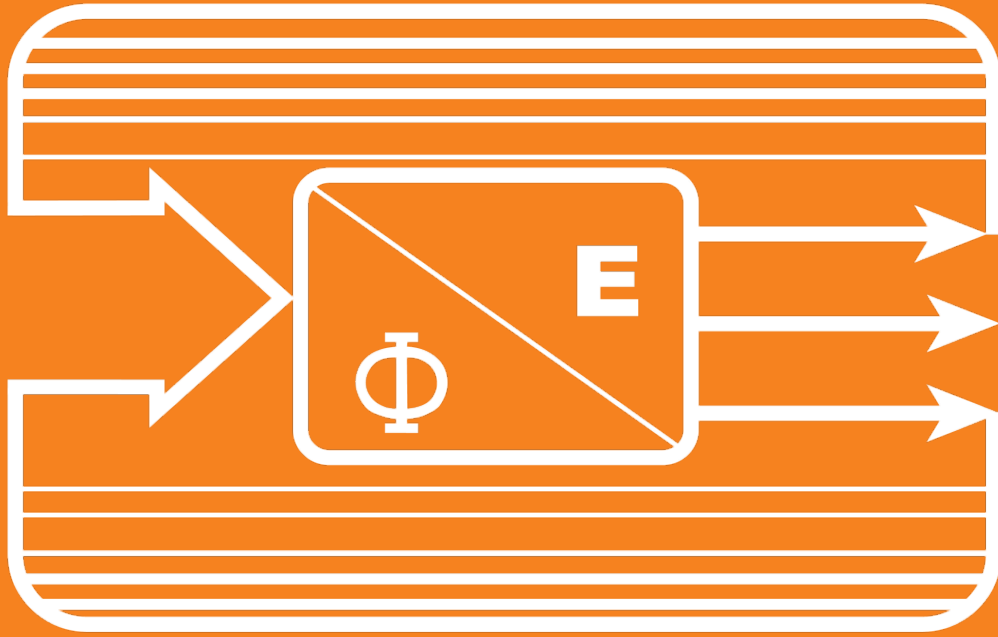


Tagungsband



Sensoren und Messsysteme 2026

23. ITG/GMA-Fachtagung

09. – 10. Juni 2026,
Nürnberger CongressCenter

www.sensoren2026.de

VDE ITG

Tagungsband

23. ITG/GMA-Fachtagung Sensoren und Messsysteme 2026

Bei diesem Band handelt es sich um den Kongressband der
23. ITG/GMA-Fachtagung Sensoren und Messsysteme 2026.

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



ISBN 978-3-910600-11-9

© AMA Service GmbH, 2026

Die 23. ITG/GMA Fachtagung „Sensoren und Messsysteme“ findet am 9. und 10. Juni 2026 parallel zur Messe Sensor+Test im Nürnberg Congress Center statt. Die Veranstaltung hat sich seit ihrer Begründung 1982 als zentrales deutschsprachiges Forum für den fachlichen und interdisziplinären Austausch im Bereich Sensorik und Messtechnik etabliert.

In diesem Jahr übernimmt Prof. Stefan Rupitsch, Lehrstuhl für Elektrische Mess- und Prüfverfahren am IMTEK, Universität Freiburg, die wissenschaftliche Leitung der Tagung und folgt damit auf Prof. Reindl. Wir möchten Prof. Reindl für seine jahrelange Arbeit danken und Prof. Rupitsch herzlich willkommen heißen.

Das Programm der Tagung bildet die breite Entwicklung der Sensorik und Messsysteme ab. Thematische Schwerpunkte sind akustische und resonante Sensoren, optische und infrarote Sensorsysteme sowie klassische Themen wie die Zustandsüberwachung technischer Systeme, die Messung mechanischer Größen und Gassensorik. Ergänzend gibt es Sessions zu aktuellen Spezialthemen: Positions-, Schwingungs- und Akustikmessung mit Drohnen, Event-basierte Bildverarbeitung („Event-based Vision“) sowie einen Vortragsblock zum Sonderforschungsbereich „EcoSense“ der Universität Freiburg.

Die Plenarvorträge decken zentrale aktuelle Entwicklungen im Bereich Sensoren und Messsysteme ab. Katrin Gajo (IST AG) beschäftigt sich mit kapazitiven Feuchtesensoren im industriellen Einsatz und beleuchtet dabei aktuelle Herausforderungen sowie mögliche Innovationspfade. Gerhard Lammel (Bosch Sensortec) stellt dar, wie ultrakleine MEMS-Sensoren dazu beitragen, Alltagsaufgaben zu vereinfachen und die Sicherheit in unterschiedlichen Anwendungsfeldern zu erhöhen. Svenja Knappe von der University of Colorado Boulder gibt Einblicke in die funktionelle Bildgebung durch Quantengasmagnetometer und verdeutlicht, wie mit dieser Technologie neue Möglichkeiten in der nicht-invasiven Analyse von Gehirnaktivitäten entstehen. Stefan Seelecke (Universität des Saarlandes) zeigt, wie dielektrische Elastomer-Sensoren aus der wissenschaftlichen Forschung heraus in die praktische Anwendung geführt werden und damit neue Wege für flexible und vielseitig einsetzbare Sensorsysteme eröffnen.

Die Tagung wird von der Informationstechnischen Gesellschaft im VDE (ITG) und der VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA) gemeinsam getragen. Wir danken allen Autorinnen und Autoren für die eingereichten Beiträge sowie allen, die zum Gelingen der Veranstaltung beitragen.



Prof. Jürgen Wöllenstein
Fraunhofer IPM und Universität Freiburg
IMTEK



Prof. Stefan J. Rupitsch
Universität Freiburg, IMTEK

Kongresskomitee

Wissenschaftliche Tagungsleiter

Prof. Dr. J. Wöllenstein

Prof. Dr. S. Rupitsch

Albert-Ludwigs-Universität Freiburg und Fraunhofer IPM

Albert-Ludwigs-Universität Freiburg

Programmausschuss

Prof. Dr.-Ing. J. Auge

Dr.-Ing. M. Bartholmai

Univ.-Prof.

Dr. H. Bosse

Prof. Dr. Ing. J. Czarske

Dr.-Ing. M. Ebermann

Dr. S. Eichstädt

Prof. Dr.-Ing. G. Fischerauer

Prof. Dr. H. Fritze

Prof. Dr.-Ing. T. Fröhlich

Prof. Dr. M. Gerken

Prof. Dr.-Ing. G. Gerlach

Prof. Dr.-Ing. M. Heizmann

Prof. Dr.-Ing. B. Henning

Dr.-Ing. D. Imkamp

Prof. Dr. B. Jakoby

Prof. Dr. Y. Joseph

Dr. U. Kaiser

Prof. Dr.-Ing. O. Kanoun

Prof. Dr.-Ing. A. König

Prof. Dr. rer. M. Kreutzbruck

Prof. Dr.-Ing. P. Lehmann

Prof. Dr. P. Lieberzeit

Dr. V. Migunov

Prof. Dr.-Ing. R. Moos

Dr.-Ing. U. Neuschaefer-Rube

Prof. Dr.-Ing. C. Rembe

Prof. T. Sauerwald

Dr. A. Schäfer

Dr. B. Scherer

Univ. Prof. Dr. U. Schmid

Dr. K. Schmitt

Univ.-Prof. Dr.-Ing. G. Scholl

Prof. Dr. G. Schrag

Prof. Dr. A. Schütze

Prof. Dr. G. Sessler

Dr. E. Starke

Prof. U. Steinmann

Prof. Dr.-Ing. M. Vellekoop

Prof. Dr.-Ing. T. Wendt

Prof. Dr.-Ing. Dr. B. Zagar

Prof. Dr.-Ing. S. Zimmermann

Dr. J. Zosel

Hochschule Magdeburg-Stendal

BAM Bundesanstalt für Materialforschung

Technische Universität Graz

Physikalisch-Technische Bundesanstalt

Technische Universität Dresden

InfraTec GmbH

Physikalisch-Technische Bundesanstalt

Universität Bayreuth

Technische Universität Clausthal

Technische Universität Ilmenau

Christian-Albrechts-Universität Kiel

Technische Universität Dresden

Karlsruher Institut für Technologie KIT

Universität Paderborn

Carl Zeiss Industrielle Messtechnik GmbH

Johannes-Kepler-Universität Linz

Techn. Universität Bergakademie Freiberg

Endress + Hauser

Technische Universität Chemnitz

RPTU Kaiserslautern-Landau

Universität Stuttgart

Universität Kassel

Universität Wien

ABB Corporate Research

Universität Bayreuth

Physikalisch-Technische

Technische Universität Clausthal

Fraunhofer-Institut - IVV -

Hottinger Brüel & Kjaer GmbH

Endress+Hauser

Technische Universität Wien

Albert-Ludwigs-Universität Freiburg

Universität der Bundeswehr Hamburg

Technische Universität München

Universität des Saarlandes

Technische Universität Darmstadt

SICK Engineering GmbH

Otto-von-Guericke-Universität Magdeburg

Universität Bremen

Hochschule Offenburg

Johannes-Kepler-Universität Linz

Leibniz Universität Hannover

Kurt-Schwabe-Institut für Mess- und Sensortechnik e.V. Meinsberg

Plenarvorträge

PT1	Wie ultrakleine MEMS Sensoren den Alltag erleichtern und Leben retten	13
	G. Lammel, Bosch Sensortec, Nürnberg	

Special Sessions

S2: Hoch selektiv, schnell und günstig: GC-MOS-Sensorsysteme

S2-2	Störstoffüberwachung beim Kunststoffrecycling mittels Online-Gaschromatografie	18
	W. Reimringer, A. Schütze, Universität des Saarlandes, Saarbrücken, T. Sauerwald, Fraunhofer IVV, Freising	

S3: Innovative Sensor-Intelligence Devices für Anwendungen von morgen

S3-1	Aktuelle Arbeiten zu Inspektions- und Monitoring- Techniken zur Erhaltung der Standsicherheit kritischer Infrastruktur am Beispiel Brückenbauwerke	25
	D. Koster, S. Klein, J. Emge, M. Schäfer, C. Weingard, Fraunhofer IZFP, Saarbrücken	
S3-2	Sensor-Intelligence Devices für die multimodale Datenerfassung in der Industrie 4.0: Anwendungen, Chancen und Potenziale	34
	H. Jacob, T. Müller, J. Oswald, D. Koster, H.-G. Herrmann, Fraunhofer IZFP, Saarbrücken	
S3-3	Konzeption und Implementierung einer automatisierten Prozessführung und -überwachung in Bioreaktoren am Beispiel der Kultivierung von humanen Stammzellen	38
	M. M. Gepp, R. Ruff, P. Woll, G. Aydogan, D. Amelin, T. Knoll, H. Zimmermann, S. Wagner, T. Velten, Fraunhofer IBMT, Sulzbach/Saar	
S3-4	From Source to Feature: Multimodal Data Preparation for Machine Learning	42
	D. Amelin, M. Jochem, G. Aydogan, M. M. Gepp, R. Ruff, Fraunhofer IBMT, Sulzbach	
S3-5	Measurement System for Investigating the Long-Term Stability of Permanent Magnets for Sensor and Actuator Systems	45
	M. Lanz, G. Martinek, T. Kresse, G. Schneider, D. Goll, Aalen University, Aalen, A. Georgiadis, Leuphana University, Lüneburg	

S4: Positions-, Schwingungs- und Akustikmessung mit Hilfe von Drohnen

S4-1	Der Einfluss der Abtasttakt drift auf unterabtastete Laser-Doppler-Vibrometer-Signale	52
	R. D. Sando Fagno, N. Neumann, C. Rembe, Technische Universität Clausthal, Clausthal-Zellerfeld	
S4-2	UAV-basierte Nahfeld-zu-Fernfeld-Transformation zur detaillierten Charakterisierung großer Emissionsquellen unter Betriebsbedingungen	59
	J. Gundlach, R. R. Subramanyam, D. Ulm, C. Gerasch, C. van Aswegen, K. Baaske, S. Jacob, T. Kleine-Ostmann, Physikalisch-Technische Bundesanstalt (PTB), Braunschweig	
S4-3	Effect of UAV Size on EKF-Based Position Uncertainties and the Performance of UAV-Mounted RIS	67
	D. Müller, M. Mönnigmann, Ruhr-Universität Bochum, Bochum, K. Weinberger, A. Sezgin, Ruhr-Universität Bochum, Bochum	
S4-4	Reliable measurements for safe rescue operations after building collapses (ZeMeRGe)	75
	D. M. Spale, G. Gidion, J. Schiffert, S. J. Rupitsch, University of Freiburg, Freiburg, E. Maier, K. Schmitt, Fraunhofer IPM, Freiburg	

S5: Event-Based Vision

S5-1	Real-Time Multi-Particle Tracking at 10 kHz: Leveraging Binary Frames for Event-Based Vision	82
	J. Baßler, M. Heizmann, Karlsruhe Institute of Technology, Karlsruhe, P. Bäcker, Fraunhofer IOSB, Karlsruhe	
S5-2	Characterization of Event-Camera Biases and Operating Conditions for High-Speed Sorting.....	90
	P. Bäcker, G. Maier, T. Längle, Fraunhofer IOSB, Karlsruhe, J. Baßler, M. Heizmann, Karlsruhe Institute of Technology, Karlsruhe	
S5-3	SNR Enhancement in Event-Based Grayscale Image Reconstruction via Pixel Sensitivity Characterization.....	98
	A. L. Delucchi, J. Baßler, M. Heizmann, Karlsruhe Institute of Technology, Karlsruhe, P. Bäcker, Fraunhofer IOSB, Karlsruhe	
S5-4	Data-Efficient Optical 4D Measurements of Moving Objects with Stereo Event Cameras and Tailored Structured Illumination	105
	C. Freitag, A. T. Khusainova, S. Heist, P. Kühmstedt, G. Notni, Fraunhofer IOF, Jena, A. Stark, C. Franke, Friedrich-Schiller-Universität Jena, Jena	

S6: Eco Sense

S6-1	Entwicklung einer neuartigen photoakustischen Boden- sonde für Kohlenstoffdioxid	113
	J. Kapp, C. Weber, J. Wöllenstein, Universität Freiburg, Freiburg, J. Brozon, H. Schack-Kirchner, Universität Freiburg, Freiburg	
S6-4	Resonanter Photoakustik-Sensor für die Detektion von Isopren im mittleren Infrarotbereich	119
	S. Schnell, C. Hild, H. Lee, K. Schmitt, J. Wöllenstein, Universität Freiburg, Freiburg, S. Hugger, Fraunhofer IAF, Freiburg	

Vorträge**D1-1: Zustandsüberwachung**

D1-1.1	System zum störungsreduzierten Messen von Struktur- schwingungen mithilfe eines Laser-Doppler-Vibrometers und drohnengetragenen Spiegels	127
	R. Zimmermann, U. Durak, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Braunschweig, F. Wang, C. Rembe, Technische Universität Clausthal, Clausthal-Zellerfeld	
D1-1.2	Voruntersuchung zum Einsatz von kostengünstigen MEMS- Beschleunigungssensoren für die kontinuierliche Bestimmung der Lageänderung von Brückenbauwerken	135
	D. Thiem, A. Hennig, Hochschule Ruhr West, Mülheim an der Ruhr, G. Schiele, Universität Duisburg-Essen, Duisburg, M. Krückhans, Stadt Essen, Essen	
D1-1.3	Miniaturisierter Fluoreszenzsensor auf Ru(dpp)₃²⁺-Basis zur Echtzeitbestimmung der Atemsauerstoffkonzentration	143
	J. Steinebrunner, C. Weber, M. Willaredt, K. Schmitt, J. Wöllenstein, Universität Freiburg, Freiburg	

D1-2: Kraft-, Drehmoment-, und Druck

D1-2.1	Differential Torque Measurement Concept for Condition Monitoring of a Cycloidal Gearbox in a Flight Control Actuator	149
	M. Schmidt, DLR, Braunschweig	

D1-2.2	Joint Simulation–Measurement Framework for Estimation and Compensation of Distributed Parasitics in a Flexible PCB Sensor Matrix	157
	X. Ma, A. Fischerauer, G. Fischerauer, University of Bayreuth, Bayreuth, G. Wu, O. Kanoun, Technische Universität Chemnitz, Chemnitz	
D1-2.3	Integration siliziumbasierter Dehnmessstreifen in Keramik zur Realisierung von Nass-Nass Differenzdrucksensoren	163
	C. Kleinholz, T. Frank, A. Cyriax, S. Jagomast, C. Maier, M. Hintz, T. Ortlepp, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt, A. Schroeter, U. Krieger, VIA electronic GmbH, Hermsdorf	
D1-3: Gassensoren		
D1-3.1	Temperature-Compensated Hydrogen Sensing with In-Fiber Dual-FBG Fabry-Pérot Structures and Palladium-Based Thin Films.....	168
	A. Braunmüller, F. Buchfellner, J. Roths, Munich University of Applied Sciences, Munich, M. Becherer, Technical University of Munich, Munich	
D1-3.2	Qualifizierung eines Gas-Sensors für den Einsatz von VOC- Messungen bei Abluftreinigungen.....	174
	A. Braun, R. Moos, G. Hagen, Universität Bayreuth, Bayreuth, H. Sauter, PlasmaAir AG, Weil der Stadt	
D1-3.3	Charakterisierung und Anwendung von additiv gefertigten Metallkomponenten in der Gasmesstechnik.....	177
	M. Stopp, C. Bur, Universität des Saarlandes, Saarbrücken, O. Maurer, D. Bähre, Universität des Saarlandes, Saarbrücken	
D2-1: Akustische und resonante Sensoren		
D2-1.1	Ultraschall-Messmethode zur Analyse des Heilungsfortschritts von Knochenfrakturen.....	185
	E. N. Link, J. Lützelberger, K. S. Drese, Hochschule für angewandte Wissenschaften Coburg, Coburg	
D2-1.2	Ultraschall-Durchflussmessung mit geführten akustischen Wellen: Der Einsatz von Zylindermoden in Rohrsystemen	192
	A. Backer, K. S. Drese, Hochschule für angewandte Wissenschaften Coburg, Coburg	
D2-1.3	Einzel- und Doppelmikrofonmessungen in einem Helmholtz-Resonator zur Bestimmung der CO₂-Konzentration in trockener Stickstoffatmosphäre	197
	S. Kindorf, C. Weber, M. Benkendorf, B. Bierer, J. Wöllenstein, F. Jost, Fraunhofer IPM, Freiburg, M. Eberl, Infineon AG, Neubiberg	
D2-1.4	Triboakustischer Pin: ein neuer Sensor für die in-situ Schmierfilmüberwachung in tribologischen Kontakten	201
	M. Schirru, F. Tatzgern, M. Varga, AC2T research GmbH, Wiener Neustadt (Österreich)	
D2-2: Analytische Messsysteme		
D2-2.1	Compact Ion Mobility Spectrometer with Two Parallel Drift Tubes for Simultaneous Detection of Positive and Negative Ions.....	208
	T. Ostermeier, T. Jähner, A. Nitschke, M. Hitzemann, S. Zimmermann, Leibniz University Hannover, Hannover	
D2-2.2	Coated Blade Spray Ion Mobility Spectrometry for Medical Applications	210
	S. Deicke, C. Thoben, T. Ostermeier, F. A. Reddner, S. Zimmermann, Leibniz University Hannover, Hannover	
D2-2.3	Entwicklung und Evaluierung kosteneffizienter Peak- Detektor-Schaltungen für das Online-Monitoring von Teilentladungen an Mittelspannungskabeln.....	212
	F. Grimm, F. Kolodziej, F. Wein, C. Andres, U. Steinmann, Otto-von-Guericke-Universität Magdeburg, Magdeburg	
D2-2.4	Nullfeld-Magnetfeldkamera	220
	R. Rasser, P. A. Koss, K. Buse, Fraunhofer IPM, Freiburg im Breisgau, S. Knappe, University of Colorado, Boulder (USA)	

D2-3: Biosensoren

- D2-3.2 Identification of Bacteria by Pyrolysis-Gas Chromatography-Ion Mobility Spectrometry** 224
D. Röckrath, M. Hitzemann, A. Höhne, P. Timmermann, S. Zimmermann, Leibniz University Hannover, Hannover, A. Vargas-Valderrama, Royal Military Academy, Brussels (Belgium)
- D2-3.3 Label-Free Electrochemical Biosensor for Real-Time Monitoring of Erythropoiesis via Glycophorin A (CD235a) Detection** 226
S. M. Hesari, L. Deckert, U. Steinmann, Otto-von-Guericke-Universität Magdeburg, Magdeburg, C. von Heckel, J. N. B. Rios, Otto-von-Guericke-Universität Magdeburg, Magdeburg
- D2-3.4 Modular, 3D-Printed Organ-on-Chip Platform with Integrated Optical Sensing for Real-Time Monitoring of the Tumor Microenvironment** 233
M. Wilfred, L. Deckert, U. Steinmann, Otto-von-Guericke-Universität Magdeburg, Magdeburg

D3-1: Temperatursensoren

- D3-1.3 Ultradünne Lithiumtantalat-Chips mit Absorptionsschicht für hochempfindliche pyroelektrische Infrarotsensoren** 240
A. Eydam, G. Gerlach, Technische Universität Dresden, Dresden
- D3-1.4 Verbesserte Schätzung der Temperaturverteilung in gasdurchströmten Festbetten mittels elektrischer Kapazitätstomografie** 246
M. Weiss, G. Fischerauer, Universität Bayreuth, Bayreuth, A. Jess, Universität Bayreuth, Bayreuth
- D3-1.5 Hochtemperatursensoren auf Basis von Keramik- Kompositwerkstoffen** 254
B. Feng, H.-P. Martin, P. Gierth, L. Rebenklau, A. Michaelis, Fraunhofer IKTS, Dresden

D3-2: Messunsicherheit und Selbstvalidierung

- D3-2.1 Representing Measurement Uncertainty by Gaussian Mixtures** 258
L. Hoffmann, M. Y. Alayoubi, M. Heizmann, Karlsruhe Institute of Technology, Karlsruhe, A. Darijani, J. Beyerer, Fraunhofer IOSB, Karlsruhe
- D3-2.2 Towards Hardware Implementation of Homogeneous Multi-Sensor Self-X System through Simulation-Based Virtual Sensor Modeling** 264
E. Gerken, A. König, University of Kaiserslautern-Landau, Kaiserslautern
- D3-2.3 Kalibrierungsfreier Stromsensor für Hochspannungsfreileitungen** 270
P. Arnold, C. Brockmann, H. Pötter, U. Ganesh, Fraunhofer IZM, Berlin
- D3-2.4 Automatic Sensor Data Fusion for the Detection of Bone Layer Transitions and Channel Length Measurement in Axial Oscillation Bone Drilling** 278
A. Boaron, D. Stöbener, A. Fischer, University of Bremen, Bremen

D3-3: Optische und akustische Messverfahren

- D3-3.1 Quantitative Analyse von Störungen der Periodizität in strukturierten Materialien mittels breitbandiger geführter Wellen** 285
M. Wippermann, J. Hölscher, L. Claes, Universität Paderborn, Paderborn
- D3-3.2 Improvements of the Accuracy of Generic Extrinsic Camera Calibration with ICP** 291
C. Schmerbeck, L. Dau, M. Heizmann, Karlsruhe Institute of Technology, Karlsruhe
- D3-3.3 Vision Guided Laser Tracking System for Dynamic Vibrometry using a Scanning Laser Doppler Vibrometer and ArUco Marker** 297
A. Talukder, C. Rembe, Technische Universität Clausthal, Clausthal-Zellerfeld
- D3-3.4 Robuste Befestigung von Ultraschallwandlern: Prüfstrategien für kryogene Einsatzbedingungen** 303
D. Pollmann, A. Backer, K. S. Drese, Hochschule für angewandte Wissenschaften Coburg, Coburg

D3-3.5	Integrated photonic and electronic building blocks for NV-diamond magnetometry	308
	L. Schäd, S. Eppel, M. Kaschel, K. Edelmann, G. Ronniger, J. Anders, Institut für Mikroelektronik Stuttgart, Stuttgart, T. Hoose, M. Thiel, Nanoscribe GmbH & Co. KG, Eggenstein-Leopoldshafen, M. Kern, Institute of Smart Sensors, Stuttgart	
M1-1: Optische und IR-Sensoren 1		
M1-1.1	Pyroelektrische DLaTGS-Detektoren für FTIR-Spektrometer – Moderne Technologie oder alter Hut?	316
	M. Ebermann, D. Mutschall, E. Müller, InfraTec GmbH Infrarotsensorik und Messtechnik, Dresden	
M1-1.3	Tuned Sideband Diode Laser Absorption Spectroscopy	324
	K. Wu, S. Palzer, Technische Universität Dortmund, Dortmund	
M1-1.4	Resonant-Akustischer Sensor für die gleichzeitige Detektion von Sauerstoff und Kohlenstoffdioxid im Lebensmittelbereich	326
	C. Weber, K. Schmitt, J. Wöllenstein, Fraunhofer IPM, Freiburg	
M1-2: Energie- und Umwelttechnik		
M1-2.1	Abbaubare Sensorknoten zur Überwachung der Bodenfeuchte im Smart Farming	332
	S. Böller, B. Loepp, T. Greuter, Fraunhofer IMS, Duisburg, A. Grabmaier, Universität Duisburg-Essen, Duisburg	
M1-2.2	Inductive Harvester and Data Transmission on Medium-Voltage High Power Cable Shields for Autonomous Sensor Nodes	338
	M. Lenzhofer, D. Holzmann, Silicon Austria Labs GmbH, Villach (Austria), D. Sturm, KNG-Kärnten Netz GmbH, Klagenfurt am Wörthersee (Austria)	
M1-2.3	Instrumentierung und Stabilisierung von Lichtbögen zur Plasma-Schmelzreduktion	343
	T. Voglhuber-Brunnmaier, B. Lehner, J. Stanek, Johannes Kepler Universität, Linz (Österreich), L. Ecker, Johannes Kepler Universität, Linz (Österreich), E. K. Reichel, Metallurgisches Kompetenzzentrum K1-met, Linz (Österreich)	
M1-2.4	CNN-based classification of scale deposits and damages in drainage pipes using a mobile inspection rover	351
	V. Putz, C. Hofer, S. Fleischanderl, L. B. Hörmann, J. Schröck, W. Hofer, Linz Center of Mechatronics GmbH, Linz (Austria), H. Hölzl, T. Schachinger, V. M. Batka, ÖBB-Infrastruktur AG, Wien (Austria)	
M1-3: Autonome Systeme und Robotics		
M1-3.1	Additive Manufactured Sensitive Robotic Gripper Based on the Wedge-Hook Mechanism	359
	N. Hangst, S. Wieber, T. M. Wendt, Offenburg University of Applied Sciences, Offenburg, S. J. Rupitsch, University of Freiburg, Freiburg	
M1-3.2	Robot-Assisted Test Stand for Calibrating 3-D-Printed Torque Sensors	363
	L. Stiglmeier, T. M. Wendt, Hochschule Offenburg, Offenburg, S. J. Rupitsch, Albert-Ludwigs-Universität Freiburg, Freiburg	
M1-3.3	Shared Control in Assistive Robotics: Characteristics Influencing User Acceptance	370
	A. Fischer-Janzen, T. M. Wendt, Offenburg University of Applied Science, Offenburg	
M1-3.4	Additively Manufactured Calibration Target for Intensity-Based LiDAR Calibration in Industrial Robotics	372
	T. M. Wendt, S. Süme, A. Schnebel, Offenburg University of Applied Sciences, Offenburg, S. J. Rupitsch, University of Freiburg, Freiburg	

M2-1: Optische und IR-Sensoren 2

M2-1.1	Analyse der UAV-basierten Methanferndetektion unter variierenden Reflexionseigenschaften von Oberflächen	377
	E. Maier, T. Strahl, J. Wöllenstein, K. Schmitt, Universität Freiburg, Freiburg, J. Herbst, S. Rademacher, Fraunhofer IPM, Freiburg, D. M. Spale, S. J. Rupitsch, Universität Freiburg, Freiburg	
M2-1.2	Mikrobolometerplattform mit anpassbaren Metamaterialabsorbern zur multispektralen Bildgebung im Infrarotbereich	385
	A. Litke, S. Weyers, A. L. Schall-Giesecke, Fraunhofer IMS, Duisburg	
M2-1.4	Anwendung konfokaler Abstandssensoren zur Spaltnessung in einem aktiven Spaltnessmechanismus in einem elektrischen Antriebwerk	390
	N. Herter, A. Hupfer, Universität der Bundeswehr München, Neubiberg	

M2-2: Modellbildung und Signalverarbeitung

M2-2.1	Enhancing Radar-Based Pedestrian Detection with Artificial Intelligence for Multipath Mitigation.....	398
	I. Hoyer, S. Böller, K. Seidl, Fraunhofer IMS, Duisburg	
M2-2.2	Metamodel-based Optimization of Coil Units of Axial-Flux Machines and their Validation Using a Test Bench	404
	F. Schwinning, J. Ihrens, T. A. Kern, Hamburg University of Technology, Hamburg	
M2-2.3	Ein integriertes LSPIV Fernerkundungssystem, Sensordatenfusion und Bildverarbeitung für Assistenzsysteme im regulären Einsatz der Fährschiffahrt	412
	N. Rüssmeier, A. Ngome, Jade Hochschule, Wilhelmshaven	
M2-2.4	Enable Transferability of interpretable ML Models for Bearing Diagnostic in Changing Operating Conditions	420
	J. Schauer, H. El Moutaouakil, Centre for Mechatronics and Automations Technology, Saarbrücken, A. Schütze, Saarland University, Saarbrücken	

M2-3: Prozessmesstechnik

M2-3.3	Fabrication and Characterization of a Sapphire Fabry-Pérot Fiber Sensor for High-Temperature Measurement up to 1000 °C	426
	P. Soni, A. Roehrl, F. Buchfellner, J. Roths, Munich University of Applied Sciences, Munich	
M2-3.4	Hyper-fast Gas Chromatography Coupled with Ion Mobility Spectrometry.....	431
	J. A. Wittwer, A. Nitschke, M. Hitzemann, J. Winkelholz, A. T. Kirk, C. Schaefer, T. Kobelt, C. Thoben, M. Lippmann, S. Zimmermann, Leibniz University Hannover, Hannover	

M3-1: Mechanische Größen

M3-1.1	Evaluation of Time Synchronization Mechanisms in Legal Metrology	433
	M. Esche, F. Reichmuth, R. Meyer, M. Nischwitz, Physikalisch-Technische Bundesanstalt, Berlin	
M3-1.2	Strukturmonitoring mit passivem RFID-Dehnungssensor: Flexible Lösung für Industrie und Infrastruktur	440
	T. Frank, S. Hermann, A. Grün, M. Kermann, M. Hintz, A. Cyriax, K. Ettrich, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt, P. Peitsch, R. Jurisch, S. Kahl, Micro-Sensys GmbH, Erfurt	
M3-1.4	Medienresistenter Absolutdrucksensor ohne Ölfüllung und frontbündigem Medienkontakt	444
	T. Frank, S. Hermann, A. Grün, M. Kermann, A. Cyriax, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt	

M3-2: Modellbildung und Signalverarbeitung 2

M3-2.1	Monocular Inter-UAV Distance Estimation Using Machine Learning and Geometric Scaling	450
	R. Pommeranz, K. Tebbe, J. Fischer, R. Heynicke, G. Scholl, Helmut-Schmidt-University, Hamburg	
M3-2.2	A Data-Driven Multi-Configuration Parameter based Characterization of Eddy Current Inductive Sensors' Signals using Machine Learning	456
	A. Khan, A. Hetzner, M. Kagerer, T. Greiner, Pforzheim University, Pforzheim	
M3-2.3	A Dataset for Condition Monitoring of Screwdriver Bits based on a Piezoelectric Accelerometer for Machine Learning	463
	H. El Moutaouakil, J. Schauer, Centre for Mechatronics and Automations Technology, Saarbrücken, A. Schütze, Saarland University, Saarbrücken	
M3-2.4	Analysis of the Intrinsic Noise Spectrum of MEMS Microphones	469
	E. Wallesch, G. Rodriguez Gutierrez, S. Palzer, Technische Universität Dortmund, Dortmund	

M3-3: Sensornetzwerke

M3-3.1	Energieautarkes LoRaWAN-Sensornetzwerk und integrierte IoT-Datenmanagementplattform für Smart-City- und Umweltmonitoring – Pilotprojekt „IoT-Lab“ der TU Bergakademie Freiberg	471
	P. Arki, Y. Joseph, Technische Universität Bergakademie Freiberg, Freiberg, T. Schumann, Technische Universität Bergakademie Freiberg, Freiberg, J. Kunze, K. Voigtländer, Advanced Data Processing Dresden GmbH, Dresden	
M3-3.2	Ultra-Low-Power Wireless Temperature Measurement for High-Temperature Rotating Equipment.....	477
	L. B. Hörmann, B. Uhl, V. Putz, Linz Center of Mechatronics GmbH, Linz (Austria), G. Minelgaite, J. Barrenetxea-Arando, N. Montes, Fundación Inatec (OTUA), Legutio (Spain)	
M3-3.3	HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks	485
	Keine Autoren angegeben	
M3-3.4	A Comparative Study on Measurement Accuracy: Wireless Sensor Networks vs. Conventional Systems for the Measurement of Large-Scale Geometric Structures.....	491
	L.-M. Bretthauer, R. Heynicke, G. Scholl, Helmut Schmidt University, Hamburg	

Poster

P01	Direct 3-D Printing of Conductive Polymers on Metal Sheets: A Novel Electrical Connection Strategy	498
	S. Schröder, T. M. Wendt, Offenburg University of Applied Sciences, Offenburg, S. J. Rupitsch, University of Freiburg, Freiburg	
P02	Photoakustisches Messsystem zur Produktüberwachung	502
	H. Wünscher, T. Frank, C. Heinze, T. Schildhauer, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt	
P04	Assessing the Controllability of Modular UAV Swarms over 5G Networks via ROS 2	506
	K. Tebbe, R. Pommeranz, G. Kapoor, R. Heynicke, G. Scholl, Helmut-Schmidt-University, Hamburg	
P05	Neuartiger Sensor zur Messung der Schmierfilm- Eigenschaften im Zahnradkontakt.....	514
	C. Schneidhofer, M. Schirru, N. Bernegger, AC2T research GmbH, Wiener Neustadt (Österreich)	
P06	Ein neuartiger Sensor für Öl-Condition Monitoring in extremen Umgebungsbedingungen	520
	M. Schirru, AC2T research GmbH, Wiener Neustadt (Österreich)	
P07	Benutzungsgerechte Untersuchung der Wirkung von Gleitsichtgläsern	525
	F. Fleischmann, D. Hilbig, Hochschule Bremen, Bremen	

P08	Prototyp eines piezoelektrischen Energy Harvesters zur Umwandlung menschlicher Bewegungsenergie für die Versorgung von Wearables	530
	B. Keune, M. Delle, C. Dirks, N. Boushaba, D. Brakensiek, A. Möller, THGA, Bochum	
P09	Von der Elektrode zur Cloud: Leckagemonitoring mit kapazitiver Sensorik und mioty®	536
	J. Ehrlich, B. Brunner, T. Nierla, M. Benkert, G. Domann, Fraunhofer ISC, Würzburg, F. Kemeth, A. Jarresch, A. Meldau, Fraunhofer IIS, Nürnberg	
P10	Latenz- und Energievergleich drahtloser Protokolle für haptische VR-Systeme	542
	A. Kohlhaupt, U. Steinmann, IFAT, Magdeburg	
P11	ThinkIsense – KI in batteriebetriebenen Sensorknoten für die Prozessautomatisierung	548
	S. von Dosky, S. Klehr, G. A. Erdler, Siemens AG, Karlsruhe	
P12	Optimierung der Sensorleistung: Einfluss von Schutzmembranen auf die Ansprechzeit	550
	M. Schindler, M. Khadilkar, B. Hirler, E. Anar, P. Hils, W.L. Gore & Associates GmbH, Putzbrunn	
P13	Körperschallanalyse zur Fehlererkennung bei Fahrzeugkomponenten	556
	M. Bussas, Trout GmbH, Kassel	
P14	Development and validation of a multimodal sonar measurement system for object detection in water bodies.....	560
	S. Lennartz, M. Schneider, H. Mantz, T. Walter, University of Applied Sciences Ulm, Ulm	
P16	Comparative study of 10-mm length commercial erbium doped optical fibers for temperature sensing with a fluorescence decay method.....	568
	R. Kruglov, K. Rostan, R. Engelbrecht, Technische Hochschule Nürnberg Georg Simon Ohm, Nürnberg	
P17	Tool-Center-Point-Kalibrierung eines Werkzeuges für Kaltschmiedeprozesse	572
	B. Loris, M. Gapp, T. Weiser, Hochschule Kempten, Marktoberdorf, S. J. Rupitsch, Universität Freiburg, Freiburg	
P18	Development of an Interactive Search System for Standardised Documents in Legal Metrology and Comparison with a Retrieval-Augmented Generation Approach	576
	Z. Wang, M. Esche, M. Nischwitz, Physikalisch-Technische Bundesanstalt, Berlin, H. Grönniger, Ostfalia University of Applied Sciences	
P19	Optimierung der Lage der Roboterbasis durch Kombination verschiedener Messverfahren	584
	M. Gapp, T. Weiser, Hochschule für angewandte Wissenschaften Kempten, Marktoberdorf	
P20	Measurement Concept for the Deflection Behaviour of Shape Memory Alloy Actuators in Robotics.....	589
	A. Dierle, T. M. Wendt, T. Seifert, Offenburg University of Applied Sciences, Offenburg	
P21	Eye-Tracking Accuracy in Robot Controls: Separating Genuine Issues from Artificial Barriers	594
	A. Fischer-Janzen, J. Roederer, T. M. Wendt, Offenburg University of Applied Science, Offenburg	
P23	Waferintegrierte PowderMEMS®-Mikromagnete: CMOS-kompatible Stützfeldquellen für moderne Sensoranwendungen	596
	O. Behrmann, T. Lisec, M. T. Bodduluri, D. Cichon, B. Gojdka, Fraunhofer ISIT, Itzehoe, Fraunhofer IIS, Erlangen	
P26	Sensorbasierte Zustands- und Positionsüberwachung von Umformwerkzeugen im Bereich der Innenhochdruckumformung – Erste Ergebnisse	599
	T. Reuter, P. Fischer, A. Grundmann, ICM e.V., Chemnitz, T. Frank, CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt, K. Nestler, KOMITEC electronics GmbH, Zwickau	

Wie ultrakleine MEMS Sensoren den Alltag erleichtern und Leben retten

Dr. Gerhard Lammel, Director System Architecture, Bosch Sensortec

Sensoren und Messsysteme 2026, Nürnberg
09. Juni 2026



Agenda

1. Wie klein können Sensoren werden?
2. Optische Partikelsensoren und Waldbrände
3. Hochgenaue Drucksensoren und Notruf
4. Robotik
5. Zusammenfassung



Who we are Invented for life

Robert Bosch GmbH

Four business sectors

Mobility Solutions, Industrial Technology, Energy & Building Technology, Consumer Goods



Mobility Electronics

The electronic heart of Bosch

Electronic Control Units & Semiconductor components



Bosch Sensortec

Sensing solutions for consumer electronics

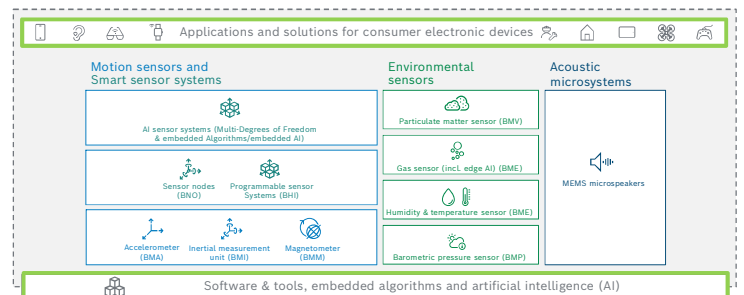
Motion & environmental sensors, smart sensor systems, optical microsystems, acoustic microsystems



3 Internal | Bosch Sensortec | BISTEAPI BISTEAPI PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any design, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



The large spectrum of MEMS devices And their combinations



4 Internal | Bosch Sensortec | BISTEAPI BISTEAPI PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any design, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

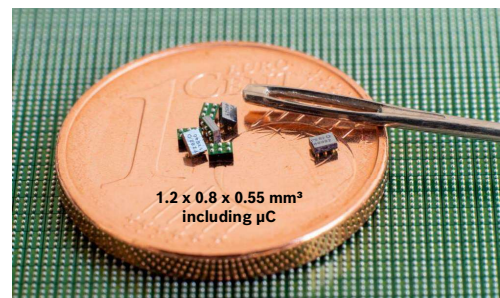


1

Wie klein können Sensoren werden?



The world's smallest Accelerometers BMA580/BMA530 in Wafer Level Chip Scale Package (WLCSPP)

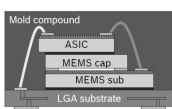


5 Internal | Bosch Sensortec | BISTEAPI BISTEAPI PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any design, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

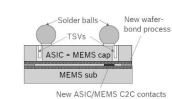


Size matters: The world's smallest CE accelerometer Enabled by our cutting-edge ASICap technology

Classical LGA
4 mm²
Accelerometer



ASICap
Chip Scale Package
1 mm²
Accelerometer



Footprint 1 mm² and height 0.55 mm:
The smallest Bosch product ever



Main features

- No mold compound
- No bond wires
- No substrates



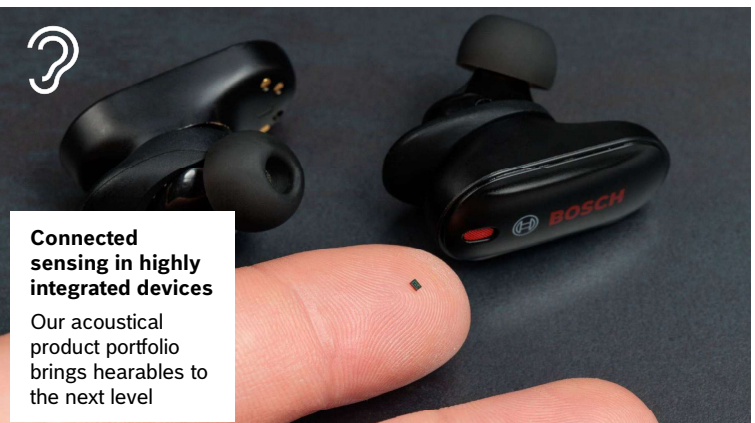
Main technology milestones

- ASIC TSVs
- ASIC/MEMS wafer bonding



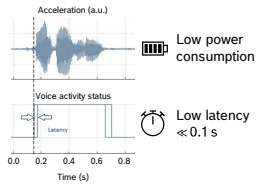
Enabler for the Internet of Things:
µController, ASIC, and MEMS in a single tiny package.

6 Internal | Bosch Sensortec | BISTEAPI BISTEAPI PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any design, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



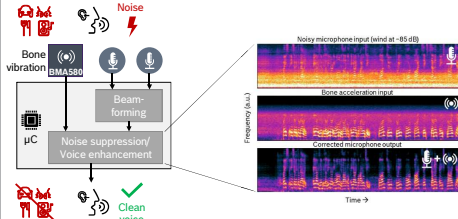
High-bandwidth accelerometer Enabling many voice enhancing functions

Integrated μ Controller:
On-sensor **voice wake-up** feature



Highly integrated sensor:
MEMS, ASIC, and μ Controller

Not sensitive to acoustic noise: **Noise suppression!**



Integrated μ Controller makes
configuration & customization
highly flexible

Enabler of new use cases to
make life more convenient

9 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.

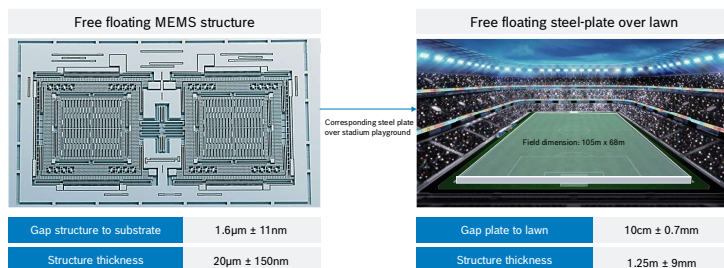


By the way: we are on Mars, too.

The BMI160 takes flight stability to new spheres.



MEMS sensing solutions Highest precision in MEMS processing



11 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.



2

Optische Partikelsensoren und Waldbrände



Every 5 seconds
one human being
is killed by fine dust.



Source: World Health Organization, <https://www.who.int/news/photo/2017/1/candle>

14 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.



Particulate Matter Common Sources of Air Pollution and 'Fine Dust'

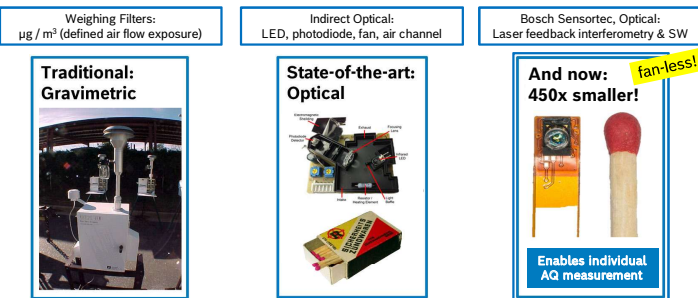


Awareness to air quality continues to increase => Particulate Matter seen as the proxy indicator

14 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.



Particulate Matter Sensor Fine Dust Measuring Devices



2 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.



Particulate Matter Sensor BMV080 Working principle

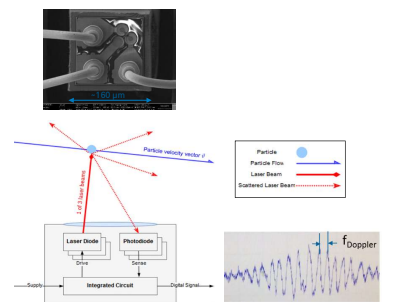
Novel Measurement Technology

- Laser beams of VCSELs with integrated photodiodes (IR-VIP, 850nm)
- Signals detected by the photodiodes:
 - Intensity & duration of light pulses backscattered by particles
 - Doppler shift of Laser frequency of these light pulses
- Calculation of particle features by system (velocity, size, ...)
- Model-based derivation of PM2.5 concentration by system

Key Features

- Ultra-compact form factor & low power consumption
- Accurate, robust against ambient light
- Fan-less measurement in free space

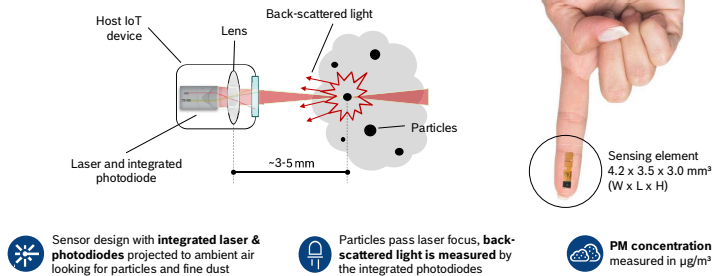
Enables integration into ultra-compact IoT devices
such as air quality monitors, smart thermostats,
smart air purifiers and all kind of wearables.



5 Internal | Bosch Sensortec | BISTEAP1 BISTEAP1-FC | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patents, registrations, trademarks, and/or other intellectual property rights.



BMV080 uses the Doppler effect with back-scattered light from moving particles.

