wechselwirkt mit der Probe, wodurch Resonanzfrequenz und -amplitude zu höheren Werten verschoben (links) bzw. verringert (rechts) werden (nach [6]).

Dieses Model wird nun an den Messaufbau und die Sensoren dieser Arbeit angepasst. So ist hier die Abweichung von einer senkrecht zur Probenoberfläche wirkenden Antastkraft, d. h. der Winkel zwischen Sensor- und Probenoberfläche mit ca. 1° vernachlässigbar gering. In diesem Fall hat das Feder-Dämpfer-Glied parallel zur Oberfläche keine Auswirkung auf die Schwingung und kann somit vernachlässigt werden. Stattdessen muss aber die Luftdämpfung berücksichtigt werden, da sie sich drastisch auf die Resonanzgüte auswirkt [7] und somit zur Auswertung gemessener Amplituden unabdingbar ist. Des Weiteren wird entsprechend des Messaufbaus die Schwingungsanregung an der Basis des Sensors in das Modell aufgenommen.
Das Ergebnis ist ein Schwingungsmodell, das auf Grundlage physikalischer Zusammenhänge das Verhalten aller out-of-plane-Schwingungen des Mikrotasters beschreibt. Wird dieses Modell nun durch die Sensitivitäten des Sensors und des Piezoaktors aus den jeweiligen Datenblättern erweitert und mit Messungen verglichen, zeigen sich schnell einige Probleme: Einerseits

kann das Modell die gemessenen Resonanzamplituden ohne Kenntnis der Luftdämpfung nicht nachbilden, andererseits können gemessene Resonanzfrequenzen, sowohl freischwingend als auch in Kontakt mit der Probe, durch das Modell nicht exakt reproduziert werden. Es ist somit nötig, so viele Parameter des Modells wie möglich zu überprüfen und anhand von Referenzmessungen so anzupassen, dass das Modell ein gemessenes Verhalten möglichst vollständig wiedergibt.

Kalibrierung des Schwingungsmodells

Die erste zu überprüfende Größe ist die Empfindlichkeit des Sensors. Diese wird gemessen, indem der Sensor quasistatisch gegen einen harten, nicht verformbaren Probenkörper, wie z. B. einen Siliziumchip, ausgelenkt wird, während gleichzeitig seine Position und Ausgangsspannung aufgezeichnet werden. Ein entsprechendes Messergebnis ist in Abbildung 4 dargestellt.

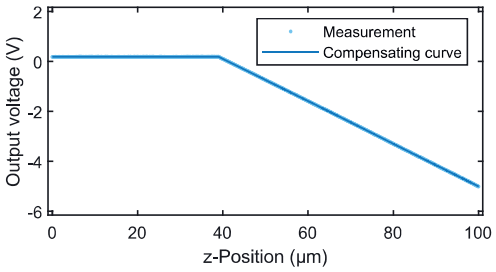


Abb. 4: Kennlinie der statischen Mikrotaster-Auslenkung. Gemessene Werte (Punkte) und Ausgleichskurve (Linie) sind nicht getrennt sichtbar. Eine konstante Ausgangsspannung liegt vor, bis der Sensor bei z ≈ 39 µm Kontakt mit der Probe herstellt. Danach fällt die gemessene Spannung nahezu linear ab. Die Steigung hier entspricht der Empfindlichkeit.

Laut Datenblatt sollte die Ausgangsspannung des Mikrotasters in dieser Messung mit ca. 85 mV/µm abfallen, sobald Kontakt hergestellt ist. Der gemessene Verlauf der Empfindlichkeit in Abhängigkeit von der Cantileverauslenkung, dargestellt in Abbildung 5, weicht davon jedoch um bis zu 1,5 mV/µm ab.

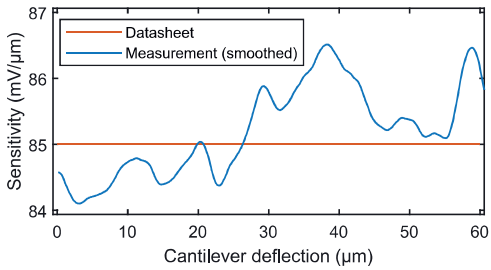


Abb. 5: Empfindlichkeit des Sensors in Abhängigkeit seiner Auslenkung. Wert laut Datenblatt (rot) und geglätteter gemessener Verlauf (blau).

Die mittlere gemessene Empfindlichkeit des Sensors liegt mit ca. 85,2 mV/µm nahe an dem nominellen Wert, es zeigt sich jedoch auch ein Anstieg der Empfindlichkeit mit der Auslenkung. Die gemessene Auslenkungsabhängigkeit der Empfindlichkeit muss im Modell berücksichtigt werden.

Nachdem die Empfindlichkeit des Sensors bestimmt ist, werden Messungen der freischwingenden out-of-plane-Schwingungsmoden durchgeführt. Abbildung 6 zeigt beispielsweise die erste out-of-plane-Schwingungsmode. Hier können über die jeweilige Resonanzfrequenz, Resonanzgüte und Resonanzamplitude Informationen über Cantilever-Geometrie, Luftdämpfung und Empfindlichkeit der Anregung erlangt werden.

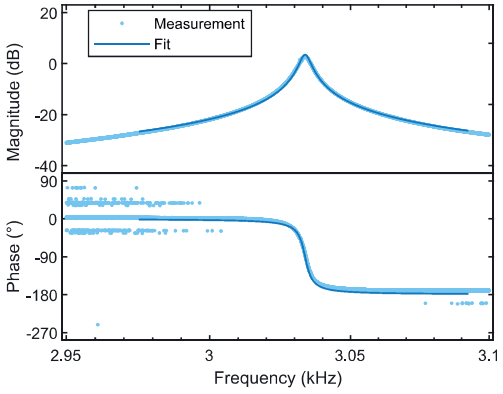


Abb. 6: Amplituden- (oben) und Phasengang (unten) des freischwingenden Sensors gemessen nahe der ersten out-of-plane-Schwingungsmode: gemessener Verlauf (Punkte) und Fit des Schwingungsmodells an die Messwerte (Linie).

Zunächst werden die gemessenen Resonanzfrequenzen genutzt, um die nominelle Geometrie des Cantilevers zu überprüfen. Da Länge und Breite des Cantilevers in der Fertigung durch Lithographie sehr präzise eingestellt werden können, ist davon auszugehen, dass hier nur geringe Abweichungen zwischen Datenblatt und tatsächlichen Werten bestehen. Vielmehr sind die Abweichungen zwischen den Resonanzfrequenzen des Modells und den gemessenen Werten durch eine Abweichung der tatsächlichen Dicke des Cantilevers vom Nominalwert zu erklären.

Wie Abbildung 7 zeigt, liegen die Resonanzfrequenzen des Modells unter Annahme der nominellen Geometrie stets unter den gemessenen

Werten. Durch Erhöhung der Dicke des Cantilevers im Modell auf ca. 55,3 µm nehmen die Abweichungen zwischen Modell und Messungen auf weniger als 3% ab. Diese verbleibenden Abweichungen können durch Idealisierungen des Modells erklärt werden, die der Realität nur näherungsweise entsprechen. Um dies zu berücksichtigen wird eine Ausgleichskurve durch die verbleibenden Abweichungen gelegt und zur Umskalierung der Frequenzachse verwendet. So bildet das Modell alle gemessenen Resonanzfrequenzen nach. Dies ist für die weitere Auswertung wichtig, da es andernfalls nicht möglich wäre das Modell an gemessene freischwingende Resonanzmoden zu fitten.

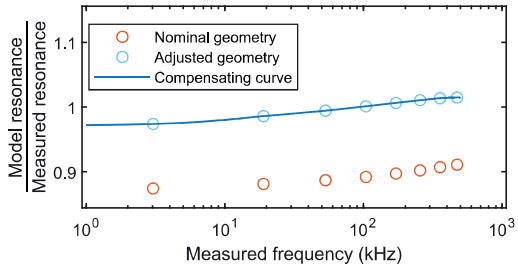


Abb. 7: Modell-Resonanzfrequenz bezogen auf gemessene Resonanzfrequenz für die ersten acht out-of-plane-Schwingungsmoden eines Cantilevers unter Annahme der nominellen Sensorgeometrie laut Datenblatt ([4], rote Kreise) und nach Anpassung der Cantileverdicke (blaue Kreise) mit einer Ausgleichskurve (blaue Linie), die in die weitere Modellrechnung eingeht.

Nachdem sichergestellt ist, dass die gemessenen Resonanzfrequenzen mit denen des Modells übereinstimmen, kann das Modell, wie in Abbildung 6 zu sehen ist, an die Messungen gefittet werden. Für jede Resonanzmode wird so jeweils ein Wert der Luftdämpfung und einer der Empfindlichkeit der Anregung bestimmt.
Die resultierenden Werte der Luftdämpfung sind in Abbildung 8 aufgetragen. Es zeigt sich ein monotoner Verlauf, in dem die Luftdämpfung mit steigender Frequenz steigt. Da Werte der Luftdämpfung a priori nicht bekannt sind [5], können hieraus jedoch keine weiteren Schlüsse gezogen werden.

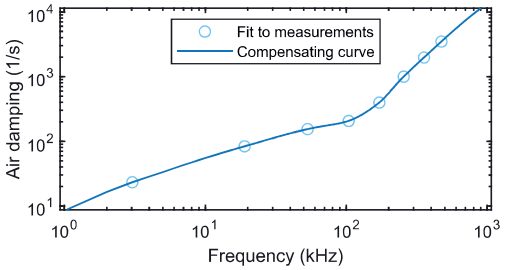


Abb. 8: Frequenzverlauf der Luftdämpfung: Bestimmt durch Anpassungen des Schwingungsmodells an Messungen des freischwingenden Sensors in den Frequenzbereichen um die ersten acht out-of-plane-Schwingungsmoden herum.

Im nächsten Schritt werden Resonanzmoden des Sensors in Kontakt mit einer Probenoberfläche aufgenommen. Ein Beispiel für eine solche Messung ist in Abbildung 9 dargestellt.

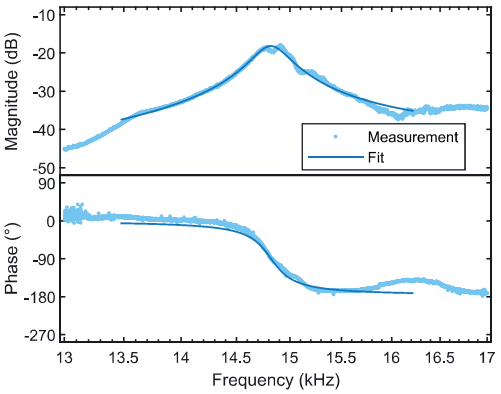


Abb. 9: Amplituden- (oben) und Phasen-Frequenzgang (unten) bei Kontakt des Sensors mit einer Probe in der Umgebung der ersten out-of-plane-Schwingungsmode: gemessener Verlauf (Punkte) und Fit des Schwingungsmodells an die Messwerte (Linien).

Aus diesen Kurven werden durch Anpassung anschließend jeweils ein Wert für die Kontaktsteifigkeit, die Dämpfungskonstante und die Empfindlichkeit der Anregung bestimmt. Werden mehrere Moden in Kontakt unter gleichen Bedingungen untersucht, so sollte die bestimmte Kontaktsteifigkeit für alle Moden identisch sein [8]. Ist dies nicht der Fall, so muss die Position der Tastspitze L_1 im Modell angepasst werden.
Aus den Messungen der freischwingenden Resonanzmoden und der Kontaktresonanzmoden kann nun der Frequenzverlauf der Empfindlichkeit des Piezoaktors bestimmt werden. Wie in Abbildung 10 dargestellt, liegen die bestimmten Werte sämtlich unter dem nominellen Wert aus

dem Datenblatt. Dies ist möglicherweise durch Schwingungen der Mikrotasterträger-Platine zu erklären. Es ist jedoch unklar, ob dieser Effekt allein die Abweichungen von bis zu zwei Größenordnungen erklärt. Trotz verbleibender Unklarheiten führt die Berücksichtigung der aus den Anpassungen ermittelten Korrekturwerte dazu, dass das Modell die Messkurven besser nachbildet, weshalb darauf nicht verzichtet wird.

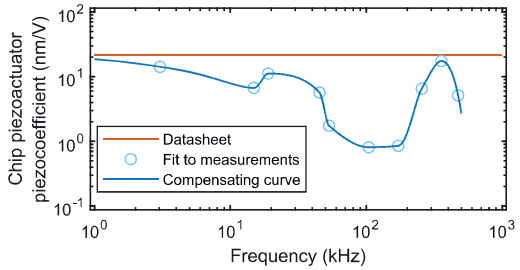


Abb. 10: Frequenzverlauf des Piezokoeffizienten des verwendeten Piezoaktors: Wert aus Datenblatt ((Miniatur-Piezoaktor, 5 × 5 × 2 mm³, PL 055.30 PICMA®Chip Actuator, www.piceramic.de), rot), Verlauf bestimmt durch Anpassungen des Schwingungsmodells an Messungen nahe den ersten acht out-of-plane-Schwingungsmoden bei freischwingendem Sensor und der ersten beiden out-of-plane-Schwingungsmoden bei Kontakt mit einer Probe (blaue Kreise) mit Ausgleichskurve (blaue Linie).

Der letzte Schritt der Kalibrierung ist die Bestimmung der Position der Tastspitze L_1 entlang der Cantileverachse. Die Notwendigkeit dieses Schrittes ist nicht durch Abweichungen zwischen nominellen Werten und realem Sensor, sondern durch Idealisierungen im Modell zu erklären, die der Realität nur näherungsweise entsprechen [8]. Um dies zu berücksichtigen, werden mindestens zwei Kontaktresonanzmoden des Tasters in Kontakt mit der gleichen Probe unter gleichen Bedingungen aufgenommen. Anschließend werden verschiedene Werte für L_1 angesetzt, das Modell wird mit allen hier beschriebenen Methoden auf diese Spitzenposition kalibriert und die durch Anpassung bestimmten Kontaktsteifigkeiten der gemessenen Resonanzmoden werden verglichen. Wird dies für eine ausreichend hohe Anzahl an Spitzenpositionen durchgeführt, ergibt sich ein Verlauf entsprechend Abbildung 11.

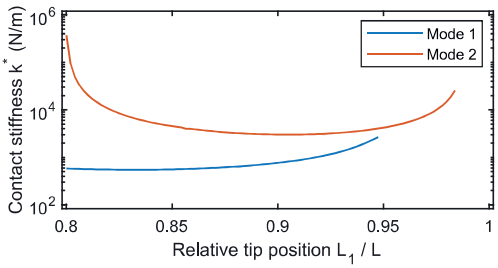


Abb. 11: Kontaktsteifigkeit in Abhängigkeit von der relativen Position der Spitze bezogen auf die Länge des Cantilevers, berechnet aus den gemessenen Resonanzfrequenzen in der ersten (blau) und zweiten (rot) out-of-plane-Schwingungsmode. Sobald L_1/L den Wert 0,95 im Fall der ersten bzw. 0,98 im Fall der zweiten Mode überschreitet, kann die gemessene Resonanzfrequenz nicht mehr nachgebildet werden.

Im Fall der ersten Kontaktresonanzmode führt die Verschiebung der Tastspitze in Richtung des freien Endes zu steigenden Kontaktsteifigkeiten. Überschreitet L_1/L den Wert 0,98, so kann das Modell die gemessene Resonanzfrequenz, wie in Abbildung 9 dargestellt, nicht mehr nachbilden. Bei Betrachtung der zweiten Kontaktresonanzmode fällt die Kontaktsteifigkeit zu Beginn und steigt später wieder an. Hier kann die gemessene Resonanzfrequenz ab einer relativen Spitzenposition von 0,98 nicht nachgebildet werden. Mit der nominellen Spitzenposition L_1 = 4,9 mm (L_1/L = 0,98) kann folglich nur die zweite gemessene Resonanzfrequenz reproduziert werden.

Entgegen der Vermutung gemäß [8] schneiden sich beide Kurven jedoch nicht. Deshalb wird nicht der Schnittpunkt sondern der Punkt der geringsten Abweichung zwischen beiden Kurven für die Bestimmung der Spitzenposition verwendet. Dieser liegt bei L_1 = 4,74 mm (L_1/L = 0,948).

Das kalibrierte Modell kann nun genutzt werden um Messungen auszuwerten. In Abbildung 12 ist beispielsweise das Ergebnis von Messungen an einer dünnen Polymerschicht auf einem Siliziumsubstrat dargestellt. In diesen Messungen wurde die Resonanzfrequenz der ersten out-of-plane-Schwingungsmode bei verschiedenen Antastkräften untersucht. Die Antastkraft wurde anschließend über die in Abbildung 4 dargestellte Kurve in die Deformation der Probe umgerechnet. Mit steigender Deformation wird zunehmend der Einfluss des Siliziumsubstrats unter der Polymerschicht gemessen, wodurch die Kontaktsteifigkeit steigt.

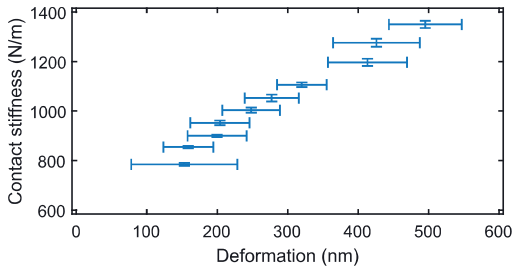


Abb. 12: Kontaktsteifigkeit in Abhängigkeit von der Deformation der Probe, gemessen in der ersten out-of-plane-Schwingungsmode an einer 1400 nm dünnen Schicht von mr-NIL210 (micro resist technology GmbH) Nanoimprintlack auf einem Siliziumsubstrat.

Fazit und Ausblick

In dieser Arbeit wurde ein Kontaktresonanz-Schwingungsmodell aus der Rasterkraftmikroskopie an einen deutlich größeren Cantilever-sensor angepasst. In diesem Zuge wurde eine Methode vorgestellt, das Modell auf Basis von Referenzmessungen systematisch zu kalibrieren, sodass Amplituden-/Phasenfrequenzgänge im Bereich der resonanten out-of-plane-Schwingungsmoden möglichst genau nachgebildet werden. Dafür wurden Nichtlinearitäten des Sensors, Abweichungen der Sensorgeometrie vom Datenblatt, Auswirkungen der Luftdämpfung sowie Frequenzabhängigkeiten der Anregung berücksichtigt. Die vorgestellte Prozedur bildet die Grundlage dafür, Kontaktresonanz-Messungen an dünnen Schichten auf Werkstücken präzise und reproduzierbar auswerten zu können. Als Anwendungsbeispiel wurde eine ausgewertete Messung an einer dünnen Polymerschicht aufgeführt.

In weiteren Analysen müssen allerdings noch verbleibende Unschlüssigkeiten geklärt werden: Einerseits muss das Ausmaß der Frequenzabhängigkeit des Piezoaktors untersucht werden, um zu überprüfen, ob die hier bestimmten Werte der Realität entsprechen. Andererseits müssen weitere Kontaktresonanzmessungen an Referenzproben durchgeführt werden, um die durch die Position der Tastspitze gegebene effektive Cantileverlänge für die relevanten Schwingungsmoden konsistent zu bestimmen.

Daraufhin müssen die Kontaktsteifigkeiten und Dämpfungskonstanten, die mit dem in dieser Arbeit beschriebenen Modell an Dünnschichtproben bestimmt werden, mit Hilfe von Referenz-Messverfahren auf physikalische Größen wie das (komplexe) Elastizitätsmodul zurückgeführt werden.

Finanzierung

This project has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme under no. 17IND05 MicroProbes.

Literaturnachweis

[1] Hofmann, E., Rüsch, M. Industry 4.0 and the current status as well as future prospects on logistics. *Computers in Industry* 89, 23–34 (2017). https://doi.org/10.1016/j.compind.2017.04.002

[2] Wasisto, H. S.; Doering, L.; Brand, U.; Peiner, E. Ultra-high-speed cantilever tactile probe for high-aspect-ratio micro metrology. In *2015 Transducers - 2015 18th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, Anchorage, AK, USA, 21–25 June 2015, pp. 1061–1064. IEEE. https://doi.org/10.1109/transducers.2015.7181109

[3] EMPIR Project "17IND05 MicroProbes - Multifunctional ultrafast microprobes for on-the-machine measurements", https://www.ptb.de/empir2018/microprobes/

[4] https://www.cismst.org/loesungen/mikrotast-spitzen/. [Zugriff: 29. 04. 2019]

[5] Rabe, U. Atomic force acoustic microscopy. *Applied Scanning Probe Methods II* 37–90 (2006). Springer Berlin Heidelberg

[6] Kocun, M., Ohler, B. Exploring Nanoscale Viscoelastic Properties. *Imaging & Microscopy* 15 No. 4 22-24 (2013).

[7] Lübbe. J., Schnieder, H., Reichling, M. Pressure Dependence of the Q-factor of Cantilevers Used for NC-AFM. *E-Journal of Surface Science and Nanotechnology* 9 30–33 (2011)

[8] Hurley, D. C. Contact resonance force microscopy techniques for nanomechanical measurements. *Applied Scanning Probe Methods XI* 97–138 (2009). Springer Berlin Heidelberg

DOI 10.5162/sensoren2019/P2.12

Tastsystem zur Bestimmung der instrumentierten Eindringhärte

Th. Frank, C. Maier, M. Kermann, A. Grün, A. Cyriax, J. Patel, Th. Ortlepp
Forschungsinstitut für Mikrosensorik und Photovoltaik GmbH, Konrad-Zuse-Straße 14, D-99099 Erfurt, Deutschland

Zusammenfassung

Die mechanische Charakterisierung dünner Schichten ist für viele Anwendungen wichtig. Zur Bestimmung der Härte stehen verschiedene Verfahren zur Verfügung. Im vorliegenden Beitrag wird die Nanoindentierung beschrieben, auch instrumentierter Eindringversuch genannt. Der Eindringkörper ist eine Diamantspitze mit bekannter Geometrie, welche in die zu prüfende Oberfläche hineingedrückt wird. Aus der kontinuierlich gemessenen Kraft in Abhängigkeit vom Eindringweg können die Größen: Härte und reduziertes Elastizitätsmodul bestimmt werden. Beschrieben wird eine neue Sensoranordnung. Zwei übereinander liegende Weg-Kraftsensoren nehmen die Kennlinie auf. Dadurch, dass beide Messstrecken und die Messkraft auf einer Linie liegen, genügt die Anordnung dem „Abbeschen Komparatorprinzip". Als Sensor wird ein siliziumbasierter 3D-Kraftsensor mit 4 Messbrücken verwendet.

Keywords: Instrumentierte Eindringhärte, Kraftsensor, piezoresistiv, Silizium, Nanoindenter

Einleitung

Die mechanische Charakterisierung dünner Schichten ist für viele Anwendungen wichtig. Zur Bestimmung der Härte stehen verschiedenen Verfahren zur Verfügung. Die Härteprüfung nach Vickers oder nach Rockwell sind in der Industrie die bekanntesten [1]. Weniger häufig verwendet wird die Universalhärte nach Martens. Bei diesem Verfahren werden während der Belastungs- und Entlastungsphase kontinuierlich die Kraft und die Eindringtiefe gemessen. Die Nanoindentierung, auch instrumentierter Eindringversuch, ist diesem Verfahren angelehnt, findet aber in viel kleinerem Maßstab statt. Der Eindringkörper ist eine Diamantspitze mit bekannter Geometrie, welche in die zu prüfende Oberfläche hineingedrückt wird. Aus der kontinuierlich gemessenen Kraft in Abhängigkeit vom Eindringweg können die Größen:

- die Härte eines Materials als Quotient der Maximallast und der projizierten Kontaktfläche,
- der reduzierte Elastizitätsmodul und
- das Kriechen

bestimmt werden.

Durch den nur minimalen Eindruck ist es möglich ein Mapping zur Untersuchung der Variation der mechanischen Eigenschaften über einen größeren Probenabschnitt durchzuführen. Insbesondere können die mechanischen Eigenschaften von Korn- und Phasengrenzen bestimmt werden.

Verfügbare Messgeräte besitzen deshalb hochauflösende Kraft- und Wegsensoren. Wegen den hohen Anforderungen sind diese sehr aufwendig gestaltete und benötigen zudem sehr präzise Führungen in einem sehr steifen Gestell.
Der Beitrag stellt einen miniaturisierten Kraftwegsensor zur Bestimmung der instrumentierten Eindringhärte vor.

Das Funktionsprinzip

Der Messbereich für die Prüfkraft reicht bis 100 mN, für den Wegsensor sind maximal 5 µm erforderlich. Das Messsystem wird mit zwei kombinierten Kraft-Weg-Sensoren realisiert. Diese sind übereinander angeordnet und durchdringen sich gegenseitig. Der Eindringkörper, ein Berkovich-Diamant auf einem Stahlstab, ist am oberen Kraftsensor befestigt. Er durchdringt den unteren Kraftsensor berührungslos, welcher nur als Wegsensor verwendet wird. Der obere Sensor misst die auf den Diamanten wirkende Kraft und den von System zurückgelegten Weg, abzüglich der Eindringtiefe. Der untere Sensor und auch der nachgiebigere misst nur den Gesamtweg. Aus beiden Daten kann die gesamte Kraft-Eindringtiefe-Kurve berechnet werden. Auch der Aktor zur Generierung der Kraft kann über den unteren Sensor geregelt werden.

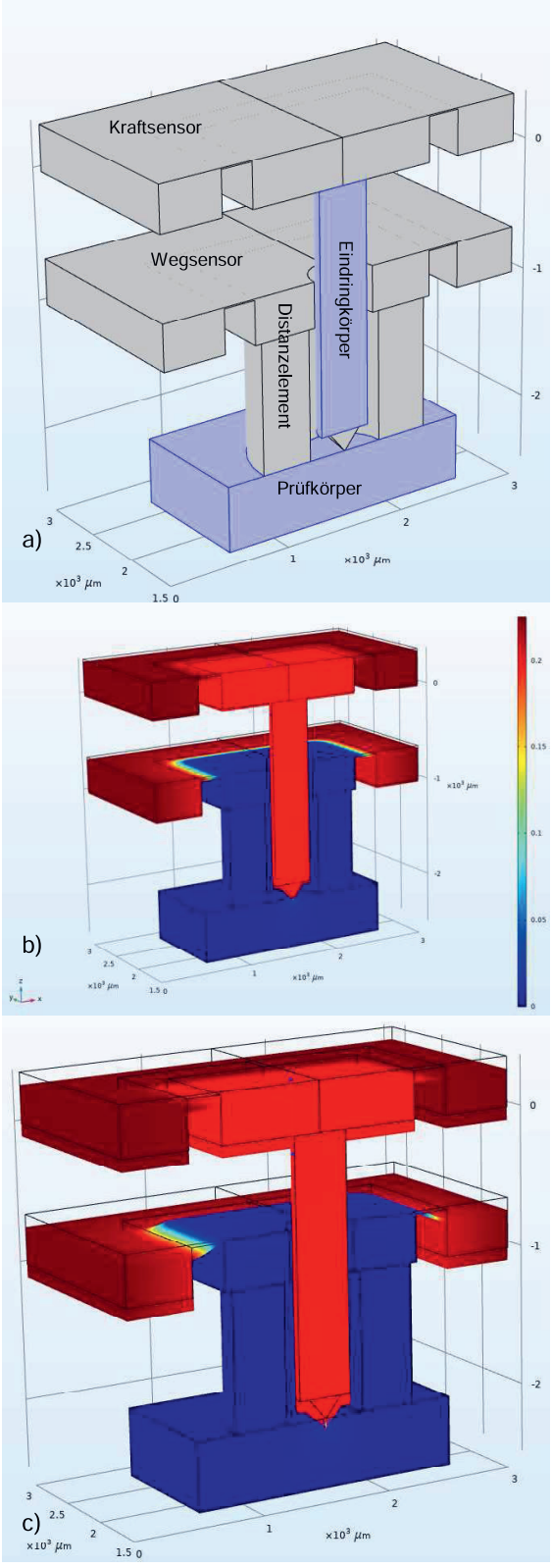


Abb. 1: Illustration des Funktionsprinzips, a) geometrischer Aufbau, b) Anfang der Bewegung, c) Abschluss des Eindringversuchs.

Der Vorteil dieser Lösung ist der kompakte und doch hoch präzise Aufbau.

Dadurch, dass beide Messstrecken und die Messkraft auf einer Linie liegen, genügt die Anordnung dem „Abbeschen Komparatorprinzip". Benannt wurde es nach Ernst Abbe, der es als Erster beschrieben hat.

Liegen die am Prüfling zu messende Strecke und der Maßstab (Maßverkörperung) des Messgerätes nicht in einer Linie, können im Messgerät Kippfehler (sogenannte Kippfehler erster Ordnung) auftreten. Diese hängen vom Spiel der Führung quer zur Bewegungsrichtung ab und sind prinzipiell nicht zu vermeiden. Die beiden Sensoren sind ausschließlich durch ein Festkörpergelenk spielfrei geführt.
Die Kraftsensoren sind aus Silizium aufgebaut und beinhalten vier piezoresistive Messbrücken an jeder Seite einer Biegeplatte mit einem biegesteifen Zentrum. Diese können sowohl Kraft und Weg in Richtung der Eindringtiefe als auch die Verkippung messen und so die Qualität der Messung bewerten.
Dadurch, dass die beiden Sensoren auf engstem Raum angeordnet sind, werden Abweichungen, die außerhalb entstehen nahezu vollständig kompensiert. Dies sind im Wesentlichen Abweichungen in der Führung und in der Positionssteuerung.
Die Abb. 1 zeigt den prinzipiellen Aufbau des Systems „Nanoindenter". Dieses System ist in der Lage Betrag und Winkel eines Kraftvektors zu bestimmen. Durch die Dicke der Biegeplatte wird die Empfindlichkeit auf Weg, bzw. Kraft definiert.

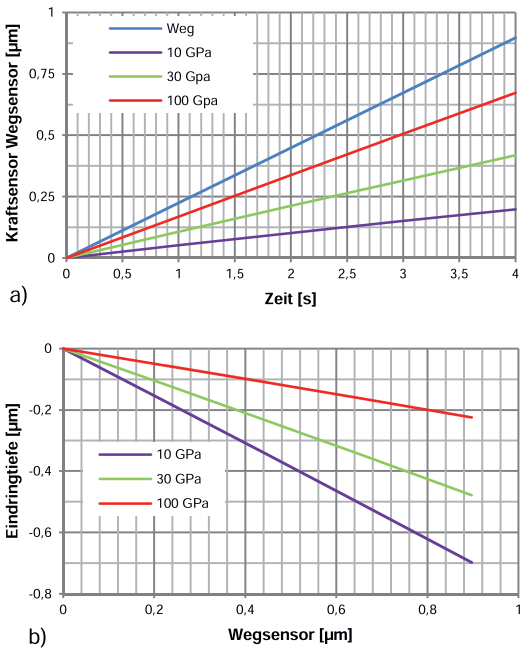


Abb. 2: Messvorgang Nanoindenter

Der Wegsensor ist in Richtung hoher Nachgiebigkeit, der Kraftsensor in Richtung hoher Stei-

figkeit optimiert. Beide Sensoren sind fest über einen Verdrahtungsträger mit einander verbunden. Das Antriebssystem ist so in der Lage beide Sensoren gleichförmig auf den Prüfkörper zu bewegen. Dabei ist im biegesteifen Zentrum des Kraftsensors der Eindringkörper mit der Diamantspitze verankert. Der Wegsensor ist ein Distanzelement, welches mit dem Prüfkörper in Kontakt steht.

Der Messvorgang beginnt, sobald Diamantspitze und Distanzelement den Prüfkörper berühren. Der Wegsensor nimmt den Gesamtweg des Systems auf. Der Kraftsensor hingegen nur die Differenz aus dem Gesamtweg und der Eindringtiefe des Diamanten. Die Abb. 2 zeigt den Messvorgang am Beispiel variierender Prüfkörper hinsichtlich ihres Elastizitätsmoduls. Abb. 2a zeigt die Rohdaten bei identischer Zeitbasis. Der Weg ist bei allen Prüfungen identisch. Anschließend werden die Daten, die Differenz aus dem Gesamtweg und der Eindringtiefe des Diamanten, über den Weg aufgetragen. Zu erkennen ist die wegabhängige Eindringtiefe. Nach einer Kraft-Kalibrierung ist auch eine Darstellung der Eindringkraft über den Weg möglich.
Die Abb. 3 zeigt die Rohsignale des ersten unkalibrierten Messdurchgangs. Das Signal wird geglättet, wenn das Signal des Kraftsensors (Indenter) über dem Gesamtweg aufgetragen wird.

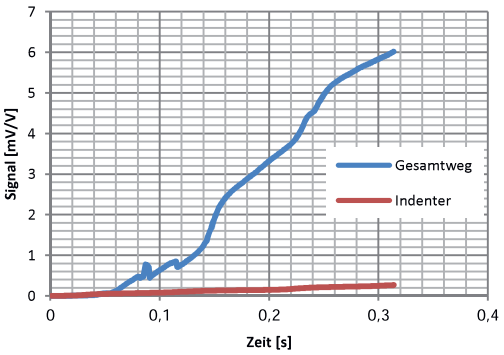


Abb. 3: Erster Messversuch ohne Kalibrierung

Aufbau Nanoindenter

Die Abb. 4 zeigt den verwendeten Aufbau des Indenter-Sensors im Detail. Auf einem Verdrahtungsträger ist von unten der Wegsensor durch Thermokompressionsbonden [2] gefügt und elektrisch kontaktiert. Wegen der guten mechanischen Eigenschaften dieses Verfahren zeigte es sich als optimal, da neben der elektrischen Kontaktierung, eine hohe positionstreue und eine hohe mechanische Festigkeit gefordert wurde.

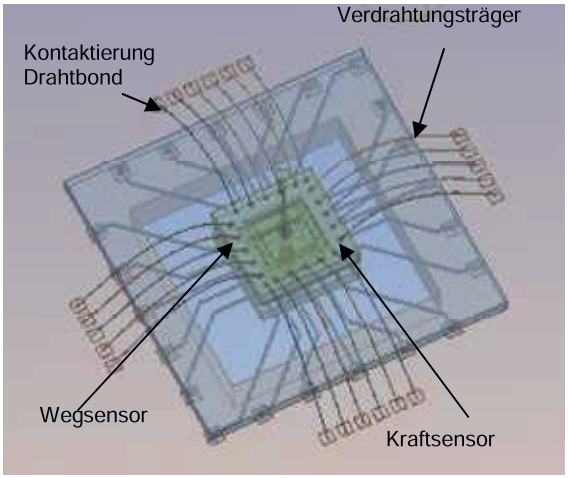


Abb. 4: Einzelheit: Aufgebaute Sensoren auf dem Verdrahtungsträger

Auf Underfill-Material konnte deshalb vollständig verzichtet werden. Die Gold-Stud-Bumps sind in der Abb. 5 gezeigt.

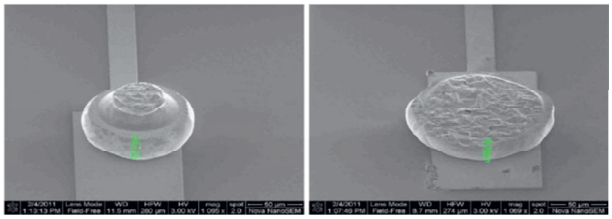


Abb. 5: Gold-Stud-Bumbs, vor dem Fügen (links), nach dem Fügen (rechts)

Durch die hochohmigen in Reihe geschalteten Messwiderstände ist eine Fehleranalyse durch reine Widerstandsmessung der Schaltung nicht effektiv. Daher wurde der *Flip-Chip*-Aufbau mit 22 Kontakten mit einem Ultraschallmikroskop überprüft. Dieses Verfahren eignet sich hervorragend zur Abschätzung der Zuverlässigkeit. Ein fehlerhafter Kontakt ist in der Abb. 7 gezeigt.

Der obere Wegsensor wird auf den Verdrahtungsträger geklebt. Über ein Flip-Chip-Verfahren wird der Verdrahtungsträger auf die Trägerleiterkarte gefügt. Der Wegsensor wird über lange Drahtbondbrücken mit der Leiterkarte (PCB) verbunden, siehe Abb. 4.

In der Abb. 8 ist die am Halter montierte Leiterkarte zu sehen

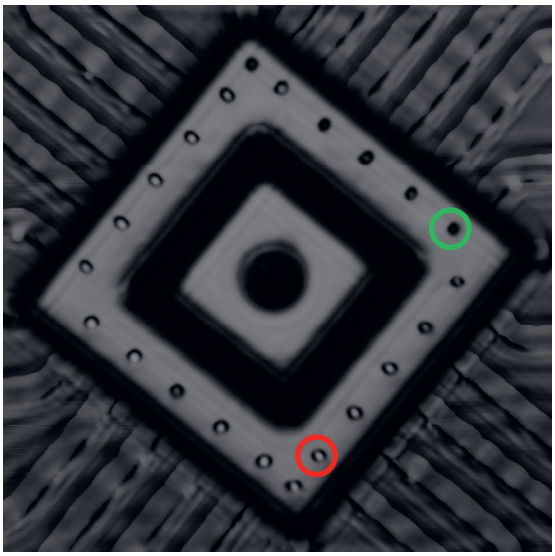


Abb. 6: Ultraschalluntersuchung zur Bestimmung der Bondqualität, eine helle Stelle bedeutet eine fehlerhafte Verbindung (rot), dunkle Stellen sind in Ordnung (grün)

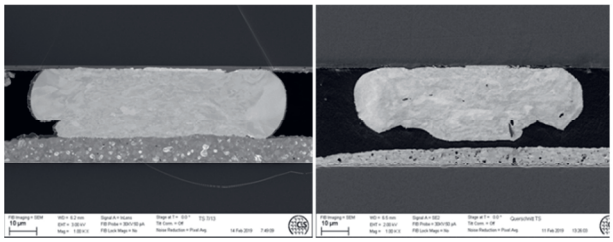


Abb. 7: Vergleich gute und unzureichende Bondung

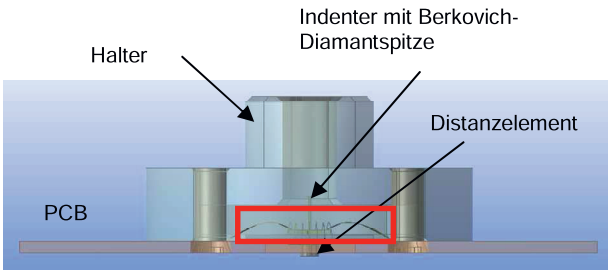


Abb. 8: Aufbau des Messsystems, Details sind in Abb. 9 gezeigt

Die Abb. 9a zeigt den realen Aufbau des Kraftsensors mit der Bohrung zur Aufnahme des Indenters mit der Berkovich-Diamantspitze. Abb. 9b zeigt den realen Aufbau des Wegsensors mit dem zylindrischen Distanzelement. Das Messsystem besitzt folgende Daten:

- Messspanne 30 mV/V
- Sensitivität 5 mV/(V µm)
- Auflösung Weg 60 pm
- Auflösung Kraft 10 µN

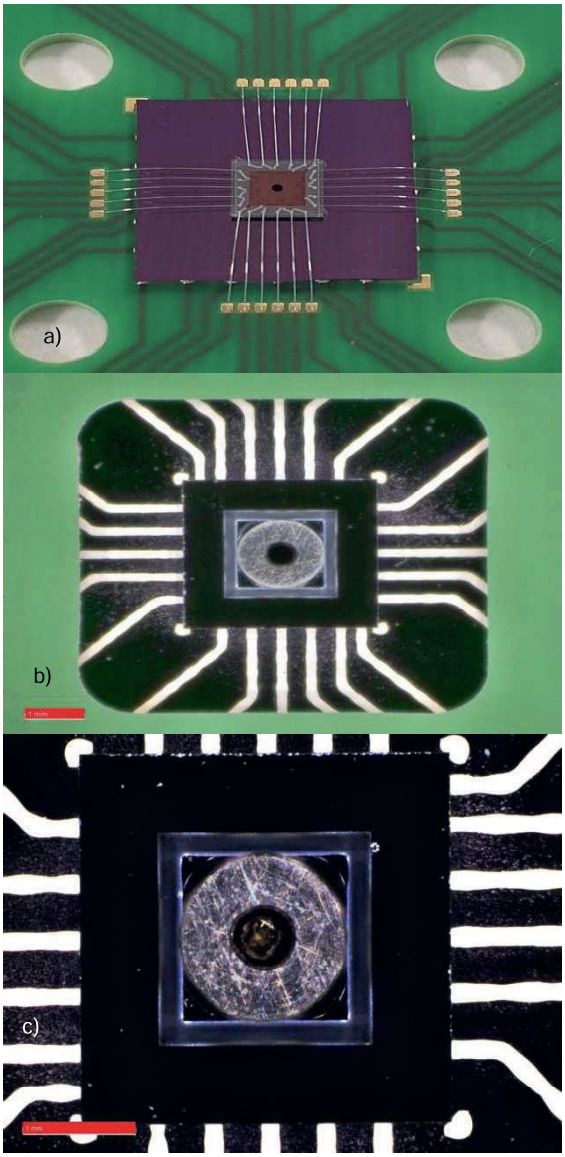


Abb. 9: a) Kraftsensor geklebt und drahtgebondet, b) Wegsensor mit Distanzelement, c) Sicht auf den Diamanten

Literaturnachweis

[1] https://de.wikipedia.org/wiki/Härte
[2] S. Lohse, R. Schachler, Thermokompressionsbonden, https://www.finetech.de/de/wissen/technical-paper-bonding/ thermokompressionsbonden/

Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Energie (BMWi) und dem für Bundesministeriums für Bildung und Forschung (BMBF) für die finanzielle Unterstützung der Projekte AC-Brücke (MF150187), Dumbo (MF160051), Flip-Chip-DDS (49MF180034), SISI (MF150021) und 3D-Kraftsensor (MF140005).

DOI 10.5162/sensoren2019/P2.13

Korrelation von wiederholten Messungen in einer Modenverwirbelungskammer

C. Cammin[1], R. Heynicke[1], G. Scholl[1]
[1]Professur für Elektrische Messtechnik,
Helmut-Schmidt-Universität, Universität der Bundeswehr Hamburg,
Holstenhofweg 85, 22043 Hamburg, Deutschland

Zusammenfassung

Modenverwirbelungskammern (MVKn) stellen vielfältig verwendbare, reproduzierbare Messumgebungen dar. Grundsätzlich handelt es sich dabei um elektromagnetisch geschirmte Volumina, in denen sog. Modenrührer eingebracht sind, die ein elektromagnetisches Feld „verwirbeln“, sodass ein Prüfling einer definierten aber statistisch variierenden Umgebung ausgesetzt wird. Die Modenrührer in MVKn können schrittweise zwischen den einzelnen Messzeitpunkten oder kontinuierlich verfahren werden. In diesem Beitrag werden die jeweiligen Eigenschaften von wiederholten Messungen, die sowohl bei kontinuierlichem als auch schrittweisem Betrieb der Modenrührer aufgenommen wurden, betrachtet. Mit Hilfe von Korrelationskoeffizenten wird jeweils die Wiederholbarkeit der aufgenommenen Messwertsequenzen insgesamt untersucht. Darüber hinaus wird die Wiederholbarkeit von Fadingeinbrüchen gesondert betrachtet, da diese einen wesentlichen Einfluss auf Funksysteme haben, die mittels einer MVK charakterisiert werden können. Als Ursache für Abweichungen bei Wiederholungen wird abschließend das zeitliche Verhalten bzw. die Streuungen zwischen den Messzeitpunkten betrachtet.

Keywords: Modenverwirbelungskammer, Korrelation, Drahtlose Sensor/Aktor Netzwerke, drahtlose Sensorsysteme, Test

Einleitung

Modenverwirbelungskammern (MVKn) werden als reproduzierbare Messumgebungen für Messungen zur Elektromagnetischen Verträglichkeit [1] sowie für Antennen und drahtlose Kommunikationsgeräte [2] genutzt. MVKn bestehen grundsätzlich aus einem elektromagnetisch abgeschirmten Volumen, in das ein Prüfling eingebracht wird. Neben mindestens einer Antenne zur Ein- sowie Auskopplung eines elektromagnetischen Feldes befindet sich in einer MVK mindestens ein sogenannter Modenrührer. Häufig sind die Modenrührer als mechanisch bewegte oder rotierende Metallplatten oder als Drehteller für den Prüfling ausgeführt, um die Ausbreitungsbedingungen der elektromagnetischen Welle im Prüfvolumen zu verändern. Für eine bestimmte Anzahl von unterschiedlichen Positionen des bzw. der Modenrührer werden einzelne Messungen durchgeführt. Über eine vollständige Sequenz dieser einzelnen Messungen ist der Prüfling einem statistisch quasi-isotropen und quasi-homogenen elektromagnetischem Feld mit Rayleigh-verteilter Amplitude ausgesetzt, unabhängig von der Position oder Ausrichtung des Prüflings innerhalb des Prüfvolumens. Zudem kann eine MVK als Kanalemulator genutzt werden, um Fadingeinbrüche eines Funkkanals zu emulieren [3,4,5]. Typischerweise sind die Mittelwerte der Einzelmessungen von Interesse zur Beurteilung des Prüflings, womit die Abfolge bzw. Sequenz der Einzelmessungen meistens irrelevant ist [1,2].

MVKn können grundsätzlich entweder bei kontinuierlichem oder schrittweisem bzw. quasi-statischem Betrieb der mechanischen Modenrührer eingesetzt werden [1]: Kontinuierlicher Betrieb bedeutet, dass während der Bewegung der Modenrührer zeitlich unsynchronisiert eine definierte Anzahl an Einzelmessungen bzw. Frequenzsweeps (z.B. Streuparametermessungen mittels Vektor-Netzwerkanalysator (VNA)) durchgeführt werden.
Schrittweiser Betrieb bzw. quasi-statisch bedeutet, dass zunächst eine bestimmte, definierte Position jeweils durch die Modenrührer mechanisch angefahren wird und nach Erreichen dieser Position eine Einzelmessung oder ein Frequenzsweep (z.B. Streuparametermessung mittels VNA) gestartet wird. Nach Abschluss einer Einzelmessung bzw. eines Frequenzsweeps wird eine neue

Position durch die Modenrührer mechanisch angefahren und danach eine weitere Messung durchgeführt.

In diesem Beitrag werden wiederholte Messungen des gleichen Prüflings, die sowohl kontinuierlich als auch schrittweise in einer MVK aufgenommen wurden, hinsichtlich ihrer statistischen Eigenschaften und ihrer Wiederholbarkeit untersucht.

Messaufbau

In einer MVK wurden die komplexwertigen Streuparameter zwischen einer fest installierten Anordnung aus drei umschaltbaren Antennen und einer Discone-Antenne als Prüfling mit einem VNA im Frequenzbereich von 2,4 GHz bis 2,5 GHz in $L = 1001$ Frequenzschritten gemessen. Die Messbandbreite des VNA betrug stets 100 kHz.

Die verwendete MVK ist vom Typ RTS60 des Herstellers Bluetest [6]. Die Innenabmessungen betragen ca. 1,8 m (Höhe), 1,2 m (Tiefe) und 1,9 m (Länge). Als Modenrührer sind zwei Bleche die jeweils entlang der Decke und einer Seitenwand sowie ein Drehteller für das Equipment Under Test (EUT) eingebaut, die allesamt über Schrittmotoren bewegt werden. Zudem ist hinter einem festen Blech eine Antennenanordnung, bestehend aus drei zueinander orthogonal angeordneten Antennen fest installiert, wovon über eine Umschalteinrichtung jeweils eine Antenne zur Anregung des elektromagnetischen Feldes in der MVK ausgewählt wird. Jenes Blech dient dabei zur Unterdrückung einer direkten Sichtverbindung zwischen der festen Antennenanordnung und dem EUT [4].
Insgesamt war die MVK derart mit Absorbern beladen, dass das die RMS-Delayaufspreizung (engl. RMS Delay Spread) des Verzögerungsleistungsprofils (engl. Power Delay Profile, PDP) ca. 50 ns betrug. Ein Wert in dieser Größenordnung ist typisch für Funkkanäle in industriellen Umgebungen im 2,4 GHz Band [7].
Die Konfiguration der MVK, d.h. die Beladung mit Absorbern sowie die Position der Prüflingsantenne wurde für alle Messungen beibehalten. In Abbildung 1 ist ein Teil des Innenraums mit dem Drehteller, der Discone-Antenne (als EUT) sowie eingebrachten Absorbern gezeigt.
Sowohl bei kontinuierlichem Betrieb als auch bei schrittweisem Betrieb der Modenrührer wurden jeweils $N = 600$ Frequenzsweeps bei verschiedenen Kombination der Modenrührer und $A = 3$ festinstallierten umschaltbaren Antennen gespeichert.



Abb. 1: Foto des Innenraums der MVK.

Diese Messungen wurden jeweils viermal wiederholt, sodass jeweils $M = 4$ Messwert-sequenzen für jede Konfiguration und Einstellung aufgenommen wurden.

Im schrittweisen Betrieb wurde jeweils eine von insgesamt $K = 200$ Kombinationen von Positionen der Modenrührer an Wand und Decke sowie eine Drehtellerposition angefahren und auf eine $(a = 1)$ der drei fest installierten Antennen geschaltet. Danach wurde ein Frequenzsweep des VNA gestartet. Anschließend wurde auf die zweite fest installierte Antenne $(a = 2)$ umgeschaltet und ein erneuter Frequenzsweep des VNA gestartet. Danach wurde auf die dritte fest installierte Antenne $(a = 3)$ umgeschaltet und erneut ein Frequenzsweep gestartet. Im Anschluss wurde eine neue Kombination von Modenrührerstellungen angefahren und die jeweiligen Messungen sowie Umschaltungen wurden fortgesetzt, bis insgesamt $N = A \cdot K = 600$ Frequenzsweeps aufgenommen wurden.

Beim kontinuierlichen Betrieb wurde zu Beginn auf die erste fest installierte Antenne $(a = 1)$ geschaltet und alle Modenrührer in Bewegung versetzt. Mit dieser Antenne wurden während der Bewegungen $K = 200$ Frequenzsweeps aufgenommen. Nach Erreichen von jeweils $K = 200$ Frequenzsweeps wurde auf die nächste fest installierte Antenne $(a = 2; 3)$ umgeschaltet und die Messungen bei bewegten Modenrührern fortgesetzt, bis insgesamt $N = A \cdot K = 600$ Frequenzsweeps aufgezeichnet wurden. Die minimale Zeit zwischen den einzelnen Frequenzsweeps wurde dabei auf nominell 100 ms eingestellt, damit sich die Modenrührer bis zum nächsten Frequenzsweep ausreichend weit bewegen konnten.
Für beide Betriebsmodi wurden jeweils die Zeitpunkte zu Abschluss einer Messung in einer Log-Datei protokolliert. In Tabelle 1 sind die Messparameter aufgelistet.

Tab. 1: Zusammenfassung der Messparameter.

Lauf-variable	Gesamt-zahl	Beschreibung
a	$A = 3$	Feste Antennen
k	$K = 200$	Kombinationen der Modenrüher-positionen
l	$L = 1001$	Frequenzschritte
m	$M = 4$	Messwert-sequenzen
n	$N = A \cdot K = 600$	Antennen- und Modenrüher-kombinationen bzw. aufgenommene Frequenzsweeps

In Tabelle 2 sind jeweils die Messparameter fett dargestellt, über die in den folgenden Abschnitten die gemessenen S_{21}-Parameter gemittelt werden.

Tab. 2: Kennzeichnung der Messparameter, über die gemittelt wird.

Bezeichnung	Mittelung über alle
$\overline{S_{21}}(\boldsymbol{n}, l, m)$	Antennen und Modenrührer-kombinationen
$\overline{S_{21}}(n, \boldsymbol{l}, m)$	Frequenzen
$\overline{S_{21}}(n, l, \boldsymbol{m})$	Messwert-sequenzen

Mittlere Leistungsübertragung

Die Abbildungen 2 und 3 zeigen jeweils die Mittelwerte der gemessenen S_{21}-Parameter über der Frequenz für den kontinuierlichen (Abb.2) sowie den schrittweisen (Abb.3) Betrieb der Modenrührer, d.h. $|\overline{S_{21}}(\boldsymbol{n}, l, m)|^2$.
Beim schrittweisen Betrieb der Modenrührer sind die Mittelwerte $|\overline{S_{21}}(\boldsymbol{n}, l, m)|^2$ identisch über die jeweiligen Messwertsequenzen, d.h. über M. Im Vergleich dazu unterscheiden sich die Mittelwerte bei kontinuierlichem Betrieb jeweils geringfügig. Insgesamt sind die Mittelwerte $|\overline{S_{21}}(\boldsymbol{n}, l, m)|^2$ im untersuchten Bereich jeweils nahezu frequenzflach, da in den Abbildungen 2 und 3 jeweils nur ein kleiner Bereich dargestellt ist. In Tabelle 3 sind die Mittelwerte über alle Modenrührerkombinationen sowie über alle betrachteten Frequenzen, d.h. $|\overline{S_{21}}(\boldsymbol{n}, \boldsymbol{l}, m)|^2$, dargestellt. Diese Werte sind für alle Messsequenzen faktisch identisch.