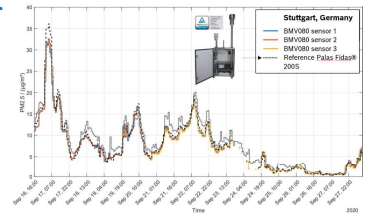
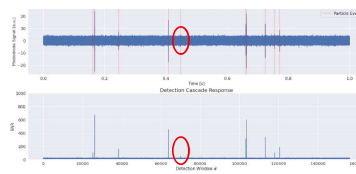


17 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Particulate Matter Sensor BMV080 Find the needle in the haystack...

Signal for typical particle size distribution @ 300 µg/m³
(Duration: 1s, sampling rate: 10 MHz)



- Test location: Outdoor (Stuttgart Feuerbach)
- Test duration: 11 days
- Reference Equipment: Palas Fidas® 2005 - optical PM reference equipment used by the German state of Baden-Württemberg

7 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH



BOSCH

Wildfires Arcachon, France

- This once was a restaurant.

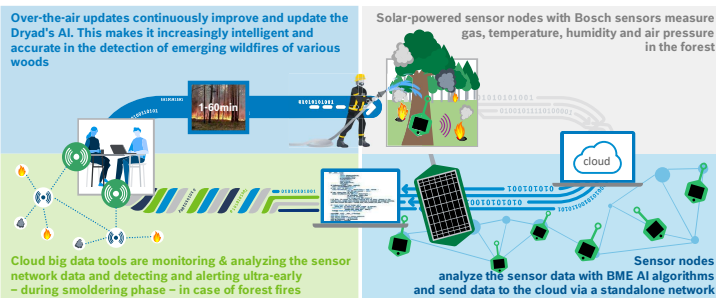


photo: Gerhard Lattner

20 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Environmental sensing Detecting wildfires before they spread



21 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

Wildfire detection Unauthorized Wildfire in detected in 30 Minutes

- Dryad's **ultra-early wildfire detection system**, detected a small illegal fire swiftly, allowing for prompt response measures.
- The sensor network detected a change in air composition through the **Bosch BME688 gas sensor**, triggering an alert through the mesh to the customer.
- Pilot deployment since January 2023 with 91 sensors covering 78-hectare area in the forest adjacent to Deir Mar Moussa.



source: <https://www.dryad.net/post/dryad-silvanet-detects-unauthorized-wildfire-in-lebanon>

22 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

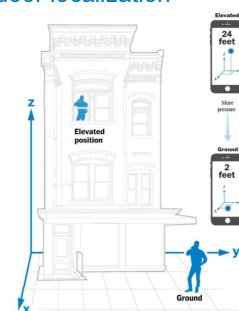
BOSCH

3

Hochgenaue Drucksensoren und Notruf

BOSCH

Pressure Sensor Use cases Indoor localization



Current situation

- Increasing usage of wireless phones for emergency calls, particularly from indoors where traditional localization technologies don't work efficiently or fail completely (e.g. GPS)

Initiative

- FCC (Federal Communications Commission) issued E911 rule in the US: The location of the 911 caller will automatically be transmitted to responder

Challenge

- Development of new technology to implement this feature as a standard in future wireless phone generations in the US is finished now.

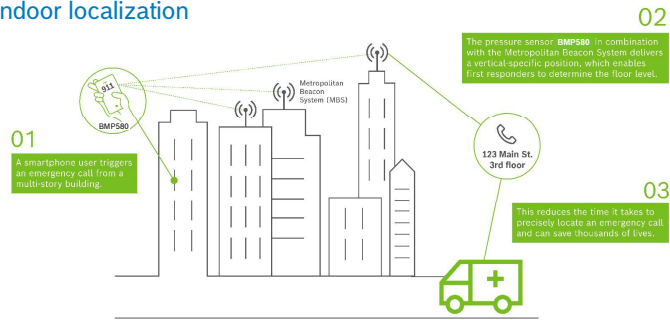


Source: NextNav

24 Internal | Bosch Sensortec | BSE0AP1 BSE0PU PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any disclosed, unpublished, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.

BOSCH

E911 emergency call initiative Indoor localization



25 Internal | Bosch Sensortec | BSEAPP_820PU1 PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patent, trademark, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



Barometric pressure sensor Indoor localization



Solution

- FCC requires wireless providers in the US to adopt a z-axis location accuracy metric of ± 3 meters relative to the handset for 80 percent of indoor wireless 911 calls.
- Bosch and NextNav LLC, a 3D geolocation service provider, have created a solution that provides highly accurate z-axis capabilities indoors.
- Combining barometric pressure sensors inside smartphones with NextNav's Metropolitan Beacon System (MBS) z-axis service determines three-dimensional location and positioning for enhanced emergency calls (E911).

The FCC estimates that the additional vertical information has the potential to save up to **10,000** lives per year just in the USA alone.*

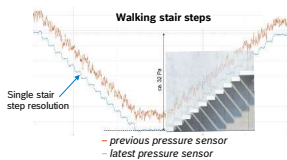
*Source: <https://docs.fcc.gov/public/attachments/DOC-380516A1.pdf>

26 Internal | Bosch Sensortec | BSEAPP_820PU1 PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patent, trademark, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



Hidden high-tech in consumer sensor What does accuracy of 17.5cm mean?

Walking the stairs or measurement to the extreme



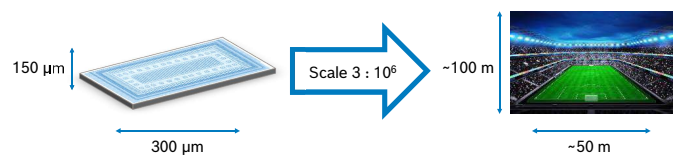
Equivalent of

	~17.5 cm step height, Air column 13000.000 m vs. 13000.175 m
	~2.1 Pa pressure difference 101302 vs 101300 Pa
	/ 300 ~7.6 µg weight change of air on membrane
	/ 1000000 ~12.6 aF capacitance change

27 Internal | Bosch Sensortec | BSEAPP_820PU1 PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patent, trademark, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



Hidden high-tech in consumer sensors Deflection of pressure sensor membrane



- Size of membrane 300x150µm²
- Deflection change for one step height (17.5cm): 5.9µm

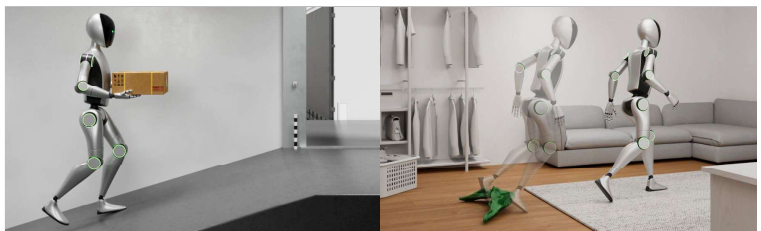
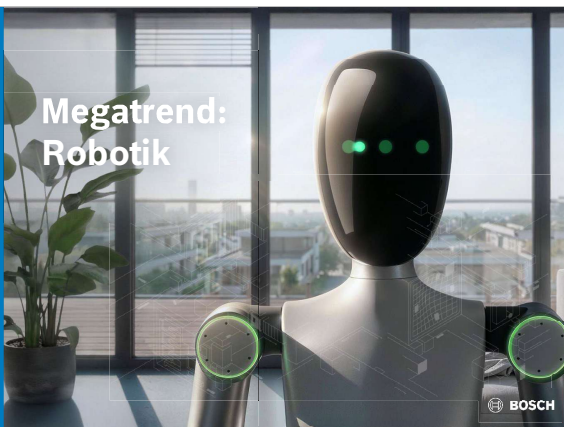
- Size of soccer field ~ 100x50m²
- deflection change at the center: 2µm

28 Internal | Bosch Sensortec | BSEAPP_820PU1 PE | 2026-06-09
© Bosch Sensortec GmbH 2026. All rights reserved. Also regarding any patent, trademark, reproduction, editing, distribution, as well as in the event of applications for industrial property rights.



4

Megatrend: Robotik



Motion control and balancing

Maintaining balance, body orientation, and precise limb control is essential for humanoid and social robots. Bosch MEMS sensors enable the robot to continuously sense and correct its posture and motion.



Dynamic motion control and balancing

The robot detects when it is falling or changing height and reacts to stay safe thanks to Bosch's MEMS sensors. IMUs and accelerometers capture real-time dynamic motion such as falling, slipping, or other unwanted accelerations or displacements. Pressure sensors precisely measure barometric air pressure to monitor subtle height changes. Together, they give the robot a sense of movement and orientation in space.

Sensor	Type	Purpose
BMP581	Ultra-low noise pressure sensor	Height tracking, navigation enhancement
BMP585	Gel-filled pressure sensor	Height tracking, surface pressure detection, navigation enhancement
BM1088	6-axis IMU	Platform stabilization and basic motion control, navigation and SLAM, orientation tracking
BM1423	Standard IMU	Basic motion control, stabilization, orientation tracking
BM1563	High performance IMU	Platform stabilization and basic motion control, navigation & SLAM, orientation tracking



Limb positioning

Encoders are a key technology for precise limb positioning in robots. IMUs provide additional motion data to enhance positioning and motion control. Smart IMUs integrated into body parts such as arms, legs, and the waist monitor and control movements even under dynamic conditions. Magnetometers add absolute heading information, ensuring stable and accurate orientation tracking. Combined with motor control and joint angle measurements from encoders, robots can detect unintended motions such as slippage and adapt to dynamic situations.

Sensor	Type	Purpose
BM1563	High performance IMU	Platform stabilization, real-time motion and orientation tracking
BH1360	Programmable smart IMU	Context recognition, programmability, built-in algorithms for orientation tracking
BM1350	Magnetometer	Absolute heading and orientation immediately on start-up, supporting drift compensation when used with IMU



Environmental sensing

To act intelligently, robots must sense more than just motion or visuals. Environmental sensors from Bosch, such as air quality and humidity sensors, enable context-aware behavior. Humanoid robots can detect poor air quality, smoke, and gases, suggest ventilation, and provide real-time feedback to users as ambient-aware companions.

PM sensor Gas sensor Pressure sensor

Sensor	Type	Purpose
BME690	Gas, humidity, temperature, pressure sensor	Monitors environmental comfort, indoor air quality, and specific gases with fast response time
BMV080	Particulate matter sensor	Detects dust, smoke, and fine particles in real-time
BMP581	Ultra-low noise pressure sensor	Altitude tracking and weather conditions
BMP585	Gel-filled pressure sensor	Media robust, track water pressure, altitude and weather conditions

BOSCH

Tactile Sensing

The robot uses gel-filled air pressure sensors and integrated IMUs in its hands and fingers to detect pressure intensity, distribution, contact force, and 3D hand orientation in real time. This high-resolution tactile sensing enables dexterous manipulation and basic motion control—such as grasping, releasing, and grip adjustment—even without continuous visual input. Vibration and acceleration patterns further support slippage detection.

Pressure sensor Magnetometer IMU

Sensor	Type	Purpose
BM1088	High performance IMU	Basic motion control, bias stability, vibration robustness
BM1423	Standard IMU	Basic motion control, stabilization, orientation tracking
BM1563	Next gen high performance IMU	Platform stabilization and basic motion control, navigation & SLAM, orientation tracking
BMM350	Geomagnetic, high accuracy sensor	3D hand orientation, absolute orientation tracking
BMP581, BMP585	Pressure sensors	Fine surface pressure detection for touch-sensitive control and gripping

BOSCH

Tactile sensing is a major driver for humanoid robots.
What does a tactile sensor look like?

Pressure sensors have been designed to measure fluid pressure but are excellent tactile sensors.

Deformation → Pressure increase

Elastomer Pressure Sensor Press

BOSCH

Mechanical robustness of CE sensors
Life (of a sensor) can be harsh...

Sensors are exposed to various kinds of mechanical stress...

...yet, they must sense extremely small movements!

BOSCH

From use-cases to requirements to tests
Characterization & qualification in BST test center

Characterization	Temporarily applied stress	Stress over lifetime
Characterization tests (performance at t=0) - Offset - Sensitivity - Noise - Accuracy - Current consumption	Sensor specific tests (performance in specific stress situations) - Bending - Short-term Temperature stability - Vibration robustness - Pressure Stability - Installation robustness	Qualification tests (performance over lifetime) - Environmental (temperature, humidity,...) - Mechanical (shocks, drops) - Chemical (chemicals, water,...) - Electrical (ESD, solderability)

BOSCH

5

Zusammenfassung

BOSCH

Sense #LikeABosch
Conclusions

1. Engineers can save lives (when working on the right product)
2. MEMS sensors can contribute significantly to sustainability.
3. Interdisciplinary teamwork is the key to success.

BOSCH

Thank You!

IPCEI Microelectronics and Communication Technologies



The IPCEI ME/CT project is supported by the Federal Ministry for Economic Affairs and Climate Action on the basis of a decision by the German Parliament, by the Ministry for Economic Affairs, Labor and Tourism of Baden-Württemberg based on a decision of the State Parliament of Baden-Württemberg, the Free State of Saxony on the basis of the budget adopted by the Saxon State Parliament, the Bavarian State Ministry for Economic Affairs, Regional Development and Energy and financed by the European Union - NextGenerationEU.

BOSCH

Störstoffüberwachung beim Kunststoffrecycling mittels Online-Gaschromatografie

Wolfgang Reimringer¹, Andreas Schütze¹, Tilman Sauerwald^{2,3}

¹Universität des Saarlandes, Lehrstuhl für Messtechnik, 66123 Saarbrücken

²Fraunhofer IVV, Abteilung Sensorische Analytik & Technologien, 85354 Freising

³Universität des Saarlandes, Fachrichtung Systems Engineering, 66123 Saarbrücken

w.reimringer@imt.uni-saarland.de

Zusammenfassung

Dieser Beitrag beschreibt beispielhaft eine Systementwicklung für den Einsatz von Gaschromatografie-(GC)-Sensor-Systemen in der industriellen Prozessüberwachung. Aufbauend auf Arbeiten zur Verwendung von Metalloxidhalbleiter-(MOS)-Gassensoren als GC-Detektor und einer experimentellen Untersuchung der Prozessatmosphäre wurde ein automatisierter Demonstrator für die Störstoffüberwachung im Polypropylen-Recycling entwickelt. Das System umfasst die vollständige Messkette von der Probenahme bis zur Quantifizierung. Zentrale Herausforderungen sind dabei die Temperaturführung von Probennahme und GC-Ofen sowie die konstruktive Integration und der Betrieb des Sensor-Detektors. Der realisierte Demonstrator arbeitet mit einem Messzyklus von 300 s, einschließlich eines 180-sekündigen Temperaturprogramms zwischen 50 und 200 °C an einer 5 m langen DB-FFAP-Säule. Dabei wird ein Retentions-Index-(RI)-Bereich von 600 bis 1620 mit einer Empfindlichkeit von bis zu 0,5 ng abgebildet. Sensoren und Aktoren werden über eine Industriesteuerung eingelesen und angesteuert; diese stellt auch eine webbasierte grafische Bedienoberfläche zur Verfügung, um die Anbindung in verschiedene relevante Datenverarbeitungskontexte zu zeigen. Insgesamt wird eine Vielzahl von Herausforderungen gelöst, die auf andere Anwendungsfelder übertragbar sind und als Ausgangspunkt für die Weiterarbeit dienen können.

Keywords: Gaschromatographie, Sensor-Detektor, Metalloxidhalbleiter-Gassensor, Kunststoffrecycling, Prozessüberwachung

Hintergrund

Die Kombination aus gaschromatografischer (GC) Trennung und MOS-Sensor als Detektor ist eine vielseitig einsetzbare Möglichkeit, diese kostengünstigen Breitband-Sensoren mit einer zuverlässigen physikalischen Trennung der Analyten zu verbinden. Aktuelle Arbeiten beschäftigen sich mit der Charakterisierung der Detektoreigenschaften [1], mit dem Entwurf der GC-Sensor-Kombination [2] und Kompensationsstrategien für preisgünstige GC-Konzepte [3].

Für die praktische Einsatzfähigkeit in einer Industrieanwendung muss jedoch die Messkette vollständig implementiert und automatisiert werden. Gleichzeitig ist eine hinreichend schnelle Messung erforderlich, um eine echtzeitfähige Prozessbeobachtung zu gewährleisten.

Die hier gezeigte Entwicklung zielt auf den Anwendungshintergrund des Kunststoff-Recyclings ab, bei dem unterschiedlichste Störstoffe aus der vorangegangenen Nutzung eine heterogene Materialqualität bedingen, die bislang

nur unzureichend über Stichprobenmessungen bewertet werden kann. Für die Online-Messung wurde in einer grundlegenden Untersuchung die Vakuum-Entgasung des Compoundier-Extruders als geeignete Messstelle identifiziert und die dort vorliegende Atmosphäre charakterisiert [4].

Die darauf aufbauende Systementwicklung berührt zahlreiche Aspekte, die in anderen Anwendungsfeldern ebenfalls relevant sind. Hervorzuheben sind hierbei das Ofenkonzept und die Optimierung des Detektors.

GC-Sensor-Demonstrator

Auf Basis der in [5] beschriebenen Probenahme-Vorrichtung in Verbindung mit den in [4] vertieften Erkenntnissen wurde die automatisierte Anlage in Abb. 1 konzipiert, die in die Vakuum-Entgasung (3) des Extruders (1) zwischengeschaltet ist. Die Probenahme erfolgt nach einer Feuchteabscheidung (7) auf den kalten Thermodesorber (8), der ein Sorbensbett aus Tenax TA enthält. Durch geeignete

Probennahmezeiten und -temperaturen kann ein Kompromiss für das breite Spektrum polarer und unpolarer Analyten erzielt werden [4].

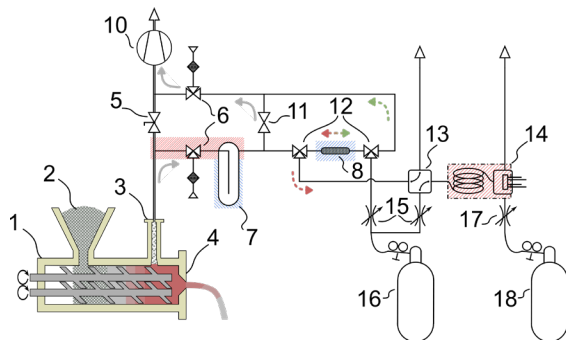


Abb. 1: Gesamtaufbau des Probennahme-GC-Systems am Compound-Extruder. Erweitert nach [6].

Nach der Probennahme wird der Entgasungsstrom über (11) umgeleitet, der Thermodesorber mit (12) abgesperrt und aufgeheizt. Die desorbierten Analyten werden mit einem Stickstoff-Trägergasstrom (16) zur Injektionsbaugruppe (13) transportiert und dort definiert auf das GC-Sensor-System (14) gegeben. Zum Betrieb des MOS-Sensors wird als Makeup-Gas (18) synthetische Luft verwendet, alle Gasströme werden über Massenflussregler (15, 17) eingestellt.

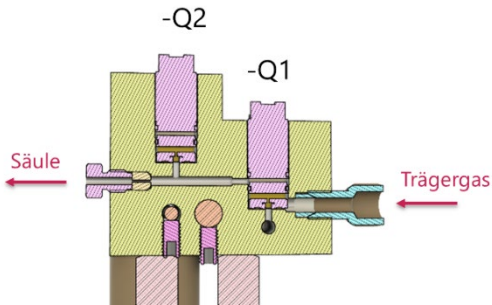


Abb. 2: Injektionsbaugruppe mit temperiertem Block für zwei Kartuschenventile, aus [6].

Der gesamte Aufbau ist modular konzipiert, konstruktive Aspekte sind in [6] und [7] detailliert beschrieben. Für die Ansteuerung der ventilbasierten Injektionsbaugruppe in hat es sich als notwendig herausgestellt, statt den Spülstrom vom Thermodesorber in den Trägergasstrom zu injizieren, den Trägergasstrom während der Injektion abzuschalten. Die fluidische Restriktion der GC-Säule lässt sonst keine definierte Injektion zu, stattdessen arbeiten die Massenflussregler für Trägergas und Spülgas gegeneinander. Eine Injektionsdauer und eine Injektionsvorlaufzeit von jeweils 2 s haben sich für den Kompromiss aus Probeneintrag in die Säule und kurzem Peak bewährt.

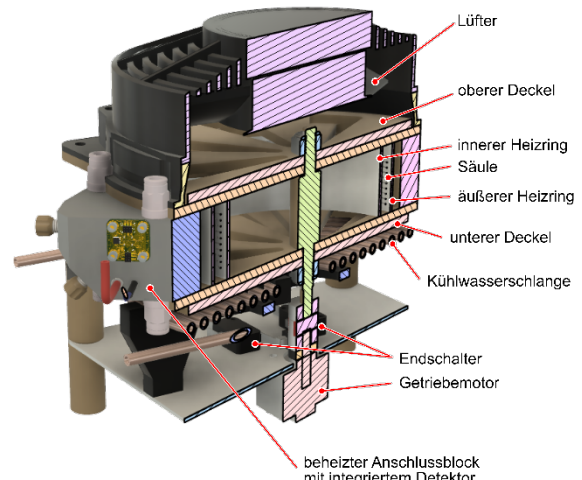


Abb. 3: Aufbau des GC-Ofens für schnelle Temperaturwechsel, aus [6].

Der Gaschromatografie-Teil des Versuchsaufbaus ist vollständig im GC-Ofen integriert, siehe Abb. 3, um Kaltstellen entlang der Säule zu vermeiden. Der Ofenanschlussblock enthält insbesondere den Detektor mit Anströmung und Zumischung des Makeup-Gasstroms.

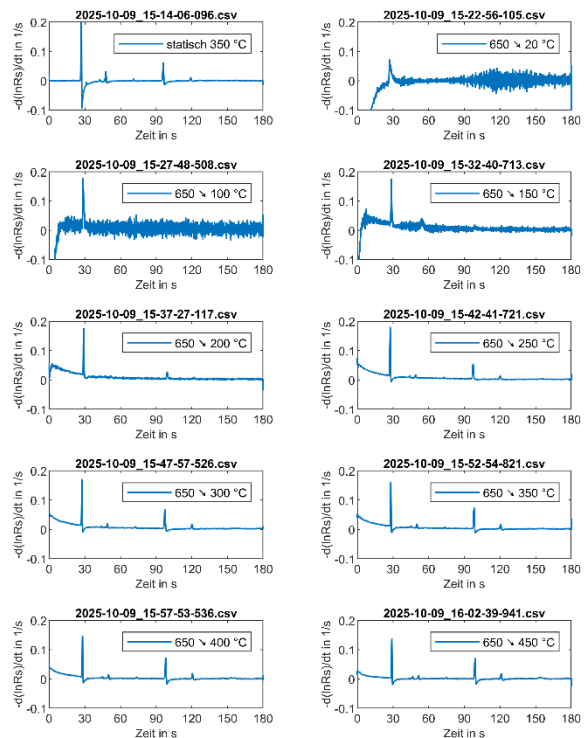


Abb. 4: Vergleichsmessungen mit der Aldehyd-Mischung nach Tab. 2, statisch und bei variierten Tieftemperaturen des DSR-Betriebs.

Charakterisierung

Neben konstruktiven Verbesserungen aus [5] wurde im Gesamtsystem der Sensorbetrieb kritisch untersucht. Ein Vergleich des DSR-

Betriebs aus [1] mit variierten Tieftemperaturen in ergab, dass beim verwendeten Gassensor (GGS1530, Umweltsensortechnik, Geratal) ein Temperatursprung keine Vorteile bringt. Vor allem für die Messung schwererer Moleküle mit höheren Retentionszeiten ist eine hohe Betriebstemperatur für die Detektion wichtig. In Testmessungen hat sich ein Betrieb des Sensors bei einer konstanten Temperatur von 450 °C als gut geeignet herausgestellt.

Tab. 1: Beobachtete Retentionszeiten ausgewählter Substanzen mit Retentionsindizes aus der Literatur [8].

Substanz	CAS	RT in s	RI _{FFAP}
n-Hexan	110-54-3	19	600
n-Heptan	142-82-5	23	700
Ethanol	64-17-5	26	948
Pentanal	110-62-3	38	974
n-Decan	124-18-5	39	1000
Hexanal	66-25-1	47	1090
n-Undecan	1120-21-4	61	1100
n-Butanol	71-36-3	66	1140
Heptanal	111-71-7	71	1192
E-2-Hexenal	6728-26-3	78	1213
Octanal	124-13-0	95	1280
Heptenal	2463-63-0	104	1353
Nonanal	124-19-06	118	1420
n-Pentadecan	629-62-9	150	1500
Buttersäure	107-92-6	160	1620

Das verwendete Temperaturprogramm des GC-Ofens mit 30 s Startphase bei 50 °C, 120 s Rampe nach 200 °C und 30 s Haltephase hat sich als geeignet erwiesen, um Analyten mit Retentionsindizes zwischen 600 und 1620 in einem 180-sekündigen GC-Lauf zu trennen. Der gesamte Ofenzyklus dauert dann 270 s, die vollständige Messung 300 s, vgl. Abb. 9. listet einige qualitativ untersuchte Substanzen mit ihren beobachteten Retentionszeiten auf.

Tab. 2: Aldehydmischung für die GC-Charakterisierung.

Substanz	CAS	n je 1 µL Lsg in nmol	m je 1 µL Lsg in ng
Ethanol	64-17-5	27500	789
Pentanal	110-62-3	275	23,7
Hexanal	66-25-1	275	27,6
Heptanal	111-71-7	275	31,4
Octanal	124-13-0	275	35,3
Nonanal	124-19-06	275	39,1
Buttersäure	107-92-6	275	24,2

Für die Charakterisierung des GC wurde eine isomolare Lösung verschiedener Aldehyde in Ethanol hergestellt, siehe Tab. 2. Damit wurde das TD-Röhrchen mittels Mikroliterspritze bei 25 °C gespiket und die Flüssigkeit mit gefilterter Luft ins Sorbens eingezogen, um eine realistische Verteilung im Röhrchen zu erzielen. Im Anschluss wurden Desorption und GC-Messung entsprechend der oben erläuterten Funktionsweise des Gesamtsystems durchgeführt.

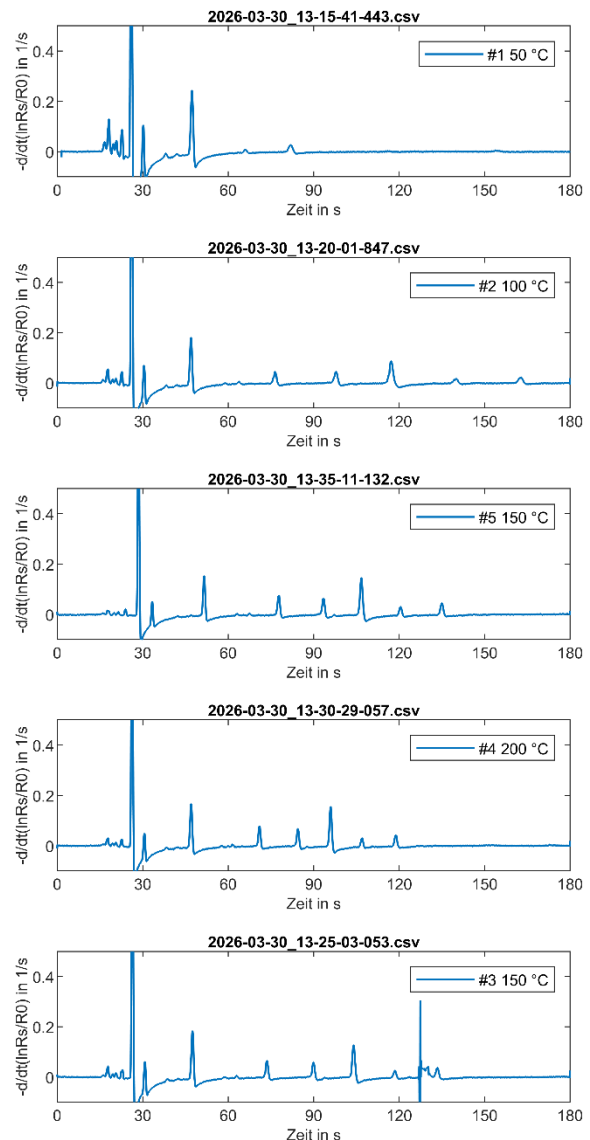


Abb. 5: Einfluss des Ofenprogramms bei Vermessung des Aldehyd-Gemisches nach Tab. 2 mit unterschiedlichen Ofenendtemperaturen.

Der Einfluss des temperaturprogrammierten Betriebs ist in der Messreihe von Abb. 5 dargestellt. Damit wird auch gezeigt, dass ohne ausreichende Betriebstemperatur der Säule manche Substanzen außerhalb des praktisch sinnvollen Zeitrahmens (2 – 3 min) eluieren. Messung #3 zeigt bei $RT = 127$ s einen

entsprechenden Peak von Analyten aus den vorangegangenen GC-Läufen #1 und #2 mit offensichtlich zu geringer Endtemperatur.

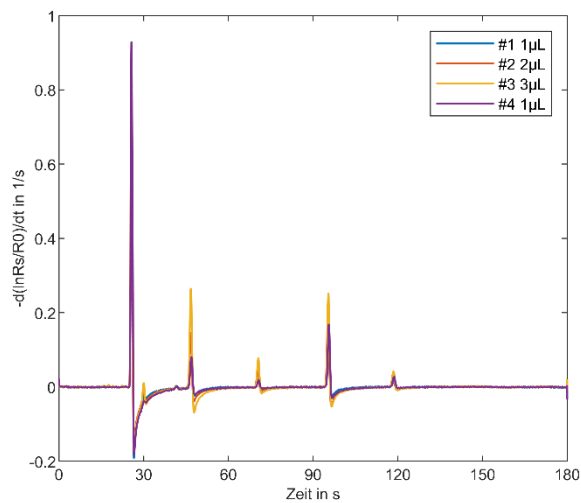


Abb. 6: Stabilität der Retentionszeiten im temperaturprogrammierten Ofen bei vier unabhängigen Messungen.

Hinsichtlich der Wiederholgenauigkeit der Ofenkonstruktion zeigt sich in Abb. 6, dass die Variation der Retentionszeiten in unabhängigen Messungen deutlich unter der mittleren Peakbreite liegen (< 380 ms bei $RT = 119$ s). Hier zeigt sich auch, dass die Retentionszeiten für Analytmengen bis zu 117 ng nicht von der aufgegebenen Menge abhängen. Das bedeutet, dass die Säule noch nicht überladen ist.

Quantifizierung

Für die Auswertung des Sensorsignals wird zunächst das Widerstandssignal differenziert, um die Gaspeaks aufzufinden und zeitlich einzuordnen. Dazu wurden verschiedene Algorithmen für Signalkonditionierung und Ableitung implementiert, auch im Hinblick auf eine mögliche Berechnung in der Industriesteuerung. Insbesondere die Varianten mit gleitendem Mittelwert und fortlaufender Differenzbildung einerseits und einem Faltungsalgorithmus mit einem Stufensignal andererseits haben sich als geeignete Verfahren herausgestellt.

Die Detektionsgrenze (LOD, *limit of detection*) des Gesamtaufbaus mit Thermodesorber, Injektor, GC-Aufbau und Sensordetektor wurde für Aldehyde bis C9 bestimmt und beträgt substanzabhängig 0,5 – 0,9 ng für $S/N = 2:1$.

Die Bestimmung der Peakfläche erfolgt dabei nicht über die (Rück-)Integration der Peaks, sondern über die Bestimmung charakteristischer Zeiten und Betrachtung des Original-Widerstandssignals. Damit kann die Peakflächenberechnung beispielsweise als arithmetische Differenzbildung des Widerstandssignals an den Wendepunkten erfolgen. Abb. 7 zeigt

beispielhaft die Signalverläufe für einen Nonanal-Peak.

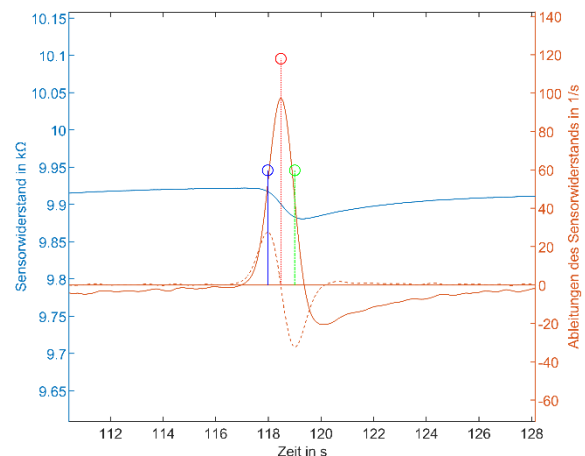


Abb. 7: Beispielverlauf der Sensorsignale an einem Gaspeak. Blau: Sensorwiderstand R_s , Rot: 1. Ableitung $d/dt R_s$ mit Maximum zur Bestimmung der Retentionszeit (roter Marker), gestrichelt: 2. Ableitung $d^2/dt^2 R_s$ zur Bestimmung der Wendepunkte (blauer und grüner Marker).

Die Darstellung der Gaspeaks kann wie in Abb. 4 – 6 mithilfe des Ableitungssignals erfolgen. Alternativ kann für eine abstraktere Darstellung das in Abb. 8 gezeigte Lollipop-Diagramm verwendet werden, dabei gibt die Lage der Markierungen die Retentionszeit an, die Höhe die Peakfläche. Durch Kalibrierung kann die Peakfläche für bekannte Substanzen in eine Konzentration umgerechnet werden.

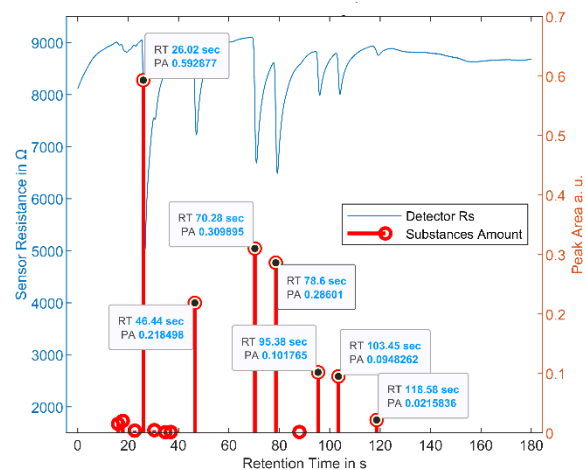


Abb. 8: Automatisierte Quantifizierung mit Bestimmung relevanter Retentionszeiten und Peakflächen.

Diskussion

Die vorstehend gezeigten Ergebnisse wurden zur quantitativen Bewertung mittels Spiking des Thermodesorbers mit isomolaren

ethanolischen Lösungen der Zielsubstanzen erzielt. Im Vergleich zur Vielzahl an Komponenten der Extruder-Entgasung (>95 in [4]) musste dafür eine Auswahl getroffen werden. Hinsichtlich der Retentionsindizes wird der in [4] beschriebene Bereich nicht vollständig wiedergegeben. Dazu müsste eine kürzere Säule eingesetzt oder der effektive Fluss auf der Säule erhöht werden. Allerdings wurde in [9] gezeigt, dass nicht zwangsläufig das vollständige Substanzspektrum für die Qualitätsbewertung von Kunststoff-Rezyklat abgebildet werden muss. Ein anwendungsgerechter Kompromiss aus Umfang des Chromatogramms und Trennleistung ist in ausführlichen Praxisversuchen zu ermitteln. Neben der grundsätzlichen Charakterisierung liegen Ansätze für die automatische Quantifizierung vor. Diese funktioniert aufgrund der guten Trennleistung des Systems für die betrachteten Substanzgemische ohne Probleme. Hier ist vornehmlich der generelle Trend der Basislinie zu beachten, für den eine Online-Kompensationsstrategie unabhängig vom vollständigen (zukünftigen) Signalverlauf gefunden werden muss. Im Falle eines überladenen Chromatogramms mit überlappenden Peaks funktioniert die vorstehend beschriebene Strategie nur eingeschränkt. Hier kommen aufgrund des speziellen Sensorverhaltens aber auch Fitting-basierte Algorithmen an ihre Grenzen. Deshalb ist eine sauber dosierte Probenahme wichtig, die darüber hinaus dafür sorgt, dass die Thermodesorption ein lineares Verhalten aufweist. In diesem Zusammenhang ist eine weitere Verbesserung der Sensitivität notwendig, die für allgemeine Anwendungen zunächst an den Totraumvolumina des Probenahmeteils ansetzen sollte, wo aktuell – aufgrund der anwendungsbedingt großen Querschnitte – eine Verdünnung der Analyten auf dem Weg vom Thermodesorber zur Injektion stattfindet.

Fazit und Ausblick

Als wesentliches Ergebnis liegt ein vollständig automatisiertes Probenahme- und Analytiksystem vor, das das GC-Sensor-Konzept für eine relevante industrielle Anwendung demonstriert. Im Zuge der Entwicklung wurden grundlegende Erkenntnisse zur Anwendung im Kunststoffrecycling sowie für die Probennahmetechnik gewonnen, weshalb eine entsprechende Schutzrechtsanmeldung durch den Projektpartner Fraunhofer Gesellschaft erfolgt ist [10]. Highlights des implementierten Systems sind Thermodesorber, GC-Ofen und Detektor, die zusammen den in Abb. 9 gezeigten Messzyklus von 300 s ermöglichen. Damit steht der Versuchsaufbau im direkten Vergleich mit dem als Benchmark zu betrachtenden Hyper-Fast-GC aus [11], das einen Messzyklus von lediglich

90 s hat und dabei Spuren von 20 Zielsubstanzen quantitativ detektiert. Dafür kommt ein hochkomplexer Aufbau mit einer direkt beheizten Kapillarsäule, konsequent optimierter Fluidik und Kühlung mittels flüssigem Kohlenstoffdioxid zum Einsatz. Das in [11] als nächst bestes referenzierte System wird in [12] beschrieben; es hat einen Messzyklus von 7 min und arbeitet mit einer Kombination aus direktbeheizter Kapillare und konventionellem GC-Ofen. Somit hat der vorliegende Demonstrator das Ziel einer für die Echtzeitfähigkeit gewünschten, möglichst kurzen Messperiode erfüllt – ohne dabei die Anwendbarkeit durch spezielle Säulentypen oder Hilfsmedien zu beschränken.

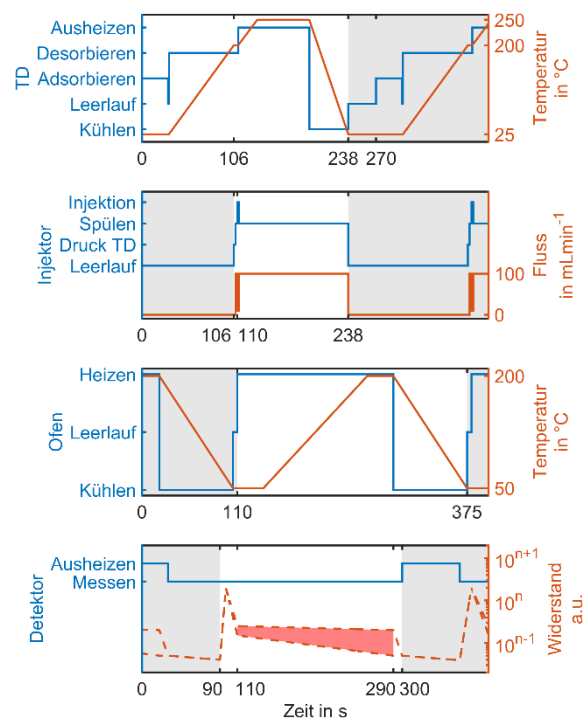


Abb. 9: Zeitabläufe am Probenahme-GC-Sensor-Aufbau, aus [7].

Für die gezeigte Anwendung am Compoundier-Extruder soll in Folgeprojekten der langfristige Messeinsatz erforscht werden. Ausführliche Untersuchungen zu der im Regelbetrieb auftretenden Atmosphäre sollen dabei eine konsequente Weiterentwicklung des Systems ermöglichen. Dies betrifft die Optimierung der Probennahmekonstruktion und -methode, den Betrieb des Gesamtsystems und die Quantifizierung der relevanten Substanzen. Letztlich muss in Fortführung der gezeigten Quantifizierungsansätze auch ein Modell etabliert werden, mit dem aus der Messung auf das Residuum der Substanzen im Extrudat geschlossen werden kann. Im Zuge der Übertragung auf andere Anwendungen ist es von Vorteil, dass das Demonstrationssystem modular aus eigens konstruierten oder markt gängigen Komponenten

zusammengesetzt ist. Das erlaubt eine Weiterarbeit in Form des Probetriebs von geänderten oder alternativen Systemteilen, den Austausch spezifischer Komponenten wie TD-Sorbens, Säulenmaterial oder Sensor-Detektor, sowie alternative Implementierungen von Ablaufsteuerung und Online-Auswertung.

Konkrete Anwendungen, für die bereits eine Übertragung der Erkenntnisse begonnen sind, liegen in den Bereichen Nahrungsmittel und Medizintechnik [13], [14], [15].

Verbesserungen der vorliegenden Komponenten betreffen den Thermodesorber, die Anbindung über den Injektor zur Säule und den Einbau der Säule im GC-Ofen: Der Thermodesorber hat aufgrund der Notwendigkeit zur Wasserkühlung eine große thermische Masse, die durch konstruktive Änderungen reduziert werden kann, während gleichzeitig die Führung von Kühlwasser und Heizungsanschluss überarbeitet werden sollte. Die Anbindung zum Injektor muss, wie bereits erwähnt, hinsichtlich der auftretenden Toträume optimiert werden; der Übergang zur Säule sollte für eine bessere Abdichtung von Ventil und Säulenanschluss revidiert werden, da je nach Säulenlänge hier ein deutlicher Überdruck auftritt. Schließlich ist der Austausch der Säule in der vorliegenden Konstruktion sehr zeitaufwendig, da die Säule manuell mit Abstandhaltern in einer gleichmäßigen, einlagigen Helix äquidistant zwischen die Ofenheizungen eingelegt werden muss. Das erschwert Versuche mit wechselnden Säulenmaterialien oder den perspektiven Austausch als Verbrauchsmaterial in der Anwendung.

Förderhinweis

Die hier vorgestellten Arbeiten wurden im Rahmen des Innovationslabors KIOptiPack vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) unter Förderkennzeichen 033K116 gefördert.

Literatur

- [1] T. Baur, C. Schultealbert, A. Schütze, and T. Sauerwald, „Novel method for the detection of short trace gas pulses with metal oxide semiconductor gas sensors“, *Journal of Sensors and Sensor Systems*, Bd. 7, Nr. 1, S. 411–419, Mai 2018, doi: 10.5194/jsss-7-411-2018.
- [2] O. Brieger, C. Bur, A. Schütze, and T. Sauerwald, „Setup for characterization of MOS gas sensors as GC detector“, *Sensors and Actuators B: Chemical*, Bd. 444, S. 138396, Dez. 2025, doi: 10.1016/j.snb.2025.138396.
- [3] M. Koehne, C. Schmidt, S. Singh, A. Grasskamp, T. Sauerwald, and G. Zeh, „Development of a gas chromatography system coupled to a metal-oxide semiconductor (MOS)

sensor, with compensation of the temperature effects on the column for the measurement of ethene“, *Journal of Sensors and Sensor Systems*, Bd. 12, Nr. 2, S. 215–223, Aug. 2023, doi: 10.5194/jsss-12-215-2023.

- [4] W. Reimringer, H. Haug, and T. Sauerwald, „Inline gas-phase sampling methods for contaminant monitoring in polyolefin recycling“, *JSSS*, Bd. 15, Nr. 1, S. 35–46, März 2026, doi: 10.5194/jsss-15-35-2026.
- [5] W. Reimringer, H. Haug, L. Seifert und T. Sauerwald, „A4.4 – Probennahmeverrichtung zur Kontaminantenüberwachung im Polyolefinrecycling“, 22. GMA/ITG Fachtagung Sensoren und Messsysteme 2024, Nürnberg, Juni 2024, S. 108–111. doi: 10.5162/sensoren2024/A4.4.
- [6] W. Reimringer, H. Haug und T. Sauerwald, „Entwicklung einer automatisierten Probennahme zur Online-Störstoffüberwachung im Polyolefinrecycling“, in *Langfassungsband*, Dresden: DECHEMA, Nov. 2024, S. 95–100. doi: 10.5162/17dss2024/P05.
- [7] W. Reimringer, A. Schütze und T. Sauerwald, „Demonstrator für zyklische Probennahme und schnelle Gaschromatografie in der Kunststoffaufbereitung“, in *tm – Technisches Messen*, Chemnitz, Sep. 2025, S. 37–42. doi: 10.1515/teme-2025-0058.
- [8] J. Kreissl, V. Mall, P. Steinhaus, and M. Steinhaus, „Leibniz-LSB@TUM Odorant Database, Version 1.2. Leibniz Institute for Food Systems Biology at the Technical University of Munich: Freising, Germany“. 2022. <https://www.leibniz-lsb.de/en/databases/leibniz-lsb-tum-odorant-database>
- [9] H. Haug, Z. Catruc, M. Stock, T. Sauerwald, and A. T. Grasskamp, „Assessing and correlating analytical and human sensory data from virgin polyolefins and regrinds to predict their odor properties“, *Resources, Conservation and Recycling*, Bd. 222, S. 108479, Aug. 2025, doi: 10.1016/j.resconrec.2025.108479.
- [10] W. Reimringer, H. Haug und T. Sauerwald, „Verfahren und Anlage zur Detektion von mindestens einem flüchtigen Störstoff in einem schmelzbaren Material“, WO/2025/238150, 20.11.2025 <https://patentscope.wipo.int/search/de/detail.jsf?docId=WO2025238150>
- [11] M. D. Chopra, P. J. Müller, J. Leppert, M. Wüst, and P. Boeker, „Residual solvent analysis with hyper-fast gas chromatography-mass spectrometry and a liquid carbon dioxide cryofocusing in less than 90 s“, *Journal of Chromatography A*, Bd. 1648, S. 462179, Juli 2021, doi: 10.1016/j.chroma.2021.462179.
- [12] M.-J. Rocheleau, M. Titley, and J. Bolduc, „Measuring residual solvents in pharmaceutical samples using fast gas chromatography techniques“, *Journal of Chromatography B*, Bd. 805, Nr. 1, S. 77–86, Juni 2004, doi: 10.1016/j.jchromb.2004.02.018.

- [13] O. Brieger *u. a.*, „Schlussbericht zum IGF-Projekt SENT-GC-MOS: Systementwicklungstool zur effizienten Entwicklung tragbarer GC-Systeme mit MOS-Detektor für die Kontrolle der Lebensmittelqualität“, Zenodo, Apr. 2026.
doi: 10.5281/zenodo.19481019.
- [14] O. Brieger, D. Arendes, and C. Bur, „Advancements in Drug Monitoring: A GC-MOS Approach for Propofol Measurement“, in *2024 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN)*, Mai 2024, S. 1–3.
doi: 10.1109/ISOEN61239.2024.10556018.
- [15] T. Maus *et al.*, „Relationship Between Propofol in Breath Condensate and Its Volatile Fraction in an Isolated Porcine Lung Model“, *ASA - ANESTHESIOLOGY* 2025, Okt. 2025.
<https://www.abstractsonline.com/pp8/#!/20301/presentation/33174>

Aktuelle Arbeiten zu Inspektions- und Monitoring-Techniken zur Erhaltung der Standsicherheit kritischer Infrastruktur am Beispiel Brückenbauwerke

Dirk Koster¹, Samuel Klein¹, Julia Emge¹, Marius Schäfer¹, Christoph Weingard¹

*¹Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP,
Campus E3 1 | 66123 Saarbrücken | Deutschland*

Zusammenfassung

Die Ermittlung des Zustands von Bauwerken gewinnt in Deutschland wegen der alternden Infrastruktur eine immer höhere Bedeutung, vor allem für die öffentliche Sicherheit und aus volkswirtschaftlichen Gründen. Besonders Brücken nehmen hier eine herausragende Rolle ein: In jüngster Vergangenheit mussten mehrere Brücken wegen strukturellen Schäden am Bauwerk für den Betrieb gesperrt werden, in Dresden ist es am 11.09.2024 sogar zum Einsturz des Brückenzug C der Carolabrücke gekommen. Mobile Inspektionen und kontinuierliches Monitoring mit Sensoren können einen entscheidenden Beitrag zur Standsicherheit kritischer Infrastruktur leisten, indem den zuständigen Bauingenieuren und Bauwerksprüfer Sensordaten schnell und auch vor Ort zugänglich gemacht werden. So können direkt weitere Schritte eingeleitet werden. Welche Sensordaten benötigt werden, muss individuell für jedes Bauwerk unter verschiedenen Randbedingungen entschieden werden. Das Fraunhofer IZFP hat deshalb verschiedene Technologien für das kontinuierliche Monitoring und die wiederkehrende Inspektion entwickelt. Dabei kommen moderne Techniken wie Hardware-Beschleuniger für sensornahe KI-Datenverarbeitung sowie die Integration der Sensordaten ins Building Information Modeling (BIM) zum Einsatz. Auf zwei Technologien wird in dieser Veröffentlichung näher eingegangen: die Multimodale Autarke Sensor-Plattform (MAUS) für energieeffizientes kontinuierliches Monitoring und das Sensor-EDGE als Komplettlösung für die umfassende Strukturdiagnose. Beide ermöglichen eine unkomplizierte, sichere und energieeffiziente Bereitstellung von Inspektions- und Überwachungsdaten und erleichtern damit eine genauere Zustandsbewertung von Infrastrukturbauwerken sowie die vorausschauende Instandhaltung.

Keywords: Bauwerksprüfung, zerstörungsfreie Prüfung im Bauwesen (ZfPBau), Zustandsüberwachung, Structural Health Monitoring (SHM), Building Information Modeling, Edge Computing

23. ITG/GMA-Fachtagung, Sensoren und Messsysteme 2026
Nürnberg, 09. Juni 2026

Aktuelle Arbeiten zu Inspektions- und Monitoring-Techniken zur Erhaltung der Standsicherheit kritischer Infrastruktur am Beispiel von Brückenbauwerken

*Dirk Koster¹, Samuel Klein¹, Julia Emge¹, Marius Schäfer¹, Christoph Weingard¹,
¹Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP, Saarbrücken*

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer
IZFP
Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP

Sensor und Datensysteme für
Sicherheit, Nachhaltigkeit und Effizienz

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Motivation: Schutz kritischer Infrastruktur und Aufrechterhaltung öffentlicher Versorgungssicherheit



Dresden: Einsturz von Brückenzug C der Carolabrücke am 11.09.2024



Berlin: Vollsperrung der sogenannten Ringbahnbrücke der Stadtautobahn A100 am 19.03.2025 wegen eines ausbreitenden Risses im Tragwerk.



Bonn: Vollsperrung der Rheinbrücke Bonn-Nord (Friedrich-Ebert-Brücke) seit 03.06.2026 wegen strukturellen Schäden am Tragwerk der linksrheinischen Vorlandbrücke.

- Regelmäßige Überwachung und Prüfung von Ingenieurbauwerken im Straßen- und Wegenetz werden durchgeführt (DIN 1076).
- Frühzeitige Schadenserkenennung mit zeitnaher Analyse, Bewertung und Instandsetzung kann Kosten in Milliardenhöhe vermeiden.
- Die Bereitstellung einer umfassenden Daten- und Informationsgrundlage am Bauwerk selbst ist dabei von entscheidender Bedeutung.

Lösung: Zustandsbewertung durch kontinuierliches Monitoring und wiederkehrender Inspektion direkt am Bauwerk

Seite 2

10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Aber welche Technologie wird benötigt?



Seite 3

10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Herausforderungen

Die Implementierung von Monitoring-Systemen unterliegt unterschiedlichsten Randbedingungen

- Spannungsversorgung:
 - Kabelgebunden
 - Autarke Energieversorgung (Solar)
- Kommunikation:
 - Lokale Infrastruktur vorhanden (LoRaWAN; WLAN/LAN usw.)
 - Lokale Infrastruktur nicht vorhanden (GSM, LTE, usw.)
- Sensordaten:
 - Welche Daten sollen ermittelt werden? Rohdaten oder bereits vorverarbeitete Informationen?
- Bewertung:
 - Wie sieht das Zielsystem aus? User-Interface? Wer administriert das Zielsystem und welche Daten sollen gespeichert werden?



Seite 4

10.06.2026

© Fraunhofer IZFP

Fraunhofer
 Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

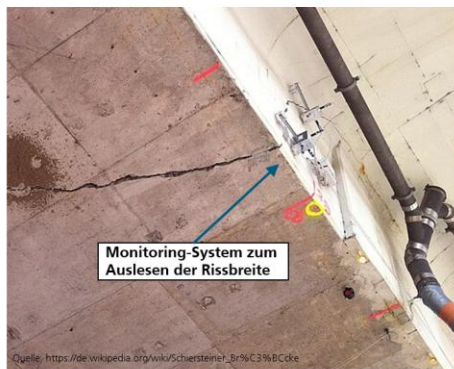
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Inspektions- und Monitoringtechniken für verschiedene Anforderungen kritischer Infrastruktur

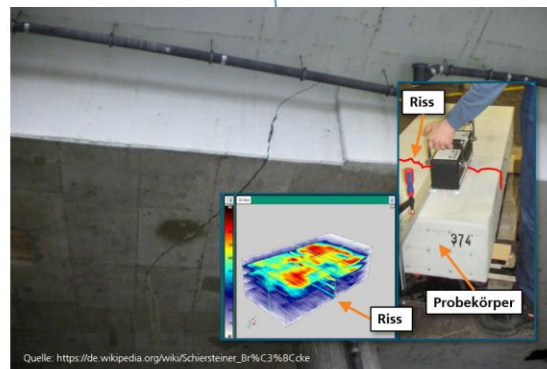
Anforderungen der Betreiber aus der Überwachung und Prüfung nach DIN 1076

 Anforderungen
 Produktebene
 Sensoren usw.

Kontinuierliches Monitoring



Wiederkehrende Prüfung / Inspektion

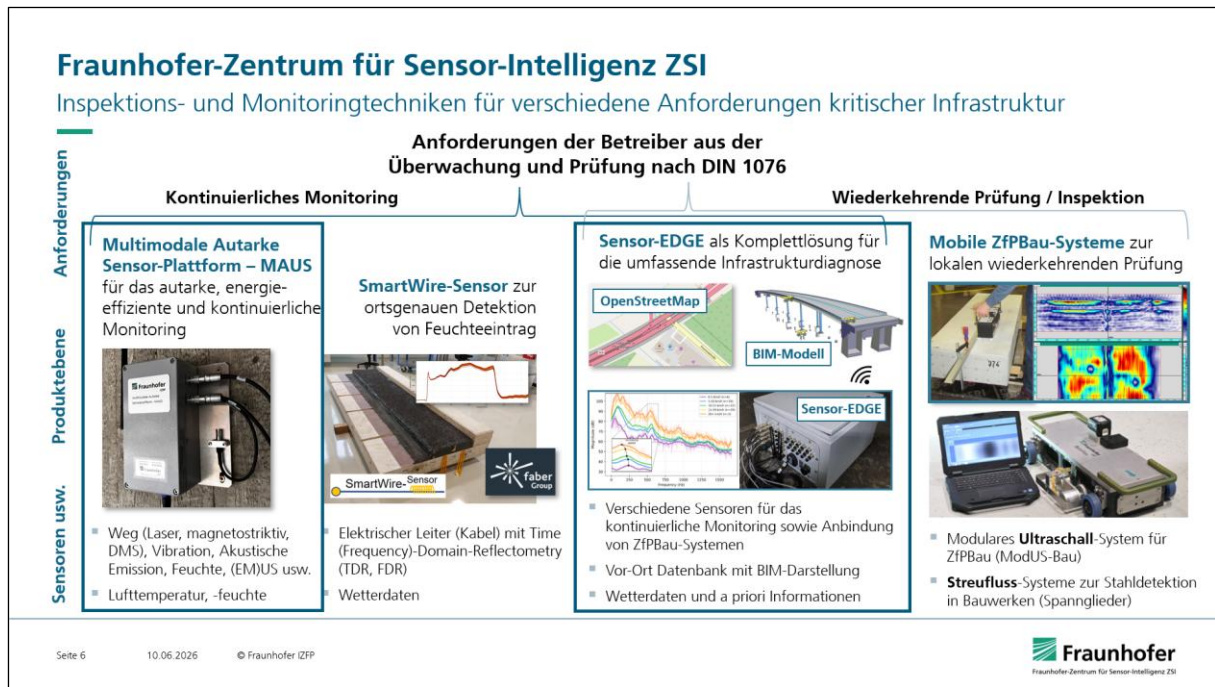


Seite 5

10.06.2026

© Fraunhofer IZFP

Fraunhofer
 Fraunhofer-Zentrum für Sensor-Intelligenz ZSI



Fraunhofer IZFP

Fraunhofer-Institut für zerstörungs-freie Prüfverfahren IZFP

Sensor und Datensysteme für Sicherheit, Nachhaltigkeit und Effizienz

Multimodale Autarke Sensor-Plattform – MAUS für das energieeffiziente und kontinuierliche Monitoring

Seite 7 10.06.2026 © Fraunhofer IZFP

Fraunhofer
 Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

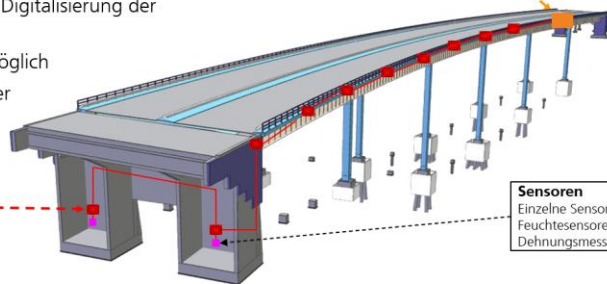
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Multimodale Autarke Sensor-Plattform – MAUS

Energieeffizientes und kontinuierliches Monitoring

- Energieeffizientes und -autarkes System (Solarpanel)
- Flexible Sensorenerweiterung und reduzierter Aufwand für die Verkabelung durch den Einsatz einer Daisy-Chain-Topologie (Serielle Verbindung von Sensor-Frontends)
- Störuneempfindlich durch frühzeitige Digitalisierung der Sensorsignale direkt am Sensor
- Fusion verschiedener Sensordaten möglich
- Datentransport direkt an die Betreiber

Sensor-Frontend (Mini-MAUS)
Sensornähe Elektronik zum Auslesen von Einzelsensoren und sehr nahen Vorverarbeitung.



Sensoren
Einzelne Sensorelemente wie Feuchtesensoren, akustische Emission, Dehnungsmessstreifen und viele weitere.

Datentransport
an die Betreiber

MAUS-Basiseinheit
Elektronik zum Auslesen, speichern, verarbeiten und versenden von Sensordaten.

Seite 8

10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

MAUS: Konfigurationsbeispiel - Wegmessung



Messung der Lagerwege an einer Brücke

- Kontinuierliche Ermittlung der Lagerwege unter Einbeziehung von Umweltdaten (Lufttemperatur und -feuchte)
- Einsatz verschiedener Sensoren (Laser, magnetostruktiv)
- Kontinuierliche Energieversorgung über 230VAC
- Datenübertragung über das LoRaWAN-Netz der Stadtwerke



Seite 9

10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

MAUS: Konfigurationsbeispiel - Wegmessung



Seite 10 10.06.2026 © Fraunhofer IZFP

Fraunhofer
IZFP
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Sensor-EDGE als Komplettlösung für die umfassende Infrastrukturdiagnose

Ergebnisse aus dem öffentlichen Projekt ImaB-Edge

Fraunhofer
IZFP

Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP

Sensor und Datensysteme für
Sicherheit, Nachhaltigkeit und Effizienz

Seite 11 10.06.2026 © Fraunhofer IZFP

Fraunhofer
IZFP
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

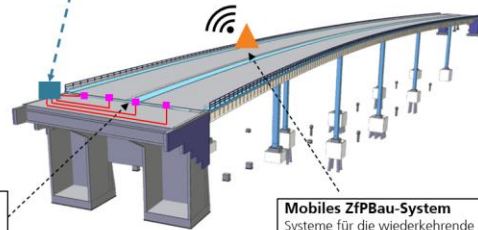
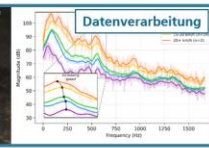
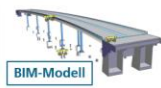
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Sensor-EDGE

Komplettlösung für die umfassende Infrastruktundiagnose

- Hochperformantes System (perspektivisch mit Solarpanel)
- Vielzahl von Sensoren, auch mit sehr hoher Datenrate, anbindbar (parallele Verbindung).
- Datenbank, unterstützt durch KI-Algorithmen, für eine maximale Datenreduktion.
 - Einlernphase mit hoher Datenrate zur Algorithmenentwicklung
 - Betriebsphase mit minimaler Datenrate
- Ergänzung von Sensordaten aus der wiederkehrenden Prüfung (mobiles ZfPBau-System) zur signifikanten Verbesserung der Infrastruktundiagnose.
- Ermöglicht eine vor Ort Bewertung verschiedener Bauwerke durch die Vorhaltung eines digitalen Sensorabbilds

Datentransport
an die Betreiber



Sensoren
Einzelne Sensorelemente,
auch mit hoher Datenrate.

Mobiles ZfPBau-System
Systeme für die wiederkehrende
Prüfung mit z.B. Radar oder Ultraschall

Seite 12

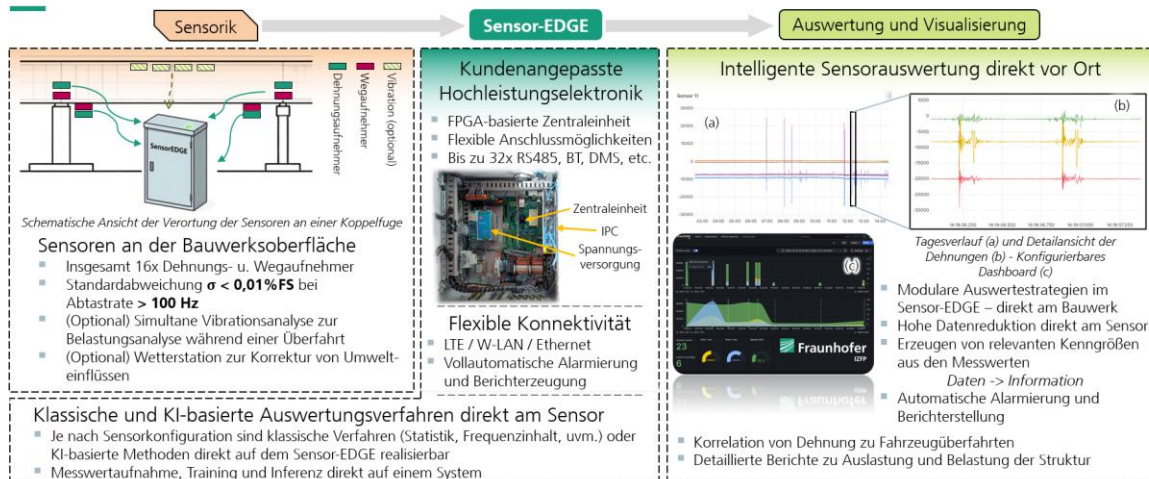
10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Sensor-EDGE: Überwachung von Dehnungs- und Lagerwegverläufen an einem Brückenbauwerk



Seite 13

10.06.2026

© Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Sensor-EDGE: Überwachung von Dehnungs- und Lagerwegverläufen an einem Brückenbauwerk



Seite 14 10.06.2026 © Fraunhofer IZFP

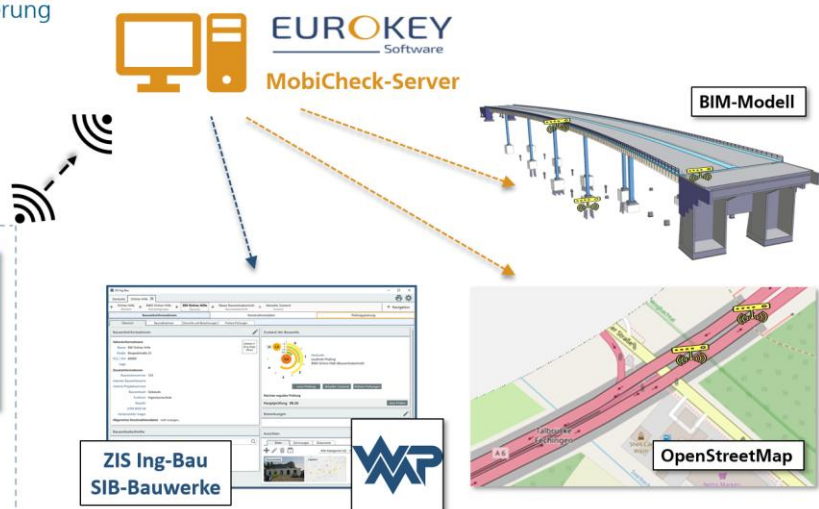
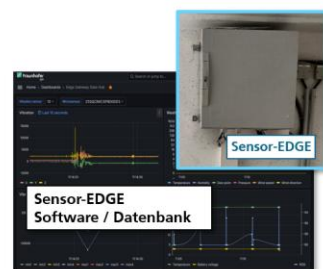
Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Fraunhofer-Zentrum für Sensor-Intelligenz ZSI

Datenhandling und -visualisierung

Digitalisierung und
Datenvisualisierung

Sensor-EDGE / MAUS
/ SmartWire-Sensor



Seite 15 10.06.2026 © Fraunhofer IZFP

Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI



Fraunhofer
Fraunhofer-Zentrum für Sensor-Intelligenz ZSI



Fraunhofer
IZFP

Fraunhofer-Zentrum für
Sensor-Intelligenz ZSI

Fraunhofer-Institut für Zerstörungs-
freie Prüfverfahren IZFP

Sensor und Datensysteme für
Sicherheit, Nachhaltigkeit und Effizienz

Vielen Dank für Ihre Aufmerksamkeit

Dirk Koster, M.Sc.
Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP
Chief Scientist

Campus E3 1
66123 Saarbrücken

Telefon +49 (0) 681 9302-3894
Mobil +49 172 378 8091
Mail dirk.koster@izfp.fraunhofer.de

Sensor-Intelligence Devices für die multimodale Datenerfassung in der Industrie 4.0: Anwendungen, Chancen und Potenziale

Heiko Jacob^{1,2}, Thorsten Müller¹, Jan Oswald¹, Dirk Koster¹, Hans-Georg Herrmann¹

¹Fraunhofer-Institut für Zerstörungsfreie Prüfverfahren IZFP

, Campus E3 1 | 66123 Saarbrücken | Deutschland

²Hochschule für Technik und Wirtschaft des Saarlandes htw saar | Waldhausweg 14 | 66123 Saarbrücken | Deutschland

Zusammenfassung

Die vorliegende Publikation präsentiert das Konzept der Sensor-Intelligence Devices (SIDs) zur multimodalen Datenaufnahme in Industrie 4.0-Umgebungen. Der Fokus liegt auf der Integration Datenerfassung und Edge-KI-Technologien für intelligente und vernetzte Sensorsysteme. Eine offene Systemstruktur ermöglicht die Berücksichtigung von A-priori-Informationen zur Optimierung der Prozesssteuerung sowie zur Realisierung prädikativer Instandhaltungsstrategien. Anhand eines prototypischen Demonstrators wird das zugrundeliegende Funktionsprinzip erläutert. Die implementierten SIDs basieren auf kostengünstigen Sensorsystemen, die definierte Kriterien intelligenter Sensorik erfüllen. Das konzipierte System muss Datensicherheit und -souveränität gewährleisten. Diese Infrastruktur ermöglicht eine umfassende Datenerfassung und -archivierung in Produktionsprozessen und gewährleistet die vollständige Rückverfolgbarkeit relevanter Prozessparameter – ein kritischer Aspekt bei der Aufarbeitung von Anomalien in Produktionsumgebungen. Das SID-Konzept bietet einen systematischen Ansatz zur Integration dieser Anforderungen innerhalb einer kohärenten Systemarchitektur.

Keywords: Smart Sensor, Industrie 4.0, Prozessüberwachung, Prädikative Instandhaltung, Edge Computing

Hintergrund, Motivation und Zielsetzung

Die Entwicklung neuer Konzepte der Datenerfassung ist im Kontext von Industrie 4.0 dringend erforderlich, wenn komplexe und unter anderem global vernetzte Prozesse aufeinander abgestimmt werden müssen. Eine zentrale Fragestellung ist hierbei, wie jeder einzelne Produktionsschritt sinnvoll dokumentiert und bewertet werden kann, aber auch welche Implikationen die Informationen aus vorherigen Prozessschritten bereitstellen. Sensor-Intelligence Devices (SIDs) können in diesem Umfeld eine entscheidende Rolle spielen, indem sie die Erfassung und Verarbeitung multimodaler Daten ermöglichen. Die Bedeutung der Integration verschiedener Datenquellen für die Verbesserung von Entscheidungsprozessen wurde bereits umfassend dargelegt [1, 2]. Die Verwendung von Daten ist insbesondere für die Entscheidungsfindung bei der prädikativen Instandhaltung als Maßnahme gegen Werkzeug- und Maschinenverschleiß von zentraler Bedeutung [3-6]. Diese Arbeit zielt darauf ab, die

Fähigkeiten von SIDs hinsichtlich intelligenter Datenerfassung und Prozessoptimierung zu untersuchen.

Technologische Grundlagen

Diese Studie stellt einen neuartigen Ansatz zur Nutzung von Sensor-Intelligence Devices vor, einer Eigenentwicklung des Fraunhofer-Zentrum für Sensor-Intelligenz (ZSI). Dabei handelt es sich um ein Sensornetzwerk, das aus Sensorknoten und einem EDGE-Gateway besteht und für den Einsatz in der Zustandsüberwachung von Infrastrukturen sowie vernetzten Produktionsumgebungen konzipiert wurde [7, 8]. Obwohl SIDs primär für die Erfassung von Langzeitdaten in Infrastrukturen entwickelt wurden, erstreckt sich ihr Anwendungsbereich weit darüber hinaus und bietet erhebliches Potenzial für Fertigungsumgebungen, insbesondere durch die Implementierung von Machine-Learning-Techniken. Durch den Einsatz geeigneter Datenverarbeitungsmethoden wie der Hauptkomponentenanalyse (PCA) und der Zeitreihenanalyse können SIDs wertvolle Erkennt-

nisse über Produktionsprozesse generieren. Dies umfasst auch das Potenzial für prädikative Instandhaltungsstrategien auf Basis der durch SIDs erfassten Daten.

Das SID-Konzept zeichnet sich durch seine Übertragbarkeit aus und kann in verschiedenen Fertigungsszenarien implementiert werden. Grundlage der SIDs ist die multimodale Datenaufnahme über Sensorknoten, die mit Temperatur, Druck, Bewegungs- und Luftfeuchtigkeitssensoren ausgestattet sind und je nach Anwendungsszenario um weitere Sensormodalitäten ergänzt werden können.

Sensor-Intelligence Devices (SIDs):

Das SID-Konzept basiert auf einem Sensornetzwerk mit einem einfachen Kommunikationsmodell. Jedes SID repräsentiert einen Sensorknoten, der lokale Informationen im jeweiligen Messszenario erfasst. Im Kontext von Produktionsszenarien umfasst dies lokale Informationen aus mehreren Prozesszonen, räumlich verteilte Produktionsschritte, die Überwachung verschiedener voneinander abhängiger Fertigungsprozesse sowie Kombinationen dieser Anwendungsfälle. Jedes SID verfügt über eine eindeutige Identifikationsnummer (ID), sodass Daten in Abhängigkeit vom spezifischen Standort des jeweiligen Knotens respektive Produktionsschritts zurückverfolgt werden können.

Die Koordination der Knoten erfolgt durch einen Master (EDGE-Gateway), der die Datenverwaltung übernimmt. SIDs nutzen Edge-KI-Technologien, wodurch die Daten vorverarbeitet und mittels Edge-Computing einer geeigneten Daten- bzw. Dimensionsreduktion unterzogen werden können, um ausschließlich relevante Informationen mit hoher Signifikanz für das jeweilige Messproblem zu speichern.

Die Kommunikation erfolgt über Ethernet oder WiFi mit UDP als primärem Kommunikationsprotokoll; je nach Anwendungsfall ist alternativ TCP/IP und OPC UA möglich. Die Kommunikation basiert auf dem Master-Slave-Prinzip mit einem zentralen Master und einem oder mehreren Sensorknoten als Slaves. Der Master initiiert die Kommunikation über Broadcast oder Multicast zur Synchronisation aller Knoten im Subnetz. Die Sensorknoten kommunizieren über Unicast, um gezielte und kollisionsfreie Statusaktualisierungen zu

gewährleisten.

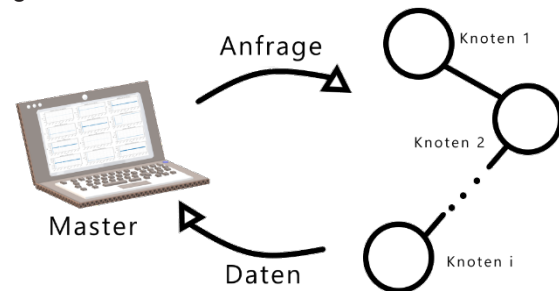


Abbildung 1 Einfache Skizze des SID-Kommunikationsprotokolls

Implementierung der SIDs:

Die Integration der SIDs folgt dem dargestellten Iterationszyklus (vgl. Abb. 2). Dabei wird jeder Sensorknoten systematisch im Kontext der vernetzter Produktionsumgebungen entwickelt.

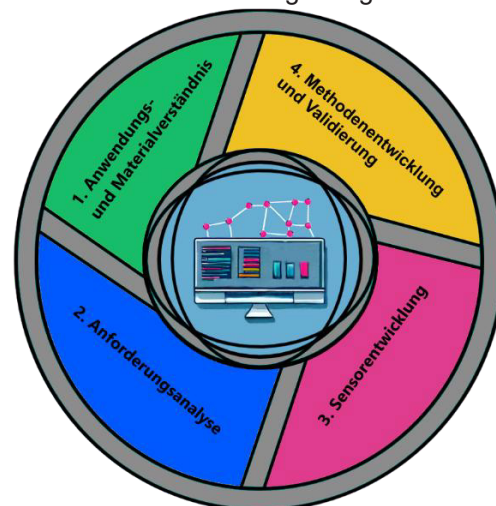


Abbildung 2 Die vier Phasen der SID-Integration

Die Integration von SIDs erfolgt im Allgemeinen in aufeinanderfolgenden Phasen: Zunächst wird in der Phase „Anwendungs- und Materialverständnis“ der Einsatzkontext definiert, d. h. Zielapplikation, relevante Rahmenbedingungen sowie potenzielle Material- oder Technologieeinsätze. Basierend darauf folgt die Phase „Anforderungsanalyse“, in der funktionale, technische und operationale Anforderungen abgeleitet, Messgrößen festgelegt, der Einsatz weiterer prozessbezogener Kenngrößen evaluiert (z. B. Drehzahl, Motorleistung, Abrichtintervalle usw.), Akzeptanzkriterien definiert und Prioritäten gesetzt werden, um anschließend Sensorentwicklung und Integration externer Daten abzuleiten. Der wichtigste Punkt bildet die Ableitung der relevanten Zielgrößen (z. B. Detektion des Werkzeugverschleißes in einem Drehprozess). In der dritten Phase „Sensorentwicklung“ erfolgt die architektonische Ausgestaltung der SIDs mitsamt des angegliederten digitalen Ökosystems, einschließlich der Auswahl von

Sensorik, Edge-Komponenten, Datenformaten und Vorverarbeitungsstrategien. Das Grundgerüst für diesen Aufbau liefert ein Prototyp, der je nach Anforderung weiterentwickelt, getestet und gegebenenfalls angepasst werden kann. Die vierte Phase „Methodenentwicklung und Validierung“ widmet sich der Implementierung in Prozessketten oder Einsatzumgebungen durch Analyse- und Modellierungsansätze, Edge-KI, Datenverarbeitung, Dimensionsreduktion und prädiktive Methoden, deren Validierung, Bewertung und Dokumentation auf Basis der Anforderungsanalyse erfolgt. Zentraler Bezugspunkt des Prozesses ist das EDGE-Gateway als Integrations- und Koordinationspunkt. Er sammelt die mittels der Edge-KI vorverarbeiteten und transformierten Daten aus allen Phasen, konsolidiert Ergebnisse und erzeugt Feedback. Mit dieser Herangehensweise wird das SID-Konzept schrittweise an die Anforderungen angepasst, datenbasiert validiert und so zu einer robusten, effizienten Lösung weiterentwickelt, die sich zuverlässig in komplexe Produktionsprozesse integrieren lässt.

Diese systematische Vorgehensweise folgt konzeptionell dem PDCA-Zyklus [9] (Plan-Do-Check-Act) und lässt sich methodisch in das V-Modell XT [10] einbetten, das mit seinem flexiblen, produktzentrierten Ansatz den gesamten Systemlebenszyklus abdeckt.

Anwendungsbeispiel: Außenlängsdrehen

Am Beispiel eines Drehprozesses zur Herstellung von Wellen wird die Integration von SIDs exemplarisch demonstriert. Die Zielsetzung umfasst die Echtzeitüberwachung des Prozesszustands, die frühzeitige Detektion von Werkzeugverschleiß sowie die Implementierung einer prädiktiven Instandhaltungsstrategie zur Vermeidung ungeplanter Stillstandzeiten und von Ausschuss.

Phase: Anwendungs- und Materialverständnis

Im betrachteten Beispiel, einem Drehprozess, werden Stahlwellen (42CrMo4) mit einer CNC-Drehmaschine bearbeitet. Typische kritische Einflussfaktoren sind hierbei Werkzeugverschleiß, thermische Effekte, Prozessstabilität und Umgebungseinflüsse.

Phase: Anforderungsanalyse

Basierend auf dem Prozess kann als zentrale Zielgröße die Klassifikation des Werkzeugverschleißzustands in die Kategorien „Ok“ und „kritisch“ definiert werden. Die kontinuierliche Überwachung erfordert die Erfassung von Vibrationsdaten und Temperaturdaten. Für die relevanten Messgrößen eignen sich drei-

achsige Vibrationen mittels MEMS-Beschleunigungssensoren, für Werkzeug-, Umgebungs- und Maschinentemperaturen Pt-100, Umgebungsbedingungen sowie externe Maschinendaten wie Spindeldrehzahl und Motorstrom sollten ebenfalls integriert werden. Die funktionale Anforderung besteht in der Detektion kritischer Verschleißzustände mit einer definierten Vorlaufzeit (z. B. 25 Bearbeitungszyklen) bei einer Kommunikationslatenz unter 100 ms.

Phase: Sensorentwicklung

Für den einfachen Drehprozess können drei Sensorknoten strategisch positioniert werden: SID-01, werkzeugnah zur Erfassung von Vibrationen und ggf. Abwärme des Werkzeugs, SID-02 am Maschinenrahmen für strukturelle Schwingungen und Maschinentemperatur sowie SID-03 im Arbeitsraum für Umgebungsbedingungen. Die Kommunikation erfolgt koordiniert durch ein EDGE-Gateway als zentraler Master. Auf Sensorknoten-Ebene werden Rohdaten bereits vorverarbeitet (Mittelung von Bewegungsdaten, 5 dominante Frequenzen über Fast-Fourier-Transformation, etc.), um eine sinnvolle Datenreduktion zu gewährleisten.

Phase: Methodenentwicklung und Validierung

Für die Modellentwicklung wird eine repräsentative Anzahl von Bearbeitungszyklen mit systematischem Werkzeugwechsel benötigt. Visuelle Inspektion sowie Rauheitsmessungen indizieren die Verschleißkategorie, die entsprechend gelabelt werden muss. Sowohl ein gewisser Kantenverschleiß als auch unzureichende Oberflächengüte müssen mit „kritisch“ gelabelt werden. Die erfassten Sensor-Features können mittels Hauptkomponentenanalyse auf eine definierte Anzahl Hauptkomponenten reduziert werden, oder über ein Varianzkriterium in der Hauptkomponentenanalyse gewichtet werden. Somit sollte sich eine deutliche Cluster-Bildung je nach Verschleißzustand zeigen. Es empfiehlt sich eine Validierung der entwickelten Methode nach einem Toleranzkriterium für Fehlalarmraten.

Ausblick

Die demonstrierten Fähigkeiten der SIDs sind auf verschiedene Fertigungsverfahren übertragbar. Besonders relevant sind Anwendungen in mehrstufigen Produktionsketten: Die eindeutige ID-Zuordnung der Sensorknoten ermöglicht die durchgängige Verfolgung von

Prozessparametern über mehrere Fertigungsschritte hinweg. Die Kommunikationsarchitektur erlaubt dabei den Informationsaustausch sowohl in Vorwärts- als auch in Rückwärtsrichtung entlang der Prozesskette. Dies ermöglicht neben der Detektion von Qualitätsabweichungen die prozessübergreifende Nutzung von Zustandsinformationen für adaptive Regelungsstrategien. Die resultierende durchgängige Rückverfolgbarkeit bildet die Grundlage für eine ganzheitliche Optimierung der Wertschöpfungskette.

Über die Zustandsüberwachung hinaus eröffnen SIDs erweiterte Anwendungsfelder: Prognosemodelle zur Vorhersage verbleibender Werkzeugstandzeiten dienen nicht nur prädiktiven Instandhaltungsstrategien, sondern öffnen auch weitere Anwendungen, wie die Abschätzung von Qualitätsmerkmalen, sprich die Identifikation kausaler Zusammenhänge zwischen Werkzeugdegradation und Produktmerkmalen. Die Integration in die übergeordnete Produktionsplanung eröffnet Potentiale zur Bestimmung der Produktionsauslastung und Maschinenverfügbarkeit.

Ein wesentlicher Vorteil des SID-Konzepts liegt in seiner Eignung für Retrofitting-Anwendungen in Brownfield-Anlagen. Die modulare Architektur und einfache Kommunikationsstruktur ermöglichen die kosteneffiziente Nachrüstung bestehender Maschinen ohne (bauliche) Eingriffe oder produktionsrelevante Stillstandzeiten. Die einfache, nicht-invasive Installation der Sensorknoten und die steuerungsunabhängige Datenerfassung erlauben die Integration älterer Anlagen ohne herstellerspezifische Schnittstellen in Industrie-4.0-Umgebungen. Dies bietet insbesondere kleinen und mittleren Unternehmen einen niederschweligen Zugang zur digitalen Transformation ihrer Produktion bei überschaubarem Investitionsaufwand.

Danksagung

Diese Arbeit wurde vom Ministerium für Wirtschaft, Innovation, Digitales und Energie des Saarlandes (MWIDE) und der Fraunhofer-Gesellschaft (Leistungszentrum „Sensor-Intelligenz“) unterstützt.

Literaturnachweis

[1] Lebedynska, Y. (2011). *Entwicklung eines Informationssystems mit Reifegradmanagement für automatisierte Schraubprozesse*. BTU Cottbus-Senftenberg.

[2] Zhang, Y., Cheng, Y., Wang, X. V., Zhong, R. Y., Zhang, Y., & Tao, F. (2019). Data-driven smart production line and its common factors. *The International Journal of Advanced Manufacturing Technology*, 103, 1211-1223.

[3] Chien, C. F., & Chen, C. C. (2020). Data-driven framework for tool health monitoring and maintenance strategy for smart manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, 33(4), 644-652.

[4] Lu, B., Chen, Z., & Zhao, X. (2021). Data-driven dynamic predictive maintenance for a manufacturing system with quality deterioration and online sensors. *Reliability Engineering & System Safety*, 212, 107628.

[5] Lee, J., Ni, J., Singh, J., Jiang, B., Azamfar, M., & Feng, J. (2020). Intelligent maintenance systems and predictive manufacturing. *Journal of Manufacturing Science and Engineering*, 142(11), 110805.

[6] Farooq, B., Bao, J., Li, J., Liu, T., & Yin, S. (2020). Data-driven predictive maintenance approach for spinning cyber-physical production system. *Journal of Shanghai Jiaotong University (Science)*, 25, 453-462.

[7] Klein, S., Emge, J., & Koster, D. (2024). Using edge AI: Continuous monitoring of critical infrastructure. *inno*.

[8] Koster, D., Oswald, J., Jungmann, C., Klein, S., & Weingard, C. (2024). Erweiterung des Sensordatenraums zur Ermittlung der Datenaufnahmequalität für Wirbelstromanwendungen. In *Anwenderseminar Wirbelstromprüfung 2024*.

[9] Realyvásquez-Vargas, A., Arredondo-Soto, K. C., Carrillo-Gutiérrez, T., & Ravelo, G. (2018). Applying the Plan-Do-Check-Act (PDCA) cycle to reduce the defects in the manufacturing industry. A case study. *Applied sciences*, 8(11), 2181.

[10] Biffi, S., Winkler, D., Höhn, R., & Wetzel, H. (2006). Software process improvement in Europe: potential of the new V-modell XT and research issues. *Software Process: Improvement and Practice*, 11(3), 229-238.

Konzeption und Implementierung einer automatisierten Prozessführung und -überwachung in Bioreaktoren am Beispiel der Kultivierung von humanen Stammzellen

Michael M. Gepp¹, Roman Ruff¹, Paul Woll^{1,2}, Gizem Aydogan^{1,2}, Dmitry Amelin¹, Thorsten Knoll¹, Heiko Zimmermann^{1,2}, Sylvia Wagner¹, Thomas Velten¹

¹ *Fraunhofer Institut für Biomedizinische Technik – IBMT, Joseph-von-Fraunhofer-Weg 1, 66280 Sulzbach/Saar*

² *Universität des Saarlandes, Molekulare und Zelluläre Biotechnologie/Nanotechnologie, Campus, 66123 Saarbrücken*

Zusammenfassung

Die Kultivierung humaner induzierter pluripotenter Stammzellen (hiPSCs) für In-vitro-Testsysteme erfordert eine präzise Prozessführung, da Schwankungen bei Nährstoffen, Temperatur und mechanischen Einflüssen die Zellqualität stark beeinflussen. Im miniaturisierten Maßstab führen manuelle Eingriffe, diskontinuierliche Analytik und fehlende Online-Überwachung zu hoher Chargenvariabilität. In dieser Arbeit soll die biologischen Hintergründe und Anforderungen für ein integriertes Konzept zur Sensorisierung, Prozessüberwachung und teilweisen Automatisierung von Mini-Bioreaktoren für die hiPSC-Kultivierung dargestellt und Ansätze zur Umsetzung vorgestellt werden.

Im Fokus stehen die perfusionsbasierte Kultivierung von hiPSC-Sphäroiden, die Erfassung metabolischer und physikalischer Prozessparameter sowie eine automatisierte morphologische Qualitätsbewertung durch bildbasierte Analytik und Machine-Learning-Methoden. Erste Ergebnisse zeigen, dass ein manuell durchgeführte Medienwechsel zur Simulation einer Perfusion stabilere Kulturbedingungen ermöglichen und die metabolische Prozessführung verbessern. Gleichzeitig beeinflussen geringfügige Änderungen am System das Zellverhalten, weshalb zellbiologische Experimente frühzeitig in den Implementierungsprozess integriert werden sollten.

Keywords: Bioreaktoren, Zellkultivierung, Zellkulturautomatisierung, Biosensorik

Biologischer Hintergrund

Humane induzierte pluripotente Stammzellen haben sich seit ihrer Entdeckung [1] in den letzten Jahren als zentrale Modellplattform für Disease Modeling, Wirkstoffscreenings und regenerative Medizin etabliert [2]. Ihr großes Potenzial, z.B. In-vitro-Zellmodelle für Herzmuskelgewebe künstlich im Labor zu züchten [3], geht jedoch mit hohen Anforderungen an die Prozessführung einher: Geringfügige Abweichungen bei Nährstoffverfügbarkeit, pH-Wert, Temperatur oder mechanischer Belastung können Wachstum, Viabilität und Pluripotenz maßgeblich beeinflussen. In der klassischen 2D-Kultur werden diese Schwankungen meist durch einmal täglich durchgeführte manuelle Mediumwechsel kompensiert, was jedoch zu abrupten Konzentrationsänderungen und einer eingeschränkten Prozesskontrolle führt. Miniaturisierte Bioreaktorsysteme hingegen bieten die Möglichkeit, 3D-Sphäroidkulturen unter definierten und

skalierbaren Bedingungen zu etablieren sowie Zellen zu expandieren und zu differenzieren [4]. Der Medienwechsel erfolgt hier in vielen Forschungslaboren noch immer analog zur 2D-Kultur über manuelle Wechsel. Insbesondere die Perfusionskultur erlaubt perspektivisch einen kontinuierlichen Mediumswechsel und damit stabilere Kulturbedingungen, die einen direkten Effekt auf z.B. Proliferation oder Differenzierung der Zellen haben [5]. Eine Übersicht der externen Einflüsse auf hiPSCs ist in Abbildung 1 dargestellt. Der volle Nutzen dieser Systeme wird jedoch erst durch geeignete Sensorik und automatisierte Datenauswertung erschlossen [6]. Neben der klassischen Sensorik (z.B. biochemische Sensoren für Glukose) werden für Zellsysteme wie hiPSCs zunehmend Methoden der künstlichen Intelligenz übertragen bzw. angewendet [7]. Hierfür werden insbesondere qualitativ hochwertige und annotierte Datensätze benötigt, um Modelle zu trainieren, die dann z.B. für die Klassifizierung von Zellzuständen

eingesetzt werden können. Die Modifikation bestehender Bioreaktorsysteme durch zusätzliche Funktionalitäten stellt dabei stets eine Herausforderung dar.

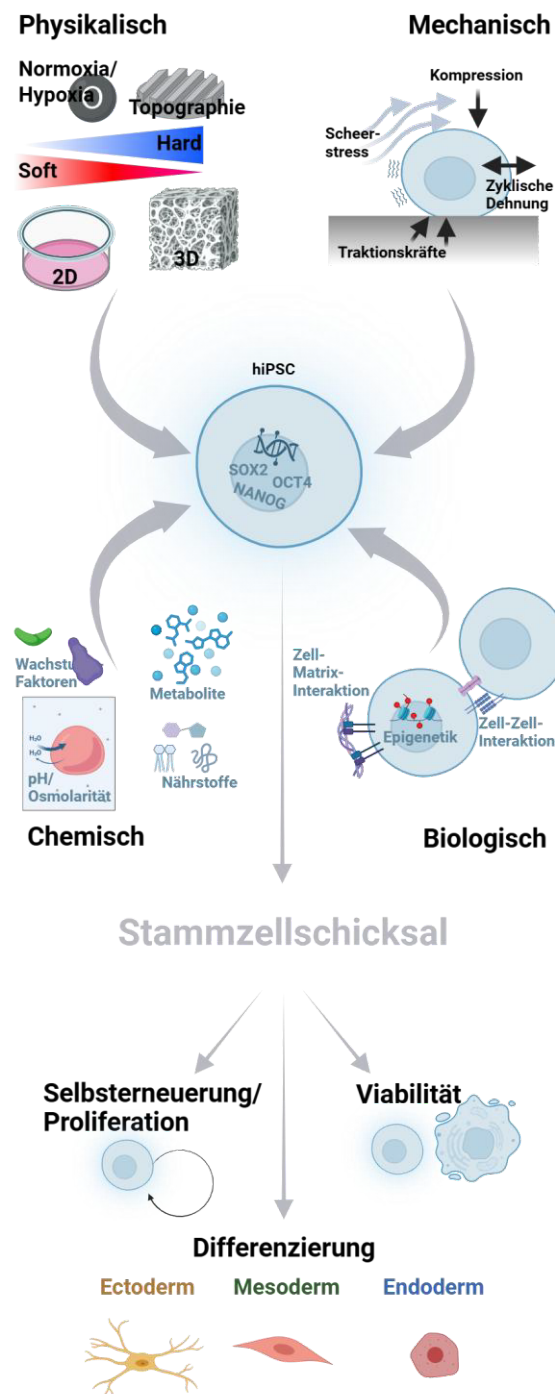


Abb. 1: Biologischer Hintergrund: Beeinflussung des Schicksals der Stammzellen durch externe Stimuli, Kultivierungsbedingungen und Umgebungseigenschaften. Graphik erstellt mit Biorender.com

Vor diesem Hintergrund adressiert die vorliegende Arbeit die Frage, wie Sensorisierung und

datenbasierte Prozessüberwachung in Mini-Bioreaktoren konkret umgesetzt und für die Kultivierung von hiPSCs nutzbar gemacht werden können.

Anforderungsprofil

Das Anforderungsprofil (vereinfachte Darstellung) eines (miniaturisierten) Bioreaktors adressiert zentrale biologische und prozesstechnische Randbedingungen kultivierter Zellsysteme. Zur Minimierung schädlicher Scherkräfte werden rotierende Reaktionsgefäße mit integrierten Strombrechern eingesetzt. Eine konstante Kultivierungstemperatur von 37°C wird durch einen temperaturkontrollierten Inkubationsraum in Kombination mit einer variablen Perfusionsrate sichergestellt. Die Stabilisierung des pH-Werts im Zellkulturmedium erfolgt über Puffersubstanzen und eine einstellbare CO₂-Konzentration im Inkubationsraum sowie eine kontinuierliche Perfusion zum Abtransport metabolischer Säuren wie Laktat. Für die Sterilisierbarkeit des Systems kommen biokompatible, autoklavierbare Materialien zum Einsatz, die bevorzugt additiv gefertigt werden.

Implementierung

Miniaturisierter Bioreaktor und Perfusion

Als Kultivierungssystem wurde ein miniaturisierter Suspensionsbioreaktor (Basis: OMNI-Life Science GmbH & Co.KG, Bremen; CERO-Bioreaktorsystem) eingesetzt, der Reaktorgefäße mit einem Arbeitsvolumen von bis zu 50 ml aufnimmt. Diese Bioreaktoren ermöglichen sowohl die Skalierung von Prozessen als auch das Screening verschiedener Bedingungen durch parallele Nutzung.

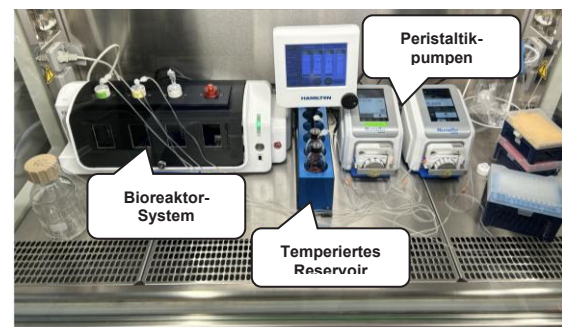


Abb. 2: Aufbau des modifizierten Bioreaktorsystems mitsamt Fluidik unter einer Sterilbank.

Zur Realisierung einer kontinuierlichen Perfusion wurde das Bioreaktorsystem durch einen 3D-gedruckten Deckel sowie modulare Fluidik-Einsätze erweitert. Die Einsätze erlauben es, Zu- und Abflüsse des Kulturmediums

voneinander zu trennen und sind zudem so gestaltet, dass sie mit zusätzlichen Sensoren ausgestattet werden können (siehe Abbildung 2).