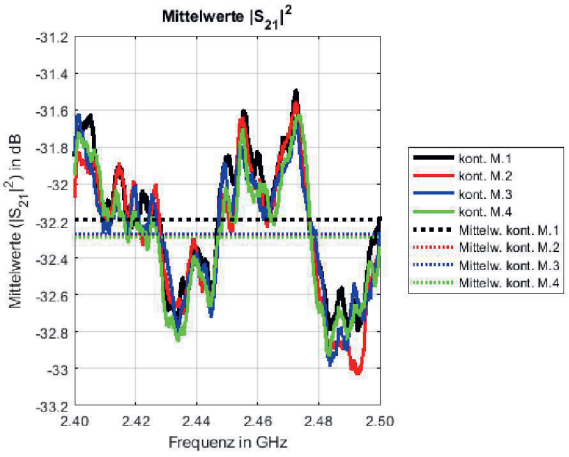


Abb. 2: Mittelwerte der Beträge der S_{21}-Parameter in dB bei kontinuierlichem Betrieb der Modenrührer.

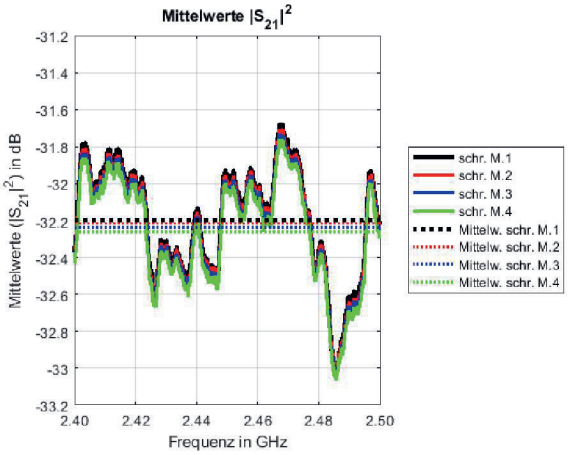


Abb. 3: Mittelwerte der Beträge der S_{21}-Parameter in dB bei schrittweisem Betrieb der Modenrührer.

Tab. 3: Mittelwerte der Beträge der S_{21}-Parameter in dB über alle Antennen-/ Modenrührerkombinationen N sowie über alle betrachteten Frequenzen L für die jeweiligen Messsequenzen.

Name	Betrieb der MVK	$\overline{\lvert S_{21}\rvert^2}$ in dB
kont. M.1	kontinuierlich	-31,14
kont. M.2	kontinuierlich	-31,22
kont. M.3	kontinuierlich	-31,22
kont. M.4	kontinuierlich	-31,24
schr. M.1	quasi-statisch	-31,15
schr. M.2	quasi-statisch	-31,17
schr. M.3	quasi-statisch	-31,19
schr. M.4	quasi-statisch	-31,22

Korrelation der Messwertsequenzen

In [8,9] wurde die Nutzung des frequenzabhängigen Korrelationskoeffizienten zur Detektion von Abweichungen einer

Referenz, wie etwa zur Detektion abweichender Polarisationen oder eines Defekts von Antennen gezeigt.
In gleicher Weise wird im Folgenden der frequenzabhängige (Leistungs-) Korrelationskoeffizient zwischen zwei Messwertsequenzen genutzt, um Divergenzen zwischen diesen zu verdeutlichen. Der (Leistungs-) Korrelationskoeffizient $\rho_P(i,j)$ zwischen zwei Messwertsequenzen i und j von $|S_{21}|^2$-Parametern $(i,j \in M;\ i \neq j)$ ist definiert als

$$\rho_P(i,j) = \frac{\sum_{n=1}^{N}\left[\left(|S_{21,i,n}|^2 - \overline{|S_{21,i}|^2}\right)\cdot\left(|S_{21,j,n}|^2 - \overline{|S_{21,j}|^2}\right)\right]}{\sqrt{\sum_{n=1}^{N}\left(|S_{21,i,n}|^2 - \overline{|S_{21,i}|^2}\right)^2 \cdot \sum_{n=1}^{N}\left(|S_{21,j,n}|^2 - \overline{|S_{21,j}|^2}\right)^2}}, \quad (1)$$

wobei $N = 600$ der Gesamtzahl der Kombinationen der Modenrührer und Antennen entspricht. Neben diesem (Leistungs-) Korrelationskoeffizienten wird zudem der komplexe Korrelationskoeffizent $\rho_C(i,j)$, bezogen auf die komplexwertigen Streuparameter zwischen zwei Messwertsequenzen i und j untersucht. Dieser ist definiert als

$$\rho_C(i,j) = \frac{\sum_{n=1}^{N}\left[\left(S_{21,i,n} - \overline{S_{21,i}}\right)^* \cdot \left(S_{21,j,n} - \overline{S_{21,j}}\right)\right]}{\sqrt{\sum_{n=1}^{N}\left|S_{21,i,n} - \overline{S_{21,i}}\right|^2 \cdot \sum_{n=1}^{N}\left|S_{21,j,n} - \overline{S_{21,j}}\right|^2}}, \quad (2)$$

wobei mit dem $*$-Symbol die konjugiert komplexe Zahl bezeichnet wird.
Die Abbildungen 4 und 5 zeigen jeweils den (Leistungs-) Korrelationskoeffizienten der aufgenommen Messwertsequenzen von Streuparametern für je zwei Messreihen bei kontinuierlichem (Abb. 4) bzw. schrittweisem (Abb. 5) Betrieb der Modenrührer.

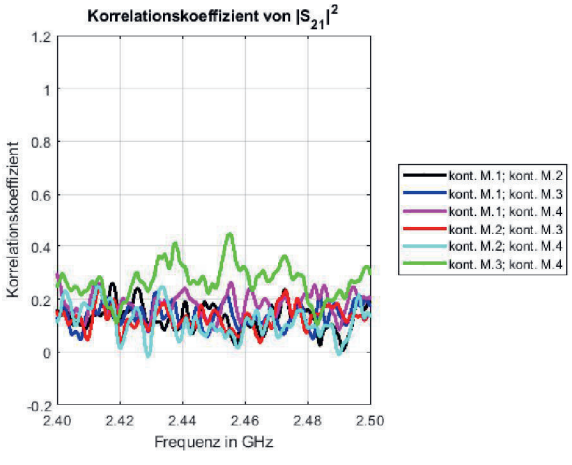


Abb. 4: Korrelationskoeffizienten über der Frequenz für jeweils zwei aufgenommene Messwertsequenzen bei kontinuierlichem Betrieb der Modenrührer.

Bei kontinuierlichem Betrieb der Modenrührer (Abb. 4) liegen die Korrelationskoeffizienten der aufgenommenen Messwertsequenzen im Mittel über der Frequenz bei ca. 0,2 bis 0,3 und sind somit jeweils als unkorreliert zu betrachten. Für den schrittweisen Betrieb (Abb. 5) liegen die Korrelationskoeffizienten über alle Frequenzen und alle Kombination von Messreihen bei 1 und sind somit stark korreliert.

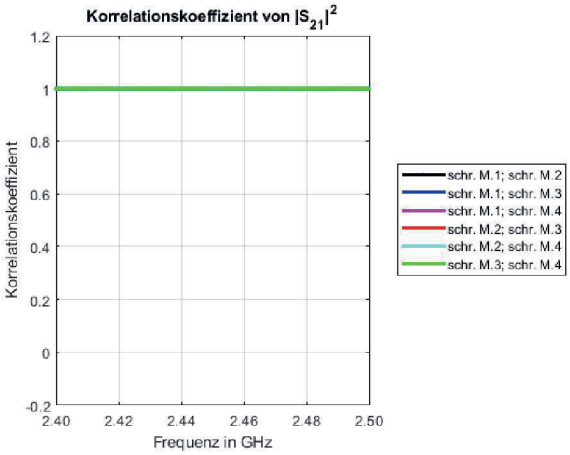


Abb. 5: Korrelationskoeffizienten über der Frequenz für jeweils zwei aufgenommene Messwertsequenzen bei schrittweisem Betrieb der Modenrührer.

Zusätzlich zum (Leistungs-) Korrelationskoeffizenten wird im folgenden der komplexe Korrelationskoeffizient nach Gl. (2) ausgewertet, da dieser zudem Unterschiede der Phase der gemessenen S_{21}-Parameter darstellen kann. In den Abbildungen 6 und 7 sind jeweils die Real- und Imaginärteile des komplexen Korrelationskoeffizenten als Funktion der Frequenz gezeigt.

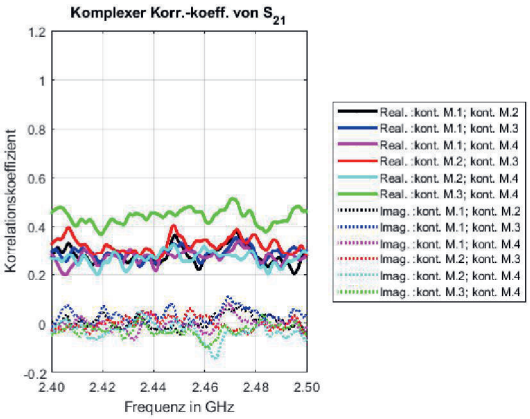


Abb. 6: Real- und Imaginärteil des komplexen Korrelationskoeffizienten über der Frequenz bei kontinuierlichem Betrieb der Modenrührer.

Bei kontinuierlichem Betrieb der Modenrührer (Abb. 6) liegen die Realteile jeweils zwischen 0,2 und 0,5 und die Imaginärteile jeweils ungefär bei 0. Bei schrittweisem Betrieb der Modenrührer (Abb. 7) liegen die Realteile jeweils für alle Kombinationen der untersuchten Messsequenzen bei 1, wohingegen die Imaginärteile allesamt bei 0 liegen. Folglich ist jeweils die Phasendifferenz der aufgenommenen Sequenzen bei schrittweisem Betrieb zu vernachlässigen.

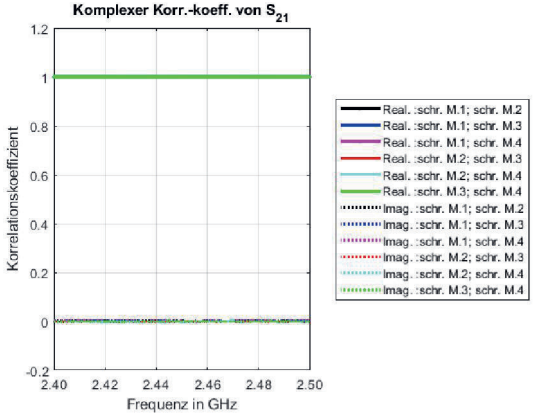


Abb. 7: Real- und Imaginärteil des komplexen Korrelationskoeffizienten über der Frequenz bei schrittweisem Betrieb der Modenrührer.

Wiederholbarkeit der Fadingeinbrüche

In [7] wurde bereits für diese MVK (bei schrittweisem Betrieb) der Modenrührer gezeigt, dass die Beträge der S_{21}-Parameter tatsächlich Rayleigh-verteilt sind. Für solche Rayleigh-verteilten $|S_{21}|$-Parameter bzw. für einen solchen Rayleigh-Funkkanal sind die Empfangsleistungen bzw. $|S_{21}|^2$ exponentialverteilt.

Für (Schmalband-) Funkanwendungen ist die Wahrscheinlichkeit von tiefen Fadingeinbrüchen von besonderem Interesse, da diese zu einem Übertragungsfehler führen können [5,10]. Im Folgenden wird daher die Wiederholbarkeit dieser Fadingeinbrüche bzw. von übertragener Leistung, die unterhalb der mittleren übertragenen Leistung, d.h. $|\overline{S_{21}}(\boldsymbol{n}, l, m)|^2$, liegt, untersucht.

In den folgenden beiden Abbildungen 8 und 9 sind jeweils die empirischen Verteilungsfunktionen der Empfangsleistung über alle Frequenzen und Modenrührer- sowie Antennenkombinationen für kontinuierlichem als auch schrittweisem Betrieb der Modenrührer dargestellt. Dabei wurden alle $|S_{21}|^2$-Werte von jeder Messwertsequenz auf deren Mittelwert, d.h. $|\overline{S_{21}}(\boldsymbol{n}, \boldsymbol{l}, m)|^2$, normiert. Die Farben der einzelnen Messwertsequenzen sind jeweils transparent, sodass sich in deren Überlappung eine petrol-grüne Färbung ergibt.

Insgesamt sind die jeweiligen Verteilungsfunktionen nahezu identisch, lediglich für tiefe Fadingeinbrüche ab etwa 30 dB unterhalb der Mittelwerte $|\overline{S_{21}}(\boldsymbol{n}, \boldsymbol{l}, m)|^2$ gibt es geringfügige Differenzen. Diese sind auch darauf zurück zu führen, dass die absolute Anzahl der Werte in diesem Bereich sehr gering ist.

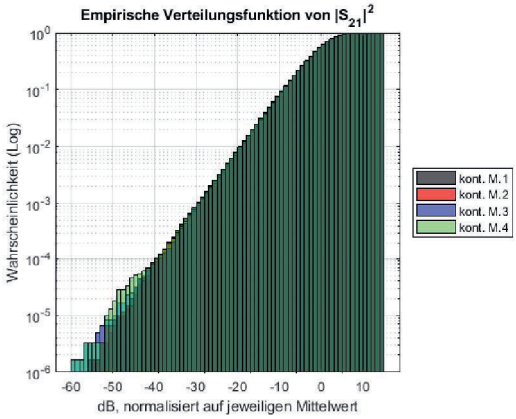


Abb. 8: Empirische Verteilungsfunktionen der Empfangsleistung bei kontinuierlichem Betrieb über alle Frequenzen und Antennen- / Modenrüherkombinationen, normiert auf den jeweiligen Mittelwert der Messwertsequenz.

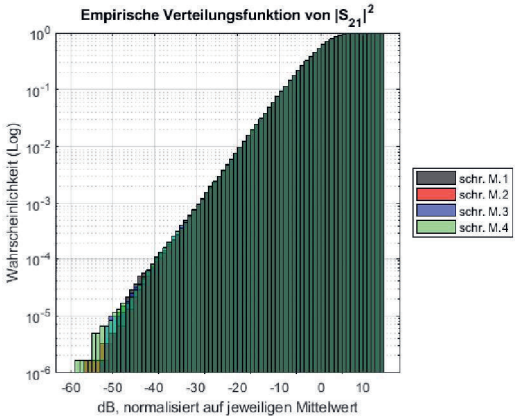


Abb. 9: Empirische Verteilungsfunktionen der Empfangsleistung bei schrittweisem Betrieb über alle Frequenzen und Antennen- / Modenrüherkombinationen, normiert auf den jeweiligen Mittelwert der Messwertsequenz.

Zur Verdeutlichung des tiefen Fadingbereichs ist in den Abbildungen 10 und 11 jeweils die absolute Anzahl der Fadingeinbrüche über alle Frequenzen und Modenrührer- sowie Antennenkombinationen dargestellt, die jeweils im Bereich von (mindestens) 20 dB bis (mindestens) 35 dB unter dem Mittelwert der

jeweiligen Messwertsequenz, d.h. $|\overline{S_{21}}(\boldsymbol{n}, \boldsymbol{l}, m)|^2$, liegen, da dieser Bereich für den typischen Betrieb von (Schmalband-) Funksystemen relevant ist.
Generell ist jeweils die absolute Anzahle der Fadingeinbrüche bei kontinuierlichem als auch bei schrittweisem Betrieb nahezu identisch. Bei kontinuierlichem Betrieb der Modenrührer (Abb. 10) wird die absolute Differenz zwischen der Anzahl der Fadingeinbrüche für tiefere Einbrüche geringer. Bei schrittweisem Betrieb der Modenrührer (Abb. 11) sind die Differenzen zwischen der absoluten Anzahl der Fadingeinbrüche insgesamt geringer als bei kontinuierlichem Betrieb.

Zudem wurden jeweils die Indices (über Frequenz und Modenrüher-/Antennenkombination bzw. Samplenummer, d.h. n, l) der Messwertsequenzen, die jeweils zu einem Fadingeinbruch von mindestens 1, 2, 3, ..., 35 dB unter dem Mittelwert führten, in einem Array gespeichert. Somit entstanden 35 Arrays von Indices je Messwertsequenz und je kontinuierlicher / schrittweiser Betrieb der MVK. Von den jeweils viermal 35 Arrays bzw. Mengen von Indices je Betriebsmodus wurden jeweils 35 Schnittmengen und Vereinigungsmengen über die Wiederholungen, d.h. M, gebildet.
Die Schnittmengen haben die Bedeutung, dass bei allen Messwertsequenzen für die enthaltenen Indices ein entsprechend tiefer Fadingeinbruch erfolgte.
Die Vereinigungsmengen haben hingegen die Bedeutung, dass für jeden enthaltenen Index (n, l) es mindestens eine Messwertsequenz für diesen Betriebsmodi gibt, bei dem ein entsprechend tiefer Fadingeinbruch erfolgte. Die jeweilige Anzahl der Elemente in den Schnitt- und Vereinigungsmengen ist ebenfalls in den Abbildungen 10 und 11 gezeigt.

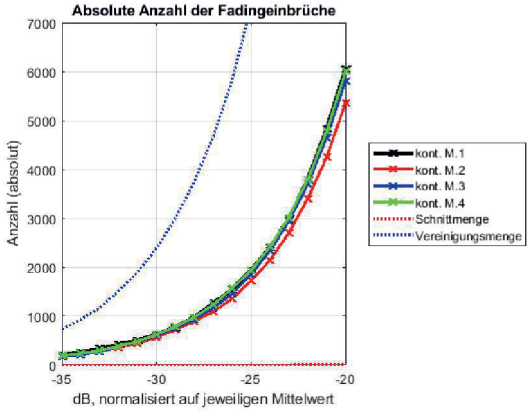


Abb. 10: Absolute Anzahl der Fadingeinbrüche bei kontinuierlichem Betrieb der Modenrührer.

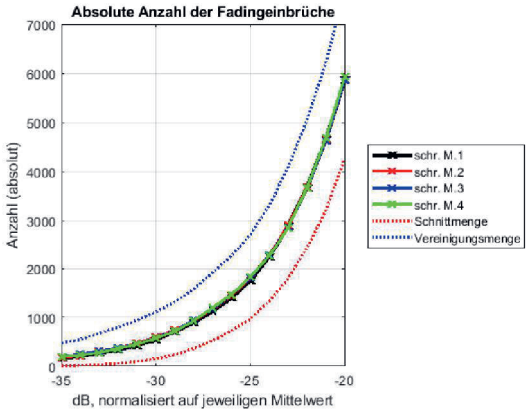


Abb. 11: Absolute Anzahl der Fadingeinbrüche bei schrittweisem Betrieb der Modenrührer.

An diesen jeweiligen Indices (n, l) wurden zudem die Messwertsequenzen verglichen, d.h. der Mittelwert von den Differenzen der Fadingtiefe über die Anzahl der Indices. Die Differenz der Fadingtiefe bezeichnet hier die Differenz zwischen dem Maximum (über m) und dem Minimum (über m) von $|\overline{S_{21}}|^2(n, l, m)$ für diesen Fadingeinbruch.
In den Abbildungen 12 und 13 sind diese mittleren Differenzen in dB zwischen den vier Wiederholungen (d.h. M) über der Fadingtiefe dargestellt. Generell nehmen die mittleren Differenzen der absoluten Fadingeinbrüche für tiefere Fadingeinbrüche zu.
Für den kontinuierlichen Betrieb der Modenrührer (Abb. 12) ist die Schnittmenge der Indices, die zu einem Fadingeinbruch von mindestens 28 dB unter dem Mittelwert führen, leer, weshalb in diesem Bereich die Differenz nicht bestimmt werden kann. Die mittleren Differenzen für die Vereinigungsmenge der Indices sind stets höher als für die Schnittmenge der Indices. Ab einer Fadingtiefe von mindestens 10 dB unter dem Mittelwert verläuft die mittlere Differenz in dB für die Vereinigungsmenge nahezu linear mit steigender Einbruchtiefe, wobei die mittlere Differenz jeweils ca. 5 dB höher ist als die Einbruchtiefe unterhalb des Mittelwertes.
Dies bedeutet beispielsweise, dass für einen Index, der in mindestens einer Messwertsequenz einen Fadingeinbruch von 30 dB repräsentiert, im Vergleich zu den anderen Messwertsequenzen mittlere Abweichungen von 35 dB zu erwarten sind. Folglich sind diese tiefen Fadingeinbrüche bei kontinuierlichem Modenrührerbetrieb nicht wiederholbar.
Für den schrittweisen Betrieb der Modenrührer (Abb. 13) sind sowohl die Schnittmenge als auch die Vereinigungsmenge in dem dargestellten Bereich stets nicht leer.

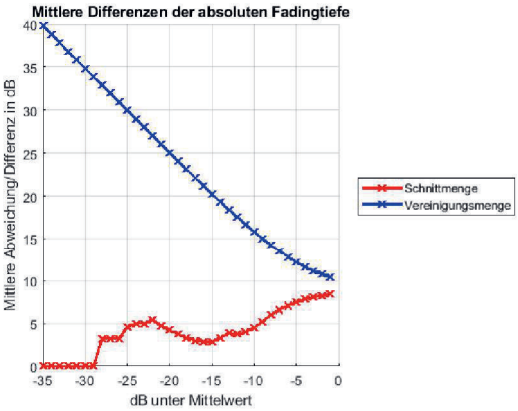


Abb. 12: Mittlere Differenzen in dB der absoluten Fadingtiefe bei kontinuierlichem Betrieb der Modenrührer.

Zudem sind die mittleren Differenzen in dB deutlich geringer als bei kontinuierlichem Betrieb der Modenrührer. Die mittleren Differenzen verlaufen für die Schnitt- sowie die Vereinigungsmenge der Indices nahezu identisch bis zu einer Fadingtiefe von ca. 20 dB unter dem Mittelwert. Für tiefere Fadingeinbrüche weichen die mittleren Differenzen bereichsweise voneinander ab. Bei einer Fadingtiefe von 30 dB unter dem Mittelwert beträgt die mittlere Differenz jeweils ca. 10 dB für beide Mengen. Dies bedeutet, dass für einen Index, der einen Fadingeinbruch von 20 dB repräsentiert, im Vergleich zu den anderen Messwertsequenzen mittlere Abweichungen von ca. 5 dB zu erwarten sind. Als Konsequenz können die Fadingeinbrüche bei schrittweisem Betrieb der Modenrührer und beibehaltender Konfiguration der MVK als wiederholbar angesehen werden.

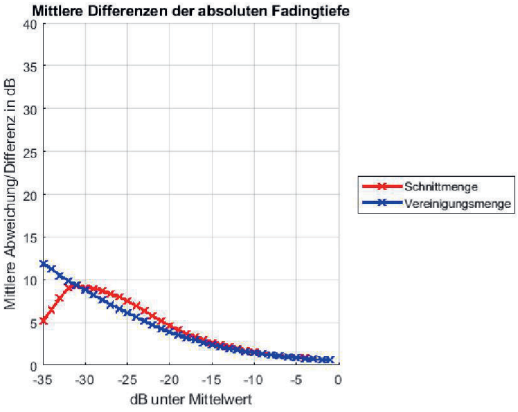


Abb. 13: Mittlere Differenzen in dB der absoluten Fadingtiefe bei schrittweisem Betrieb der Modenrührer.

Zeitliches Verhalten der Messungen

In diesem Teil werden die Log-Dateien, welche die Messzeitpunkte protokollieren, untersucht. Dabei sind insbesondere die Zeitdifferenzen zwischen den einzelnen Ereignissen der Messwertaufnahme von Interesse. In Tabelle 4 sind Minimum, Mittelwert, Maximum und Streuung der Zeitdifferenzen in Sekunden zwischen den einzelnen Frequenzsweeps für den kontinuierlichen Betrieb der Modenrüher angegeben. Dabei wurden jeweils die Zeiten für das Umschalten zur nächsten fest installierten Antenne von ca. 0,4 – 0,5 s nicht in die Betrachtungen mit einbezogen.
Bei allen vier Wiederholungen liegen die Werte dicht beieinander. Der minimale zeitliche Abstand zwischen zwei aufeinander folgenden Frequenzsweeps betrug bei allen Wiederholungen jeweils knapp 90 ms und lag somit unter der nominell eingestellten Dauer von 100 ms. Im Mittel wurde jeweils die nominell eingestellte Dauer überschritten.
Die maximalen Werte betrugen jeweils etwa das Doppelte des Mittelwerts.
Obwohl die Unterschiede zwischen den Mittelwerten der Zeitdifferenzen gering sind, bedeutet diese Abweichung, dass sich bei insgesamt $N = 600$ Frequenzsweeps die Gesamtdauer der Messungen zwischen den Wiederholungen um mehr als 2 s unterscheiden. Dies entspricht wiederum der vielfachen Dauer eines Frequenzsweeps. Somit begründet sich, wie zuvor bereits gezeigt, dass die Messwertsequenzen bei kontinuierlichem Modenrührerbetrieb unkorreliert sind.

Tab. 4: Zeitdifferenzen in s zwischen den einzelnen Frequenzsweeps bei kontinuierlichem Betrieb der Modenrührer.

Zeit in s	kont. M.1	kont. M.2	kont. M.3	kont. M.4
Min.	0,086	0,088	0,088	0,087
Mittelwert	0,111	0,115	0,115	0,114
Max.	0,213	0,212	0,232	0,234
Streuung	0,015	0,015	0,016	0,015

In Tabelle 5 sind Minimum, Mittelwert, Maximum sowie Streuung der Zeitdifferenzen für den schrittweisen Betrieb der Modenrührer und die vier Messwertsequenzen angegeben.
Bei jeder dieser schrittweisen Messwertsequenzen wird eine Modenrührerkombination angefahren und bei dieser Position dann zu den weiteren Antennen umgeschaltet, wie im Abschnitt „Messaufbau“ beschrieben. Daher sind die Mittelwerte und Streuungen für das jeweilige Umschalten sowie das Fahren der Modenrührer zusätzlich getrennt angegeben.

Das mechanische Fahren der Modenrührer dauert im Mittel ca. 1,8 s, wohingegen das Umschalten zur nächsten Antenne inklusive eines Frequenzsweeps in ca. 0,4 s abgeschlossen ist. Auffällig ist zudem, dass die Streuungen der Umschaltzeiten ohne ein Fahren der Modenrührer ca. 10 ms betragen und damit geringer sind als die Streuungen für den kontinuierlichen Modenrührerbetrieb (ohne ein Umschalten der Antennen).

Tab. 5: Zeitdifferenzen in s zwischen den einzelnen Frequenzsweeps bei schrittweisem Betrieb der Modenrührer.

Zeit in s	schr. M.1	schr. M.2	schr. M.3	schr. M.4
Min.	0,416	0,417	0,415	0,408
Mittelwert	0,879	0,880	0,880	0,866
Max.	1,876	1,848	1,887	1,851
Streuung	0,635	0,634	0,634	0,640
Umschalten zu Ant. 2: Mittelwert / Streuung	0,436 / 0,010	0,437 / 0,012	0,437 / 0,011	0,417 / 0,005
Umschalten zu Ant. 3: Mittelwert / Streuung	0,436 / 0,007	0,437 / 0,009	0,435 / 0,006	0,420 / 0,010
Fahren und Umschalten zu Ant. 1: Mittelwert / Streuung	1,770 / 0,145	1,771 / 0,145	1,770 / 0,146	1,765 / 0,145

Fazit

In diesem Beitrag wurde gezeigt, dass bei einer MVK zunächst unabhängig davon, ob die Modenrührer kontinuierlich oder schrittweise bewegt werden, die gleichen mittleren Leistungsübertragungsfunktionen und gleichen Fadingeinbrüche erzielt werden. Jedoch sind bei zeitlich nicht-synchronem, kontinuierlichem Betrieb der Modenrührer die aufgenommenen Messwertsequenzen zwischen den Wiederholungen unkorreliert. In Folge dessen sind bei kontinuierlichem Betrieb für Wiederholungen der Messsequenzen zwar die statistischen Eigenschaften konstant, nicht jedoch die Abfolge der Messwerte. Somit ist hierbei für bestimmte Samplenummern das spezifische frequenzselektive Verhalten nicht reproduzierbar.

Im Gegensatz dazu sind bei schrittweisem Betrieb der Modenrührer die Messwertsequenzen stark korreliert. Darüber hinaus sind für eine bestimmte Frequenz und Modenrüherkombination auch die Fadingeinbrüche gut reproduzierbar. Daher können die Messwertsequenzen bei Wiederholungen als nahezu identisch angesehen werden.

Insgesamt sollte folglich dessen eine MVK ausschließlich schrittweise betrieben werden, wenn eine Reproduzierbarkeit der Messwertsequenzen oder Fadingeinbrüche gefordert ist.

Literaturnachweis

[1] DIN EN 61000-4-21:2011-12; VDE 0847-4-21:2011-12, Elektromagnetische Verträglichkeit (EMV) - Teil 4-21: Prüf- und Messverfahren – Verfahren für die Prüfung in der Modenverwirbelungskammer (IEC 61000-4-21:2011); Deutsche Fassung EN 61000-4-21:2011, 2011

[2] 3GPP TR 25.914 V15.0.1, 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Measurements of radio performances for UMTS terminals in speech mode (Release 15)

[3] C. L. Holloway, et al.: On the Use of Reverberation Chambers to Simulate a Rician Radio Environment for the Testing of Wireless Devices, *IEEE Transactions on Antennas and Propagation* 54 (11), 3167-3177 (2006), DOI: 10.1109/TAP.2006.883987

[4] P.S. Kildal, et al.: Characterization of Reverberation Chambers for OTA Measurements of Wireless Devices: Physical Formulations of Channel Matrix and New Uncertainty Formula, *IEEE Transactions on Antennas and Propagation* 60 (8), 3875-3891 (2012), DOI: 10.1109/TAP.2012.2201125

[5] C. Cammin, D. Krush, R. Heynicke, G. Scholl: Test method for narrowband F/TDMA-based wireless sensor/actuator networks including radio channel emulation in severe multipath environments, *J. Sens. Sens. Syst.*, 7, 183-192 (2018), DOI: 10.5194/jsss-7-183-2018

[6] Bluetest AB: *RTS60 Reverberation Test System*, Datasheet, Rev. G, 2014. URL: https://bluetest.se/sites/default/files/RTS60_BTD-10-019_Rev_G.pdf (Download: 20.04.2019)

[7] C. Cammin, D. Krush, R. Heynicke, G. Scholl: Charakterisierung einer Modenverwirbelungskammer als Testumgebung für drahtlose Sensor/Aktor-Module, *tm - Technisches Messen* 85(6), 420-433 (2018) DOI: 10.1515/teme-2018-0003

[8] C. Cammin, D. Krush, R. Heynicke, G. Scholl: Employing Correlation for Wireless Device Characterization in Reverberation Chambers, 19. ITG/GMA Fachtagung Sensoren und Messsysteme, 2018, Nürnberg, erschienen in *ITG-Fb. 281: Sensoren und Messsysteme,* VDE Verlag

[9] C. Cammin, D. Krush, R. Heynicke, G. Scholl: Employing correlation for wireless component and device characterization in reverberation chambers, *J. Sens. Sens. Syst. 8*, 185-194 (2019), DOI: 10.5194/jsss-8-185-2019

[10] M. Pätzold: *Mobilfunkkanäle*, Vieweg+Teubner Verlag, Wiesbaden, 1999. ISBN: 978-3-663-05656-0

DOI 10.5162/sensoren2019/P2.14

Intelligenter Zahnriemen

Dominik Großkurth[1], Georg Martin[2]
[1]Fachgebiet Integrierte Elektronische Systeme, TU Darmstadt, Merckstraße 25, 64283 Darmstadt
[2]Fachgebiet Produktentwicklung und Maschinenelemente, TU Darmstadt, Otto-Berndt-Straße 2, 64287 Darmstadt

Zusammenfassung

In dieser Arbeit wird ein Messsystem beschrieben, mit dem es möglich ist Messwerte innerhalb eines Zahnriemens zu erheben und diese, während des Betriebs des Zahnriemens, an eine externe Stelle zu senden.
Ein solches Messsystem ermöglicht die Zustandsüberwachung des Zahnriemens im Betrieb und ist damit ein weiterer Schritt hin zu einer vollständig vernetzten Produktion. Das Hauptaugenmerk wird dabei auf den Verschleiß des Zahnriemens gelegt, da dieser aktuell nur umständlich durch Personal vor Ort ermittelt werden kann. Als Indikator dienen dabei die Vorspannkraft des Zahnriemens, welche indirekt aus seiner Eigenfrequenz ermittelt wird, und die Temperatur im Inneren des Zahnriemens. Zum Aufnehmen dieser beiden Messgrößen wurde ein Sensorknoten entwickelt, der klein genug ist, um in einen Zahn des Zahnriemens eingesetzt werden zu können.
Erste Messungen bestätigen, dass es möglich ist mittels Betriebsschwingungsdaten Rückschlüsse auf die Vorspannkraft und damit auf den Verschleiß des Zahnriemens zu ziehen.
Zudem wird die mechanische Belastung des Sensorknotens näher untersucht und eine verbesserte Version vorgestellt. Die neue Version benötigt signifikant weniger Bauraum und kann voraussichtlich besser mit den beschrieben Belastungen umgehen.

Keywords: Zustandsüberwachung, Zahnriemen, Schwingungsmessung

Problemstellung

In Industrieanlagen und Produktionsmaschinen wird im Rahmen von Industrie 4.0 zunehmend eine zustandsorientierte Instandhaltung umgesetzt, die auf der Zustandsüberwachung (Condition Monitoring) von Bauteilen beruht. Dies ermöglicht es, Wartungskosten zu reduzieren und ungeplante Maschinenausfälle zu vermeiden.
In vielen industriellen Anlagen werden Zahnriemengetriebe mit Hochleistungszahnriemen aus Polyurethan eingesetzt, um Bewegungen zu übertragen. Sie bestehen typischerweise aus Zugträgern, meist Stahlseile, und einer Matrix aus Polyurethan, die auch die Zähne bildet, vgl. Abbildung 1.
Im einfachsten Fall besteht ein Zahnriemengetriebe aus einer Antriebsscheibe und einer Abtriebsscheibe, die vom Zahnriemen umschlungen werden. Die Betriebskraft wird in den Zahnscheiben über die PU-Matrix in den Riemen übertragen, im Riemenabschnitt zwischen zwei Scheiben (sog. Trum) wird die Kraft hauptsächlich durch die Zugträger geleitet.

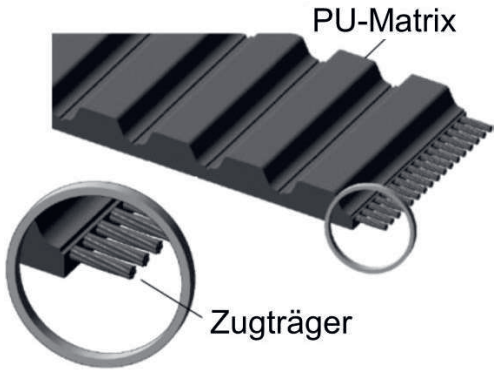


Abbildung 1. Aufbau eines Polyurethan-Zahnriemens, Abb. nach [1]

Ein Zahnriemen benötigt prinzipbedingt eine Vorspannkraft F_{TV}, um effektiv und effizient zu funktionieren, vgl. Abbildung 2. Bei zu geringer Vorspannkraft findet der Zahneingriff in der Riemenscheibe nicht mehr korrekt statt, was zu Schwingungen im Riemen und der umgebenden Maschine sowie zu erhöhtem Verschleiß führt.

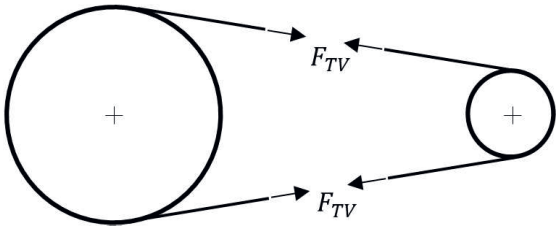


Abbildung 2. Vorspannkraft in einem Zahnriemengetriebe

Im Verlauf der Betriebszeit verschleißen die Zugträger, infolgedessen nimmt die Steifigkeit des Zahnriemens ab. Dies führt zu einer Abnahme der Vorspannkraft, einer verschlechterten Laufruhe und einem erhöhten Verschleiß, sodass bei starkem Zugträgerverschleiß der Zahnriemen ausgetauscht werden muss [2]. Wenn der Riemen nicht rechtzeitig getauscht wird, kann der Zugträgerverschleiß fortschreiten, bis es zum Riss des Zahnriemens und damit zum Ausfall der gesamten Maschine kommt.

Der Zugträgerverschleiß begrenzt in vielen industriellen Anwendungen die Lebensdauer von PU-Zahnriemen. Allerdings existiert kein Modell zur Berechnung der Lebensdauer eines Zahnriemengetriebes, wie dies z.B. für Stirnradgetriebe etabliert ist. Aufgrund dieser Unsicherheit bezüglich der Lebensdauer treten einerseits immer wieder ungeplante Ausfälle auf, andererseits werden viele Zahnriemen vor Erreichen ihres tatsächlichen Lebensdauerendes vorsorglich getauscht. Eine Zustandsüberwachung, welche das Erreichen eines kritischen Verschleißmaßes erkennt, kann deshalb einen wichtigen Beitrag zu einer kosten- und ressourceneffizienteren Produktion leisten. Allerdings sind bisher, anders als z.B. für Wälzlager, noch keine Zustandsüberwachungssysteme für Zahnriemen verfügbar. Die Überwachung des Riemenzustands kann bisher nur durch die manuelle Kontrolle der Vorspannkraft erfolgen.

Gemessen wird diese Kraft, indem der Riemen durch manuelles Anschlagen breitbandig zu Schwingungen angeregt wird und aus der Schwingungsantwort die Eigenfrequenzen bestimmt werden. Die Eigenfrequenzen f_λ sind, wie bei einer Gitarrensaite, durch die Beziehung

$$f_\lambda = \lambda \cdot \sqrt{\frac{F_{TV}}{4 \cdot m' \cdot l^2}}$$

mit der Vorspannkraft verknüpft. Dabei ist m' das Längengewicht in kg/m, l die frei schwingende Trumlänge zwischen zwei Scheiben und F_{TV} die Vorspannkraft. λ bezeichnet die Ordnung der Eigenfrequenz, d.h. $\lambda = 1$ bezeichnet die erste Eigenfrequenz mit einem Schwingungsbauch, $\lambda = 2$ die zweite Eigenfrequenz mit zwei Schwingungsbäuchen und einem Schwingungsknoten, etc. Dieses Messverfahren ist etabliert und gilt als das Vorspannkraft-Messverfahren mit der höchsten Genauigkeit [1]. Entsprechende Messgeräte sind von diversen Herstellern auf dem Markt verfügbar.

Die regelmäßige manuelle Kontrolle der Vorspannkraft ist aufgrund des Personalaufwands und der schlechten Zugänglichkeit in den meisten Fällen nicht möglich. An der TU Darmstadt wird deshalb in einem Kooperationsprojekt mit BRECO Antriebstechnik Breher GmbH & Co. KG und CONTECS Engineering Services GmbH ein Zustandsüberwachungssystem entwickelt, das auf einer schwingungsbasierten Vorspannkraftmessung ohne externe Anregung, nur anhand der Betriebsschwingungen, basiert. Diese Zustandsüberwachung soll nicht nur ungeplante Ausfälle verhindern, sondern auch eine optimierte und damit kosteneffizientere Auslegung des Riemengetriebes ermöglichen, sowie Folgeschäden an der Maschine oder eine mangelhafte Qualität von Fertigungsprozessen vermeiden, indem die Laufruhe des Riementriebs sichergestellt wird.

Lösungsvorschlag

Zur Messung von Parametern, die potentiell auf den Verschleiß eines Zahnriemens hinweisen, wurde daher ein Sensorknoten entwickelt. Dieser ermöglicht das Erfassen von Beschleunigungen sowie der Temperatur und ist in der Lage diese Messwerte drahtlos an eine Basisstation zu übertragen.

Dabei wird die Beschleunigung an zwei möglichst weit voneinander entfernten Punkten (35mm) in allen drei Raumrichtungen aufgenommen. Die Position der Beschleunigungsmesser auf der Platine ist in Abbildung 4 zu sehen.

Zur besseren Beschreibung der Schwingungsrichtungen wird ein Koordinatensystem für den Zahnriemen definiert, zu sehen in Abbildung 3, auf das sich im Folgenden bezogen wird. Mithilfe der Beschleunigungssensoren ist es möglich die Transversalschwingungen des Zahnriemens zu messen. Diese treten in z-Richtung auf und werden durch den Polygon-Effekt angeregt. Dieser Effekt tritt auf wenn ein Zugmittel über ein Antriebsrad formschlüssig angetrieben wird. Zugmittel wie Zahnriemen oder Ketten bestehen aus einzelnen verhältnismäßig starren Segmenten, wodurch sie nicht kreisrund auf- bzw. ablaufen können. Hierdurch verhält sich das runde Antriebsrad wie ein Polygon dessen Seitenzahl von dessen Durchmesser und der Größe der Segmente des Zugmittels abhängt. Durch die Polygonform der Zahnscheiben ist der wirksame Radi-

us, mit dem der Zahnriemen um die Scheibe läuft, während einer Umdrehung nicht konstant. Deshalb variiert, auch bei gleichförmiger Winkelgeschwindigkeit, die Geschwindigkeit des Zahnriemens und sein Abstand zur Scheibenachse. Diese Ungleichförmigkeiten regen u.a. transversale Schwingungen des Riemens an.
Des Weiteren können Stöße beim Anlaufen an die Bordscheiben erkannt werden, da diese maßgeblich in y- Richtung auftreten. In x-Richtung können Stöße des Antriebs und des Abtriebes erkannt werden, die die anderen Messungen stören können.



Abbildung 3. Koordinatensystem bezogen auf einen Punkt auf dem Zahnriemen

Der zweite Beschleunigungssensor dient dazu, Torsionsschwingungen des Zahnriemens um die Achse der Laufrichtung (x-Achse) erkennen zu können, um zu verhindern, dass Torsionseigenfrequenzen fälschlicherweise als Saiten-Eigenfrequenzen interpretiert werden. Durch diese Torsionsschwingungen käme es zu einer sich periodisch ändernden Phasenverschiebung der z-Beschleunigungssignale der beiden Sensoren, während reine Transversalschwingungen durch gleichphasige Signale charakterisiert sind. Aufgenommen werden die Beschleunigungen mittels eines kompakten und Energie sparenden MEMS Beschleunigungssensors. Dieser erfasst Beschleunigungen mit einer Amplitude von bis zu ±400g, löst diese mit 12Bit auf und stellt sie digitalisiert über eine SPI Schnittstelle zur Verfügung.
Die Zustandsüberwachung basiert darauf, dass ein Zahnriemen beim Zahneingriff zu Schwingungen angeregt wird. Dabei wirken mehrere Anregungsmechanismen, hervorzuheben ist die Anregung durch den Polygoneffekt. Wenn die Schwingungsanregung im Betrieb breitbandig ist, kann die Saiten-Eigenfrequenz anhand dieser gemessenen Betriebsschwingungen bestimmt werden.
Kommt es durch Verschleiß zur Längung des Zahnriemens, so stimmt der Abstand der Zähne nicht länger mit dem Abstand der Lücken auf der Zahnscheibe überein. Während dieses Prozesses erhöht sich allmählich die Reibung der Zahnflanken an den Zahnscheiben und führt so zu einer steigenden Erwärmung des Zahnriemens. Aus diesem Grund wurde als zweite Messgröße die Temperatur ausgewählt. Aufgenommen wird sie durch einen separaten Sensor, dieser ermöglicht die Messung von Temperaturen von –40°C bis +125°C, löst diese mit 12 Bit auf und kann über eine I2C Schnittstelle ausgelesen werden.
Als Mikrocontroller wurde der BGM121 [3] der Firma Silicon Labs eingesetzt. Dieser ist ein System in Package (SiP), er setzt auf einen ARM Cortex-M4 und enthält bereits passive Komponenten sowie eine integrierte Antenne über die er mittels Bluetooth Low Energy mit der Außenwelt kommunizieren kann. Dies ist nötig, da der Zahnriemen im Betrieb mit hohen Geschwindigkeiten rotiert und nicht drahtgebunden kontaktiert werden kann. Aus demselben Grund wurde für die Energieversorgung während der Messungen ein Li-Polymer-Akku eingesetzt.

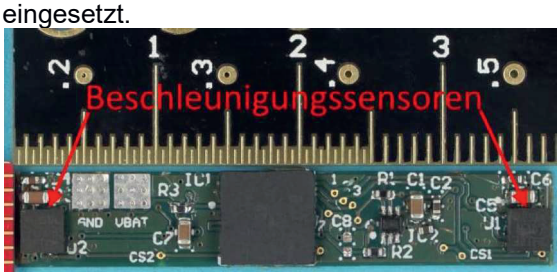


Abbildung 4. Sensorknoten zur Schwingungsmessung innerhalb von Zahnriemen

Abbildung 4 zeigt die bestückte Platine des Sensorknotens. An den äußeren Rändern befinden sich die beiden Beschleunigungssensoren, als größte Komponente ist der Mikrocontroller in der Mitte angeordnet und rechts neben ihm befindet sich der sechsbeinige Temperatursensor.

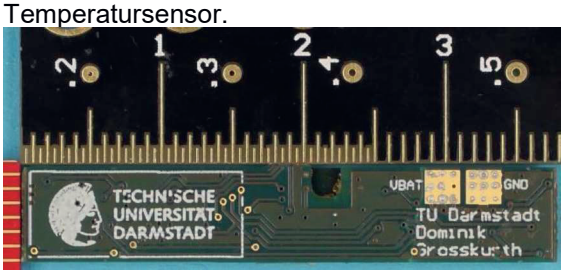


Abbildung 5. Unterseite des Sensorknotens

Als größte Herausforderung stellte sich die Unterbringung des Sensorknotens innerhalb des Zahnriemens heraus. Da der Wunsch bestand, Zahnriemen auch auf Prüfständen mit Rückenrolle einsetzen zu können, konnten keine Komponenten auf der Rückseite des Riemens angebracht werden. Damit verblieb als einzige Option die Einbringung innerhalb eines Zahnes. Daher wurde eine etwa 4x7x47mm große Aussparung in zwei Zähne des Zahnriemens gefräst, um in einem der Zähne den Sensorknoten und in dem anderen Zahn den Li-Polymer-Akku unterzubringen. Verbunden wurden beide Zähne mit Leitern,

die nachträglich im Zwischenraum der Zugträger verlegt wurden. Verfüllt wurde der mit dem Sensorknoten bestückte Zahn mit einem Schmelzklebstoff aus Ethylen-Vinylacetat-Copolymer (EVA), da dieser eine ähnliche Steifigkeit und Verarbeitungstemperatur wie das Matrixmaterial Polyurethan der Zahnriemen aufweist. In Abbildung 6 ist die so vergossene Elektronik zu sehen, bevor das überschüssige Vergussmaterial entfernt wurde.
Eine Vorverarbeitung der Sensorsignale innerhalb des Sensorknotens ist nicht gewünscht, so ergibt sich die nötige Datenrate aus den drei Achsen der zwei Beschleunigungssensoren und deren Abtastrate. Die Messwerte werden zur leichteren Handhabung auf 16Bit erweitert und die Abtastrate wird auf 800Hz festgelegt, da mit Frequenzen <200Hz gerechnet wird.
Zum Erreichen einer möglichst großen Datenrate wurde die Paketgröße der Funkübertragung von standardmäßig 23Byte auf 241Byte erhöht und das Verbindungsintervall auf 7,5ms reduziert. Somit war eine stabile Übertragung von allen Sensorwerten gewährleistet (2·3·16Bit·800Hz = 76,8kBit/s).

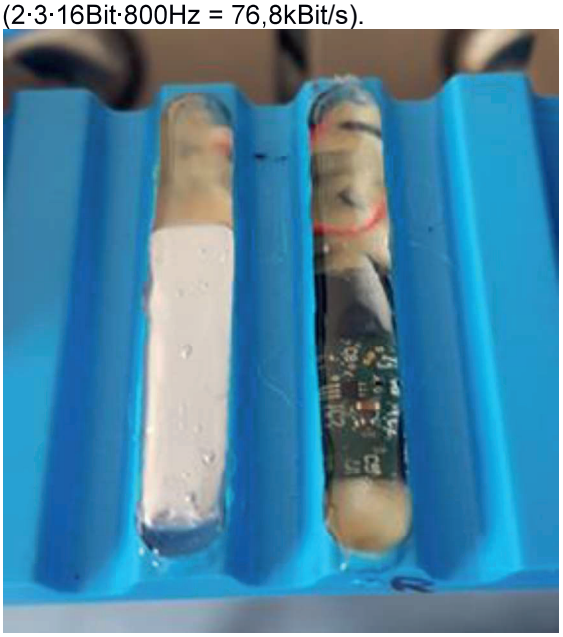


Abbildung 6. In den Zahnriemen eingebetteter Sensorknoten und Energiespeicher

Empfangen werden die Daten ebenfalls von einem Mikrocontroller der Firma Silicon Labs, dieser übermittelt die Daten anschließend per USB an einen Computer, wo die Daten von einem Programm entgegengenommen werden. Dieses Programm ermöglicht dem Anwender die ankommenden Daten als FFT-Plot zu sehen oder die Rohdaten abzuspeichern. Außerdem ist eine Konfiguration der Beschleunigungssensoren möglich. Sie bieten die Möglichkeit den Messbereich zu ändern (±400g, ±200g, ±100g), um kleinere Amplituden präziser auflösen zu können und zudem einen internen Hochpassfilter zuzuschalten (0Hz, 2,5Hz, 5Hz, 10Hz, 20Hz).

Experimentelle Untersuchungen

Um die grundsätzliche Eignung einer riemenintegrierten Schwingungsmessung für die Zustandsüberwachung zu überprüfen, wurde ein Versuch aufgebaut. Dabei wurden zwei Fragestellungen untersucht: Erstens sollte geprüft werden, ob die Schwingungsanregung so breitbandig ist, dass die Eigenfrequenzen im Spektrum sichtbar sind. Zweitens sollte die Haltbarkeit der Sensorelektronik unter den mechanischen Beanspruchungen, insbesondere durch die Zentripetalkraft und die Biegung beim Scheibendurchlauf, untersucht werden.
Der Versuchsaufbau besteht aus einem einfachen Riementrieb mit zwei Scheiben. Eine Scheibe ist angetrieben, die zweite Scheibe stellt lediglich eine Umlenkung ohne anliegendes Drehmoment dar. Der beschriebene Versuchsaufbau wurde zum einem vor Ort an der TU Darmstadt aus Aluminium-System-Profilelementen aufgebaut und entspricht einer Maschine mit geringer Steifigkeit. Zum anderem wurde ein Prüfstand der Firma BRECO Antriebstechnik verwendet der sowohl eine hohe Masse als auch eine hohe Steifigkeit aufweist.
Während der ersten Messungen traten nach wenigen Minuten Verbindungsabbrüche auf. Nachdem ausgeschlossen wurde, dass Metallteile die Funkstrecke blockieren, wurde der Riemen gezielt mechanischen Belastungen in verschiedenen Richtungen ausgesetzt. Hierbei zeigte sich, dass die Verbindungsprobleme bei Torsion des Zahnriemens oder Biegen entlang eines Zahnes (y-Richtung) auftreten. Die Platine wurde anschließend außerhalb eines Riemens diesen Belastungen ausgesetzt. Es stellte sich heraus, dass sich die Platine gleichmäßig über ihre gesamte Länge biegt. Das mit 6,5mm Kantenlänge verhältnismäßig lange System in Package (SiP, BGM121) verformt sich bedingt durch sein aus Duroplast bestehendes Gehäuse jedoch nicht. Dies führt zu einer mechanischen Belastung der Lötstellen unterhalb des SiP und zu deren Bruch. In Abbildung 9 sind diese Stellen, sowie potentiell problematische Stellen an den Beschleunigungssensoren markiert. Abbildung 7 zeigt ein Foto eines gebogenen Sensorknotens. Die Biegung entsteht hierbei durch das aufbringen einer Kraft an den Enden der Platine mittels einer Klemmvorrichtung. Eine Vergrößerung des kritischen Abschnitts ist in Abbildung 8 zu sehen.



Abbildung 7. Durch Einspannen verformter Sensorknoten

Die Ursprüngliche Vergussmasse aus EVA weist eine relativ hohe Steifigkeit auf und zwingt den Sensorknoten sich im gleichen Maße wie den Zahnriemen zu verbiegen. In weiteren Versuchen wurde daher der Zahn mit weicherem Silikon (1-K-Polysiloxan) vergossen, um die mechanischen Belastungen zu reduzieren. Zudem wurden die Stellen an denen die Beschleunigungssensoren und das SiP platziert sind mit Epoxidharz versteift. Auch hier erwiesen sich die mechanischen Belastungen im Betrieb als zu hoch.



Abbildung 8. Bruch der Lötstellen unterhalb des System in Package

Verstärkt wurde diese Problematik durch die Schwächung der Platine unterhalb des Mikrocontrollers. Diese wurde vorgenommen, um die Abstrahlcharakteristik der darüber liegenden Antenne im SiP nicht zu verändern. Abbildung 5 zeigt die Elektronik von unten, dort ist die Schwächung zwischen der 2cm und der 2,2cm Markierung an der oberen Seite zu erkennen.

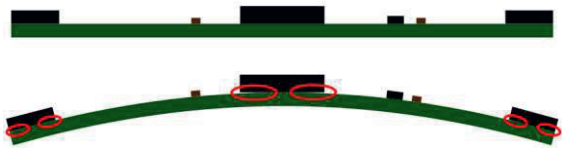


Abbildung 9. Belastung der Lötstellen des Sensorknotens während des Betriebs

Erst das Fixieren des Sensorknotens am Zahngrund mit einem flexiblen Kontaktklebstoff (Polychloropren) und der Verzicht auf das Verfüllen des Zahns führte dazu, dass über mehrere Stunden stabil gemessen werden konnte.

Abbildung 10 zeigt das Spektrogramm eines Hochlaufs bis 340U/min. Deutlich sichtbar ist, dass das Schwingungsverhalten insgesamt von der Zahneingriffsfrequenz und ihren höheren Ordnungen dominiert wird. Aber auch die erste Eigenfrequenz und höhere Eigenfrequenzen bis zur 6. Ordnung (λ=6) werden über den gesamten Drehzahlbereich angeregt. In den Resonanzdurchfahrten, d.h. wenn die Zahneingriffsfrequenz einer Eigenfrequenz entspricht, treten deutliche Amplitudenüberhöhungen auf. Diese Ergebnisse deuten darauf hin, dass im Betrieb eine ausreichend starke und breitbandige Schwingungsanregung stattfindet, um die Eigenfrequenz des Riemens zu bestimmen.

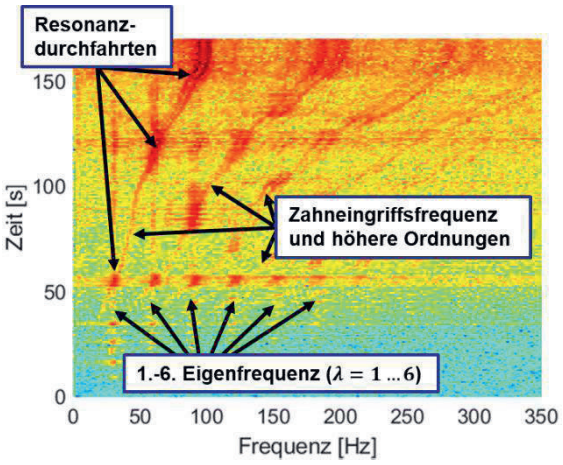


Abbildung 10. Spektrogramm des sensorintegrierten Beschleunigungssensors bei einem Hochlauf

Ausblick

Mit diesen Ergebnissen wurde eine neue Version des Sensorknotens entwickelt. Bei dieser wird auf einen zweiten Beschleunigungsmesser verzichtet. Zudem wird ein Mikrocontroller verwendet, der keine integrierte Antenne besitzt und im Wafer Level Package vorliegt (WLCP). So kann die Größe der Elektronik von 6,5x37x2mm ≈ 480mm³ auf 3,2x25x1,75mm ≈ 130mm³ um den Faktor 3,7 reduziert werden.
Abbildung 11 zeigt den neuen Sensorknoten eingelegt in einen AT10 Zahnriemen der Firma BRECO Antriebstechnik. Dies ermöglicht die Unterbringung in kleineren Zähnen, wodurch es möglich wird kleinere Zahnriementypen zu untersuchen.

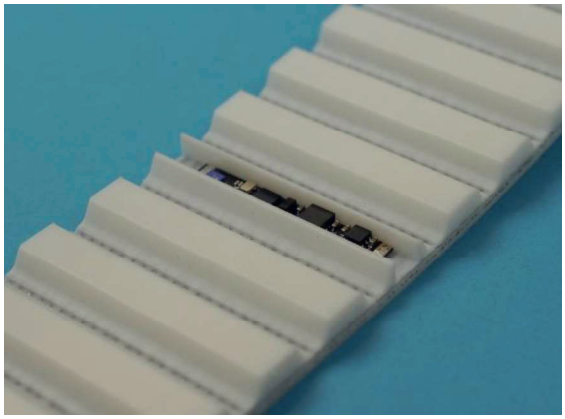


Abbildung 11. Neuer Sensorknoten in einem AT10 Zahnriemen der Firma BRECO Antriebstechnik

Außerdem wurde bei dieser Version eine 0,2mm dicke, flexible Leiterplatte aus Polyimid verwendet. Die Bauteile wurden dabei in kleinen Gruppen angeordnet, sodass unterhalb der Bauteile jeweils eine 0,5mm dicke Versteifung aus FR4 angebracht werden kann, um diese Bereiche vor mechanischen Belastungen zu schützen. In Abbildung 12 ist ein Foto mit einer verformten, unbestückten Platine zu sehen bei der die entsprechenden Biegestellen markiert wurden.

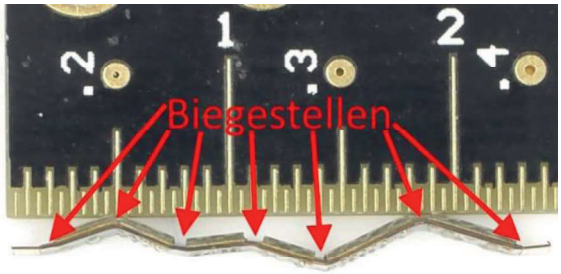


Abbildung 12. Flexible Leiterplatte mit Versteifungen

Mittels dieser Elektronik müssen nun weitere Versuche in verschiedenen Produktionsanlagen stattfinden, um die mechanische Stabilität dieses Ansatzes zu verifizieren.
Die bei diesen Versuchen gewonnen Messergebnisse dienen im weiterem zur Verifikation der Annahme, dass der Verschleiß des Zahnriemens mit an der Abnahme der Vorspannkraft abgelesenen werden kann und dass diese indirekt über die Eigenfrequenz, während des Betriebes ermittelt werden kann.

Literaturnachweis

[1] T. Nagel, *Zahnriemengetriebe: Eigenschaften, Normung, Berechnung, Gestaltung*. München: Hanser, 2008.

[2] R. Perneder, *Handbuch Zahnriementechnik: Grundlagen, Berechnung, Anwendungen*. Berlin: Springer, 2009.

[3] „BGM121/BGM123 Blue Gecko Bluetooth SiP Module Data Sheet“. Silicon Laboratories Inc.

DOI 10.5162/sensoren2019/P2.15

Energieautark arbeitendes Indoor-Smart-Gardening-System mit drahtlosem Monitoring und automatisierter Bewässerung

Elke Mackensen, Axel Rombach, Aaron Spitznagel, Julius Klose
Hochschule Offenburg; Fakultät Elektrotechnik, Medizintechnik und Informatik;
Badstraße 24; 77652 Offenburg

Zusammenfassung

Smart Home-/Smart-Building-Anwendungen sind ein stetig wachsender Markt. Smart Gardening ist ein Beispiel dafür, Nutzern mehr Komfort und eine bessere Lebensqualität zu Hause oder in Bürogebäuden zu ermöglichen. Im Rahmen dieses Beitrags wird die Entwicklung eines Indoor-Smart-Gardening-Systems mit dem Fokus auf energieautarkes Arbeiten vorgestellt. Herzstück des Systems ist ein 3D-gedruckter Blumentopf für einzelne Pflanzen mit integrierter Elektronik zum Monitoring der wichtigsten Pflanzenparameter und einem integrierten Wasserreservoir mit Tauchpumpe für das automatisierte Bewässern der Pflanze. Energy Harvesting per Solarzellen ermöglicht ein energieautarkes Arbeiten des Blumentopfes. Eine selbstentwickelte Low-Power-Funkschnittstelle im Blumentopf und ein externes Gateway ermöglichen die drahtlose Vernetzung mehrerer Pflanzen. Das Gateway dient zur Auswertung der Pflanzenparameter, der Ansteuerung der im Netzwerk vorhandenen Blumentöpfe und als Benutzerinterface.

Keywords: Smart Home, Smart Gardening, Wireless Sensor Nodes, Energy Harvesting, Low-Cost-Füllstandssensor

Einführung und Stand der Technik

Auf Grund verfügbarer Funktechniken und Fortschritte in der Informationstechnik sind Smart-Home-/Smart-Building-Applikationen ein stark wachsender Markt [1]. Smart Gardening ist ein typisches Beispiel einer Smart-Home-/ Smart-Building-Anwendung. Im Detail bedeutet Smart Gardening per Messung diverse Parameter der Pflanze (z.B. Bodenfeuchte, Licht- und Temperatur-verhältnisse, Luftfeuchtigkeit, Nährstoffgehalt usw.) zu erfassen, auszuwerten und entsprechende Aktionen auf Grund der Messparameter einzuleiten. Dazu gehören z.B. eine automatisierte Bewässerung der Pflanze und/oder den Benutzer darüber zu informieren, welche Aktionen für eine einzelne Pflanze aus einem Pool von *n* Pflanzen notwendig sind.

Smart-Gardening-Systeme stehen im Fokus der Forschung, erste Systeme sind am Markt verfügbar [2]-[9]. Man unterscheidet zwischen Indoor- und Outdoor-Gardening-Systemen, die aus mehreren vernetzten Sensorkonten bestehen, welche die unterschiedlichen Pflanzenparameter messen. In vielen Fällen – insbesondere im Outdoor-Bereich – werden die gemessenen Parameter per Funk zu einem Gateway oder direkt in eine Internet-Cloud übertragen. Software-Anwendungen auf dem Gateway ermöglichen es, die Daten dem Nutzer auf entsprechenden Endgeräten wie PC, Smartphone oder Tablets verfügbar zu machen. Indoor-Smart-Gardening-Systeme unterscheiden sich zumeist in den folgenden Merkmalen (siehe Tab. 1): Art der Pflanzenparameter, die gemessen werden, Grad der Automatisierung (z.B. automatische Bewässerung und Düngung), Art der Energieversorgung und Art der drahtlosen Kommunikationstechnik, welche einen Einfluss auf die Kommunikationsreichweite in einem Netzwerk von *n* Pflanzen und den Energiebedarf der Sensorknoten hat

In der Literatur beschriebene Indoor-Smart-Gardening-Systeme haben oftmals die Nachteile eines externen Wassertanks zur Bewässerung der Pflanze, zudem besteht die Energieversorgung des Systems zumeist aus Primärbatterien, die nach einer bestimmten Zeit ausgewechselt werden müssen. Außerdem wird eine Vernetzung eines Pools aus *n* Planzen nur bedingt unterstützt.

Tab. 1: Vergleich verschiedener Indoor-Smart-Gardening-Systeme

	System 1 [6]; Forschung	**System 2 [7]; Forschung**	**System 3 [8]; Kommerziell**	**System 4 [9]; Kommerziell, aber nicht mehr verfügbar**
Gemessene Pflanzen-parameter	Bodenfeuchte, Temperatur, Luftfeuchtigkeit,	Bodenfeuchte, Licht	Bodenfeuchte	Bodenfeuchte, Temperatur, Licht, Nährstoffgehalt
Ansteuerung von Aktuatoren	Automatische Pflanzen-bewässerung	Automatische Pflanzen-bewässerung	Automatische Pflanzen-bewässerung	Automatische Pflanzen-bewässerung und Düngung
Kommentar	Sensoren/Aktoren sind direkt mit einem Arduino Board verbunden	Externes Wasser-reservoir, Sensoren/Aktoren sind direkt mit einem Raspberry-Pi-Board verbunden	Externes Wasser-reservoir	Die gesamte Elektronik ist in einen Blumentopf integrierte
Energie-versorgung und energetische Lebensdauer	Keine Informationen verfügbar	12 VDC; Keine Informationen zur energetischen Lebensdauer	2 AAA Batterien; Keine Informationen zur energetischen Lebensdauer	4 AA Batterien; energetische Lebensdauer 6 Monate
Drahtlose Kommuni-kationstechnik von den Sensoren/ Aktoren zu einem Gateway und Netzwerk Funktionalität	Kein Gateway; Arduino-Boards sind zu einer Thingspeak Cloud via Ethernet oder Wi-Fi konnektiert, auf die Benutzer per Spartphoe-App zugreifen können	Kein Gateway; Keine Netzwerk-funktionalität. Benutzer bekommen Informationen per email	Keine Netzwerk-funktionalität. Eine LED informiert über zu wenig Wasser im Wasserreservoir	Bluetooth V4.0 BLE; Kein Gateway; Die Elektronik des Blumentopfes kann per Smart Phone App konnektiert werden

In Rahmen von diesem Beitrag wird die Entwicklung eines innovativen Indoor-Smart-Gardening-Systems vorgestellt, dessen Fokus auf energieautarkem Arbeiten, einer einfachen Bedienbarkeit für den Benutzer und die optimale Pflege individueller Pflanzen liegt.

Im Folgenden wird zunächst auf die Architektur und die Funktionalität des Systems eingegangen, danach werden Messergebnisse des Indoor-Smart-Gardening-Systems präsentiert. Ein Schwerpunkt des Beitrags ist insbesondere die Betrachtung des energieautarken Arbeitens des Systems, das durch Energy-Harvesting und eine selbstentwickelte Low-Power-Funkschnittstelle erreicht wird.

Architektur und Funktionalität des Indoor-Smart-Gardening-Systems

Das entwickelte System besteht aus drei Hauptbestandteilen (siehe Abb. 1): Wichtigster Bestandteil ist das Netzwerk aus 3D-gedruckten intelligenten Blumentöpfen für individuelle Pflanzen. In dem Blumentopf ist eine Low-Power-Funkschnittstelle integriert. Ein Raspberry-Pi3-System dient als Gateway für das Netzwerk von *n* Blumentöpfen. Das Gateway dient zur Auswertung der Pflanzenparameter, der Ansteuerung der Blumentöpfe im Netzwerk und als Benutzerschnittstelle für Software-APPs auf Smartphone, PC oder ähnlichem.

Der intelligente Blumentopf (siehe Abb. 2) besteht aus einem integrierten Wasserreservoir mit einer Tauchpume für die automatische Bewässerung der Pflanze und einer integrierten Elektronik zum Monitoring verschiedenen Pflanzenparametern und des Wasserstandes im Wasserreservoir. Energy-Harvesting per Solarzellen ermöglichen ein energieautarkes Arbeiten des Blumentopfes.

Eine detaillierte Architektur der integrierten Elektronik des intelligenten Blumentopfes ist in Abb. 3 dargestellt. Die geerntete Energie wird in einem Akku mit einer Kapazität von

350 mAh gespeichert. Die Elektronik besteht aus diversen Sensoren. Ein Multisensorchip ermöglicht die Erfassung der Umgebungsbedingungen der Pflanzen. Der zum Einsatz kommende Multisensorchip ermöglicht die Messung der Umgebungs-temperatur, der Luftfeuchtigkeit und des Luftdrucks. Ein weiterer Sensor ermöglicht die Messung der Bodenfeuchte, so dass sichergestellt werden kann, dass die Pflanze optimal mit Wasser versorgt wird. Zudem kommt ein selbstentwickelter Sensor zum Einsatz, der den Wasserstand im Wasserreservoir überprüft.

Ein Mikrocontroller und ein Power-Management sind weitere Bestandteile der integrierten Elektronik. Das Power-Management ermöglicht energieautarkes Arbeiten des Blumentopfes. Verschiedene Methoden werden angewendet, um den Energieverbrauch der Elektronik zu minimieren:

- Ausnutzung der Low-Power-Modes der einzelnen Komponenten.
- Einsatz von Komponenten mit niedrigen Leckströmen.
- Verwendung einer selbstentwickelten energieoptimierten Funkschnittstelle für die drahtlose Kommunikation mit dem Gateway, die bei 868 MHz arbeitet.
- Implementierung einer Ein-/Ausschaltlogik für die einzelnen Komponenten. Einzige wirklich dauerhaft betriebene Komponente des Systems ist eine Real-Time-Clock.

Der eingesetzte Mikrocontroller verfügt über einen FRAM-Speicher für die Speicherung der Sensordaten und der Firmware. Der Mikrocontroller erfasst alle Sensordaten, wertet diese aus, initiiert eine Bewässrung und überträgt relevante Daten zu dem Raspberry-PI3-basiernden Gateway. Der Raspberry-PI3 wurde auf Grund seiner vielen standardisierten Schnittstellen ausgewählt. Das Gateway wird mit einer konstanten 5-V-DC-Spannung versorgt. Das Gateway wurde durch die gleiche 868-MHz-Funkschnittstelle erweitertet, welche auch in der integrierten Elektronik des Blumentopfes enthalten ist. Das Gateway empfängt von allen vorhandenen Blumentöpfen die übermittelten Daten und macht sie für weitere Softwareanwendungen verfügbar. Ein lokaler Node-Red-Server ermöglicht den Zugriff auf die Daten per WiFi oder Ethernet.

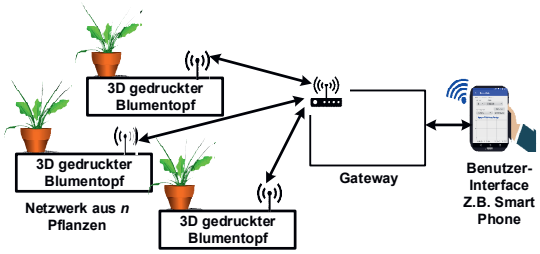


Abb. 1: Vereinfachte Architektur des Indoor-Smart-Gardening-Systems

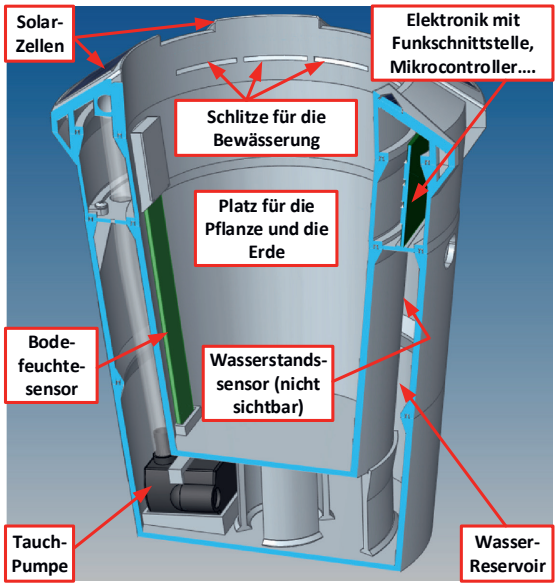


Abb. 2: Querschnitt durch den 3D-gedruckten intelligenten Blumentopf

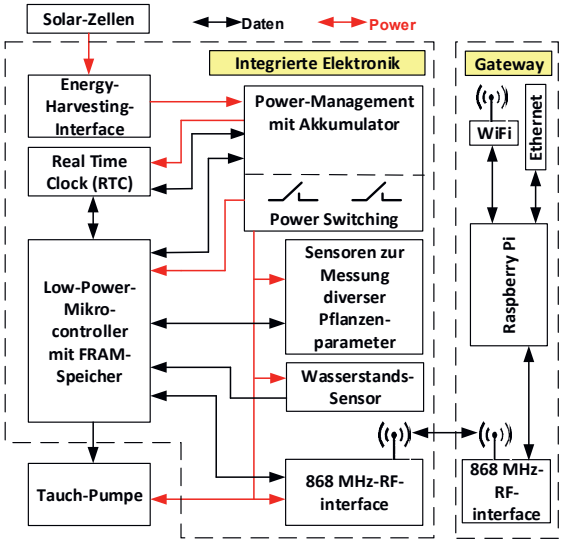


Abb. 3: Detaillierte Architektur des Indoor-Smart-Gardening-Systems

Messdatenerfassung des Blumentopfes

Wie bereits beschrieben, lassen sich mit dem eingesetzten Multisensorchip Temperatur, Luftdruck und Luftfeuchtigkeit erfassen. Abb. 4 zeigt als Beispiel die Messung der Temperatur in einem Innenraum in der Nähe eines Fensters am 28. August 2018.

Die Messung der Bodenfeuchte basiert auf einem kapazitiven Sensor. Die sich auf Grund der Bodenfeuchte ändernde Kapazität ist in einen Oszillatorschaltkreis eingebunden. Dessen sich ändernde Frequenz wird mit dem Mikrocontroller gemessen. Um genaue Messungen zu erhalten, muss der Sensor in Abhängigkeit der Erde bzw. des Pflanzensubstrates, in welchem die Pflanze eingebettet ist, kalibriert werden. Abb. 5 zeigt die Messung der Bodenfeuchte mit dem Smart-Gardening-System. Zu erkennen sind zwei Bewässerungs-Vorgänge. Die eher langsamen Messwertänderungen resultieren aus der langsamen Ausbreitung des Wassers in der Erde.

Zur Messung des Wasserstandes in dem Wasserreservoir des Blumentopfes kommt ein selbstentwickelter Sensor zum Einsatz, mit dem sich sechs unterschiedliche Wasserhöhen messen lassen (siehe Abb. 6). Der Sensor besteht aus einem resistiven Spannungsteiler. Der Spannungsteiler besteht aus Goldkontakten, die durch den elektrischen Widerstandes des Wassers geschlossen werden. D.h. der komplette Widerstand des Spannungsteilers ändert sich auf Grund der Höhe des Wasserstandes im Wasserreservoir. Eine nachfolgende Schaltung verstärkt den Spannungsabfall an dem resistiven Spannungsteiler. Nach einer Digitalisierung der analogen Spannung erfolgt eine Auswertung der Messwerte mit dem Mikrocontroller. Abb. 7 zeigt die verstärkte Ausgangsspannung des Sensors in Abhängigkeit des Wasserstandes.

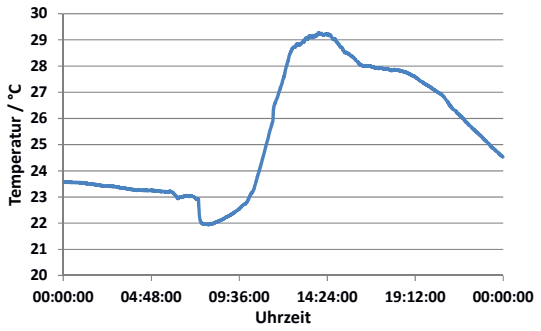


Abb. 4: Gemessene Umgebungstemperatur mit dem Smart-Gardening-System

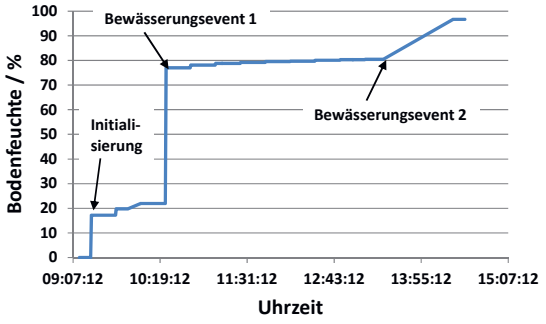


Abb. 5: Gemessene Bodenfeuchte mit dem Smart-Gardening-System

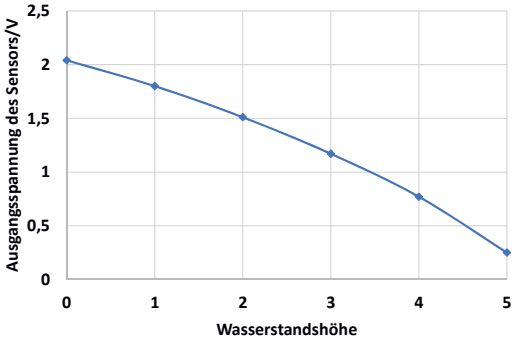


Abb. 7: Gemessene Ausgangsspannung des Wasserstandssensors in Abhängigkeit der Wasserstandshöhen im Wasserreservoir

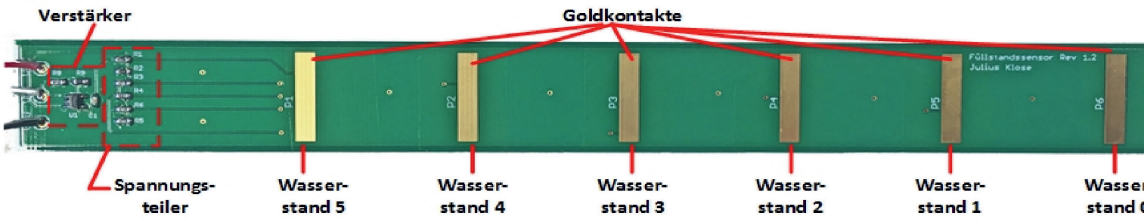


Abb. 6: Wasserstandssensor mit Goldkontakten

Drahtlose Kommunikation zwischen Blumentopfes und Gateway

Die Entwicklung energieautarker, drahtlos kommunizierender Smart-Home-Anwendung ist nur möglich bei der Realisierung einer entsprechenden Low-Power-Funkschnittstelle. Das Smart-Gardening-System nutz für die drahtlose Kommunikation einen 868-MHz-Transceiver von ST Microelectronics. Dieser Transceiver wurde ausführlich evaluiert, um die Funkschnittstelle so zu konfigurieren, dass dieser möglichst wenig Energie benötigt [10]-[11]. Im Vergleich zu anderen Kommunikationstechnologien [12]-[13], hat die drahtlose Schnittstelle des Blumentopfes einen recht niedrigen Energieverbrauch [10]. Die Kommunikation des Gateways und der Blumentöpfe ist als Stern-Topologie implementiert. Die einzelnen Blumentöpfe haben einen zugewiesenen Zeit-Slot für die Kommunikation, damit lässt sich die Kommunikation am energiesparendsten realisieren. Insgesamt lassen sich an einem Gateway bis zu 20 Blumentöpfe betreiben, prinzipiell könnte man das Netzwerk auch noch auf mehr als 20 Blumentöpfe erweitern.

Die erreichbare Kommunikationsreichweite zwischen den Blumentöpfen und dem Gatway ist eine wichtige Größe für das Indoor-Smar-Gardening-System. Die mögliche Kommunikationsreichweite wurde in einem realistischen Einfamilienhaus gemessen (siehe Abb. 8). Für die Messung wurde das Gateway recht zentral im Haus platziert, die Blumentöpfe wurden auf andere Räume verteilt. An jedem Blumentopf wird die Sendeleistung soweit reduziert, bis das Gateway die Daten des Blumentopfes nicht mehr korrekt empfangen konnte.

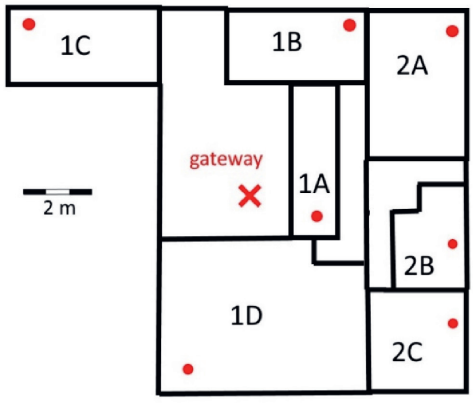


Abb. 8: Szenario in einem Haus, um die möglichen Kommunikationsabstände zwischen Gateway (rotes Kreuz) und den Blumentöpfen (rote Punkte) zu evaluieren

Die Ergebnisse der Messung sind in Tab. 2 dargestellt. Wie erwartet, haben die Anzahl der Wände und sonstigen Hindernisse natürlich einen Einfluss auf die fehlerfreie Übertragung. Die Ergebnisse zeigen jedoch, dass in einem Haus mit vielen Räumen und Türen, die Sendeleistung der drahtlosen Schnittstelle zumeist nicht mehr als 0 dBm betragen muss. Diese geringe Sendeleistung ermöglicht enorme Energieeinsparungen bei gleichzeitig zuverlässiger Kommunikation.

Tab. 2: Gemessene minimale notwendige Sendeleistung für das Messszenario aus Abb. 8

Blumen-topf Nr.	Direkte Reich-weite zwischen Blumen-topf und Gateway	Anzahl Wände zwischen Blumen-topf und Gateway	Minimale Sende-leistung für eine noch fehler-freie Kommu-nikation
1A	2,0 m	1	-30 dBm
1B	5,7 m	2	-22 dBm
1C	8,0 m	2	-8 dBm
1D	5,3 m	1	-20 dBm
2A	7,6 m	3	-16 dBm
2B	6,0 m	4	-3 dBm
2C	7,0 m	5	5 dBm

Energetische Betrachtung des Gesamtsystems

Das energetische Verhalten des intelligenten Blumentopfes wurde für ein bestimmtes Anwendungsszenario untersucht. Die Rand-bedingungen des Anwendungsszenarios sind wie folgt: Alle 10 Minuten führt das System mit den verschiedenen Sensoren Messungen aus und sendet die Messdaten an das Gateway. Die Sendeleitung des Transceivers beträgt 11 dBm, 62 Byte Payload-Daten werden an das Gateway gesendet, die Tauchpumpe läuft maximal 30 Sekunden pro Tag. Tab. 3 stellt die benötigte Energie der integrierten Elektronik inklusive Tauchpumpe pro Tag für das beschriebene Anwendungs-szenario dar. Die Ergebnisse basieren auf ausführlichen Messungen aus [11]. Es ist zu erkennen, dass durch das Ein-/Ausschalten der Komponenten versucht wird, möglichst große Standby-Zeiten der einzelnen Komponenten zu erreichen. Das System benötigt pro Tag in der Summe 8,28 J. Der verwendete Akkumulator hat eine Kapazität von 350 mAh bei einer Spannung von 3,7 V. Unter der Annahme, dass 50 % seiner Kapazität genutzt werden können, d.h. 175 mAh oder 2331 J, beträgt die energetische Lebensdauer des Systems ca. 280 Tage.

Tab. 3: Benötigte Energie des intelligenten Blumentopfes

System-Komponente	Aktive Arbeitszeit pro Tag/s	Benötigte Energie für die Aktive Arbeitszeit/J	Standby-Zeit pro Tag/s	Benötigte Energie für die Standby-Zeit/J
Block Power-Management	86400	6,5707200	0	0
Tauchpumpe	30	0,015069000	86370	0
Real Time Clock	5,76	0,000384912	86394,24	0,125444436
Mikrokontroller und Umgebungssensor	14,4	0,039024000	--	--
Mikrokontroller und Bodenfeuchtesensor	14,4	0,085536000	--	--
Mikrokontroller und Wasserstandssensor	14,4	0,000950400	--	--
Mikrokontroller und Funkschnittstelle im Sendemodus	72	0,599040000	--	--
Mikrokontroller und Funkschnittstelle im Empfangsmodus	72	0,599040000	--	---
Mikrocontroller inaktiv	--	--	8684,8	0,000003960
Funkschnittstelle inaktiv	--	--	86256	0,241948080
Summe benötigter Energie pro Tag/J		*7,909764312*		*0,367396476*

Um zu überprüfen, ob die implementierten Solarzellen ausreichend Energie pro Tag liefern können, wurde die Ausgangsleistung der Solarzellen gemessen. Für die Messung wurde der Blumentopf an ein Fenster mit eher schlechten Lichtverhältnissen platziert. Die Messergebnisse sind in Abb. 9 dargestellt. Während der Messung von drei Stunden konnte eine Energie von insgesamt 547,2 J geerntet werden. Somit sind die Solarzellen ausreichend, um den Blumentopf energieautark zu betreiben.

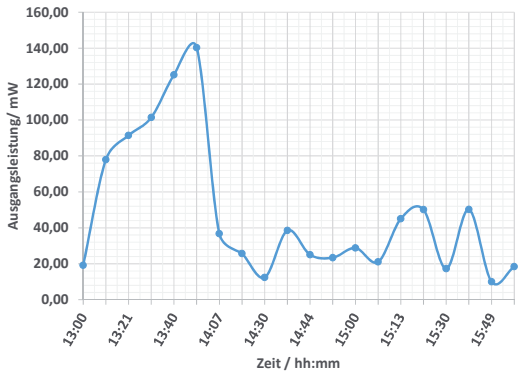


Abb. 9: Gemessene Ausgangsleistung der Solarzellen während drei Stunden

Zusammenfassung

Im Rahmen dieses Beitrags wurde die Entwicklung eines Indoor-Smart-Gardening-Systems mit dem Fokus auf energieautarkes Arbeiten vorgestellt. Herzstück des Systems ist ein 3D-gedruckter Blumentopf für einzelne Pflanzen mit integrierter Elektronik zum Monitoring der wichtigsten Pflanzenparameter und einem integrierten Wasserreservoir mit Tauchpumpe für das automatisierte Bewässern der Pflanze. Energy Harvesting per Solarzellen ermöglicht ein energieautarkes Arbeiten des Blumentopfes. Die wichtigsten Eigenschaften des Smart-Gardening-System sind in Tab. 4 aufgeführt. Gegenüber existierenden und in der Literatur beschriebenen Ansätzen für Smart-Gardening-Anwendungen, hat das System mehrere Vorteile:

- Das System lässt sich sowohl für Indoor- als auch Outdoor-Anwendungen einsetzen.
- Die integrierte Elektronik könnte auf Grund des energieeffizienten Hardwaredesigns für eine große Anzahl anderer Anwendungen im Smart-Home-Bereich eingesetzt werden.
- Der größte Vorteil des präsentierten Systems ist, dass sogar unter Worst-Case-bedingungen (+ 11 dBm Sendeleistung und wenig vorteilhaften Lichtverhältnissen), eine energetische Lebensdauer von mehreren Jahren erreicht werden kann, ohne Batterien zu ersetzen.

Tab. 4: Eigenschaften des in diesem Beitrag präsentierten Indoor-Smart-Gardening-Systems

Eigenschaften	Beschreibung
Gemessene Pflanzen-parameter	Bodenfeuchte, Umgebungstemperatur, Luftfeuchtigkeit, Luftdruck
Ansteuerung von Aktuatoren	Automatisierte Bewässerung
Kommentar	Die gesamte Elektronik ist in einen Blumentopf integriert
Energie-versorgung und energetische Lebensdauer	Integrierte Elektronik des Blumentopfes: Energy Harvesting mit Solar-zellen. Energetische Lebensdauer unendlich; Gateway (ein Raspberry Pi3): 5-V-DC-Netzteil
Drahtlose Kommuni-kationstechnik von den Sensoren/ Aktoren zu einem Gateway und Netzwerk Funktionalität	868-MHz-Funkschnitt-stelle: Line-of-Sight-Kommunikationsreich-weite: 80 m @ 5 dBm; Bis zu 20 Blumentöpfe können an das Gateway konnektiert werden. Das Gateway kann die Daten über eine Internet-Cloud, Smartphones oder ähnlichem zur Verfügung stellen.

Zukünftig soll das System dahingehend erweitert werden, dass damit das Umgebungslicht und der Nährstoffgehalt des Bodens gemessen werden kann. Zudem soll eine Datenbank programmiert werden, in der für unterschiedliche Pflanzentypen genaue Randbedingungen für die Pflege der Pflanze enthalten sind. Aufgrund dieser Datenbank soll die Elektronik des Systems für jede individuelle Pflanze die optimalen Parameter für die Pflege der Pflanze einstellen (z.B. Menge der Bewässerung), oder den Anwender darüber informieren, falls ein Optimum nicht erreicht ist (z.B. zu viel Schatten).

Literaturnachweis

[1] J.R. Rosslin, K. Tai-hoon, “Applications, systems, and methods in smart home technology: a review” International Journal of Advanced Science and Technology, Vol. 15, 2010.

[2] P.-P. Ray: Internet of things for smart agriculture: Technologies, practices and future direction. In Journal of Ambient Intelligence and Smart Environments, vol. 9, no. 4, pp. 395-420, 2017; DOI: 10.3233/AIS-170440

[3] A. H. Abbas, M. M. Mohammed, G. M. Ahmed, E. A. Ahmed and R. A. A. A. Abul Seoud, "Smart watering system for gardens using wireless sensor networks," 2014 International Conference on Engineering and Technology (ICET), Cairo, 2014, pp. 1-5.; DOI: 10.1109/ICEngTechnol.2014.7016780

[4] F. Caetano, R. Pitarma and P. Reis, "Intelligent management of urban garden irrigation," 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), Barcelona, 2014, pp. 1-6.; Doi: 10.1109/CISTI.2014.6876907

[5] G. Carrión, M. Huerta and B. Barzallo, "Monitoring and irrigation of an urban garden using IoT," 2018 IEEE Colombian Conference on Communications and Computing (COLCOM), Medellin, 2018, pp. 1-6.; DOI: 10.1109/ColComCon.2018.8466722

[6] V. Biswal, H. M. Singh, W. Jeberson, A. S. Dhar, "Greeves: A Smart Houseplant Watering and Monitoring System", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 4, Issue 7, July 2015 ISSN: 2278 – 7798, 2015

[7] S. N. Ishak, N. N. N. A. Malik, N. M. A. Latiff, N. E. Ghazali and M. A. Baharudin, "Smart home garden irrigation system using Raspberry Pi," 2017 IEEE 13th Malaysia International Conference on Communications (MICC), Johor Bahru, 2017, pp. 101-106; DOI: 10.1109/MICC.2017.8311741.

[8] Daisy.si: http://daisy.si/dv1/Daisy_dv1_EN_rv0.2.pdf (User Manual; Last Access: 05.03.19).

[9] Parrot: https://www.parrot.com/global/connected-garden/parrot-pot#app (Last Access: 18.10.18).

[10] P. Moser: „Entwicklung einer drahtlosen Sensor-Netzwerk-Plattform für den Low-Power-Anwendungsbereich“, Bachelor Thesis.Offenburg, 2017.

[11] S. Möhringer, P. Moser, E. Mackensen, “Indoor Smart Gardening based on an Energy Autonomous Wireless Network Platform” In: Wireless Congress 2018: Systems & Application, Conference Proceedings, ICM - International Congress Center München, Munich, Germany, November 2018.

[12] E. Mackensen, M. Lai and T. M. Wendt, "Bluetooth Low Energy (BLE) based wireless sensors," SENSORS, 2012 IEEE, Taipei, 2012, pp. 1-4. DOI: 10.1109/ICSENS.2012.6411303

[13] T. Wendt, F. Volk and E. Mackensen, "A benchmark survey of long range (LoRaTM) spread-spectrum-communication at 2.45 GHz for safety applications," 2015 IEEE 16th Annual Wireless and Microwave Technology Conference (WAMICON), Cocoa Beach, FL, 2015, pp. 1-4. DOI: 10.1109/WAMICON.2015.7120312

DOI 10.5162/sensoren2019/P2.16

Construction of an universal blood collecting system with integrated analytics

Susanne Schlegel[1], Eric Krumnow[1], Steffen Zinn[1], Markus Riester[2], Marilena Minieri[3] and Andreas H. Foitzik[1]

[1] Technical University of Applied Science Wildau (THWi) – Hochschulring 1 15745 Wildau, Germany
[2] midge medical GmbH, Bessemerstrasse 22, 12103 Berlin, Germany
[3] Università degli Studi di Roma "Tor Vergata" (URTV) – Via Orazio Raimondo 18 – 00173 Roma, Italy

Abstract

This paper presents the construction of a new medical application system (MAS), unifying blood collecting, sample preparation and analytics in one device. Additionally, the implemented analytical device should be a commercially available lateral flow assay (LFA). Therefore, a universal case has to be designed. A needle system should extract the blood from the patient, while capillary forces transfer the sample through internal feeding into a reservoir. Afterwards, a buffer flushes the reservoir and carries the analyte through the migration distance to the test row. The developed setup utilizes microfluidic mixing structures as for further sample preparation. Thereby, the risk of contamination for the patient and medical personnel could be significantly reduced. Consequently, trained medical personal should not be necessary to operate this setup and increase the overall filed of application. This is especially fundamental when considering an application in case of epidemic events whereas a reliable medical treatment is often not reliable.

Keywords: blood collection, sampling, analysis, LFA, medical device

Introduction

Due to the proceeding of globalization, from time to time the epidemic rise of diseases occurs all over the world. Tropical diseases, for example, are not anymore just a problem of the country of origin [1]. Since more than half of the world population still do not have access to essential medical care [2], the risk of a pandemic outbreak increases even further. In order to prevent such a disastrous event, it is required to identify spreading diseases in a fast and reliable manner. Usually, the sample preparation for biological and chemical probes involves following a strict workflow to eliminate any contamination beforehand [3,4]. Furthermore, it must be carried out by trained personnel such as a nurse or other health care personnel, which is expensive as well as time consuming. Therefore, a new medical application system (MAS) is necessary to provide medical care and clarification in terms of a pandemic outbreak.

Actual situation

For the analysis of the medical conditions saliva and blood samples can be used [3], which should be regarded as samples of high complexity. Well-equipped laboratories could for example prepare hemograms to provide an overview about the general medical status of the patient. [4] However, more than half of the world population has no access to essential medical care [2], consequently those complex medical analytics are not available in many regions, especially in developing countries. Hence, simple analytical methods are needed.

A simple method to perform blood analysis is the lateral flow assay (LFA) test. The LFA test has the advantage of low costs as well as a shorter execution period compared to other methods. Furthermore, the technique doesn't require other additional sophisticated equipment [5]. However, the blood sampling and its analysis should only be performed by medical trained personal, since the blood sampling must follow an exact operating procedure. Not carrying out the procedure correctly could result in an infection of the patient or contamination of the sample itself. Furthermore, a non-adherence of the operating procedure could influence the result of the test. False negative or false positive results lead to a faulty rehabilitation of the patient. Therefore, the procedure of the LFA test should generally be performed by medical trained personal.

Although the LFA test is simpler and more cost-efficient compared to other methods, it still requires medical resources, execution time and funding. This may be problematic in case of an upcoming epidemic. As medical care is not comprehensively assured, the necessary medical resources should be transferred to the affected regions. For that reason, medical personal and tests are sent out with high expense. It should also be considered, that the need of medical trained persons, equipment, financial funding and execution time rise with the increasing patient count. [5]

Requirements for the medical analytical system

The lack of the useful medical analytic systems is setting the requirements of prospective MAS. An improved usability provides a better reliability as well as improved overall performance. For this reason, financial effort could be reduced since medical trained personal is not necessary any more. Thereby, with the help of simpler and faster MAS a high number of patients could be tested in a short period of time and identified an epidemic in an early stage. Therefore, the MAS should also contain a form of identification system in order to provide a clear apportionment of the obtained data. Figure 1 displays schematically the concept of such a MAS.

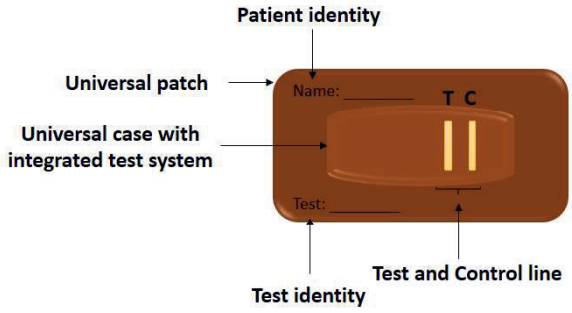


Figure 1: Schematically patch structure of a prospective MAS (top view)

Generally, the MAS should be small in size and cost-efficient. Furthermore, the operation procedure of the MAS should be simple and intuitive. Another important requirement of the MAS is the functionality in case of extreme climatic conditions, so that it can be utilized in a variety of climatic regions. However, the most important requirement is the combination of the blood sampling procedure and the subsequent examination in one system. Therefore, the blood must be transferred to the test strip directly after the extraction. Figure 2 schematic displays the functional unit of such a system with an integrated LFA as well as an extraction via needle.

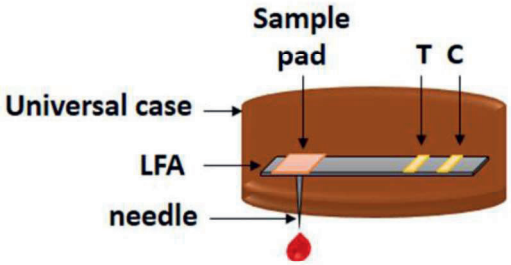


Figure 2: Schematically functional unit of a prospective MAS (top view)

The procedure of blood sampling should be automated and minimal invasive, whereby scars on the skin can be reduced and the risk of successive infections on sampling point will be minimized. The blood sampling procedure must be designed simplistic as well as as single-use in order to prevent transmission of infections between several patients. Therefore, a medical trained person is not required to perform the sample extraction. A temperature, dust and humidity resistant case protects the test from detrimental environmental effects in order to provide the function of the assay under different climatic conditions.

The intended functionality of the MAS assures reliable results because of the fully automated operating procedure, which further minimize potential errors. Such a system could be arranged in size and structure as an adhesive tape. [6]

Realization strategy

The overall development strategy is separated into four phases. Hereby, the first phase of development is focused on the conception and realisation of the housing. In general, the housing of the MAS is determined by the size of the integrated functional components, the material properties and requirements for the internal mounts. Furthermore, the housing should be sealed in order to protect the sensitive interior from detrimental environment (dust, humidity, etc.). Regarding the design of the superficies of the MAS a good handling and space-saving geometry should be considered, which resulted in two general geometries.

The first approach is based on a circularly cylindrical geometry and provides the advantage of an overall good handling but a rather limited installation space. In comparison, the second concept utilizes a rectangular cylindrical geometry. However, considering the

shape of the LFA test stripes as well as possible mounting strategies the first approach has been selected as the first designs of the housing. A first concept is schematically shown in figure.

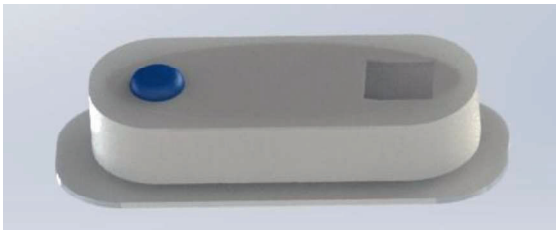


Figure 3: Side view of possible housing design from the prospective MAS

The base plate with the adhesive tab s located at the bottom to position the MAS directly on the skin of the patient. Furthermore, it exceeds the limits of the rest of the whole device in order to increase the contact area. Therefore, the MAS could be positioned in a stable position directly on the skin of the patient. During the blood sampling a fixed position is essential to minimize the lesion of the skin. A shift of the blood sampling system during the penetration of the skin could otherwise cause an increased lesion and therefor enhances the risk of an infection.

Another element of the housing is an integrated window, though which an optical readout of the LFA strip inside the MAS is provided.

Additionally, a mechanical switch is integrated into the top of the MAS to initiate the sample extraction which can be either carried out via a needle or a lancet. The realisation of the blood sampling system is part of the second development phase. It contains the extraction device, a mechanical release mechanism as well as a safety-mechanism.

In regard to the aspired field of application it is intended to extract the blood samples directly via the patient skin [6]. The skin is the largest organ in the human body and consist of three layers, the epidermis, the dermis and the subcutaneous layer. Furthermore, the thickness of the skin depends on the actual skin type and the specific location on the body [7]. Considering the intended application, the extraction of a blood sample should be reproducible and not impaired by varying thickness, which would otherwise affect the subsequent analysis [8]. Additionally, the lesion caused by the penetration of the skin should be as small as possible to decrease the risk of an infection. The challenge is to adjust these two demands of the blood sampling system. Figure 4 displays the intended process chain for the MAS.

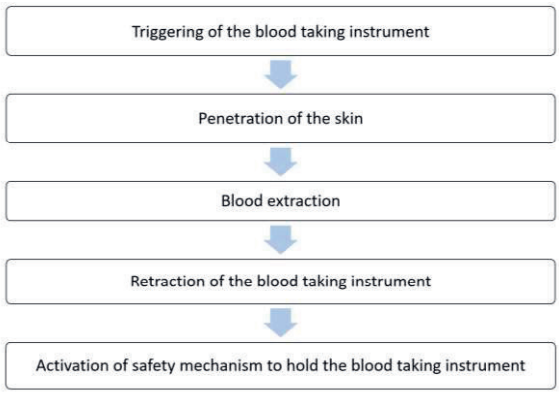


Figure 4: Process flow diagram of the blood sampling procedure

The first step is the activation of the extraction procedure by pressing the mechanical switch on top of the MAS. This motion is directed onto a mechanical transmitter in order to limit the movement of the extraction element. Otherwise, a reliable penetration of the skin cannot be guaranteed which could cause lesions or an insufficient extraction. In the next step, the extraction element is pressed downwards and penetrates the skin. This position is hold for certain period of time until enough blood has been extracted. Afterwards, the extraction element is retraced and secured inside the housing in order to prevent a second skin penetration or further usage of the MAS. Therefore, the expulsion mechanism contains a mechanical fuse which will block the overall movement of the extraction element after the retraction.

For the penetration needles or lancets could be used [8]. Future subsequent experiments will show which kind of blood taking instrument fulfils the functional requirements. For further consideration a needle as extraction element has been be taken.

The third development phase is focused on the transport of the sample to the integrated analytical component. Currently, it is intended to utilize a microfluidic cartridge which contains multiple chambers as well as passive valves and venous valves to retain a backflow of the sample. An exemplary structure is shown in figure 6. In this approach, capillary forces are utilized to transfer the samples though the channels and accumulate the extracted volume in the first so called “waste” chamber. If this first chamber is completely filled, the blood continuous to flow into the actual “sample chamber”. Using this approach, it is intended to provide a defined sample volume for further analysis, otherwise an inadequate amount of blood could inaccurate the outcome. As last

step of the process, buffer will flush the sample chamber and transport the samples to the analytical system.

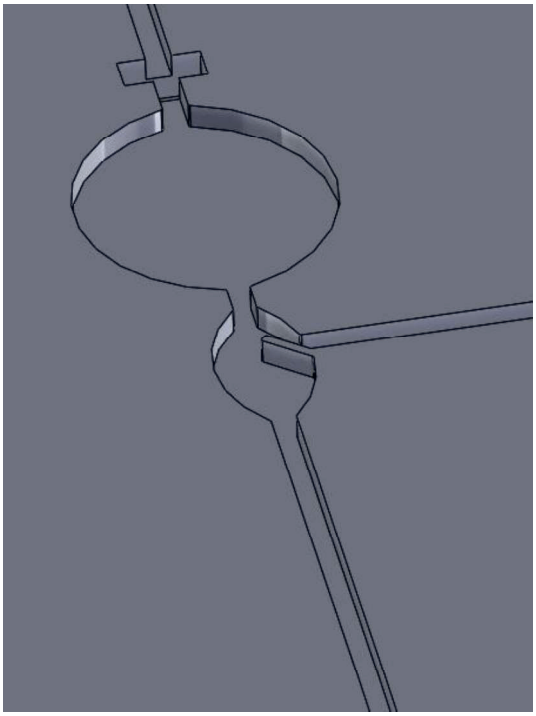


Figure 5: Microfluidic structure form the sample chamber with passive and venous valves

The last element of the development strategy contains the analytical setup. This is currently based on a commercial sandwich LFA test. The general functionality of such a sandwich LFA test is shown in Fig.6. In case of a successful analysis usually two separate lines are visible (the test and the control line). If the blood sample is negative the test only shows the so called control line. LFA tests are usually regarded invalid when no line or only the test line is visible. [5]

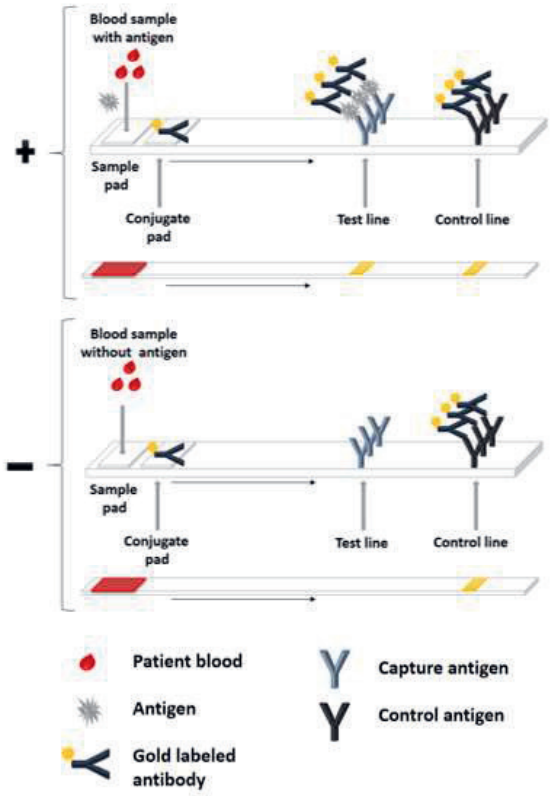


Figure 6: Schematic representation of a sandwich lateral flow assay (LFA). The first LFA shows two lines (the test and the control line) and is positive (+). The second test shows only one line (the control line) and it is negative (-).