Zellkultur, Prozessführung und Sterilität

Die Stammzellen (hiPSCs-Linie UKKi011-A, ebisc.org/UKKi011-A) wurden für die Experimente entweder als 2D-Monolayer oder als Sphäroide in mTeSR1-Medium (STEMCELL Technologies, Vancouver) kultiviert (siehe hierzu auch [4]). Für die Sphäroidkultur wurden Einzelzellen mit definierten Inokulationsdichten in den Bioreaktor eingebracht. Die Kultivierung erfolgte unter kontinuierlicher Rotation, initialer Zugabe von Apoptose-Inhibitoren (ROCK-Inhibitor) und anschließend entweder mit diskontinuierlichem mehrfachen Mediumwechsel oder perfusionsbasierter Prozessführung. Die Sterilität der Komponenten wurde durch sterile Verbrauchsmaterialien, eine nachträgliche Sterilisation durch Autoklavieren oder Ethanolbehandlung sichergestellt. Darüber hinaus wurde der modifizierte Bioreaktor in einer Sicherheitswerkbank aufgebaut, um das Risiko einer Kontamination während der Kultivierung zu minimieren. Die Implementierung eines Perfusionsystems erfolgte über Peristaltikpumpen (Masterflex Ismatec, Gelsenkirchen) mit definierten Flussraten im Bereich von $\leq 0,2$ ml pro min. Das Kulturmedium wurde aufgrund der Wachstumsfaktoren in einem Metallblock gekühlt (KK4TBE Kühlgerät, Labortechnik Medingen, Arnsdorf). Schlauchmaterialien und Innendurchmesser wurden hinsichtlich Autoklavierbarkeit, Totvolumen und Langzeitstabilität systematisch evaluiert.

Prozessüberwachung und Analytik

Die Prozesscharakterisierung umfasste Temperaturmessungen zur Bestimmung der Wärmeübertragung im modifizierten Bioreaktor sowie zur Abschätzung maximal zulässiger Perfusionsraten. Die metabolische Analytik (Glukose, Laktat) erfolgte mittels Offline-Referenzmessungen sowie einer explorativen morphologischen Analyse von Sphäroiden (Größe, Rundheit, Kompaktheit) durch automatisierte Bildauswertung. Die Pluripotenzbewertung erfolgte durch qPCR, Durchflusszytometrie sowie ergänzend durch Lebendzellfärbungen. Ebenso wurden bildbasierte Datensätze für das Training von Machine-Learning-Modellen (CNNs) erzeugt, um morphologische Merkmale automatisiert mit dem Pluripotenzstatus zu korrelieren.

Ergebnisse und Diskussion

Charakterisierung des perfusionsfähigen Mini-Bioreaktors

Das modifizierte Bioreaktorsystem besitzt einen 3D-gedruckten Deckel, der die fluidischen Komponenten mit den vier Bioreaktorgefäßen verbindet. Um eine sterile Prozessführung zu gewährleisten und das Kontaminationsrisiko zu minimieren, befindet sich der gesamte Perfusionsaufbau in einer Sterilbank unter laminarer Luftströmung. Durch die Modifikation des Systems zeigte sich anhand der Zellzahlen, dass die Reproduzierbarkeit zwar mit dem Referenzsystem vergleichbar ist, jedoch insgesamt eine reduzierte Zellzahl erzielt wird (siehe Abbildung 3). Temperaturmessungen zeigten, dass bei geeigneten Flussraten eine konstante Kulturtemperatur (> 36 °C) aufrechterhalten werden kann, obwohl das zugeführte Medium gekühlt gelagert wird. Die Durchmischung im Reaktorgefäß entsprach bei Rotation weitgehend dem theoretisch erwarteten Verhalten, während ohne Rotation Hinweise auf einen fluidischen Kurzschluss beobachtet wurden.

Die Anpassung des Mini-Bioreaktors ermöglichte in einigen Durchläufen eine stabile Perfusionskultur über mehrere Tage. In anderen Fällen traten jedoch Schwierigkeiten bei der gleichmäßigen Zufuhr und Abfuhr des Mediums auf, die in der Konsequenz ein Redesign von manchen Komponenten erfordert.

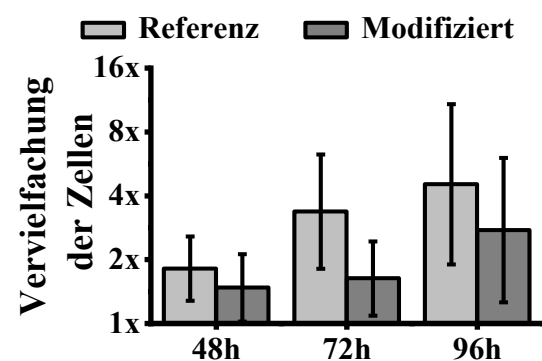


Abb. 3: Veränderung der Zellzahl dargestellt als Vervielfachung gegenüber 24 h mit ($n=3$) und ohne ($n=4$) fluidischem Einsatz. Es sind Mittelwert und Standardabweichung dargestellt.

Analyse von Metaboliten und Morphologie

Sowohl in 2D- als auch in 3D-Kulturen zeigte sich, dass ein kontinuierlicher bzw. häufigerer Mediumwechsel (manuell) zu einem erhöhten kumulierten Glukoseverbrauch und einer stabileren Laktatentwicklung führt. In 3D-Sphäroidkulturen kam es ohne regelmäßigen Mediumwechsel zu einem frühzeitigen Abfall der Wachstumsrate und teilweise zu negativen Wachstumsraten, insbesondere bei hohen Inokulationsdichten. Diese Effekte konnten durch

einen mehrfach durchgeführten manuellen Medienwechsel deutlich abgeschwächt werden. Abbildung 4 zeigt exemplarisch die Korrelation von Zellzahlentwicklung, Glukosekonzentration mit und ohne Medienwechsel.

Die Offline-Analyse der Glukose- und Laktatverläufe zeigte dabei eine Korrelation mit der Zellzahl, insbesondere bei höheren Inokulationsdichten. Allerdings war die Vorhersagekraft einzelner Metaboliten für z.B. einen Softsensor limitiert, was auf die Notwendigkeit multivariater Modelle hinweist. Erfolgt kein Medienwechsel, stagniert die Zellzahl bereits nach 48 h.

Die morphologische Analyse ergab, dass pluripotente Sphäroide eine hohe Rundheit, homogene Textur und begrenzte Größenverteilung aufweisen. Differenzierte oder gestresste Kulturen zeigten hingegen unregelmäßige Formen und erhöhte Variabilität. Machine-Learning-Modelle konnten diese Unterschiede in ersten Ansätzen zuverlässig klassifizieren und zeigten eine gute Übereinstimmung mit etablierten biologischen Endpunktanalysen.

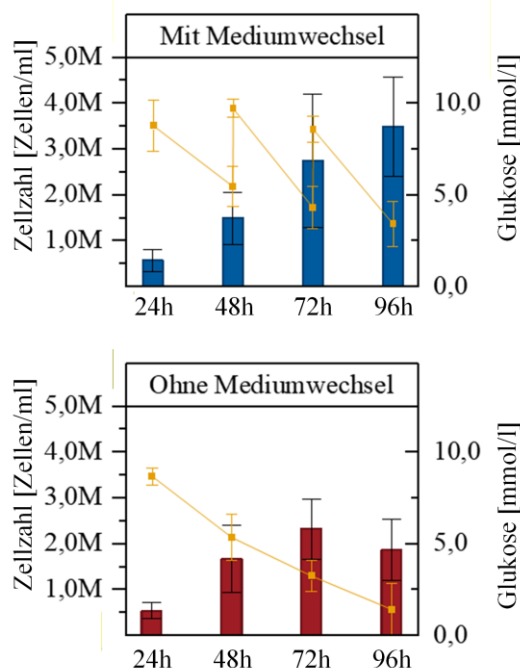


Abb. 4: Expansion in Suspensionsbioreaktoren von hiPSC bei Inokulation von 750.000 Zellen/ml und Verlauf der Glukosekonzentrationen mit und ohne Medienwechsel.

Die Ergebnisse verdeutlichen, dass miniaturisierte Bioreaktoren ein hohes Potenzial für die standardisierte hiPSC-Kultivierung besitzen, dieses jedoch maßgeblich von der verfügbaren Prozessüberwachung abhängt. Die Perfusionsskultur reduziert diskontinuierliche Schwankungen, ersetzt jedoch keine aktive Prozesskontrolle. Metabolische Parameter wie Glukose und

Laktat eignen sich als globale Indikatoren des Kulturzustands, sind aber allein nicht ausreichend, um komplexe Qualitätsmerkmale wie Pluripotenz zuverlässig abzubilden. Die Kombination aus metabolischer Analytik, physikalischer Sensorik und automatisierter Bildauswertung stellt daher einen vielversprechenden Ansatz dar. Insbesondere KI-gestützte Methoden bieten die Möglichkeit, subtile morphologische Veränderungen frühzeitig zu erkennen und mit biologischen Qualitätskriterien zu verknüpfen. Es zeigt sich, dass eine solche multimodale Sensorisierung den Übergang von reaktiver zu prädiktiver Prozessführung ermöglichen kann.

Ausblick

Zukünftige Forschungs- und Entwicklungsarbeiten werden die hier vorgestellten Ansätze und Erfahrungen in einem vollständig integrierten System zusammenführen. Dazu zählen robuste In-line- oder At-line-Messungen metabolischer Parameter, die Echtzeit-Auswertung morphologischer Daten sowie deren Kopplung an automatisierte Regelungsstrategien, etwa zur dynamischen Anpassung der Perfusionsrate. Langfristig kann so eine teil- bis vollautomatisierte Plattform entstehen, die eine reproduzierbare, skalierbare und regulatorisch anschlussfähige Kultivierung humaner Stammzellen im Miniaturmaßstab ermöglicht.

Danksagung

Diese Arbeit wurde vom Ministerium für Wirtschaft, Innovation, Digitales und Energie des Saarlandes (MWIDE) und der Fraunhofer-Gesellschaft (Leistungszentrum „Sensor-Intelligenz“) unterstützt.

Literaturnachweis

- [1] K. Takahashi et al., *Cell* 131, 861–872 (2007); doi: 10.1016/j.cell.2007.11.019
- [2] S. Yamanaka, *Cell Stem Cell* 33, 372–381 (2026); doi: 10.1016/j.stem.2026.02.003
- [3] B. Fischer et al., *Journal of Biomedical Materials Research Part A* 111, 1600–1615 (2023); doi:10.1002/jbm.a.37558
- [4] C. K. Kwok et al., *Reproductive Toxicology* 112, 23–35 (2022); doi:10.1016/j.reprotox.2022.05.007
- [5] C. Kropp et al., *Stem Cells Translational Medicine* 5, 1289–1301 (2016); doi:10.5966/sctm.2015-0385
- [6] S. J. Reyes et al., *Processes* 10, 189 (2022); doi:10.3390/pr10020189
- [7] Q. D. Vo et al., *PLoS ONE* 19, e0302537 (2024); doi: 10.1371/journal.pone.0302537

From Source to Feature: Multimodal Data Preparation for Machine Learning

Dmitry Amelin¹, Matthias Jochem¹, Gizem Aydogan^{1,2}, Michael M. Gepp¹, Roman Ruff¹

¹Fraunhofer Institute for Biomedical Engineering IBMT, 66280 Sulzbach, Germany

²Saarland University, 66123 Saarbrücken, Germany

Abstract

Three-dimensional spheroid and organoid cultures are increasingly used in biomanufacturing due to their ability to mimic in-vivo tissue behavior, yet current analytical methods remain labor-intensive and insufficient for real-time process understanding. In this work, we present a multimodal data-ingestion and preparation framework that integrates sparse morphological features extracted from YOLO-based microfluidic imaging with dense electrochemical sensor measurements and associated metadata. Dedicated ingestion pipelines for imaging and biochemical data were developed, followed by structured preprocessing steps that harmonize these heterogeneous streams into a unified temporal representation. A human-in-the-loop labeling workflow ensures high-quality annotations for model development. This approach lays the groundwork for scalable, automated spheroid production by enabling early detection of abnormal growth, supporting rapid decision-making in the laboratory, and improving reproducibility and quality assurance in organoid manufacturing.

Keywords: bioprocess monitoring, organoids, multimodal data, data ingestion

Motivation

Three-dimensional (3D) spheroid and organoid cultures have become indispensable model systems in drug discovery, toxicology, and regenerative medicine due to their ability to replicate key aspects of in-vivo tissue physiology, including cell to cell interactions, nutrient gradients, and complex self-organized morphology [1], [2], [3]. As these systems move toward automated and scalable bioprocessing, the need for non-invasive, high-frequency monitoring technologies is increasing. Such approaches must capture dynamic growth behavior and morphological evolution without disturbing the culture.

Conventional analytical methods (e.g., manual microscopy, offline biochemical assays, or endpoint destructive measurements) are labor-intensive, disruptive, and poorly suited to real-time process understanding. This hampers optimization and limits the deployment of advanced 3D culture workflows at scale.

At the Fraunhofer Center for Sensor Intelligence (ZSI), integrated optical, microfluidic, and electrochemical sensor modules have been developed to enable continuous, AI-supported monitoring of spheroids in bioreactors [4]. Microfluidic sampling allows the extraction and imaging of spheroids without spheroids loss and culture conditions interruption. Parallel electrochemical sensors measure metabolic

analytes (glucose, lactate, oxygen, pH) at approximately 2 min intervals. Together, these systems generate rich multimodal data streams representing both morphological and metabolic states.

However, these data sources differ significantly in structure and sampling frequency. Their integration into predictive bioprocess models therefore remains a major challenge. Specifically, such a framework could enable:

- Early detection of aberrant growth trajectories,
- Scheduling of media exchanges,
- Real-time decision support,
- Ensuring reproducibility and quality in organoid- and spheroid-based manufacturing.

Because biological growth is driven by dynamic interactions between environmental, metabolic, and morphological variables, this becomes a time-series forecasting problem requiring the fusion of heterogeneous inputs. Modern machine learning models (e.g., RNNs, GRUs/LSTMs, TCNs, transformers) provide a promising solution.

To address these challenges, this work focuses on establishing coordinated data-ingestion workflows and a structured preparation framework that can incorporate YOLO-derived morphological features, dense electrochemical time-series measurements, and optional

bioprocess metadata. This structured approach creates the foundation on which forecasting models can be trained, enabling prediction of morphological characteristics of spheroids.

Data Acquisition

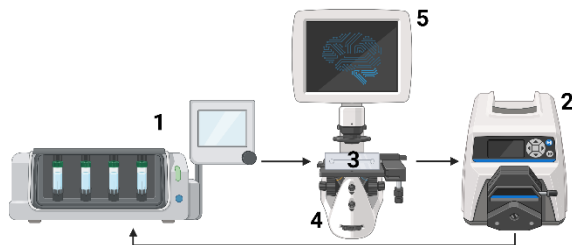


Figure 1 Schematics representation of the camera-based sensor. (The diagram was created with BioRender.com)

Figure 1 schematically depicts the experimental setup. Spheroid cultures were maintained in a bench-scale suspension bioreactor (1) equipped with an automated at-line sampling system. A peristaltic pump (2) transfers spheroids into a transparent channel (3), where they are imaged by a microscope (brightfield, phase-contrast, or fluorescence) (4). After acquisition, the spheroids are returned to the bioreactor. The images were processed on an Nvidia Jetson Nano single-board computer (5) in real-time during acquisition using a model based on You Only Look Once object segmentation architecture [5]. The model generated segmentation masks for spheroids, which can be seen on Figure 2. From each mask, morphological characteristics were computed, including area, diameter, perimeter, and circularity. Electrochemical sensors integrated into the bioreactor or alternatively integrated in-line with the imaging system provided continuous measurements of glucose, lactate, oxygen, and pH at around 2 min intervals. Figure 3 shows the compact measurement electronics for the electrochemical sensors with a diameter of 30 mm.

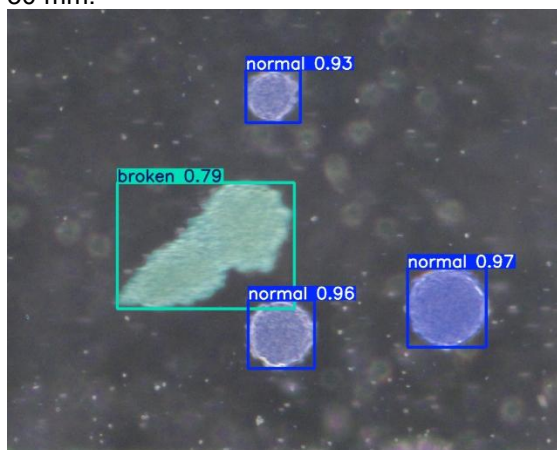


Figure 2 An example of spheroids segmentation

Because the electrochemical and imaging subsystems differ in throughput and metadata requirements, two separate data pipelines were implemented. A graphical user interface application for the electrochemical sensors receives data from the electronics via Bluetooth Low Energy and forwards it to the database. For microscopic pictures, due to the high volume of image data and additional metadata needs, a dedicated gRPC-based image transfer pipeline was developed. Each image and its metadata were serialized and sent to a middleware service running on the database server. The middleware validated metadata and orchestrated storage, separating raw images (stored in a dedicated high-volume repository) from structured metadata stored in PostgreSQL.

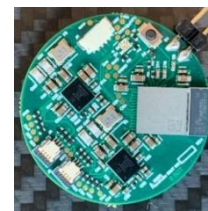


Figure 3 Electrochemical measurement electronics, 30 mm in diameter

Data Storage

All structured data, including sensor values, metadata, sampling times, and experimental parameters, were stored in PostgreSQL. Time-series measurements were organized in a dedicated measurement table, supplemented by a sensor metadata table for fast contextual lookup. Images were not stored as binary objects in PostgreSQL due to database bloat and slow retrieval associated with large-object storage [6]. Instead, only file paths and metadata were stored in the database, ensuring efficient and scalable data management. Additionally, all information about spheroids detected on images were stored in a separate table.

Both the electrochemical and image related measurement tables contain a foreign-key reference to a table with the experiment metadata. This way we can find related measurements for ML model training and validation.

Labeling

As previously stated, in addition to storing raw images, the segmentation output for each detected spheroid, including the contour, bounding-box, and class information, is recorded in a dedicated table within the database. To support systematic validation and correction of these annotations, a web-based interface was developed that allows biologists

to review segmentation and classification results. If the ML model produces incorrect or incomplete segmentations, users can flag the corresponding images directly through the interface. Flagged samples are then exported to dedicated labeling software for manual correction. After review, the corrected segmentation masks and class labels are re-imported and used to update the database records.

This workflow enables efficient human-in-the-loop curation of the dataset and ensures high-quality ground truth labels for both model training and subsequent validation cycles.

Outlook

The next major challenge is identifying the most suitable model architecture for combining sparse imaging data with dense electrochemical measurements to predict future morphology. We are going to focus on the following two approaches.

To synchronize the datasets, temporal windows of electrochemical data preceding each imaging event will be extracted (e.g., 12 h, 24 h, 48 h). Each window will be encoded into a compact representation using a dedicated time-series encoder such as one-dimensional convolutional neural network (1D CNN), Long short term memory (LSTM), gated recurrent units (GRUs), or Temporal Convolutional Network (TCN).

For each imaging timepoint, t , the morphology vector, $M(t)$, will be concatenated with the electrochemical embedding, $E(t)$, to form a joint feature vector:

$$X(t) = [M(t), E(t)]$$

As an alternative, a Cross-Attention Fusion strategy will be evaluated. Here, the full electrochemical time series is retained, sparse and dense measurements are encoded separately, and a cross-attention layer identifies relevant metabolic patterns conditioned on morphological features.

In both cases, the fused representation will be passed to a multi-layer perceptron trained to predict the spheroid morphology at the next imaging time point.

With this work, we aim to establish a foundation for scalable spheroid production by enabling early prediction of abnormal growth trajectories, supporting biologists in making faster and more informed decisions, and ultimately improving reproducibility and quality assurance in organoid manufacturing.

Acknowledgement

This work was supported by the Saarland Ministry of Economic Affairs, Innovation, Digital and Energy (MWIDE) and the Fraunhofer

Gesellschaft (High-Performance Center 'Sensor- Intelligence').

Bibliography

- [1] S. Chen, X. Chen, Z. Geng, and J. Su, "The horizon of bone organoid: A perspective on construction and application," *Bioact Mater*, vol. 18, pp. 15–25, Dec. 2022, doi: 10.1016/j.bioactmat.2022.01.048.
- [2] W. Hu and M. A. Lazar, "Modelling metabolic diseases and drug response using stem cells and organoids," *Nat Rev Endocrinol*, vol. 18, no. 12, pp. 744–759, Dec. 2022, doi: 10.1038/s41574-022-00733-z.
- [3] O. Kopper *et al.*, "An organoid platform for ovarian cancer captures intra- and interpatient heterogeneity," *Nat Med*, vol. 25, no. 5, pp. 838–849, May 2019, doi: 10.1038/s41591-019-0422-6.
- [4] D. Amelin *et al.*, "Bioreactor Intelligence for Real-Time Monitoring and Automation of Stem Cell Cultivation," *Lectures*, pp. 137–138, May 2025, doi: 10.5162/SMSI2025/C2.2.
- [5] G. Jocher, J. Qiu, and A. Chaurasia, *Ultralytics YOLO*. (2023). [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [6] R. Sears, C. Ingen, and J. Gray, "To BLOB or Not To BLOB: Large Object Storage in a Database or a Filesystem?," *Computing Research Repository - CORR*, Jan. 2007.

Measurement System for Investigating the Long-Term Stability of Permanent Magnets for Sensor and Actuator Systems

M.Lanz^{1,2}, G.Martinek¹, T.Kresse¹, A.Georgiadis², G.Schneider¹ and D. Goll¹

¹ *Materials Research Institute, Aalen University, Aalen, Germany*

² *Faculty of Management and Technology, Leuphana University, Lüneburg, Germany*

The efficient long-term use of permanent magnets requires reliable prediction of temperature-, field-, and time-dependent magnetization losses. This study presents an automated Helmholtz coil measurement setup to quantify time- and temperature-dependent polarization changes in NdFeB samples of various grades and geometries. The measurements distinguish time- and temperature-dependent changes between reversible and irreversible contributions, enabling determination of a temperature-based magnetic viscosity parameter S_v^T for prediction of future losses. In addition, temperature-based magnetic stabilization effectively eliminates irreversible losses and enables constant magnetic stability. A combination of both approaches supports optimizing the design of magnetic systems.

Keywords: Permanent magnets, long-term prediction, viscosity parameter, magnetic stabilization

1. Introduction

Permanent magnets are essential in applications such as electric motors, generators, sensors, and actuators [1]. They all share a common challenge: their magnetization must remain stable under varying field and temperature conditions over time, as instability can lead to reduced precision in sensors or decreased efficiency in, e.g., electric motors. Accounting for such magnetic losses typically relies on the magnetic viscosity parameter S_v , the ratio of time-dependent to direct field- or temperature-induced losses [2]. The value of S_v enables future loss prediction through the use of the fluctuation field [3]. However, the determination of S_v , mainly by vibrating sample magnetometry (VSM) is costly, technically demanding, and sensitive to uncertainties especially when adjusting from external to internal fields [4], [5]. As a simple alternative approach, a novel measurement approach based on temperature rather than field as the influencing quantity has been developed [6]. It enables quantification of temperature-driven polarization changes and their time-dependent decreases, yielding a temperature-based viscosity parameter S_v^T as an alternative to the established field-based S_v approach [4]. Using only temperature as the influencing quantity simplifies interpretation and improves applicability for system designers. Although the viscosity parameter enables future loss prediction, it does not suppress time-dependent losses. Thermal stabilization through pre-aging offers a complementary approach by shifting the magnet into a lower but stable polarization

state prior to application [7], [8]. This work therefore presents the measurement setup, the determination and application of the temperature-based viscosity parameter, and the complete suppression of time-dependent losses through thermal stabilization.

2. Background

In contrast to soft magnetic materials such as steel or electrical sheet, permanent magnets are designed for high remanent polarization J_r , the polarization at zero external field and high coercive field strength H_{cj} , the field-strength at $J = 0$. These properties make them an effective alternative to electromagnets, which require soft magnetic cores, copper windings, and external power supply. In open-circuit conditions (e.g. in air at zero external field), permanent magnets exhibit its self-demagnetizing field [9] quantified as

$$H_{\text{self}} = -\frac{N \cdot J}{\mu_0} \quad (1)$$

where N is the geometry-dependent demagnetization factor [10], J is the magnetic polarization of the sample, and μ_0 is the vacuum permeability. The presented setup targets open-circuit applications, using temperature variations as the influencing factor resulting in $J(T, t)$ curves showing polarization as a function of temperature T and time t . These polarization states correspond to the intersections of the material $J(H_{\text{int}})$ curves at specific temperatures with the sample load line as visualized in Fig. 1.

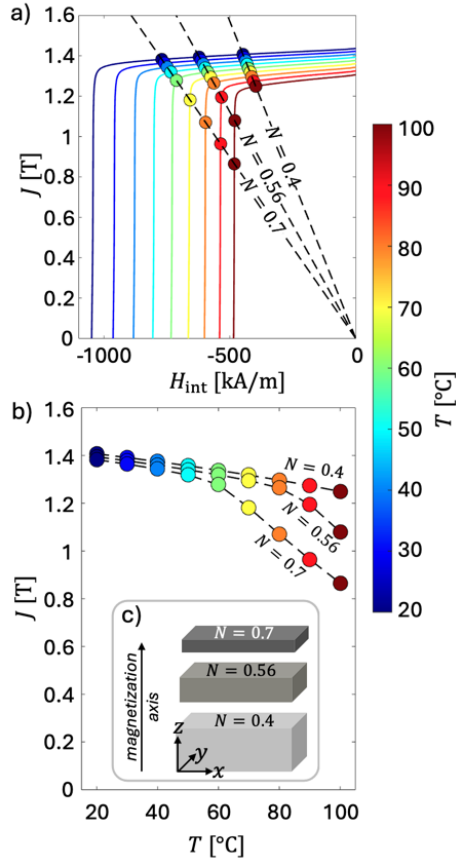


Fig. 1: $J(H)$ curves of an N52 sample illustrating the influence of different sample temperatures and sample geometries in a) (demagnetization factors $N = 0.4$ ($20 \times 10 \times 10 \text{ mm}^3$), 0.56 ($20 \times 10 \times 5 \text{ mm}^3$), 0.7 ($20 \times 10 \times 2 \text{ mm}^3$), overview in c). In b), the resulting $J(T)$ curves are obtained from the intersections of the three load lines with the $J(H)$ curves at increasing temperatures from a).

Fig. 1a shows $J(H_{\text{int}})$ curves at temperatures from 20 to 100 °C together with three sample load lines for different sample geometries with $N = 0.4$, $N = 0.56$ and $N = 0.7$, as shown in Fig. 1c. Their intersections, indicated by colored circles, define the open-circuit $J(T)$ curves in Fig. 1b. Polarization decreases with increasing temperature, with samples of lower demagnetization factors (e.g. $N = 0.4 < 0.56 < 0.7$) or higher coercivity retaining higher polarization at elevated temperatures. For quantifying the maximum temperature stability, the transition temperature T_t , marked in Fig. 2a, is defined as the tangent intersection between the reversible and irreversible regions of the $J(T)$ curve. Starting at room temperature, temperature-induced changes are mostly reversible below T_t . Beyond T_t , losses become predominantly irreversible. In addition to these instantaneous changes, irreversible time-

dependent polarization decreases are critical for long-term magnet performance. Fig. 2b illustrates this effect as the decay from the instantaneous polarization states to the time-dependent polarization decrease $\Delta J^{t.d.}$, representing polarization reduction during prolonged exposure times at the respective temperature due to thermal activation (magnetic viscosity). Consequently, temperatures above T_t should be avoided, as even small variations can induce significant irreversible polarization losses with pronounced time-dependent behavior.

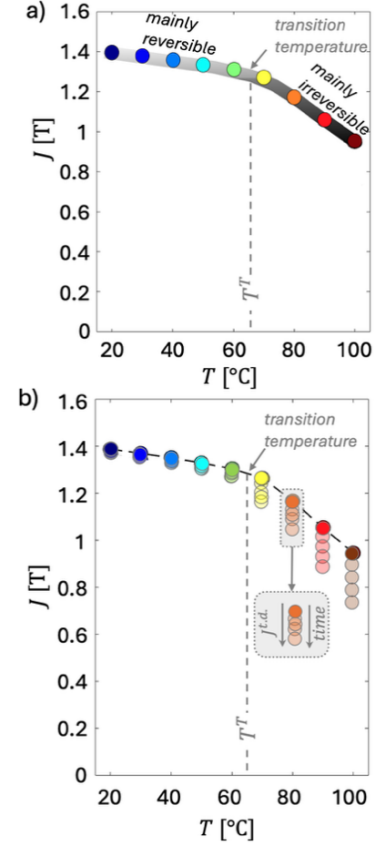


Fig. 2: Schematic $J(T)$ curve illustrating the stability of polarization. a) Section of the $J(T)$ curve highlighting a linear region with mainly reversible changes and an irreversible region beyond the transition temperature with mainly irreversible losses. b) Starting at the transition temperature, additional time-dependent decreases can be expected for longer exposure durations at each temperature step.

3. Experimental set up

3.1 Measurement Setup

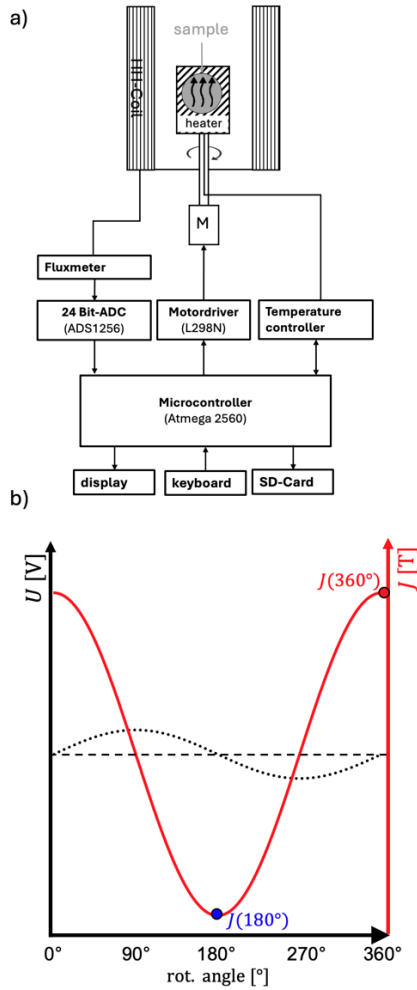


Fig. 3: Measurement setup with a) temperature-controlled sample rotating inside a Helmholtz measurement coil with digital measurement and control circuitry. b) shows the induced voltage and time-integrated polarization signal over one rotation. The induced voltage is integrated by a fluxmeter, digitized, and stored with process parameters on a SD card. Stepwise motor rotation enables discrete data acquisition and angle-dependent polarization extraction by averaging the minimum and maximum extrema, e.g. $(J(180°) + J(360°))/2$ (red curve in b)).

The measurement setup for determining $J(T, t)$ values, shown in Fig. 3a, uses a permanent magnet sample rotated along its magnetization direction inside a Helmholtz measurement coil (MS210, Magnet-Physik GmbH, Germany) [6]. The rotating sample induces a voltage in the measurement coil, which is integrated over time by a fluxmeter (EF-14, Magnet-Physik) to obtain

the magnetic moment. The fluxmeter signals are digitized by a 24-bit ADC and processed by an Atmega 2560 microcontroller, which also controls the stepper motor and communicates with the temperature controller. The stepper motor, with a 1:10 gear, provides 1800 steps per rotation and, together with the ADC, achieves a noise level of approximately 0.00015 T at room temperature and 0.0002 T at 100 °C. The sample rotation, as shown in Fig. 3b, further improves angular accuracy and reduces fluxmeter drift by averaging signal peaks compared to standard manual Helmholtz coil measurements. Heating is achieved via a PTC heating foil (12 V, 40 W) on an aluminum plate, reaching 260 °C with ± 0.1 °C stability.

3.2 Magnetic long-term stability

All permanent magnets exhibit time-dependent effects. The closer the operating temperature is to the transition temperature, the higher the time-dependent losses as shown in Fig. 2b. The measured temperature- and time-dependent polarization decrease enables prediction of future magnetization states, based on the value of the temperature-based viscosity parameter S_v^T presented in section 3.2.2 and the fluctuation temperature T_f , presented in section 3.2.3. Another approach is to eliminate time-dependent losses through magnetic stabilization, presented in the following section 3.2.1. Here, the time-dependent decrease represents the instability region, which is reduced by driving the magnet into a lower, but stable magnetization state. Since this pre-aging step lowers the magnetization level, it is not suitable for all applications, particularly those requiring maximum polarization values or where small time-dependent losses are acceptable.

3.2.1 Magnetic stabilization

Magnetic stabilization, or pre aging, refers to the controlled partial demagnetization of a permanent magnet to a lower but stable polarization state prior to use [7], [8]. This prevents further irreversible losses, including direct and time-dependent effects. Fig. 4 shows a schematic polarization-to-time $J(t)$ course in a) and the set temperature profile $T(t)$ in b). Without stabilization, the sample polarization decreases from point A to B at the maximum application temperature T^{app} . Applying a stabilization temperature $T^{stab} > T^{app}$ reduces the polarization to a lower but stable state, e.g. J^c at point C. Further temperature variations up to this stabilized temperature result in only reversible changes during operation without additional irreversible losses.

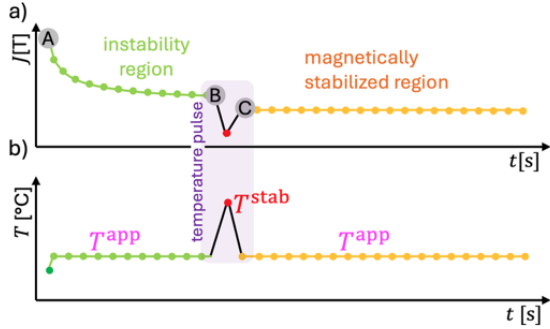


Fig. 4: Schematic visualization of the magnetic stabilization procedure. a) Polarization time course $J(t)$. b) set temperature profile $T(t)$, illustrating the time-dependent decrease of polarization at the application temperature T^{app} within the instability region. This polarization decrease can be suppressed by a temperature pulse T^{stab} , e.g. from point B to C, resulting in a reduced but stable polarization state (stabilized region).

3.2.2 Magnetic viscosity parameter

The magnetic viscosity parameter quantifies the permanent magnet's resistance to external influences over time. Originally defined for magnetic fields as S_v [2], we adapted it for temperature, introducing the temperature-based viscosity parameter S_v^T . S_v^T is measured using a temperature ramp with stepwise heating and constant-temperature phases to determine irreversible polarization decreases. The general measurement procedure for determining S_v^T is shown in Fig. 5.

First, the temperature is increased from T^A to T^B , denoted as $\Delta T^{A,B}$. This causes an immediate polarization change ΔJ^{tot} , which defines the total temperature-dependent response:

$$\frac{\Delta J^{\text{tot}}}{\Delta T^{A,B}} = \frac{J^A - J^B}{T^A - T^B} \quad (2)$$

During the subsequent constant-temperature phase at T^B , the decrease of the irreversible time-dependent polarization $\Delta J^{\text{t.d.}}$ from J^B to J^C is used to determine the magnetic viscosity S as

$$S = \frac{\Delta J^{\text{t.d.}}}{\ln\left(\frac{t}{t_0}\right)}, \quad (3)$$

where t is the exposure time and t_0 the reference time marking the onset of time-dependent demagnetization. The reversible temperature response is measured by a small temperature change between T^C to T^D and is given by

$$\frac{\Delta J^{\text{rev}}}{\Delta T^{C,D}} = \left| \frac{J^D - J^C}{T^D - T^C} \right| \quad (4)$$

The irreversible temperature response is then calculated as:

$$\frac{\Delta J^{\text{irr}}}{\Delta T} = \frac{\Delta J^{\text{tot}}}{\Delta T^{A,B}} - \frac{\Delta J^{\text{rev}}}{\Delta T^{C,D}} \quad (5)$$

Finally, the temperature-based magnetic viscosity parameter S_v^T is calculated as:

$$S_v^T = \frac{S}{\left(\frac{\Delta J^{\text{irr}}}{\Delta T}\right)} [^\circ\text{C}] \quad (6)$$

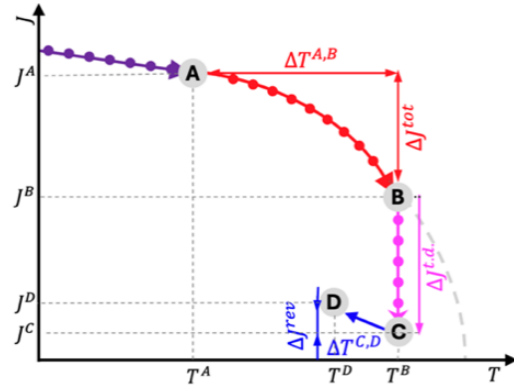


Fig. 5: Schematic measurement procedure for determining the temperature-dependent viscosity parameter S_v^T . The temperature increase $\Delta T^{A,B}$ from point A to point B is used to determine the corresponding total polarization decrease ΔJ^{tot} from J^A to J^B . At constant temperature T^B , the time-dependent decrease $\Delta J^{\text{t.d.}}$ leads from J^B (point B) to J^C (point C) over time t . Finally, a small temperature increase $\Delta T^{C,D}$ from point C to point D is applied to determine the reversible polarization change ΔJ^{rev} .

3.2.3 Prediction of long-term losses

Based on the value of the temperature-dependent magnetic viscosity parameter S_v^T , future polarization states can be estimated using a fictitious fluctuation temperature, analogous to the fluctuation field concept as described in [11] according to Néel [3]. The time-dependent polarization decrease $\Delta J^{\text{t.d.}}$, for example shown in Fig. 6 from J^B (point B) to J^C (point C), can also be described by adding an additional fictitious fluctuation temperature $T_f(t)$ to the maximum temperature T^B .

The fluctuation temperature can be quantified as the product of S_v^T and the logarithm of the exposure time t , defined as the duration of maximum temperature exposure, with t_0 denoting the reference time:

$$T_f(t) = S_v^T * \ln(t/t_0) \quad (7)$$

Adding $T_f(t)$ to the maximum temperature T_{max} , here T^B results in the effective temperature $T_{\text{eff}} = T^E$:

$$T_{\text{eff}}(t) = T_{\text{max}} + T_f(t) \quad (8)$$

The polarization value at T^E is J^E and therefore identical to the time-dependent decrease, i.e. $J^C = J^E$. Consequently, instead of measuring the time-dependent polarization decrease over

a long time, the fluctuation temperature approach allows determining the polarization state at present to predict its behavior under long-term future temperature exposures.

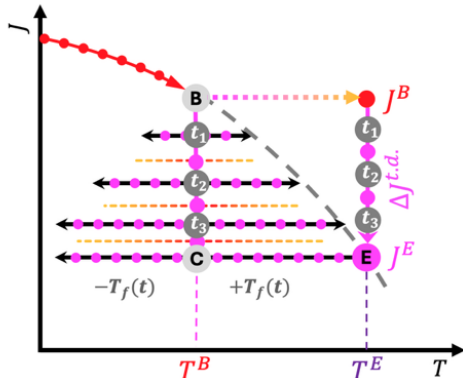


Fig. 6: Schematic visualization of the fictitious fluctuation temperature. The time-dependent decrease $\Delta J^{t.d.}$ reduces the polarization from J^B to J^C . The same polarization state can alternatively be reached by applying an effective temperature, defined as the maximum temperature T^B plus the fictitious fluctuation temperature $T_f(t)$ as a function of exposure time t , resulting in T^E with $J(T^E) = J^E$, which is equal to J^C .

4 Results and Discussions

4.1 Application of Permanent magnets for sensor systems

For the application of permanent magnets, two material grades were selected: (A1) N52 ($H_{cj} = 955$ kA/m) and (B) N52-SH ($H_{cj} = 1590$ kA/m), both with $J_r = 1.4$ T in disc geometry with a diameter of 40 mm and a height of 4 mm corresponding to a demagnetization factor of $N = 0.8$, while for comparing different geometries sample A2, is realized in block geometry with dimensions of $40 \times 10 \times 4$ mm³ and $N = 0.65$, as schematically shown in Fig. 7.

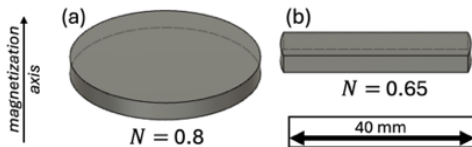


Fig. 7: Overview of the sample geometries. a) Disc sample with a diameter of 40 mm and a height of 4 mm, $N = 0.8$. b) Block sample with dimensions of $40 \times 10 \times 4$ mm³ and $N = 0.65$. In both cases, the magnetization direction corresponds to the 4 mm dimension.

The resulting $J(T)$ curves of these samples are shown in Fig. 8, together with the corresponding

transition temperatures T_t and the time-dependent decrease $J^{t.d.}$, measured over 18 min (1080 s) at selected temperatures.

It can be observed that sample A1 shows a transition temperature of around 25 °C, sample A2 around 50 °C, and sample B around 90 °C. The transition temperature marks the shift from mainly reversible to irreversible losses. Consequently, the time-dependent decrease also increases above the transition temperature, as shown by the $\Delta J^{t.d.}$ measurements: in the linear region, where mainly reversible losses occur, time dependency is nearly zero within the observed time frame, while a strong decrease appears beyond the transition temperature. The magnitude of $\Delta J^{t.d.}$ therefore reflects the time stability at each temperature: smaller time-dependent decreases correspond to higher stability of the sample polarization.

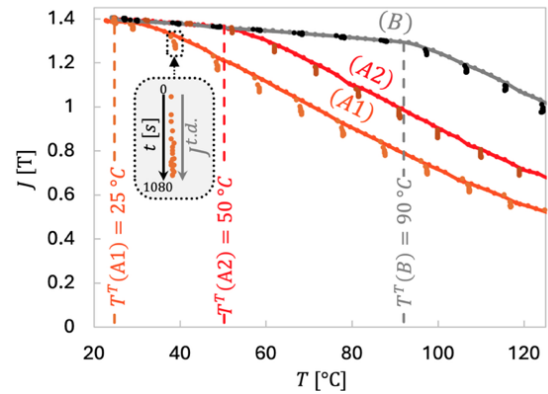


Fig. 8: $J(T)$ curves of selected materials A1 ($N = 0.8$, N52), A2 ($N = 0.65$, N52) and B2 ($N = 0.8$, N52-SH) with indicated transition temperatures. Time-dependent losses are indicated by $J^{t.d.}$ during 18 min of constant temperature exposure. An increasing time dependence is observed at temperatures above the transition temperature.

4.1.1 Magnetic stabilization

Since sample A1 exhibits a transition temperature of only 25 °C and sample A2 of 50 °C, both samples were temperature-stabilized using a temperature pulse of 65 °C (as an example), as shown in Fig. 9.

After stabilization, both samples show increased temperature resistance, with a new transition temperature of approximately 65 °C, effectively minimizing irreversible and time-dependent decreases up to this temperature. The polarization level of sample A1 was reduced to approximately 1.10–1.07 T, while that of sample A2 was reduced to approximately

1.30–1.27 T. The nearly zero time-dependent decrease within the extended reversible region confirms the effectiveness of the stabilization process.

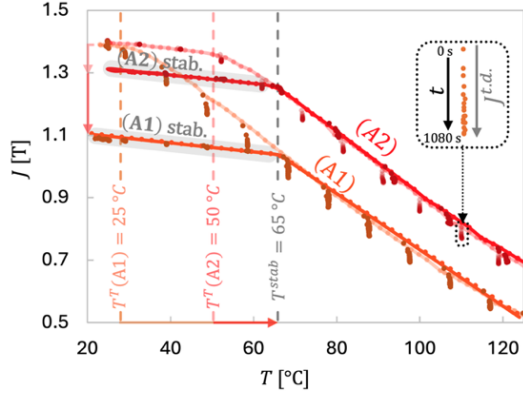


Fig. 9: $J(T)$ material curves of N52 samples A1 ($N = 0.8$), A2 ($N = 0.65$) illustrating the effect of temperature stabilization with 65 °C, extending the initial transition temperatures T_t from 25 °C for A1 and 50 °C for A2 to 65 °C, thereby eliminating irreversible time-dependent losses ($J^{t.d.}$) below this temperature.

4.1.2 Long-term prediction

The procedure for determining the temperature-based magnetic viscosity parameter described in Section 3.2.2 was repeated several times over a temperature range of 45 to 160 °C with increasing temperature of ΔT^{tot} and $t = 18$ min per cycle for samples A and B. The measurement results of S_v^T [°C] are shown in Fig. 10.

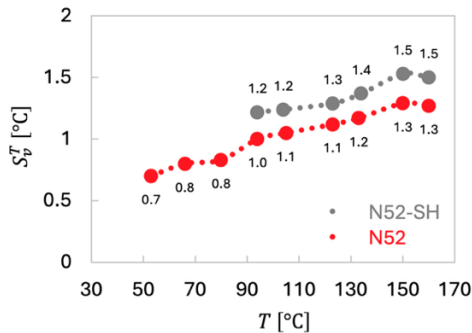


Fig. 10: Measurement results of S_v^T for the N52 sample (red) and the N52-SH sample (grey) over a temperature range from 45 °C to 160 °C. For the N52 sample, S_v^T varies from 0.7 °C at 50 °C to 1.3 °C at 160 °C. For the N52-SH sample, S_v^T varies from 1.2 °C at 100 °C to 1.5 °C at 160 °C.

The low-coercivity N52 sample exhibits S_v^T values of 0.7 °C at approximately 45 °C, increasing to 1.3 °C at approximately 160 °C, while the high-coercivity N52-SH sample shows

$S_v^T = 1.2$ °C at approximately 90 °C, rising to 1.5 °C at approximately 160 °C.

Missing values for N52-SH are mainly attributed to the difficulty of detecting the comparatively small time dependence at lower temperatures. The values of S_v^T can then be used, together with the previously described fluctuation temperature (section 3.2.3), to predict future polarization states, as shown in Fig. 11, based on the $J(T)$ curves from Fig. 8. The colored dashed lines with marked peak times ranging from 10 s up to 10 years are linked to the $J(T)$ material curves by corresponding colors. Their intersections with each effective temperature (as a function of exposure time) indicate the predicted polarization states based on $T_{\text{max}} = 40$ °C. At this maximum temperature, sample A1 is already above its transition temperature T_t and therefore does not provide time stability. Sample A2 can be used with only minor expected time-dependent irreversible losses for durations of up to approximately one week at T_{max} , as the resulting temperature remains just below the transition temperature. For longer durations, such as one year, the temperature exposure becomes critical, as the temperature exceeds the transition temperature, resulting in significant time-dependent irreversible losses. In contrast, sample B demonstrates good performance even after 10 years at T_{max} , with a projected effective temperature of maximum 65 °C, which is far below the transition temperature and therefore suitable for long-term use.

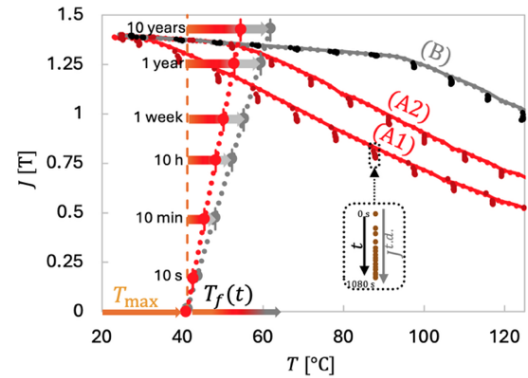


Fig. 11: Application of the fluctuation temperature $T_f(t)$ to predict long-term polarization states at a maximum temperature of $T_{\text{max}} = 40$ °C for exposure times t ranging from 10 s to 10 years. The colored dashed lines correspond to the respective $J(T)$ material curves. The value of $T_f(t)$ at a given exposure time t on the $J(T)$ curve indicates the predicted polarization state after that time.

5 Summary and Conclusions

A novel measurement setup has been developed to characterize the long-term stability of permanent magnets under open-circuit conditions and thermal influence, enabling evaluation of both direct and time-dependent polarization changes. Two complementary strategies are presented: thermal stabilization and long-term state prediction. For assessing long-term stability in magnetic systems, the following steps can be considered:

1. **Evaluation:** Determine whether the sample magnetization state can be reduced while still meeting the requirements of the intended application.
2. **Stabilization:** If reduction is possible, magnetic stabilization effectively eliminates time-dependent losses by operating the sample at a reduced polarization state.
3. **Viscosity-based prediction:** If a high magnetization state is required, the temperature-dependent viscosity parameter and the resulting fluctuation temperature allow straightforward prediction of future magnetization states of the magnetic material.

The use of both strategies provides a practical framework to mitigate time-dependent demagnetization, enabling informed material selection and efficient use of permanent magnets across a wide range of applications.

Acknowledgment

The research was funded by the Federal Ministry of Research, Technology and Space (BMBF/BMFTR), Germany, within the scope of the project Smart-MAG (grant no. 13FH4I08IA).

References

- [1] O. Gutfleisch, M. A. Willard, E. Brück, C. H. Chen, S. G. Sankar, and J. P. Liu, "Magnetic Materials and Devices for the 21st Century: Stronger, Lighter, and More Energy Efficient," *Advanced Materials*, vol. 23, no. 7, pp. 821–842, Feb. 2011, doi: 10.1002/adma.201002180.
- [2] R. Street and J. C. Woolley, "A Study of Magnetic Viscosity," *Proc. Phys. Soc. A*, vol. 62, no. 9, p. 562, Sep. 1949, doi: 10.1088/0370-1298/62/9/303.
- [3] L. Néel, "Le traînage magnétique," *J. Phys. Radium*, vol. 12, no. 3, pp. 339–351, 1951, doi: 10.1051/jphysrad:01951001203033900.
- [4] T. Kresse, G. Martinek, G. Schneider, and D. Goll, "The Field-Dependent Magnetic Viscosity of FeNdB Permanent Magnets," *Materials*, vol. 17, no. 1, Art. no. 1, Jan. 2024, doi: 10.3390/ma17010243.
- [5] M. Lanz, G. Martinek, T. Kresse, G. Schneider, and D. Goll, "Magnetic Viscosity Characterization of Fe–Nd–B Permanent Magnets Using a Novel Measurement Approach," *IEEE Transactions on Magnetics*, vol. 61, no. 4, pp. 1–5, Apr. 2025, doi: 10.1109/TMAG.2025.3544856.
- [6] M. Lanz, G. Martinek, T. Kresse, A. Georgiadis, G. Schneider, and D. Goll, "Application-Oriented In-Situ Testing Methods for Permanent Magnet Characterization," *SMSI-Nuremberg 2025*, vol. AMA-Science, pp. 327–328, 2025, doi: 10.5162/SMSI2025/P48.
- [7] M. Haavisto, H. Kankaanpää, T. Santa-Nokki, S. Tuominen, and M. Paju, "Effect of Stabilization Heat Treatment on Time-Dependent Polarization Losses in Sintered Nd-Fe-B Permanent Magnets," *EPJ Web of Conferences*, vol. 40, p. 06001, 2013, doi: 10.1051/epjconf/20134006001.
- [8] C. E. Webb, "The stability of permanent magnets," *Proceedings of the IEE - Part C: Monographs*, vol. 108, no. 14, pp. 317–324, Sep. 1961, doi: 10.1049/pi-c.1961.0043.
- [9] R. Hilzinger and W. Rodewald, *Magnetic materials: fundamentals, products, properties, and applications*. Erlangen : Hanau, Germany: Publicis ; VAC, VACUUMSCHMELZE, 2013.
- [10] A. Aharoni, "Demagnetizing factors for rectangular ferromagnetic prisms," *Journal of Applied Physics*, vol. 83, no. 6, pp. 3432–3434, Mar. 1998, doi: 10.1063/1.367113.
- [11] M. Lanz, G. Martinek, T. Kresse, G. Schneider, and D. Goll, "Evaluation of Time-Dependent Magnetic Losses of Permanent Magnets," *Metals*, vol. 16, no. 3, p. 322, Mar. 2026, doi: 10.3390/met16030322.

Der Einfluss der Abtasttakt drift auf unterabgetastete Laser-Doppler-Vibrometer-Signale

Rosver Daniel Sando Fogno¹, Niels Neumann¹, Christian Rembe¹
¹ TU Clausthal, IEI, 38678 Clausthal-Zellerfeld, Leibnizstraße 28, Germany
¹ sandofogno@iei.tu-clausthal.de

Zusammenfassung

Die Laser-Doppler-Vibrometrie (LDV) hat sich als Goldstandard für hochpräzise, berührungslose Schwingungsmessungen in der Strukturanalyse und der Zustandsüberwachung von Bauwerken etabliert. Ihr Einsatz in mobilen oder unbemannten Flugsystemen (UAS) ist jedoch durch den hohen Rechen- und Energiebedarf für die Verarbeitung hochfrequenter Signale zur Erzielung einer höheren Auflösung begrenzt. Die Bandpassunterabtastung ist eine vielversprechende Lösung, um die zu verarbeitende Datenmenge und den Energiebedarf zu reduzieren. Ein weiterer Vorteil besteht darin, dass keine analoge Mischstufe mehr notwendig ist, da die gewünschte Zwischenfrequenz direkt durch die Abtastung im geeigneten Nyquist-Bereich entsteht. Sie stellt jedoch strenge Anforderungen an die Stabilität des Abtasttakts. Diese Stabilität lässt sich auf UAS-Plattformen nur schwer gewährleisten, da hochstabile Oszillatoren aufgrund von Kosten, Baugröße und Leistungsaufnahme meist nicht einsetzbar sind. In dieser Arbeit wird der Einfluss einer Abweichung der Abtastfrequenz in einem Unterabtastungsszenario untersucht. Analytisch wird gezeigt, dass die daraus resultierende Frequenzabweichung im durch eine FFT berechneten Spektrum linear mit der Nyquist-Zonenordnung zunimmt. Dieses Verhalten wurde sowohl durch Simulationen als auch experimentell verifiziert und verdeutlicht eine kritische Verwundbarkeit von Unterabtastung-Verfahren in praxisnahen, ressourcenbeschränkten Anwendungen. Die in diesem Beitrag gezeigten Ergebnisse liefern ein grundlegendes Verständnis der Unsicherheitseinflüsse in unterabgetasteten Systemen und bilden eine wesentliche Grundlage für die Entwicklung robuster Kompensationsverfahren.

Keywords: Laser-Doppler-Vibrometrie, Bandpass-Unterabtastung, Abtasttakt drift, Nyquist-Zonen, Frequenzrekonstruktionsfehler, unbemannte Flugsysteme (UAS)

Einleitung

Die präzise Erfassung mechanischer Schwingungen ist eine zentrale Voraussetzung für die Zustandsüberwachung von Bauwerken im Rahmen des Structural Health Monitoring (SHM). Die Laser-Doppler-Vibrometrie (LDV) hat sich dabei als hochgenaue, berührungslose Messmethode etabliert, da sie Schwingungsgeschwindigkeiten mit hoher Orts- und Frequenzauflösung erfassen kann und gegenüber kontaktbehafteten Sensoren keine Massebelastung verursacht [1].

Mit dem zunehmenden Einsatz unbemannter Flugsysteme (Unmanned Aerial Systems, UAS) für die automatisierte Inspektion von Infrastrukturen wächst der Bedarf an kompakten und energieeffizienten LDV-Systemen [2]. Die Integration auf mobilen Plattformen ist jedoch mit erheblichen Herausforderungen verbunden. Insbesondere erfordert die Verarbeitung der hochfrequenten LDV-Signale Abtastraten im zweistelligen

Megahertz-Bereich, um die notwendige zeitliche Auflösung sicherzustellen [3]. Dies führt zu großen Datenmengen, hohem Rechenaufwand und signifikantem Energiebedarf, was die Einsatzdauer und Echtzeitfähigkeit auf ressourcenbeschränkten UAS-Plattformen limitiert [4], [5].

Ein vielversprechender Ansatz zur Reduktion der Datenrate und Systemkomplexität ist die Bandpass-Unterabtastung (Bandpass Subsampling). Dabei wird das hochfrequente Signal gezielt in eine niedrigere Nyquist-Zone abgebildet, sodass eine deutlich reduzierte Abtastfrequenz ausreicht, ohne Informationsverlust zu verursachen, sofern die Bandpass-Bedingungen des Nyquist-Shannon-Theorems eingehalten werden [4]. Zusätzlich kann auf eine analoge Mischstufe verzichtet werden, was die Hardwarekomplexität und den Energiebedarf weiter reduziert.

Die Leistungsfähigkeit dieses Ansatzes ist jedoch maßgeblich von der Stabilität des Abtasttakts abhängig. Während in stationären

Messsystemen hochstabile Oszillatoren (z. B. OCXOs) eingesetzt werden können, sind diese für UAS-Anwendungen aufgrund von Größe, Kosten und Leistungsaufnahme häufig ungeeignet. Abweichungen der Abtastfrequenz führen in unterabgetasteten Systemen zu systematischen Abweichungen in der Frequenzrekonstruktion. Dieser Effekt unterscheidet sich grundlegend von konventioneller Abtastung und verstärkt sich mit steigender Ordnung der genutzten Nyquist-Zone. Eine Nyquist-Zone umfasst den Frequenzbereich zwischen zwei aufeinanderfolgenden Vielfachen von $\frac{f_s}{2}$ und besitzt somit die Breite $\frac{f_s}{2}$. Die Zonenordnung wird mit N_Z bezeichnet.

In diesem Beitrag wird der Einfluss von einer Abtasttakt drift auf unterabgetastete LDV-Signale systematisch untersucht. Es wird analytisch hergeleitet, dass die resultierende Frequenzabweichung linear mit der Nyquist-Zonennummer zunimmt. Dieses Verhalten wird durch Simulationen sowie experimentelle Messungen verifiziert.

Die vorgestellten Ergebnisse liefern ein fundiertes Verständnis der Grenzen von Bandpass-Unterabastungsverfahren in mobilen LDV-Systemen und bilden die Grundlage für die Entwicklung robuster Kompensationsstrategien in energieeffizienten UAS-Anwendungen.

Laser-Doppler-Vibrometrie (LDV) und Signalcharakteristik

In einem typischen Heterodyn-LDV-System wird ein Laserstrahl in einen Referenz- und einen Messstrahl aufgeteilt. Der Messstrahl wird auf die schwingende Oberfläche gerichtet, wobei das reflektierte Licht aufgrund des Doppler-Effekts eine frequenzabhängige Verschiebung erfährt, die proportional zur Oberflächengeschwindigkeit ist [6].

Zur Erzeugung eines heterodynen Schwebungssignals wird die Frequenz des Referenzstrahls mithilfe einer Bragg-Zelle (akusto-optischer Modulator, AOM) um eine feste Frequenz f_c verschoben. Diese Trägerfrequenz liegt typischerweise im Bereich von 40 MHz bis 80 MHz.

Die Überlagerung von Referenz- und Messstrahl am symmetrischen Photodetektor führt zu dem elektrischen Ausgangssignal [7]

$$i_{BD} = K_{PD} \sqrt{P_r P_m} \varepsilon \cos(2\pi f_c t + \varphi(t)), \quad (1)$$

wobei K_{PD} die Gesamtverstärkung aus Photodetektor und nachfolgender Transimpedanzverstärkung beschreibt. P_r und P_m bezeichnen die optischen Leistungen des Referenz- bzw. Messstrahls, ε den Interferenzkontrast und $\varphi(t)$ die durch die Bewegung des Messobjekts hervorgerufene Dopplerphase.

Die Gesamtphase des Signals ergibt sich zu

$$\theta(t) = 2\pi f_c t + \varphi(t) \quad (2)$$

Durch zeitliche Ableitung der Gesamtphase erhält man die Momentanfrequenz des LDV-Signals

$$f_{mom}(t) = \frac{1}{2\pi} \dot{\theta}(t) = f_c + f_D(t), \quad (3)$$

wobei f_D die Dopplerfrequenz bezeichnet. Diese ist gegeben durch

$$f_D(t) = \frac{2 \cdot v(t)}{\lambda} = \frac{1}{2\pi} \dot{\phi}(t), \quad (4)$$

wobei λ die Laserwellenlänge und $v(t)$ die Oberflächengeschwindigkeit des Messobjekts beschreibt.

Die Momentanphase $\varphi(t)$ steht dabei in direktem Zusammenhang zur gesuchten Auslenkung $s(t)$ über

$$\phi(t) = \frac{4\pi}{\lambda} s(t). \quad (5)$$

Durch die Braggzelle wird im Heterodyn-Verfahren das gesuchte Nutzsignal somit von der Nullfrequenz gezielt in einen hohen Frequenzbereich auf die Zwischenfrequenz f_c geschoben. Der Einsatz der Braggzelle hat dabei zwei wichtige Vorteile.

Erstens erlaubt dieses Vorgehen nun auch die Detektion der Richtung der Bewegung, was im homodynen Fall ($f_c = 0$ Hz) nicht möglich wäre. Zweitens wird der störende Einfluss von niederfrequentem Rauschen und Gleichanteilen reduziert, welche die Messgenauigkeit andernfalls beeinträchtigen würden [8].

Bei konventioneller Direktabtastung ergibt sich die erforderliche Abtastfrequenz zu

$$f_s \geq 2 \cdot (f_c + f_D^{max}). \quad (6)$$

Bei einer typischen Trägerfrequenz von 40 MHz und einer Dopplerbandbreite im Megahertz-Bereich ergeben sich erforderliche Abtastraten im zweistelligen Megasample-Bereich. Dies

führt zu hohen Datenraten sowie einem erhöhten Rechenaufwand für die digitale Signalverarbeitung.

Die hohen Abtastraten stellen eine wesentliche Herausforderung für den Einsatz von LDV-Systemen auf ressourcenbeschränkten Plattformen wie unbemannten Flugsystemen (UAS) dar. Der Energiebedarf von Analog-Digital-Wandlern steigt mit zunehmender Abtastrate, während gleichzeitig hohe Datenraten die Anforderungen an Signalverarbeitung, Speicherung und echtzeitfähige Übertragung signifikant erhöhen. Daraus ergibt sich ein grundlegender Zielkonflikt zwischen der hohen Trägerfrequenz des LDV-Signals und den begrenzten Ressourcen hinsichtlich Energie, Rechenleistung und Gewicht auf mobilen Plattformen.

Obwohl die eigentliche Informationsbandbreite des Signals vergleichsweise gering ist, erzwingt die hohe Trägerfrequenz bei konventioneller Abtastung unverhältnismäßig große Abtastraten. Beispielsweise ergibt sich für eine maximale Dopplergeschwindigkeit von $v_D^{max} = 40 \frac{\text{mm}}{\text{s}}$ und eine Schwingungsfrequenz von $f_{vib} = 100 \text{ Hz}$ gemäß DIN 4150-3 bei Verwendung eines grünen Lasers mit $\lambda = 532 \text{ nm}$ eine maximale Dopplerfrequenz von:

$$f_D^{max} = \frac{2 \cdot v_D^{max}}{\lambda} \approx 150,94 \text{ kHz} \quad (7)$$

und unter Anwendung der Carson-Regel eine effektive Signalbandbreite [7]:

$$B = 2 \cdot (f_{vib} + f_D^{max}) \approx 501,98 \text{ kHz} \quad (8)$$

Bandpass-Unterabtastung

Unterabtastung (*Subsampling*), auch Bandpass-Unterabtastung oder harmonische Abtastung genannt, ist eine fortgeschrittene Technik, die speziell für die Digitalisierung von bandbegrenzten Hochfrequenzsignalen eingesetzt wird. Im Gegensatz zur Nyquist-Abtastung, die eine Abtastfrequenz von $f_s \geq 2 \cdot f_{max}$ (bezogen auf die höchste Signalfrequenz) erfordert, wird bei der Unterabtastung ein Signal mit einer Abtastfrequenz digitalisiert, die kleiner als $2 \cdot f_{max}$ [9]. Hierbei wird das Phänomen des Aliasing gezielt genutzt, um das hochfrequente Bandpasssignal in eine tiefere Nyquist-Zone zu überführen, von wo es ohne Informationsverlust rekonstruiert werden kann.

Das Eingangssignal muss zunächst durch einen analogen Bandpassfilter (BPF) begrenzt werden, siehe Abb. 1. Dieser Filter stellt sicher, dass das Signal vollständig innerhalb einer einzigen Nyquist-Zone liegt und unerwünschte Aliasing-Effekte von außerhalb des gewünschten Bandes minimiert werden, siehe Abb. 2.

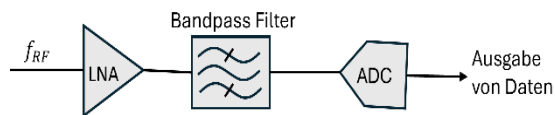
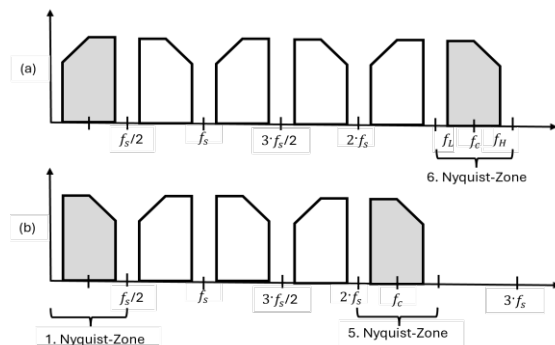


Abb. 1: Prinzipielle Struktur eines unterabtastenden Empfangssystems mit Bandpassfilter (BPF), Verstärkung (LNA) und Digitalisierung (ADC). Der BPF stellt sicher, dass das Signal innerhalb einer Nyquist-Zone liegt.



Frequenzlage abgebildet werden, sodass eine deutlich reduzierte Abtastrate ausreichend ist. Zusätzlich ist zu beachten, dass die Lage des abgebildeten Spektrums von der Parität der Nyquist-Zone abhängt. Während in geraden Nyquist-Zonen (siehe Abb. 2(a)) Spektralspiegelung auftritt, wird das Signal in ungeraden Nyquist-Zonen (siehe Abb. 2(b)) nicht invertiert abgebildet. Dies führt dazu, dass die Richtung der Frequenzverschiebung zwischen aufeinanderfolgenden Zonen wechseln kann. Der beschriebene Effekt ist deterministisch und kann bei bekannter Nyquist-Zone mathematisch durch eine einfache Invertierung eindeutig korrigiert werden, sodass er die grundsätzliche lineare Abhängigkeit des Fehlers nicht beeinflusst.

Damit keine Überlappung der Spektralrepliken auftritt, muss die Abtastfrequenz so gewählt werden, dass das gesamte Signalband eindeutig in eine Nyquist-Zone abgebildet wird. Für ein Signal mit unterer und oberer Grenzfrequenz f_L und f_H ergeben sich die zulässigen Abtastfrequenzen zu [9]

$$\frac{2 \cdot f_H}{N_Z} \leq f_s \leq \frac{2 \cdot f_L}{N_Z - 1} \quad (11)$$

und die Nyquist-Zone muss die Bedingung erfüllen

$$2 \leq N_Z \leq \left\lfloor \frac{f_H}{B} \right\rfloor. \quad (12)$$

Diese Bedingung stellt sicher, dass die durch Aliasing entstehenden Spektralanteile nicht miteinander interferieren und die ursprüngliche Information vollständig erhalten bleibt.

N_Z	f_c (MHz)	$ \Delta f_a = f_{a_{ist}} - f_{a_{soll}} $ (kHz)	$\Delta f_{a,N_Z} - \Delta f_{a,N_Z-2}$ (kHz)
1	0.275	$ 274.850 - 275 = 0.150$	-
2	0.825	$ 286.891 - 275 = 11.891$	11.741
4	1.925	$ 299.078 - 275 = 24.078$	12.187
6	3.025	$ 310.935 - 275 = 35.935$	11.857
8	4.125	$ 323.123 - 275 = 48.123$	12.188
10	5.225	$ 334.98 - 275 = 59.98$	11.857
12	6.325	$ 346.838 - 275 = 71.838$	11.858
$\Delta f_s = f_s - f'_s = 11.989 \text{ kHz}$			
Sollträger $f_a = 275 \text{ kHz}$			

Um sicherzustellen, dass die Trägerfrequenz im Zentrum einer Nyquist-Zone liegt, gilt nach Gleichung [10]

$$f_s = \frac{4 \cdot f_c}{2 \cdot N_Z - 1}. \quad (13)$$

Diese Wahl ist besonders vorteilhaft, da sie den maximalen Abstand zu den Zonengrenzen gewährleistet und dadurch die Empfindlichkeit gegenüber Überlappung reduziert.

Einfluss von Abtasttakt drift

In praktischen Abtastsystemen ist die Abtastfrequenz f_s nicht exakt konstant, sondern unterliegt Abweichungen infolge von Taktinstabilitäten des verwendeten Oszillators. Die tatsächliche Abtastfrequenz kann daher als

$$f'_s = f_s + \Delta f_s \quad (14)$$

ausgedrückt werden.

Wird ein sinusförmiges Hochfrequenzsignal $x(t) = A \cos(2\pi f_c t)$ mit einer Samplingfrequenz $f_s < 2 \cdot f_c$ und bei der geplanten gezielten Positionierung in der Mitte der jeweiligen Nyquist-Zone, d.h. gemäß Gl. (13), abgetastet, ergibt sich für die ideale Abbildungsfrequenz nach Unterabtastung

$$f_a = |f_c - n \cdot f_s|. \quad (15)$$

Ist nun die reale Abtastfrequenz durch Taktinstabilität verändert und beträgt stattdessen f'_s , so verschiebt sich das Aliassignal im Spektrum zu

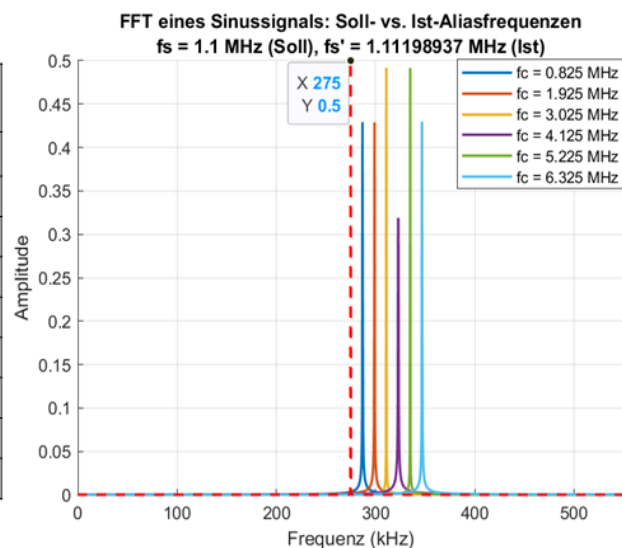


Abb. 3: Simulation der Frequenzverschiebung bei Bandpass-Unterabtastung infolge von Abtasttakt drift ($f_s = 1,1 \text{ MHz}$, $f'_s = 1,111989 \text{ MHz}$). Mit steigender Nyquist-Zonenordnung N_Z verschiebt sich die Aliasfrequenz systematisch, während die Signalfrequenz im Spektrum erhalten bleibt. Die Frequenzabweichung nimmt dabei linear mit N_Z zu ($\Delta f_a \propto N_Z$).

$$f'_a = |f_c - n(f_s + \Delta f_s)|. \quad (16)$$

Die Abweichung Δf_a zwischen idealer und realer Spektralposition ergibt sich dann zu

$$\Delta f_a = |f_c - n(f_s + \Delta f_s) - (f_c - n \cdot f_s)| = |-n \Delta f_s|. \quad (17)$$

Diese exakte Entwicklung zeigt: Die Verschiebung im Spektrum ist linear proportional zur Abtastfrequenz-Abweichung Δf_s und zum Index n der Aliasing-Zone. Unter Verwendung der in Gl. (13) definierten Wahl der Abtastfrequenz ergibt sich für zwei Nyquist-Zonen N_Z und $N_Z + 2$

$$n = \left\lfloor \frac{f_c}{f_s} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{2N_Z - 1}{4} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{2N_Z + 1}{4} \right\rfloor$$

$$n' = \left\lfloor \frac{2(N_Z + 2) - 1}{4} + \frac{1}{2} \right\rfloor = \left\lfloor \frac{2N_Z + 5}{4} \right\rfloor.$$

Dann gilt

$$\Delta f'_a - \Delta f_a = |n' \cdot \Delta f_s| - |n \cdot \Delta f_s| = |n' - n| \cdot \Delta f_s \quad (18)$$

$$\Delta f'_a - \Delta f_a = \left| \left\lfloor \frac{2N_Z + 5}{4} \right\rfloor - \left\lfloor \frac{2N_Z + 1}{4} \right\rfloor \right| \cdot \Delta f_s$$

$$\Delta f'_a - \Delta f_a = \Delta f_s \quad (19)$$

mit $\lfloor x + 1 \rfloor - \lfloor x \rfloor = 1$ für alle $x \in \mathbb{R}$.

Damit zeigt sich mit der Gl. (19): Die Frequenzverschiebung nimmt pro zwei Nyquist-Zonen um genau Δf_s zu.

Ergebnisse: Simulation und Experiment

Um die theoretisch hergeleitete Abhängigkeit des Frequenzfehlers von der Nyquist-Zone zu validieren, wurden sowohl numerische Simulationen als auch experimentelle Messungen durchgeführt. Die Untersuchungen konzentrierten sich auf den Vergleich der Soll-Trägerfrequenz mit der durch Instabilitäten verschobenen Ist-Trägerfrequenz im rekonstruierten Spektrum.

Für die Simulation wurde ein sinusförmiges Signal mit variabler Trägerfrequenz f_c betrachtet, das einer Bandpass-Unterabtastung unterzogen wird. Die nominale Abtastfrequenz wurde dabei auf $f_s = 1,1$ MHz festgelegt, während die tatsächliche Abtastfrequenz $f'_s = 1,111989$ MHz betrug, was einer Differenz von $\Delta f_s = 11,989$ kHz entspricht. Die Trägerfrequenz wurde gemäß Gl. (13) gezielt so gewählt, dass sie sich für unterschiedliche Werte von N_Z im Zentrum der jeweiligen Nyquist-Zonen befindet. Dadurch wird sichergestellt, dass die beobachteten Effekte ausschließlich auf die Abtastabweichung zurückzuführen sind.

Die Simulationsergebnisse sind in der rechten Darstellung von Abb. 3 dargestellt. Es ist deutlich erkennbar, dass sich die spektrale Lage des Signals mit steigender Nyquist-Zonenordnung systematisch verschiebt. Während bei niedrigen Nyquist-Zonen nur geringe Abweichungen auftreten, wächst die Frequenzverschiebung bei höheren Zonen deutlich an. Die Verschiebung erfolgt dabei entlang der Frequenzachse, während die Signalförmigkeit im Spektrum erhalten bleibt.

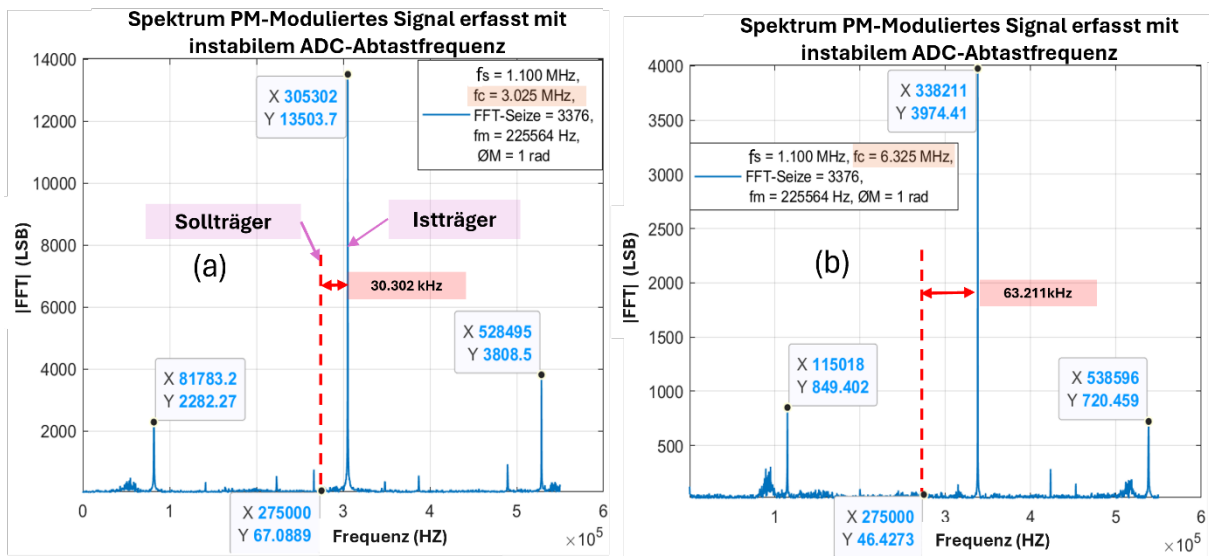


Abb. 4: Frequenzabweichung des Trägersignals bei steigenden Nyquist-Zonen: (a) $f_c = 3.025$ MHz ($6. N_Z$), Abweichung $\Delta f_a = 30.302$ kHz; (b) $f_c = 6.325$ MHz ($12. N_Z$), Abweichung $\Delta f_a = 63.211$ kHz. Die Richtung der Verschiebung bleibt gleich, die Abweichung nimmt jedoch mit wachsendem N_Z deutlich zu.

Ein quantitativer Vergleich ist in der Tabelle links in Abb. 3 dargestellt. Ausgehend von einem Sollträger bei $f_a = 275$ kHz ergibt sich für $N_Z = 6$ eine Frequenzabweichung von 35,935 kHz, während diese für $N_Z = 12$ auf 71,838 kHz ansteigt. Gleichzeitig zeigt die Differenz $\Delta f_{N_Z} - \Delta f_{N_Z-2}$ zwischen zwei aufeinanderfolgenden Nyquist-Zonen eine nahezu konstante Zunahme von etwa 11,8 kHz. Es ist zu beobachten, dass dieser Wert geringfügig von der theoretisch erwarteten Abweichung der Abtastfrequenz Δf_s differiert. Diese Diskrepanz lässt sich primär auf die endliche Frequenzauflösung der diskreten Fourier-Transformation (DFT) zurückführen. Bei einer FFT-Länge von 3376 Punkten und einer Abtastrate von 1,1 MHz beträgt die Bin-Breite ca. 325,8 Hz.

Ergänzend zur Simulation wurden experimentelle Messungen durchgeführt, um die beobachteten Effekte unter realen Bedingungen zu verifizieren. Hierzu wurde ein phasenmoduliertes Signal mithilfe eines Signalgenerators (Rohde & Schwarz SML03) erzeugt und über einen 16-Bit-Analog-Digital-Wandler des Arduino Portenta H7 digitalisiert. Die Abtastfrequenz des Systems weist dabei, bedingt durch die interne Taktquelle, eine Abweichung von der nominalen Einstellung auf. Weitere Details zu den Versuchsparametern sind in Abb. 4 dargestellt.

Die experimentellen Ergebnisse bestätigen die Simulation. In Abb. 4(a) ist der Fall für $N_Z = 6$ dargestellt, bei dem sich eine Frequenzabweichung zwischen Soll- und Istträger von etwa 30,3 kHz ergibt. Für den Fall $N_Z = 12$, dargestellt in Abb. 4(b), steigt diese Abweichung auf etwa 63,2 kHz an. Damit zeigt sich eine deutliche Zunahme der Abweichung mit steigender Nyquist-Zonenordnung. Weiterhin ist zu erkennen, dass die Differenz der Frequenzabweichung zwischen zwei aufeinanderfolgenden Nyquist-Zonen näherungsweise konstant ist. Dies bestätigt den analytisch hergeleiteten linearen Zusammenhang:

$$\Delta f_a \propto N_Z$$

Die beobachteten leichten Abweichungen von einer ideal konstanten Differenz lassen sich im experimentellen Aufbau durch zusätzliche Effekte erklären. Im Gegensatz zur Simulation unterliegt die reale Abtastfrequenz zeitlichen Schwankungen (Taktjitter), wodurch die effektive Frequenzabweichung nicht strikt konstant ist. Darüber hinaus beeinflussen

Rauschen sowie die endliche FFT-Auflösung die Bestimmung der Spektralpeaks.

Die Richtung der Frequenzverschiebung bleibt dabei über alle betrachteten Nyquist-Zonen hinweg konstant, während sich der Betrag des Fehlers weiterhin näherungsweise linear vergrößert. Die Ergebnisse zeigen eine sehr gute Übereinstimmung zwischen Theorie, Simulation und Experiment.

Zusammenfassung und Ausblick

Die vorliegenden Untersuchungen zeigen, dass die Bandpass-Unterabtastung zwar eine effektive Reduktion der Abtastrate und damit des Energiebedarfs ermöglicht, jedoch gleichzeitig eine inhärente Sensitivität gegenüber Abtasttakt drift aufweist. Es wurde analytisch hergeleitet und sowohl durch Simulation als auch experimentell bestätigt, dass der resultierende Frequenzfehler proportional zur Nyquist-Zonenordnung anwächst. Damit führt ein zunehmender Unterabtastungsfaktor direkt zu einer Verstärkung von Abweichungen der Frequenzlage im Spektrum.

Die Ergebnisse verdeutlichen, dass bereits kleine Taktabweichungen bei höheren Nyquist-Zonen zu signifikanten Abweichungen in der Frequenzrekonstruktion führen können. Dieser Effekt ist insbesondere für LDV-Anwendungen kritisch, da die physikalische Messgröße unmittelbar aus der Frequenzinformation abgeleitet wird. In ressourcenbeschränkten Systemen, wie sie beispielsweise auf unbemannten Flugsystemen eingesetzt werden, stellt dies eine fundamentale Einschränkung dar, da dort häufig keine hochstabilen Taktquellen verfügbar sind.

Insgesamt ergibt sich somit ein grundlegender Zielkonflikt zwischen der Reduktion der Abtastrate durch Unterabtastung und der erreichbaren Messgenauigkeit. Die Wahl der Nyquist-Zone stellt dabei einen zentralen Entwurfsparameter dar, der sowohl die Energieeffizienz als auch die Robustheit gegenüber Abtasttakt drift maßgeblich beeinflusst.

Zukünftige Arbeiten werden sich daher auf die Entwicklung und Evaluierung geeigneter Kompensationsverfahren konzentrieren, mit dem Ziel, die durch Abtasttakt drift verursachten Abweichungen zu reduzieren. Insbesondere erscheint die Nutzung zusätzlicher Referenzsignale sowie modellbasierter Korrekturansätze vielversprechend, um die Vorteile der Bandpass-Unterabtastung auch in

driftbehafteten, mobilen Systemen zuverlässig nutzbar zu machen.

Literaturnachweis

- [1] C. Providakis, M. Moustarakis, und G. Providaki, „Operational Modal Analysis of Historical Buildings and Finite Element Model Updating Using α Laser Scanning Vibrometer“, *Infrastructures*, Bd. 8, S. 37, Feb. 2023, doi: 10.3390/infrastructures8020037.
- [2] C. Rembe, B. J. Halkon, und M. A. A. Ismail, „Measuring Vibrations in Large Structures with Laser-Doppler Vibrometry and Unmanned Aerial Systems: A Review and Outlook“, *Advanced Devices & Instrumentation*, Bd. 6, S. 0103, Sep. 2025, doi: 10.34133/adi.0103.
- [3] H. Volkers, „On the calibration of laser Doppler vibrometers using simulated excitation“, *Measurement: Sensors*, Bd. 38, S. 101748, Mai 2025, doi: 10.1016/j.measen.2024.101748.
- [4] C.-H. Tseng, „Bandpass sampling criteria for nonlinear systems“, *IEEE Transactions on Signal Processing*, Bd. 50, Nr. 3, S. 568–577, März 2002, doi: 10.1109/78.984739.
- [5] H. Wang u. a., „An Elegant Solution: An Alternative Ultra-Wideband Transceiver Based on Stepped-Frequency Continuous-Wave Operation and Compressive Sensing“, *IEEE Microwave Magazine*, Bd. 17, Nr. 7, S. 53–63, Juli 2016, doi: 10.1109/MMM.2016.2549138.
- [6] C. Rembe und A. Dräbenstedt, „D1.1 - Speckle-Insensitive Laser-Doppler Vibrometry with Adaptive Optics and Signal Diversity“, in *Proceedings SENSOR 2015*, Nürnberg, Germany: AMA Service GmbH, Von-Münchhausen-Str. 49, 31515 Wunstorf, Germany, 2015, S. 505–510. doi: 10.5162/sensor2015/D1.1.
- [7] C. Rembe u. a., „Laser Doppler vibrometry“, *Nat Rev Methods Primers*, Bd. 6, Nr. 1, S. 30, Apr. 2026, doi: 10.1038/s43586-026-00487-2.
- [8] H. Tabatabai, D. E. Oliver, J. W. Rohrbaugh, und C. Papadopoulos, „Novel Applications of Laser Doppler Vibration Measurements to Medical Imaging“, *Sens Imaging*, Bd. 14, Nr. 1, S. 13–28, 2013, doi: 10.1007/s11220-013-0077-1.
- [9] R. G. Vaughan, N. L. Scott, und D. R. White, „The theory of bandpass sampling“, *IEEE Transactions on Signal Processing*, Bd. 39, Nr. 9, S. 1973–1984, Sep. 1991, doi: 10.1109/78.134430.
- [10] G. Giovannetti u. a., „Application of undersampling technique for the design of an NMR signals digital receiver“, *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, Bd. 29B, Nr. 3, S. 107–114, 2006, doi: 10.1002/cmr.b.20065.

UAV-basierte Nahfeld-zu-Fernfeld-Transformation zur detaillierten Charakterisierung großer Emissionsquellen unter Betriebsbedingungen

Jonathan Gundlach, Ramesh Raja Subramanyam, David Ulm, Criostoir Gerasch, Christian van Aswegen, Kai Baaske, Stefan Jacob, Thomas Kleine-Ostmann
Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig
+49 531-592-2215, jonathan.gundlach@ptb.de

Zusammenfassung

Die Beschreibung der Wellenausbreitung im Raum und insbesondere die Charakterisierung von Emissionsquellen ist von großem Interesse in verschiedenen physikalischen Disziplinen. Große Emissionsquellen, akustischer oder elektromagnetischer Natur, können oft nicht in herkömmlichen Laboren untersucht werden. Wir entwickeln eine innovative Messtechnik, die es ermöglicht, die abgestrahlte Leistung und die Richtcharakteristiken großer Emissionsquellen während ihres Betriebs ohne Modifikation zu charakterisieren. Diese Methode basiert auf der Verwendung von Uncrewed Aerial Vehicles (UAV), die mit modernster Messtechnik ausgestattet sind und sich frei im Nahbereich der Emissionsquellen bewegen können. Anschließend wenden wir eine spezielle Transformationsrechnung an, um aus den diskreten Messdaten auf einer Hüllfläche um die Emissionsquelle, deren Abstrahlungseigenschaften zu extrahieren. Die Prinzipien der wellenförmigen Ausbreitung unterliegen in verschiedenen Disziplinen ähnlichen mathematischen Gesetzmäßigkeiten. Daher ist unsere Methode leicht von elektromagnetischen auf akustische Felder übertragbar. In unserem Projekt demonstrieren wir diese Interdisziplinarität und erwarten Ergebnisse, die in vielfacher Hinsicht von großem Interesse sind. Basierend auf diesen Untersuchungen entwickeln wir Messplattformen unterschiedlicher Komplexität, speziell in Bezug auf Position, Zeit und Messgrößen in Abwägung zu den zu erwartenden Genauigkeiten. Darüber hinaus diskutieren wir potenzielle Schwächen und die Skalierbarkeit der Methode hinsichtlich der Anzahl der eingesetzten UAV und der verwendeten Messtechnik.

Keywords: UAV, Infraschall, elektromagnetische Felder, Antennenmesstechnik, Nahfeld- Fernfeld Transformation

Einleitung

UAVs sind eine aufstrebende Technologie, die besonders rasche Fortschritte in wichtigen Bereichen wie der Quantenkommunikation [1], der Landwirtschaft [2], dem Transportwesen [3], und dem Katastrophenmanagement ermöglicht [4]. Obwohl diese Anwendungen oft in gewissem Maße Überwachung und Sensorik beinhalten, steckt die Verwendung von UAVs als Instrumente für rein messtechnische Zwecke noch in den Kinderschuhen. Präzise Messungen liegen jenseits der Möglichkeiten der meisten kommerziell erhältlichen UAV-Lösungen. Dies liegt vor allem an den technischen Beschränkungen, die luftgestützten Messplattformen im freien Raum auferlegt sind, wie z. B. begrenzte Energieversorgung und Messdauer, kleine zulässige Nutzlasten und ungenaue Positionierung. Die Umgebungsbedingungen in

der Luft sind im Allgemeinen rau, mit stark schwankenden Umgebungsbedingungen wie Temperatur, Feuchtigkeit, Wind und Luftdruck. Außerdem erzeugt der Antrieb von Propellern unweigerlich Störungen wie Luftströmungsturbulenzen, Vibrationen, Lärm und elektromagnetische (EM) Störungen.

Eine potenziell wichtige messtechnische Anwendung für UAV ist die Kartierung physikalischer Felder mit hoher Genauigkeit in unzugänglichen oder abgelegenen Gebieten. Die Kartierung physikalischer Felder ist in vielen wissenschaftlichen Disziplinen wichtig, z. B. elektromagnetische Felder von Antennen, Schallfelder, Windgeschwindigkeit und Temperaturverteilungen. Obwohl Forscher bereits Hubschrauber [5] und Flugzeuge [6] mit Besatzung für die Feldkartierung eingesetzt haben, kann die Datenqualität aufgrund der noch größeren Störungen und Rückwirkungen, die von den bemannten Flugzeugen

ausgehen, beeinträchtigt werden. Außerdem ist der Betrieb von bemannten Flugzeugen teuer. Jüngste Experimente mit UAVs zur Messung des elektrischen Feldes [7], zur Messung der Luftqualität [8] und zur Lokalisierung akustischer Quellen [9] unterstreichen das große Potenzial der UAV-gestützten Messtechnik zur Erfassung von Daten, die mit anderen Methoden nicht gewonnen werden können.

Ein interessanter Anwendungsfall ist die Vorhersage von Fernfeldern aus Nahfeldmessungen stationärer harmonischer Wellenfelder unter Verwendung von Oberflächenintegralformulierungen (surface integral formulations – SIFs). SIFs werden in der Physik und im Ingenieurwesen mit Anwendungen in der Elektromagnetik [10, 11], Optik [12,13], und Akustik [14,15] verwendet. Mit ihnen lassen sich Fern- und Nahfelder für harmonische Wellenphänomene genau vorhersagen, indem nur die Feldvariablen entlang einer geschlossenen Fläche um die Quellen verwendet werden. Aufgrund dieser Eigenschaft werden SIF in effizienten numerischen Methoden wie der Boundary-Element-Methode [16] und der Momentenmethode genutzt [17].

Wir nutzen ein Verfahren, das zur Messung von Wellenfeldern über große Flächen verwendet werden kann, um die Fernfeld-Wellenabstrahlung mit SIF vorherzusagen. Die Methode wird durch die jüngsten Fortschritte in der UAV-technologie in Bezug auf Flugzeit, Stabilität und Positionierung ermöglicht [18].

Die vorgeschlagene Methode baut auf der experimentellen Nahfeld-zu-Fernfeld-Transformation (NFFT) auf, die üblicherweise verwendet wird, um die Fernfeldabstrahlung elektromagnetischer Antennen für kleine und mittelgroße Antennen aus Nahfeldmessungen in Laborumgebungen zu extrahieren [19]. NFFT wird verwendet, wenn direkte Auswertungen von Nahfeldmessungen nicht empfehlenswert sind, weil sich das Wellenfeld noch nicht vollständig entwickelt hat. In solchen Fällen kann die Richtwirkung der Quelle die lokalen Messergebnisse erheblich beeinflussen. NFFT-basierte Methoden lösen dieses Problem durch die Anwendung eines zweistufigen Verfahrens. Zunächst wird das Wellenfeld an diskreten Punkten entlang einer Fläche gemessen, die die Antenne im Nahfeld umschließt. Diese Daten werden zur numerischen Anpassung einer äquivalenten Quellendarstellung auf der Oberfläche verwendet. Im zweiten Schritt wird diese äquivalente Quellenverteilung für eine SIF zur Berechnung der Fernfeldstrahlung verwendet.

Erst kürzlich wurde die UAV-gestützte NFFT eingeführt und erfolgreich an großen Antennen getestet [7, 20], was das große Potenzial des Verfahrens für eine genaue Fernfeldvorhersage belegt. Dabei wurde ein UAV mit Hochfrequenz-Glasfaser-Verbindungen zu einem bodengestützten Netzwerk-Vektoranalysator (VNA) verwendet, um die Amplitude und Phase der elektrischen Feldstärke im Nahfeld zu erfassen. Trotz vielversprechender Ergebnisse schränken die optischen Fasern die Auswahl der Flugbahnen ein. Außerdem lässt sich die Methode nicht ohne Weiteres auf mehrere UAV übertragen, da eine aufwändige Verfolgungs- und Analyseausrüstung erforderlich ist. Andere Ansätze verwenden phasenlose Messverfahren in Kombination mit einem Real-Time-Kinematic(RTK)-System, das den Bedarf an teuren und einschränkenden Komponenten reduziert, aber Messungen an mindestens zwei Oberflächen [21]. Im akustischen Bereich werden Verfahren eingesetzt, die ähnliche Methoden wie die elektromagnetische NFFT verwenden, um äquivalente Quellen und Fernfelder zu berechnen. Tatsächlich weisen die in diesen Methoden verwendeten Gleichungen viele Ähnlichkeiten mit ihren elektromagnetischen Gegenstücken auf, wie von anderen Gruppen gezeigt wurde [22-24]. Beispielsweise können äquivalente Quellen durch akustische Nahfeldholographie-Methoden [25] oder Multimode-Techniken bestimmt werden [26]. Diese Techniken beinhalten eine systematische Abtastung einer geschlossenen Oberfläche (oder einer "Holographie") um eine Schallquelle herum. Bei den meisten Anwendungen erfolgt die Abtastung der Oberflächen mit Mikrofonarrays oder Roboterarmen. Dies schränkt die Anwendbarkeit der derzeitigen Methoden auf kleine und mittelgroße Emissionsquellen ein. Sobald eine äquivalente Quellendarstellung bestimmt ist, kann das Fernfeld mit den Kirchhoff-Helmholtz-Integrals oder ähnlichen akustischen Analogien berechnet werden [27,28].

Infraschall und Niederfrequenter Schall

Eine Vielzahl industrieller Großgeräte und Prozesse, insbesondere in strategischen Sektoren wie der Energieversorgung (Windkraftanlagen, Biogasanlagen), dem Transportwesen (Schiffe, Züge), dem Katastrophenschutz (Erdbeben, Stürme) sowie der Verteidigung (Explosionen, Raketenstarts), erzeugt niederfrequenten Schall (<200 Hz) und Infraschall (<20 Hz). Diese Schallquellen können aufgrund ihrer Größe und der

erforderlichen Umgebungsbedingungen nicht im Labor vermessen werden.