The concept for a mount as well as connection with the sample transport will be constructed with computational aided design (CAD). Afterwards, the sample transfer from the extraction point to the analytical assay will be simulated in order to identify technical difficulties and to optimize the overall flow conditions.

The designed MAS is currently realized utilizing 3D-printing technologies. However, after a successfully verification of the fitting accuracy the individual components will be manufactured out via injection molding. Subsequent performance experiments are afterwards also intended.

Summary

A combination of preparation techniques and analysis systems could improve the overall process of blood sampling. This is especially important when considering the situation of sample analysis in case of pandemic prediction

or other catastrophic events. However, the preparation for such sensitive sample requires a strict workflow in order to eliminate the risk of contamination. By utilizing microfluidic sample and preparation systems the risk of contamination for the patient and medical personnel could be reduced significantly. Furthermore, MAS are less time consuming and don't require medical trained personnel.

A theoretical approach utilizing a LFA test as an analytical element has been drafted. The designed microfluidic device is currently manufactured utilizing 3D-printing and will be later on carried out via injection molding. Subsequent experiments are also intended in order to verify the overall performance of the realized designs.

References

[1] Diesfeld, H.J., Krause, G., Teichmann, D.: Praktische Tropen- und Reisemedizin. Thieme, Stuttgart 2003

[2] Information on: http://www.who.int/mediacentre/factsheets/fs395/en/, actualization on Dec. 2017 from WHO, date of access: 12.01.2018

[3] J. Pawliszyn, "Sampling and sample preparation for field and labatory: fundamental and new directions in sample preparation," in hardbound, 2nd ed. vol. 37, J. Peters, Ed. D. Barceló, Amsterdam: Elsevier Science B.V. 2002, pp. 799–825.

[4] L. Thomas, Labor und Diagnose, Th-Books, 6. Auflage, 2005

[5] Development and application of lateral flow test strip technology for detection of infectious agents and chemical contaminants: a review, Babacar Ngom, Yancheng Guo, Xiliang Wang, Dingren Bi, June 2010, Volume 397, Issue 3, pp 1113–1135

[6] Personalization of phlebotomy using smart systems integration, Dr. Markus Riester, meisterwerk ventures GmbH (DE), Jörn Bungartz, midge medical GmbH Berlin (DE), Preprint of Smart Systems Integration Conference, Cork, Ireland, 8 – 9 March 2017

[7] Shier, D., Butler, J. & Lewis, R. in Hole's Human Anatomy and Physiology 8th Edn 171–183 (McGraw Hill, 1999)

[8] WHO guidelines on drawing blood: best practices in phlebotomy. Switzerland 2010. Bookshelf ID: NBK138661

DOI 10.5162/sensoren2019/P2.18

Characterization of the single point precision of steel gear wheel measurements using tactile coordinate measurement machines in scanning mode

Andreas Michael Müller[1], Sebastian Metzner[1], Tino Hausotte[1]
[1]Institute of Manufacturing Metrology (FMT), Naegelsbachstrasse 25, 91052 Erlangen, Germany

Summary

The framework of the single point uncertainty developed at the Institute of Manufacturing Metrology (FMT) presents a methodology to determine and evaluate the local measurement uncertainty for a measurement setup by local comparison of a measurement series with an associated reference geometry. This approach, which was originally developed and optimized for the processing of complete areal measurements of work pieces, was now also extended to line scans found in dimensional testing using tactile coordinate measurement machines (CMM). Target of the investigation are straight-toothed steel gear wheels, which can be dimensionally characterized by both helix and profile scans using a CMM in scanning mode in combination with a rotatory stage. The modification of the single point uncertainty framework in order to determine the single point precision of repeated gear wheel measurements was implemented successfully but the test setup also yielded abnormally high random measurement errors, which could not fully be explained within our examinations.

Keywords: gear wheel, tactile coordinate measurement, single point precision, scanning mode

Introduction

The framework of the single point uncertainty describes a methodology to statistically evaluate the local measurement uncertainty of a measurement series consisting of *n* repeated single measurements with respect to the associated reference geometry in the sense of International Vocabulary of Metrology (VIM) [1]. Usually, a geometric registration routine is required to geometrically align each of the single measurements with the reference geometry. The surface data is represented by triangle meshes using the STL file format. Originating from homogeneously distributed surface points on the reference surface (also called sampling points), the distances to each single measured surface are calculated. Depending on the sampling strategy applied, slightly different distances are computed [2]. Finally, each sampling point is associated with *n* calculated distances (one for each measurement repetition). For each set of distances, the mean value as well as the standard deviation can be computed. In case the reference geometry is known, the local systematic deviations combined with the local random deviations represent the single point (measurement) uncertainty. In case no reference measurement exists, the nominal geometry of the work piece can be used. In that case, the mean distance value represents the combination of the systematic error and the work piece deviations and both effects cannot be separated. Nonetheless, the distribution of random measurement errors (precision) of the measurement setup can be determined.

The exact knowledge of the single point uncertainty of a measurement setup result can greatly influence the quality of the subsequent measurement data evaluations. The processing of single point precision data as weighting factors in geometry element fitting routines can lead to a more accurate determination of geometry element properties [3]. For industrial computed tomography (CT), the visual evaluation of the locally varying uncertainties gives direct insight into the underlying X-ray penetration lengths. This is because they directly affect the single point noise due to the impaired signal to noise properties of the recorded projections. The method was successfully used to correct systematic measurement errors in CT measurements by determining the systematic single point errors of a simulated measurement series and subsequently correcting these systematic measurement deviations [4]. The core routine used in this former work at the Institute of Manufacturing Metrology, which is based on ray-triangle intersection tests, provided accurate results under most conditions [5]. This algorithm computes the distance from each sampling

point of the reference / nominal geometry in the direction of the vertex normal vectors of these sampling points to the triangulated measurement geometry. This sampling strategy is called “normal vector” (Fig. 1). If a measured coordinate is trustworthy, the same applies for the normal vector of the nominal geometry and thus defines the direction in which the corresponding point on a measurement surface is most likely expected. Mathematically, this problem is described by a ray-triangle intersection test, which is a well-known problem in the field of computer graphics [6–10]. Subsequent development efforts resulted in the introduction of an alternative sampling strategy, which solves the problematic sampling of edge regions associated with the sampling in the direction of the surface normal vector [2]. This sampling strategy calculates the shortest distance from a sampling point to the target surface and is therefore called “shortest distance” (Fig. 1). Additionally, this sampling exhibits superior run time properties compared to the existing ray-tracing solution. The determination of the single point uncertainty makes it then possible to pass information about the measurement uncertainty to complex extended tolerance analysis methods and therefore consider the uncertainty inherently associated with any measurement [11].

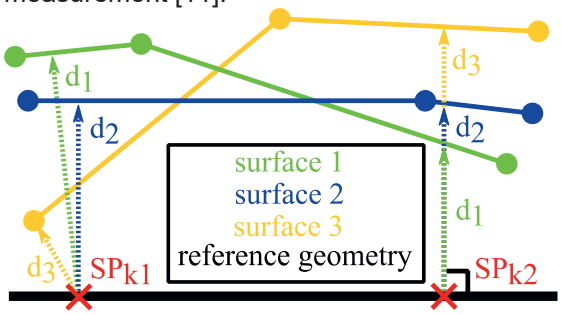


Fig. 1: Visualization of the sampling strategies “shortest distance” (left) and “normal vector” (right).

So far, we only determined and evaluated the single point uncertainties for measurement systems producing an areal measurement result represented as triangle meshes. These are mainly CT and structured-light scanning [12]. This contribution presents various suitable adjustments to the single point uncertainty framework in order to also be able to process line scans from coordinate measurement machines (CMM) using the example of gear wheel measurements.

Measurement data

For the following demonstration purposes, a wire eroded, straight-toothed steel gear wheel characterized by 17 teeth, a width of 8 mm and tip circle diameter of 19.4 mm is used. The measurement setup consists of the tactile CMM UPMC 1200 CARAT S-ACC with an included rotatory stage. The measurement data evaluation was done using the software tools Zeiss Calypso 5.6 and Zeiss Gear Pro 5.9.0.2. Contrary to areal measurements, which are typically represented by or easily converted into a triangle mesh representation, CMM line scans consist of point clouds with additional meta-information (e.g. probing vectors). With respect to gear wheel inspection, the guideline VDI/VDE 2612 states, *“Unless agreed otherwise, the profile is measured in a transverse plane approximately in the middle of the face width.”* [13]. The profile measurement is complemented by the helix measurement: *“The helix preferably is measured on the diameter of the V-cylinder.”* [13]. Summarizing, each gear flank is described by one profile scan and one helix scan. The exact scan trajectory in interaction with the rotatory stage within the presented examinations were determined by the software Zeiss Gear Pro. All controllable filter operations on the measurement data were switched off, the ball tip diameter was 0.8 mm.

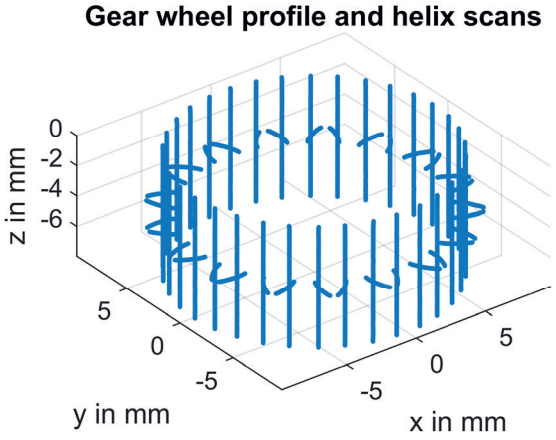


Fig. 2: Visualization of a complete tactile gear wheel measurement consisting of profile and helix scans.

Figure 2 shows the results of a complete gear wheel measurement using the described tactile measurement setup. In order to apply the single point uncertainty framework, the complete gear wheel measurement was repeated 20 times (n = 20), following the recommendation in the norm ISO 15530-3:2011 [14]. Thus, this setting theoretically results in a relative uncertainty of the investigations, caused by the limited sampling, of 16 %, assuming normally distributed uncertainties [15].
At the beginning of the measurement, the coordinate system of the gear wheel with respect to the coordinate system of the CMM was defined. First, the rotation axis of the gear wheel was de-

termined by two circle scans. The remaining rotatory symmetry was dissolved by centering of the probe stylus between two gear teeth. The measurements were performed in a temperature-controlled environment with temperatures set to 20 °C ±0.2 K (temperature classification A according to VDI/VDE 2627-1 [16]) with a relative humidity of 45 % ±10 % (humidity classification A according to VDI/VDE 2627-1 [16]).

Calculating the single point precision

As mentioned above, the single point uncertainty framework was developed to evaluate the uncertainty parameters from repeated areal measurements of measurement objects. Contrary to that, the tactile gear evaluation is characterized by line scans, thus the sampling strategy using the normal vector of the sampling start surface was not feasible. The reason for that is that in general a ray-tracing algorithm can only test the intersections of a ray with areal targets which is not the case for line scans. Consequently, the sampling strategy “shortest distance” had to be used. The measurement results of a tactile line scan are represented by the recording of m measurement points $\{X_i\}_{i=1}^{m} \subseteq \mathbb{R}^3$ as well as the associated probing vectors $\{V_i\}_{i=1}^{m} \subseteq \mathbb{R}^3$ and the distances $\{d_i\}_{i=1}^{m} \subseteq \mathbb{R}^1$ from the nominal geometry to $\{X_i\}_{i=1}^{m}$ in the direction of the probing vectors $\{V_i\}_{i=1}^{m}$. Although the nominal geometry is not given explicitly in the coordinate system of the measurement system, the nominal geometry coordinates $\{N_i\}_{i=1}^{m} \subseteq \mathbb{R}^3$ belonging to each measurement coordinate $\{X_i\}_{i=1}^{m}$ can be reconstructed as follows (1):

$$N_i = X_i - V_i \cdot d_i \tag{1}$$

In the following, $\left\{\{N_i^k\}_{i=1}^{m}\right\}_{k=1}^{n} \subseteq \mathbb{R}^3$ denotes the nominal geometry of the same line scan of all n measurements of the measurement series. Assuming a perfect measurement, the superposition of all calculated sets of nominal geometry coordinates $\left\{\{N_i^k\}_{i=1}^{m}\right\}_{k=1}^{n}$ should result in the same nominal geometry by definition. However, Fig. 3 shows that this is not the case for real measurements due to various kinds of error sources influencing the measurements. To determine the single point uncertainties, the common nominal geometry must be known, because it represents the definition of the sampling points. As a result of the pre-knowledge, that the gear wheel is straight-toothed, we know that each nominal helix geometry must be a straight line. Consequently, one possibility to reconstruct the nominal helix scan line from the repeatedly measured scans is to calculate the solution for a line regression model (Fig. 3, green line). This was done based on the singular value decomposition (SVD) of all coordinates $\left\{\{N_i^k\}_{i=1}^{m}\right\}_{k=1}^{n}$ [17], which results in a least squares solution for the scan direction vector. The position of that scan line is determined by the mean value of $\left\{\{N_i^k\}_{i=1}^{m}\right\}_{k=1}^{n}$.

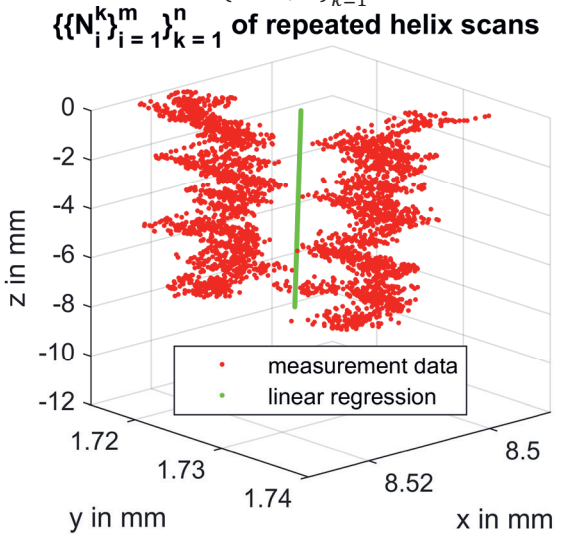


Fig. 3: Visualization of the calculated nominal geometry from repeated helix scans and the result of the linear regression. Caution: Axes are not equally scaled!

The z-components e_z of all probing vectors $\{V_i\}_{i=1}^{m}$ recorded during the measurements were always zero, thus all probing vectors are perpendicular to the scan direction e_z.

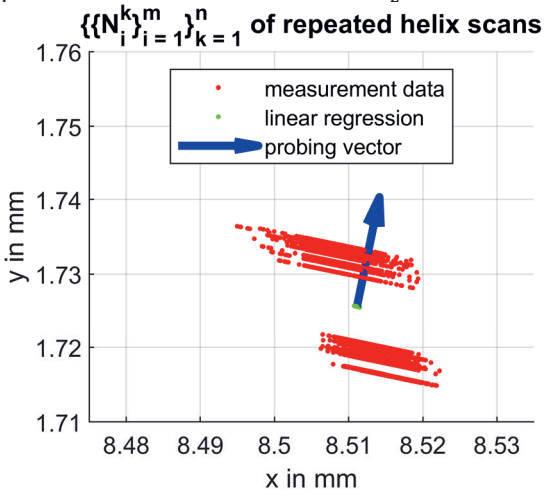


Fig. 4: Projection of the calculated nominal geometry into the xy-plane. Measurement deviations perpendicular to the probing vector (z component is zero) are not recorded.

That means that deviations from the nominal geometry perpendicular to both the probing vector and the scan direction can in principle not be

recorded for straight-line scans using only one vector $\{V_i\}_{i=1}^m$, which is also observable in Fig. 4. Consequently, the regression analysis was necessary here to reconstruct the nominal helix, which is required to sample the single point scatter. Figure 4, which contains the same measurement data as Fig. 3, also shows that the repeated scans are affected by some kind of offset relative to each other, which indicates the influence of a dominant error source within the measurement chain. After determining the nominal helix scan, sampling points were spaced out equally on that line. After that, the shortest (signed) Euclidean distances from each sampling point to each repeatedly measured helix scan defined by $\{X_i^k\}_{i=1}^m$ were calculated, resulting in the distance sets $\{t_k\}_{k=1}^n \subseteq \mathbb{R}^1$ now associated with each sampling point. The calculation also allowed for intersections between sampling points using linear interpolation and is therefore not equal to a simple nearest neighbor search. The "shortest distance" sampling also additionally requires the vertex normal vectors for each sampling point in order to decide if deviations are counted as positive or negative values. These vertex normal vectors for each sampling point, are defined by the mean vector of $\{V_i\}_{i=1}^m$. Next, the single point scatter for each sampling point was defined by the standard deviation of $\{t_k\}_{k=1}^n$. The described approach was then repeated for all helix scans. A very similar data processing pipeline was implemented to determine the nominal profile scan geometry. Here, the SVD was used to identify the two main axes of the point cloud $\left\{\{N_i^k\}_{i=1}^m\right\}_{k=1}^n$, which was then followed by a regression analysis using a high degree polynomial ($\mathbb{R}^2$) model.

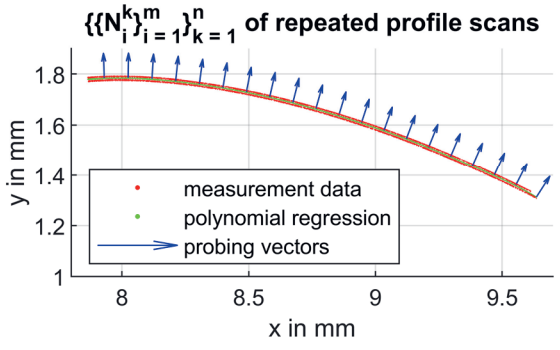


Fig. 5: Visualization of the calculated nominal geometry from repeated profile scans and the result of the high order polynomial regression (between the red lines).

Figure 5 shows the determination of the nominal profile scan line (embedded between the red lines). Note that the recorded z-components e_z of the probing vectors $\{V_i\}_{i=1}^m$ during the measurements are again always zero. After that, equally spaced sampling points are defined on that nominal geometry and the associated vertex normal vectors are calculated by derivation of the fitted function (z-components e_z set to zero). Finally, the single point scatter could then be calculated the same way as described above, which was also repeated for each of the profile scans afterwards.

Results

Figure 6 shows the calculated single point precision for all gear wheel profile and helix scans. The observed precision values represent the superimposition of all uncorrected measurement error contributions along the complete measurement chain. This includes the definition of the coordinate systems and additional components like to rotatory stage. The profile scans exhibit systematically lower precision values in the range between 2 µm and 10 µm compared to the helix scans (between 10 µm and 25 µm). Additionally, the precision noticeably fluctuates along the helix scans. For the examined measurement setup, the precision of the rotatory stage can be regarded as especially important Because of that, the rotatory stage was undertaken additional examinations, which are described in the subsequent section. Finally, the findings are discussed in the section after that.

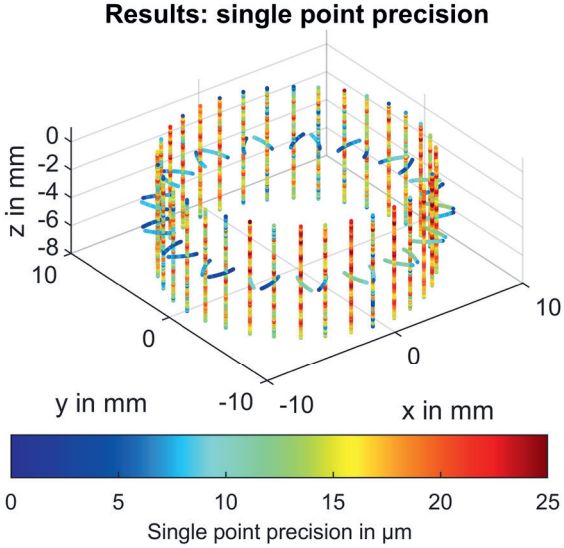


Fig. 6: Single point precision for helix and profile scans of repeated gear wheel measurements.

Testing the precision of the rotatory stage

The target of the following examination was to determine, if the error characteristics of the rotatory stage could account for the observed single point precision values. For that purpose, a

flick normal (precision cylinder with a flattening of 15 µm) was repeatedly measured 20 times. A single measurement consisted of two circle scans of the cylinder shell surface, which were used to determine the cylinder axis. Another circle scan characterizes the front surface of the cylinder and the fourth circle scan, measures the flick (flattened area). Figure 7 shows the distance of the measurement coordinates from the common Gaussian circle in Polar coordinates. The location of the flick is clearly visible at around 10° (see also Fig. 8). The flick normal was rotated by the rotatory stage during the measurement with a fixed position of the probe stylus. This approach makes the verification of the precision of the rotatory stage possible, because the measured coordinates were then effectively represented by polar coordinates (angle in xy-plane recorded by the rotatory stage and radius by the CMM). The coordinate system was arranged such that the z-axis was represented by the determined flick normal cylinder axis. Then, the measured coordinates, which are part of the flick, were extracted by an iterative distance and angle based search algorithm.

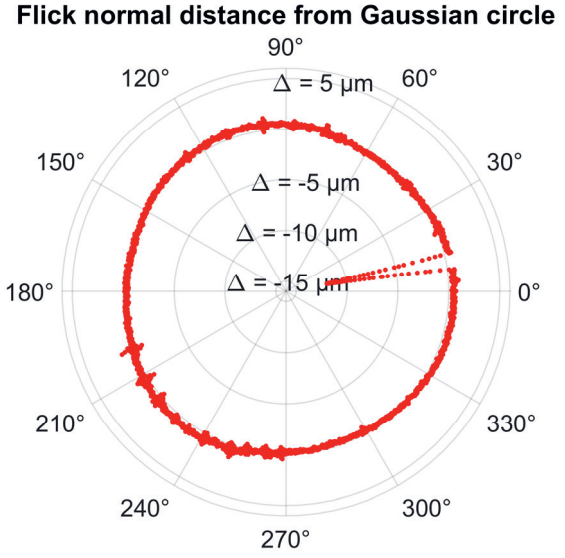


Fig. 7: Circle scan containing the flick (height 15 µm): Distances from Gaussian circle. See also coordinates in Fig. 8

Afterwards, the position of the flick was determined by robustly fitting a polynomial function (degree 1), which can be written as (2), into the selected part of the circle scan (Fig. 8).

$$f(x) = ax + b \tag{2}$$

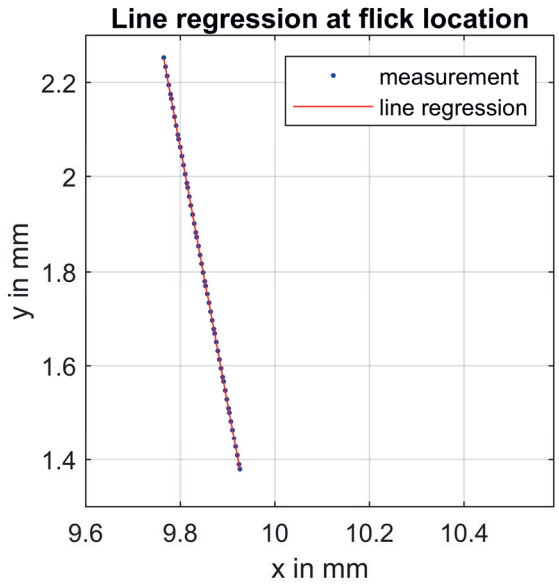


Fig. 8: Robust line regression at the flick location, same coordinate system as Fig. 7.

The angular representation of the slope (3) of that line then represents the angular position of the flick.

$$\theta = \arctan(a) \tag{3}$$

Consequently, the scatter of that angle observed over multiple repeated measurements describes the angular positioning precision of the rotatory stage.
Assuming an underlying normal distribution and targeting a level of confidence of approx. 99 %, the value of three standard deviations [15] of the angular position was determined to be 83.7 arcsec. Additionally, the probing repeatability of the measurement chain was tested by determining the scatter of the perpendicular distance of the fitted polynomial function to the cylinder axis of the flick normal. Three standard deviations were calculated as 0.41 µm. During the calibration of the rotatory stage, a wobbling angle of 24 arcsec was determined. Additionally, the angle between the rotatory axis and the CMM coordinate system was characterized by 6 arcsec. These two error sources are of systematic nature and were therefore corrected by the measurement software.

Discussion

As already mentioned during the discussion of Fig. 4, the measurement chain was affected by various sources of uncertainty which lead to the observation of scatter values up to 20 µm – 25 µm when determining the single point scatter. This could not be explained by the probing scatter of the CMM itself, which is verified by regular calibration of the measurement device,

as well as by the observed positional repeatability (0.41 µm with 99 % confidence) of the perpendicular distance between the flick feature to the cylinder axis. A substantial part of the observed single point scatter could be explained by the lack of precision of the rotatory stage. The remaining part of the observed single point scatter is caused by various other influences. It is possible, that the centering operation during the definition of the coordinate system of the gear wheel was unstable enough to be responsible for some the observed effects. Furthermore, the surface roughness is regarded as an important influence factor on the achieved measurement uncertainties, which is reflected by the fact that the surface roughness value R_z directly contributes to the measurement uncertainty using the "Virtual CMM" used for the uncertainty calculation of the used system [18]. Additionally, dynamic effects during the scan can reduce the achieved precision, which can also be observed in Fig. 6 for the helix scans, where the precision fluctuates along the scan trajectory.

In this contribution, an adjustment of the single point uncertainty framework, which had primarily been developed to evaluated areal measurements, was presented. Now, locally resolved uncertainty values can also be calculated for CMM examinations. Repeated measurements of a steel gear wheel were used to generate the measurement data. The adjustment was characterized by the additional reconstruction of the underlying nominal geometry from the measurement data. The measurement series exhibited unexpectedly large scatter, which is in general untypical for tactile CMM measurements. A large part of the observed scatter could be assigned to the uncertainty of the used rotatory stage, although the observation could not be fully explained.

Finally, the single point uncertainty framework was successfully used to determine the scatter of the complete measurement chain of repeated CMM measurements. Future research efforts need to focus on the verification of these results by examination of different measurement objects.

Acknowledgments

The authors would like to thank the German Research Foundation (DFG) for supporting the research project "FOR 2271 process-oriented tolerance management based on virtual computer-aided engineering tools" under grant number HA 5915/9-1. The authors thank the Institute of Manufacturing Technology (LFT, FAU Erlangen-Nuremberg) for providing the steel gear wheel used for demonstration purposes in this article.

References

[1] International vocabulary of metrology: Basic and general concepts and associated terms (VIM) German-English version ISO/IEC-Guide 99:2007, Corrected version 2012, 4th ed. Berlin: Beuth, 2012.

[2] A. M. Müller and T. Hausotte, "Comparison of different measures for the single point uncertainty in industrial X-ray computed tomography," in 9th Conference on Industrial Computed Tomography, 2019.

[3] A. M. Müller and T. Hausotte, "Utilization of single point uncertainties for geometry element regression analysis in dimensional X-ray computed tomography," in 9th Conference on Industrial Computed Tomography, 2019.

[4] A. M. Müller, F. Wohlgemuth, and T. Hausotte, "Simulation-based correction of systematic errors for CT measurements," in 8th Conference on Industrial Computed Tomography, 2018.

[5] M. Fleßner, A. M. Müller, D. Götz, E. Helmecke, and T. Hausotte, "Assessment of the single point uncertainty of dimensional CT measurements," in NDT.net issue Vol.21 No.02, 2016.

[6] A. S. Glassner, An introduction to ray tracing. San Francisco, Calif.: Kaufmann, 2007.

[7] H. H. Hu, Ray tracing from the ground up: A K Peters, 2007.

[8] P. Shirley and R. K. Morley, Realistic ray tracing, 2nd ed. Natick, Mass.: AK Peters, 2003.

[9] T. Möller and B. Trumbore, "Fast, Minimum Storage Ray-Triangle Intersection," Journal of Graphics Tools, vol. 2, no. 1, pp. 21–28, 1997.

[10] C. Schlick and G. Subrenat, "Ray Intersection of Tessellated Surfaces: Quadrangles versus Triangles," in Graphics Gems 1993, pp. 232–241.

[11] A. M. Müller, T. Oberleiter, K. Willner, and T. Hausotte, "Implementation of parameterized work piece deviations and measurement uncertainties into performant meta-models for an improved tolerance specification," Proceedings of the International Conference on Engineering Design, ICED, 2019.

[12] A. M. Müller, S. Metzner, T. Hausotte, D. Schubert, and D. Drummer, "Separation of locally determined work piece deviations and measurement uncertainties for structured-light scanning of customized polymer gear wheels," in 20. GMA/ITG-Fachtagung Sensoren und Messsysteme, 2019.

[13] Measurement and testing of gears, VDI/VDE 2612 Part 1, 2018.

[14] Geometrical product specifications (GPS) -- Coordinate measuring machines (CMM): Technique for determining the uncertainty of measurement, ISO 15530-3:2011, 2018.

[15] Joint Committee for Guides in Metrology (JCGM/WG 1), Guide to the expression of uncertainty in measurement: JCGM 100:2008 - GUM 1995 with minor corrections, 2008.

[16] Measuring rooms - Classification and characteristics - Planning and execution, VDI/VDE 2627 Part 1, 2015.

[17] C. M. Shakarji and V. Srinivasan, "Theory and algorithms for weighted total least-squares fit-

ting of lines, planes, and parallel planes to support tolerancing standards," Journal of Computing and Information Science in Engineering, vol. 13, no. 3, 2013.

[18] F. Wäldele and H. Schwenke, "Automatische Bestimmung der Messunsicherheiten auf KMGs auf dem Weg in die industrielle Praxis (Automated Calculation of Measurement Uncertainties on CMMs – Towards Industrial Application)," tm - Technisches Messen Plattform für Methoden, Systeme und Anwendungen der Messtechnik, vol. 69, no. 12/2002, p. 550, http://www.degruyter.com/downloadpdf/j/teme.2002.69.issue-12_2002/teme.2002.69.12.550/teme.2002.69.12.550.xml, 2002.

DOI 10.5162/sensoren2019/P3.3

Modellbasierte Analyse von Stahlbau-Strukturen, am Beispiel Maßprüfung und Vollständigkeitsprüfung an Schienenfahrzeug-Wagenkästen mittels 3D-Scan und Software-Algorithmus

Dr.-Ing. Uwe Jurdeczka,
ALSTOM Transport Deutschland GmbH,
Linke-Hofmann-Busch-Straße 1, 38239 Salzgitter
uwe.jurdeczka@alstomgroup.com

Zusammenfassung

Die Analyse von Stahlbaustrukturen (wie zum Beispiel Wagenkästen für Schienenfahrzeuge) bezüglich
- Maßprüfung gemäß DIN 25043-2:2012, und
- Vollständigkeitsprüfung

ist eine Herausforderung wegen der Abmessungen des Prüfobjektes (ca. 20 x 3 x 3m) und wegen der Anzahl der Anbauteile (mehrere hundert).

Es gelingt, mit dem Einsatz von 3D Scannern ausreichend umfangreiche Information über die Ist-Konfiguration (Maße und Vollständigkeit) der zu prüfenden Stahlbaustruktur zu gewinnen.
Für die Maßprüfung ergeben sich Vorteile im Prüfablauf. Zusätzlich wird ein detailreiches Abbild gewonnen, welches für spätere Auswertungen genutzt werden kann.
Mittels eines angepassten Algorithmus wird auf das jeweilige 3D Modell als Soll Zustand mit Differenzbildern überlagert. Diese Differenzbilder zeigen eventuelle Maß-Abweichungen und auch eventuelle Fehlteile an, mindestens jedenfalls die suspekten Stellen. Damit lassen sich die Nachteile der bisher praktizierten ausschließlich manuellen Vollständigkeitsprüfung, nämlich Ermüdung und Geschwindigkeit der Prüfer, weitgehend umgehen.

Keywords: Stahlbaustrukturen, Maßprüfung, Vollständigkeitsprüfung, 3D Laserscanner, Auswertealgorithmus, Differenzbilder.

Motivation

Bisher werden pro Wagenkasten ca. einhundert diskrete Maße gewonnen durch Laser-Tracker oder Schnurgerüst, ergänzt um weitere Messungen mit Winkelmesser und Innenmeßschieber. Den formalen Anforderungen der DIN 25043-2:2012 wird damit sicher genügt. Die Analyse im Falle unerwarteter Montageprobleme wird so allerdings nicht unterstützt. Gerade diese Analyse von Merkmalen zu einem späteren Zeitpunkt bei fortgeschrittener Montage eröffnet Potential. Nebenher kann so auch die Statistische Prozeßregelung mit ausreichend Eingangsinformationen gespeist werden.
Die Taktzeit und die konstruktive Variantenvielfalt mit geringer Stückzahl völlig baugleicher Wagenkästen sind die motivierenden Randbedingungen für die Vollständigkeitsprüfung. Mittels MES – Manufacturing Execution System beherrscht man diese Variantenvielfalt in der Fertigung. Die entsprechende fertigungsbegleitende Prüfung erfordert ebenso Aufmerksamkeit. Die Variantenvielfalt besteht in der Art der Anbauteile, ihrer Anzahl und in ihrer Positionierung. Circa dreihundert Anbauteile (Halter, Laschen, Erdungsanschlüsse u.ä.) sind innerhalb der Taktzeit von einer Schicht (ca. 8h) auf Vollständigkeit zu prüfen. Die Entwicklung von gerätegestützten Vollständigkeitsprüfungen schreitet entsprechend voran. Erfahrungsberichte für erste Entwicklungen hin zu einer gerätegestützten und modellbasierten Vollständigkeitsprüfung, für kleinere Prüfvolumen als die hier vorgestellten, liegen vor [1], [2], [3] und [4]. Diese Entwicklung wird ergänzt durch Handlungsanweisungen an Montage- und Prüfpersonal, indem Sollpositionen von Merkmalen per Laserlicht auf der Montageposition oder per Bildschirm im CAD-Datensatz angezeigt werden [5], [6]. Die

bekannten Ansätze für die Vermessung von Konstruktionen mit Hauptabmessungen größer 10m und für die Vollständigkeitsprüfungen sind also weiterzuführen. Modellbasierte Prüfungen versprechen nun die Möglichkeit, auf die Variantenvielfalt der Konstruktion und Fertigung mit einfach anzupassenden Prüfplänen und Merkmalslisten einzugehen.



Abb. 1: Ein Wagenkasten als Beispiel für Prüfobjekt im Prüfstand

Schmitt et.al. [7] geben den Überblick für die derzeit zur Verfügung stehenden Meßverfahren für große Prüfobjekte (Large-Scale). Kersten et.al. [8] sowie auch Lichtenberger und Wintjens [9] vertiefen dies für den Terrestrischen Laserscanner.

In [10] und [11] ist beschrieben das Generieren von Differenzbildern aus dem Vergleich 3D CAD Modell (Soll) zu 3D Daten (Ist). Hier geht man auch auf die der Fertigung zugestandenen Toleranzen und deren Ausnutzung ein. In [12] wird eine Lösung entwickelt, für das grundsätzliche Problem, dass die reale Länge des Wagenkastens mit dem Nennmaß aus dem CAD-Datensatz nur innerhalb der Toleranzen von mehreren Millimetern übereinstimmt. So führt beispielsweise eine Längenabweichung der realen Stahlbaukonstruktion versus Datensatz zu einer Anzeige im Differenzbild. Allerdings muss eine solche Abweichung nicht unbedingt ein Fehler sein, sondern kann eine als noch i.O. bewertete Fertigungsstreuung sein. So könnte eine ganze Seite einer kubischen Stahlbaukonstruktion (z.B. Seitenwand eines Schienenfahrzeuges) als fehlend/abweichend dargestellt werden. Dies, obwohl kein Fehler vorliegt, sondern die Toleranz auf das Nennmaß ausgenutzt wird.

Die vorhandenen Geräteklassen und Auswerteverfahren wurden also modifiziert bezüglich:

- der Erfassung des Ist-Zustandes des Prüfobjektes (Punktwolke, 3D Scan), und
- des Auswertealgorithmus.

Dies gelang durch die Verwendung des Terrestrischen Laser Scanners 100HSX-SR [13] und durch Implementierung eines Algorithmus in die Software PolyWorks Inspector Premium [14].

Bei der Anwendung des 3D Scannens für die Generierung von Maßaussagen bestehen gegenüber der Vollständigkeitsprüfung zusätzliche Schwierigkeiten. Diese sind am Ende des nächsten Kapitels aufgeführt.

Erfassung des Ist-Zustandes, Maße

Es wird also hier ein terrestrischer 3D Laser Scanner (TLS, auch HDS - High Definition Surveying genannt) zur Datenakquise (3D Scan, Punktwolke) genutzt. Schnelligkeit, vergleichsweise einfache Bedienung und handhabbare Punktwolken sind die Vorteile aus der Verwendung eines TLS/HDS zur Gewinnung des Abbildes des Prüfobjektes.

Ein typischer Wagenkasten eines Regionalzuges kann mit vier Sensorpositionen außen und weiteren fünf bis sechs Sensorpositionen im Innenraum soweit vollständig gescannt werden, um die Abschattungen durch Säulen im Inneren zu umgehen.



Abb. 2: Beispiel Punktwolke. Dargestellt ist die Intensität der Reflexion (Remision) als Grauwert. Dadurch eine fotorealistische Ansicht der Punktwolke.

Da aus verschiedenen Sensor-Positionen aufgenommen wird, sind zunächst die Einzelscans zu registrieren und zu einer gemeinsamen Punktwolke zusammen zu fassen. Durch Vernetzen können die Punkte sinnvoll ausgedünnt und einzelne Fehlabtastungen u.a. der Hallendecke eliminiert werden. Die Punktwolke aus den einzelnen 3D

Scans ist dazu aufzubereiten, nämlich durch Registrierung (Zusammenführen) der Einzelscans aus den verschiedenen Sensorpositionen und Vermaschung/Vernetzung. Beispielparameter für die Vermaschung: Netzgröße maximal 4mm und Abweichung Einzelpunkt 0,1mm. Optional kann auch eine Beurteilung der Punkte nach Scanner Abstand und besonders nach

Zielmarken (Zielmarken z.B. in Abb. 2) gibt die Software anschliessend eine Statistik bzgl. der 3D-Genauigkeit aus. Die maximale Abweichung bei 20 Antastungen lag bei ca. 0,5mm, die Standardabweichung bei ca. 0,25mm.
Abbildung 3 zeigt die registrierte Gesamt-Punktwolke eines Wagenkastens. In Abbildung 3 sind auch bereits Maße eingetragen, nämlich

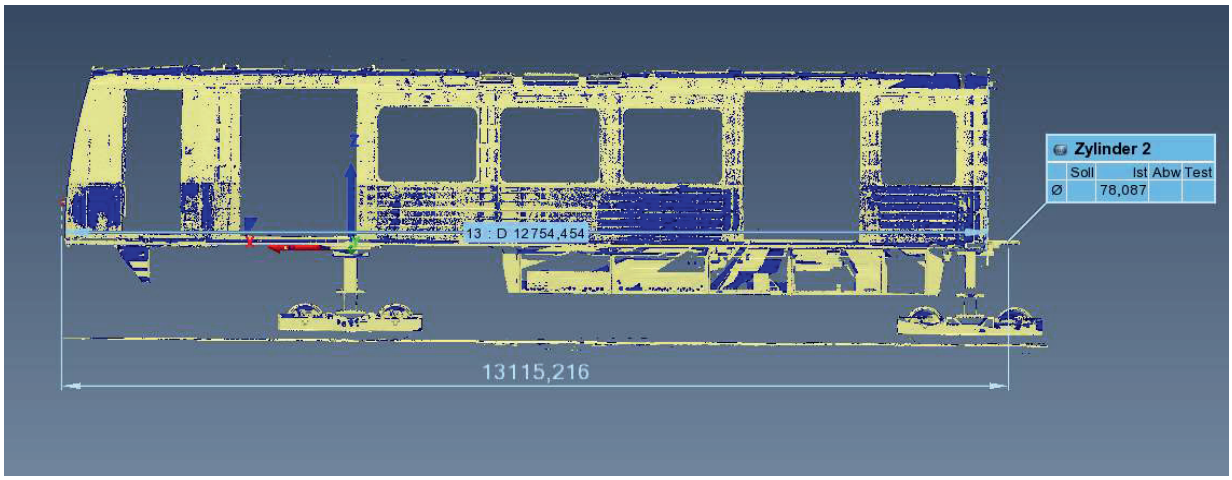


Abb. 3: Gesamt-Punktwolke eines Wagenkastens. Abgeleitet zwei Längenmaße und der Durchmesser einer Lagerbohrung.

Auftreffwinkel erfolgen. Die so mit den Scannern und der nachfolgenden Aufbereitung gewonnene Punktwolke ist eine detailreiche Abbildung des tatsächlich gebauten Zustandes (Ist-Zustand).

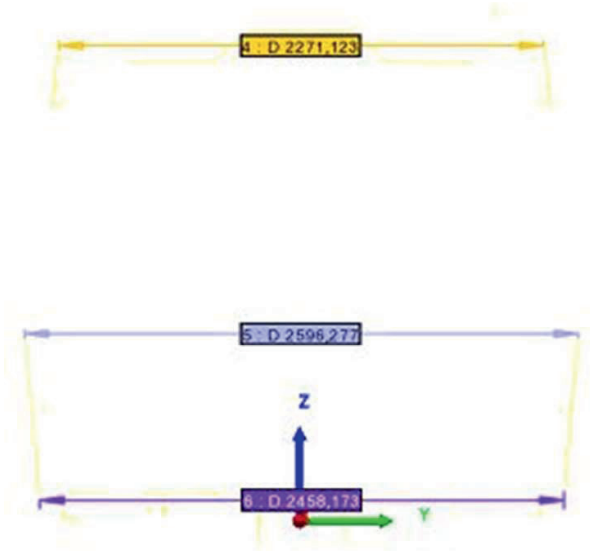


Abb. 4: Beispiel Referenzebenen im Wagenkasten.

Zur Reproduzierbarkeit der Messungen wurde eine Voruntersuchung durchgeführt. Bei der automatischen Registrierung der Scans mittels

Längemaße mit Bezugspunkt
i. Rückwand und
ii. Mittelpunkt der Lagerbohrung.
Letztere ebenfalls aus der Punktwolke bestimmt. Die Auswertung erfolgt entsprechend der Messblätter nach DIN EN 25043-2. Zunächst war es notwendig, Ebenen als Referenzebenen in die registrierten Scandaten zu definieren. Ein Beispiel für Referenzebenen ist in Abbildung 4 zu sehen.
Tabelle 1 führt die Beispiele für Maße in den drei Referenzebenen (Abb. 4) gemäß DIN 25043-2:2012 der Gesamt-Punktwolke (Abb. 3) auf.

Tab. 1: Beispiel für Maße

Name	SOLL	IST	ABWEICHUNG
Y32L+R	2280,00	2271,12	-8,88
Y41L+R	2600,00	2596,28	-3,72
Y30L+R	2458,00	2458,17	0,17

Abbildung 5 zeigt eine zusätzlich eingerechnete Fahrzeugmittenebene (Symetrieebene) und die darauf bezogenen Breitenmaße, nämlich Y-Maße. Zum Vergleich können die Y 41L und 41R Maße aus Tabelle 1 und Abbildung 4 herangezogen werden.

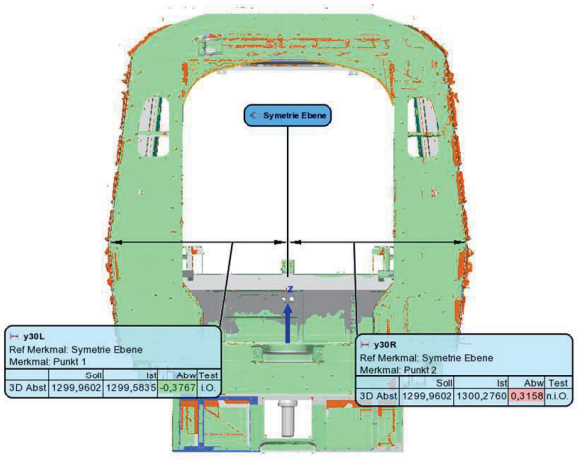


Abb. 5: Zusätzlich eingerechnete Fahrzeugmittenebene (Symetrieebene) und die darauf bezogenen Breitenmaße, nämlich Y-Maße.

Durch Filtereinstellung am Scanner ist die Akquise der Punktwolken so zu beeinflussen, dass Pseudopoints (u.a. sogenannte Gardinen) weitgehend ausgeschlossen werden. Verbleibende Pseudopoints stören den Algorithmus für die Vollständigkeitsprüfung (beschrieben im folgenden Abschnitt) nicht wesentlich. Für die Maßprüfung sind sie allerdings sehr störend. Ihrer Unterdrückung ist also Aufmerksamkeit zu widmen. Beispiele für Pseudopoints in Abbildung 6 (unregelmäßiges Gebirge auf Rückwand, Gardinen in den Fensterausschnitten).

marken an den Stirnwänden ermittelt. Die Standpunkte wurden von jedem Scan zum Folgescan um ca. 1m verändert. Eine Registrierung ging hier also noch nicht ein.

Tab. 2: Vergleich zweier Messkampagnen

Kampagne 1	Länge in mm	Mittelwert in mm	Standard abweich.
Scan01	12100,78	12100,78	0,12
Scan02	12100,76		
Scan03	12100,61		
Scan04	12100,83		
Scan05	12100,99		
Kampagne 2			
Scan01	12100,76	12100,76	0,16
Scan02	12101,14		
Scan03	12100,66		
Scan04	12100,92		
Scan05	12100,92		

Zur Gewinnung der Gesamtpunktwolke aus den einzelnen Scans ist eine Registrierung nötig. Bei der Registrierung mittels automatisch erkannter Zielmarken werden angegeben 0,3mm Standardabweichung und 0,7mm maximale Abweichung eines Einzelpunktes. Dies stellt für die Vernetzung gute Daten bereit.

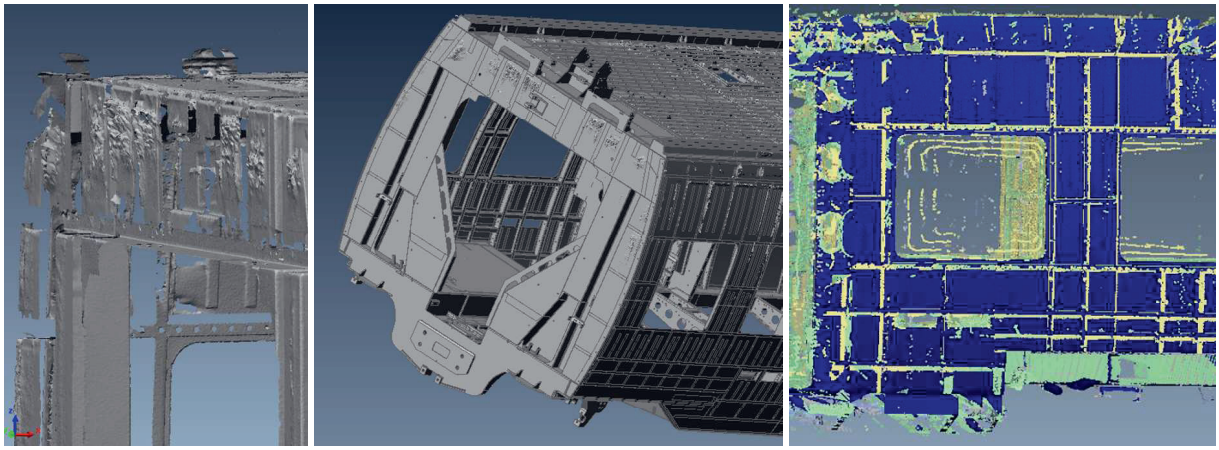


Abb. 6: Beispiel für Pseudopoints an der Rückwand und in den Fensterausschnitten eines Wagenkastens (Software ProcessC3DShell).

Erreichte Genauigkeit

Die Aufnahme der Punktwolke mit dem Phasenscanner erfolgt mit Genauigkeiten im Sub-Millimeter Bereich. Die Tabelle 2 gibt den Vergleich der Wagenkastenlänge in zwei Meßkampagnen in einem Wagenkasten wieder. Die Werte wurden durch Abtastung von Ziel-

Zum Algorithmus der Vollständigkeitsprüfung

Innerhalb des zu untersuchenden Bereichs wird die Punktwolke im Best Fit auf den Datensatz gelegt.
Eine zusätzliche Möglichkeit ist der RANSAC-Algorithmus, welcher oft verwendet wird, um

Best Fits robust gegen Ausreißer zu machen. Hierbei wird die Ausgleichung mit stark reduzierter Punktanzahl durchgeführt und die erhaltenen Parameter werden auf die verbleibenden Punkte angewendet. Dies kann zum einen die benötigte Rechenzeit senken und zum anderen (und wichtiger) vermeiden, dass Ausreißer-Punkte den Best Fit stören und so das Auffinden der Ausreißer nach dem Restresiduum erschweren. Sehr umfangreich wurde der Algorithmus zur Vollständigkeitsprüfung bereits beschrieben in [10], aber auch in [12] und [18], weswegen an dieser Stelle dahin verwiesen wird. Bestandteil des Algorithmus ist die Definition eines Wertes für ε. Dies ist der Abstand zwischen CAD Modell und überlagerter Punktwolke, innerhalb dessen eine Relation zwischen CAD Element und Punktwolke angenommen wird.
In Abbildung 7 werden Elemente mit Punktabständen größer ε farbmarkiert. Die Betrachtung der Seitenwand Nähe Rückwand zeigt rote Punkte. Es sind dies Schutzstöpsel aus Plastik, die zum Zeitpunkt des Scannens bereits angebracht waren und natürlich nicht zum Konstruktionsdatensatz des Wagenkastens gehören. Das Auffinden dieser Ø8x20 mm bzw. Ø10x20mm großen Merkmale zeigt die potentielle Sensibilität des Prüfsystems. Unterschiede zwischen 3D Model und Punktwolke werden durch Anwendung des Algorithmus auffindbar sichtbar gemacht. Solch kleine Merkmale werden also erfaßt und ausgewertet.

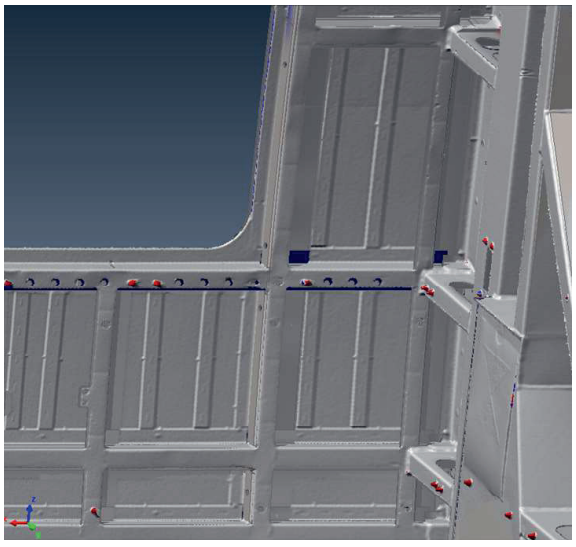


Abb. 7: Ausschnitt Seitenwand. Dargestellt ist der 3D Datensatz, die Farbmarkierung codiert den Abstand der korrespondierenden Punkte der Punktwolke (Software PolyWorks).

Erfahrungen und Ausblick

Mit der Verwendung eines TLS mit angepaßtem Fokusbereich zur Akquise von Punktwolken und der darauffolgendem Bearbeitung der Punktwolken mit einem angepaßten Allgorithmus steht nun ein weiteres Meßverfahren für Stahlbaustrukturen bis 20m zur Verfügung. Ein Proof of Concept wurde positiv validiert (Alstom interner Bericht).

Mit der Ausgabe von Differenzbildern und den darin markierten Fehlteilen gelingt eine von Ermüdung und Erfahrung des Prüfers weitgehend unabhängige Vollständigkeitsprüfung [15], [16] und [18]. Von den ca. dreihundert Anbauteilen werden viele als richtig und vorhanden erkannt. Detektierte Fehler werden innerhalb des Modells nachvollziehbar und gut auffindbar dargestellt. Eine manuelle Nachkontrolle an diesen vergleichsweise wenigen Stellen separiert dann echte Fehleranzeigen von den gelegentlichen Fehlern höherer Ordnung.