Akustische Messdrohnen eignen sich besonders für die Erfassung niederfrequenten Schalls großer Quellen (und nicht für hochfrequenten Schall). Dies hat einen wesentlichen Grund: Der Eigenlärm der Drohne für Frequenzen bis etwa 100 Hz vernachlässigbar. Dies zeigen unsere Untersuchungen sowohl im Freifeld als auch im Labor [29, 30]. Beispielsweise erzeugen Windkraftanlagen tonale Komponenten bei der Blattdurchlaufrfrequenz von etwa 1 Hz sowie deren Harmonischen, ebenso wie bei der Drehfrequenz von Getriebe und Generator (20–50 Hz) [31].

In unserem Projekt demonstrieren wir, dass Drohnen tatsächlich für akustische Messungen niederfrequenten Schalls eingesetzt werden können. Hierzu haben wir zunächst den Eigenlärm der Drohne in einem Freifeldraum präzise vermessen (Abbildung 1) [30]. Es zeigte sich, dass die Drohne im niederfrequenten Bereich keinen signifikanten Eigenlärm erzeugt. Erst ab der ersten Blattdurchlaufrfrequenz tritt Eigenlärm auf, der akustische Messungen überlagern kann. Abhängig von Größe und Beladung der Drohne sind somit Messungen bis etwa 150–200 Hz möglich.

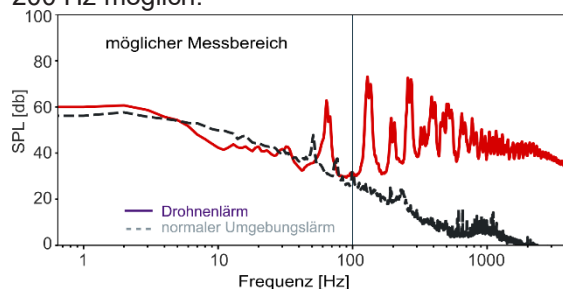


Abbildung 1: Akustischer Eigenlärm einer Drohne, verglichen mit typischen atmosphärischen Hintergrundlärm an einem ruhigen, windstillen Tag.

Auch wenn der Eigenlärm der Drohne im niederfrequenten Bereich vernachlässigbar ist, stellt der durch atmosphärischen Wind induzierte Lärm in der praktischen Anwendung ein wesentliches Problem dar. Experimente mit verschiedenen Windschutzkonzepten haben gezeigt, dass die Verwendung eines Mikrofonpaares in Kombination mit einer Kohärenzanalyse die effektivste Windunterdrückung ermöglicht [32].

Ein möglicher Versuchsaufbau ist in Abbildung 2 dargestellt. Der Abstand der Mikrofone betrug in diesem Aufbau etwa 2 m und ermöglichte damit eine effektive Windunterdrückung bis ungefähr 10 Hz. Mit diesem Aufbau konnten präzise Schalldruck-

und Schalleistungsmessungen aus der Luft durchgeführt werden. Hierzu wurde eine spezielle Schallquelle installiert, bei der während des Betriebs die abgegebene Schalleistung bei einer Frequenz von 10 Hz bestimmt werden kann [33]. Eine Abbildung des Versuchsaufbaus mit Messdrohne und Schallquelle ist in Abbildung 3 dargestellt. In diesem Versuch konnte die Schalleistung mit einer Unsicherheit von <1 dB bestimmt werden.



Abbildung 2: Versuchsaufbau einer Hexacopter-Drohne mit Mikrofonpaar.



Abbildung 3: Messung der Schalleistung einer kalibrierten Schallquelle.

Nach der Verifizierung der Messmethode wurden erste Messungen an realen Emittoren durchgeführt. Abbildung 4 zeigt eine Messdrohne auf Nabenhöhe einer Windkraftanlage. Dieses Experiment diente der Aufnahme von Richtcharakteristiken des Generatorlärms, der diskrete Töne bei etwa 20 Hz emittierte. Die Durchführung sowie die Ergebnisse des Tests sind detailliert in Ref. [30] beschrieben.



Abbildung 4: Messung des Lärms einer Windkraftanlage, hier gemessen aus der Luft auf Nabenhöhe.

Elektromagnetische Emitter

Die Charakterisierung von elektromagnetischen Emittlern bzw. Antennen im Labor wird bereits seit langem erfolgreich durchgeführt. Zu diesem Zweck werden Antennenscanner verwendet, welche die Übertragung zwischen der zu vermessenden Antenne (Antenna under Test - AUT) und der Messantenne bzw. Sonde messen. Diese Messung erfolgt wiederholt während eine der Antennen von einem Positioniersystem auf einer Kugel- oder Zylindermantelfläche oder einer Ebene relativ zu der anderen Antenne bewegt wird. Aus diesen Transmissionsmessungen und deren Position kann mittels Nahfeld-Fernfeld-Transformation (NFFT) nun die Richtcharakteristik der zu Untersuchenden Antenne rekonstruiert werden [34]. Unter Laborbedingungen können sowohl die Positions- und Lage-, als auch die Transmissionsmessungen zwischen den Antennen mit höchster Genauigkeit durchgeführt werden. Durch eine Computersteuerung finden die Messungen autonom über viele Stunden statt. Probleme durch Reflexionen und Mehrwegausbreitung der EM-Wellen zwischen den Antennen können durch EM-Absorber an allen relevanten Oberflächen um die Messstrecke herum auf ein Minimum reduziert werden. Die Messung unter Laborbedingungen ist jedoch nicht möglich, wenn die zu vermessende Antenne zu groß für den jeweiligen Antennenscanner wird oder die Antenne an ihrem Bestimmungsort charakterisiert werden soll, z.B. um Ihre Eigenschaften unter Einfluss aller realen Umgebungsbedingungen zu bestimmen. Zur Lösung dieses Problems bietet es sich an, moderne UAVs zu nutzen. In diesem Projekt soll das erprobte Prinzip des Antennenscanners aus dem Labor auf eine fliegende Plattform im Freien übertragen werden. Auch wenn das grundlegende Messprinzip bereits erprobt ist, ergeben sich diverse neue Herausforderungen, die im Folgenden näher diskutiert und

Lösungsansätze vorgestellt werden. Es wird angestrebt den Frequenzbereich von 100 MHz bis 3 GHz abzudecken.

Um eine Nahfeld-Fernfeld-Transformation durchführen zu können, muss an diversen, genau bekannten Positionen im Raum Betrag und Phase des EM-Feldes ermittelt werden. Bei einer Frequenz von 3 GHz ergibt sich eine Wellenlänge von ca. $\lambda=10$ cm. Für eine erfolgreiche NFFT ist eine Positionsgenauigkeit von $< \lambda/20$ anzustreben [35], was 0,5 cm entspricht. Gleichzeitig ergibt sich durch die Ausbreitung der EM-Welle mit der Lichtgeschwindigkeit c eine nötige Genauigkeit der Zeitsynchronisierung von 17 ps für die Synchronisation der Messung von ausgesendeter zu empfangener Welle. Verfahren zu entwickeln diese Genauigkeiten auf einer fliegenden Messplattform unter Einfluss von Umweltbedingungen zu erreichen ist einer der Schwerpunkte des Forschungsprojektes.

Die Transmissionsmessung zwischen den Antennen erfolgt mittels eines Vektor Netzwerk Analysators (VNA), welcher sowohl das Signal für die Sendeantenne mit bekanntem Betrag und bekannter Phase generiert als auch das Signal der Empfangsantenne in Betrag und Phase misst. Signallaufzeiten und Dämpfungen der Kabel können, sofern sie konstant sind, mittels einer Systemfehlerkorrektur bestimmt und korrigiert werden. Die Übertragung des Messsignals von der an dem UAV montierten Antenne zum bodengebunden VNA ist nicht mit einem klassischen Koaxialkabel möglich, da dieses zu schwer und unflexibel ist. Stattdessen wird das Signal mittels elektrooptischer Konverter in ein optisches Signal gewandelt, welches per Glasfaserkabel übertragen, und anschließend wieder in ein elektrisches Signal konvertiert wird. Die Kabelverbindung ist nach aktuellem Kenntnisstand nötig, um kohärente Messungen zu ermöglichen.

Positionsbestimmung

Kenntnis über die Position des UAV im Raum ist sowohl für die Steuerung und Flugführung als auch in Form der eigentlichen Messung bzw. zur Interpretation der Feldmessungen nötig. Da diese Anwendungen unterschiedliche Anforderungen haben, werden zwei Positionsmesssysteme eingesetzt. Diese sind ein Real Time Kinematik Global Navigation Satellite System (RTK-GNSS) und ein Lasertracker. Das UAV ist mit einem dualen RTK-fähigen GNSS ausgestattet welches zwei Antennen und Empfangskanäle hat. Der Aufbau des Messsystems ist in Abbildung 5 und die geometrischen Verhältnisse am UAV sind

in Abbildung 6 dargestellt. Am Boden befindet sich ein weiteres GNSS welches als lokale RTK-Basisstation dient. Diese bestimmt aus ihren gemessenen GNSS-Positionsdaten und ihrer bekannten realen Position Korrekturdaten, welche zum GNSS des UAVs geschickt werden. Dort kann mittels der Korrekturdaten die Position des UAV mit einer Unsicherheit von wenigen cm bestimmt werden. Aus der Positionsdivergenz der beiden Antennen des UAV kann außerdem die Orientierung um die Hochachse bestimmt werden, was ein Magnetometer zur Richtungsbestimmung ersetzt. Falls keine Korrekturdaten zur Verfügung gestellt werden, erhöht sich die Positionsunsicherheit auf, für normales GNSS übliche Werte von wenigen Metern [36], während die sonstige Funktion beibehalten wird. Das System ist robust gegenüber Umgebungseinflüssen und liefert Positionsdaten direkt an den Flugcontroller im UAV.

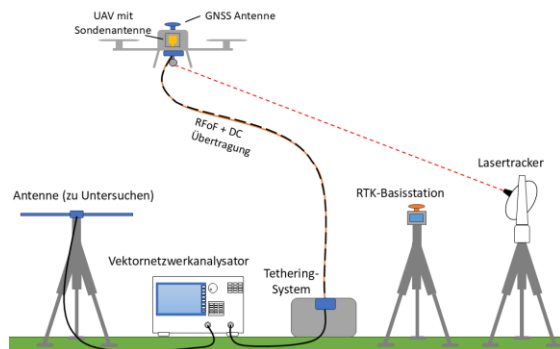


Abbildung 5: Messaufbau für Elektromagnetische Emittier.

Da die Genauigkeit des RTK-GNSS für die Auswertung der Messergebnisse bei EM-Messungen ab ca. 500 MHz nicht ausreicht, wird die Position des UAV für diese Messungen zusätzlich noch mit einem Lasertracker mit einer Unsicherheit von deutlich unter einem mm bestimmt. Dieser erfordert eine direkte Sichtverbindung zum, am UAV montierten Reflektor und hat in unseren Tests eine maximale Reichweite von bis zu 70 m. Außerdem hat der Retroreflektor nur einen typabhängigen, eingeschränkten Sichtbereich und die Messdatenerfassung des Lasertrackers erfolgt am Boden. Aufgrund dieser Einschränkungen muss das Messsystem Unterbrechungen der Positionsmessung per Lasertracker in bestimmten Manövern tolerieren können. In solchen Fällen wird, die per RTK-GNSS ermittelte Position genutzt, um den Lasertracker automatisch und schnell wieder zum Reflektor auszurichten.

Für die eigentliche Auswertung der Messungen ist die Position des Phasenzentrums der Messantenne (P_{SONDE}) und nicht die des Laserreflektors (P_{LT}) nötig. Somit muss noch eine Korrektur, der Positionsmessung mit einem Differenzvektor ($\vec{v}_{\text{LT, SONDE}}$) und einer Rotation dessen um die drei Lagewinkel des UAVs, dem Roll-, Nick- und Gierwinkel erfolgen, welche vom Flugcontroller ausgelesen werden können.

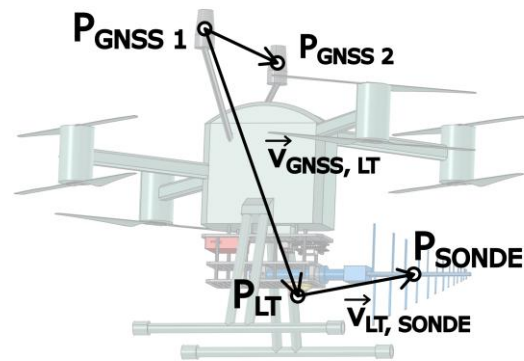


Abbildung 6: Geometrische Verhältnisse am UAV.

Zur Verifikation der beiden Positionsmesssysteme und zur Bestimmung des Offsetvektors zwischen dem Antennenreferenzpunkt (ARP) des GNSS ($P_{\text{GNSS 1}}$) und dem optischen Zentrum des Laserreflektors (P_{LT}) an dem UAV wurde eine Vergleichsmessung durchgeführt. Beide Positionsmessungen wurden sowohl am Boden als auch in der Luft sehr oft durchgeführt, wobei signifikante Rotationen um alle Achsen vorgekommen sind. Zur Auswertung stehen somit sehr viele Messpunkte zur Verfügung, welche beide Positionen zur gleichen Zeit bei diversen Orientierungen des UAV beschreiben. Anschließend wird mittels Optimierungsalgorithmus ein Offsetvektor ($\vec{v}_{\text{GNSS, LT}}$) im Koordinatensystem des UAV bestimmt, der beide Punkte verbindet und entsprechend der Lagewinkel des UAVs rotiert wird. Mit diesem Offsetvektor wird aus jeder Position der RTK-GNSS Messung eine Schätzung der Position des Laserreflektors bestimmt. Es verbleibt eine residuale Differenz zwischen zeitgleich per RTK-GNSS geschätzter und per Lasertracker gemessener Position des Laserreflektors. Mittels des Optimierungsalgorithmus wird der Offsetvektor so optimiert, dass über alle Messpunkte, d.h. auch über alle Lagewinkel eine möglichst kleine residuale Differenz zur jeweiligen Positionsmessung des Lasertrackers verbleibt. Konstante Differenzen, d.h. solche unabhängig von Fluglage und Messzeitpunkt werden durch

konstante Korrekturen der Lasertrackerposition im globalen Weltkoordinatensystem ausgeglichen. Das arithmetische Mittel aller residualen Differenzen ist somit Null.

Der ermittelte Offsetvektor beschreibt dann den realen Offset zwischen ARP und Laserreflektor. Die verbleibende residuale Differenz ist ein Indikator für die kombinierte Messunsicherheit von RTK-GNSS-, und Lasertrackerposition sowie den Lagewinkeln des UAV. Dieser Vergleich in Abbildung 7 zeigt, dass die residualen Differenzen der Positionsmessungen mehrheitlich deutlich kleiner als 5 cm sind.

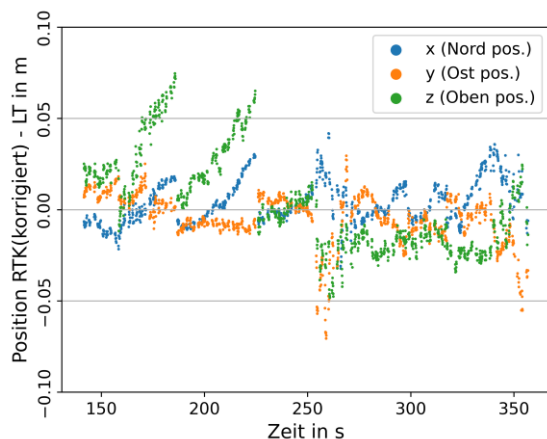


Abbildung 7: Vergleich von Lasertracker und RTK-GNSS Positionsmessung.

Die Standardabweichung liegt für die horizontalen x- und y-Richtungen im globalen Koordinatensystem bei 1,2 bzw. 1,4 cm und für die vertikale z-Position bei 2,6 cm. Diese Werte entsprechen der zu erwartenden Messunsicherheit für RTK-GNSS [36], sowohl in ihrer Größenordnung als auch der Tatsache, dass die Höhenmessung mittels GNSS ungenauer ist als die der lateralen Position.

Dieser Vergleich zeigt, dass RTK-GNSS- und LT-Positionsmessung sowie die Lagesensoren des Flugcontrollers entsprechend ihrer technischen Möglichkeiten funktionieren. Die residuale Differenz entsteht vermutlich größtenteils durch die Unsicherheit der RTK-GNSS Messung. Auch die Bestimmung des Offsetvektors zwischen GNSS-Antennenreferenzpunkt und Laserreflektor war erfolgreich und liefert zuverlässig plausible Ergebnisse. Mit einem ähnlichen Verfahren kann folglich in den Antennenmessungen aus der LT-Positionsmessung und dem Offsetvektor zur Sondenantenne ($\vec{r}_{LT,SONDE}$), die genaue Position des Phasenzentrums der Sondenantenne am im Raum bestimmt werden. In Kombination mit den Lagewinkeln des Flugcontrollers die, entsprechend dem

Versuch als ausreichend genau eingeschätzt werden können, stehen somit alle wichtigen Geometriedaten für die NFFT zur Verfügung.

Zusammenfassung

Erste Ergebnisse unseres Projekts zeigen, dass Drohnen zur Bestimmung von Wellenfeldern eingesetzt werden können. Diese Felder können sowohl elektromagnetischer als auch akustischer Natur sein. Die wichtigsten Erkenntnisse sind:

1. Aufgrund der schwierigen Messbedingungen aus der Luft müssen gezielte Maßnahmen zur Reduktion von Messfehlern ergriffen werden.
2. Ein entscheidender Faktor ist die Lagegenauigkeit. Mithilfe moderner RTK-Technik kann eine Positionsgenauigkeit im Bereich weniger Zentimeter erreicht werden.
3. Für die Messung akustischer Felder stellt atmosphärischer Wind die größte Störgröße dar, die durch den Einsatz von Mikrofonpaaren reduziert werden kann.
4. Erste Testflüge zeigen, dass Messungen im Nahfeld Rückschlüsse auf die Quellcharakteristik erlauben und somit eine Abschätzung des Fernfelds ermöglichen.

Wir erwarten, dass unsere Messmethoden eine breite Anwendung sowohl in der Wissenschaft als auch in der Industrie finden werden. Beispiele hierfür sind die Validierung atmosphärischer Ausbreitungsmodelle, die Untersuchung komplexer Quellen sowie die Planung und Umsetzung von Großprojekten wie Windparks und Antennenfeldern.

Danksagung

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) Projekt – G:(GEPRIS) 541019206.

Literaturnachweise

- [1] Kumar, A., Augusto de Jesus Pacheco, D., Kaushik, K. & Rodrigues, J. J. P. C. Futuristic view of the Internet of Quantum Drones: Review, challenges and research agenda. Vehicular Communications 36, 100487 (2022). doi:10.1016/j.vehcom.2022.100487

- [2] Murugan, D., Garg, A. & Singh, D. Development of an Adaptive Approach for Precision Agriculture Monitoring with Drone and Satellite Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10, 5322–5328 (2017). doi:10.1109/JSTARS.2017.2746185
- [3] Kellermann, R., Biehle, T. & Fischer, L. Drones for parcel and passenger transportation: A literature review. *Transportation Research Interdisciplinary Perspectives* 4, 100088 (2020). doi:10.1016/j.trip.2019.100088
- [4] Mohd Daud, S. M. S. et al. Applications of drone in disaster management: A scoping review. *Science & Justice* 62, 30–42 (2022). doi:10.1016/j.trip.2019.100088
- [5] Stehle, R. H., DesBrisay, G. A., Whitson, A. L. & Hagn, G. H. RELEDOP: a full-scale antenna pattern measurement system for high-powered HF transmitting antennas. *IEEE Transactions on Broadcasting* 34, 210–220 (1988).
- [6] Bärffuss, K., Schulz-Stellenfleth, J. & Lampert, A. The Impact of Offshore Wind Farms on Sea State Demonstrated by Airborne LiDAR Measurements. *Journal of Marine Science and Engineering* 9, 644 (2021). doi:10.3390/jmse9060644
- [7] Punzet, S. et al. Fully Coherent UAV-Based Near-Field Measurement and Transformation of the S67-15 m Ground Station Antenna at the German Space Operations Center in Weilheim. in 2022 16th European Conference on Antennas and Propagation (EuCAP) 1–5 (2022). doi:10.23919/EuCAP53622.2022.9769018.
- [8] Rohi, G., Ejofodomi, O. & Ofualagba, G. Autonomous monitoring, analysis, and countering of air pollution using environmental drones. *Heliyon* 6, e03252 (2020). doi:10.1016/j.heliyon.2020.e03252
- [9] Go, Y.-J. & Choi, J.-S. An Acoustic Source Localization Method Using a Drone-Mounted Phased Microphone Array. *Drones* 5, 75 (2021). doi:10.3390/drones5030075
- [10] Neitz, O., Mauermayer, R. A. M., Weitsch, Y. & Eibert, T. F. A Propagating Plane-Wave-Based Near-Field Transmission Equation for Antenna Gain Determination from Irregular Measurement Samples. *IEEE Trans. Antennas Propagat.* 65, 4230–4238 (2017). doi:10.1109/TAP.2017.2712180
- [11] Xu, Y., Tan, Q., Erricolo, D. & Uslenghi, P. L. E. Fresnel-Kirchhoff integral for 2-D and 3-D path loss in outdoor urban environments. *IEEE Transactions on Antennas and Propagation* 53, 3757–3766 (2005). doi:10.1109/TAP.2005.858572
- [12] Aalipour, R. Calculation of Fresnel diffraction from 1D phase step by discrete Fourier transform. *Optics Communications* 382, 651–655 (2017). doi:10.1016/j.optcom.2016.08.031
- [13] Sabatyan, A. & Tavassoly, M. T. Application of Fresnel diffraction to nondestructive measurement of the refractive index profiles of optical fibers. *Optical Engineering* 46, 128001 (2007). doi:10.1117/1.2819745
- [14] Dolcetti, G., Alkmim, M., Cuenca, J., De Ryck, L. & Krynkina, A. Robust reconstruction of scattering surfaces using a linear microphone array. *Journal of Sound and Vibration* 494, 115902 (2021). doi:10.1016/j.jsv.2020.115902
- [15] Evans, D., Hartmann, M. & Delfs, J. Beamforming for point force surface sources in numerical data. *Journal of Sound and Vibration* 458, 303–319 (2019). doi:10.1016/j.jsv.2019.05.030
- [16] Kirkup, S. The Boundary Element Method in Acoustics: A Survey. *Applied Sciences* 9, 1642 (2019).
- [17] Gibson, W. C. The Method of Moments in Electromagnetics. (Chapman and Hall/CRC, 2021). doi:10.1201/9780429355509.
- [18] Emimi, M., Khaleel, M. & Alkrash, A. The Current Opportunities and Challenges in Drone Technology. *International Journal of Electrical Engineering and Sustainability (IJEES)* 74–89 (2023). doi:10.65998/ijeess.v1i3.47
- [19] Cicchetti, R. et al. Near-Field to Far-Field Transformation Techniques with Spiral Scannings: A Comprehensive Review. *International Journal of Antennas and Propagation* 2014, e143084 (2014). doi:10.1155/2014/143084

- [20] Eibert, T. F., Punzet, S., Mittereder, T., Faul, F. T. & Paulus, A. H. UAV-Based Near-Field Measurements at a Doppler Very High Frequency Omnidirectional Radio Range. in 2023 17th European Conference on Antennas and Propagation (EuCAP) 1–5 (2023). doi:10.23919/EuCAP57121.2023.10133603
- [21] Fernández, M. G., Álvarez López, Y. & Andrés, F. L.-H. Advances in Antenna Measurement and Characterization Using Unmanned Aerial Vehicles. in 2019 13th European Conference on Antennas and Propagation (EuCAP) 1–5 (2019).
- [22] Wei, L. & Rodríguez-Fortuño, F. J. Far-field and near-field directionality in acoustic scattering. *New J. Phys.* 22, 083016 (2020). doi:10.1088/1367-2630/ab9fb
- [23] Sun, T., Li, G. & Wang, K. A Near-to-far-field Transformation Method for Acoustic Scattering Problems. in 2007 International Conference on Mechatronics and Automation 304–308 (2007). doi:10.1109/ICMA.2007.4303559
- [24] Plotnick, D. S., Marston, P. L. & Marston, T. M. Fast nearfield to farfield conversion algorithm for circular synthetic aperture sonar. *The Journal of the Acoustical Society of America* 136, EL61–EL66 (2014). doi:10.1121/1.4885486
- [25] Maynard, J. D., Williams, E. G. & Lee, Y. Nearfield acoustic holography: I. Theory of generalized holography and the development of NAH. *The Journal of the Acoustical Society of America* 78, 1395–1413 (1985). doi:10.1121/1.392911
- [26] Lavrentjev, J. & Åbom, M. Characterization of fluid machines as acoustic multi-port sources. *Journal of Sound and Vibration* 197, 1–16 (1996). doi:10.1006/jsvi.1996.0514
- [27] Godin, O. A. The Kirchhoff–Helmholtz integral theorem and related identities for waves in an inhomogeneous moving fluid. *The Journal of the Acoustical Society of America* 99, 2468–2500 (1996). doi:10.1121/1.415524
- [28] Kirchhoff, G. Zur Theorie der Lichtstrahlen. *Annalen der Physik* 254, 663–695 (1883). doi:10.1002/andp.18832540409
- [29] Raja Subramanyam, Ramesh, Rafael Castro-Mota, Kevin Picker, Volker Wittstock, and Stefan Jacob. ‘Can Drones Serve as Carriers for Acoustic Measurement Equipment?’ 27 November 2024. doi:10.17866/rd.salford.27913320.v1
- [30] Subramanyam, Ramesh Raja, Johan Christian van Aswegen, Criostoir Gerasch, Volker Wittstock, and Stefan Jacob. ‘A Comprehensive Approach to Airborne Low-Frequency Acoustic Field Measurement Using a Drone-Based System’. *Sensors International* 7 (January 2026): 100384. doi:10.1016/j.sintl.2026.100384
- [31] Blumendeller, Esther, Ivo Kimmig, Gerhard Huber, Philipp Rettler, and Po Wen Cheng. ‘Investigations on Low Frequency Noises of On-Shore Wind Turbines’. *Acoustics* 2, no. 2 (2020): 343–65. doi:10.3390/acoustics2020020
- [32] Ramesh Raja Subramanyam, Criostoir Gerasch, Stefan Jacob. ‘Application of Atmospheric Wind Filter Arrays for Acoustic Monitoring UAVs’. *Forum Acusticum, Euronoise conference, 2025, Malaga, Spain*. doi:10.61782/fa.2025.0141
- [33] Subramanyam, Ramesh Raja, Rafael Castro-Mota, Volker Wittstock, and Stefan Jacob. ‘Synergizing Mode Decomposition and Numerical Analysis for Creating Sound Sources with Known Power Output: Application to an Outdoor Infrasound Source’. *Measurement* 255 (November 2025): 117956. doi:10.1016/j.measurement.2025.117956
- [34] Faul, F., Near-Field Measurement and Transformation of Modulated Electromagnetic Fields using Uncrewed Aerial Vehicles. *Dissertation, Technische Universität München* (2024), 80-81.
- [35] Ulm, D., Kleine-Ostmann, T. & Schrader, T. Antenna calibration based on near-field to far-field transformation algorithms. 2019 13th European Conference on Antennas and Propagation (EuCAP) (2019).
- [36] Datenblatt: mosaic-X5 Compact multi-constellation GNSS receiver, Septentrio (2025).

Effect of UAV Size on EKF-Based Position Uncertainties and the Performance of UAV-Mounted RIS

David Müller¹, Kevin Weinberger², Aydin Sezgin², Martin Mönnigmann¹

¹*Automatic Control and Systems Theory, Ruhr-Universität Bochum, Bochum, Germany*

²*Digital Communication Systems, Ruhr-Universität Bochum, Bochum, Germany*

Abstract

Reconfigurable Intelligent Surfaces (RIS) are able to extend the range of radar applications, allowing coverage of areas that were previously impossible or hard to reach. This is achieved by dynamically manipulating the electromagnetic wave properties with multiple adjustable reflecting elements in real time. This capability proves most effective when the RIS is positioned in line-of-sight (LoS) and proximity to the transmitter and receiver. To meet these criteria in real-world applications, the RIS is mounted on an unmanned aerial vehicle (UAV). However, a UAV is exposed to many disturbances during flight, which compromise the performance of the RIS-enabled link. Our previous work has shown that dynamically adjusting the RIS using the UAV position estimates from an extended Kalman filter (EKF) can effectively counteract disturbances and significantly enhance the channel link quality compared to a static RIS configuration. We quantified the uncertainties regarding the position estimation of the RIS in experiments. Data have so far been collected using a laboratory-scale UAV, i.e. a UAV with a rotor-to-rotor diameter of 500mm.

In this work, we deploy a UAV with a rotor-to-rotor diameter of 1060mm to investigate the position uncertainties of the UAV-mounted RIS and their impact on the channel link quality for a UAV that is capable of more demanding real-world missions. Building on this, we conduct a comprehensive comparison with our laboratory-scale UAV to assess the scalability of the proposed measurement methodology.

Keywords: Unmanned Aerial Vehicle (UAV), Reconfigurable Intelligent Surface (RIS), Uncertainty Propagation, Experimental Measurements, Prototypes

Introduction

Reconfigurable intelligent surfaces (RISs), originally designed for 6G wireless networks [1], have attracted attention for sensing applications due to their ability to control the propagation environment. By adjusting the phase, amplitude, and frequency of the incoming signals at each element, RISs are capable of extending the range of radar applications, allowing coverage of areas that were previously inaccessible [2].

A key performance metric for sensing is the effective end-to-end channel quality between the transmitter (Tx)-RIS-receiver/target (Rx), which directly influences the achievable signal-to-noise ratio. This effective channel and the resulting radar performance depend on proper configuration of the RIS as well as its position, where maintaining a line-of-sight (LoS)-link is particularly important. However, environmental factors such as blockages or changing terrain heights often prevent these conditions from being satisfied. To circumvent this problem, unmanned aerial vehicles (UAVs) have been

proposed to carry the RIS, such that the LoS-link can be guaranteed. Because of its low weight and power efficient design, RISs are excellent alternatives to heavy and power-hungry base stations for use with UAVs.

While mounting the RIS on a UAV enables reposition, the UAV is subject to external disturbances and internal sources of uncertainty. These uncertainties lead to constant changes in the position of the UAV, even during hover flight. Consequently, the RIS-reflected path is subject to these changes, affecting signal transmission. Due to the resulting continuous changes in the channel link, it becomes evident that a constant, predetermined RIS configuration does not suffice [3].

Our previous works have shown that using UAV onboard sensors in combination with an extended Kalman Filter (EKF) for state estimation of the UAV, followed by adjusting the RIS configuration based on the UAV state, can significantly improve the channel link quality compared to a static, time-invariant configuration [4,5]. These performance gains

persist despite the presence of estimation uncertainties. However, prior experiments have been limited to deployment of a laboratory-scale UAV, restricting conclusions regarding real-world performance.

In this work, we extend our analysis by deploying a larger UAV with a rotor-to-rotor diameter of 1060 mm to investigate the effect of UAV size on position uncertainties and their impact on the channel link quality. Similar to previous works, we follow the *Guide to the expression of uncertainty in measurement (GUM)* [6] to analyze the uncertainties. Based on the results obtained, we perform a comprehensive comparison with a laboratory-scale UAV to evaluate the scalability of the proposed measurement methodology.

System and Channel Model

The system model in this paper assumes a single-antenna Tx transmitting signals to a single-antenna Rx on the ground. It is further assumed that the direct Tx-Rx link on the ground is blocked and a UAV-mounted RIS is used to establish a LoS link. Each RIS consists of M elements, capable of altering the phase of the reflected signal. The reflected signals from the RIS are then received by the Rx. The effective Tx-RIS-Rx channel can be expressed as

$$h^{eff} = \sum_{m=1}^M h_m \theta_m g_m,$$

where h_m and g_m represent the channel coefficients between the m -th RIS element and the Tx/Rx, respectively. The reflect coefficient $\theta_m = A_m(\varphi_m)e^{j\varphi_m}$ is defined by the tunable phase $\varphi_m \in [0, 2\pi)$ induced to the reflected signal at each reflecting element m , and the phase-dependent reflect amplitude $A_m \in [0, 1]$. The cascaded channel coefficient between Tx-RIS-Rx over the m -th RIS element can be formulated as $h_m^{casc} = h_m g_m$.

We use pose data obtained from experimental measurements to compute the channel coefficients. We determine the cascaded channel components for each reflecting element m based on the geometry of the LoS path, where d_m^h (d_m^g) is defined as the distance between the Tx (Rx) and the m -th RIS element. The cascaded Tx-RIS-Rx channel through the m -th RIS element is given by [7]

$$h_m^{eff} = \left[\frac{c}{4\pi\nu d_m^h} e^{j\frac{2\pi}{\lambda} d_m^h} \right] \left[\frac{c}{4\pi\nu d_m^g} e^{j\frac{2\pi}{\lambda} d_m^g} \right], \quad (1)$$

where c is the speed of light, $\nu = 5.53\text{GHz}$ the frequency and λ the corresponding wavelength.

Experimental Setup

We mounted RIS prototypes [8] to two UAVs varying in size (see Fig. 1) and recorded their orientation and position both as estimated with the EKF and measured with a high-precision motion capture system (MCS) during outdoor hover flight. By comparing these datasets, we investigate the effect of UAV sizes on the uncertainties in the EKF position estimation. The impact on the channel link quality is subsequently analyzed.

The larger UAV used for the experiments is a PM-X6 hexacopter with a rotor-to-rotor diameter of 1060mm and a net weight of 5kg. The laboratory-scale UAV is a customized Holybro X500 quadcopter with a rotor-to-rotor diameter of 500mm and a net weight of 1.8kg. Both UAVs are controlled by a Pixhawk CubeOrange flight controller powered by a STM32H753 microcontroller to eliminate any influence from processing performance. The flight controller holds three different inertial measurement units (IMU), which are used for redundancy (specifically, the primary IMU is a ICM20602, the second one a ICM20948 and the third one a ICM20649¹). The open-source software ArduPilot² (V4.2.0) is installed on the flight controller, running an individual EKF for every IMU. This enables the UAV to use lane-switching³, a transition between multiple EKFs in case variances exceed a certain threshold. The active EKF is selected in the same order as the IMUs. The primary and secondary IMU are mounted on a separate temperature-controlled, vibration-isolated board, while the third IMU is located directly on the flight controller. Three different IMUs are installed on the flight controller, as they have different frequency responses and will therefore respond differently to vibrations. We tuned filter parameters for better filtering of vibrations by running an in-flight Fast Fourier Transform on the UAV and analyzing spectra for characteristic frequencies. Measurements of the IMU are then fused with other sensors by the EKF. We install a Here3 GNSS antenna on the UAV, for position estimates in the x-y-plane. The Here3 is running a ublox M8P-2 GNSS⁴ module allowing for position measurements with a precision of 2.5m. We use the barometer installed on the CubeOrange flight controller for measuring the

¹ <https://ardupilot.org/copter/docs/common-thecubeorangeoverview.html>

² <https://ardupilot.org/>

³ <https://ardupilot.org/dev/docs/common-ek3-affinity-laneswitching.html>

⁴ <https://docs.cubepilot.org/user-guides/here-3/here-3-manual>

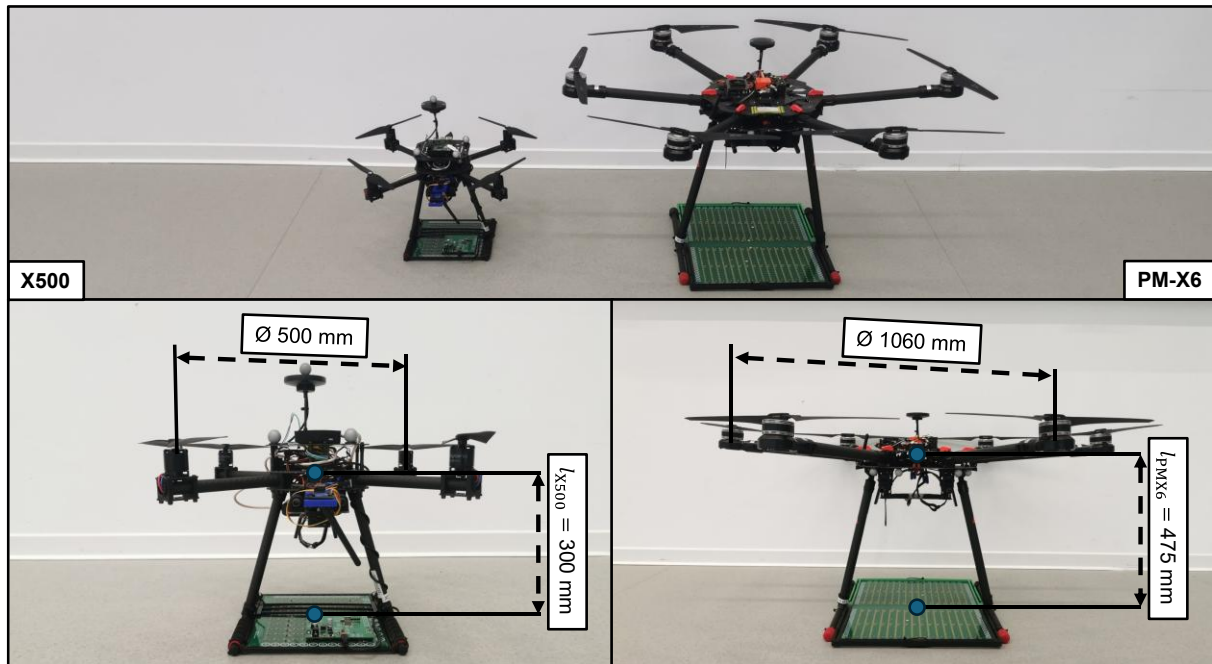


Figure 1: UAV setup for flight experiments. Blue dots represent the origin of the UAV's body frame (top) and the center of the RIS (bottom).

height over ground. Specifically, we use a MS5611 barometer, which is mounted on the same temperature and vibration-isolated board as the first and second IMU. To maintain the heading of the UAV, we fuse the IMU measurements with ICM20948 magnetometer data.

For the larger UAV (PM-X6), the RIS is mounted $l_{\text{PMX6}} = 475\text{mm}$ underneath the origin of the UAV's body frame (see Fig. 1), with no offsets in the x-y-plane. Similarly, the RIS is mounted $l_{\text{X500}} = 300\text{mm}$ underneath the origin of the laboratory scale UAV (X500). A Raspberry Pi, acting as the RIS controller, is installed on the UAVs. Note that the RIS mounted on the larger UAV is scaled to approximately four times the size of the laboratory-scale counterpart. This is possible because the larger UAV can support a larger RIS, which is beneficial for system performance and reflects realistic deployment scenarios. However, due to the RIS shape, it can interact with the rotor airflow, potentially contributing to system uncertainties. Given the larger rotor-to-rotor diameter, the increased RIS size on the larger UAV ensures a fair comparison with the laboratory-scale platform.

All experiments are conducted outdoors on an open field under the same environmental conditions. We perform 10 hover flights, each lasting 30s, with both UAVs. The UAVs are maneuvered to an altitude of approximately 3.5m using the stick inputs of a Hex Herelink 1.1

controller. The position is then maintained autonomously by the flight controller (position hold mode⁵). We restart the UAVs and calibrate the IMUs on the flight controller before every flight to ensure the same initial conditions. The EKF state estimates are recorded on an onboard SD-Card with a frequency of 100Hz. Reference data for the UAV position and orientation is measured and recorded with an MCS. Specifically, 10 Vicon Vantage V5 infrared cameras and the Software Tracker 3 were installed. This system is able to track objects equipped with an asymmetric pattern of reflective markers with a root mean squared error of less than 0.2mm with frequencies of up to 420Hz⁶. We calibrate the motion capture system once, after it has reached its working temperature. The motion capture system sampling frequency is set to 100Hz, i.e., the same logging frequency as the flight Controller.

We synchronize the motion capture system and UAV data prior to evaluation to remove the temporal phase shift between the two sets of data. Subsequently, we discard the first and last 5s of the recorded data, as these intervals correspond to takeoff and landing. Within the remaining 20s, we randomly extract a contiguous interval of 10s of hover flight. The corresponding 1000 data points are then used for investigating the uncertainty in estimating the position in the following section.

⁵ <https://ardupilot.org/copter/docs/poshold-mode.html>

⁶ <https://www.vicon.com/>

Uncertainty in RIS-Position Estimation

We characterize the uncertainties of the RIS position estimation with the EKF for both UAVs to evaluate the effect of UAV size on the estimation accuracy of the EKF, which is crucial for adjusting the RIS configuration. We evaluate the position uncertainties of the attached RIS by propagating the UAV position and orientation uncertainty. Orientations and their uncertainties enter because of the offset between the center of the RIS and the origin of the UAV body frame as described in the previous section.

We characterize the difference of the UAV position and orientation $x_{k,MCS}$, $y_{k,MCS}$, $z_{k,MCS}$, $\Phi_{k,MCS}$ (roll), $\theta_{k,MCS}$ (pitch), $\psi_{k,MCS}$ (yaw) measured with the MCS at time point k to the position and orientation $x_{k,EKF}$, $y_{k,EKF}$, $z_{k,EKF}$, $\Phi_{k,EKF}$, $\theta_{k,EKF}$, $\psi_{k,EKF}$ estimated with the EKF. Because the EKF introduces errors that are not mean free, we need to characterize these means. Let $x_{k,UAV}^e = x_{k,MCS} - x_{k,EKF}$, refer to the absolute error of the x -position of the UAV at time point k and let $y_{k,UAV}^e$, $z_{k,UAV}^e$, $\Phi_{k,UAV}^e$, $\theta_{k,UAV}^e$, and $\psi_{k,UAV}^e$ be defined accordingly. In order to avoid repeating expressions for x , y , z , Φ , θ , and ψ , we use q to refer to any of the six components (e.g., $\bar{q}_{UAV}^e$ in (2) may be replaced by $\bar{x}_{UAV}^e$, $\bar{y}_{UAV}^e$, or any other component). We characterize the absolute position and orientation errors with the arithmetic mean

$$\bar{q}_{UAV}^e = \frac{1}{n} \sum_{k=1}^n q_{k,UAV}^e, \quad (2)$$

the experimental variance of observations

$$s_{q_{UAV}^e}^2 = \frac{1}{n-1} \sum_{j=1}^n (q_{j,UAV}^e - \bar{q}_{UAV}^e)^2, \quad (3)$$

the experimental standard deviation $s_{q_{UAV}^e}$, the variance of the mean

$$s_{\bar{q}_{UAV}^e}^2 = \frac{s_{q_{UAV}^e}^2}{n}, \quad (4)$$

and the experimental standard deviation of the mean $s_{\bar{q}_{UAV}^e}$.

After calculating $\bar{q}_{UAV}^e$ with (2), we can calculate our best estimate for each component, i.e.

$$q_{k,EKF} - \bar{q}_{UAV}^e, \quad (5)$$

which is mean-free by construction and where the two terms have the uncertainties

$$u_{q_{UAV}^e}, \quad (6)$$

and

$$u_{\bar{q}_{UAV}^e} = s_{\bar{q}_{UAV}^e}, \quad (7)$$

respectively, with $s_{\bar{q}_{UAV}^e}$ found with (5), and where $u_{\bar{q}_{UAV}^e}$ is called the standard uncertainty. The uncertainty in (6) represents that of a single

EKF estimate. Unlike (7) and the experimental variance of the mean in (4), it must not be scaled down by the factor $1/n$ that results from n repeated measurements. Consequently, the reduction factor $1/n$ in (4) must be dropped, which implies $u_{q_{UAV}^e}^2 = s_{q_{UAV}^e}^2$.

Assume a quantity of interest $M \in \mathbb{R}^3$ ('the measurand') cannot be measured directly but determined from measurable quantities q_i with a measurement function f , i.e.,

$$M = f(q_1, q_2, \dots, q_n). \quad (8)$$

Then the uncertainty of M can be characterized with the combined standard uncertainty $u_{c,M}$, which is the positive square root of the combined variance [6, Sec. 5]

$$u_{c,M}^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial q_i} \right)^2 u_{q_i}^2, \quad (9)$$

where the partial derivative is evaluated at the estimates of the input quantities and where u_{q_i} is the standard uncertainty of the input quantity q_i .

For our system, three measurands, the components of the estimated position of the RIS center $p_{RIS} = [x_{RIS}, y_{RIS}, z_{RIS}]^T$ must be treated. The three instances of the measurement function (8) read

$$p_{RIS} = P_{UAV} + R(\Phi, \theta, \psi) l_{RIS}, \quad (10)$$

where $l_{RIS} = [0, 0, -l_m]$ describes the spatial offset between UAV center of mass and the RIS, where $m \in \{\text{PMX6}, \text{X500}\}$. The position of the UAV is given by $p_{UAV} = [x_{k,EKF} - \bar{x}_{UAV}, y_{k,EKF} - \bar{y}_{UAV}, z_{k,EKF} - \bar{z}_{UAV}]^T$ as in (6) with uncertainties (7) and (8), and $R(\Phi, \theta, \psi) \in \mathbb{R}^{3 \times 3}$ is the product of the rotation matrices

$$R(\Phi, \theta, \psi) = R_z(\psi) R_y(\theta) R_x(\Phi).$$

Evaluating (9) for the three measurands $p_{RIS} = [x_{RIS}, y_{RIS}, z_{RIS}]^T$ and their measurement functions (10) yields the combined variances

$$\begin{aligned} u_{c,x_{RIS}}^2 &= u_{x_{UAV}}^2 + u_{\bar{x}_{UAV}}^2 + \sum_{q_i} \left(v_x^T \frac{\partial R}{\partial q_i} \Big|_{\Phi_0, \theta_0, \psi_0} l_{RIS} \right)^2 u_{q_i}^2 \\ u_{c,y_{RIS}}^2 &= u_{y_{UAV}}^2 + u_{\bar{y}_{UAV}}^2 + \sum_{q_i} \left(v_y^T \frac{\partial R}{\partial q_i} \Big|_{\Phi_0, \theta_0, \psi_0} l_{RIS} \right)^2 u_{q_i}^2 \\ u_{c,z_{RIS}}^2 &= u_{z_{UAV}}^2 + u_{\bar{z}_{UAV}}^2 + \sum_{q_i} \left(v_z^T \frac{\partial R}{\partial q_i} \Big|_{\Phi_0, \theta_0, \psi_0} l_{RIS} \right)^2 u_{q_i}^2 \end{aligned} \quad (11)$$

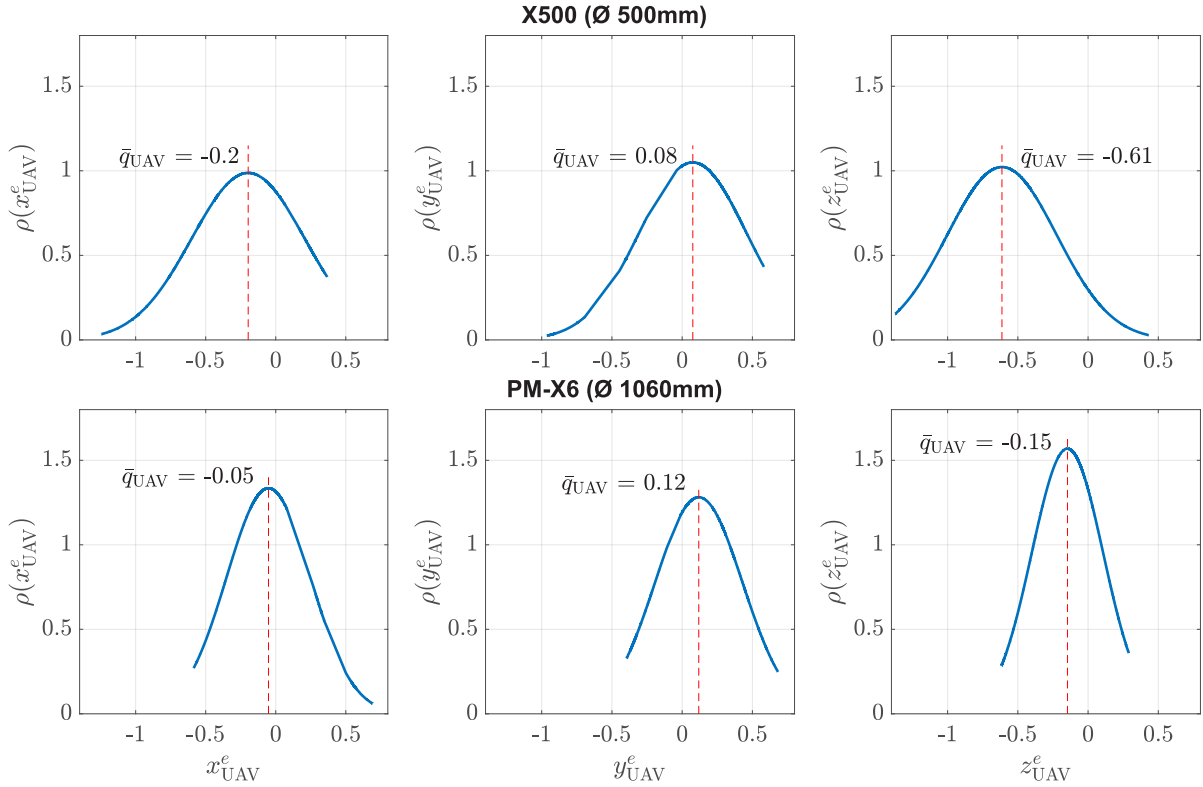


Figure 2: Normal distributions of the position errors for the laboratory-scale UAV X500 (top) and the larger UAV PM-X6 (bottom). Red lines indicate arithmetic means.

where sums are carried out for $q_i \in \{\Phi, \theta, \psi\}$, v_x , v_y , and v_z are the cartesian unit vectors, and $\Phi_0 = \theta_0 = \psi_0 = 0$ because we are treating hover flight. Moreover, $u_{\bar{x}_{UAV}}$ is as introduced in (7) and $u_{\bar{y}_{UAV}}$ and $u_{\bar{z}_{UAV}}$ are defined accordingly. The resulting combined standard uncertainties $u_{c,x_{RIS}}$, $u_{c,y_{RIS}}$, and $u_{c,z_{RIS}}$ are stated in Tab. 1.

Table 1 also lists the expanded combined uncertainty U_c for completeness. The expanded uncertainty is an interval of the form

$$q - U_c \leq q \leq q + U_c, \quad (12)$$

with $U_c = bu_{c,m}$. We choose $b = 2$, which corresponds to U_c defining an interval with an approximate 95% confidence level, commonly used in UAV and networking applications [9, 10].

Given the results in Tab. 1, the influence of orientation uncertainty for the larger UAV is found to be relatively small, similar to the laboratory-scale UAV, making the position uncertainty the predominant source of uncertainty. Although the larger RIS increases the effective surface area exposed to wind disturbances and alters the airflow around the UAV, the increased mass and inertia of the UAV itself appear to counteract these effects, resulting in orientation uncertainties of the same order for both systems. In fact, based on the findings in [4], these orientation uncertainties

are sufficiently small that they do not impact the channel link quality.

Tab. 1: Uncertainties in the UAV and RIS position estimation of the laboratory-scale UAV X500 (first values) and the larger UAV PM-X6 (second values).

	X500 /PM-X6		
GUM params	Φ in °	θ in °	ψ in °
$\bar{q}_{UAV}^e$	0.23 /-0.001	0.22 /-0.005	-0.06 /0.06
$u_{\bar{q}_{UAV}^e}$	0.005 /0.002	0.005 /0.002	0.002 /0.004
$u_{q_{UAV}}$	0.49 /0.20	0.48 /0.18	0.18 /0.43
	x in m	y in m	z in m
$\bar{q}_{UAV}^e$	-0.20 /-0.05	0.08 /0.12	0.61 /-0.15
$u_{\bar{q}_{UAV}^e}$	0.004 /0.003	0.004 /0.003	0.004 /0.003
$u_{q_{UAV}}$	0.40 /0.30	0.40 /0.30	0.38 /0.25
$u_{c,q_{RIS}}$	0.43 /0.30	0.43 /0.31	0.38 /0.25
$U_{c,q_{RIS}}$	0.86 /0.60	0.86 /0.62	0.76 /0.50

For the position uncertainty of the UAVs, the differences become more pronounced. For better comparison, Fig. 2 shows the normal distributions of the position uncertainties in each spatial direction for both UAVs. It can be observed that the larger UAV exhibits significantly smaller position uncertainties than the laboratory-scale UAV, with reductions of approximately 25% in the x- and y-directions and about 35% in the z-direction. Since the UAV position uncertainty represents the predominant contribution to the RIS position uncertainty, similar improvements are observed for the RIS position uncertainty. With more than 2.7 times the mass, and thus greater inertia, disturbances like wind result in slower position changes for the larger UAV compared to the laboratory-scale UAV. This behavior facilitates tracking by the EKF, resulting in more accurate position estimates. Based on these results, we expect an improvement in channel link quality compared to the laboratory-scale UAV, even though both UAVs are equipped with the exact same sensor setup. However, the remaining position uncertainties with an expanded uncertainty of about 60cm may still be too significant to achieve optimal channel quality at this frequency.

Comparison of the Inference of the Channel Link Quality

In this section, we assess how different uncertainties affect the channel link quality by comparing one representative flight of each of the UAVs. Furthermore, we show that, although the uncertainties associated with the larger UAV are between 25% and 35% smaller than those of the laboratory-scale UAV, they remain too significant to achieve optimal channel quality.

The Tx antenna is positioned at (0,0,0) below the center of the RIS with the Rx antenna positioned 2m away along the x-axis of the coordinate system from the Tx antenna. We calculate the current RIS position $p_{k,\text{RIS}}$ at time step k according to (10) for both UAVs based on data from an extra flight, which was performed and processed in the same way as the other 10 flights, but was not used for calculating the uncertainties. We assume $\phi = \theta = \psi = 0^\circ$ for the RIS orientation at all times, in order to isolate the effect of position fluctuations. Although the RIS carried by the larger UAV has been scaled from laboratory dimensions to a size suitable for real-world missions in order to capture airflow-related uncertainties, we additionally assume that both RISs have identical dimensions and the same number of elements. Specifically, we set $M =$

120, such that only position uncertainties affect the channel link quality.

Under these assumptions, we can calculate the distances d_m^h and d_m^g , i.e., between each reflecting element m for the Tx-RIS and RIS-Rx link, respectively. Consequently, we can determine the cascaded channel coefficients h_m^{casc} for each time step k with (1), which is prerequisite for the optimization process.

With the above assumptions, it becomes possible to compute the optimal RIS configurations based on the values of h_m^{casc} in the observed interval. Because this optimization must be performed in real-time, i.e. at every time step k , we need to carefully consider the computational cost when designing the optimization algorithm. Furthermore, the RIS is limited to applying only 180° phase shifts, i.e., $\varphi_m \in \{0^\circ, 180^\circ\}$, and suffers from a -3 dB loss in reflection amplitude A_m for $\varphi_m = 180^\circ$, consistent with the behavior of its physical counterpart.

Due to the blocked path between Tx and Rx, the phase of the effective channel h^{eff} can be selected arbitrarily during the optimization process. This can be expressed as

$$\varphi_m^*(C) = C - \varphi'_m, \quad (13)$$

where $C \in \mathbb{R}$ denotes an arbitrary value for the desired phase of h^{eff} and $\varphi'_m = \frac{2\pi}{\lambda}(d_m^h + d_m^g)$. We select C from a finite set of to ensure the optimization to be fast enough for use in real time. In the present paper, we use the values

$$C \in \zeta = \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}\right\}, \quad (14)$$

which results in a suitable trade-off to identify a good solution within the available time [8]. It will also be evident from the results in the following section that this restriction to $|\zeta| = 8$ values does not constitute any limitation.

After optimizing the phase shift φ_m^* for every patch $m = \{1, \dots, M\}$, we need to determine which of the patch states, $\varphi_m^* = 0^\circ$ or 180° , is the optimal choice. We choose a fast rounding operation in light of the real-time requirements. Note that the same approach would still hold if continuous phase shifts were available. In fact, the problem would become simpler, as removing the hardware-imposed constraint of selecting only discrete phase states allows for more flexible optimization and can potentially improve performance. For the resulting binary RIS configurations, we choose the best performing one by assessing and comparing the simulated channel quality of the resulting RIS-facilitated links. This procedure can mathematically be formulated as maximizing h^{eff} for all $C \in \zeta$

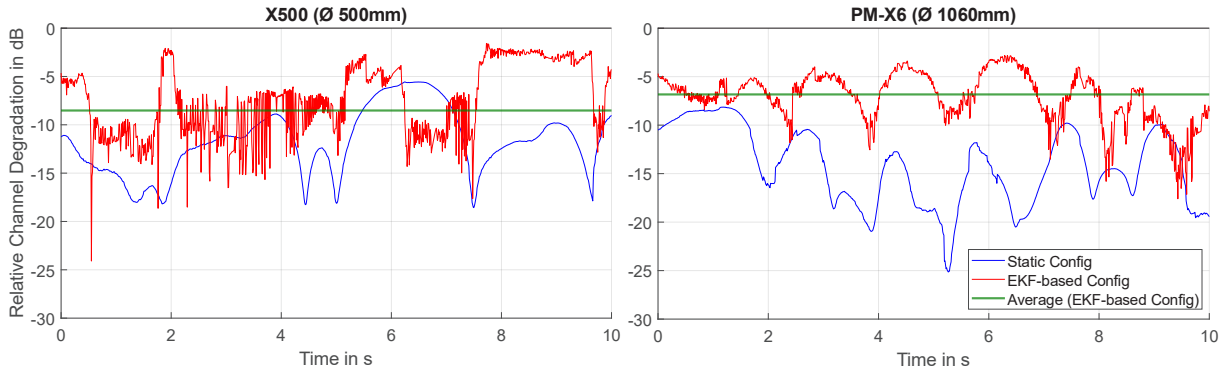


Figure 3: Relative channel degradation of the laboratory-scale UAV (X500 – left) and the larger UAV (PM-X6 – right). The base line (0dB) is the optimal channel quality achieved by using MCS data (i). Red lines indicate relative channel degradation for the EKF-based RIS configuration (ii), where the green lines show the average relative channel degradation for (ii). Blue lines show the average relative channel degradation for the static RIS configuration (iii).

$$h_c^{eff}(C) = \sum_{m=1}^M h_m^{casc} A_m \left(\text{rd}(\varphi_m^*(C)) \right) e^{j \text{rd}(\varphi_m^*(C))}$$

$$\text{with } A_m(\tau) = \begin{cases} 0.5012 \text{ (-3dB)}, & \text{if } \tau = \pi \\ 1, & \text{otherwise} \end{cases}$$

$$\text{rd}(\tau) = \begin{cases} \pi, & \text{if } \frac{\pi}{2} \leq \tau \leq \frac{3\pi}{2} \\ 0, & \text{otherwise} \end{cases}$$

$$C \in \zeta = \left\{ 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4} \right\} \quad (15)$$

in every time step k . This optimization can be carried out efficiently, because the optimal C and $\varphi_m^*(C)$ can be found independently for each patch by comparing the $|\zeta| = 8$ choices and rounding subsequently.

Finally, we can compare how different UAV sizes and their uncertainties affect the channel quality h^{eff} when optimizing the RIS with the respective EKF data. While realistic propagation environments introduce additional effects such as multipath fading and scattering, which may degrade performance, the real-time reconfigurability of the LoS path of the RIS provides performance gains that outweigh these effects, leading to significant observable improvements in practice [3].

We recall that we want to analyze the impact of the different EKF-based RIS position uncertainties and thus assume the RIS orientation to be constant and the UAVs carrying identical RIS. We evaluate the optimal effective channel h^{eff} for three different cases of RIS position information for each UAV: (i) based on the UAV's position and orientation measured with the motion capture system for reference, (ii) based on the UAV's position and orientation estimated with the EKF, (iii) under the assumption that only the target position is available, where (iii) is equivalent to assuming a fixed known x-, y-, z-position for all times. For cases (i) and (ii), the optimal RIS configuration is determined according to (15). For case (iii), we optimize the RIS for $x = y = 0$ and $z = 3.5$

once and maintain the resulting RIS configuration for all k .

Figure 3 shows the optimal effective channel h^{eff} based on the MCS measurements from the flight experiments conducted with the laboratory-scale X500 (left) and the larger PM-X6 (right). The static (iii) and EKF-based (ii) setups are compared against the MCS configuration (i), which acts as a reference. The figure shows that the static and EKF-based RIS configurations deteriorate in performance when compared to the MCS-based configurations for both UAVs, which highlights the impact of the RIS position estimation uncertainties on the channel gain. However, when comparing the impact of the position uncertainties on the channel gain between the larger PM-X6 and the X500, it becomes evident that the PM-X6 significantly outperforms the X500. Specifically, configuring the RIS based on the EKF information from the PM-X6 degrades the channel gain by 6.75dB on average, whereas reconfiguring the RIS based on X500 data degrades the channel by 8.5dB, corresponding to an approximately 21% higher loss in dB. The relative loss is consistent with the uncertainties determined in the previous section, indicating that the reduction in uncertainty is directly impacting the quality of the effective channel. A similar trend is observed for the maximum channel degradation. The estimates from the PM-X6 degrade the channel by up to 17.6dB, while the estimates from the X500 reach a maximum degradation of 24.1dB. This corresponds to a difference of approximately 27%, aligning with the previously determined uncertainties. These results reveal that scaling the proposed measurement technology to larger platforms is not only feasible but also beneficial, as it reduces uncertainty and improves potential measurement accuracy. Note that in practice, differences in the absolute channel gain may be even more pronounced,

since larger UAVs are capable of carrying significantly larger RIS arrays.

Conclusion

We investigated the effect of UAV size on the position uncertainties and the channels of UAV-mounted RIS, by conducting flight experiments with a laboratory-scale UAV and a UAV more than twice its size. For investigation of the position uncertainties, we followed the *Guide to the expression of uncertainty in measurement (GUM)* [6]. Despite the larger surfaces exposed to wind and the increased airflow disturbances caused by the larger RIS mounted on the larger UAV, the higher mass and inertia of the larger UAV mitigate these effects. This results in slower dynamics, which improve the tracking performance of the EKF and lead to reduced position uncertainties compared to the laboratory-scale setup.

We subsequently analyzed the impact of the uncertainties on the channel link quality by adjusting the RIS configuration based on real EKF data from both UAVs through numerical simulations. Consistent with the results from the uncertainty analysis, the deployment of the larger UAV results in reduced channel degradation relative to the optimal RIS configuration compared to the laboratory-scale UAV, indicating that scaling the proposed system to larger platforms is beneficial. However, despite the reduced uncertainties achieved with the larger UAV, they remain too high to match the channel quality when adjusting the RIS configuration based on ground truth data. Consequently, system upgrades are still necessary, where additional sensors like real-time kinematic GNSS could be particularly useful. These upgrades also become essential when considering more dynamic flight maneuvers, as position uncertainty is likely to increase due to more rapid changes in the estimated state.

Acknowledgment

This work was funded by the German Research Foundation (DFG) in the course of the project SPP2433 under the project no. 541021107 (Measurement Technology on Flying Platforms) under grant MO 1086/17-1, grant SE 1697/22-1, and grant INST 213/1028-1 FUGB.

References

- [1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M. Alouini, and R. Zhang, "Wireless communication through reconfigurable intelligent surfaces," in *IEEE Access*, vol. 7, pp. 116753-116773, 2019; doi: 10.1109/ACCESS.2019.2935192
- [2] S. Buzzi, E. Grossi, M. Lops and L. Venturino, "Radar Target Detection Aided by Reconfigurable Intelligent Surfaces," in *IEEE Signal Processing Letters*, vol. 28, pp. 1315-1319, 2021, doi: 10.1109/LSP.2021.3089085
- [3] K. Weinberger, S. Tewes, J. Müller, R. Dyrka, M. Mönnigmann and A. Sezgin, "Flying Intelligent Surfaces: Joint Adjustment of Position and Configuration for UAV-Mounted RIS," *2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)*, 2023, pp. 1-5, doi: 10.1109/VTC2023-Spring57618.2023.10200263
- [4] D. Müller, K. Weinberger, R. Dyrka, A. Sezgin and M. Mönnigmann, "Assessing EKF-based Orientation Uncertainties and its Impact on the Channels of UAV-mounted RIS," *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, 2024, pp. 1-5, doi: 10.1109/VTC2024-Fall63153.2024.10757906
- [5] D. Müller, K. Weinberger, A. Sezgin and M. Mönnigmann, "The Impact of EKF-Based Position Uncertainties on the Performance of UAV-Mounted RIS," *2025 25th International Conference on Control, Automation and Systems (ICCAS)*, 2025, pp. 1339-1344, doi: 10.23919/ICCAS66577.2025.11301115
- [6] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, and OIML, "Evaluation of measurement data — Guide to the expression of uncertainty in measurement." URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008E.pdf/cb0ef43f-baa5-11cf-3f85-4dcd86f77bd6.
- [7] A. Goldsmith, "Wireless communications," *Cambridge university press*, 2005, doi: 10.1017/CBO9780511841224
- [8] K. Weinberger, S. Tewes, M. Heinrichs, R. Kronberger and A. Sezgin, "RIS-Based Channel Modeling and Prototypical Validation," *WSA & SCC 2023; 26th International ITG Workshop on Smart Antennas and 13th Conference on Systems, Communications, and Coding*, 2023, pp. 1-6, ISBN: 978-3-8007-6050-3
- [9] K. Xue, J. Rodríguez-Piñero, Y. Yu, J. Hong, X. Yin and X. Shunqin, "Performance and Reliability of 5G Communications for USV-UAV Critical Applications," *2023 17th European Conference on Antennas and Propagation (EuCAP)*, 2023, pp. 1-5, doi: 10.23919/EuCAP57121.2023.10132977
- [10] N. E. -D. Safwat, I. M. Hafez and F. Newagy, "Air-To-Air Channel Model for UAVs In Dense Urban Environments," *2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT)*, 2021, pp. 204-209, doi: 10.1109/ICEECCOT52851.2021.9707927

Reliable measurements for safe rescue operations after building collapses (ZeMeRGe)

Dominik M. Spale¹, Gunnar Gidion¹, Joel Schiffert¹, Eric Maier², Katrin Schmitt², Stefan J. Rupitsch¹

¹Department of Microsystems Engineering (IMTEK), University of Freiburg, Germany

²Fraunhofer Institute for Physical Measurement Techniques (IPM), Germany

dominik.spale@imtek.uni-freiburg.de

Abstract

This contribution presents an overview of the ZeMeRGe project, which aims to develop reliable sensor systems for search-and-rescue operations in disaster scenarios. In particular, a radar-based system for contactless vital sign detection is investigated, enabling the detection of respiration in survivors trapped under debris using an unmanned aerial vehicle (UAV).

While the feasibility of radar-based vital sign detection has been demonstrated in previous work, its practical deployment requires a systematic understanding of measurement uncertainty. This paper introduces the underlying measurement principles and discusses the main factors influencing detection reliability, including physiological variability, UAV motion, environmental effects such as debris composition, and radar system design parameters.

Different radar approaches, including stepped-frequency continuous-wave (SFCW) and frequency-comb continuous-wave (FCCW) systems, are presented and compared in the context of UAV-based operation. In addition, simulation methods based on finite-difference time-domain techniques are introduced as a tool for controlled uncertainty analysis.

Experimental results obtained under laboratory and field conditions illustrate the impact of target orientation and platform motion on signal quality and detectability. The presented results provide a foundation for future quantitative uncertainty budgets and contribute to the development of robust and trustworthy radar-based systems for supporting first responders.

Keywords: radar, vital sign detection, measurement uncertainty, unmanned aerial vehicles

Introduction

In disaster scenarios such as earthquakes or gas explosions, emergency personnel must rapidly assess both environmental hazards and the presence of survivors. Currently, rescue dogs are widely used for victim detection. However, their deployment is limited in duration due to fatigue, and they cannot be used in hazardous environments, such as areas with potential gas leaks.

These limitations motivate the development of complementary technological solutions. Within the ZeMeRGe project, a radar-based approach for contactless vital sign detection for deployment on unmanned aerial vehicles (UAVs) is investigated at the Department of Microsystems Engineering (IMTEK). This so-called bioradar exploits the Doppler modulation caused by periodic thoracic motion to detect human respiration, as illustrated in Fig. 1.

While previous studies have demonstrated the feasibility of radar-based vital sign detection, it has become evident that detection performance strongly depends on environmental and

system-related factors. These include debris composition, the orientation of the survivor, UAV motion, and radar system parameters. As a result, first responders consistently raise a critical question: How reliable are the measurements?

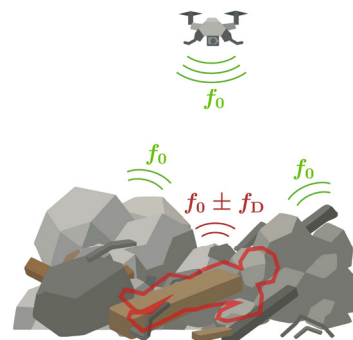


Fig. 1: Radar-based vital sign detection from a UAV. While clutter in the environment does not modulate the radar echo, thoracic movement introduces a Doppler shift.

To address this, the ZeMeRGe project aims to systematically quantify the measurement uncertainty of both the bioradar and a complementary gas sensing system, following the GUM (Guide to the Expression of uncertainty in measurement) methodology [1]. For the radar system, this requires understanding how physical, environmental, and algorithmic factors influence the detectability of respiration signals.

Bioradar measurement principle

Radar-based vital sign detection relies on the fact that human respiration induces small, quasi-periodic movements of the chest. These movements modulate the reflected electromagnetic wave and can be detected remotely.

In a bioradar system, electromagnetic waves in the low GHz range are transmitted towards a target area. The waves propagate through the environment, including debris, and are partially reflected by objects such as the human body. The periodic displacement of the chest induces a Doppler shift in the backscattered signal, thereby enabling the extraction of vital-sign information. The received signal therefore contains information about motion-induced changes in phase and amplitude, as illustrated in Fig. 2. This principle is well-known in medical monitoring of vital signs [2] as well as for search-and-rescue (SAR) missions [3], as is the case for this project. While radar-based vital sign detection, especially in the medical context, is an active research field, little is known about measurement uncertainty in SAR applications, especially on UAVs.

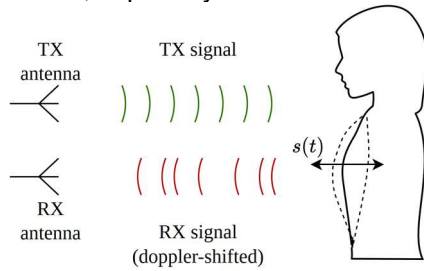


Fig. 2: Working principle of the bioradar. The echo signal is modulated in both frequency/phase and amplitude by the chest movement.

The thoracic motion can be approximated as a quasi-periodic displacement

$$s(t) = s_{pp} \sin(2\pi f_{resp} t) \quad (1)$$

with the breathing frequency f_{resp} and the peak-to-peak chest displacement s_{pp} . The assumption of quasi-periodicity is justified by

the relatively stable respiration rate over short observation intervals.

For adults at rest, the respiration rate typically lies between 12 and 20 breaths per minute (0.2 to 0.33 Hz), although significant variability can occur in practice [4]. This frequency range is important for system design, as it enables the discrimination of respiration from other motion sources, such as environmental clutter or mechanical vibrations. In addition to phase modulation (Doppler effect), respiration also alters the radar cross-section of the chest due to volumetric changes. Both effects can be exploited to estimate the breathing frequency.

Radar signal chain

To better understand how this information is extracted, it is useful to consider the underlying signal model.

The transmitted radar signal $r_{TX}(t)$ propagates through the environment and is reflected by a target at distance R . The received signal can be approximated as

$$r_{RX}(t) = \alpha \cdot r_{TX}(t - \tau) + n(t), \quad (2)$$

where α represents attenuation due to propagation and reflection, τ is the round-trip delay, and $n(t)$ denotes noise and interference. For a breathing human target, the radar distance $R(t)$ determining the round-trip delay is time-dependent with

$$R(t) = R_0 + s(t), \quad (3)$$

where $s(t)$ describes the respiration-induced chest displacement as described in Eq. (1) and R_0 is a time-independent radar distance component. This results in a time-varying phase modulation of the received signal $r_{RX}(t)$. For a single frequency component with wavelength λ , the phase φ is given by

$$\varphi = \frac{4\pi}{\lambda} s(t). \quad (4)$$

Thus, even small chest movements on the order of millimetres lead to measurable phase changes in the radar signal, which form the basis for respiration detection. Based on this signal model, the choice of radar waveform determines how spatial (range) and temporal (motion) information can be extracted from the received signal.

Radar types in CW vital sign detection

In its simplest form, a continuous-wave (CW) radar operating at a single frequency, as used in simple motion sensors in smart home applications, is sufficient to detect motion

through Doppler modulation. However, such systems cannot provide any information about the distance of the detected motion. This limitation is critical in search-and-rescue scenarios, where multiple motion sources may be present. Without spatial resolution, it is impossible to distinguish between relevant signals (e.g., a buried survivor) and irrelevant ones (e.g., moving vegetation), if they happen to exhibit similar Doppler shift frequencies. Fundamentally, the range resolution ΔR of a radar is given by

$$\Delta R = \frac{c}{2 \cdot BW} \quad (5)$$

with the speed of light c and the bandwidth of the excitation signal BW . This range resolution relation follows from the fact that radar distinguishes targets by their propagation delay. The smallest resolvable delay difference $\Delta\tau$ is inversely proportional to the signal bandwidth BW by

$$\Delta\tau \approx \frac{1}{BW}. \quad (6)$$

Since radar measures round-trip delay, range R and delay are related by

$$R = \frac{c\tau}{2}. \quad (7)$$

Hence, the corresponding range resolution takes the form of Eq. (5). A larger bandwidth therefore improves the ability to separate targets in distance. Intuitively, a larger bandwidth enables finer discrimination between different propagation delays and thus between different target distances.

To achieve this required bandwidth, different radar concepts have been developed, each with specific advantages and trade-offs in terms of complexity, measurement speed, and robustness.

One widely used approach is frequency-modulated continuous-wave (FMCW) radar. In FMCW systems, the transmit frequency is continuously swept over time. Due to the propagation delay, the received signal exhibits a frequency offset relative to the transmitted signal, which is proportional to the target distance.

An alternative approach is stepped-frequency continuous-wave (SFCW) radar. Here, the required bandwidth is synthesised by frequency hopping, i.e., sequentially transmitting discrete frequency tones. The phase differences across these frequencies are then combined to reconstruct a range profile using Fourier-based processing. While SFCW offers high flexibility

and accurate range estimation, this approach is inherently time-consuming due to the sequential frequency sweep. After each frequency hop, the synthesiser requires a finite settling time before a phase-stable measurement can be acquired. Since this stabilisation delay occurs at every frequency step, the total sweep duration increases with the number of sampled frequencies.

To circumvent this issue, one approach currently being investigated in the ZeMeRGe project employs a frequency-comb continuous-wave (FCCW) radar. In this system, multiple frequency components are transmitted simultaneously. This enables instantaneous wideband operation, significantly faster acquisition, and reduced synchronization requirements compared to FMCW and SFCW. These properties make FCCW particularly attractive for UAV-based bioradar applications, where measurement time, system flexibility, and robustness against motion are critical.

In practical implementations, the range information obtained from both SFCW and FCCW is discretized into so-called range bins, where each bin corresponds to a specific distance interval from the radar. By transforming the multi-frequency signal into the spatial domain, the received signal can be decomposed into multiple distance-dependent components. Each range bin can then be analysed individually in the time domain, allowing the separation of relevant signals, such as human respiration, from other motion sources located at different distances. The spatial extent of each range bin is determined by the range resolution ΔR given by Eq. (5). For the FCCW radar system with a bandwidth of approximately 60 MHz, this results in a range bin width of approximately $\Delta R \approx 2.5$ m.

After range separation, each range bin provides a time-dependent signal corresponding to a specific spatial region. These signals are typically analysed in the slow-time domain to extract periodic components associated with respiration. In this contribution, spectral methods such as the fast Fourier transform (FFT) and time-frequency methods such as the continuous wavelet transform (CWT) are used to estimate the respiration frequency. These methods allow robust extraction of periodic components even in the presence of noise and non-stationary effects.

Sources of measurement uncertainty

The measurement uncertainty of bioradar-based vital sign detection is influenced by multiple factors, which can be grouped into four main categories.

First, the respiration rate is not constant. Its frequency and amplitude are variable over time. Additionally, the orientation of the body relative to the radar significantly affects the observable motion component.

When operating from a hovering UAV, small translations or rotations during the measurement alter the distance between radar and target, creating phase offsets indistinguishable from human motion unless corrected with additional hardware.

Material, geometry and properties of the debris site are highly unpredictable in practical scenarios and influence the measurement results. For materials with high electromagnetic transparency, e.g., wood, the attenuation or power loss due to the reflections of the radar waves are significantly weaker compared to materials with metal components, e.g., reinforced concrete. Any claim about the measurement uncertainty is thus only valid for a certain material composition, and for different compositions, the assessed uncertainties will differ.

Finally, many design choices of the radar hardware and processing affect the final measurement uncertainty. For example, the choice of operating frequency in the low GHz range represents a trade-off between penetration capability and system size. Lower frequencies exhibit reduced attenuation in common building materials and therefore improve the ability to detect signals through debris. However, the wavelength λ increases with decreasing frequency, requiring larger antennas, which is a critical constraint for UAV-based systems. In practice, frequencies in the range of 1 to 3 GHz provide a suitable compromise between penetration depth and system integration. Transmit power affects the signal-to-noise ratio and thus the minimum detectable motion, but at the same time, the available transmit power is limited both by the hardware and by regulatory and medical limits for maximum radiation exposure. The choice of signal processing algorithms influences frequency resolution via the time-frequency trade-off, as well as by their efficiency at extracting the respiratory rate. Additionally, measurement duration directly influences the achievable spectral precision.

A key parameter influencing all these contributions is the signal-to-noise ratio (SNR). Due to path loss, material attenuation, and imperfect reflections, the received signal amplitude decreases significantly with increasing distance and obstruction. At the same time, system noise and environmental clutter remain present, reducing the effective SNR. As a result, respiration-induced

modulations may become indistinguishable from noise, directly impacting both detection reliability and frequency estimation accuracy.

Previous research at the Department of Microsystems Engineering

In the preceding FOUNT2 [5] and SORTIE projects [6], the foundation for the current bioradar system was developed. In [7], a frequency-comb continuous-wave radar (FCCW) is introduced as a bioradar, enabling rapid software-defined radio (SDR) prototyping since no transmit-receive synchronization is required. In [8], this bioradar platform was extended with the use of prominent features in the FFT and CWT spectra of the radar signal. A linear support-vector machine was trained on these features to automatically detect survivors. However, this performance does not directly translate to real deployments, where a quantitative evaluation of the system is still to be conducted. Hence, ZeMeRGe shifts the focus from classification accuracy to measurement uncertainty quantification.

Simulation approaches

To enable systematic uncertainty analysis, simulation-based approaches are employed, allowing controlled variation of environmental and system parameters. In order to efficiently simulate the system, different approaches are possible. In [9], a computationally efficient framework of the vital sign detection radar with the finite-difference time-domain (FDTD) method for sparse frequency excitations is presented, building on the simulation of a beat period of the radar signal instead of a full period. Instead of simulating continuous motion, multiple static configurations corresponding to different chest positions are computed and combined to reconstruct the temporal signal. By using the FDTD method, it is possible to incorporate complex materials, e.g., by using Debye dispersion. However, it inherently neglects certain real-world effects, such as noise, and therefore provides insight into fundamental limits rather than complete system performance.

In addition, a simulation framework for mono-frequency radar systems such as in [10] provides valuable insight into the underlying mechanisms of vital sign detection. Although mono-frequency models cannot capture range-dependent effects or multi-target scenarios, they provide a controlled environment to investigate fundamental performance limits and sensitivities of radar-based vital sign detection. As such, they form an essential building block for understanding and validating more complex multi-frequency radar systems. This separation

of effects enables a structured decomposition of uncertainty sources. While mono-frequency simulations primarily capture physical and environmental contributions, multi-frequency simulations extend the analysis to include range-dependent effects and signal processing contributions. Together, these approaches provide a hierarchical framework for uncertainty analysis, ranging from fundamental physical limits to full-system behaviour.

Measurement with a stepped-frequency continuous-wave (SFCW) radar under laboratory conditions

To illustrate the discussed principles and their practical implications, the following section presents experimental results obtained with a stepped-frequency continuous-wave (SFCW) radar system under controlled laboratory conditions.

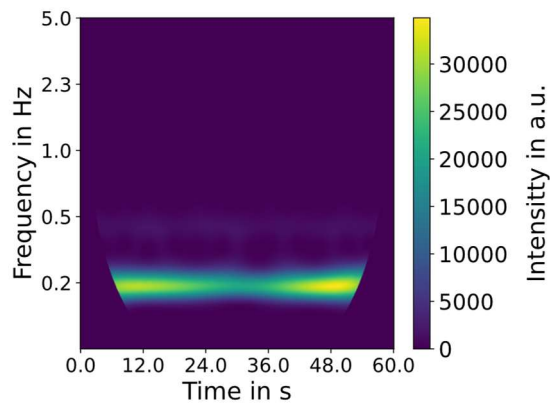


Fig. 3: SFCW-based respiration rate measurement with a controlled breathing dummy (mechanically moving plate).

Three examples are given in Fig. 3 to Fig. 5, using an SFCW radar operating at 2.4 to 2.5 GHz in an on-site RF testing chamber. This specific radar allows for the adjustment of the frequency steps, per-frequency time-domain sampling points, internal range-bin processing and adjustable transmit power and receiver attenuation. In the example, a test subject was hidden behind a pile of building debris of an approximate depth of 1 m consisting of ceramic tiles, with the radar scanning the subject in forward- and backward-facing positions; all other parameters remained unchanged. The spectrogram was obtained using a continuous wavelet transform.

In Fig. 3, a measurement was performed using a moving plate with a fixed rotational frequency, mimicking the respiratory chest excursion induced by survivors. Here, the “respiration rate” of $f_{\text{resp}} = 0.2$ Hz can be clearly seen in the

spectrogram for the entire duration of the measurement.

In Fig. 4 and Fig. 5, measurements were performed with an adult male test subject instead of the dummy. It can be observed that the orientation of the test subject significantly influences the detectability of the respiration signal. In the backward-facing position, the breathing frequency of the test subject is, while still recognizable, subject to a weaker SNR and thus harder to detect compared to the forward-facing measurement. In quantifying the measurement uncertainty, this is one of many relevant factors.

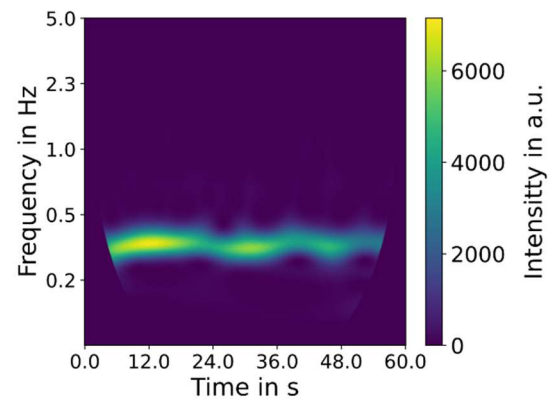


Fig. 4: SFCW-based respiration rate measurement for a forward-facing anteroposterior position.

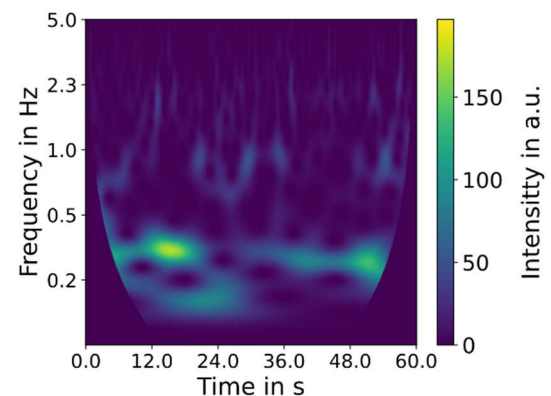


Fig. 5: SFCW-based respiration rate measurement for a backward-facing anteroposterior position.

Field measurement with FCCW radar

Additional experiments were carried out under field conditions using the FCCW radar system described in [7], as depicted in Fig. 6. The radar system was mounted on a DJI Matrice 400. A measurement time of 55 s was used. In the following experiments, the analysis is performed on individual range bins. In

particular, range bin 1 corresponds to the closest spatial interval (covering the range from 0 to 2.5 m), where the strongest reflections from nearby targets are expected. Therefore, a detectable respiration signal in range bin 1 indicates that the target is located in close proximity to the radar.



Fig. 6: Field tests with a DJI Matrice 400 drone, using the mechanically moving aluminium-coated plate as a breathing dummy. Here, bags with ceramic tiles are used as obstructions.

Fig. 7 depicts the FFT spectra for different range bins obtained from a measurement with a stationary UAV placed on the table. The breathing dummy respiration rate was set to $f_{\text{resp}} = 0.5$ Hz. This respiration rate is successfully detected, as evidenced by the prominent spectral peak in range bin 1, covering a range of approximately 0 to 2.5 m.

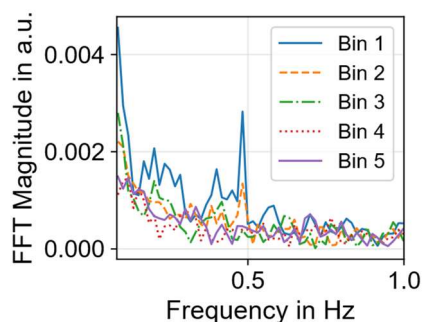


Fig. 7: FFT spectra of a measurement with the FCCW radar for a stationary UAV placed on the table with wooden/Styrofoam obstructions.

This breathing frequency becomes obscured once the UAV is in motion, as depicted in Fig. 8. Due to the UAV motion, the radar cannot clearly extract the breathing frequency from the echo signal. Without motion correction, the measurement uncertainty increases significantly, as motion-induced phase variations mask the respiration signal.

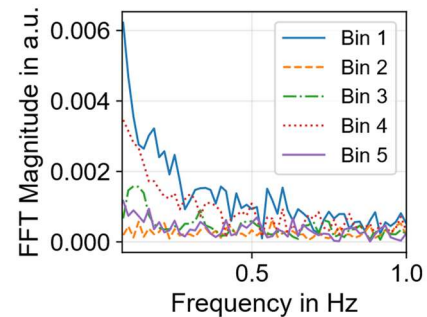


Fig. 8: FFT spectra of a measurement with the FCCW radar for a UAV at an altitude of 5 m above the table with ceramic tile bags as obstructions.

Expected impact

By producing quantitative uncertainty budgets, ZeMeRGe enables meaningful reliability statements (e.g., confidence that a detected respiration signature corresponds to a human), optimisation trade-offs for UAV operation (speed vs. accuracy vs. battery life), and safer and more efficient deployment for first responders. The project's results will make bioradar technology both more effective and scientifically verifiable, turning it into a trustworthy decision-support tool for disaster relief. In particular, the presented analysis highlights the need for motion compensation and scenario-dependent calibration to achieve reliable operation in real-world deployments.

Acknowledgment

This contribution was partially funded by the German research project SORTIE with support from the German Federal Ministry of Education and Research (BMBF) under grant number 13N15189 and by the German Research Foundation (DFG) under the project number 541028991 "ZeMeRGe".

The authors acknowledge support by the state of Baden-Württemberg through bwHPC.

Literature

- [1] *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*, 2008. doi: 10.59161/JCGM100-2008E.
- [2] G. Paterniani *et al.*, "Radar-Based Monitoring of Vital Signs: A Tutorial Overview," *Proceedings of the IEEE*, vol. 111, no. 3, pp. 277–317, Mar. 2023, doi: 10.1109/JPROC.2023.3244362.
- [3] Y. Rong, R. Gutierrez, K. V. Mishra, and D. W. Bliss, "Noncontact Vital Sign Detection With UAV-Borne Radars: An Overview of Recent Advances," *IEEE Vehicular Technology Magazine*, vol. 16, no. 3, pp. 118–128, Sep. 2021, doi: 10.1109/MVT.2021.3086442.
- [4] J. Badawy, O. K. Nguyen, C. Clark, E. A. Halm, and A. N. Makam, "Is everyone really breathing 20 times a minute? Assessing epidemiology and variation in recorded respiratory rate in hospitalised adults," *BMJ Qual Saf*, vol. 26, no. 10, pp. 832–836, Oct. 2017, doi: 10.1136/bmjqs-2017-006671.
- [5] G. Gidion, D. Shi, T. Aftab, D. Grgic, P. Tengse, and D. Vössing, "Verbundprojekt FOUNT2: Fliegendes Lokalisierungssystem für die Rettung und Bergung von Verschütteten : Abschlussbericht : 15.10.2016-14.10.2019," Albert-Ludwigs-Universität Freiburg, 2020. doi: 10.2314/KXP:1735491330.
- [6] G. Gidion *et al.*, "Verbundprojekt SORTIE: Sensor-Systeme zur Lokalisierung von verschütteten Personen in eingestürzten Gebäuden - Abschlussbericht," Institut für Mikrosystemtechnik (IMTEK), Lehrstuhl für Elektrische Mess- u. Prüfverfahren (EMP) and Institut für Informatik (IIF) - Lehrstuhl für Autonome Intelligente Systeme (AIS), 2023.
- [7] D. Shi, G. Gidion, T. Aftab, L. M. Reindl, and S. J. Rupitsch, "Frequency Comb-Based Ground-Penetrating Bioradar: System Implementation and Signal Processing," *Sensors*, vol. 23, no. 3, p. 1335, Jan. 2023, doi: 10.3390/s23031335.
- [8] D. Shi, G. Gidion, L. M. Reindl, and S. J. Rupitsch, "Automatic Life Detection Based on Efficient Features of Ground-Penetrating Rescue Radar Signals," *Sensors*, vol. 23, no. 15, p. 6771, Jul. 2023, doi: 10.3390/s23156771.
- [9] D. M. Spale, G. Gidion, T. Schaechtle, and S. J. Rupitsch, "Numerical Modeling of Radar-based Vital Sign Detection in Debris Sites for Sparse Frequency Excitations," in *2025 22nd European Radar Conference (EuRAD)*, Sep. 2025, pp. 299–302. doi: 10.23919/EuRAD65285.2025.11234190.
- [10] D. M. Spale, G. Gidion, E. Maier, K. Schmitt, and S. J. Rupitsch, "A Study on Clutter Rejection in Monochromatic Radar-based Breathing Activity Detection with Finite Difference Time Domain Simulations," in *2026 17th German Microwave Conference (GeMiC)*, Mar. 2026, pp. 200–203. doi: 10.1109/GeMiC71240.2026.11516338.

Real-Time Multi-Particle Tracking at 10 kHz: Leveraging Binary Frames for Event-Based Vision

Johannes Baßler¹; Paul Bäcker²; Michael Heizmann¹ 

¹Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT), Karlsruhe

²Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe

Abstract

While event cameras provide microsecond temporal resolution, purely asynchronous tracking algorithms remain strongly load-dependent, scaling poorly with both instantaneous event-rate bursts and the number of simultaneous particles. We present a deterministically scalable framework that converts this variable tracking workload into a bounded, GPU-parallel detection task by aggregating events into ultra-short 100 μ s binary frames. Validated in a physical free-fall sorting scenario, our architecture sustains a strict 10 kHz tracking loop on standard consumer hardware. It scales to 4,000 simultaneous trajectories within the 100 μ s deadline, achieving an effective $\mathcal{O}(1)$ core detection latency to guarantee the execution determinism required for industrial automation.

Keywords: Event-based vision, real-time multi-object tracking, binary frames, GPU acceleration, industrial automation

1 Introduction

The automated monitoring and sorting of rapidly moving objects – such as bulk materials traversing a free-fall section – poses severe challenges for classical machine vision. Standard area-scan cameras suffer from motion blur at exposure times required for adequate light collection, while operating them at ultra-high frame rates generates prohibitive volumes of redundant data. Neuromorphic event cameras overcome these limitations by reporting only asynchronous, pixel-level brightness changes. Their microsecond-scale temporal resolution and high dynamic range theoretically provide the perfect sensory foundation for high-speed tracking applications [5]. Despite their theoretical advantages, deploying event cameras in real-time industrial environments remains rare. The core bottleneck lies in the computational paradigm. Purely asynchronous tracking algorithms process the data stream sequentially on an event-by-event basis. Consequently, when dozens of objects simultaneously enter the field of view, the resulting burst of millions of events per second forces the data association complexity to explode [1, 2, 3]. This strict dependency on the visual scene dynamics makes guaranteeing a constant Worst-Case Execution Time (WCET) practically impossible on standard Von Neumann architectures [5].

To bypass this sequential bottleneck, recent literature frequently transitions back to synchronous processing by accumulating events into dense representations such as time surfaces or voxel grids. However, to extract meaningful geometric features for classical computer vision algorithms, these approaches often require extended accumulation windows spanning tens of milliseconds [6, 7]. In

highly dynamic scenarios, this prolonged aggregation reintroduces motion blur, effectively nullifying the defining temporal advantage of the event sensor. While natively asynchronous Spiking Neural Networks (SNNs) offer a promising alternative, they currently lack the mature inference hardware and unified toolchains necessary for reliable, low-latency integration into industrial pipelines [13, 12].

In this paper, we present a deterministic, frame-based architecture for real-time multi-particle tracking. Building upon the concept of efficient binary representations [11], we restrict the event accumulation window to an ultra-short duration of 100 μ s. Because this 10 kHz quantization threshold is comparable to the hardware's inherent timestamp jitter and refractory period [16, 17], virtually no physically relevant temporal information is lost. By leveraging a decoupled, multi-threaded architecture with zero-overhead GPU execution paradigms, our system transforms the unpredictable, event-rate-dependent tracking workload into a highly parallelizable task with effective $\mathcal{O}(1)$ latency scaling for the core detection. We demonstrate that robust, microsecond-scale multi-object tracking is achievable on standard consumer hardware, providing a viable blueprint for deploying event-based vision in strict real-time settings.

2 Related Work

Purely asynchronous tracking algorithms aim to preserve the extreme temporal resolution of the sensor by processing each incoming event individually as it occurs. Prominent methodologies in this domain include event-based clustering, such as the asynchronous mean-shift algorithm proposed by Barranco et al. [1], and the *AEB tracker* in-

troduced by Wang et al. [2]. The *AEB tracker*, for instance, associates individual events to spatio-temporal Gaussian models using nearest-neighbor classification, whereas feature-driven methods like *HASTE* (Alzugaray & Chli) [3] process single events via continuous multi-hypothesis evaluation. While this event-by-event paradigm minimizes latency for sparse scenes with a limited number of targets, it exhibits inherent scalability constraints. As detailed by Gallego et al. [5, Sec. 4.1 and 4.2], the purely asynchronous event-by-event processing paradigm limits parallelizability, as the computational complexity of data association intrinsically scales with the dynamically varying event rate. Alternatively, Spiking Neural Networks (SNNs) offer a natively asynchronous parallel processing paradigm [4]. However, simulating the irregular firing patterns of SNNs on standard GPUs causes severe control-flow divergence [12], while dedicated neuromorphic processors currently lack the unified compiler toolchains and hardware maturity required for strict industrial integration [13].