Die modellbasierte Sollbeschreibung (mittels 3D CAD Daten Modell) führt zu hoher Flexibilität bei Änderung an den Prüfobjekten (Wagenkästen) und somit zu einer besonderen Eignung für variantenreiche Montageprozesse. Konstruktive Änderungen im CAD Modell können so auch kurzfristig in den Prüfplan (Liste der zu prüfenden Merkmale) einfließen. Besonders wichtig für den Umfang der Rechenoperationen ist die Struktur der 3D-Modell Daten. Soweit es gelingt, durch sinnvolle Strukturierung jeweils ein sogenanntes Light Model mit den Anbauteilen separat vom kompletten Konstruktionsdatensatz auszugeben, stehen geeignete Eingangsinformationen über den Soll Zustand zur Verfügung. Das Light Model lässt sich als Auszug aus der Stückliste verstehen, das also alle Anbauteile aufführt. Der Im- und Export von Prüfmerkmalsplänen aus dem Light Model kann dann gut unterstützt werden. So können also Prüfmerkmalspläne weitgehend automatisiert entstehen (Anbauteil-Name, Identnummer, weiteres).

Bestandteil des Prüfablaufes ist eine Einteilung des Prüfobjektes in Bereiche und zwar sowohl im Datensatz als auch in der Punktewolke. Für diese geometrischen Bereiche können für minimales Rauschen optimale geometrische Bereichsgrößen ermittelt werden. Hier sind weitere Untersuchungen nötig.

Die hier skizzierte Lösung der Einführung eines zweiten Koordinatensystems (lokal für den Bereich der betrachteten Anbauteile,

zusätzlich zum globalen Koordinatensystem für den gesamten Wagenkasten) stellt allgemein ausreichend Freiheitsgrade zur Ausrichtung bereit. Diese werden auch bei Ausrichtvorgängen zwischen den Arbeitsschritten, speziell hybriden Verfahren, der Additiven Fertigung benötigt [17].

Dieses Projekt erhielt Förderung durch das Innovationsprogramm Horizon 2020 der Europäischen Gemeinschaft unter Fördernummer 723826.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No. 723826.



Auch wenn die tatsächlichen Abmessungen in anderen Wertebereichen liegen, kann die Herangehensweise übertragen werden, um die Ungenauigkeit und die Streuung solcher Ausrichtvorgänge zu minimieren.

Literaturnachweis

[1] Dirk Berndt, Ralf Warnemünde: Flexible modellbasierte Montageprüfung. http://www.iff.fraunhofer.de/de/geschaeftsbereiche/messtechnik-prueftechnik/montagepruefung-modellbasiert.html; 2012.

[2] Kelch, J.: Vollautomatisierte 100% Inspektion von Karosserieteilen-Aus allen Perspektiven. Carl-Hanser Verlag München, QZ Jahrgang 60 (2015) 05; Seite 62.

[3] Bevor Fehler entstehen, 3D Bildverarbeitungssystem prüft Schaumstoffteile. Carl-Hanser Verlag München, QZ 60 (2015) 11; Seite 64.

[4] Cem Tarih: Aktueller Stand der Automatisierung von Sichtprüfungen, speziell Vollständigkeitsprüfungen an großen Stahlbaustrukturen am Beispiel Schienenfahrzeugwagenkästen. Studienarbeit an der Ostfalia Braunschweig/ Wolfenbüttel, 2016.

[5] T. Drechsel: Laserprojektion im Lokomotivbau. Positionieren von Anbauteilen in der Produktion der Lokomotive Vectron. DVV Media Group / Eurailpress, EI-Eisenbahningenieur Jahrgang 67 (2016) 04, Seite 16.

[6] Der "Schlaue Klaus" Anwendungen für intelligente Bilderkennung. Optimum GmbH, Karlsruhe, https://www.optimum-gmbh.de/der-schlaue-klaus.html, am 15.08.2016.

[7] R. Schmitt, M. Peterek, E. Morse, W. Knapp, M. Galetto, F. Härtig, G. Goch, B. Hughes, W. Estler: Advances in large scale metrology - review and future trends. CIRP Annals - Manufacturing Technology 65 (2016), 643-665, doi.org/10.1016/j.cirp.2016.05.002.

[8] Th. Kersten, H.-J. Przybilla, R. Lichtenberger, P. Wintjens: Zum Genauigkeitspotential der terrestrischen 3D-Nahbereichscanner Surphaser® 75USR und 100HSX-SR. in Beiträge der Oldenburger 3D-Tage 2018, ISBN 978-3-87907-643-7.

[9] Lichtenberger, R. und Wintjens, P.: Laserscanner Surphaser – effiziente und präzise Digitalisierung für Qualitätskontrolle und Reverse Engineering. in Luhmann, T.; Schumacher, Ch.: Beiträge der Oldenburger 3D-Tage 2017, ISBN 978-3-87907-625-3.

[10] Jurdeczka, U.: Concept for completeness checking of joined structures exemplified on rail vehicle car body shells. J. Sens. Sens. Syst., 6, 53-63, 2017; auch unter http://www.j-sens-sens-syst.net/6/53/2017/, DOI 10.5194/jsss-6-53-2017.

[11] Jurdeczka, U.: Vollständigkeitsprüfung an Schienenfahrzeug-Wagenkästen. Tagungsband der 3D-Nordost 2016, Berlin Adlershof, ISBN 978-3-942709-16-3.

[12] Jurdeczka, U.: Beitrag zur modellbasierten Vollständigkeitsprüfung von Anbauteilen an Stahlbau-Strukturen, vertieft an Schienenfahrzeug-Wagenkästen mit ca. 300 Haltern, Laschen und Erdungskontakten, Messvolumen 20m x 3m x 3m. Dissertation an der TU Braunschweig. Shaker Verlag 2018, ISBN 978-3-8440-5801-7.

[13] http://www.surphaser.com/pdf/ Surphaser%20100 HSX.pdf, am 30.10.2018.

[14] http://www.duwe-3d.de/de/products/polyworks-inspector, am 30.10.2018.

[15] Jurdeczka, U.: Optisches Prüfsystem für variantenreiche Montageprozesse. Carl-Hanser Verlag München, QZ Jahrgang 61 (2016) 07; Seite 46.

[16] Jurdeczka, U.: Konzeption zur Vollständigkeitsprüfung von gefügten Strukturen am Beispiel Schienenfahrzeugwagenkästen. Postervortrag auf der 18. GMA/ITG-Fachtagung Sensoren und Messsysteme 2016, Nürnberg, DOI 10.5162.

[17] P. Penchev, K. Essa, U. Jurdeczka, A. Mehmeti, N. Adkins, L. Carter, S. Dimov, J. Maurath, N. Maillol, J. Bajolet, D. Bhaduri: System-level integration tools for hybrid laser-based powder bed fusion manufacturing platforms. 2018, Journal of Manufacturing Systems, Journal of Manufacturing Systems 50 (2019) 87–102, https://doi.org/10.1016/j.jmsy.2018.12.003.

[18] Jurdeczka, U.: Modellbasierte Analyse von Stahlbau-Strukturen, am Beispiel Maßprüfung und Vollständigkeitsprüfung an Schienenfahrzeug-Wagenkästen. Tagungsband der 3D-Nordost 2018, Berlin Adlershof, ISBN 978-3942709194.

DOI 10.5162/sensoren2019/P3.4

Dimensionale Messtechnik in der digitalen Produktion – Aufgaben, Vernetzung und Schnittstellen

Dietrich Imkamp[1]
[1] Carl Zeiss Industrielle Messtechnik GmbH, Carl-Zeiss-Straße 22, D-73447 Oberkochen

Zusammenfassung

Digitalisierung ist heute ein prägender Begriff auch in der industriellen Produktion. Daraus hat sich der Begriff der „Digitalen Produktion“ entwickelt, Darunter wird einer Produktionsweise verstanden, in der die Informationen durchweg in digitaler Form erzeugt, verwaltet, gespeichert, verteilt, verarbeitet und kommuniziert werden [1]. Für die dimensionale Messtechnik, als der „Informationslieferant“ zur Qualität der Produktion, hat dieses Thema eine besondere Bedeutung. Die digitale Verarbeitung von Messergebnissen ist bereits weit verbreitet. Allerdings beschränkt sich die digitale Weiterverarbeitung dieser Informationen außerhalb der Messtechnik auf einzelne Anwendungsfelder. Die Übernahme von digitalen Informationen in die Messtechnik hat sich dagegen in den letzten Jahren erheblich weiterentwickelt. Darüber hinaus werden sich Geräte für dimensionale Messaufgaben mit allen relevanten Daten wie Werkzeugmaschinen in die digital gesteuerte Produktion integrieren.

Keywords: Dimensionale Messtechnik, Digitalisierung, Industrie 4.0

Einführung

Die Messtechnik wird in der Produktion vorwiegend zur Prüfung der hergestellten Produkte eingesetzt, wobei die Prüfung von geometrischen Maßen und deren Toleranzen [2] dominiert. Sie ist dazu informationstechnisch mit vielen anderen Bereichen des industriellen Produktionsprozesses verknüpft. An der Aufgabenstellung ändert sich in der digitalen Produktion nichts. Aber durch die Art der Verknüpfung mit digitalen Schnittstellen eröffnen sich neue Möglichkeiten der Kommunikation.

Der Schwerpunkt dieses Beitrags liegt bei der Koordinatenmesstechnik. Dabei wird die Produktoberfläche mit einem Messsystem punktweise erfasst, und die bezüglich eines geeigneten Koordinatensystems aufgenommen Punkte mit Hilfe von Verfahren der Ausgleichsrechnung zu Parametern der Produktgestalt (z. B.: Abstände, Parameter einer Flächenbeschreibung) verdichtet [2]. Sowohl zur Punkterfassung als auch zur anschließenden Auswertung werden Digitalrechner verwendet, so dass sich eine digitale Integration dieser Geräte in die Produktion anbietet.

Trotzdem ist zu ergänzen, dass heute auch andere dimensionale Messgeräte in der Produktion mit digitalen Schnittstellen ausgerüstet werden, um besonders Messergebnisse rechnergestützt zu übertragen, so dass auf Papiermessprotokolle verzichtet werden kann. Diese Schnittstellen sind heute auch bei vielen einfachen Handmessmittel (z. B. Messschieber) verfügbar. Eine Integration dieser Messgeräte ist wesentlich einfacher als die Integration der komplexen Messsysteme in der Koordinatenmesstechnik.

Der Prozess der Koordinatenmesstechnik und seine digitalen Schnittstellen

Der Prozess der Koordinatenmesstechnik besteht aus drei Schritten: Messablauferzeugung, Messung und Auswertung [2] (Abbildung 1).

Bei der Messablauferzeugung wird festgelegt, welche Punkte zur Messung der Produktmerkmale erfasst werden. Bei einem Koordinatenmessgerät, mit dem ein einen Sensor bewegendes System bezeichnet wird [3], werden bei der Ablauferzeugung die Sensor-Verfahrbewegungen definiert. Für ein optisches Koordinatenmesssystem werden Ansichten bestimmt, um die erforderlichen Punkte zu erfassen [4]. Als Basis dient dazu in der digitalen Produktion ein zur Messsoftware übertragenes CAD-Modell des Produktes, das heute auch die Produktmerkmale enthält (Abbildung 2). Zur Übertragung der Produktmerkmale dienen mit dem geometrischen Produktmodell verknüpfe

Attribute, die als „Product Manufacturing Information" (PMI) bezeichnet werden [5].

Für Koordinatenmessgeräte existiert zur digitalen Übertragung von Messabläufen das DMIS-Format. Sein funktionaler Umfang deckt jedoch nur Punkt- und Liniensensoren ab [6]. Herstellerspezifische Erweiterungen zur Vergrößerung des Funktionsumfangs schränken die Übertragungsmöglichkeiten an unterschiedliche Geräte stark ein.

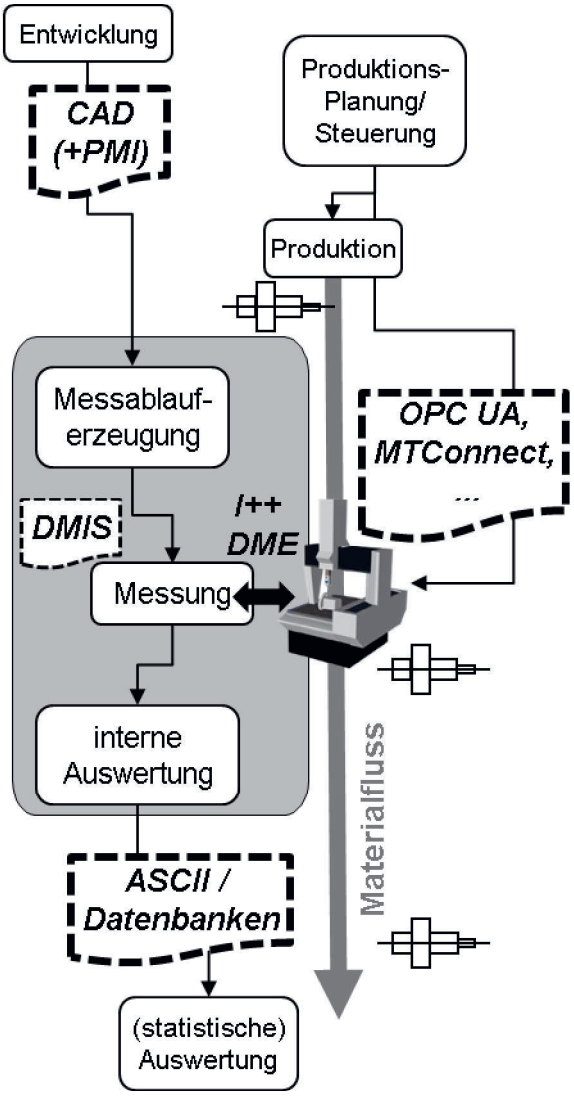


Abb. 1: Prozess der Koordinatenmesstechnik mit digitalen Schnittstellen zur Integration in die Produktion

Die Messung wird durch ein Koordinatenmesssystem ausgeführt [2]. Im Bereich der Koordinatenmessgeräte gibt es zur Kommunikation zwischen Gerätesteuerung und Messsoftware das standardisierte Format „I++ DME" [7]. Es allerdings genauso wie DMIS funktional auf bestimmte Sensoren beschränkt.

Die Ergebnisse der Koordinatenmessung werden nach der internen Auswertung (Verdichtung der aufgenommenen Koordinatenpunkte zu Parametern der Produktgestalt, Filterung der Punkte, usw. [2]) vorwiegend in verschiedenen ASCII-Formaten (z. B.: DMIS-Out, Q-DAS) oder über geeignete Softwareschnittstellen direkt an Datenbanken digital weitergegeben. Ein allgemein anwendbarer Standard hat sich an dieser Stelle nicht durchgesetzt [6].

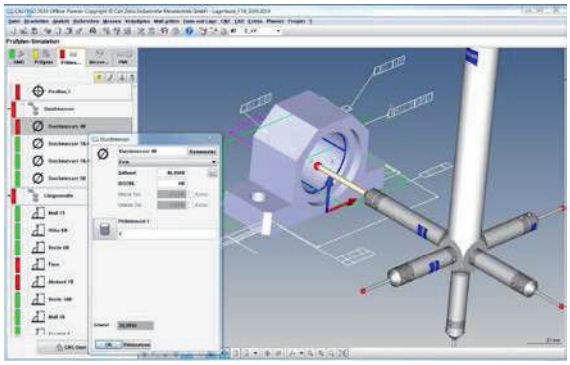


Abb. 2: CAD-Modell mit Maßen und Toleranzen in einer Software für Koordinatenmessgeräte mit Visualisierung des Tasters

Die Messergebnisse werden einerseits genutzt, um die Konformität mit den Produktanforderungen zu prüfen und andererseits, um in vorgelagerten Bereichen Verbesserung durchzuführen, so dass Regelkreise entstehen [2]. Dabei werden unterschiedliche Übertragungsformate und Formen der Ergebnisdarstellung benutzt. Bei einzelnen Produktionsverfahren existieren auch automatische Verfahren zur Rückführung auf Basis digitaler Informationen (z. B.: Voreinstellung von Erodier-Maschinen auf Basis von Koordinatenmessungen) [6].

Die Beschreibung des Prozesses der Koordinatenmesstechnik verdeutlicht die Komplexität der dazu erforderlichen Schnittstellen und Datenstrukturen [8]. Mit dem QIF (Quality Information Framework) Standard [9, 10] wird durch einen ganzheitlichen Ansatz versucht, diese Komplexität zu beherrschen. Der Standard besteht aus 6 Anwendungsmodellen, einer von allen Modellen genutzten Bibliothek sowie einer Verknüpfung zu dem hier zuvor bereits erwähnten DMIS Standard. Informationen zur Produktspezifikation einschließlich der Produktgestalt werden von einem CAD-Modell mit PMIs (s.o.) bezogen. Als Datenformat wird XML benutzt.

Digitaler Schatten und Digitale Zwilling der Koordinatenmesstechnik

Unter einem Digitalen Schatten wird in der Produktionstechnik ein Abbild der Prozesse in der Produktion, der Entwicklung und angrenzenden Bereichen verstanden mit dem Zweck, eine echtzeitfähige Auswertungsbasis aller relevanten Daten zu schaffen. Der Digitale Schatten überführt zunächst den realen Prozess in die virtuelle Welt. Der Digitale Zwilling kann darauf aufbauend durch ein Prozessmodell und Simulation ein möglichst identisches Abbild der Realität liefern [11].

Im Zusammenhang mit der Koordinatenmesstechnik spielen Modelle des Messprozesses zur Korrektur systematischer Abweichung schon seit ihren Anfängen eine wichtige Rolle [12]. Heute kommen darauf aufbauend Modelle zur Abschätzung der Messunsicherheit durch Simulation, die unter dem Begriff der „virtuellen Messgeräte“ bekannt geworden sind, zur Anwendung [13].
Weiterhin werden bei der Off-Line Programmierung von Messabläufen digitale Zwillinge der Produkte mit ihren Aufspannvorrichtungen, die als CAD Modelle importiert werden, zusammen mit digitalen Zwillingen der Geräte zur Sensor-Verfahrweg-Festlegung und Kollisionskontrolle durch Verfahrweg-Simulation genutzt (Abbildung 3).

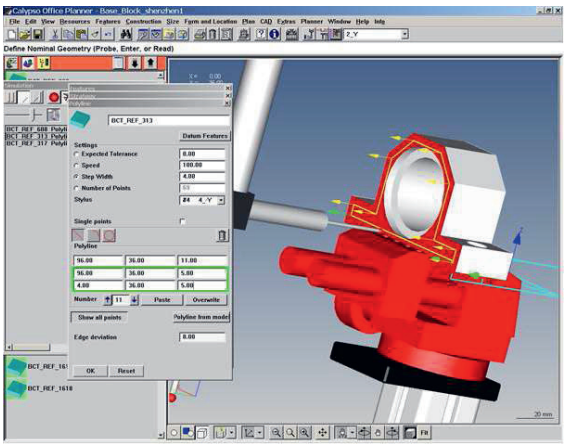


Abb. 3: Kollisionskontrolle durch Verfahrweg-Simulation an einem Digitalen Zwilling des Produktes mit Aufspannvorichtung

Diese digitalen Zwillinge ermöglichen eine effizientere Nutzung der Koordinatenmesstechnik durch die Reduzierung und Bestimmung der Messunsicherheit sowie die Off-Line Programmierung zur Messablauferstellung ohne Messgerät. Zur Integration der Koordinatenmesstechnik in ein digitales Abbild der Produktion dienen diese Zwillinge aber nicht.

Digitale Integration der Koordinatenmesstechnik

Für die Integration der Koordinatenmesstechnik in das digitale Abbild der Produktion und damit auch in Produktionsplanung und Steuerung (Abbildung 1) ist ein digitaler Schatten der real verfügbaren Koordinatenmesssysteme entsprechend der zuvor beschriebenen Abgrenzung zum digitalen Zwilling erforderlich. Dazu sind alle relevanten Daten eines Koordinatenmesssystems zu definieren, und sie über eine geeignete Schnittstelle digital zur Verfügung zu stellen. Als Schnittstelle bieten sich dazu die neuen Entwicklungen im Bereich der Produktionstechnik wie MTConnect [14] aus den USA oder OPC UA (Open Platform Communications Unified Architecture) [15] aus Europa an. Beide Standards bieten unterschiedliche Modelle zur Beschreibung von Systemen in der Produktion an. Für Koordinatenmesssysteme gibt es bis auf eine prototypenhafte Implementierung für MTConnect mit beschränktem Funktionsumfang keine Vorlage [16]. Daher wird hier ein hierarchisches Datenmodell für ein Koordinatenmesssystem aufgestellt und in Form eines einfachen Entity-Relationship Diagramms dargestellt (Abbildung 4).

Das Modell lehnt sich an die MTConnect Beschreibung eines „Devices“ an. Es beschreibt ein Koordinatenmesssystem als Gerät (Device) mit seinen Eigenschaften und untergeordneten Komponenten (Components). Diese Daten bilden die weitgehend unveränderlichen Stammdaten des Systems.
Zusätzlich werden der aktuelle Gerätezustand einschließlich Informationen über die Umgebungsbedingungen und die Fähigkeit bestimmte Messabläufe unter Berücksichtigung der Anforderungen an die Messunsicherheit auszuführen dargestellt. Über die ausführbaren Messabläufe wird eine Verbindung zu den Produkten und ihren Produktmerkmalen hergestellt. Diese Daten verändern sich während des Geräteeinsatzes und werden laufend aktualisiert. Diese Aktualisierung erfolgt automatisch beispielsweise bei den Umgebungsbedingungen, die durch Sensoren erfasst werden, oder durch manuelle Eingabe beispielsweise in Form der Bestätigung einer erfolgreich durchgeführten Kalibrierung.

Schließlich speichert das Gerät Daten zur Historie beispielsweise zu Kollisionen, um den

Gerätezustand kontinuierlich zu dokumentieren. Diese Informationen können dazu dienen, die Richtigkeit von Messergebnissen aus der Vergangenheit zu bestätigen, oder vorausschauende Wartungen („Predictive Maintenance“) zu planen [17].

Die Kategorisierung der Daten des Systems in Stammdaten, Bewegungsdaten und Historiendaten orientiert sich an den Datenstrukturen, die schon heute typischerweise in der Prüfmittelüberwachung verwendet werden [2]. Mit der automatisierten digitalen Bereitstellung dieser Daten lässt sich die Überwachungsdichte einerseits verdichten und andererseits effizienter gestalten.

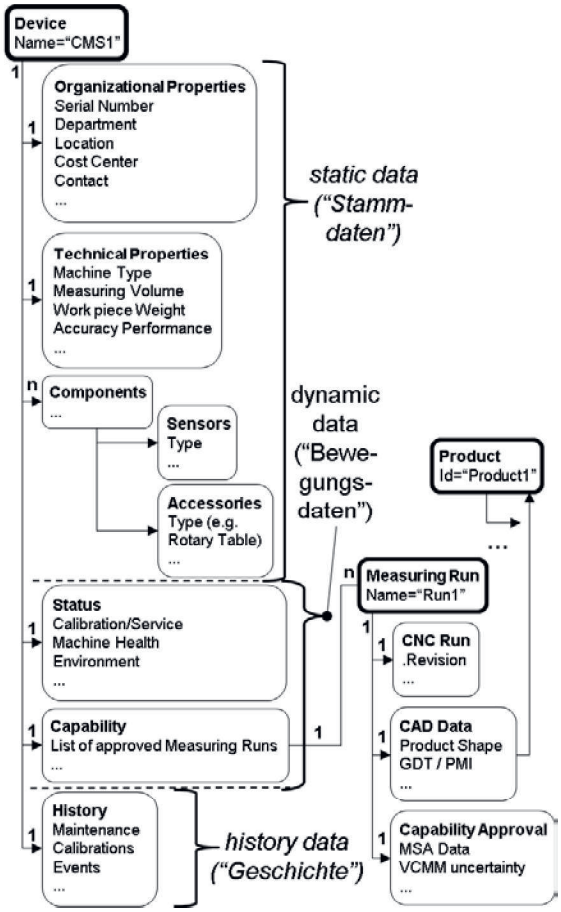


Abb. 4: Datenstruktur für einen digitalen Schatten eines Koordinatenmesssystems

Die Daten dieses digitalen Schattens können durch die Produktion genutzt werden, um ein geeignetes Messgerät zur Prüfung eines bestimmten Produktes zu finden. Das Prüfmittelmanagement kann auf Basis der Daten Wartungen und Kalibrierungen planen.
Dieser Ansatz eines Datenmodells für einen digitalen Schatten eines Koordinatenmesssystems ist nicht vollständig, und die Anwendungsszenarien für die Daten werden Ergänzungen und zusätzliche Verknüpfungen erfordern. Dennoch machen die Anwendungsbeispiele deutlich, dass ein Koordinatenmesssystem wie eine Werkzeugmaschine in die digitale Produktion über die gleichen Schnittstellen integrierbar ist.

Zusammenfassung

Der digitale Informationsfluss für ein Koordinatenmesssystem und die mögliche Darstellung seiner Datenstruktur für einen digitalen Schatten zeigen die Möglichkeiten der Einbindung von Koordinatenmesssystemen in die digitale Produktion auf.

Literaturnachweis

[1] Westkämper , E. Spath, D., Constantinescu, C., Lentes, J.: Digitale Produktion, Verlag: Springer, 2013.
[2] Pfeifer, T., Schmitt, R.: Fertigungsmesstechnik, Oldenbourg Verlag, München 2010.
[3] DIN EN ISO 10360-1 Geometrische Produktspezifikation (GPS), Annahmeprüfung und Bestätigungsprüfung für Koordinatenmess-geräte (KMG), Teil 1: Begriffe Juli 2003.
[4] Imkamp, D., Wanner, J.: Vom Messablaufplan über die Programmierung, Durchführung und Auswertung bis zur Ergebnis-darstellung, in: Weckenmann, A. (Hrsg.): Koordinatenmess-technik: Flexible Strategien für funktions- und fertigungsgerechtes Prüfen. Carl Hanser Verlag, München 2012.
[5] Imkamp, D., Gabbia, A.: Nutzung des PMI Standards zur automatisierten Übertragung von Merkmalsdaten in die dimensionale Messtechnik, in: ITG-Fachbericht 250, Sensoren und Messsysteme 2014 zur 17. ITG/GMA-Fachtagung Sensoren und Messsysteme 2014 vom 03. bis 04. Juni 2014 in Nürnberg.
[6] Pfeifer, T., Imkamp, D.: Koordinatenmesstech-nik und CAx-Anwendungen in der Produktion – Grundlagen, Schnittstellen, Integration, Carl Hanser Verlag, München 2004.
[7] Imkamp, D., Resch, J., Puntigam, W.: Nutzung der neuen hersteller-neutralen I++ DME Schnittstelle bei der Prüfung der Messabweichungsangaben von Koordinatenmessgeräten, in: Vorträge der 13. ITG/GMA-Fachtagung, Sensoren und Messsysteme in Freiburg/Breisgau, Germany, 13./14. März 2006, VDE Verlag GmbH, Berlin 2006.
[8] Evans, J., Frechette, S., Horst, J., Huang, H., Kramer, T., Messina, E., Proctor, F., Rippey, B., Scott, H., Vorburger, T., Wavering, A.: Analysis of Dimensional Metrology Standards (NISTIR 6847), National Institute of Standards and Technology (NIST), USA 2001
[9] ANSI/DMSC QIF 3.0 Quality Information Framework. Digital Metrology Standards Consortium, Inc. (DMSC), USA, 2018.

[10] Zhao, Y. F., Horst, J. A., Kramer, T. R., Rippey, W., Brown, R. J.: Quality Information Framework – Integrating Metrology Processes, in: IFAC Proceedings Volumes 45, Issue 6, 23–25 May 2012, Pages 1301-1308, Internetadresse: https://doi.org/10.3182/20120523-3-RO-2023.00113.
[11] Bauernhansl, T.; Krüger, J.; Reinhart, G. et al. (Hrsg.: Wissenschaftliche Gesellschaft für Produktionstechnik Wgp e.v.): Wgp-Standpunkt Industrie 4.0, Internetadresse: www.wgp.de/uploads/media/WGP-Standpunkt_Industrie_4-0.pdf. Zuletzt aufgerufen am 25.04.2019.
[12] Wäldele, F.: Genauigkeitssteigerung von Koordinatenmessgeräten durch numerische Fehlerkorrektur. In: Neumann, H. J. (Hrsg.): Präzisionsmesstechnik in der Fertigung mit Koordinatenmessgeräten, Expert Verlag, Ehningen bei Böblingen 1988.
[13] Schmitt, R., et. al.: Virtuelle Messgeräte: Definition und Stand der Entwicklung, in: tm - Technisches Messen, Oldenbourg Industrieverlag, 75 (2008) 5, S. 298-310.
[14] MTConnect Standard, Version 1.4.0 (ANSI/MTC1.4-2018). MTConnect Institute, AMT - The Association For Manufacturing Technology, USA (22.04.2019: https://www.mtconnect.org/standard-download20181).
[15] IEC 62541 OPC Unified Architecture. Part 1-14, 100, 2010-2018.
[16] MTConnectToolbox: https://github.com/johnmichaloski/MTConnectToolbox Zuletzt aufgerufen am 25.04.2019.
[17] Imkamp, D., Grieser C.: Onboard Diagnostics - Ein innovatives System zur Maschinenüberwachung. In: Sensoren und Messsysteme 2004, Tagung Ludwigsburg 15.-16.3.2004 (VDI-Berichte 1829), VDI-Verlag, Düsseldorf 2004, S.785-789.

DOI 10.5162/sensoren2019/P3.6

Untersuchung eines Prozessroboters als Messroboter in der Montage

Prof. Dr.-Ing. Rainer Müller[1,2], Dr.-Ing. Matthias Scholer[2], Ali Kanso M.Sc.[2], Anne Blum M.Sc.[1]
[1]Universität des Saarlandes, Lehrstuhl für Montagesysteme, Eschberger Weg 46, 66121 Saarbrücken, Deutschland
[2]ZeMA - Zentrum für Mechatronik und Automatisierungstechnik gemeinnützige GmbH, Gruppe Montageverfahren und –automatisierung, Eschberger Weg 46, 66121 Saarbrücken, Deutschland

Zusammenfassung

Im Zuge der Produktionsdigitalisierung werden mehr Daten aufgenommen und ausgewertet. Neue Roboter besitzen eine integrierte Drehmomentensensorik. Diese erfasst nicht nur interne Roboterdaten, sondern auch externe Einflüsse. Diese Daten können zur Inbetriebnahme von Prozessen oder zur Aussage der Qualität verwendet werden. Prozesse werden somit transparenter und quantifizierbarer. Um den Roboter als Messgerät einzubinden, bedarf es einer Untersuchung der Prüfprozesseignung. Zur Untersuchung wurde eine neuartige Vorgehensweise entwickelt, welche zwei vorhandene Methoden aus unterschiedlichen Disziplinen miteinander kombiniert. Zum einen wird die Methode DMAIC aus dem Qualitätsmanagement zur Strukturierung der Untersuchung und zum anderen wird der Eignungsnachweis von Prüfprozessen zur inhaltlichen Ermittlung von Wissen verwendet. Durch einen Versuchsaufbau werden Ergebnisse identifiziert, um die systematische und zufällige Messabweichung des Roboters zu bestimmen. Ziel dieses Papers ist die Analyse für den Einsatz eines Roboters als Messgerät im Montageprozess um somit neue Anwendungsfelder zu ermöglichen.

Keywords: Roboter, Drehmomentensensor, Montage, Messunsicherheit

Einleitung und Motivation

In der Produktion nimmt die Anzahl der Prüfprozesse durch steigende Qualitätsanforderungen stetig zu. Ziel ist Wissen über das Produkt und den Prozess zu generieren. Nach der Gesellschaft für Mess- und Automatisierungstechnik (GMA) des VDI/VDE ergeben sich daraus neue Anforderungen für die Messtechnik: „*Schneller, Sicherer, Genauer und Flexibler*“ [1]. Um den Anforderungen gerecht zu werden, gibt es seit Jahren einen Trend zur Integration der Messsysteme in die Produktion. Messungen finden nicht mehr im separaten Messraum, sondern in der Montagelinie statt. Durch die Integration wird eine 100%-Prüfung der Produkte ermöglicht. Einerseits sind die Auswahl und das Anwendungsspektrum von Messsystemen enorm, andererseits wird in vielen Unternehmen die Messtechnik als ein reiner Kostenfaktor gesehen. [1] Bei dem Einsatz in der Produktion ist zusätzlich ein Prüfmittelmanagement notwendig. Dieses sorgt dafür, dass das richtige Prüfmittel zum richtigen Zeitpunkt, am richtigen Ort, in einem einsatzbereiten Zustand zur Verfügung steht [2]. Dies ist mit zusätzlichen Kosten, Zeit und Ressourcen verbunden. Gleichzeitig schränkt es die Flexibilität ein, welche heutzutage immer bedeutender wird.

Im Zuge von Industrie 4.0 werden für eine angepasste Automatisierung Mensch-Roboter-Kooperationen (MRK) in der Montage eingesetzt. Die verwendeten Leichtbauroboter besitzen eine integrierte Sensorik, wie z.B. Drehmomentensensoren in den Roboterachsen. Diese Sensoren dienen zum Schutz, um Kollisionen rechtzeitig zu erkennen und externe Drehmomente zu ermitteln. Das Lean Management handelt von der Vermeidung von Verschwendung in der Produktion. Prozesse sollen schlank, beispielsweise ohne Wartezeit, sinnlose Transporte oder Nacharbeit erfolgen. [3] Beeinflusst durch diesen Gedanken, d.h. der konstanten Verbesserung, ist es sinnvoll den vorhandenen Prozessroboter auch gleichzeitig als Messgerät zu nutzen. Während des eigentlichen Montageprozesses kann so parallel eine In-Prozess-Messung erfolgen. Anhand der gemessenen

Drehmomente in den Roboterachsen können die Masse und der Massenschwerpunkt eines Roboterwerkzeuges sowie von einem zu greifenden Bauteil im Montageprozess bestimmt werden. Dieser Ansatz spart Ressourcen sowie Kosten und steigert die Wandlungsfähigkeit im Prozess, da keine zusätzlichen Betriebsmittel implementiert werden müssen. Um diese Idee in der Praxis umzusetzen, bedarf es einer ausführlichen Betrachtung des Messroboters, um herauszustellen wie zuverlässig dessen Messergebnisse sind.

Stand der Technik

Prüfprozesseignung

Laut der GUM (Guide to the Expression of Uncertainly in Measurement) existiert ein wahrer Wert der Messgröße. Grundsätzlich ist dieser unbekannt. Die Messunsicherheit ist ein Maß dafür, wie gut der Mensch glaubt, den wahren Wert zu kennen. [4] Es gibt zwei Arten von Messabweichungen, die systematischen und die zufälligen. Die systematische Messabweichung ist die Differenz (Offset) zwischen dem Referenzwert eines Normals und dem Mittelwert der gemessenen Werte. Beschrieben wird sie durch den sogenannten Bias. Durch nicht beherrschte Einflussgrößen erfolgt bei einer wiederholten Messung eine andere Messwertausgabe. Diese zufällige Messabweichung lässt sich nicht vorhersehen und wird durch die Standardabweichung beschrieben. [5,6,7]

Zur Feststellung der Messunsicherheit gibt es standardisierte Vorgehensweisen. Dazu zählen Unternehmensrichtlinien, wie von Bosch oder Ford, die Measurement System Analysis (MSA) von der Automotive Industry Action Group (AIAG) und der Verband der Deutschen Automobilindustrie (VDA) mit dem Band 5 Prüfprozesseignung [6,7]. Die Vorgehensweise zur Beurteilung eines Messsystems ist bei allen Verfahren ähnlich. Unter realen Bedingungen erfolgen Messungen. Die gewonnenen Messwerte werden grafisch und numerisch analysiert und Kennwerte bestimmt, welche mit vorgegebenen Grenzwerten verglichen werden. Der Unterschied der Verfahren liegt in der Berechnung der Kennwerte sowie der Betrachtung der Anzahl der Einflussgrößen. [8] Die MSA hat einen hohen Bekanntheitsgrad und ist Basis für viele Firmenrichtlinien. In der MSA erfolgt die Berechnung der GRR-Wertes (GageRepeatility and Reproducibility). Die Firmenrichtlinien erweitern dies durch die Bestimmung der Cg- bzw. Cgk-Werte, welche die systematische Messabweichung untersucht (Verfahren 1). Der VDA Band 5 hingegen basiert auf der ISO-Norm sowie der GUM. Das Beurteilungskriterium ist ein Eignungskennwert QMP (Capability Ratio Process), der die doppelte erweiterte Messunsicherheit ins Verhältnis zur Toleranz setzt. [8]

Messungen in der Montage

In der VDI/VDE erfolgt die Unterteilung nach dem Grad der Integration in die Produktion [1]. Die Unterteilung kann in diesem Fall, wie in Abb. 1 dargestellt auf die Montage übertragen werden. Messungen können in einem separaten Messraum (Montagefern), in der Umgebung der Montagelinie (Montagenah), in einer Messstation der Montagelinie (Inline), vor oder nach der Montage (Pre- bzw. Post-Prozess) oder während der eigentlichen Montage (In-Prozess) erfolgen. Durch die zunehmende Integration in die Montage sinkt die Reaktionszeit in den Prozess einzugreifen. Mit der In-Prozess-Messung erfolgt eine 100% Prüfung. Einerseits werden so keine fehlerhaften Produkte an den Kunden weitergegeben, was zugleich einer Null-Fehler-Strategie entspricht [9]. Andererseits steigen die Einfluss- bzw. Störgrößen auf die eigentliche Messung. In der Montage herrschen nicht immer konstante Bedingungen, so dass Störgrößen wie beispielsweise die Umwelt berücksichtigt werden müssen.

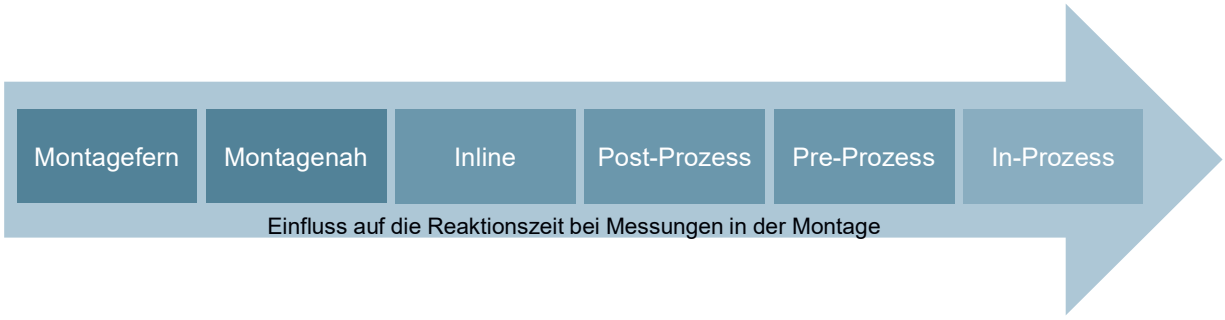


Abb. 1: Integrationsgrad der Messung in der Montage (in Anlehnung an [1])

Roboter

Zur Erprobung der Forschungsidee und Validierung der entwickelten Vorgehensweise wurde ein möglichst geeigneter Roboter ausgewählt.

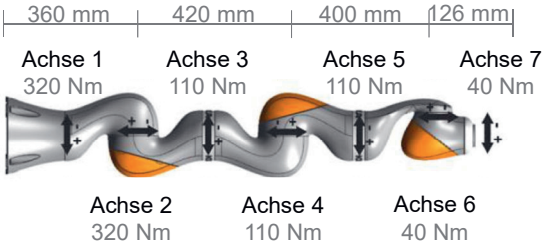


Abb. 2: Roboter KUKA LBR iiwa 14 R820 mit sieben Drehmomentensensoren

Der Roboter KUKA LBR iiwa 14 R820 ist ein Leichtbauroboter (LBR) und ein „intelligent industrial work assistant“ (iiwa). Mit einer Traglast von 14 kg und einer Reichweite von 820 mm eignet er sich besonders für sensitive Fügeaufgaben sowie die Zusammenarbeit mit dem Menschen. Ermöglicht wird dies durch die besondere interne Sensorik und die Regelmechanismen. Dafür besitzt der LBR iiwa in allen Achsen Winkel- und Gelenkmomentensensoren. Neben der Kollisionserkennung kann die Sensorik auch für weitere Anwendungen verwendet werden. Greift der Roboter ein Werkstück, übt dies eine zusätzliche Kraft aus und hat somit Einfluss auf die gemessenen Momente in den Gelenkmomentensensoren. Der Roboter prüft den ermittelten Wert aus der Gelenkmomentensensor mit dem erwartetenen externen Drehmoment, das für eine Achse aufgrund des Robotermodells und der angegebenen Lastdaten berechnet wird. Die Drehmomentensensoren für den LBR iiwa 14 sind, wie in Abb. 2 dargestellt, unterschiedlich festgesetzt: 320 Nm in Achse 1 und 2, 176 Nm in Achse 3 und 4, 110 Nm in Achse 5 sowie 40 Nm in Achse 6 und 7. Die achsspezifische Momentengenauigkeit wird mit ± 2 % angegeben. [10]

Methodische Vorgehensweise

Ziel dieses Papers ist die Analyse für den Einsatz eines Roboters als Messgerät im Montageprozess. Dazu zählt eine erste Voruntersuchung wie genau die gemessenen Drehmomente sind. Für die Vorgehensweise werden zwei bestehende Ansätze aus der Literatur kombiniert. Die Methode DMAIC (Define, Measure, Analyse, Improve, Control) aus dem Qualitätsmanagementansatz Six Sigma wird verwendet [11]. DMAIC ist in der Produktion eine etablierte Methodik, vor allem wird sie in Kontinuierlichen Verbesserungsprozessen (KVP) eingesetzt. Die einzelnen DMAIC-Phasen dienen dem Prozessablauf der Untersuchung. Die Innovation ist, dass diese Phasen inhaltlich mit den Methoden der Prüfprozesseignung kombiniert werden. Durch die Verbindung der zwei Methoden, kann ein korrekter Ablauf der Untersuchung erfolgen. Die entwickelte Vorgehensweise ist schematisch in Abb. 3 dargestellt, beginnend mit der Beschreibung der Zielstellung, der Angabe zum Messsystem sowie einer Versuchsplanung. Anschließend folgt die Messdurchführung, Analyse der Messwerte sowie das Generieren von Verbesserungsideen für den Versuchsaufbau oder für spätere Applikationen, die Dokumentation sowie Kontrolle. Die Messungen am Roboter sowie die Analyse der gewonnenen Daten, geben Rückschlüsse auf die Streuung des Drehmomentes und auf die Wiederholbarkeit in den einzelnen Achsen. Dabei wird die Messunsicherheit des Roboters mit integrierten Drehmomentensensoren bestimmt. Bei der allgemeinen Prüfprozesseignung liegt ein Anwendungsfall mit bestimmten Toleranzangaben für ein

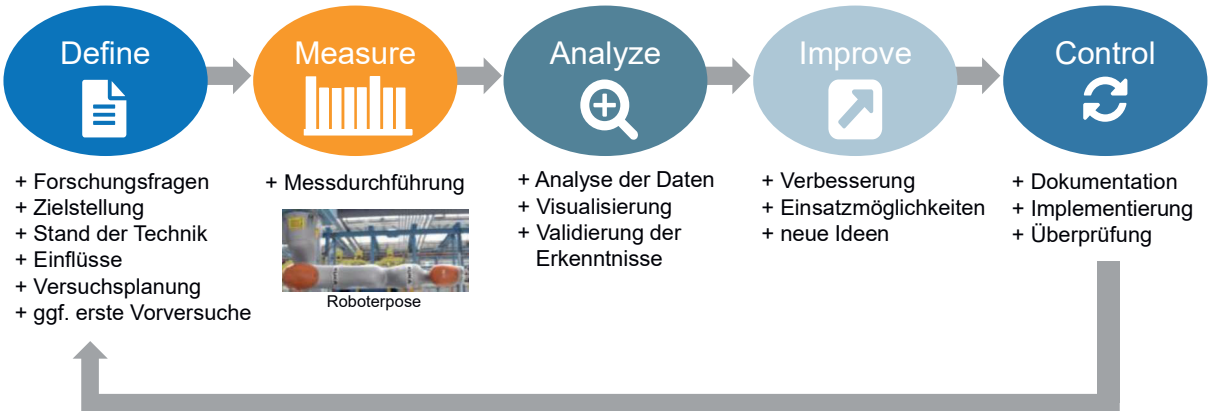


Abb. 3: Vorgehensweise der Untersuchung eines Prozessroboters als Messroboter

Produktmerkmal vor. Eine Toleranzgrenze ist in diesem Fall nicht gegeben. Aus diesem Grund erfolgen Messungen mit Lasten über die gesamte Nutzlast am Roboterflansch, um dabei die Drehmomente in den Roboterachen zu beobachten und ein aussagekräftiges Ergebnisse zu erhalten.

Umsetzung und Validierung

In der **Define-Phase** wird die Zielstellung, Einflussmöglichkeiten auf die Messung und Messplanung definiert. Der Versuchsaufbau wird mit möglichst geringen Einflussgrößen gewählt. Durch Lasten werden Drehmomente erzeugt, aufgenommen und analysiert. Dazu werden am Roboterflansch Gewichte der gesamten Nutzlast, d.h. von wenigen Gramm bis zur Nennlast befestigt um eine vollständige Aussagekraft zu erhalten. Das Drehmoment in den Achsen wird während der Messungen aufgezeichnet und kann im Anschluss ausgewertet werden. Der Roboter befindet sich während der Messung in einer definierten Pose, d.h. Achse 2 bei 90 Grad und die restlichen Achsen bei 0 Grad (siehe Abb. 3).

Da keine Messung exkat ist, sondern das Ergebnis vom Messsystem, dem Messverfahren, der Qualifikation des Bedieners, der Umgebung und anderen Einflüssen abhängt, werden diese Einflüsse für den Versuchsaufbau dokumentiert [5,6]. Ein Ishikawa-Diagramm in Abb. 4 zeigt die Einflusskomponenten auf das Messystem. Es ist ein Ursachen-Wirkungs-Diagramm zur systematischen Suche nach den Ursachen [12]. Es zeigt Einflüsse auf das Messergebnis durch die Messmethode, Messgerät, Auswertung, Mensch, Messobjekt und die Umwelt auf. Die Messmethode beeinflusst beispielsweise das Messergebnis, durch die Dauer und Anzahl der Messungen sowie die Messfrequenz. Durch Filter, Messwertverknüpfungen oder Wahl der Daten einer spezifischen Roboterachse kann die Auswertung ebenfalls das Messergebnis verändern. Natürlich ist die Umwelt mit der Temperatur oder Schwingungen nicht zu vernachlässigen. Ebenso ist der Mensch ein Einflussfaktor, da die Wahl des Roboters und die Sorgfalt der Durchführung eine Auswirkung hat. Diese gesamten Einflüsse auf das Messergebnis wirken sich meistens als zufällige Abweichungen auf das Messergebnis aus.

Die eigentliche Messdurchführung am Roboter erfolgt in der **Measure-Phase**. Das Drehmoment (M) setzt sich dabei aus der Gravitationskraft (F) der gewählten Last und der Länge des Hebelarms (r), d.h. von der Achse zum Roboterflansch, zusammen (Gl.1).

$$M = r \cdot F \qquad (1)$$

Bei der **Analyze-Phase** erfolgt das Aufbereiten, Auswerten und Visualisieren der Daten, um schlussendlich Wissen zu generieren. Bei der Betrachtung der Daten zeigt sich, dass durch die Anordnung der Achsen in der gewählten Pose im Messaufbau ein Hebelarm auf die Achsen 2, 4 und 6 entstehen und diese als sinnvolle Messergebnisse verwendet werden können. Auf den Achsen 1, 3, 5 und 7 wirkt hingegen kaum ein messbares Drehmoment.

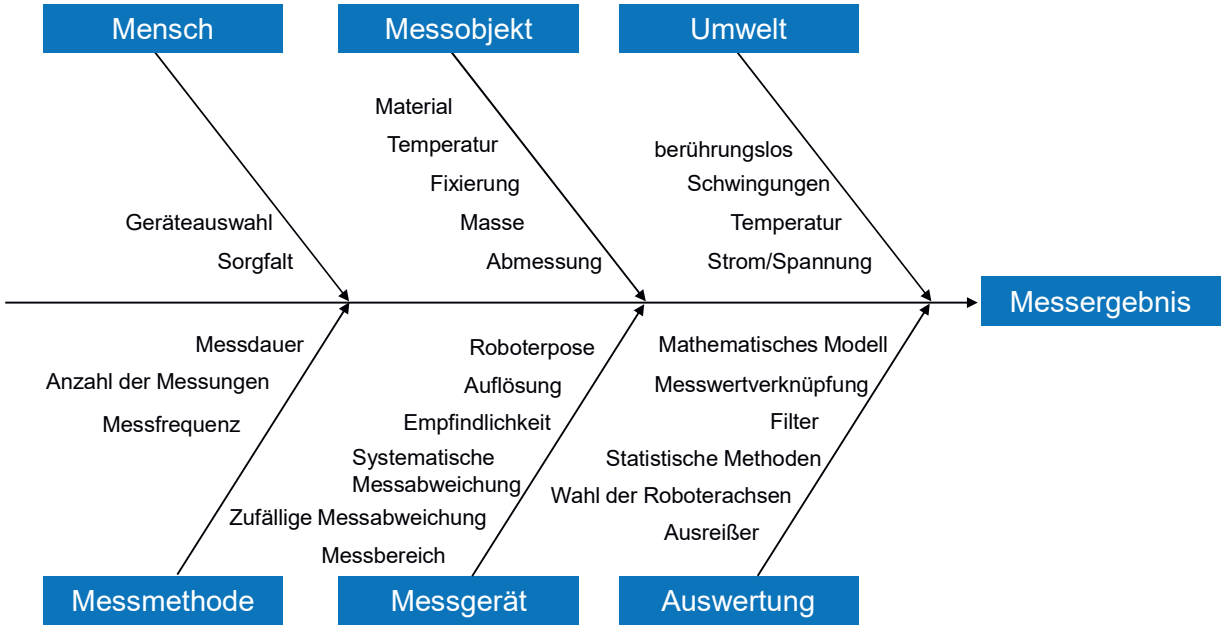


Abb. 4: Ursachen-Wirkungs-Diagramm der Einflüsse auf das Messergebnis

Nachfolgend werden einige Ergebnisse vorgestellt:

- systematische Messabweichung
- zufällige Messabweichung
- Reproduzierbarkeit
- Empfindlichkeit

Die systematische Messabweichung, auch Bias genannt, ist die Abweichung zwischen dem Mittelwert der gemessenen Werte bei wiederholtem Messen $(\overline{x}_g)$ und dem Referenzwert des Merkmals (x_m) (Gl.2) [13].

$$Bi = \left|\overline{x}_g - x_m\right| \qquad (2)$$

In der Achse 2 zeigt sich ohne Last ein Drehmoment von 82,10 Nm. Bei einer zusätzlichen Last von 10 kg am Roboterflansch ergibt sich ein Drehmoment von 172,89 Nm. Die Drehmomentdifferenz liegt folglich bei 90,79 Nm. Der errechnete Referenzwert, aus Hebelarm multipliziert mit der Last ergibt annähernd den selben Wert. Daraus folgt ein Bias von 0,13 Nm. Um die Linearität der Drehmomentsensoren zu bestimmen, wurden Messungen mit <100 g, 500 g, 1 kg, 5 kg, 10 kg und 14 kg mit der Anzahl von n = 25 durchgeführt. Die Messdaten in Abb. 5 zeigen eine annähernd lineare Kennlinie. Bei Gewichten unter 100 g ergibt sich ein max. Fehler von 15 g. Es zeigt sich, dass ein geringer systematischer Messfehler vorliegt, der durch ein Offset ausgeglichen werden kann.

Für die zufällige Messabweichung wird die Standardabweichung verwendet. Sie ist das häufigst verwendete Maß um die Streubreite der Daten um den Mittelwert zu quantifizieren. Die Standardabweichung der Messungen sind unter 0,03 Nm. Beispielsweise liegt in der Achse 2 ohne Last am Roboterflansch eine Standardabweichung von 0,018 Nm und bei einer Last von 14 kg eine Standardabweichung von 0,016 Nm vor.

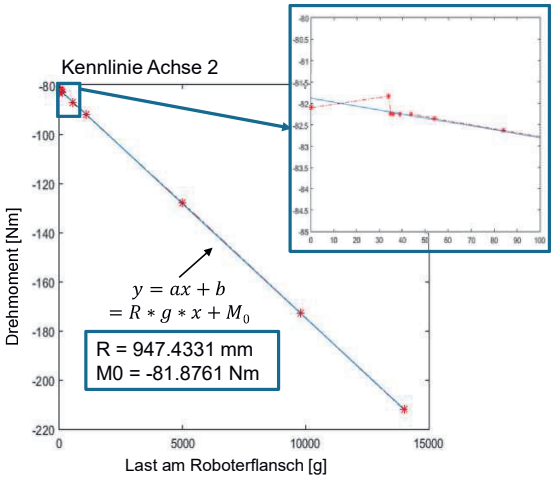


Abb. 5: Linearität in Achse 2

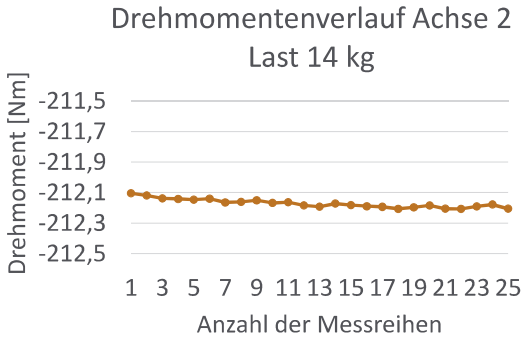


Abb. 6: Drehmoment in Achse 2 bei 25 Messungen bei einer Last von 14 kg

Die Reproduzierbarkeit ist definiert als die Variation des Mittelwertes bei Messungen [7]. Bei dem wiederholten Messen mit einer Anzahl von n = 25 unter denselben Messbedingungen, zeigen sich verschiedene Mittelwerte. Die Mittelwerte zeigen eine max. Drehmomentendifferenz von 0,10 Nm. In Abb. 6 ist der Verlauf des Drehmomentes bei 25 Messungen für die Last von 14 kg dargestellt. Des Weiteren zeigte sich, dass die Standardabweichung der Mittelwerte nach 25 Messungen konstant verläuft.

Für die Empfindlichkeit, d.h. den Einsatz des Roboters bei geringen Lasten zu bestimmen, wurden Gewichte von 5 g, 10 g, 20 g und 50 g am Flansch befestigt. Die Messungen wurden mit einer Anzahl n = 25 durchgeführt. Es zeigt sich, dass die geringen Lasten in den Daten zu erkennen ist. Dennoch ist in Abb. 7 abzulesen, dass Lasten unter 20 g nicht differenziert werden können. Die Streuung der Messdaten ist zu hoch und lässt keine Unterscheidung zu. In Achse 2, 4 und 6 können ab 20 g eine nachvollziehbare Aussage getroffen werden.

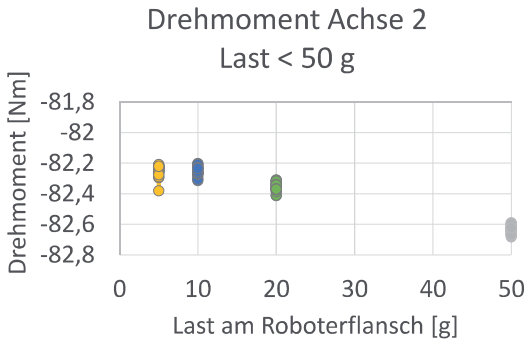


Abb. 7: Drehmoment in Achse 2 bei 25 Messungen mit geringen Lasten

Grundsätzlich ist zwischen den sieben Roboterachsen zu unterscheiden. Achse 2 ist mit einem größeren Drehmomentsensor ausgestattet als beispielsweise Achse 7. Zu beachten ist, dass abhängig von der Pose, ein Hebelarm auf bestimmte Achsen wirkt. Dieser Hebelarm liefert Messdaten, die genutzt werden können. Die Messdaten haben eine zufällige Messabweichung, die durch verschiedene Faktoren beeinflusst werden. Die Messwerte streuen mit einer geringen Standardabweichung um den Mittelwert. Die systematische Messabweichung ist dahingegen gering. Eine lineare Kennlinie bei Messungen von verschiedenen Lasten ist gegeben, jedoch existieren geringe Abweichungen. In den Versuchen konnten Lasten ab 20 g detektiert werden.

Diskussion

Auf Basis dieser Ergebnisse können in der **Improve-Phase** zum einen die Verbesserung der Messergebnisse erfolgen und zum anderen eine konkrete Einordnung dieser Applikation in die Montage erfolgen.

Das Messergebnis kann in einem gewissen Rahmen verbessert werden. Dazu wird das Ishikawa-Diagramm (Abb. 4) aus der Define-Phase verwendet. Die hier beschriebenen Einflüsse können zum Teil reduziert werden, z.B. durch das Filtern der Daten. Ein gleitender Mittelwert ist eine Methode zur Glättung der Datenreihe und somit zur Reduktion der Streuung. Auch durch die Wahl der entsprechenden Roboterachse, z.B. nur Achse 2 oder Achse 2 und Achse 4, kann das Messergebnis verbessert werden.

Nachdem eine Betrachtung der Fähigkeiten des Roboters als Messgerät erfolgte, ist im nächsten Schritt das Konkretisieren der Anwendung in der Montage notwendig. Die Grundidee ist, die In-Prozess-Messung möglichst schlank („lean“), d.h. ohne zusätzliche Betriebsmittel, zu realisieren. Die allgemeinen Aufgaben im Montageprozess sind Handhaben, Fügen, Inbetriebnahme (bspw. durch Funktionsprüfung), Hilfsprozesse (bspw. durch Kontrollieren) und Sonderoperationen (bspw. durch Verpacken) [14]. Durch den Roboter können diese Aufgaben, vor allem der Hilfsprozess „Kontrollieren“, übernommen werden, auch vor dem Hintergrund, dass dieser Prozess in der Montage nicht wertschöpfend ist. Unter Kontrollieren gehört laut der VDI Norm 2860 unter anderem das Prüfen von Gewicht. In Form eines Ist- und Sollabgleiches von der Masse und Massenschwerpunkt (XYZ) können durch den Roboter Abweichungen ermittelt werden. Beispielhaft können bei Montagebeginn durch die integrierte Sensorik die statische Identifikation der Werkzeugdaten eines Robotergreifers erfolgen. Ebenso ist ein Gewichtsabgleich der Bauteile bei variantenhohen Produktionen vorstellbar, um somit eine Verwechselung auszuschließen.
In der **Control-Phase** erfolgen zum einen die Dokumentation der bisherigen Erkenntnisse sowie die Validierung. Dies bedeutet die Untersuchung im konkreten Anwendungsfall an einem Normal und eine vollständige Prüfprozesseignung durchzuführen. Diese Validierung ist Grundlage für die spätere Implementierung in der Montage.

Zusammenfassung und Ausblick

Ziel war es den Roboter als Messmittel zu klassifizieren und herauszufinden mit welcher Genauigkeit dieser eingesetzt werden kann. Mit dem Roboter soll das Prüfen von Gewicht und Massenschwerpunkt eines Objektes in der Montage erfolgen. Das Objekt kann dabei ein Werkzeug, zum Beispiel ein Greifer, oder ein Bauteil sein. Zur strukturierten Vorgehensweise wurden die DMAIC-Phasen (Define, Measure, Analyze, Improve, Control) aus dem Qualitätsmanagement für die Untersuchung des Roboters verwendet. Ausgefüllt wurden die Phasen mit den Ansätzen aus der Prüfprozesseignung. Diese neu entwickelte Methodik dient dazu einen Roboter als Messgerät zu untersuchen und kann ebenfalls auf andere Anwendungsfälle übertragen werden. Mithilfe eines Ishikawa-Diagramms konnten zuvor Einflüsse auf das Messergebnis ermittelt werden. Die Messungen wurden in einer Roboterpose über die gesamte Nutzlast durchgeführt, so dass Wissen über die systematische und die zufällige Messabweichung generiert wurde.

Die Untersuchung mit der entwickelten Vorgehensweise zeigte einen Vertrauensbereich für den Roboter auf. Im nächsten Schritt erfolgt die Validierung der Erkenntnisse an einem konkreten Anwendungsfall. Dazu zählt ein praxistauglicher Messablauf mit optimierter Messdauer, Bestimmung der Roboterachsen als Messwertgeber, Filterung der Daten usw. um schlussendlich die gesamte Messunsicherheit zu bestimmen.

Danksagung

Dieser Beitrag entstand im Rahmen des Forschungsprojektes „messtechnisch gestützte Montage“, welches aus Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) gefördert wird.

Literaturnachweis

[1] VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA), *Fertigungsmesstechnik 2020 - Technologie-Roadmap für die Messtechnik in der industriellen Produktion*, Düsseldorf, (2011).

[2] CP. Keferstein, M. Marxer, C. Bach, *Fertigungsmesstechnik,* Springer Vieweg, Wiesbaden,. 279–298 (2018); doi: 10.1007/978-3-658-17756-0-6

[3] T. Ohno, W. Hof, E. Stotko, *Das Toyota-Produktionssystem: das Standardwerk zur Lean Production,* 3. Auflage, Campus-Verlag, Frankfurt am Main, (2013).

[4] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *Evaluation of measurement data – An Introduction to the „Guide to the expression of uncertainty in measurement” and related documents*, Joint Committee for Guides in Metrology, JCGM 104:2009*,* (2009).

[5] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, *International vocabulary of metrology – Basic and general concepts and associated terms (VIM)*, Joint Committee for Guides in Metrology, JCGM 200:2008, 3. Auflage, (2012).

[6] Verband der deutschen Automobilindustrie e. V. (VDA), *Prüfprozesseignung – Eignung von Messsystemen, Mess- und Prüfprozessen, Erweiterte Messunsicherheit, Konformitätsbewertung*, Band 5, Frankfurt/M., 2. Auflage, (2011).

[7] A.I.A.G. – Chrysler Group LLC, Ford Motor Company, General Motors Corporation, *Measurement System* Anal*ysis (MSA) Reference Manual*, 4. Auflage, Michigan, (2010).

[8] E. Dietrich, *Wo liegen die Unterschiede?: Vergleich von MSA und VDA Band 5*. QZ Qualität und Zuverlässigkeit, 30–34 (2011).

[9] T. Bauernhansl, M. ten Hompel, B. Vogel-Heuser, *Industrie 4.0 in Produktion, Automatisierung und Logistik*: *Anwendungen, Technologien, Migration*, Springer Fachmedien Wiesbaden, (2014); doi: 10.1007/978-3-658-04682-8

[10] KUKA Roboter GmbH*, Sensitive Robotik_LBR iiwa*. https://www.kuka.com/de-de/produkte-leistungen/robotersysteme/industrieroboter/lbr-iiwa, (2016).

[11] ML. George, D. Rowlands, M. Price, J. Maxey, *Das Lean Six Sigma Toolbook: Mehr als 100 Werkzeuge zur Verbesserung der Prozessgeschwindigkeit und –qualität*, 1. Auflage., Verlag Franz Vahlen, (2016).

[12] W. Geiger, W. Kotte, *Handbuch Qualität: Grundlagen und Elemente des Qualitätsmanagements: Systeme, Perspektiven*, 5. Auflage, Vieweg & Teubner Verlag, Wiesbaden, (2008); doi: 10.1007/978-3-8348-9429-8

[13] E. Dietrich, *Leitfaden zum „Fähigkeitsnachweis von Messsystemen“*, Birkenau, Version 2.1, (2002).

[14] G. Pahl, W. Beitz, J. Feldhusen, K-H. Grote, *Konstruktionslehre: Grundlagen erfolgreicher Produktentwicklung: Methoden und Anwendung*, 7. Auflage, Springer Vieweg, Heidelberg, (2007); doi: 10.1007/978-3-540-34061-4

DOI 10.5162/sensoren2019/P3.7

Überprüfung und Bewertung des Schriftbildes sowie des Anlagenzustandes industrieller Signieranlagen mit maschinellem Lernen

Nicole Immesberger[1], *Steffen Klein*[1], *Martin Jochum*[2], *Andreas Schütze*[1]
[1] ZeMA – Zentrum für Mechatronik und Automatisierungstechnik gGmbH
Eschberger Weg 46, 66121 Saarbrücken, Deutschland
[2] Aktien-Gesellschaft der Dillinger Hüttenwerke
Werkstraße 1, 66763 Dillingen/Saar, Deutschland

Zusammenfassung

Ungünstige Prozesssituationen bedingt durch einen fehlerbehafteten Anlagenzustand zeigen sich bei Markieranlagen nicht nur in den Signalen interner Anlagensensorik, sondern auch in Anomalien des Schriftbildes der Markierung. In dieser Arbeit wird daher ein Ansatz zur Schriftbildkontrolle und Anlagenzustandsüberwachung von Blechmarkieranlagen vorgestellt. Auf Basis von Kameraaufnahmen wird das Signierergebnis automatisch gemäß Markiervorgabe auf Vollständigkeit und Qualität beurteilt. Die Aufnahmen durchlaufen dazu eine Signalverarbeitungskette, welche sich aus Schritten der digitalen Bildverarbeitung und Methoden des maschinellen Lernens zusammensetzt. Unerwünschte Effekte können dabei durch gezielte Merkmalsextraktion und -selektion von Bildattributen detektiert und eingeordnet werden. Aus der jeweiligen Klassenzugehörigkeit lassen sich prädiktive Maßnahmen für die Erhaltung eines einwandfreien Anlagenzustandes und optimalen Signierergebnisses ableiten.

Keywords: Bildverarbeitung, Zustandsüberwachung, Maschinelles Lernen, Schriftbildkontrolle, Punkt-Matrix-Schrift

Motivation und Einleitung

Anlagenzustandsüberwachung mit dem Ziel der frühzeitigen Fehlererkennung in Prozessen und Ableitung von Maßnahmen zur bedarfsgerechten Instandhaltung ist von zentraler Bedeutung in der Industrie. Unvorteilhafte Prozesssituationen und für die Maschine u. U. kritische Betriebszustände können durch kontinuierliche Beobachtung von Sensorsignalen sowie daraus resultierenden Handlungsempfehlungen systematisch vermieden und hohe Qualitätsstandards erreicht werden [1]. In der Wertschöpfungskette ist dies speziell auch für Markieranlagen zur individuellen Produktkennzeichnung von Interesse, deren Signatur einerseits die betriebsinterne Rückverfolgbarkeit des Produkts sicherstellt sowie andererseits die eindeutige Identifikation der Warenlieferung und den hochwertigen Markenauftritt beim Endkunden garantiert.

Digitalbilder der Signierung, die z. T. direkt von modernen Markieranlagen archiviert werden, dienen als Basis zur Bewertung des Schriftbildes sowie als Indikator für Fehlerfälle der Anlage. Bei der abgelichteten Markierung (s. Abb. 1) handelt es sich um einen alphanumerischen Punkt-Matrix Aufdruck in weißer Farbe auf einem Grobblech, wobei der Signiertext jedes Bleches einzigartig ist und sich ein Signierblock aus maximal vier Signierzeilen zusammensetzt. Hierbei können negative Einflüsse auf das Signierergebnis, wie beispielsweise Farbmangel oder -überschuss, sowohl lokal stark begrenzt als auch über größere Bildbereiche auftreten, was eine Analyse des Signierergebnisses auf verschiedenen Bildebenen erfordert.

Methodik

Zur Detektion unerwünschter Effekte werden aus den Bildern daher zuerst sukzessive kleinere Bildausschnitte isoliert und daraus u. a. statistische und formbeschreibende Bildmerkmale extrahiert. Da jede Signatur individuell ist, wird eine für das maschinelle Lernen nutzbare Datenbasis aus den wiederkehrenden Schriftzeichen aggregiert. Die Einzelzeichen werden dafür über ein mehrstufiges Segmentierverfahren extrahiert. Durch Kenntnis der Textvorgabe und der herstellerspezifischen Schriftart können darüber hinaus Bildmerkmale auf Signierzeichen- oder -zeilenebene synthetisch generiert werden. Durch die modellbasierte Vorhersage und Differenzbildung von Ortsreihen können Bildattribute, wie z. B. Farb-

homogenität oder Positionstreue der Signierpunkte, objektiv bemessen werden und die Datenbasis gezielt mit für das maschinelle Lernen relevanten Mustern ergänzt werden.
Die weitere Verarbeitung der Bildsignale erfolgt in den für das maschinelle Lernen typischen Schritten der Merkmalsextraktion, Merkmalsselektion, Modellbildung (Training) und Modellvalidierung (Test). Großes Potenzial liegt hierbei in der signaturunabhängigen Feststellung und Lokalisierung sowohl diskreter Farbeffekte als auch Farbgradienten des Schriftbildes, die von Oberflächendefekten und Umgebungseinflüssen abzugrenzen sind. Entsprechend des Klassifikationsergebnisses und der Anlagensystematik können automatisch typische Ursachen der auftretenden Bildfehler diagnostiziert und Markierungen, die nicht den Anforderungen entsprechen, detektiert werden.

Markieranlage

Bei der betrachteten industriellen Signieranlage handelt es sich um einen Portal-Aufbau in der Linie oberhalb der Fördereinrichtung der Produkte, was eine Farbsignierung der Bleche im Produktionsfluss gewährleistet. Die Markiereinheit der Anlage befindet sich auf Schlitten an einer Traverse und ist quer zur Transportrichtung des Bleches verfahrbar. Die insgesamt vier Markierköpfe der Markiereinheit können jeweils einer Signierzeile zugeordnet werden. Die Digitalkamera zur Aufnahme der Fotos ist starr an der Markiereinheit der Anlage befestigt und senkt sich mit der Markiereinheit auf eine feste Höhe über die Oberfläche ab. Während des Blechtransports werden mehrere Einzelbilder im Signierprozess angefertigt, die ausschnittsweise die Signatur zeigen und softwaregestützt zur Abbildung der vollständigen Signatur zusammengesetzt werden. Die verfügbare Bilddatenbasis enthält die Archivaufnahmen der Signaturen.

Blechfarbmarkierung

Die im Fokus der Untersuchung stehende Signatur dient der farblichen Oberflächenmarkierung von Blechen. Die weiße Signatur erstreckt sich über bis zu drei in Leserichtung aufeinanderfolgende Blöcke mit maximal vier Zeilen, die jeweils bis zu 52 Zeichen umfassen können. Der zur Verfügung stehende Zeichensatz ist alphanumerisch und setzt sich aus Groß- und Kleinbuchstaben des lateinischen Alphabets und Sonderzeichen zusammen. Es handelt sich um eine herstellerspezifische Schriftart, wobei ein einzelnes Zeichen in einer Punktmatrix des Formats 10 x 6 (bzw. 10 x 11 Halbe) abgebildet wird. Die tatsächliche Zeichenhöhe und -breite ergibt sich aus der Anordnung der Punkte zur Darstellung des Zeichens. Der Abstand der Punkte in der Ebene ist innerhalb des Zeichens konstant. Das Druckraster sieht eine äquidistante Anordnung der Zeichen für eine Signatur vor, d. h. die Zeichen werden zeilenweise gleichmäßig untereinander ausgerichtet. Die Breite eines einzelnen Blockes resultiert aus dessen längster Signaturzeile. Die Größe der Zwischenräume der Signierblöcke ist konstant. Die Markierung der Bleche dient zur eindeutigen Identifikation, weshalb jede Signatur einzigartig ist und folglich nur einmal in der Datenbank vorkommt.

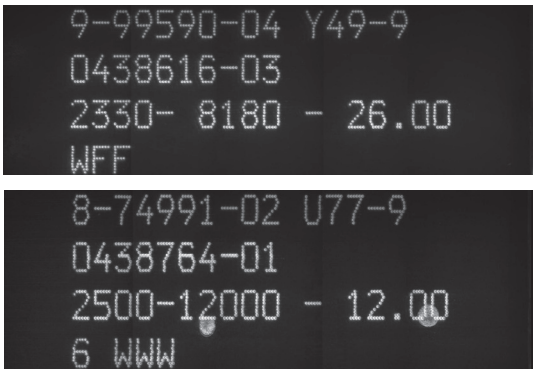


Abb. 1: Beispielhafte Kameraaufnahmen der Blechmarkierung; oben: Signatur ohne Auffälligkeiten, unten: Signatur mit Störeffekten (s. Zeile 3)

Bildverarbeitung

Laut [3] ist „die industrielle Bildverarbeitung angesichts steigender Anforderungen an Produktqualität und -dokumentation zu einer Schlüsseltechnologie geworden“. Die digitale Bildverarbeitung von Bilddaten stellt typischerweise eine Kette von Bildoperationen dar, die auf den jeweiligen Anwendungsfall spezialisiert ist. Die hier zur Vorbereitung für das maschinelle Lernen ausgewählten Schritte der Bildverarbeitung wurden in der *MATLAB* Entwicklungsumgebung implementiert und sind nachfolgend kurz genannt. Allgemein werden zur Oberflächeninspektion der Produktmarkierung mehrere Bildvorverarbeitungsschritte zur Erhöhung der Zuverlässigkeit und Stabilität bei der Objektlokalisation und -qualitätsbestimmung durchgeführt, wie sie auch für die Erkennung von Barcodes in [3] beschrieben wird. Zuerst wird eine Bildglättung vorgenommen, welche die Struktur des Bildes weichzeichnet und die Bildverarbeitung in der weiteren Abfolge unempfindlicher gegenüber Bildrauschen macht. Weiterhin kann durch eine Normierung des Bildes in Form einer Histogrammspreizung und -verschiebung eine Kontrastverbesserung erreicht werden. Der hierbei verwendete Normierungsfaktor, ein Maß für die Bildhelligkeit, der potenziell

relevante Information für die spätere Klassifikation der Bilddaten enthält, wird als Einzelmerkmal gespeichert.

Zeichensegmentierung

Prinzipiell werden für das maschinelle Lernen Datenbestände benötigt, die für je eine Klasse ähnliche Merkmale aufweisen und eine Mustererkennung der Daten zulassen, sodass eine Abgrenzung zwischen den Klassen vollzogen werden kann. Für die hier betrachtete Anlagenzustandsüberwachung bedeutet dies, dass für die verschiedenen Prozesssituationen ausreichend Bildaufnahmen vorliegen müssen. Da allerdings jede der Markierungen individuell ist und inhaltlich, d. h. auch optisch, stark variiert, ist eine Verallgemeinerung der Bilddaten ungeeignet. Demzufolge ist auf Basis der Signaturaufnahmen vorerst keine Trendanalyse sinnvoll. Um dennoch auf einen hinreichenden Probenumfang für das Training beim maschinellen Lernen zurückgreifen zu können, werden die Einzelzeichen der Signaturen extrahiert und verglichen. Durch Betrachtung der Bildausschnitte der Signierzeichen bleibt der Eindruck der Schriftbildqualität erhalten und die zur Verfügung stehende Datenbasis wird vervielfältigt und abstrahiert.

Segmentierung

Für die Segmentierung der Zeichen wird die Information zur Lage des Objekts innerhalb der Bildebene gesucht. Da die gedruckte Markierung grundsätzlich einer festen Gestaltungsvorgabe unterliegt und die Bildauflösung konstant ist, kann bereits ein grober Umriss des Interessensbereichs, der die Signatur beinhaltet, vorgenommen werden. Die anschließende Bildvorverarbeitung dient der Vereinfachung der Trennbarkeit von Vordergrund (Signatur) und Hintergrund (Blechoberfläche). Diesbezüglich werden u. a. morphologische Bildoperationen zur Verstärkung und Zusammenführung der Konturen der Signierpunkte vorgenommen, um einzelne Zeichen als zusammenhängend zu identifizieren. Gemäß der horizontalen Ausrichtung der Signatur kann im nächsten Schritt durch zeilenweise Integration der Pixelwerte und statische Schwellwertsetzung die Anzahl und Intervalle der Zeilenindizes ermittelt werden (s. Abb. 2) Die Länge eines Intervalls entspricht dabei der jeweiligen Höhe der Signaturzeile bzw. der dazugehörigen Einzelzeichen. Es wird weiterhin kontrolliert, ob das Pixelmaß der Zeilenhöhen und der vertikalen Zeilenabständen innerhalb des plausiblen Toleranzbereichs liegt, da die detektierte Höhe einer Signaturzeile typischerweise leicht schwanken kann.

Um zur besseren Vergleichbarkeit der Zeichen nun Bildsegmente gleicher Größe aus dem Gesamtbild zu isolieren, wird nur der Zeilenindex der positiven Flanke, d. h. des theoretisch ersten Bildpixels, das die Signierzeile abbildet, übernommen und eine feste Intervalllänge angenommen. Angesicht der linksbündigen Ausrichtung und dem festen Matrix-Format der Signierzeichen kann ebenso in Spaltenrichtung zur Bestimmung der horizontalen Zeichenlage und -breite verfahren werden.

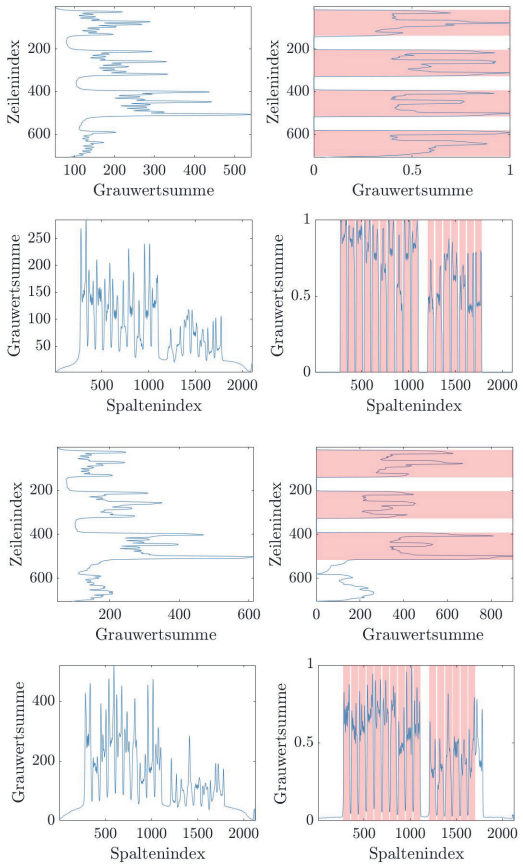


Abb. 2: Zeilen- und Spaltensignale (links) sowie deren Anpassung (rechts) zur Segmentierung der Signierblöcke in Einzelzeichen;
oben: erstes Bild aus Abb. 1, unten: zweites Bild aus Abb. 1

Jedes der dadurch gewonnenen Wertepaare mit den Vektoren der Zeilen- und Spaltenindizes impliziert eine potenzielle Position eines Bildabschnitts mit einem Signierzeichen. Da jedoch die Spalteninformation komprimiert über alle Zeilen erworben wird, d. h. eventuell nur von einem einzigen Zeichen in dieser Signierspalte stammen kann, ist nicht gegeben, dass alle potenziellen Bildsegmente mit Zeichen besetzt sind. Um dies zu überprüfen, wird für die weitere Verwendung die jeweilige Zeichenposition mit dem vorliegenden Solltext verglichen und für den Fall einer Leerstelle ver-

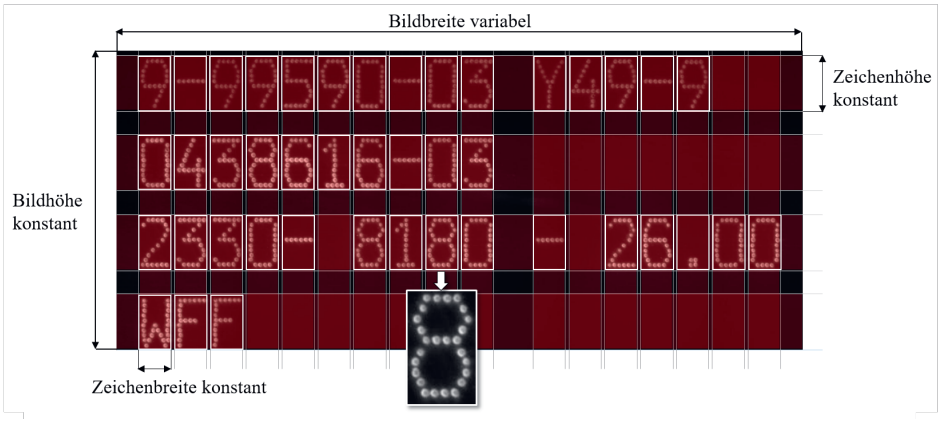


Abb. 3: Grafische Darstellung der Segmentierung von Einzelzeichen eines Signierblockes im Interessensbereich mit Kennzeichnung der Geometrieverhältnisse

worfen. Mittels der verbleibenden Lagebeschreibungen kann der relevante Bildabschnitt erschlossen und separat als Bilddatei zu einer Datenbasis hinzugefügt werden (s. Abb. 3).
Neben der Bildverarbeitung zur Objektlokalisation und Bereitstellung der Bildsegmente für das maschinelle Lernen dient der Algorithmus zur Extraktion der Einzelzeichen gleichzeitig auch als erster Indikator für die Qualität des Schriftbildes der Signatur. Da die Segmentiervorschrift an den Aufbau einer Standard-Signatur angepasst ist, verläuft die Routine nur bei erwartungsgemäßen Signaturen erfolgreich. Falls das Signaturlayout nicht dem hinterlegten Profil entspricht, d. h. Signierzeilen mit der Programmlogik nicht unmittelbar erkannt werden oder Abstände nicht der Plausibilitätsprüfung genügen, liegt offensichtlich eine Anomalie vor, die eine Qualitätseinbuße bedeutet (vgl. Abb. 2 unten). Dieses Abbruchkriterium wird als Indiz festgehalten und in Abhängigkeit dessen die möglichen Ursachen für die Abweichung bestimmt.

Erstellung einer Referenzdatenbasis

In dieser Arbeit wird ein neuartiger Ansatz zur künstlichen Generierung von Trainingsdaten verfolgt. Es wird überprüft, inwiefern sich synthetische Bilddaten, welche verschiedene Prozesslagen der Maschine möglichst optimal nachbilden, zur Kategorisierung der Archivaufnahmen eignen. Im Rahmen der Bildverarbeitung wird zu diesem Zweck ein Modell entwickelt, das entsprechend dem Muster der vorliegenden Signaturbilder möglichst realitätsnah synthetische Signaturbilder generiert. In Anlehnung an die typografische Gestaltung, das Farbschema und den geometrischen Aufbau der Blechmarkierung kann sukzessive das Einzelzeichen und weiterführend die Signaturzeile bzw. Gesamtsignatur künstlich erzeugt und zusammengefügt werden.

Modellbilder

Methodisch wird zuerst eine zweidimensionale Matrix erstellt, deren Größe sich an der Auflösung der Zeichensegmente der Archivbilder orientiert und deren Elemente den Grauwert für den Hintergrundfarbton der Signatur enthalten. Anschließend wird im Abgleich mit den vorliegenden Fotobeispielen das Punktraster ermittelt, welches die möglichen Positionen der Kreismittelpunkte für die Zeichen in Punkt-Matrix-Schriftart wiedergibt. Die benötigte Pixelanzahl für den Durchmesser wird aus den Fotobeispielen bestimmt. Die Farbverteilung der einzelnen Signierpunkte wird durch die zweidimensionale Projektion einer Halbkugel abgeleitet, deren Höhenausprägung entlang der rotationssymmetrischen Achse den Helligkeitswert angibt. Gemäß der bekannten Anzahl und Anordnung der Signierpunkte pro Zeichen werden diese iterativ im Bildraster ausgerichtet und an die passende Stelle projiziert, wo die Grauwerte der Bildmaske des Signierpunktes zu den relevanten Matrixeinträgen der Hintergrundmaske addiert werden. Das auf diese Weise modellierte Bild wird abschließend mit einem Gauß-Filter weichgezeichnet, sodass harte Konturen geglättet werden und das Bild der realen Fotoaufnahme ähnlicher erscheint. Beispielbilder sind im Abschnitt Zusammenstellung der Klassen aufgeführt.

Manipulation der Modellbilder

Die Standardausführung der Modellbilder kann nun in der Folge so manipuliert werden, dass die synthetischen Daten Signierfehler imitieren, die verschiedene Degradationsstufen der Schriftbildqualität repräsentieren. Die Bilddefekte sollen wiederum den entsprechenden kritischen Prozesssituationen zugeordnet werden können, wodurch von der Signaturqualität auf den Anlagenzustand geschlossen

werden kann. Das heißt, dass durch gezielte Variation der Modellbilder nach Signaturvorbild mangelbehafteter Exemplare, die aus seiner bestimmten Anlagenstörung resultieren, eine entsprechende Referenzdatenbasis für die Fehlerklassen generiert werden kann. Zu den skalierbaren Modellparametern der Signierpunkte zählen:

- Helligkeit
- Schattierung
- Größe
- Exzentrizität
- Homogenität
- Kontrast
- Positionstreue.

Die Helligkeit, welche der Intensität der Signierfarbe bzw. deren Reflexion gleich kommt, kann durch Vorgabe des Maximalwertes für die Bildmaske des Signierpunktes angepasst werden. Das Strukturelement für die Nachbildung des Signierpunktes, hier als Schattierung bezeichnet, kann neben der genannten Halbkugel auch einer anderen Funktion, z. B. einer Gauß-Verteilung, folgen. Die Größe eines projizierten Signierpunktes wird durch Angabe des gewünschten Radius in Pixel eingestellt. Das Format der Bildmaske muss dabei aber nicht quadratisch sein, sondern kann auch in Zeilenrichtung verzerrt sein, was einer Erhöhung der Exzentrizität, d. h. der Abbildung einer Ellipse, gleichkommt. Diese Deformation der Kreisgeometrie wird z. B. durch morphologische Dilatation erreicht. Die Homogenität eines modellierten Signierpunktes, die zum einen durch die Farbhöhen und zum anderen durch die lokale Betrachtung der Pixelumgebung bestimmt wird, kann durch die Überlagerung mit Rauschsignalen verändert werden. Der Kontrast der Signierpunkte, d. h. die Stärke des Gradienten am Rand des Objektes, welche maßgeblich für den Schärfeeindruck eines Bildes ist, kann durch die Operation des Weichzeichnens gemildert werden. Auch das Angleichen der Hintergrundhelligkeit verändert dieses Bildattribut. Zur Beeinflussung der Positionstreue der Punkte im ermittelten Raster der Punkt-Matrix-Schriftart kann eine Anpassung der Zeilen- und Spaltenindizes erreicht werden. Die Änderung des Kreisursprungs resultiert in einer Abweichung der Lage des modellierten Signierpunkts in der Zeichenebene.
Die beschriebenen Bildmanipulationen können je nach gewünschter Art des Bilddefekts einzeln oder in beliebiger Kombination durchgeführt werden. Die Ausprägung der Defekte kann innerhalb eines Toleranzbereichs zufällig erfolgen oder auch abgestimmt auf einen speziellen Fehlerfall.

Bildmerkmale

„Das Auge erfaßt Bilder eher qualitativ, während die quantitative Extraktion von Bildmerkmalen eine quantitative Analyse der Filtereigenschaften erfordert.“ heißt es in [4], was die Subjektivität des Qualitätseindrucks eines Bildes widerspiegelt und die Sorgfalt bei der Auswahl von Bildoperationen zur Merkmalsextraktion betont. Ziel bei der Dimensionsreduktion der Bilddaten ist es, charakteristische Bildeigenschaften möglichst kompakt darzustellen und gleichzeitig wenig der für die Beurteilung der Qualität relevanten Information über Filteranwendungen zu verlieren. Da die Bestimmung und Interpretation der aussagekräftigen Bildattribute jedoch Gegenstand der Forschungsarbeit ist, gilt es im ersten Iterationszyklus, möglichst viele Bildmerkmale zu bestimmen, die weiterführend analysiert und selektiert werden können.

Tab. 1: Einteilung der Bildmerkmale

Merkmalstyp	Beispiele
Spaltenmerkmal	Summen der Grauwertpixel/ Binärpixel/ Kantenpixel, Mittelwert, Varianz
Zeilenmerkmal	
Histogrammmerkmal	Bild vor/ nach Kontrastanpassung, Differenzbild aus Kontrastanpassung
Objektmerkmal	Mittelpunkt/ Form/ Fläche des Objekts
Kreismerkmal	Radius, Mittelpunkt, Sensitivität der Kreisdetektion
Sekundärmerkmal	Faktor der Kontrastanpassung, Anzahl der Objekte/ Kreise

Bei der durchgeführten Merkmalsextraktion werden entsprechend der Herkunft der Information bislang sechs verschiedene Merkmalstypen unterschieden, wie in Tabelle 1 gelistet. Spalten- bzw. Zeilenmerkmale beziehen sich auf die Verarbeitung der Ortsreihen entlang der jeweiligen Richtung. Diese werden u. a. durch statistische, formbeschreibende und spektrale Signalanalyse gewonnen und sind für beide Dimensionen äquivalent durchführbar. Im Zuge dessen können sowohl absolute als auch relative, d. h. normierte, Merkmale für die Auswertung herangezogen werden. Diese können wiederkehrend nach verschiedenen lokalen und globalen Bildoperationen für die aktuelle Ausführung des Bildes bestimmt werden, was das Angebot an potentiell wichtigen Merkmalen steigert. Gleiches gilt auch für

die Einbindung der Histogramme als densitometrische Merkmale für verschiedene Versionen des Bildes, die Aufschluss über Farbhäufungen innerhalb eines Bildes bzw. eines Teilsegments geben. Diese können chromatisch sortiert sowie kumuliert dargestellt werden. Durch beliebige Skalierung der Klassenbreite kann die Anzahl an Einträgen des Merkmalsvektors variiert werden.
Weiterhin werden Objekt- und Kreismerkmale bestimmt, die sich auf eine definierte Region beziehen. Als Objekt wird in diesem Kontext eine zusammenhängende Pixelnachbarschaft verstanden, die nach einer mehrstufigen Vorverarbeitung ein Schwellwertkriterium erfüllen. Im vorliegenden Fall der Signaturzeichen handelt es sich dabei i. d. R. um unidirektional verknüpfte Signierpunkte. Die Gestalt und Ausrichtung eines Objektes sowie der Objekte zueinander, ebenso wie u. a. die ursprünglichen Werte der assoziierten Bildpixel, bildet eine Merkmalskategorie zur Erfassung der Topologie und Struktur des Bildes ab. Weitere Merkmalstypen, die zur Veranschaulichung von Lage, Geometrie und Kolorit der Signierpunkte dienen, sind die Kreismerkmale. Da es sich bei der vorliegenden Schriftart der gedruckten Zeichen um eine Punkt-Matrix-Schrift handelt, ist demzufolge eine Isolation der runden Zeichenbausteine zur Beurteilung ihrer Ähnlichkeit naheliegend. Da in Bezug auf den Signierprozess die Punktzeilen jeweils einer eigenen Markierdüse zugeordnet werden können, d. h. im Falle von Unstimmigkeiten die wahrscheinliche Ursache eingegrenzt werden kann, ist diese Eigenschaft im Hinblick auf den Anlagenzustand von besonderem Interesse. Die Routine zur Identifikation der runden bzw. kreisähnlichen Bilderhöhungen ist der proprietären *Image Processing Toolbox* entnommen. Zur effizienten und robusten Detektion der Signierpunkte sind Vorverarbeitungsschritte, welche den Algorithmus der Kreiserkennung begünstigen, erforderlich. In einer Programmschleife werden zu diesem Zweck Kontrast- und Konturanpassungen vorgenommen, bis die gegebene Maximalzahl von Signierpunkten erreicht ist. Sind die idealen Umrisse der Signierpunkte ermittelt, kann eine Auswertung der einbeschriebenen Pixel im Originalbild (s. Abb. 4) und weiteren Bildausführungen erfolgen. Dabei impliziert die Abweichung der Radien und Koordinaten z. B. ein Maß für die Prozessstreuung respektive die Positionstreue der Signierpunkte.
Weiterhin werden im Zuge der Bildverarbeitung stellenweise Einzelwerte gesammelt, die in der Gruppe der Sekundärmerkmale zusammengefasst werden, und sich als zusätzliche Deskriptoren für die Bilder eignen. Zu diesen Metadaten zählen beispielsweise Angaben über die Anzahl der von Null verschiedenen Einträge eines Merkmalsvektors, das Integral einer Bildregion oder Filterparameter.

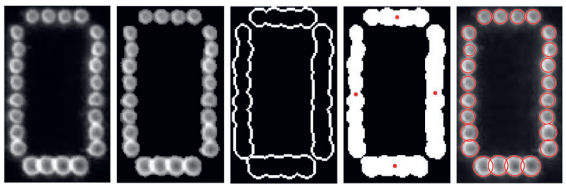


Abb. 4: Bildversionen der Bildverarbeitungskette: v.l.n.r. (1) Originalbild, (2) Freistellung des Vordergrundes, (3) Kantenbild, (4) Binärbild mit Kennzeichnung der Objektmittelpunkte, (5) Originalbild mit Projektion der detektierten Kreise

Die aufgeführten Bildmerkmale, als kompakte und charakteristische Repräsentation der Aufnahmen der Signaturzeichen, werden anschließend mit Methoden des maschinellen Lernens auf Trennbarkeit untersucht.

Maschinelles Lernen mit Bilddaten

Die Überführung der Bildinformation in unterscheidbare Merkmale dient zum einen der Reduktion des hochdimensionalen Merkmalsraums von Fotoaufnahmen und zum anderen der Projektion von Gemeinsamkeiten zur Aggregation für das maschinelle Lernen. Die aus dem Produktionsprozess verfügbaren Bilddaten der Signatur werden durch Kombination der genannten Bildverarbeitungsschritte dem maschinellen Lernen zugänglich gemacht.

Zusammenstellung der Klassen

Zunächst wird durch manuelle Vorauswahl von Bildexemplaren, die subjektiv und erfahrungsgemäß einer guten bzw. einer schlechten Signierbildqualität entsprechen, eine Datenbasis zur Erkennung von einwandfreien und kritischen Bildern zusammengestellt. Da die Information zur Wertigkeit des Markierergebnisses zwar zeichenunabhängig ist, sich die Bildinformation der Zeichen aber aufgrund der Zeichengestalt stark unterscheidet und somit keine Ähnlichkeit innerhalb der Qualitätsklassen gegeben ist, werden für das maschinelle Lernen die Signaturzeichen getrennt betrachtet. Dadurch sollen die fehlerhafte Interpretation und das Training auf systematische Zeichenunterschiede unterbunden werden. Da jedoch für das maschinelle Lernen eine möglichst große Datenbasis ideal, die manuelle Selektion aber sehr zeitaufwändig ist, soll nach dem Muster der Archivdaten eine modellbasierte Referenzdatenbasis generiert werden, die eine automatische Klassifizierung

der Zeichensegmente ermöglicht. Dazu werden zunächst Modellbilder anhand der Parameter, wie im Abschnitt Manipulation der Modellbilder beschrieben, zur Imitation der Kategorie der guten Beispiele („i. O.“) und der schlechten Beispiele („n. i. O.“) angepasst. Die Abbildungen 5 und 6 zeigen eine Teilmenge der jeweiligen Kategorien.

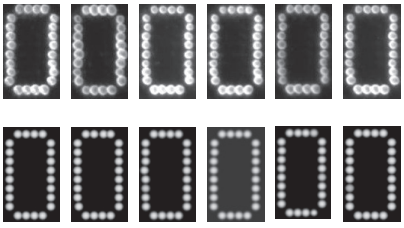


Abb. 5: Bildbeispiele der Klasse „i. O.”; oben: Originalbilder aus dem Archiv, unten: Modellbilder aus der synthetischen Datenbasis

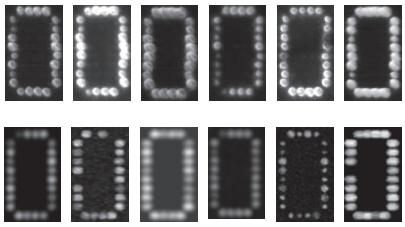


Abb. 6: Bildbeispiele der Klasse „n. i. O.”; oben: Originalbilder aus dem Archiv, unten: Modellbilder aus der synthetischen Datenbasis

Multivariate Datenanalyse

Wie u.a. in [5] erwähnt, ist für das Konzept des maschinellen Lernens bekannt, „dass die Wahl des Verfahrens immer anhand empirischer Untersuchungen und den technischen Anforderungen […] der Aufgabenstellung getroffen werden muss”. Vor diesem Hintergrund wurden mithilfe der in [2] beschriebenen Methoden Vergleichsstudien zu den für die Klassentrennung bedeutsamen Merkmalen und deren Einfluss auf das Klassifikationsergebnis durchgeführt. Für das vorliegende Szenario erfolgt das maschinelle Lernen überwacht, d. h. die Klassenzugehörigkeit der Bilder ist gegeben. Wie im Abschnitt Zusammenstellung der Klassen erläutert, gibt es testweise zwei heterogene Klassen „i. O.“ und „n. i. O.“, die aus derselben Anzahl (je 100 Exemplare) an Original- und klassenspezifischen Modellbildern bestehen.

Das für das vorliegende Szenario beste Ergebnis zur Trennung der Klassen wurde mit dem zweistufigen Konzept der Merkmalsselektion mittels Pearson Korrelation und Recursive Feature Elimination Support Vector Machines (RFESVM) erreicht. Dabei wird aus Effizienzgründen zuerst eine Vorauswahl von 500 Merkmalen getroffen [1] und diese anschließend hinsichtlich ihrer Eignung zur Trennung der Klassen schrittweise verworfen. Dabei wird die Redundanz von Merkmalen berücksichtigt. Details des Algorithmus sind [6] zu entnehmen.

Die minimale Fehlerrate bei 10-facher Kreuzvalidierung und Klassifikation mit dem k-Nächste-Nachbarn-Verfahren (k = 1) beträgt 5,75 % bei Betrachtung der zwölf besten Merkmale. Der Verlauf des Klassifikationsfehlers als Funktion der Merkmalsanzahl ist in Abb. 7 gezeigt. Zu den ausgewählten Merkmalen zählen Bildattribute aus den Kategorien Zeilen-, Spalten-, Histogramm-, Objekt- und Kreismerkmal, siehe Tabelle 1.

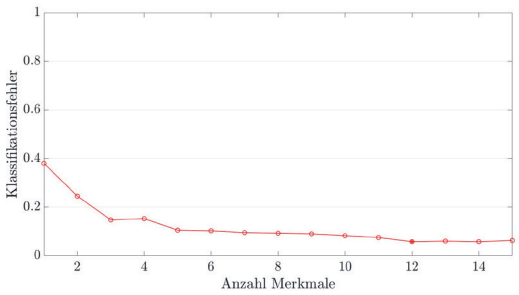


Abb. 7: Darstellung des Klassifikationsfehlers bei Klassifikation mit 1NN-Methode

Das Projektionsergebnis einer linearen Diskriminanzanalyse (LDA) basierend auf den zwölf besten Merkmalen ist in Abb. 8 dargestellt. Das Ergebnis zeigt, dass die beiden Gruppen „i. O.“ und „n. i. O.“ sowohl für Modell- als auch Originalbilder im Mittel gut getrennt werden. Auffällig ist, dass die „i. O.“ Modellbilder eine deutlich geringere Variation des Diskriminanzwertes zeigen als die drei anderen Untergruppen.

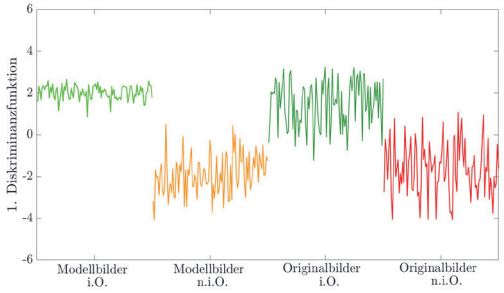


Abb. 8: Darstellung der Diskriminanzfunktion für die besten 12 Merkmale nach Vorselektion mittels Pearson Korrelation und Merkmalsselektion mittels RFESVM für die Bildgruppen der Klassen „i. O.” und „n. i. O.”

Diskussion und Ausblick

In Abbildung 8 ist zu erkennen, dass die Variablen der Originaldaten der Klassen z. T. treffend durch die Modellbilder nachempfunden werden. Jedoch wird deutlich, dass keine eindeutige Trennbarkeit erreicht wird und die Variation der Modellbilder insbesondere für die Klasse „i. O." noch nicht hinreichend dem Ideal der Originalbilder entspricht. Daher sollen die Modellbilder der Referenzdatenbasis hinsichtlich Authentizität und Repräsentationsvermögen weiter optimiert und zusätzlich die Extraktion der Bildmerkmale erweitert werden, um den Klassifikationsfehler zu minimieren.
Darüber hinaus handelt es sich bei der Modellierung der Aufgabenstellung als 2-Klassen-Problem bis dato um eine grobe, unzulängliche Formulierung der beobachteten Qualitätsabstufungen, welche mit den verwendeten Verfahren der Merkmalsselektion nur bedingt erfasst werden kann. Dies soll durch eine feinere Unterteilung der kritischen Bilderklasse „n. i. O." verbessert werden. Eine Klassenaufspaltung und Ableitung von Handlungsempfehlungen gemäß der Systemursache der Anomalien erfordert dementsprechend eine Gruppierung der Bilddefekte nach Störquellen.

Fazit

Die bisherige Forschungsarbeit zeigt ein Konzept zur Beurteilung der Schriftbildqualität von Blechmarkierungen in Punkt-Matrix-Schrift und erläutert die Modellbildung von Referenzdaten zur Erweiterung der verfügbaren Datenbasis für das Training im maschinellen Lernen. Nach einer mehrstufigen Merkmalsextraktion und -selektion wird mittels Nächste-Nachbarn-Klassifikation ein minimaler Klassifikationsfehler von 5,75 % erzielt. Aufbauend auf diesen Ergebnissen laufen aktuell Arbeiten zur anlagenzustandsadäquaten Erweiterung der Klassen und zur realitätsnahen Optimierung der Modellbilder.

Literaturnachweis

[1] VDMA, *Future Manufacturing: Magazin für intelligente Produktion*, VDMA Verlag (2017/03)

[2] T. Schneider, N. Helwig, A. Schütze, *Industrial condition monitoring with smart sensors using automated feature extraction and selection*, IOP Meas. Sci. Technol. (2018) 29 094002, doi: 10.1088/1361-6501/aad1d4

[3] C. Demant, B. Streicher-Abel, A. Springhoff, *Industrielle Bildverarbeitung – Wie optische Qualitätskontrolle wirklich funktioniert*, Springer Berlin Heidelberg (2011), doi: 10.1007/978-3-642-13097-7

[4] B. Jähne, *Digitale Bildverarbeitung und Bildgewinnung*, Springer Vieweg (2012), doi: 0.1007/978-3-642-04952-1

[5] H. Süße, E. Rodner, *Bildverarbeitung und Objekterkennung – Computer Vision in Industrie und Medizin*, Springer Vieweg (2014), doi: 10.1007/978-3-8348-2606-0

[6] I. Guyon, M. Nikravesh, S. Gunn, L. A. Zadeh, *Feature Extraction – Foundations and Applications*, Springer Berlin Heidelberg (2006), doi: 10.1007/978-3-540-35488-8

DOI 10.5162/sensoren2019/P3.8

Hardware-in-the-loop Test Concept for an Energy-optimized Process Control

Tobias Kull[1], Bernd Zeilmann[2], Gerhard Fischerauer[1]
[1]Chair of Measurement and Control Systems and Center of Energy Technology (ZET), University of Bayreuth, 95447 Bayreuth, Germany
[2]Richter R&W Steuerungstechnik GmbH, Körzendorf 52, 95491 Ahorntal, Germany
mrt@uni-bayreuth.de

Abstract

In this work, a hardware-in-the-loop test concept for the evaluation of energy-optimized process control is presented. Modern production facilities often contain distributed renewable energy systems and controllable loads. Coordinating the operation of these generators and loads can bring benefits in that the overall cost is reduced by own consumption or flexibility is provided to the distribution grid. However, these operation strategies are not yet state of the art, especially not in small and medium-sized enterprises. The model-based hardware-in-the-loop test concept that is presented in this work, supports the development of appropriate control technology for the automatisation of energy systems and processes. In this manner, control algorithms can be simulated in real-time, while being implemented on field devices including field communication. Functionalities and performance of the automation solution can be evaluated and flaws can be detected.

Keywords: hardware-in-the-loop, process control, virtual comissioning, renewable energy system

Introduction

It has become evident that future energy systems will have to cope with fluctuations in the energy supplied by regenerative sources. This could lead to temporary local congestions or overcapacities in the power distribution grid. Transmission and distribution network operators in Germany perform feed-in management to avoid congestions, curtailing the power input to the grid. Curtailment is basically undesirable as the plant owner has to be compensated for the deficit energy and renewable ressources remain unused. In 2017 the amount of curtailed energy in Germany summed up to 5,518 GWh [1]. This number will increase with a higher penetration of wind and solar energy systems in the future.

Another way to stabilize the system is to design decentralized subsystems capable of adapting their feed-in and power consumption from the grid to the grid conditions. A key benefit of such energy subsystems is the flexibility that they offer to the distribution system operator (DSO) without the need of curtailment [2].

Modern production facilities often contain distributed renewable energy systems as well as loads that can be subject to demand side management, especially in small and medium-sized enterprises (SMEs) and municipal facilities. In order to achieve the desired flexibility, these energy subsystems have to take into account the nature of the power output from renewable energy sources: quasi-deterministic (predictable to a certain amount) from photovoltaics (PV), controllable with constraints from combined heat and power (CHP) plants, and stochastic to quasi-deterministic from wind power. They also have to take into account constraints arising from different load types: continuous processes, discrete processes (e. g., purely on-off), essential loads, which cannot be switched off, less highly prioritized loads, which can be shut down at least for certain time periods, etc. The core task consists in matching the energy supply and demand by performing a demand-side scheduling and power dispatch within the decentralized energy subsystem. This may require local electrical energy storage units, such as batteries, because they serve both as load and as source.

The implementation of an energy subsystem as described above requires a local “intelligence”, i. e., a controller. This controller receives measurement, status or prediction data from heterogeneous information sources via different interfaces and in different formats, processes them and outputs commands to the controlled units in the system. Many sophisti-

cated proposals to solve this allocation problem using different methods can be found in the literature, e. g., multi-objective optimization [3] or rolling-horizon control [4]. Typically, these solutions are only validated by simulation. However, novel control algorithms cannot be implemented in the field without prior testing and characterization because untested control algorithms would pose a safety and security risk for equipment, processes, and environment. This is why we developed a hardware-in-the-loop (HiL) test concept for energy system control validation on field hardware including field communication.

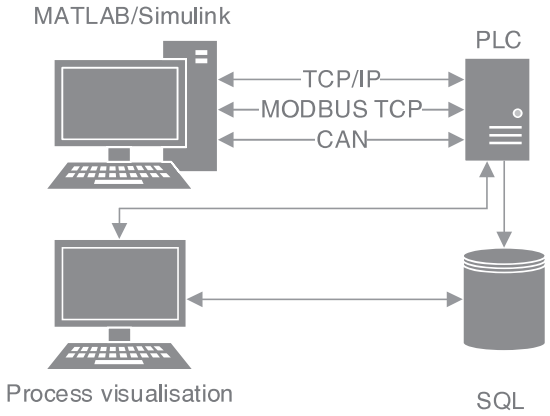


Fig. 1: Functional diagram of the HiL testbed

Methods

To test and characterize the described automation solution for energy-optimized operation we created an HiL testbed. It consists of the controller hardware, field-communication peripherals, and a model of the controlled system. The functional diagram of the testbed is depicted in Fig. 1. The controller hardware is implemented as a programmable logic controller (PLC; Wago PFC200) because this is the state of the art for industrial process control. It is extended by an SQL database for data recording and a human machine interface (HMI) in form of process visualization.

The controlled system with its many subsystems is modelled in Matlab/Simulink (Simscape Power Systems) on a desktop computer. The model environment communicates with the PLC via Modbus/TCP, CAN, and TCP/IP. The communication protocols are dictated by typical real-world environments in which the PLC is to operate.

The testbed is suited for controllers operating heterogeneous energy systems involving AC and DC energy sources, loads, and processes in an energy-optimized manner.

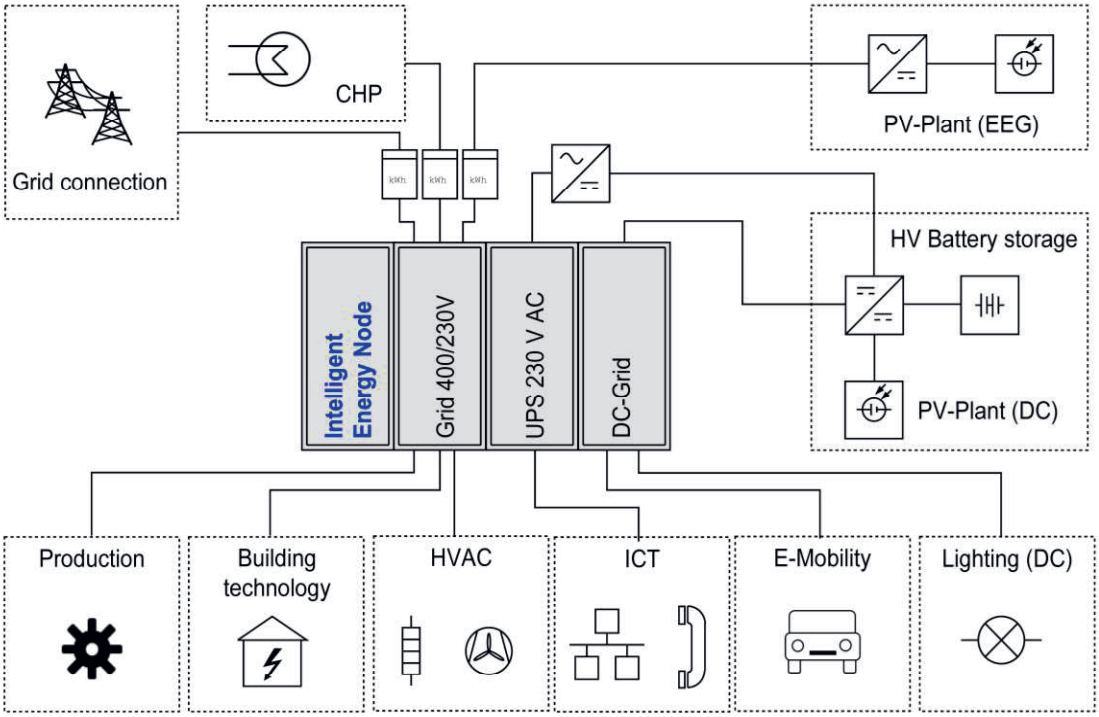


Fig. 2: Block diagram of an energy system containing typical components as well as a DC minigrid, which consists of DC loads, a high-voltage battery storage, and a DC-coupled photovoltaic power plant.

A typical heterogenous energy system is depicted in Fig. 2 on the preceding page. The energy system is connected to the public electricity grid and has three operational voltage layers. The first layer is a conventional 400/230-V three-phase AC grid, in which various AC loads and sources are connected. As a second layer, the system contains a 350-V DC subgrid, which is connected unidirectionally to the AC grid (consumption only). The DC subgrid is composed of a high-voltage battery storage, a PV power plant, and (partially refitted) DC loads. These loads are devices that run on DC internally, such as LED lighting and computers; operating them in a DC grid clearly reduces AC/DC conversion losses.

The third layer is a single-phase AC grid unidirectionally connected to the DC subgrid (consumption only). It represents an uninterruptible power supply (UPS) for key components.

Figure 3 shows the top level of the modelled energy system. It is built in a modular way so that it can be adapted and parametrized for any given AC system comprising components belonging to one of six fundamental categories (1–6). The layers two (DC grid) and three (AC UPS) are implemented in submodels (grey and orange blocks). The same remark applies to the Modbus and CAN communication interfaces (yellow and blue blocks). The characteristics of the functional units used in the model have been derived from the characteristics of physical units that the PLC is to control in the field. To this end, historic data of representative energy system components were analyzed (PV system output, plant power consumption, peak powers, power distribution, etc.).

The control algorithm running on the PLC is rule-based and implemented according to IEC 61131. The many gaps in one's knowledge about the static and dynamic characteristics of system components and the limited resources available in field devices usually force one to get by without classical approaches such as PID control, state-space control, and controller parametrization by mathematical optimization.

Historic data and model-based PV power predictions are used to improve the scheduling of flexible loads and battery power. The implementation of PV power predictions is based on the python library "PVLIB-Python" [5] and the numerical weather prediction model "COSMO-D2" of the German weather service "Deutscher Wetterdienst" (DWD) [6]. The prediction is updated every three hours with a forecast horizon of 27 hours and once a day with a 45-hour horizon. The forecast horizon of the power prediction is determined by the forecast horizon of the numerical weather prediction model.

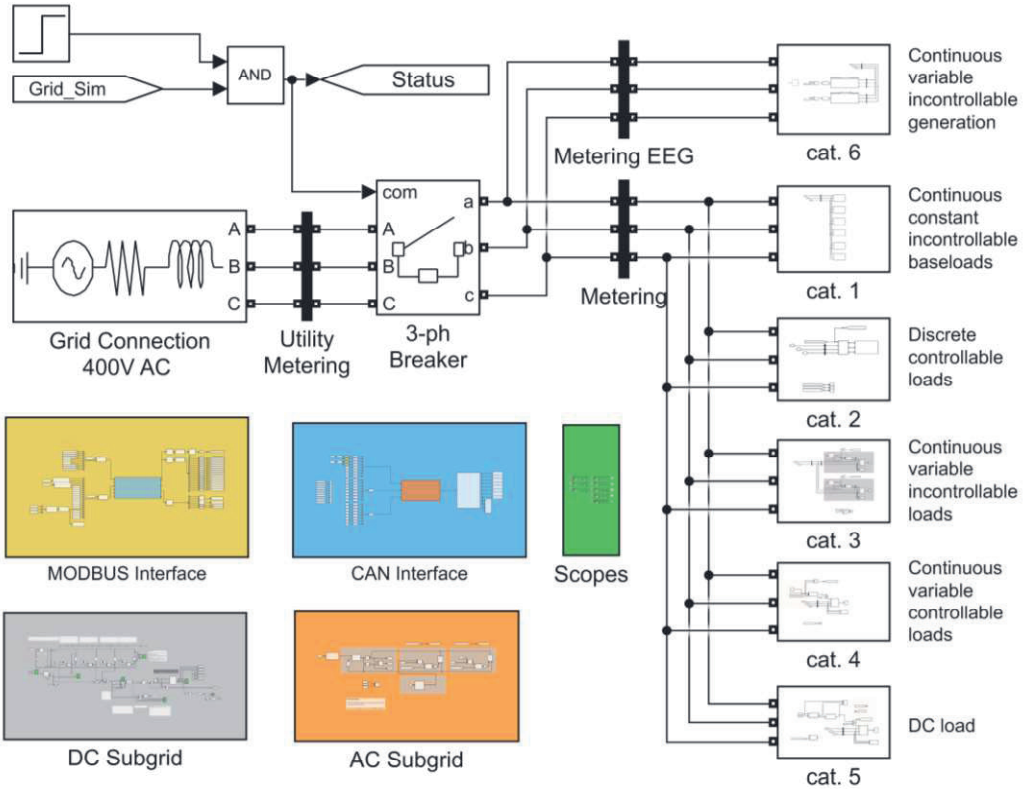


Fig. 3: Highest level of the energy system model for the real-time simulation and HiL test of the PLC-based system control.

Results

The HiL test concept allows one to validate the control quality. Control algorithms can be tested and verified, while being implemented on field hardware and using field communication. The testing possibilities include control code errors, testing and validating of the system dynamics, consideration of special cases, and the simulation of critical states.

By unit testing, single programming organization units of the PLC code can be tested. By integration testing, modules of the energy system control, such as communication modules or device-specific modules, can be evaluated. By system testing, the entirety of the modeled (virtual-reality) components controlled by the PLC can be shown to better meet the objectives of load flexibility and cost reduction than the uncontrolled system.

Figure 4 shows the result of the HiL simulation of a daytime period of the energy system. A simple control algorithm was used for the energy system in order to facilitate the evaluation of the HiL testbed. The case shown demonstrates some typical features. Whereas PV excess power is available in the first half of the day and therefore has to be curtailed, this is not the case in the afternoon (Fig. 4a). The reason is that a 6-kW load in the AC minigrid (layer 3) was switched on by the energy-system controller (Fig. 4b). As a consequence, the battery was charged in the morning (positive battery power) and discharged in the afternoon (negative battery power). This also shows in the state of charge of the battery (Fig. 4d). The resulting rectifier power at the input of the DC grid is presented in Fig. 4e.

The chart in Fig. 5 on the following page is a comparison of predicted PV power and measured output power for an existing PV power plant. The prediction was based on a 45-hour wheather forecast horizon as described above and the measured data were taken on site by a field-measurement unit. As expected, the prediction does not agree perfectly with the actually observed PV power output; after all, wheather forecasts are uncertain. Nevertheless, even a partially correct prediction of PV output power to be expected on the next day is a valuable piece of information. It enables one to better run the energy system *on the average* (not in every single case) by increasing the own consumption of electrical energy from the PV power plant.

The charts in Fig. 6 show an analysis of exemplary historic energy-system data of two full days. During the morning hours of both

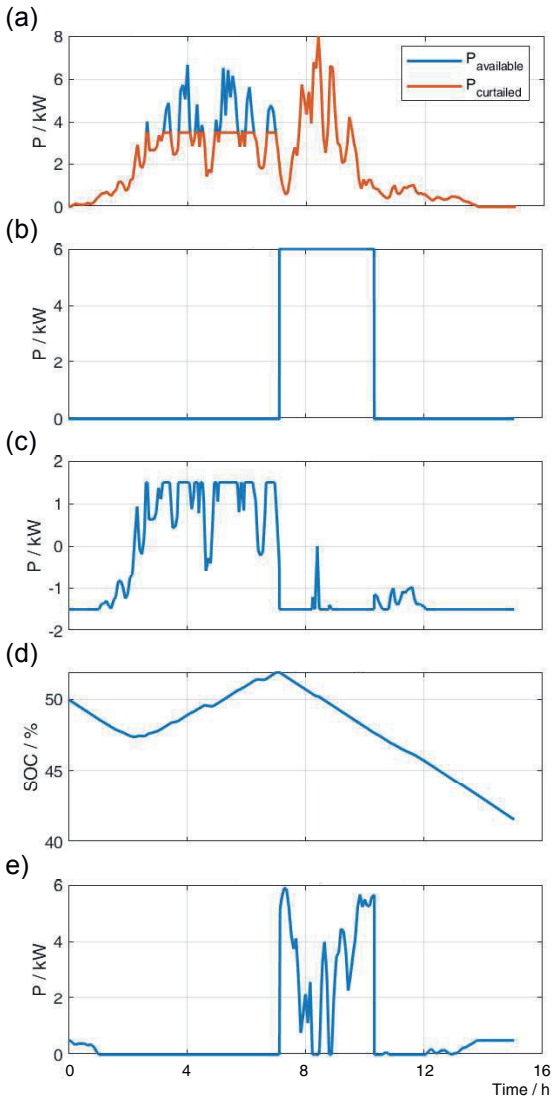


Fig. 4: Exemplary results of an energy system simulation. (a) Available (——) and curtailed (——) output power of a PV power plant. (b) Adaptive load adjustment in the AC minigrid (layer 3). (c) Resulting battery power in the DC minigrid (layer 2). (d) State of charge of the battery storage. (e) Rectifier power of the DC-grid feed-in from layer 1.

days the power demand in the energy system is higher than the PV power. During midday, the PV power exceeds the power demand. For this reason, the own consumption is limited to the instantaneous value of the power demand (Fig. 6a). As a result, power is drawn from the grid during all hours, except midday, where excess power is fed to the grid (Fig. 6b). This can also be derived from the 15-minute mean values, a common measure in energy economics for accounting purposes, of the load (Fig. 6c) and residual grid powers (Fig. 6d). By controlling the single entities in the system in an intelligent way, the residual load curve can be

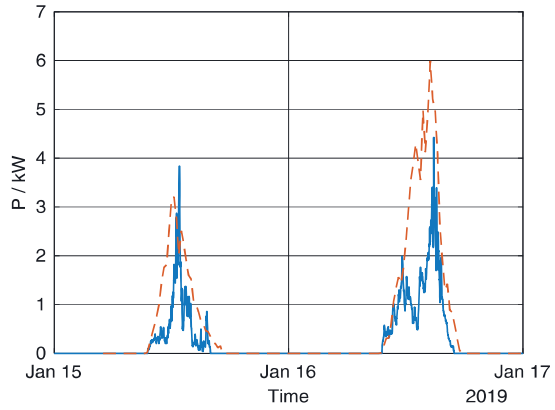


Fig. 5: Measured (——) and predicted (– –) output power of an existing PV power plant (January 15th and 16th, 2019). Prediction with a 45-hour forecast horizon based on numerical weather prediction.

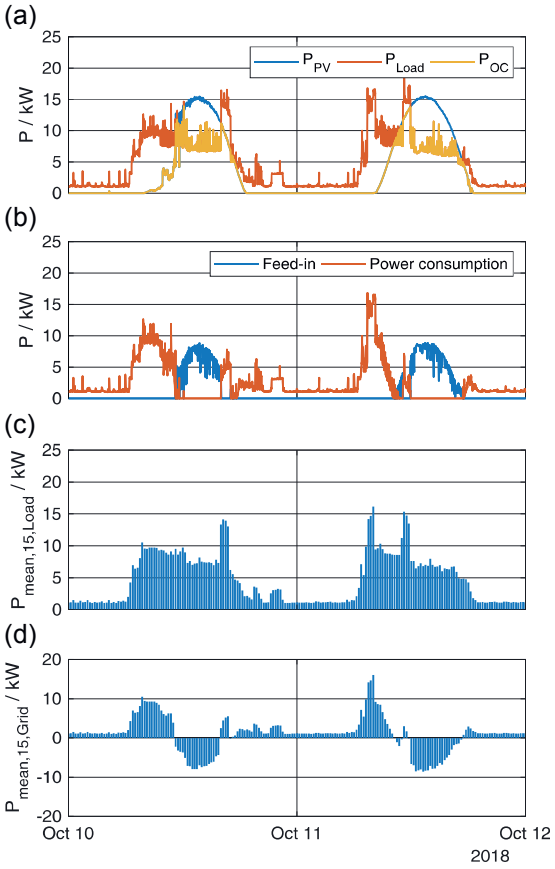


Fig. 6: Analysis of exemplary historic plant data (October 10th and 11th, 2018). (a) PV power (——), load (——) and own consumption (——). (b) Excess power fed to the grid (——), power consumption from the grid (——). (c) 15-minute mean values of power (load). (d) 15-minute mean values of power (grid).

shaped in order to avoid load peaks, to offer flexibility to the grid, and to increase own consumption. This can be seen clearly, for instance, at the three peak load situations that arise from the charging of an electric vehicle, alongside the usual daily load curve.

Discussion and Outlook

The proposed testbed is suitable to characterize and test energy-optimized process control solutions that are implemented on typical industrial control hardware, including typical field communication in real-time. Different control algorithms can now be implemented on field hardware and their performance and implementation accuracy can be tested extensively. Furthermore a virtual commissioning based on VDI/VDE 3693 can be developed [7, 8].

The current control algorithm will be evaluated extensively with the developed simulation tool. Furthermore, the control algorithm will be extended and revised. Additionally the PV power prediction accuracy will be evaluated by comparison with measurement data of a long time period including all seasons.

Acknowledgements

This work was funded by the Central Innovation Programme for SMEs (ZIM) of the Federal Ministry for Economic Affairs and Energy (BMWi) – project number ZF4152303LF7.

References

[1] Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen, *Quartalsbericht zu Netz- und Systemsicherheitsmaßnahmen,* Gesamtjahr und Viertes Quartal 2017, Bonn (2018)

[2] J. Villar, R. Bessa, M. Matos, *J. Electr. Pow. Syst. Res.* 154, 329–340 (2018); doi: 10.1016/j.epsr.2017.09.005

[3] A. Chaouachi et al., *IEEE Trans. Indust. Electron.* 60(4), 1688–1699 (2013); doi: 10.1109/TIE.2012.2188873

[4] J. Silvente et al., *Appl. Energy* 155, 485–501 (2015)

[5] W. F. Holmgren, C. W. Hansen, M. A. Mikofski, *J. Open Source Software* 3(29), *884,* (2018); doi: 10.21105/joss.00884

[6] M. Baldauf et al., *Beschreibung des operationellen Kürzestfristvorhersagemodells COSMO-D2 und COSMO-D2-EPS und seiner Ausgabe in die Datenbanken des DWD*, Offenbach (2018)

[7] VDI/VDE 3693 Part 1, Virtual commissioning – Model types and glossar (2016)

[8] VDI/VDE 3693 Part 2, Virtual commissioning – Introduction of virtual commissioning in companies (2018)

DOI 10.5162/sensoren2019/P3.10

Proof-of-concept trial of the portable electronic nose PEN3 for detection of formic acid concentration in the beehive

Antonia Genath[1)], Michael Hofmann[2)], Carlo Tiebe[2)], Ralf Einspanier[1)]
[1)]Institut für Veterinär-Biochemie, Freie Universität Berlin, Oertzenweg 19 b, 14163 Berlin
T: +49 30 838-62575, E: antonia.genath@fu-berlin.de, ralf.einspanier@fu-berlin.de
[2)]Bundesanstalt für Materialforschung und -prüfung (BAM), Unter den Eichen 44-46, 12203 Berlin
T: +49 30 8104-4518, E: michael.hofmann@bam.de, carlo.tiebe@bam.de

Abstract

Formic acid (FA) treatment as measure against the ectoparasitic mite, *Varroa destructor*, which infests the western honey bee, *Apis mellifera,* is usually performed twice a year. Although the therapeutic index is very narrow and the success of FA-treatment is highly dependent on various external factors, little is known about the local FA concentrations occuring in the hive. In this study we examined whether a commercially available odour measuring device (“electronic nose”) is suitable for measurements of high FA concentrations, as present during treatment. As measuring systems, two different setups were prepared for detection of FA concentrations in gas mixtures: A polymethyl methacrylate (PMMA) box was used to investigate whether the electronic nose is able to distinguish between different FA concentrations and a two-storied Segeberger hive to test under realistic conditions during bee-keeping. In both experimental systems, the signals from the same sensor components (sensor type no. 1, 6, 7, 8 and 9) showed an increased sensitivity towards the presence of FA in the gas mixture. Principle component analysis (PCA) and linear discriminant analysis (LDA) allowed for identification of specific patterns in the measurement signals and showed clear correlations between the signals and the present FA concentration. These results indicate that the introduced commercial multi gas sensor system is suitable for detection of high FA concentrations, and therefore could be a starting point developing a detection method of FA. The future aim is to successfully monitor the efficiency and effectiveness of this anti-varroa-treatment.

Keywords*: Apis mellifera, Varroa destructor*, Formic Acid, Metal-Oxide Gas Sensor Array

Introduction

Varroosis, caused by the ectoparasite *Varroa destructor*, is considered as one of the most important pests of the western honey bee, *Apis mellifera*. The mite causes considerable damage to its host, directly imparing the fat body tissue and indirectly by transmitting several harmful viruses and bacteria [1, 2, 3]. A colony infested with *V. destructor* will die within 2 to 3 years of the initial infestation if left without treatment [4].

Formic acid (FA) has been used mainly by beekeepers for decades as a treatment against *V. destructor* [5]. 60 % concentrated FA in water is applied in the beehive using different types of evaporators and acts in the gas phase. As FA is capable of penetrating the wax into the capped brood, it is the only miticide effective not only against the adult mites (on adult bees) but also the reproductive stages of the mite (in the sealed brood) [6]. Other advantages include that the mites are unlikely to develop resistance as well as the low risk of residue problems in bee products [1, 6, 7].

However, the therapeutic index – the range between efficiency against the mites and damage of the bees – is very narrow [8] and the therapy success is highly dependent on environmental factors, such as temperature and humidity as well as other influencing factors (e.g. bee colony size, type of application, concentration of FA and hive shape) [1].

Therefore, an accurate evaluation of the local FA concentration during treatment is necessary to either determinate the optimum efficiency of FA and/or to allow for immediate beekeeping intervention if the maximum concentration is exceeded.

The development of a screening method for estimation of actual FA concentrations in the

beehive would provide a major advance not only in honey bee health related research, where an exact determination of the FA concentration is necessary for generation of precise research results. Furthermore, professional beekeepers get the oppurtunity to monitor the therapy success during varroa control without beeing exposed to potentially harmful FA concentrations.

The long-term objective of this study is to develop a suitable and affordable device, which can be used particularly for beekeeping practices but also for scientific purposes. The commercially available electronic nose PEN3 (PEN3, AIRSENSE Analytics GmbH, Schwerin) provides the basis for further development.

The aim of the preliminary tests was to check whether PEN3 is suited to detect such high FA concentration in a gas mixture under similar conditions present in the beehive during FA treatment.

Material and Methods

Two different systems were prepared as measuring stations for detection of FA concentration in gas mixtures:

A) A PMMA box with a volume of 1.9 L, which was supplied with a gas flow of humidified synthetic air (50 mL min^{-1}, 100 mL min^{-1}, 300 mL min^{-1}), operated as a test chamber (Fig. 1). The humidity is generated by flushing synthetic air through a humidifier at 210 kPa. The humidified synthetic air is decompressed to ambient pressure. The dew point temperature of this humidified gas is 13.3 °C at ambient pressure (1013 hPa). To generate a temperature of 25 °C the box was placed on a temperature-controlled heating mat or in an oven to generate a temperature of 35 °C, respectively. A glass dish (surface of liquid phase: 9.6 cm^2) containing different concentrations of FA in deionised water (0 % to100 %) was placed within the system.
Each measurement was conducted after ten air exchanges and repeated three times, starting with blank value determination and subsequently increasing concentrations of FA in water (0 %, 20 %, 40 %, 60 %, 80 %, 100 %).

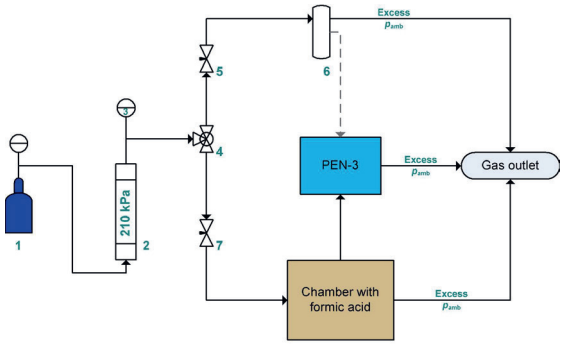


Fig. 1: Flowchart of FA detection in a measuring chamber (PMMA box). 1 – synthetic air, 2 – humidifier (water reservoir at p = 210 kPa, t = 23 °C), 3 – manometer, 4 – union tee, 5 – control valve, 6 – gas cylinder and tube to zero-air measurement, 7 – control valve.

B) A two-story polystyrene Segeberger hive compromising two boxes, a lid and a walking floor with floor grid (Fig. 2), containing eight filled honey frames per level was placed on a heat mat to generate a temperature of 25 °C. 60 % FA was applied by means of a specially designed evaporator (Nassenheider Evaporator Universal R). To avoid pollution in the surrounding area, the system was placed under a fume hood. Measuring points were drilled into the chest (three in the upper story, one in die lower story) in order to determine the FA concentration at distinct locations inside. The measuring points were connected with the analyser via PTFE (Polytetrafluoroethylene) hoses with a length of 30 cm, an outer diameter of 4 mm and a wall thickness of 0.5 mm. The gas phase in the hive was tested by PEN3 at distinct sampling times (1 h, 2 h, 3 h, 4 h, 24 h) at four sampling points until an equilibrium concentration of FA was established.

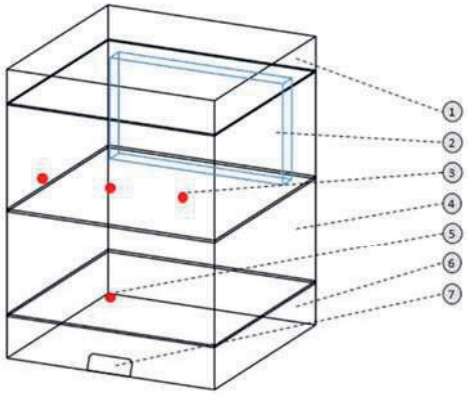


Fig. 2: Sketch showing Type B) measuring station for detection of FA concentration in gas mixtures. 1 – lid (50 cm × 50 cm × 10 cm), 2 – Nassenheider® Evaporator Universal R), 3 – measuring points upper story (1 – 3), 4 – body (50 cm × 50 cm × 22.3 cm), 5 – measuring point lower story (4), 6 – floor (50 × 50 × 10 cm), 7 – flight hole.

Tab. 1: Sensor types and detectable substance groups of the metal-oxide gas sensor array.

No.	Sensor type	Substance group
1	W1C	Aromatic compounds
2	W5S	Broadband
3	W3C	Aromatic compounds
4	W6S	Hydrogen
5	W5C	Aromatic and aliphatic compounds
6	W1S	Broadband, methane
7	W1W	Organo-sulfur compounds
8	W2S	Broadband, alcohols
9	W2W	Sulfur chlorides
10	W3S	Methane, aliphatic compounds

PEN3, an instrumented odour measuring device containing an array of chemical sensors and a pattern recognition software, was used for detection of the FA-gas-mixture. An internal pump was transferring the measuring gas mixture to the internal 10-fold metal-oxide gas sensor array (Tab. 1). For dilution purified synthetic air (system A) and ambient air (system B) were utilized as zero-gas and swept over the sensor array. The generated sensor signals describe a resistance ratio between the resistance value for each sample-gas (R) and zero-gas (R_0) measurement. Gases in varying concentrations and compositions can be recognized by their signal pattern and distinguished from another.

Results

In order to estimate the order of magnitude of the FA mass concentration in a Segeberger hive, generated by the Nassenheider Evapaporator Universal R, a theoretical consideration based on the balance equation of an ideal continuous batch reactor with constant emission source [9, 10] the following equation was used:

$$\frac{d\beta(t)}{dt} = \frac{E}{V} - n \cdot \beta(t), \qquad (1)$$

where β is the mass concentration of FA in air in mg/m³, E is the evaporation rate of FA in mg/d, n is the air change rate in 1/d, V is the free volume in m³ and t is the time in d. The solution of the differential equation under the boundary condition $\beta(0) = 0$ results as follows:

$$\beta(t) = \frac{E}{n \cdot V} \cdot (1 - e^{-n \cdot t}). \qquad (2)$$

The volume of the hive according to the internal dimensions (Fig. 2) without honey comb frames is 35.2 L. According to the data sheet of Nassenheider Evaporator Universal R [9], a 60 % FA volume of 180 mL is sufficient for an application period of 12 d. Hence, the daily consumption of 60 % FA in water solution is 15 mL/d ± 3 mL/d, which – with a FA density of 1.2273 g/cm^3 – results in an evaporation rate of 11.1 g/d ± 2.2 g/d of FA.

The bees inside the hive generate a homogeneous distribution of FA/water vapour as well as an air exchange by the bee fanning. As a result, FA will reach even the brood in covered cells and an equilibrium concentration of FA in air is established in dependence on the air change rate.

Based on equation 2 and the assumptions made, FA mass concentrations in air of 2620 mg/m^3 to 26200 mg/m^3 in the hive are established at air change rates of 0.5 h^{-1} to 5 h^{-1} (Fig. 3).

The estimated mass concentrations of FA in the air of a hive are very high and above the range of the routine measurement range of available gas sensor elements or absorption test tube methods, e.g., for monitoring of maximum workplace concentrations - MAK value, which is

9.5 mg/m³ [11]. The direct sensor-supported measurement requires a problem-solving sampling strategy for the detection of FA in air.

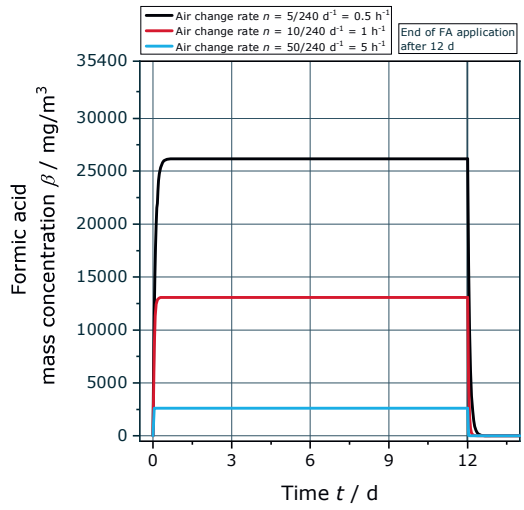


Fig. 3: Theoretical mass concentration course of formic acid over time during the FA-application.

In both experimental setups the signals from sensor components 1 (aromatic compounds), 6 (broadband, methane), 7 (organosulfur compounds), 8 (broadband, alcohol) and 9 (sulphur chlorides) showed an increased sensitivity towards the presence of FA in the gas mixture. The sensor elements 6, 8 and 9 generated the greatest resistance ratios and could therefore be assumed the most sensitive components.

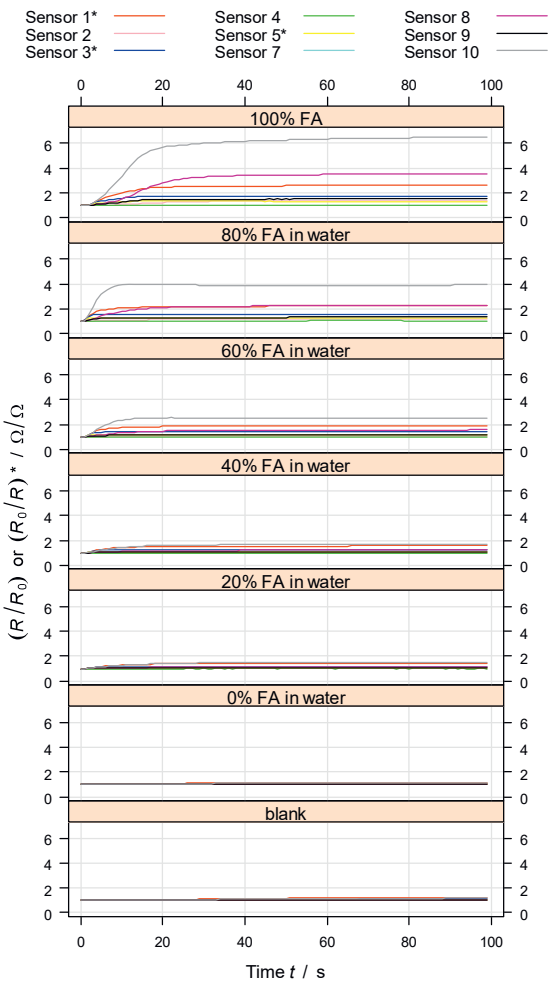


Fig. 4: Sensor sensitivities towards the presence of FA in the experimental setup type A at 35 °C.

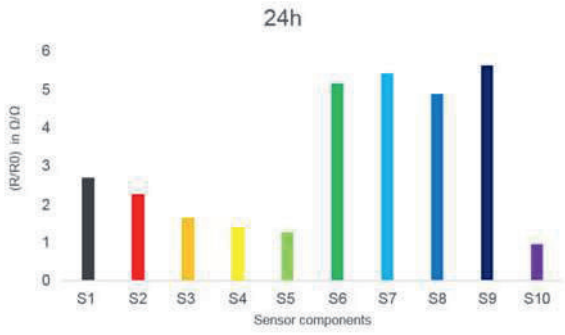


Fig. 5: FA-measurement implemented after 24 h in experimental system type B at 25 °C. In both systems signals from sensor components 1, 6, 7, 8 and 9 display clear dependence on FA concentration.

In order to identify specific patterns in the measurement signals, statistical methods like principle component analysis (PCA) and linear discriminant analysis (LDA) were conducted subsequently based on the generated data.
Figure 6 shows PCA- and LDA-plots of the most important components PC1/ PC2 and LD1/ LD2, respectively, performed on the measurement data generated during a test run in test system A with a temperature of 35°C, a gas flow rate of 300 mL min^{-1} and different concentrations of FA in water (0 - 100%).
Clearly recognizable are the dependencies of PC1 and LD1 (abscissa) on the presented FA concentration.

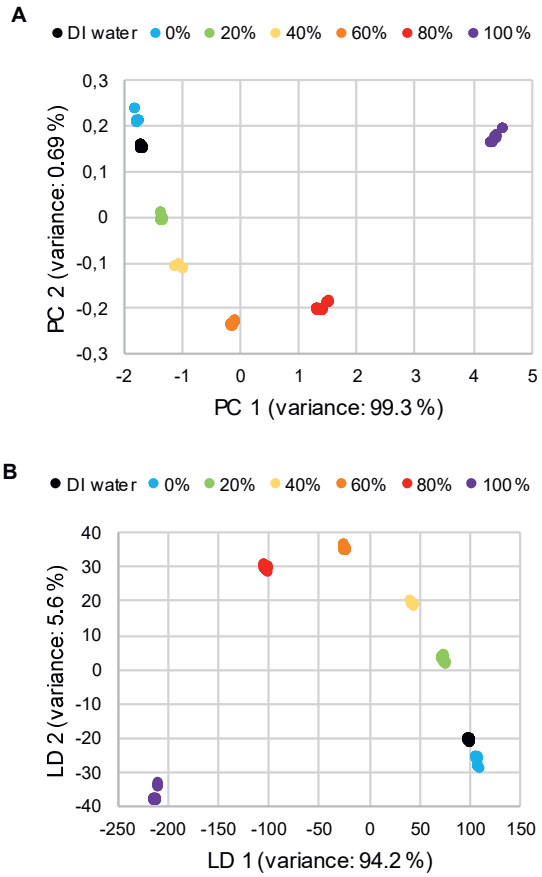


Fig. 6: PCA (A) and LDA (B) results of setup type A at 35 °C, a gas flow rate of 300 mL/min and different FA concentrations in water. Basis of this calculations are the steady state resistance values in the measurement time range 85 s to 95 s (Fig. 4).

These preliminary results demonstrate that the multi gas sensor system PEN3 is suitable for detection of high FA concentration in gas mixtures under similar conditions to those applicable in a honey bee hive during FA treatment.

Discussion/Vision

Our initial results show that metal-oxide gas sensor arrays, like those used in PEN3, are suitable for detection of FA in gas mixtures even in high concentrations, which exceed MAK value (5 µmol/mol equal to 9.5 mg/m^3). This feature allows an application in hives during FA treatment, where it cannot be excluded that exceptionally high concentrations of FA could potentially corrode or in some way harm the sensor systems.
Additionally, this type of test system is remarkable for their practical relevance due to the possibility of real-time data generation. A low cost alternative of this high-sophisticated commercially available sensor components may be of interest increasing its portability and miniaturisation.
Further investigation are required concerning the cross-sensitivities towards testparameters like temperature, humidity, air exchange rate and other chemical compounds, which frequently occur in hives.
Moreover, this system still needs to be validated, calibrated and tested for its reliability in subsequent field trials and laboratory tests according to guidelines like VDI/VDE 3518 [12-14]. Corresponding trials will be conducted in the honey bee season 2019.
A theoretical consideration based on the balance equation of an ideal continuous batch reactor with constant emission source is applicable to estimate the temporal variation of the FA mass concentration in a hive during the FA application. However, the real air change is unknown and depends on various factors, like bee fanning, temperature inside the hive, and the gas thightness of the hive.

Our findings may serve as the technical basis for the development and further improvement of a more suitable detection method of FA in terms of monitoring the efficiency and effectiveness of the FA treatment against the bee-ectoparasite *Varroa destructor*.

References

[1] P. Rosenkranz et al. “Biology and control of Varroa destructor”, *Journal of Invertebral Patholy* 103 Suppl 1, S96-S119 (2010); doi: 10.1016/j.jip.2009.07.016

[2] S. Ramsey et al. “*Varroa destructor* feeds primarily on honey bee fat body tissue and not hemolymph”, *Proceedings of the National Academy of Sciences of the United States of America* 116,5, 1792-1801 (2019); doi: 10.1073/pnas.1818371116

[3] S. Bernardi, V.Ezio “Viral epidemiology of the adult Apis mellifera infested by the Varroa destructor mite.” *Heliyon* 2,5 e00101 (2016); doi: 10.1016/j.heliyon.2016.e00101

[4] P. Neumann et al. „Varroa invasion and virus adaptation”, *Trends in Parasitology* 28,9, 353-354 (2012); doi: 10.1016/j.pt.2012.06.004

[5] P. Elzen et al. “Formic Acid Treatment for Control of Varroa destructor (Mesostigmata: Varroidae) and Safety to *Apis mellifera* (Hymenoptera: Apidae) Under Southern United States Conditions” *Journal of economic entomology* 97, 1509-1512 (2004); doi: 10.1603/0022-0493-97.5.1509

[6] J. Amrine Jr., R. Noel “Formic acid fumigator for controlling varroa mites in honey bee hives”, *International Journal of Acarology* 32,2, 115-124 (2006); doi: 10.1080/01647950608684452

[7] S. Bogdanov et al. “Determination of residues in honey after treatments with formic and oxalic acid under field conditions” *Apidologie*. 33, 399-409 (2002); doi: 10.1051/apido:2002029.

[8] R. M. Underwood, R. W. Currie "The effects of temperature and dose of formic acid on treatment efficacy against Varroa destructor (Acari: Varroidae), a parasite of Apis mellifera (Hymenoptera: Apidae)." *Experimental & Applied Acarology* 29,3, 303 (2003); doi: 10.1023/A:1025892906393

[9] S. Weiland, Nassenheider Evaporator Universal R - Long-Term Evaporator for Formic Acid, Operating Instructions, (2019-04-17), https://www.nassenheider.com/public/de/index.php?controller=attachment&id_attachment=18.

[10] E. Müller-Erlwein, Chemical reaction engineering (in German), Volume 3, Springer Fachmedien, Wiesbaden (2015); doi: 10.1007/978-3-658-09396-9.

[11] GESTIS Substance Database, Material Safety Data Sheet „Formic Acid”, (2019-04-24), http://gestis-en.itrust.de/nxt/gateway.dll/gestis_en/011490.xml.

[12] VDI/VDE 3518-1:2011-05, Multigas sensors – Terms, configuration, function and classification.

[13] VDI/VDE 3518-2:2014-12, Multigas sensors – Function, classification and assessment.

[14] VDI/VDE 3518-3:2018-12, Multigas sensors – Odour-related measurements with electronic noses and their testing.

DOI 10.5162/sensoren2019/P3.11

Recommendations for optimal sampling and filter rates in liquid chromatography

David Thomas Marehn[1,2,3], Detlef Wilhelm[1], Heike Pospisil[2], Roberto Pizzoerrato[3]
[1]AnaTox GmbH & Co. KG, Reinheimer Strasse 18b, 15517 Fuerstenwalde, Germany
[2]UAS Wildau, High Performance Computing in Life Sciences, Hochschulring 1, 15745 Wildau, Germany
[3]Università degli Studi Di Roma "Tor Vergata", Dipartimento Ingengeria Industriale, Via del Politecnico 1, 00133 Rome, Italy

Abstract

In the strong regulated field of pharmacy, all substance analysis are performed by very well defined experiments using specific techniques and parameters for each substance, which are written down in the pharmacopoeias, mainly the United State Pharmacopoeia (USP), Japanese Pharmacopoeia (JP) or European Pharmacopoeia (EP). Some of these experiments use techniques which belong to the field of chromatography especially the high performance liquid chromatography (HPLC) and gas chromatography (GC). Nearly all possible parameters are listed in the so called monographs of the active substances. Nevertheless two editable parameters of the detectors are not mentioned and not restricted here: the sampling rate and the signal filtration. Both parameters have a clear impact on the peak width, peak height and peak distortion. Therefore, it would be desirable to define recommendable setups for chromatography tests.

In this article we present how to find the optimal values for the sampling rate and signal filtration for a given substance of a monograph. Chromatograms with the highest possible sampling rate and lowest filtration have been altered by performing a sampling reduction and various signal filtration algorithms considering the detector firmware of the major HPLC manufacturers. For the evaluation, several chromatographic properties describing a chromatogram have been taken for an optimization process based on the multi objective (Pareto) optimization. The optimization has been tested on simulated and real HPLC data. The suitability has been proved by comparing the results against the usual approach of 20 points per peak.

The criterion allows to make further investigations on the substance monographs in order to develop the sampling rate and signal filtration recommendations depending on the HPLC system firmware.

Keywords: Chromatography, Pharmacopeia, Sampling Rate, Signal Filtration, Optimization

Introduction

Based on the increasing regulation in the field of pharmacy within the last century all substances that are used for medicinal products are very well defined and described in in the pharmacopoeias, whereby the United State Pharmacopoeia (USP), Japanese Pharmacopoeia (JP) and European Pharmacopoeia (EP) belong to the global player [1]. Each substance contains experiments which definitions resulted from method development processes in order to get a suitable setup for analyzing the active substances inside a medicinal product before selling it. All these setups are written down in the so called monographs and are describing specific techniques and parameters for each substance. Among other techniques some of the substances will be analyzed by the usage of chromatography like the high performance liquid chromatography (HPLC) and gas chromatography (GC). Nearly all parameters and conditions of the analytical system are defined and listed.

Nevertheless two editable parameters of the data acquiring module, the detector, are not mentioned and not restricted: the sampling rate and the signal filtration. Both parameters have a clear impact on the peak width, peak height and peak distortion [2, 3, 4]. In other words, two experiments under complete equal conditions may result in different chromatograms with respect to the mentioned detector parameters. This has a notable influence on the conclusion of a substance test and contradicts the strong legal requirement in pharmaceuticals. Therefore, it would be

desirable to define recommendable setups for chromatography tests.

Sampling Rate and Signal Filtration

Regardless of the detection type every detector used in the chromatography produces an analog, electrical signal like changes of voltage or current [5, p.19]. Usually an internally installed analog-to-digital converter (ADC) digitize the analog, continuous signal into a discrete signal and provide it to an external controller (e.g. computer) by a suitable interface. The sampling rate is a parameter of the ADC and specifies how many discrete data points per second will be sampled and provided. The unit of the sampling rate is given in Hz (1/s). Presently there are two approaches how an ADC operates. Either the analog signal will be exactly sampled by the given frequency or the highest available frequency will be used and a specific amount of data points will be averaged to one point. The second method is called data bunching [5, p.21].
The signal filtration is a routine during the data acquisition [2] or the data processing afterwards [4]. It is either based on physical hardware components, digital filters implemented in the firmware of the device or on algorithms within the chromatography data system (CDS) software installed on the computer which collects and interpret the resulting chromatogram. The filtration itself deals with noise that overlays the analog signal disturbing the chromatographic analysis [6].

Optimize Sampling Rate Reduction and Signal Filtration of Raw Data

In this article we present how to find the optimal parameters for the sampling rate and signal filtration depending on up to four criterions that are usually used for the evaluating of a chromatogram during the method development. All the experiments are based on a self written tool implementing the two sampling rate reduction approaches, several signal filtration algorithms and the continuous update method for the multi objective optimization. The input for the tool are chromatograms with the highest available sampling rate and lowest filtration because these steps will be simulated by the tool itself.
For the optimization testing several sampling rates we focused on the reduction by data bunching because this seems to be the most realistic approach due to the integrator used in Delta-Sigma analog digital converters installed in most data acquisition systems [2]. So the reduction of the sampling rate from for example 100 Hz to 20 Hz leads to an averaging of five consecutive data points to a new single one.
The choice of the implemented and tested signal filtration algorithms consider the physical and digital filters installed in the detectors of the major HPLC manufacturers. As a physical filter so called RC filters using capacitive and ohmic resistances are installed and parameterized by a time constant (τ= R*C). An estimation equation for the RC filter exists [7] and shown in equation 1.

$$s_{out} = \frac{\Delta t \cdot s_{in}(t) + RC \cdot s_{out}(t - \Delta t)}{\Delta t \ + \ RC} \qquad (1)$$

With Δt as step width between two data points, RC as time constant τ, s_{in} as the raw signal and s_{out} as the resulting, filtered signal.
All other used filters belong to the digital filters and are implemented into the firmware of the detector devices. Additional all the following filters are windowed filters. That means a window with a specific size of N data points will pass the raw data and compute a filtered data point based on a weighting function applied on the data points in the window. The windows size (WS) is odd and the main parameter of all windowed filters. The most simple filter is the moving average. Here all data points in the window have the same weighting as shown in equation 2.

$$f_i = \frac{1}{WS} \sum_{j=-WS/2}^{WS/2} s_{i+j} \qquad (2)$$

With f as the filtered signal, s as the raw signal and WS as the window size.
The Hamming filter is based on the window function by Richard Hamming and weights the data points as visible in equation 3.

$$w_n = 0.54 - 0.46 \cdot \cos\left(\frac{2\pi n}{WS - 1}\right) \qquad (3)$$

With w_n as the filter weight of the n-th data index (beginning with 0) in the window and WS as the window size.
One of the most popular filter algorithms is Savitzky-Golay. Here the weighting is based on a polynomial regression of the data points in the window. So an additional parameter for this filter is necessary, the grade of polynomial. The weighting of each data point in the window can be computed [8] or prepared in a coefficient table as we do up to a maximal window size of 29 and grade of polynomial of 15.
The last implemented filter is the moving average filter with Gaussian kernel. Here the weighting uses the Gaussian distribution (equation 4).

$$f_i = \sum_{j=-WS/2}^{WS/2} s_{i+j} \cdot e^{\frac{-\pi \cdot j^2}{SD^2}} \quad (4)$$

With f as the filtered signal, WS as the window size and SD as the standard deviation of the Gaussian equation. So the filter here depends on the window size and the standard deviation.
For the optimization of the sampling rate and signal filtration parameters the multi objective optimization (Pareto optimization) has been chosen [9]. That way it is possible to optimize the result of the filtration on criterions that evaluate the chromatogram. These criterions are influenced by the sampling rate and signal filtration and are the peak width at half height, the symmetry factor, signal-to-noise ratio and resolution. The peak width and the distance of the symmetry factor to the ideal value of 1 will be minimized as where the signal-to-noise ratio and resolution will be maximized. This approach differs from other possible optimization methods using one specific evaluation value like the Durbin-Watson value [10] or largest negative chromatogram second derivative [4].

Materials

Three different kinds of chromatograms were used for the optimization process. Two of them were simulated using the general creation equation of the HPLC simulator developed by Boswell et al. [11]. The function for one peak generation is:

$$C_i(t) = \frac{W_i}{2\sqrt{\pi}\sigma_{t,i}F} e^{-\left(\frac{t - t_{R,i}}{2\sigma_{t,i}}\right)^2} \quad (5)$$

With C_i as the molarity of compound i, W_i as the number of moles, σ as peak width, t_R as retention time and F as flow rate. Each compound i forms a peak function of t. For the final chromatogram all peak functions are summed up. That way a chromatogram containing one peak only and a sampling rate of 100 Hz was created. Another one containing 21 peaks corresponding to a separation of paracetamol, codeine, pitophenone and their impurities done by Vijta et al. [12] was created with a sampling rate of 40 Hz.
The real chromatogram of a HPLC isocratic test sample was acquired by an Agilent Technologies 1260 system that consists of a G4225A degasser, G1312B binary pump, G1367E wellplate autosampler, G1330B autosampler thermostat, G1316C column compartment and G4212B diode array detector. The isocratic test sample contains the four substances dimethyl phthalate, diethyl phthalate, biphenyl and o-terphenyl solved in methanol. The eluent consisted of a mix of 35 vol.% HPLC-grade water and 65 vol. % Acetonitrile. For the stationary phase Zorbax xDB-C8 column supplied by Agilent Technologies according to a reverse phase chromatography configuration was installed. The column had a length of 50mm, a diameter of 4.6 mm and a pore size of 1.8 µm. The analysis conditions where 1 ml/min flow, 1 µl injection volume, 40 °C column temperature, 254 nm detection wavelength and 80 Hz sampling rate.

Optimization Test Range

For the determination of the Pareto optimization front (POF) it is necessary to define a suitable test area for the parameters of the sampling rate and signal filters. For the sampling rates the common rates an Agilent Technologies detector provides were used: 100, 80, 40, 20, 10, 5, 2.5, 1.25 Hz. The given sampling rate of the raw chromatogram has been reduced only. The sampling rate has not ever been increased.
For the RC-Filter the time constant has been varied from 0 (no filtering) up to 1/2 * peak width at the half height ($W_{1/2}$) because it has been determined that the peak width exceeds an increasing of 25% at this moment when performing the filtration on an initial Gaussian peak [2]. The step of increasing the time constant depends on the current sampling period (1/sampling rate).
For all windowed filters the window size starts from 3 because a size of 1 will not filter. The size increases by 2 due to odd sizes up to a maximum that depends on the peak width at half height of the narrowest peak ($W_{1/2}$). The maximum will be calculated by equation 6.

$$WS_{max} = \frac{1}{2} \cdot sampling\ rate \cdot W_{1/2} \quad (6)$$

The Savitzky-Golay and Gaussian filter requires additional parameters with specific ranges. The grade of polynomial starts from 1 up to 1/2 * window size in order to avoid over fitting. The standard deviation of the Gaussian starts from 0.1 and increases by 0.1 as well as done by Lytle et al. [4] up to a 1/2 * window size because a higher standard deviation leads to a very close equal distribution.
The evaluation of the filtered chromatograms where done by a peak detection as a combination of local maxima detection by the persistence 1D algorithm [13] and the baseline expansion performed in the APEX Track algorithm of Waters Corporation [14]. The computations of the peak width, symmetry factor, signal-to-noise ratio and resolution are

based on the definitions in the general chapter <621> in the USP dealing with the chromatography [15].

Results and Discussion

By the usage of the self written optimization tool each kind of chromatogram visible in Fig. 1 has been processed by data bunching for reducing the sampling rate and by the filtration algorithms using the predefined parameter ranges. After the process of the raw data the peaks have been detected and evaluated by computing the values of the optimization criterions: peak width at half height, symmetry factor, signal-to-noise ratio and resolution (except for chromatogram A in Fig. 1 due to one single peak). The resulting values are used as input for the continuously updated method to find the non-dominated solutions forming the a Pareto optimization front (POF). At the end after processing all possible signal filtration parameters and sampling rates the optimal solution is determined by calculating the Euclidean distance to the utopia point. The utopia point itself corresponds to an ideal point which contains minimal or maximal values of each found solution depending on the optimization type. The Euclidean distance (ED) to the utopia point was calculated as follows:

$$ED = \sqrt{\sum_{object} \left(\frac{v_{object} - v^*_{object}}{v_{norm}} \right)^2} \quad (7)$$

Where object is the criterion like the peak width, v as the value of the object, v* as the utopia value of the object and v_{norm} as a normalization value that depends on the global minimal and maximal object values in all determined solutions. Additionally to the filter algorithms it has been determined which sampling rate is necessary to describe the narrowest peak in the chromatogram by 20 data points. The evaluation of this situation is shown in Tab. 1 beside the performed filtration algorithms. As visible in the table none of the solutions has the best value in all of the four criterions at the same time. That behavior applies for all three chromatograms. Each filter algorithm was optimized in an own run in order to get a solution one by one. A comparison of the six results by the Euclidean distance shows that for all three chromatograms the Savitzky-Golay algorithm with the given parameter lead to the best solution at all. It is presented as the rows in bold format in Tab. 1. But it has to be considered that for all three chromatograms the best grade of polynomial seems to be 1 here. That corresponds to a moving average filtration. The reason why the moving average

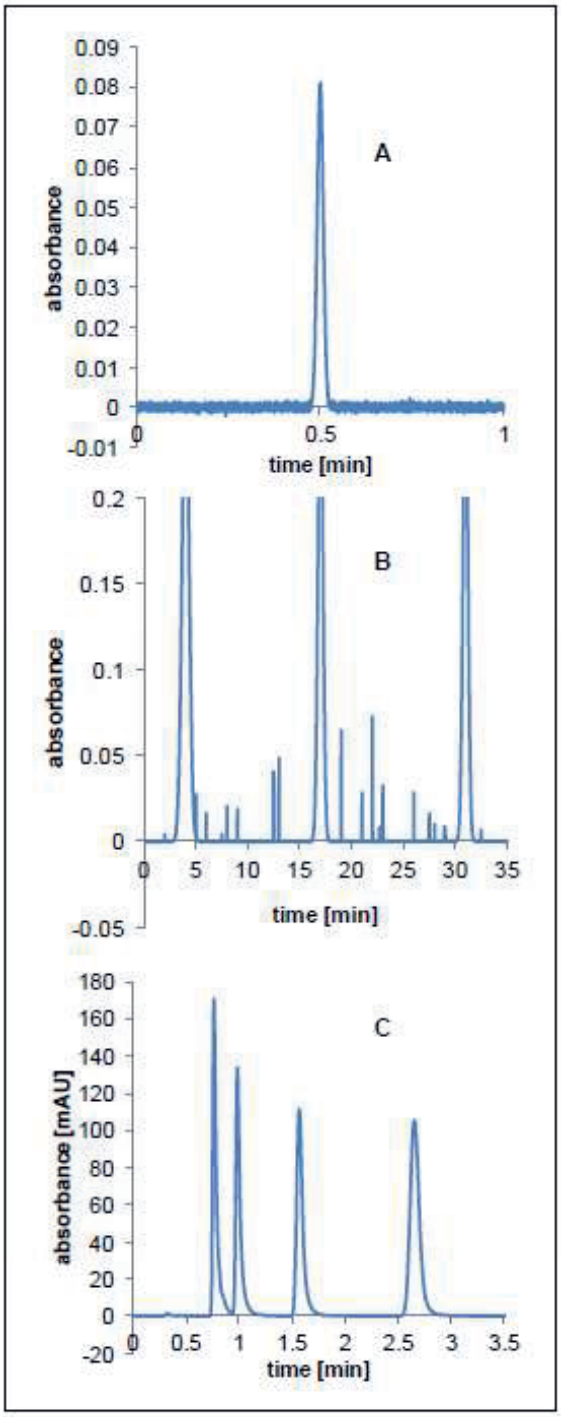


Fig. 1: A - Simulated chromatogram of one peak width sampling rate of 100 Hz, B - Simulated chromatogram based on separation of paracetamol, codeine, pitophenone and impurities in [12] with 40 Hz sampling rate, C - Real chromatogram of isocratic test sample acquired by diode array detector at 254 nm and with 80 Hz sampling rate.

filter and Savitsky-Golay did not have the same optimal solutions lies in the range of test area. The moving average has a wider range of settable window size because of the limited performance of the Savitsky-Golay filter with a maximal window size of 29. That way further solutions have been added to the POF for the moving average filter and the utopia point differs. That means comparing the final results between the algorithms is not as practicable as performing an optimization with all algorithms and parameters at once. Doing so is very time consuming and do not fulfill the demands in this article for the optimal solutions separated by each filter which corresponds with one implemented digital or installed physical filter in a data acquiring system. Nevertheless it is visible that representing the peaks by 20 data points is a good approach but if all three or four criterions are considered it is possible to find a better solution using one of the filter algorithms.

Conclusion

Using the written tool and apply the sampling reduction and all the known filtration algorithms on three test chromatograms has shown that it is possible to define a suitable test area for the filter parameter and optimize them by evaluation criterions that are usually determined during the method development and validation of a chromatographic experiment. That way it is possible to select a filter algorithm that belongs to a specific HPLC detector distributed by the different manufacturers and optimize the parameters.

For a general optimization comparing the different filter algorithms an all in one approach has to be performed by finding all solutions of the Pareto optimal front within one run.

During the current experiments the separated peaks in the chromatograms have been handled equivalent. In a following step the recommendations for the filter parameter can be improved by the distinction of the chromatograms in determination of active substance quantity or purity. This will influence the peaks that will be used for the evaluation. In Fig. 1 (b) for example a substance quantity determination would consider the three huge peaks of paracetamol, codeine and pitophenone only whereby the impurities become to the field of interest during a purity determination.

Tab. 1: Solutions of the optimization process divided by the processed chromatograms in Fig. 1 and the sampling rate and signal filtration parameters. The optimal solutions were determined by Euclidean distance to the utopia point. Each algorithm has its own utopia point. $W_{1/2}$=Peak width at half height, S/N-Ratio=Signal-to-Noise ratio, τ=time constant, WS=window size, GP=grade of polynomial, SD=standard deviation.

Sampling Rate and Filter Parameter	*∑ ($W_{1/2}$) [msec]*	*∑ |1 - Symmetry factor|*	*∑ (S/N-Ratio)*	*∑ (Resolution)*
Chromatogram A				
10.58 Hz - 20 data points	1170.233	0.007124312	156.0675	/
20 Hz - RC (τ = 150 ms)	1224.812	0.03504925	315.3293	/
40 Hz - Moving Average (WS = 11)	1173.835	0.06105919	257.8538	/
20 Hz - Hamming (WS = 5)	1170.628	0.04447394	226.5203	/
20 Hz - Savitzky-Golay (WS = 3, GP = 1)	**1168.979**	**0.006469480**	**206.7768**	**/**
20 Hz - Gaussian (WS = 5, SD = 2.3)	1170.858	0.04453622	230.6013	/
Chromatogram B				
11.55 Hz - 20 data points	134435.0	0.4056107	379.5420	1082.554
40 Hz - RC (τ = 50 ms)	125097.0	0.2609492	372.0598	1067.215
40 Hz - Moving Average (WS = 5)	126390.2	0.6281030	373.8028	1071.499
40 Hz - Hamming (WS = 5)	122863.3	0.5511815	359.7738	1079.187
40 Hz - Savitzky-Golay	**123327.1**	**0.2881788**	**362.7819**	**1078.525**

(WS = 3, GP = 1)				
40 Hz - Gaussian (ws = 5, sd = 1.6)	118718.9	0.6011666	342.7808	1084.018
Chromatogram C				
9.17 Hz - 20 data points	16984.34	7.781151	18247.30	62.51906
20 Hz - RC (τ = 100 ms)	16957.61	6.961686	22655.59	63.36911
80 Hz - Moving Average (ws = 39)	16367.06	6.170660	26634.29	76.55633
40 Hz - Hamming (ws = 19)	16175.93	6.374446	24793.98	80.99163
40 Hz - Savitzky-Golay (ws = 21, gp = 1)	**16397.73**	**6.288150**	**27415.59**	**76.67735**
80 Hz - Gaussian (ws = 51, sd = 20.9)	16206.84	6.237541	25445.40	80.37572

References

[1] Dedecke Handelsgesellschaft mbH. EP, USP und JP; *URL: http://www.dedecke-gmbh.de/epuspjp.html* (visited on 2017-10-12).

[2] M. F. Wahab, P. K. Dasgupta, A. F. Kadjo, D. W. Armstrong. Sampling frequency, response times and embedded signal filtration in fast, high efficiency liquid chromatography: A tutorial, *Analytica Chimica Acta* 907, 31-44 (2016); doi: 10.1016/j.aca.2015.11.043

[3] A. Felinger, A. Kilár, B. Boros. The myth of data acquisition rate, *Analytica Chimica Acta* 854 178-182 (2015), doi: 10.1016/j.aca.2014.11.014

[4] F. E. Lytle, R. K. Julian. Automatic Processing of Chromatograms in a High-Throughput Environment, *Clinical Chemistry* 62.1, 144-153 (2016); doi: 10.1373/clinchem.2015.238816

[5] D. Stauffer. Das Chromatogramm, *Chromatogramme richtig integrieren und bewerten*, Wiley-VCH Verlag GmbH & Co. KGaA, (2008); isbn: 978-3-527-62222-1 978-3-527-31774-5

[6] J. W. Dolan. The Role of the Signal-to-Noise Ratio in Precision and Accuracy, *LCGC Europe* 19.1, 12-16 (2006)

[7] G. Lungu. Introduction to Numerical Methods – Basic Filters #1; *URL: http://www.excelunusual.com/basic-filters-1/* (visited on 2018-04-12).

[8] P. A. Gorry. General Least-Squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method, *Analytical Chemistry* 62(6), 570–573 (1990); doi: 10.1021/ac00205a007

[9] N. Gunantara. A review of multi-objective optimization: Methods and its applications, *Cogent Engineering* 5, 1-16 (2018); doi: 10.1080/23311916.2018.1502242

[10] G. Vivó-Truyols, J. R. Torres-Lapasió, A. M. van Nederkassel, Y. Vander Heyden, D. L. Massart. Automatic program for peak detection and deconvolution of multi-overlapped chromatographic signals: Part I: Peak detection, *Journal of Chromatography A* 1096, 133-145 (2005); doi: 10.1016/j.chroma.2005.03.092

[11] P. G. Boswell, D. R. Stoll, P. W. Carr, M. L. Nagel, M. F. Vitha, G. A. Mabbott. An Advanced, Interactive, High-Performance Liquid Chromatography Simulator and Instructor Resources, *Journal of chemical education* 90, 198-202 (2013); doi: 10.1021/ed300117b

[12] J. Vojta, P. Hanzlík, A. Jedlička, P. Coufal. Separation and determination of impurities in paracetamol, codeine and pitophenone in the presence of fenpiverinium in combined suppository dosage form, *Journal of Pharmaceutical and Biomedical Analysis* 102, 85-92 (2015); doi: 10.1016/j.jpba.2014.09.002

[13] T. Weinkauf, Y. Kozlov. Extracting and Filtering Minima and Maxima of 1D Functions, 2013, *URL: https://www.csc.kth.se/~weinkauf/notes/persistence1d.html*

[14] Waters Corporation. ApexTrack Integration: Theory and Application, *Document*, (2016)

[15] The United States Pharmacopeial Convention. <621> Chromatography, *The United States Pharmacopeia and National Formulary*, (2017); isbn: 978-3-7692-6821-8

DOI 10.5162/sensoren2019/P3.12

Integration of Bragg grating sensors in components made of carbon fiber reinforced polymers

Theresia Sauer[1,], Stefan Kefer[2,*], Wolfgang Ruppert[1], Ralf Hellmann[2] and Michael Kaloudis[1]*
[1]Laboratory for Packaging and Interconnection Technology, University of Applied Sciences Aschaffenburg, Wuerzburger Strasse 45, 63743 Aschaffenburg, Germany
[2]Applied Laser and Photonics Group, University of Applied Sciences Aschaffenburg, Wuerzburger Strasse 45, 63743 Aschaffenburg, Germany
**E-mail: Theresia.Sauer@th-ab.de; Stefan.Kefer@th-ab.de*

Abstract

This contribution discusses the integration of polymer planar Bragg grating sensors (PPBG) into carbon fiber reinforced polymer (CFRP) components. For the first time, it is shown that PPBGs based on cyclic olefin copolymers can be integrated into commercial-grade composites, thereby withstanding the demanding production processes. Pre-impregnated fibers are stacked and partially modified to form a sensor pocket. Afterwards, the CFRP specimen containing the optical sensor is cured in a heated mechanical press for 2 hours at a pressure of 7 bar and a temperature of 120 °C. A subsequent evaluation of the sensor signal shows a Bragg wavelength shift of 1236 pm and a decline in signal amplitude of -2 dB. Three-point flexural tests of the cured sample reveal a linear behavior of the sensor signal towards external loads. The determined sensitivity in dependence of the CFRP specimen's maximum central deflection is -112 pm/mm, while correlation to the applied force results in a sensitivity of -5 pm/N.

Keywords: Bragg grating sensors, carbon fiber reinforced polymers, cyclic olefin copolymers, structural health monitoring

Introduction

Due to their outstanding mechanical properties - excellent strength combined with low weight and density - the usage of carbon fiber reinforced polymers (CFRP) for lightweight and high-performance applications, such as the aerospace industry, has increased dramatically over the last few years. However, in comparison to classical materials, composites tend to behave inherently different in excess load situations. In the case of metals, applied forces will cause elastic deformations which will then turn to plastic distortions when crossing the material's yield point. A further load increase will ultimately lead to cracks, fractures and the apparent breakdown of a metal workpiece. Predicting the failure process of CFRPs on the other hand is a challenging and elaborate task since even microscopically small fissures, especially at critical points, can cause sudden and unpredictable failure of a CFRP component [1]. Therefore, immense effort has to be invested in the detection of damage and material fatigue. While there are several approaches for live structural health monitoring of composites, most damage inspection methods are based on offline testing procedures. Thus, in order to reduce costs for manufacturers while simultaneously mitigating user safety concerns, alternative technologies enabling detection of structural changes in CFRPs need to be developed.

In recent years, the integration of silica-based fiber Bragg grating (FBG) sensors has proven to be a promising method for in-situ evaluation of composite materials. Albeit, their inherent one-dimensionality needs to be addressed by employing complex sensor structures, e.g. a multicore FBG, in order to enable quantification of multidimensional forces [2,3]. Polymer planar Bragg gratings (PPBG) however, offer up to three-dimensional detection of deformations [4,5]. Additionally, polymer-based optical Bragg grating sensors exhibit increased sensitivity towards stress and strain due to their distinctly low Young's modulus. However, the insufficient glass transition temperature of most PPBG-substrates prevents successful integration into commercial CFRPs, since these polymers cannot withstand harsh fabrication processes.

This challenge can be overcome by employing PPBGs based on cyclic olefin copolymers (COC). Due to outstanding glass transition temperatures of up to 250 °C, this substrate

material is able to outlast temperature measurements up to 160 °C [6,7]. Additionally, COC-PPBGs are inherently unsusceptible towards relative humidity fluctuations [8,9]. Thus, they represent a particularly auspicious sensor concept for CFRP integration.

Fundamentals

In principal, PPBGs are constituted of a polymer substrate featuring a waveguide and a Bragg grating, as depicted in Fig. 1.

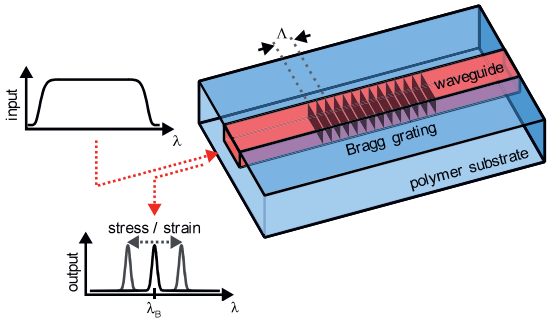


Fig. 1: Schematic of a polymer planar Bragg grating sensor.

Thereby, the Bragg grating is a periodic refractive index modulation which acts as a narrowband reflector for light traversing the waveguide. Its main reflection wavelength, denoted as Bragg wavelength λ_B, is defined by the effective refractive index n_{eff} and the grating period Λ according to Eq. (1).

$$\lambda_B = 2n_{eff}\Lambda$$

External forces will affect Λ and n_{eff} resulting in a proportional shift of λ_B, which can be detected and evaluated.

Waveguide and Bragg grating are generated simultaneously within the COC (TOPAS6017S-04, TOPAS Advanced Polymers) by irradiating the substrate with ultraviolet radiation generated by a KrF excimer laser (BraggStar Industrial, Coherent). A sophisticated discussion of the employed manufacturing method, denoted as single-writing-step procedure, is given by the authors and others elsewhere [10–12]. Fig. 2 shows the fabricated COC-PPBG before its integration into a CFRP-specimen. In order to couple the PPBG to the evaluation unit, a physical connection with a single-mode fiber (SMF) needs to be established. This is achieved by butt-coupling an optical fiber to the waveguide by means of a UV-curable adhesive (NOA 61, Norland). Evaluation of the reflected sensor signal is done by an industrial-grade interrogation unit (HYPERION si155, MICRONOPTICS).

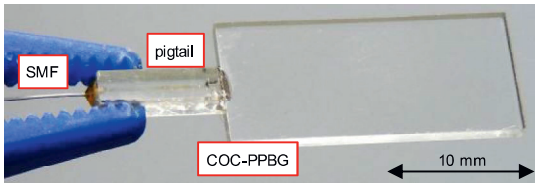


Fig. 2: COC-PPBG butt-coupled to a pigtailed standard single-mode fiber.

Carbon fiber reinforced polymers are a composite material consisting of at least two components, carbon fibers embedded in a polymer matrix (mostly epoxy resin). While the matrix serves as parent material, the fibers are responsible for absorption and redistribution of external forces, which yields components exhibiting outstanding strength-to-weight ratios. The final mechanical properties of CFRPs are thereby highly dependent on the used materials. Additionally, amount and direction of the fibers enclosed within the workpiece plays an important role. This leads to distinct anisotropic behavior of single-layer CFRP components, which is mitigated by stacking layers with alternating fiber orientations on top of each other. For commercial manufacturing of CFRPs, it is common to make use of pre-impregnated composite fibers known as prepregs, a compound of semi-cured epoxy with embedded fiber structures. Multiple prepreg layers are preformed into desired shapes and afterwards fully cured at high temperatures, and under immense pressure, thus chemically crosslinking the matrix system to generate operational CFRP components.

Sample preparation and sensor integration

Nine prepreg layers (P3252S-25, TORAYCA) with a fiber orientation of ± 45° are cut to rectangular shape and stacked on top of each other to form a CFRP specimen. While the outer two layers remain unaltered, the middle section needs to be modified in order to form a sensor pocket. This prevents uneven pressure distributions during the curing procedure and guarantees a flat surface of the specimen. Fig. 3 elucidates the preparation process. After inserting the sensor into its predefined location, the prepreg stack is positioned on a brass foil to prevent the CFRP matrix from sticking to the curing apparatus.

Most commercial manufacturers of CFRP components use autoclaves for the final curing procedure. For simple planar specimens however, it is possible to use a heated press in order to generate workpieces with comparable properties.

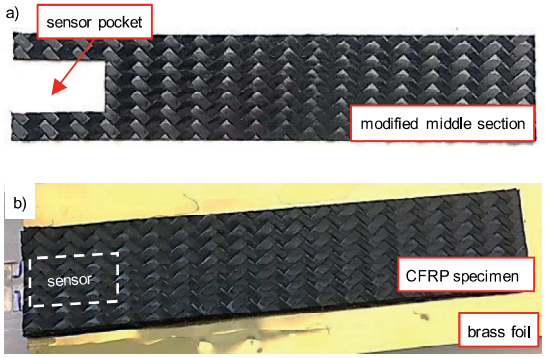


Fig. 3: a) Modified middle-section prepreg and b) CFRP sample with embedded PPBG sensor before curing.

The heated press used for curing the CFRP workpiece, equipped with a PPBG, is depicted in Fig. 4.

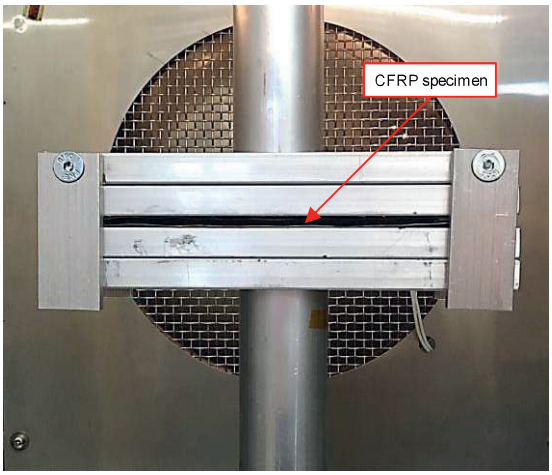


Fig. 4: Curing setup - mechanical press inside air circulated furnace.

The CFRP prepreg stack is placed in-between the plates of the press which are positioned inside an air circulated furnace. Both plates are actively controlled by a tensile test machine (Autograph AG-X 20kN, SHIMADZU). This way, constant pressure is applied to the specimen throughout the whole curing process. An overview on the applied curing parameters is given in Tab. 1.

Tab. 1: Curing parameters

Pressure	7 bar
Temperature	120 °C
Duration	2 h

Removing the specimen from the press after curing reveals that some of the epoxy has leaked from the prepreg stack. An image of the cured sample can be seen in Fig. 5 a), while Fig. 5 b) shows the PPBG signal before and after curing of the sample.

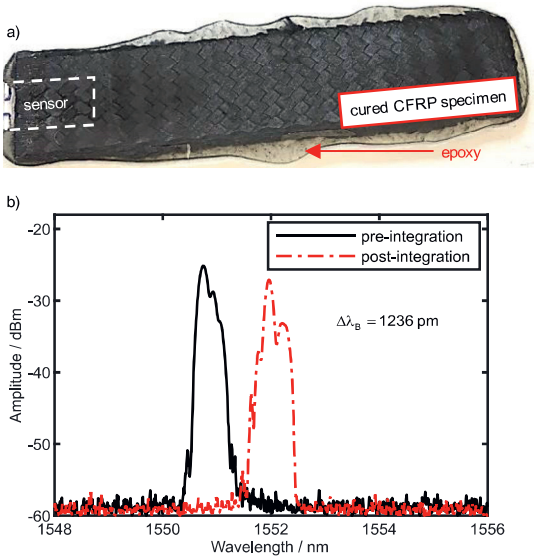


Fig. 5: a) CFRP specimen after curing and b) PPBG sensor signal pre and post curing.

Even after being exposed to the harsh curing process parameters, the sensor delivers a strong and evaluable signal. The findings demonstrate that COC-PPBGs are suitable for integration into commercial composite structures. However, the cured CFRP introduces constant force on the embedded sensor, which results in a positive Bragg wavelength shift of 1236 pm and a peak amplitude reduction of -2 dB.
In order to assess the connection of the CFRP prepregs with the integrated sensor, the specimen was examined via computed tomography (CT) scans. After finishing all experiments on the specimen, microsections were used to verify the CT scan. The results of both investigation methods are summarized in Fig. 6.

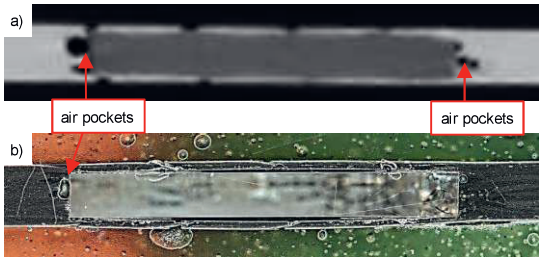


Fig. 6: a) CT scan and b) microsection of CFRP-integrated PPBG.

Both images reveal air pockets in close vicinity to the integrated sensor. Their existence can be correlated to the modification of the prepreg layers prior to embedding the sensor, which introduces air-filled voids within the specimen.

Although while cured, the epoxy becomes viscous, it is not able to completely refill these voids. This circumstance can be mitigated by adjusting the geometry of the PPBG substrate so the need for prepreg modification is reduced or even omitted completely.

Experimental results

After removing the excess epoxy on the outer edges of the test sample, examination of the integrated PPBG's sensing capabilities is done by using a three-point flexural test configuration, as shown in Fig. 7.

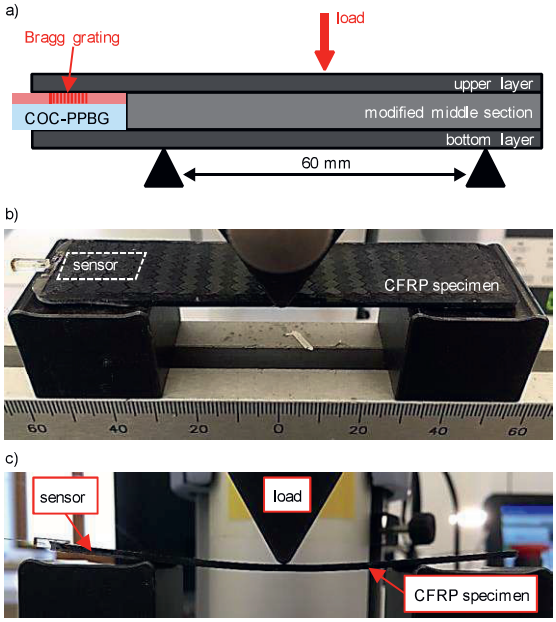


Fig. 7: a) Schematic and b) image of the employed three-point flexural test setup. Section c) shows the bent CFRP specimen at a maximum stroke of 2 mm.

Thereby, the sensor inside the specimen has no contact with the support spans and its minimum distance from the loading point is 40 mm. This configuration enables a defined bending of the specimen while it ensures that the PPBG detects deformation by flexural forces only. The maximum deflection at the center of the sample, denoted as stroke, is actively controlled by a tensile test machine (EZ-L, SHIMADZU). The system also determines the force which is necessary to reach a predefined stroke value. Fig. 8 depicts the temporal evolution of a three-point flexural test cycle. Prior to the experiment, the tensile testing machine applies a pre-load of 3 N to the specimen. Therefore, the stroke curve starts with an offset of 0.35 mm. Afterwards, the stroke is increase with a speed of 8 mm/min until it reaches its maximum at 2 mm.

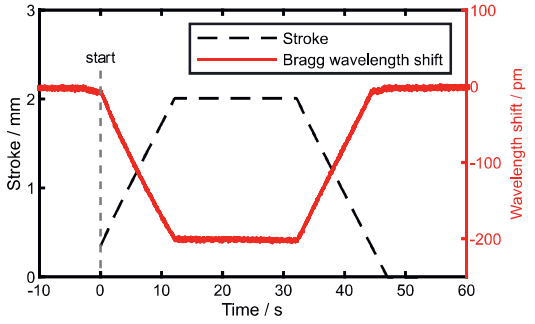


Fig. 8: Temporal evolution of the Bragg wavelength during a three-point flexural test with a maximum stroke of 2 mm.

This value is held constant for 20 s before reducing the load at similar speed until the specimen relaxes completely. Fig. 8 indicates that applying external loads to the workpiece leads to a negative shift of the sensor's Bragg wavelength. Thereby, the absolute wavelength shift can be correlated directly to the specimen's degree of bending. Since the actual sensitive element, the Bragg grating structure, is located close to the surface of the CFRP sample (c.f. Fig. 7), the bending process leads to a reduction of its grating period which in turn results in a shift of the Bragg signal towards shorter wavelengths. The maximum wavelength shift is -205 pm at a stroke of 2 mm. At the end of a cycle, it returns to its original value after removing the load from the sample.

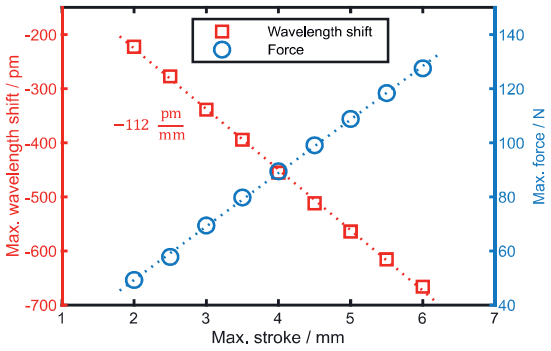


Fig. 9: Maximum wavelength shift and applied force as a function of maximum stroke during a three-point flexural testing cycle.

With increasing maximum stroke, the maximum Bragg wavelength shift follows in a linear fashion, as depicted in Fig. 9. In this configuration and within the examined range, the PPBG exhibits a sensitivity of -112 pm/mm. Thus, depending on the used interrogation device, which offer wavelength resolutions down to 1 pm, it enables detection of minor

displacements of the CFRP specimen [13]. By simultaneously tracking the force which is necessary to deflect the sample to maximum stroke, it is possible to correlate the wavelength shift to the applied load, resulting in a sensitivity of -5 pm/N.

Conclusion

In summary, it was shown that COC-based PPBGs are able to outlast the demanding production parameters of a commercial CFRP component, where the optical sensor is exposed to pressures up to 7 bar and temperatures up to 120 °C for at least 2 hours. To the best of the authors' knowledge, this is the first polymer-based optical sensor which is capable of withstanding an integration process this harsh without exhibiting noticeable losses in signal quality. Three-point flexural tests reveal a linear correlation of the sensor's Bragg wavelength shift with the applied loads. In the examined regime, up to a maximum stroke of 6 mm, the PPBG's sensitivity was determined as -112 pm/mm, or -5 pm/N, respectively. Metallographic investigations of the CFRP specimen show air pockets in close vicinity to the integrated sensor which underlines the demand for optimized sensor substrate designs for future investigations. Based on these results, COC-PPBGs constitute a promising concept for the integration into commercial CFRP products, enabling in-situ structural health monitoring of composites with outstanding sensitivity.

Acknowledgements

This work has been funded by the VDI/VDE-IT on behalf of the Bavarian Ministry of Economic Affairs, Regional Development and Energy.

References

[1] A. Zaman, S. A. Gutub, and M. A. Wafa, *Journal of Reinforced Plastics and Composites* 32, 24, 1966–1988 (2013); doi: 10.1177/0731684413492868.

[2] G. Zhou, and L. M. Sim, *Smart Materials and Structures* 11, 6, 925–939 (2002); doi: 10.1088/0964-1726/11/6/314.

[3] F. Khan, A. Denasi, D. Barrera, J. Madrigal, S. Sales, and S. Misra, *IEEE Sensors Journal*, 1 (2019); doi: 10.1109/jsen.2019.2905010.

[4] M. Rosenberger, W. Eisenbeil, B. Schmauss, and R. Hellmann, *Sensors* 15, 2, 4264–4272 (2015); doi: 10.3390/s150204264.

[5] M. Rosenberger, H. Pauer, M. Girschikofsky, H. Woern, B. Schmauss, and R. Hellmann, *IEEE Photonics Technology Letters* 28, 17, 1898–1901 (2016); doi: 10.1109/LPT.2016.2574889.

[6] G. Khanarian, *Optical Engineering* 40, 6, 1024 (2001); doi: 10.1117/1.1369411.

[7] M. Rosenberger, S. Kefer, M. Girschikofsky, G.-L. Roth, S. Hessler, S. Belle, B. Schmauss, and R. Hellmann, *Optics Letters* 43, 14, 3321–3324 (2018); doi: 10.1364/OL.43.003321.

[8] W. Yuan, L. Khan, D. J. Webb, K. Kalli, H. K. Rasmussen, A. Stefani, and O. Bang, *Optics Express* 19, 20, 19731–19739 (2011); doi: 10.1364/OE.19.019731.

[9] M. Rosenberger, S. Hessler, S. Belle, B. Schmauss, and R. Hellmann, *Sensors and Actuators A: Physical* 221, 148–153 (2015); doi: 10.1016/j.sna.2014.10.040.

[10] M. Rosenberger, G. Koller, S. Belle, B. Schmauss, and R. Hellmann, *Optics Express* 20, 25, 27288–27296 (2012); doi: 10.1364/OE.20.027288.

[11] M. Rosenberger, B. Schmauss, and R. Hellmann, *Optical Materials Express* 6, 6, 2118 (2016); doi: 10.1364/OME.6.002118.

[12] S. Hessler, M. Rosenberger, B. Schmauss, and R. Hellmann, *Optical Materials* 75, 230–235 (2018); doi: 10.1016/j.optmat.2017.10.029.

[13] D. Tosi, *Sensors* 17, 10, 2368 (2017); doi: 10.3390/s17102368.

„Self-Mixing“-Interferometrie P2.2
µKMG P1.20
3D Laserscanner P3.3
3D Printing 3.4.3
3D-AFM P1.20
3D-Integrationsverfahren 3.2.3
3-K-Laser-Doppler-Anemometrie 1.1.3
80x64 pixels 3.1.2
Absorptionsspektroskopie 4.2.1
Absorption Spectroscopy 2.1.4
Accelerators 3.4.3
Acetone P1.9
adhäsive Filmbildung P1.1
Aerosol Deposition Method (ADM P1.15
Akustik P2.3
Akzelerometer 1.3.1
analysis P2.16
Ansichtenplanung 1.1.1
Ansprechzeit 3.3.2
antenna detuning 1.3.2
Apis mellifera P3.10
Asymmetrie 1.3.1
Au-Pd 6.2.5
Au-Pd@meso-Co3O4 ..... 6.2.5
Auswertealgorithmus P3.3
Auswerteverfahren P1.13
autarke Sensorknoten 1.1.4
Autokorrelation P2.2
automatisierte Prüfung 6.3.4
Backside Illumination (BSI) 3.2.3
balanced 6.3.2
Bearing 4.3.3
Bearing faults 5.4.2
berührungslose Gas-
Temperaturmessung 2.4.2
Betriebsverhaltenskenngrößen P2.1
Bezugsnormal 1.1.3
Big Data 4.4.1
Bildverarbeitung P3.7
Bimorph-Effekt 2.3.4
blood collection P2.16
Bragg grating sensors P3.12
Brechungsindex 5.1.4
Bridge-Circuit P2.8
Brillouin-Frequenzverschiebung P2.1
bulk modulus 6.3.1
calibration procedure 5.1.3
Cantilever 6.3.3
Cantilever-Sensor P2.10
carbon fiber reinforced polymers P3.12
CD-Cantilever P1.20
characteristics 2.3.1
Chemical sensors P1.9
chemischer Sensor 4.2.4
Chromatography P3.11
Citizen Science 6.1.3, 6.1.5
classifica-tion performance 5.4.2
CMOS-Matrix 2.2.3
coaxial sensor 1.4.3
colored polymer 6.4.1
concrete 4.1.4
condition monitoring 4.3.4, 5.4.2, 6.3.1
Conductivity measurement 1.4.3
constant phase element 1.4.3
contactless sensor 3.1.3
Critical Dimension P1.20
cyclic olefin copolymers P3.12
Cyclovoltammetrie 6.2.4
DC bias field 5.1.2
De-embedding 3.2.4
defects detection in ceramics 4.1.2
Dehnmessstreifen 2.2.4
Dehungsmessstreifen 1.2.4
design study 1.2.2
Dickschicht P1.15
Dickschichttechnik P1.10
Diesel-Abgas 2.4.4
differential pressure sensor P1.5
Differenzbilder P3.3
Differenzdrucksensor P1.5
Digital Holography 1.1.2
digitale Bildverarbeitung 2.2.3
Digitalisierung P3.4
Dimensionale Messtechnik P3.4
Direct-Laser-Writing P1.14
DMS 1.2.4
domestic environments 2.3.1
Doppler radar 5.3.1
Doppler-Referenz-Verfahren 1.1.3
Drahtlose Sensor/Aktor Netzwerke P2.13
drahtlose Sensorsysteme P2.13
Drahtloses Sensornetzwerk 2.3.3
Drehmomentensensor P3.6
Drehmoment-Referenzaufnehmer 1.2.4
Drehzahlgeber P2.2
Drehzahlmessung P2.2
droplet 3.2.1
Drucksensorik 2.2.1
Dünnschichtstufennormal 5.1.4
Dynamic force 2.2.2
Dynamic force transfer standard 2.2.2
dynamic sensor systems P1.22
Dynamic Time Warping P2.9
dynamisches Sensorprinzip P1.12
Ein-Gradienten-Verfahren P1.16
electrical analysis 2.1.2
electrical impedance spectroscopy 1.4.3
Electrochemical sensors P1.7, P1.8
electromagnetic measurement 1.3.2
electro-thermal piezoresistive
cantilever sensor 3.2.4
Elektrochemische Messverfahren 6.1.3
Elektromagnetische Prüfverfahren 4.3.2
EMFC balance 6.4.5
Emission 2.1.1
Energy harvesting 2.3.1, 2.3.2, P2.15
engine oil dilution 6.3.1
Entfaltung P1.18

equivalent circuit 1.4.3
error correction 6.4.5
Ethanol P1.9
Ethen 4.2.3
Evaluierung 2.4.3
exhaust gas sensor P1.23
Farbmessung P2.4
Faserkunststoffverbunde 4.1.3
Faseroptischer Sensor P2.1
Fault Detection 4.3.3
Feature Engineering 4.3.3
Feature Extraction P2.9
feed-forward neural network 5.4.2
Feinstaub 2.1.1
Fertigungsmesstechnik 6.4.2, 6.4.3
Festelektrolyt-Detektor 3.3.1
Feuchte P2.4
fibre orientation 4.1.4
field application 5.4.4
finite element method P2.7
flow-through 6.3.2
Fluoreszenz P2.5
Fluoreszenzmesstechnik P2.5
Formic Acid P3.10
Formmodellierung 6.4.3
FPGA 4.4.3
Fruchtlager 4.2.3
FTIR-Spektrometrie 2.4.2
Füllstandmesstechnik P2.3
gas chromatography 3.3.4
Gassensor ..... 2.1.4, 4.2.4, 5.2.3, ... 6.2.1, P1.12, P1.22
Gasgemisch 6.3.3
Gasqualität 5.2.2
Gassensoren 6.2.3, 6.2.4
Gassensorik ..... 2.1.3, 3.1.1, 3.2.2, ... 4.2.1, 4.2.3, 5.2.4, 6.1.5, ... 6.3.3, P1.13
Gas-Transmissions-Spektren 2.4.2
gear wheel 6.4.1, P2.18
gedruckte Elektronik 1.3.3
Geruch 6.2.1
GPU 1.1.1, 1.1.2
Graphene Oxide (GO) P1.9
Grobblech 5.1.1
Haptik P1.19
haptische Interaktion P1.19
Hardspot 5.1.1
hardware-in-the-loop P3.8
harsh environment P1.5
Hochdruckprüfstand 5.2.2
Hochtemperaturtest 2.4.3
Hot-plate Emitter 4.2.1
hot spot temperature measurement 4.3.4
Human Robot Collaboration 1.3.4
hydrocarbon sensor P1.23
Hydrogel 2.3.4
hydrogen sensor system 3.3.4
Hysterese 2.3.4
ice and water detection 3.1.3
IC-Fassung 2.4.3
Impactschäden 4.1.3
Impedanz 2.4.4
Impedanzspektroskopie 6.2.3
InAs/GaSb-Übergitter 3.1.4
Industrie 4.0 ..... 4.4.1, 5.4.1, P3.4
Industrielle Bauteilreinigung 1.4.1
industrielles IoT 1.1.4
Infrared P2.6
infrared detection 3.1.2
infrared light sensor 3.1.3
Infrarot 5.2.2
Infrarotdetektor 3.1.4
Infrarotemitter 3.1.1
Infrarotspektroskopie 1.4.1
Inhomogenität P1.16
Inline Inspection 1.1.2
in-process measurement 5.4.4
In-Prozess-Messung 4.1.1
Instandhaltung 5.4.3
Instrumentierte Eindringhärte P2.12
interdigital transducers P2.7
Interferometer 2.2.3
Interferometrie 5.1.4
Internationales Einheitensystem .... Festvortrag
Internet of Things 6.1.5
Ionensensitive Elektroden P1.10
IR-Aktivität 5.2.2
IR-Strahler 4.2.1
Iron-oxide nanoparticles P1.7
Isopropanol P1.9
Isotopic Composition Analysis 2.1.4
Kältemittel 4.2.2
kapazitive Füllstandmessung P1.1
kapazitive Füllstandschalter P1.1
Katalytischer Sensor 6.2.5
keramische Multilayertechnologie P1.10
Kibble balance 6.4.5
Kleinstlebewesen 1.3.3
konfokale Sensorik P1.14
Kontaktresonanzspektroskopie P2.10
Körperschallsensor 5.4.3
Korrelation P2.13
Kösters-Prisma 2.2.3
Kraftsensor P2.12
Kugeldetektion 4.4.2
Kunststoffe P2.5
Kunststoffschmelze P2.4
Lagebilderfassung 2.3.3
Lambda 2.4.4
Landmarkenerkennung 4.3.1
Laser speckle photometry 4.1.2
Laserlichtschnitt 1.1.1
Layer-by-Layer Deposition P1.7, P1.9
Lebensmittelüberwachung P1.13
leistungslos 2.3.4
LFA P2.16
Light scribe Graphene P1.7

Low Hydrogen Concentration measurement sensor P1.11
Low-Cost-Füllstandssensor P2.15
Low-resolution infrared camera 4.3.4
Luftultraschall 4.3.1
Lumineszenz 5.2.3
Luminophor 5.2.3
Machine Learning 4.3.3
Machine Tool 1.1.2
magnetic particles 3.2.1
Magnetische Sensoren 3.4.4
Magnetometer 3.4.3
magnetoresistive Sensoren 3.4.4
Magnets 3.4.3
maritime Sensorsysteme 5.2.4
Maschinelles Lernen ..... 2. Plenarvortrag, ... 4.3.2, 4.4.1, 5.1.1, 5.4.1, 5.4.3, P3.7
maschinelles Lernen ..... 2. Plenarvortrag
maskenlose Lithografie P1.14
Masking P2.9
mass 3.2.1
Maßprüfung P3.3
Material parameters 5.4.4
Materialcharakterisierung 4.3.2, 5.1.1
MCT 3.1.4
measurement uncertainty 1.2.2, 6.4.1
mechanical measure 1.2.2
mechanical simulation 6.4.5
mechanische Spannung 2.2.4
medical device P2.16
Medienresistenz P1.5
medium resolution 3.1.2
medium voltage switchgear 4.3.4
Membranstrahler 4.2.1
membranbasierter isotroper Mikrotaster 6.4.2
MEMS 4.2.1
MEMS- & Nanosensoren P1.21
MEMS Sensoren 3.2.2
MEMS-Wasserstoffsensor 3.3.2
mesoporös 6.2.5
Messdatenauswertung P2.10
messtechnische Rückführung 1.1.3
Messunsicherheit 6.4.3, P3.6
Messunsicherheitsbeitrag P1.16
Metalloxid 4.2.4
Metal-organic frameworks 6.2.3
Metal-Oxide Gas Sensor Array P3.10
Meterkonvention ..... Festvortrag
Methan 4.2.1, 6.2.5
methane sensitivity P1.23
Methanol P1.9
Metrologie ..... Festvortrag
Microcantilever 3.2.1
Microhotplate 3.2.2
Mikro-environment 6.4.2
Mikroinnenverzahnungsnormal 6.4.2
Mikroschalter 2.3.4
Mikrotaster P2.10
Mikroverzahnung 6.4.2
Mikrowellen-Defektoskopie P1.18
miniaturized hydrogen sensor and mobile application P1.11
model fit 5.3.3
Modellbildung 4.2.4, P1.21
Modellierung P2.10
Modenverwirbelungskammer P2.13
Modulbaukasten 4.4.3
Monocrystalline Silicon P2.6
monotonic phase response 3.2.4
Montage P3.6
Monte-Carlo-Simulation 6.4.3, 6.4.4
multi-component transducer 1.2.2
multiharmonic sine fit 6.4.4
multivariate Statistik 5.4.1
nacelle test benches 5.1.3
Nano-CMMs P1.14
Nanoindenter P2.12
Nanotaster P1.20
Nass / Nass P1.5
Naturkonstanten ..... Festvortrag
NDIR 4.2.1, 4.2.3
Neural Networks P2.9
NFC 2.3.2
nicht-invasives Messsystem 1.3.3
Niedertemperaturnachweis 6.2.5
Nitrat 6.1.3
Nitrite detection P1.7, P1.8
noncontact monitoring 1.3.2
Non-destructive testing 4.1.2, P2.6
non-linear material parameters 5.1.2
NOx-Sensor 2.4.4, P1.12
NTC Thermistor P1.15
Ökosystem für Internet of Things ..... 2. Plenarvortrag
optical fiber sensors 2.1.2
optical measurement method 3.1.3
Optical Measuring System 2.1.4
Optical sensor system 4.1.2
Optik 4.2.3
Optimization P3.11
Optisch stimulierte Elektronenemission (OSEE) 1.4.1
Optische Lokalisierung 1.1.4
optische Messverfahren 2.1.1, 2.1.3, 3.1.1
Optische Sensoren 3.2.3
optischer Wasserstoffsensor 3.3.2
Optisches Messverfahren 4.2.4
ortsaufgelöste Brillouin-Sensorik P2.1
OT-Security ..... 2. Plenarvortrag
Packaging 3.1.1
Palladium 3.3.2
Palladium-MEMS-Sensor 3.3.1
Parallel Computing 1.1.2
parameter extractio 5.3.3
PAS-Sensor 4.2.1
PCB 3.4.3
Pellistor 6.2.5

Peltierkühler 3.1.4
Pharmacopeia P3.11
phase-locked loop 3.2.4
Phasenshift 1.1.1
Phasenübergang 3.2.2
Photoakustische Detektoren 4.2.2
photoakustischer Sensor 5.2.4
photoakustische Spektroskopie 4.2.1
Photobleichung 5.2.3
Photocatalytic effect P1.8
Photoelasticity P2.6
Photometer 4.2.2
piezoelectric materials 5.1.2
piezoelectric properties 5.1.2
piezoelectric vibration transducer 5.4.2
piezoresistiv 2.2.4, 6.3.3, P2.12
piezoresistive Dünnschicht 2.2.1
Planar capacitive sensors P2.7
Planck-Balance 6.4.4
plethysmography 1.3.2
PM10 ..... 2.1.1
Polymerabbau P2.4
Positionsmessung 3.4.4
Präzisionsbrückenverstärker 1.2.4
Präzisionsmessverstärker 1.2.4
process control P3.8
programmable reference 3.2.4
Prozessüberwachung 2.1.1
Pt | YSZ P1.12
Pt-Elektroden 6.2.4
Pt-YSZ P1.22
Punkt-Matrix-Schrift P3.7
Punktspreizfunktion P1.18
Quantifizierung 6.2.1
quantitative measurement 5.3.3
Quantum Cascade Laser 2.1.4
Querempfindlichkeit 3.3.2
R1234yf 4.2.2
R134a 4.2.2
Radio-Acoustic-Sounding System 5.3.1
radio-over-fiber 2.1.2
Raman-Sensor 3.3.1
RASS 5.3.1
Rauschunterdrückung P2.2
Rayleigh scattering 5.3.3
reduced Graphene Oxide (rGO) P1.7, P1.9
Regression 4.3.2
Rekonstruktion P1.18
renewable energy system P3.8
resistance change P2.8
resonant 6.3.2
resonant frequency 3.2.1
resonant sensor 6.3.1
Resonanz 4.1.3
respiration 1.3.2
Restschmutzmessung 1.4.1
RF modulated laser 2.1.2
RFID 1.3.3
road condition sensor 3.1.3
Roboter P3.6
Roboternavigation 4.3.1
RoHS 3.1.4
Safety 1.3.4
Safety monitoring 3.3.4
sakrale 1.3.1
sampling P2.16
Sampling Rate P3.11
Sauberkeitskontrolle 1.4.1
Sauerstoffionenleitfähigkeit 2.2.1
scanning mode P2.18
Schleifprozess 4.1.1
Schmiermittel Beruforge 152DL 5.1.4
Schriftbildkontrolle P3.7
Schülerinnen und Schüler 6.1.5
Schüttgut P2.3
Schwefelverbindungen 6.2.1
Schwingungsmessung P2.14
Screen printed electrodes (SPE) P1.8
Self-X Funktion 4.4.2
Sensor Placement 1.3.4
Sensorik 2.4.3
Sensornetzwerk 1.1.4
Sensorschnittstelle 4.4.3
Sensorsystem P1.13
Signal Filtration P3.11
Signalverarbeitung 4.4.3
Silizium 2.2.4, P2.12
Silo P2.3
Simulation 1.1.1
Simulation und Validierung P1.21
single point precision P2.18
skapulare Beschleunigung 1.3.1
Smart Gardening P2.15
Smart Home P2.15
SO2-Sensorik 5.2.4
Soft-Sensoren ..... 2. Plenarvortrag
solid-electrolyte P1.22
solid-electrolyte coulometry 3.3.4
solid-state mixed potential
electrochemical sensors P1.11
Sonarmessung 4.3.1
SPAD 3.2.3
Spannungs-Puls-Methode P1.12
Speckle-Fotografie 4.1.1
Spektrometer 4.2.3
Spektroskopie P2.4
Spillover-H2 ..... 3.3.1
Stahlbaustrukturen P3.3
stationäre Gasturbinen 2.4.2
steel fibres 4.1.4
Stoßdämpfung 1.3.1
Strahlungs-Thermometrie 2.4.2
Strain/stress characterization 4.1.2
Strömungsanalyse 1.1.3
Strömungssensoren 6.3.3
structural health monitoring P3.12
structured-light scanning 6.4.1
Synchronisation 5.4.1

tactile coordinate measurement P2.18
Taktiles Display P1.19
Temperature P2.8
Temperaturkalibrierung 3.2.2
Temperatursensor P1.15
Temperaturzyklus 6.2.1
Test P2.13
Testmethodik 2.4.3
thermoelectric gas sensor P1.23
Thermoelement P1.16
Thermopile arra 3.1.2
Thermopiles 3.1.1, P1.23
thermozyklische Betriebsweise P1.13
three-parameter sine fit 6.4.4
tibiale 1.3.1
tilt measurement 6.4.5
Titanium oxide nanoparticles P1.8
Torque calibration 5.1.3
torque transfer standard 5.1.3
torsional oscillation 6.3.2
Trägerfrequenzmesstechnik 2.2.1
true shutterless 3.1.2
U-Bahntunnel 2.3.3
UHPC 4.1.4
Ultraschall 5.4.3, 6.3.4
Ultraschall Thermografie 4.1.3
ultrasonic method 4.1.4
Ultrasound analysis 5.3.3
Umwelt P2.5
Umweltstudien 6.1.5
Uncertainty 2.2.2
Urban Gardening 6.1.3
Varroa destructor P3.10
vector network analyzer 5.4.4
vehicle safety systems P2.7
Verformungs- und Dehnungsmessung 4.1.1
verteilte faseroptische Temperaturmessung P2.1
Verwertung 6.1.3
Vibration Analysis 4.3.3
vibration sources 2.3.1
Vibrationsharvester 2.3.3
virtual comissioning P3.8
viscosity 6.3.2
Vitalparameter 1.3.3
VOC-Messung 6.2.3
Volatile Organic Compound (VOC) P1.9
Vollständigkeitsprüfung P3.3
vorausschauende Instandhaltung 4.4.1
Wafer-to-Wafer Bonding 3.2.3
Wasserstoff (H2) 5.2.2, 6.2.4
Wasserstoff-Gassensor 3.3.1
Wavelet-Transformation P2.3
Wegmessung 3.4.4
wet / wet P1.5
wetness detection P2.7
Wheatstone-Brücke 1.2.4
wind turbine test bench 1.2.2
Winkelmessung 3.4.4
Winkelsensor 2.2.3
Wireless 1.1.2, 2.1.2
wireless sensor nodes 2.3.1, P2.15
work piece deviations 6.4.1
xMR-Sensor 4.4.2
yttria-stabilized zirconia (YSZ) 3.3.4
Yttriumoxid-stabilisiertes Zirconiumdioxid (YSZ) 6.2.4
Zahnriemen P2.14
Zeitumkehr-Fokussierung P1.19
zerstörungsfreie Bauteilprüfung 6.3.4
Zerstörungsfreie Prüfung 4.3.2
Zerstörungsfreie Prüfverfahren 5.1.1
zerstörungsfreie Werkstoffprüfung P1.18
Zielerkennung 4.3.1
Zinkoxid 5.2.3
Zirkonoxid 2.2.1
Zugangskontrolle 2.3.2
Zustandsüberwachung ..... 2.2.4, ... 4.4.1, 4.4.2, 5.4.1, P2.14, P3.7
Zwei-Gradienten-Verfahren P1.16