To overcome the limitations of sequential processing, a widely adopted alternative is to aggregate asynchronous events into synchronous, dense representations. Prominent examples include event frames, Time Surfaces [6], and spatio-temporal voxel grids [7]. This transformation mitigates the data association bottleneck by enabling the direct application of mature, highly parallelizable computer vision algorithms and standard tracking-by-detection pipelines. However, a critical trade-off emerges: to extract meaningful spatial features, these representations require a sufficient spatial density of events to form recognizable structures and continuous edges. To achieve this spatial completeness, conventional methods typically rely on extended accumulation windows – often ranging from 10 ms to 50 ms depending on the scene dynamics [8, 10, 7]. In high-speed applications, such extended temporal windows inevitably accumulate overlapping particle positions, reintroducing severe intra-frame motion blur and thereby compromising the μs -scale temporal precision of event-based sensors [9, 11]. At the same time, a second trade-off arises at the system level: as the effective frame rate increases, many existing tracking pipelines are no longer fast enough to process the generated data within the available real-time budget. Thus, beyond representational fidelity, practical high-speed event-based tracking is also constrained by the real-time barrier imposed by the computational throughput of the tracking algorithm itself [5]. Our approach bridges these paradigms by explicitly prioritizing execution determinism without sacrificing temporal fidelity. Building upon the concept of efficient binary frames [11], we utilize an ultra-short accumulation window of 100 μs . By processing these frames at

10 kHz, we transform the unpredictable workload of asynchronous tracking into a deterministically parallelizable GPU task, yielding a predictable worst-case execution time (WCET) governed by fixed per-frame operations rather than the instantaneous event rate.

3 System Architecture and Methodology

At a target processing frequency of 10 kHz, the temporal budget (100 μs) for data ingestion, memory transfers, and system state updates is limited. Implementing a sequential, single-threaded processing pipeline at this frequency is highly impractical. The accumulation of blocking sensor I/O operations, host-to-device memory transfer latencies, and CPU-side CUDA driver overheads – which typically consume 5 μs to 15 μs per individual kernel launch [14] – rapidly consumes a prohibitive fraction of the available 100 μs deadline. To physically realize this frame rate and maintain execution determinism without data loss, we propose a decoupled, multi-threaded architecture supported by zero-overhead GPU execution paradigms. An overview of the system architecture is illustrated in Fig. 1.

3.1 Asynchronous Multi-Threaded Pipeline

To achieve a constant WCET and prevent the highly variable event rate from stalling the system, the architecture is divided into three asynchronous stages communicating via lock-free ring buffers:

- **Sensor I/O (Thread A):** The camera callback continuously accumulates incoming events into a bit-packed memory structure (*FastBit-Buffer*). Exactly every 100 μs , the buffer is swapped and passed to the execution queue, establishing the deterministic frame rate.
- **Detection (Thread B):** This host-side thread controls the GPU detection pipeline, launching the dense detection and centroid refinement kernels and forwarding the resulting spatial coordinates to the tracking queue.
- **Tracking (Thread C):** Operating entirely asynchronously on the CPU, the tracker consumes the detection coordinates, updates the physical state of the particles, and manages the lifecycle of the trajectories.

This pipelined approach isolates the sensor I/O from the downstream algorithmic workload. Thread A still scales with the incoming event count, but its callback is reduced to a pre-allocated bit update per event, avoiding clustering, data association, and dynamic allocation. Because the constant factor of these operations is negligibly small, the thread effortlessly absorbs extreme event bursts without exceeding its fractional time budget. This lock-free decoupling ensures that downstream computations

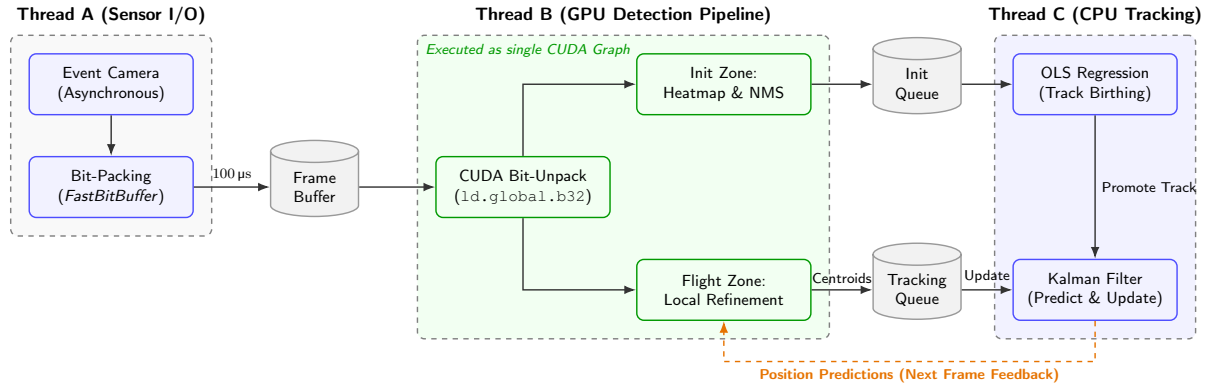


Figure 1: Overview of the decoupled, multi-threaded system architecture. The pipeline converts asynchronous events into deterministic 10 kHz binary frames (Thread A). The dense GPU detection and centroid refinement (Thread B) is fully encapsulated within a single CUDA Graph to eliminate kernel launch overhead. Thread C performs localized trajectory tracking by continuously feeding Kalman predictions back to the GPU, entirely bypassing global $\mathcal{O}(M \times N)$ data association common in conventional event-based tracking.

– such as updating the state of dozens of trajectories in Thread C – never exert backpressure on the camera SDK. The sensor interface is never blocked, guaranteeing a deterministic 10 kHz frame generation loop.

3.2 GPU Detection Pipeline

The object detection is performed solely on dense representations. However, transferring and processing full-resolution matrices at 10 kHz requires aggressive optimization.

Bit-Unpacking and Memory Access: To bypass PCIe bandwidth limitations, the frames are transferred to the GPU in their natively compressed, bit-packed format (1 bit per pixel). To minimize runtime overhead, we implement a custom CUDA unpacking kernel that reconstructs the binary frame directly in device memory. During centroid refinement, the local search window is scanned using 4-byte loads to reduce memory overhead. These loads are expressed via fixed-size `memcpy` operations rather than pointer casts, avoiding strict-aliasing violations while still allowing NVCC to emit a single 32-bit global memory load (`ld.global.b32`) instead of four separate 8-bit loads. This reduces the instruction count during local refinement and improves memory access efficiency.

Activity Heatmap and NMS: Following the frame generation, an activity heatmap is computed by aggregating spatial neighborhoods. To minimize computational overhead, this dense global search is not performed over the entire sensor plane. Instead, it is confined to a designated spatial initialization region used exclusively for spawning new trajectories (further detailed in Section 3.3). Within this

bounded region, we employ a highly parallel Non-Maximum Suppression (NMS) kernel to extract local maxima of event activity, followed by a sub-pixel centroid refinement. Because the matrix dimensions of this initialization zone are fixed, the computational cost of this dense operation remains $\mathcal{O}(1)$ relative to the scene’s activity and the incoming event rate.

CUDA Graph Execution: The most critical bottleneck at a 100 μ s time budget is the CPU-side CUDA driver overhead. A standard sequential dispatch of multiple dependent kernels (e.g., Unpack, Heatmap, NMS, and Refinement) along with their memory synchronizations incurs cumulative launch latencies and inter-kernel gaps that easily reach 20 μ s to 40 μ s, as the CPU struggles to keep the GPU queues fed [15]. At this scale, the dispatch overhead dominates the actual execution time, consuming nearly half the available frame budget before computation concludes. To eliminate this bottleneck, the entire detection sequence is pre-compiled into a *CUDA Graph*. This topology captures all kernel parameters and execution dependencies, allowing the entire pipeline to be dispatched with a single API call. By replacing the sequential driver interactions with a single graph launch, the dispatch overhead is reduced to a few microseconds.

3.3 Hybrid Trajectory Tracking

To maintain the strict 100 μ s execution cycle for dozens of concurrent particles, the CPU tracking module (Thread C) avoids standard global data association entirely. Instead, the system employs a tightly coupled, hybrid tracking architecture that conceptually divides the field of view. It applies rigorous multi-frame regression to securely initialize new targets, while established tracks in flight uti-

lize a continuous feedback loop between the CPU's Kalman filter predictions and the GPU's localized measurement extraction.

Robust Initialization via Spatial Regression: In industrial use cases like bulk material sorting, particles often follow a predictable primary motion vector, e.g., falling due to gravity, and enter the sensor's field of view at a specific boundary. Leveraging this *a priori* knowledge, we restrict track birthing to a narrow, designated spatial initialization zone spanning the entry edge (e.g., the top rows of the sensor, adapting to all edges if the entry vector is unknown). Detections in this zone are not immediately promoted to active tracks. Instead, they are aggregated into a temporal cache. A candidate is promoted to an active trajectory only if its recent history passes an Ordinary Least Squares (OLS) motion-consistency gate, requiring a minimum number of consecutive observations and a low residual Root Mean Square (RMS) error when fitting the cached detections as a linear trajectory over time. This physically constrained regression filters out spurious noise clusters without scanning the entire sensor plane for new targets.

Deterministic Flight Tracking: Once a track is established and leaves the initialization zone, global data association (e.g., Hungarian matching) is bypassed entirely. The success of this bypass relies directly on the high sampling rate. Because $\tau = 100 \mu\text{s}$ is sufficiently small relative to the particle's macroscopic velocity, the inter-frame spatial displacement of a particle's centroid is reduced to a fraction of a pixel. Instead of matching independent detections, the tracker utilizes a tightly coupled feedback loop. The CPU-based discrete Kalman filter predicts the particle's next position. These coordinates are directly fed into a secondary, localized GPU refinement kernel. By searching only within a highly constrained bounding box around the Kalman prediction, the GPU extracts the exact sub-pixel measurement and feeds it back to the CPU. This physically localized extraction handles even rapid non-linear trajectory deviations robustly and eliminates the $\mathcal{O}(M \times N)$ complexity of global association.

4 System Evaluation

This section evaluates the proposed architecture in a physical free-fall sorting setup, focusing on real-time determinism, per-thread latency, and scalability under both measured and synthetic loads.

4.1 Experimental Setup and Application Scope

To evaluate the real-time capabilities and hardware scalability of the proposed architecture, we consider

the highly dynamic use case of industrial bulk material sorting. The experimental setup consists of a vertical free-fall section where objects traverse the field of view primarily along the vertical sensor axis (denoted as the y -axis, with $y \in [0, 1279]$). To maximize the observable fall distance, the event camera is mounted rotated by 90° relative to its default landscape orientation, i.e., in a portrait-style configuration. Consequently, the sensor's long image dimension is aligned with gravity and defines the effective vertical tracking axis used throughout this section. The scene is uniformly illuminated by an LED backlight to ensure consistent contrast changes during the particles' flight.

Data acquisition is performed using an event camera (Prophesee EVK4 @ default settings) operating at its native 1280×720 resolution. The spatial initialization zone for spawning new trajectories is tightly constrained to a narrow horizontal strip spanning $y \in [200, 240]$ pixels. Once successfully initialized via the spatial regression gate, the particles' flight tracks are continuously refined until they cross the $y = 1000$ boundary. Beyond this exit zone, the particles enter the physical separation stage (e.g., a pneumatic valve array), rendering further tracking unnecessary for this application, although extended tracking would be computationally trivial.

The computational timing evaluation is based on over seven minutes of continuous event stream recordings acquired within a controlled physical laboratory setup. This extensive dataset encompasses a diverse spectrum of falling bulk materials, including small stones, wheat grains, coffee beans, fabric scraps, and dried tea leaves, generating thousands of unique object trajectories. These recordings are utilized exclusively to benchmark the system's hardware latencies, execution determinism, and scalability. While formal tracking accuracy metrics (e.g., identity switches or MOTA scores) fall outside the scope of this system architecture paper, the qualitative tracking robustness is effectively demonstrated. Even notoriously challenging targets exhibiting chaotic aerodynamics – such as irregularly shaped, fluttering dried tea leaves – are continuously tracked from entry to exit without fragmentation (see Fig. 5).

At the target 10 kHz frame rate ($\tau = 100 \mu\text{s}$), the average inter-frame displacement of a falling particle is approximately 0.5 pixels. This minimal inter-frame motion constitutes the foundational kinematic property that allows the system to bypass global data association entirely.

All computational benchmarks were executed on a standard consumer workstation equipped with an Intel Core i7-12700K CPU and an NVIDIA RTX 3060 GPU under a conventional Debian Linux environment. No real-time kernel, CPU isolation, or specialized hardware was used, so the reported latencies

reflect the determinism achieved by the software architecture itself.

4.2 Sensor I/O and Event Burst Immunity

The primary responsibility of Thread A is the continuous and deterministic ingestion of asynchronous events from the camera driver into the binary representations. The architectural details of this process are described in Section 3. Event-based sensors frequently produce temporal data bursts when dense object clusters traverse the field of view. To guarantee real-time operation, the pipeline must process these peak loads without inducing backpressure on the camera interface.

Figure 2 shows the callback-side processing latency of Thread A as a function of the incoming event rate, measured on the physical bulk-material datasets described above. Across the observed range up to approximately 25.6 Mevents/s, the CPU event-ingestion and bit-buffer update time increases only mildly and remains well below the 100 μ s frame deadline. The median latency rises from about 4 μ s at low activity to about 12 μ s in high-support bins. The 99th percentile reaches 27.0 μ s for well-supported bins and up to 36.4 μ s in rare cases. This metric excludes the separate frame-swap step and the asynchronous PCIe transfer (≈ 22 μ s), which do not compromise the pipeline's timing constraints. Overall, the measurements indicate substantial headroom before Thread A approaches the real-time limit.

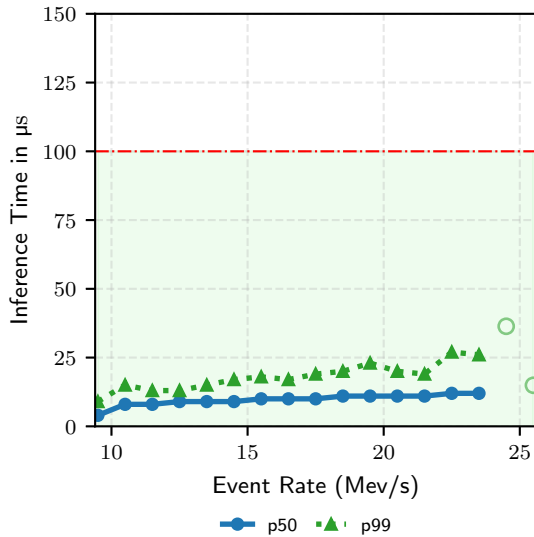


Figure 2: Callback processing latency of Thread A ($t_{a,us}$) versus incoming event rate, aggregated into 1 Mevents/s bins with median and 99th percentile latencies. The horizontal line marks the 100 μ s real-time deadline; circles denote low-support bins (< 100 frames) and their corresponding maximum values.

4.3 Compute Overheads and Tracking Scalability (Thread B)

Thread B controls the GPU-side detection pipeline, including activity-based detection and centroid calculation. The resulting particle sub-pixel coordinates are subsequently propagated to Thread C for the actual Kalman filter updates. Since the GPU kernels themselves execute within a small fraction of the available time budget, the measured end-to-end latency of Thread B is primarily determined by host-side orchestration overheads, including operating-system scheduling and CPU synchronization.

The empirical distribution of the end-to-end execution time for Thread B demonstrates this characteristic distinctly. The median total latency is 79.0 μ s, with the 99th percentile bounded at exactly 100.0 μ s (see Fig. 4). However, when isolating the raw GPU execution time for the detection and refinement branches from the total latency, the tracking computation accounts for merely 26.7 % of the time. The remaining 73.3 %, averaging 59.0 μ s per frame, is entirely consumed by operating system jitter, memory mapped I/O, and thread synchronization overheads.

To buffer against these non-deterministic scheduling spikes, the architecture uses a deep host-side ring buffer between Thread A (producer) and Thread B (inference worker), while Thread B processes the newest available frame to avoid backpressure. Analysis of the raw execution logs reveals that extreme system interrupts occasionally exceed the permissible timing constraints, forcing a frame drop. This scenario occurs in approximately one out of every 4,300 evaluated physical frames, equating to a drop rate of roughly 0.23 ‰. Because the Kalman filter algorithm in Thread C is inherently designed to predict trajectories across missing measurements, these isolated drops do not compromise the tracking integrity or the real-time reliability of the sorting process. Nevertheless, deploying the system on a real-time patched kernel with isolated CPU cores could theoretically eliminate these rare interruptions entirely. For the synthetic scalability analysis, we assumed $N_{init} = 0.1 N_{flight}$ in order to preserve a realistic workload ratio between the initialization-detection branch and the flight-refinement branch. This choice is motivated by the sorting geometry: only a small subset of particles is expected to occupy the narrow initialization zone at any given time, whereas most visible particles have already transitioned to active flight tracking. The synthetic regime is therefore intended as a scalability stress test rather than a physically realistic sorting density.

The computational headroom of the tracking implementation becomes apparent under synthetic load testing. While the physical sorting scenarios exhibit a maximum of approximately 70 simultaneous parti-

cles, the synthetic benchmark evaluated the pipeline up to 10,000 parallel flight trajectories. The end-to-end latency increases approximately linearly at a rate of roughly $0.006\ \mu\text{s}$ per particle. The system successfully maintained the $100\ \mu\text{s}$ real-time deadline for up to 4,000 simultaneously tracked objects, a volume that vastly exceeds the physical limits of the pneumatic ejection array. This is illustrated in Fig. 3.

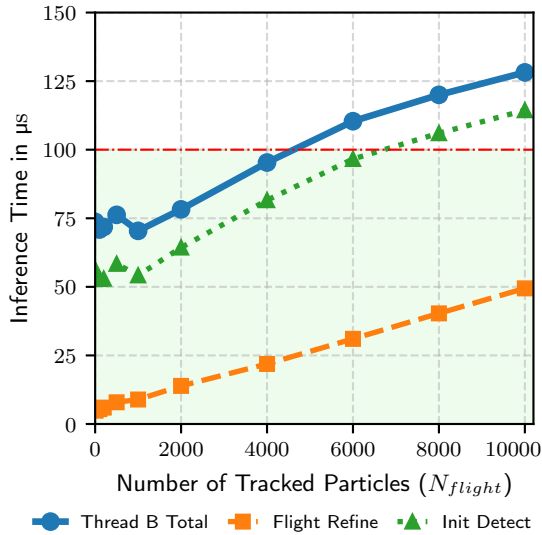


Figure 3: Mean inference time of Thread B for the physical measurements and the synthetic scalability benchmark. The curves show total Thread B latency, flight-track refinement, and initialization detection. In the synthetic benchmark, $N_{init} = 0.1 N_{flight}$ was assumed. The horizontal line marks the $100\ \mu\text{s}$ real-time deadline.

4.4 Downstream Processing and System Modularity

Thread C receives the tracking coordinates from Thread B, where the asynchronous event stream has been fully reduced to dense geometric information. The empirical execution latency of Thread C (see Fig. 4) confirms that trajectory management operates within the real-time constraints, with rare OS scheduling spikes occasionally exceeding the $100\ \mu\text{s}$ deadline without compromising overall system performance. For the evaluated sorting task, a discrete Kalman filter predicts and maintains particle trajectories from this structured coordinate data.

The optimization of this tracking backend extends beyond the scope of this work. The system architecture is designed with modularity in mind. The Kalman filter module can be substituted with alternative downstream algorithms tailored to different use cases. This decoupled design allows for flexible application logic integration without affecting the deterministic event ingestion front-end.

4.5 Overall System Performance and Algorithmic Scalability

The fundamental advantage of the proposed architecture is the decoupling of computational load from visual scene dynamics: by converting the asynchronous stream into fixed-size binary frames, the core detection achieves an effective $\mathcal{O}(1)$ runtime. While Thread A still scales marginally with the raw event volume and flight refinement scales linearly with the number of active tracks, these slopes are sufficiently flat to maintain robust real-time operation.

Figure 4 illustrates the combined latency distributions of the three processing stages, confirming that the end-to-end processing time remains bounded below the real-time deadline with rare OS jitter exceptions.

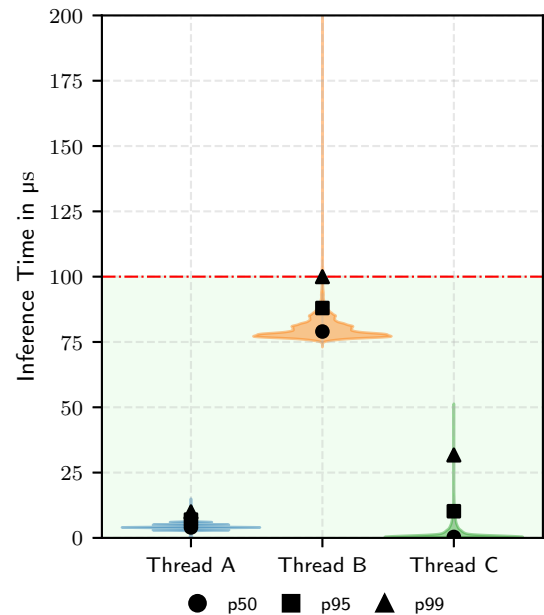


Figure 4: Per-frame inference-time distributions of Threads A, B, and C. Internal markers denote the median, 95th percentile, and 99th percentile, and the horizontal line marks the $100\ \mu\text{s}$ real-time deadline.

4.6 Temporal Fidelity and Information Retention

Despite transitioning from an asynchronous event stream to a synchronous processing pipeline, the proposed architecture discards virtually no relevant sensor information under the evaluated camera configuration. The $100\ \mu\text{s}$ accumulation window was chosen to be short relative to the effective per-pixel event rate observed at the vendor default settings. Neuromorphic sensors exhibit a bias-dependent refractory period, i.e., a recovery time before an individual pixel can register a subsequent brightness change. While aggressive bias settings may reduce this period below $100\ \mu\text{s}$ [16, 17], repeated events

from the same pixel within one accumulation window were empirically negligible in our evaluation: only 0.04‰ of all events were lost due to the binary representation.

Furthermore, while modern event cameras assign timestamps with nominal microsecond resolution, the true temporal accuracy is bounded by analog comparator delays and digital bus arbitration jitter. Recent hardware characterizations demonstrate that this inherent timestamp uncertainty consistently ranges from tens to hundreds of microseconds [17]. Therefore, precision beyond the 100 μ s quantization threshold is dominated by sensor noise rather than actual scene dynamics.

4.7 Qualitative Tracking Evaluation

While the preceding sections extensively quantify the real-time processing capabilities of the architecture, the utility of the system inherently depends on its tracking integrity. The application domain of industrial bulk sorting constitutes a highly controlled environment characterized by a static background and consistent object illumination. In addition, the sensor is deliberately mounted in a portrait-style orientation so that its long image axis follows the free-fall direction, maximizing the observable trajectory length within a single view. Consequently, the tracking module is intentionally specialized for these conditions.

Within this operational scope, the algorithmic pipeline demonstrates robust tracking capabilities despite complex object dynamics. Figure 5 visualizes the reconstructed flight paths of falling, dried tea leaves captured by the event camera. Due to their irregular geometry and aerodynamic properties, these particles exhibit chaotic rotational and translational movements during their free fall. Nevertheless, the combination of binary centroid detection and Kalman filtering successfully isolates and continuously tracks up to 18 simultaneous particles within this specific recording. The visual assessment confirms that the generated trajectories are smooth, continuous, and resistant to trajectory fragmentation. This reliable spatial representation provides the prerequisite information required for accurate downstream pneumatic ejection commands.

5 Conclusion

This paper demonstrates that event-based multi-particle tracking can be executed deterministically at 10 kHz on standard consumer hardware by converting asynchronous event streams into ultra-short binary frames. The proposed multi-threaded architecture separates event ingestion, GPU-side detection, and trajectory management, thereby preventing instantaneous event-rate bursts from directly propagating into the core detection workload. In the evaluated sorting scenario, the system maintains the 100 μ s processing deadline and scales to

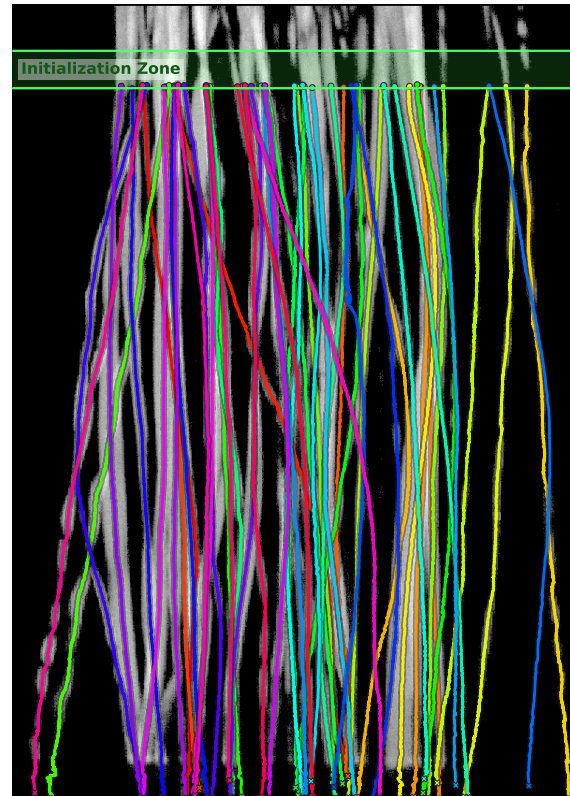


Figure 5: Composite trajectory map over time-surface-filtered event activity. Each trajectory is color-coded; the highlighted *Initialization Zone* marks the constrained region where new tracks are seeded and OLS-validated before promotion to active flight tracking. This plot covers a 7 s spanning recording of falling dried tea leaves, demonstrating the system’s ability to maintain continuous tracks despite complex aerodynamics.

4,000 simultaneously tracked particles in the synthetic stress test.

The evaluation confirms that the main remaining bottleneck is no longer the GPU-side detection computation itself, but host-side scheduling and synchronization overhead. At the same time, the current implementation is intentionally specialized for controlled optical sorting scenes, where static backgrounds, consistent illumination, and a known entry region enable simplified initialization and local refinement. Quantitative tracking-accuracy comparisons against asynchronous tracking algorithms are therefore left to future work.

Future work will focus on reducing operating-system jitter through real-time kernels and isolated CPU cores, and on extending the architecture to less constrained environments through robust background suppression and dynamic data association.

References

- [1] F. Barranco, C. Fermüller, and E. Ros, "Real-Time Clustering and Multi-Target Tracking Using Event-Based Sensors," *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018, pp. 5764–5769, doi: 10.1109/IROS.2018.8593380.
- [2] Z. Wang, T. Molloy, P. van Goor, and R. Mahony, "Asynchronous Blob Tracker for Event Cameras," *IEEE Transactions on Robotics*, vol. 40, pp. 4750–4767, 2024, doi: 10.1109/TRO.2024.3454410.
- [3] I. Alzugaray and M. Chli, "HASTE: multi-Hypothesis Asynchronous Speeded-up Tracking of Events," in *31st British Machine Vision Conference 2020, BMVC 2020*, Virtual Event, UK, September 7–10, 2020. BMVA Press, 2020.
- [4] S. Nagaraj et al., "DOTIE: Spiking Object Detection for Event Cameras," *arXiv preprint arXiv:2208.06648*, 2022. [Online]. Available: <https://arxiv.org/abs/2208.06648>
- [5] G. Gallego et al., "Event-based Vision: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, vol. 44, no. 1, pp. 154–180, Jan. 2022, doi: 10.1109/TPAMI.2020.3008413.
- [6] X. Lagorce, G. Orchard, F. Galluppi, B. E. Shi, and R. B. Benosman, "HOTS: A hierarchy of event-based time-surfaces for pattern recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 7, pp. 1346–1359, 2017, doi: 10.1109/TPAMI.2016.2574707.
- [7] D. Gehrig, A. Loquercio, K. G. Derpanis, and D. Scaramuzza, "End-to-End Learning of Representations for Asynchronous Event-Based Data," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, South Korea, 2019, pp. 5633–5643.
- [8] E. Perot, P. de Tournemire, D. Nitti, J. Masci, and A. Sironi, "Learning to Detect Objects with a 1 Megapixel Event Camera," in *Advances in Neural Information Processing Systems (NeurIPS)*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 16639–16652.
- [9] N. Messikommer, C. Fang, M. Gehrig, and D. Scaramuzza, "Data-Driven Feature Tracking for Event Cameras," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, BC, Canada, 2023, pp. 5642–5651.
- [10] A. Z. Zhu, L. Yuan, K. Chaney, and K. Daniilidis, "EV-FlowNet: Self-Supervised Optical Flow Estimation for Event-based Cameras," in *Robotics: Science and Systems (RSS)*, Pittsburgh, PA, USA, 2018.
- [11] J. Baßler and M. Heizmann, "Beyond the Sparsity Paradigm: Efficient Binary Frames for Event-Based Vision," *Proceedings of SPIE Optical Metrology*, Munich, Germany, 2025. [Online]. Available: <https://publikationen.bibliothek.kit.edu/1000188440>
- [12] M. Davies et al., "Advancing Neuromorphic Computing With Loihi: A Survey of Results and Outlook," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 911–934, 2021, doi: 10.1109/JPROC.2021.3067593.
- [13] C. D. Schuman et al., "Opportunities for Neuromorphic Computing Algorithms and Applications," *Nature Computational Science*, vol. 2, no. 1, pp. 10–19, 2022, doi: 10.1038/s43588-021-00184-y.
- [14] NVIDIA Corporation, *CUDA C++ Best Practices Guide*, Version 12.8, 2024. [Online]. Available: <https://docs.nvidia.com/cuda/cuda-c-best-practices-guide/>
- [15] A. Kerr et al., "Getting Started with CUDA Graphs," *NVIDIA Technical Blog*, Sep. 2019. [Online]. Available: <https://developer.nvidia.com/blog/cuda-graphs/>
- [16] S. McMahon-Crabtree et al., "Event-based camera refractory period characterization and initial clock drift measurement," in *Applications of Machine Learning 2023*, vol. 12693. SPIE, 2023, p. 126930V, doi: 10.1117/12.2672301.
- [17] A. Manakov, H. Hermann and B. Jähne, "Characterisation metrics for performance comparison of area scan and event-based sensors," *tm - Technisches Messen*, De Gruyter, 2025, doi: 10.1515/teme-2025-0018.

Characterization of Event-Camera Biases and Operating Conditions for High-Speed Sorting

Paul Bäcker^{1,2}; Johannes Baßler²; Georg Maier¹; Thomas Längle¹; Michael Heizmann² 

¹Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe

²Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIT), Karlsruhe

Abstract

Predictive multi-object tracking offers significant potential to improve separation accuracy in sensor-based sorting (SBS) systems. However, conventional frame-based vision pipelines are limited by exposure time, frame rate, and end-to-end latency, which reduces localization precision and actuation performance at high material throughputs. Event cameras offer an attractive alternative due to their asynchronous operation, low latency, and high temporal resolution. Yet, their performance strongly depends on sensor parameters (biases), illumination, and optical setup, and selecting suitable operating points remains a challenging multi-objective problem. In this work, we present a task-driven framework for characterizing event-camera operating points for high-speed sorting. Using a controlled rolling-sphere experiment with known geometry and an LED-based timing reference, we evaluate the influence of bias settings, illumination, and aperture on task-relevant metrics including geometric consistency, event statistics, background activity, hot-pixel behavior, and trigger-aligned latency. We validate the stability of the proposed metrics through repetition experiments and use one-dimensional parameter sweeps to reveal distinct sensitivities and trade-offs. Based on the results, we provide guidelines for selecting suitable operating points. The presented methodology provides a basis for principled manual tuning and future automated operating-point selection in event-based sorting and related low-latency vision applications.

Keywords: event cameras, bias tuning, task-driven evaluation, latency, event statistics, geometric consistency

1 Introduction

Sensor-based sorting (SBS) is widely used in industrial domains such as recycling and food processing [1]. To improve separation accuracy, predictive multi-object tracking has been proposed to compensate uncertain object motion between detection and actuation [2]. Classical frame-based vision pipelines are however limited by exposure time, frame rate, and end-to-end latency, which can degrade localization and actuation accuracy at high speeds. Event-based vision offers an alternative by reporting per-pixel brightness changes asynchronously with low latency and high temporal resolution [3, 7]. However, the event-based sensing paradigm also introduces specific challenges, that can diminish the benefits of low latency and high temporal resolution if not properly addressed:

The event rate is inherently scene dependent, with larger numbers of objects creating more events. Most asynchronous event-based tracking methods perform per-object event assignment resulting in undesirable scaling of computational complexity [8]. Even if the computational load can be handled downstream, the event-based sensor and transmission pipeline is limited by **maximum data throughput**, which can be exceeded under strong global mo-

tion and high event rates [11].

Latency is defined as the temporal difference between the time when an intensity change happens in the scene, and when it is detected by the sensor. It is dependent on the sensor operating point and illumination levels [13]. Ideally, this should be as low as possible to extract accurate object positions for precise actuation.

Shape fidelity - We want events precisely on the object boundaries in order to capture the true shape. This includes statistical error (jitter) and systematic error (varying latency between edges), which can distort the object shape and make it harder to maintain correspondences across tracking steps.

We address those challenges by investigating sensor settings in combination with different illumination and aperture conditions. We present a controlled benchmark for event cameras using a known shape prior and an LED-based latency measurement. We introduce a task-driven metric set that combines geometric, temporal, and event-statistical quality measures, and we perform bias sweeps to reveal distinct sensitivities, trade-offs, and practical operating regions across hardware settings and sensor parameters (biases). Finally, we show how these metrics support operating-point selection and can serve as a basis for later automated tuning.

This work was supported by the German Federal Ministry of Research, Technology and Space - Grant number 03DPS1146A

2 Related Work

Event cameras asynchronously report brightness changes once the change in log-intensity exceeds a polarity-dependent threshold. Foundational work on event cameras describes the pixel-level signal formation process and the asynchronous threshold-crossing principle underlying dynamic vision sensors [3]. Later characterization work measures how observable quantities such as contrast sensitivity, temporal behavior, and latency depend on the sensor and illumination conditions [12, 13].

Multiple methods address the challenges of event-based vision for tracking applications on an application level. For handling the high number of events and global motion, frame-based processing of event-data has been proposed [10, 5]. Further, event-filters can reduce the number of events, based on burst- or background suppression [5], or using a-priori global motion assumptions [9]. Although these methods reduce the computational load of the tracking algorithm, they do not change the amount of events generated at the sensor level, leading to potentially excessive data transfer and processing load, which can result in the loss of events [11].

Other approaches have altered the sensor design, enabling motion specific filtering at the sensor level [4, 6]. However, these methods have, to the best of our knowledge, not been transferred into commercially available sensors, and therefore do not yet offer a practical solution for the SBS setting. Our approach in this work is instead to optimize the sensing conditions, namely illumination, aperture, and sensor parameters. These tunable sensor parameters correspond to bias currents in the sensor pixel circuitry and are therefore referred to as **biases**.

Prior work on bias tuning ranges from empirical characterization and qualitative bias interpretation to benchmark-driven and task-driven optimization. Experimental studies investigate how bias settings affect measurable sensor properties and downstream signal quality [12, 14]. Generic benchmarking work proposes reusable tasks and metrics across different applications [15]. Other approaches adapt biases online through feedback control or black-box optimization for specific downstream tasks such as detection, place recognition, or face tracking [16, 17]. These works show that tuning is important and often task-dependent, but they do not take environmental factors and aperture settings into account. Further, we extend the task-driven evaluation by including geometric consistency and event statistics in addition to latency.

3 Background

The event-generation process is influenced by both the optical setup and several sensor biases. The optical setup determines the sharpness of the image projected onto the sensor. A small aperture in-

creases depth of field and yields sharper edges, but reduces signal strength; a large aperture increases signal strength at the cost of reduced depth of field and increased blur. Blurred edges spread intensity changes over time, which leads to uncontrollable jitter in the timing of the first event, as the event-pixel state (e.g. the current reference intensity) is to be assumed unknown. Further, inherent inhomogeneities in terms of contrast threshold, which have been reported to be as high as 30 % [3] lead to timing offsets, if the object edge is not sufficiently sharp. The pixel response can be approximated by cascaded temporal filters whose bandwidth depends on illumination and bias settings [3]. At low illumination, the photoreceptor response is slower and more strongly limited by photon flux, while at high illumination the bandwidth of the second-order low-pass filter of the feedback amplifier circuit becomes more relevant. These filters affect latency, noise sensitivity, and overshoot in the temporal response.

The biases control event generation: increasing `bias_diff_on` or `bias_diff_off` raises the ON or OFF contrast threshold, respectively, requiring stronger intensity changes to trigger an event. Increasing `bias_fo` increases high-frequency bandwidth and can reduce latency at the cost of noise; decreasing `bias_hpf` increases sensitivity to slower intensity changes.

To allow for pixel stabilization after an event, pixels are idle for a fixed time after it was triggered, ignoring any intensity changes. This time can be controlled by adjusting the `bias_refr`, where increasing `bias_refr` shortens the refractory period between successive events.

Together, optics and biases determine the trade-off between event rate, latency, noise, and geometric fidelity, motivating the task-driven characterization proposed in this work.

4 Experimental Setup

The experimental setup is shown in Fig. 1. Wooden spheres roll down a chute through the field of view of a fixed event camera, while a diffuse, uniform background illumination provides the dominant intensity contrast at the sphere boundary. From the measured traversal time of around 0.12 s for the 21 cm vertical extent, the object speed is approximately 1.7 m s^{-1} , i.e., about 5853 px s^{-1} (5.8 px ms^{-1}) in image space. Spheres are used because their silhouette remains circular for all viewing angles; deviations from circularity in the reconstructed event contour can therefore be attributed to sensing effects rather than object geometry.

We use an IDS uEye XCP-E with a Sony IMX636 event sensor (resolution 1280×720 pixels). To study the interaction between optics and sensor operating point, we vary background illumination (high/low) and aperture ($f/1.4$, $f/6$, $f/16$). For each configuration, irradiance at the sensor was recorded using the

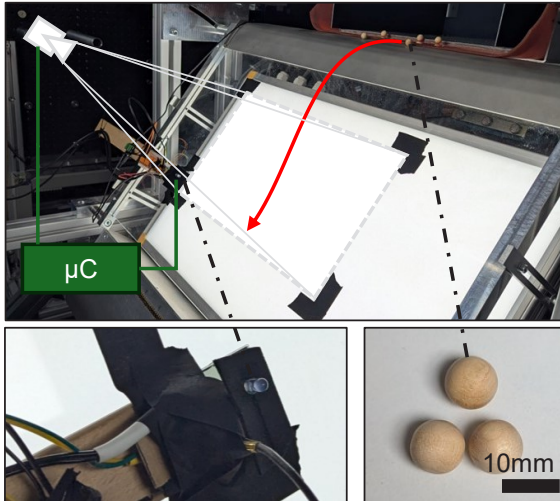


Figure 1: Experimental setup with chute, event camera, illumination, and LED trigger

Table 1: Measured sensor illumination (S. Illum.) for different background illumination and aperture settings.

ID	BG illum.	Aperture	S. Illum. (lux)
High-f/1.4	high	f/1.4	11185.27
Low-f/1.4	low	f/1.4	3026.36
High-f/6	high	f/6	894.59
Low-f/6	low	f/6	236.09
High-f/16	high	f/16	95.00
Low-f/16	low	f/16	25.00

on-sensor photodiode (Tab. 1). We identified five hot pixels at default biases, that we mask out for all recordings to avoid their influence on the metrics. A timing reference is provided by a LED that is visible in the camera image and is simultaneously connected to the camera trigger input (Fig. 1). This yields a shared external reference for the onset of an intensity change and enables direct estimation of event latency, which is later evaluated via the trigger-aligned delay distributions (Fig. 5).

5 Metric Extraction Pipeline

Fig. 2 shows the full evaluation pipeline and the per-trajectory metric extraction with two example outputs for different bias settings. The goal is to extract a set of meaningful, independent scalar metrics for each recording that can be used for bias-sweep interpretation and operating-point selection. All metrics are interpreted as error quantities, such that lower values indicate better performance. For each bias setting, we record five spheres rolling down the chute, yielding an event-cube and timestamped trigger signals (actual timing of LED activation). All events in a region of interest (ROI) around the LED are extracted and used exclusively for the **Trigger Analysis** branch, which estimates the mean ON

and OFF latency relative to the trigger signal. We automatically find the center of the LED using a heatmap of the global event activity. We consider a 5×5 pixel region around the LED center for the analysis. For each of these pixels, we extract the first positive event after each detected ON-trigger and the first negative event after each detected OFF-trigger. We repeatedly switch the LED on- and off every 10 ms. The figure shows a histogram of the extracted latencies for the two example settings, revealing a clear ON/OFF asymmetry for the setting in the top row, and a more symmetric behavior as well as much lower latencies for the setting in the bottom row. As a numerical per-sample metric $\text{Latency}_{\text{pos/neg}}$, we compute the means over all measured latencies.

The remaining events are assigned to individual objects using the Metavision Spatter **Tracker** [18]. For each detected object, the tracker yields a sequence of timestamped bounding boxes together with the corresponding object-center coordinates. These trajectory samples are converted into a structured representation and smoothed by fitting splines to the horizontal and vertical center position over time. Event assignment is then not performed by a single static crop, but by a moving ROI that follows the tracked object. Concretely, for each interval between two tracker timestamps, we retain only the events inside a bounding box enlarged by 30% around the tracked object. This produces trajectory-specific event sets that serve as the basis for the later per-object metrics.

For the **Events per Object** analysis, we sample the tracked centerline along the vertical image coordinate and identify, for each sampled row, the pixel that lies on the spline-predicted object center. At these trajectory pixels, we count the number of negative and positive events separately. This yields a compact per-trajectory measure of how often the sensor fires while the object center passes a given image location and deliberately suppresses edge-related effects away from the centerline. Across trajectories, we then aggregate the mean number of positive and negative event counts per pixel to extract the $\# \text{Events}_{\text{pos/neg}}$ metric.

In the geometric analysis branch, the per-trajectory events are first filtered to keep only the first negative and last positive event per pixel (**First/Last Filter**). This is based on the strong assumption, that the background is much brighter than the objects, so a pixel will transition from bright to dark to bright again as the object passes, leading to a negative event at the entry edge and a positive event at the exit edge. We want this analysis to be orthogonal to the event statistics, so we deliberately focus on the first and last events, which are more likely to correspond to the physical entry and exit edges of the sphere, rather than texture-related events that can

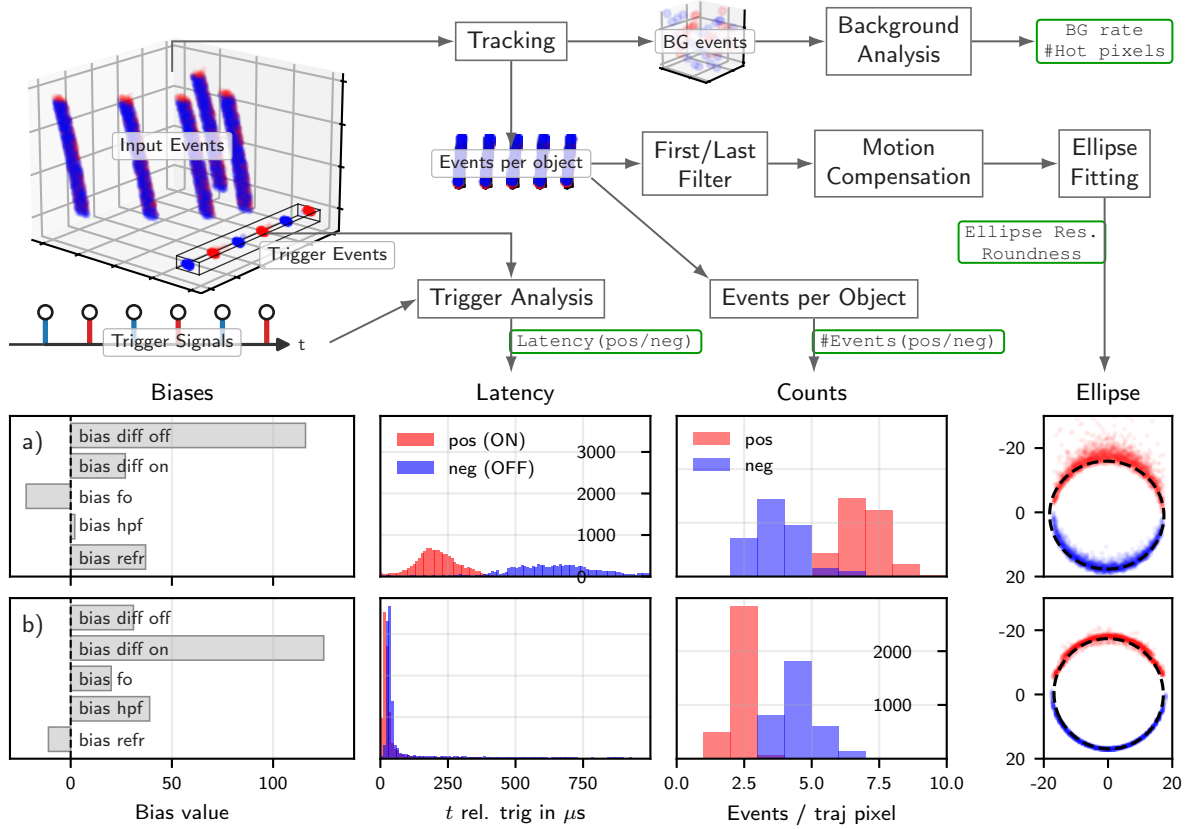


Figure 2: Full evaluation pipeline and per-trajectory metric extraction. The figure illustrates two representative bias settings a) (randomly selected) and b) (empirically tuned). The green boxes show the extracted, single scalar metrics per measurement.

occur inside the object region. For this, the implementation determines the first/last events per pixel, and ignores all other events. Then, for three evenly spaced reference positions at 25%, 50%, and 75% of each trajectory length, a **Motion Compensation** step is applied. For geometric analysis, the temporal window around a certain time reference is crucial, as larger windows introduce motion blur, but too small windows can lead to too few events for a reliable analysis. We therefore use large time windows, but compensate for the motion of the object within these windows to avoid blur. Around each reference time, we extract a temporal event window (10 ms) and fit a second-order polynomial to the event coordinates as a function of time. The resulting offsets are subtracted from the raw event positions so that the correlation between position and timestamp is minimized and the event cloud collapses toward a common spatial frame. This low-order model is consistent with the smooth, constant acceleration motion expected for spheres rolling down the chute. Finally, an **Ellipse Fit** is applied to the aligned event cloud, which yields a fitted ellipse. For a perfect circular object with perfectly aligned events, the fitted

ellipse, major (MA) and minor (ma) axes should be equal. By taking the mean over the inverse roundness ($1 - (MA/ma)$) for each of the $3 \cdot 5$ fitted ellipses we extract the `Roundness` metric for each measurement. The ellipse fit also yields a residual, which measures the average distance of the aligned events to the fitted ellipse. This residual is a measure of the spatial precision of the aligned event cloud and can be sensitive to blur or jitter in the event generation process. The mean over the residuals for all three fitted ellipses constitutes the `Ellipse Res.` metric.

Events without any present objects in the scene are used for **Background Analysis**. All events in these windows are accumulated into a per-pixel count map. Pixels with abnormally high activity are classified as hot pixels (`#Hot pixels` metric) and excluded from the final rate estimate. These are pixels whose activity exceeds ten times the median per-pixel rate and that also accumulate more than 50 events within the background window. The remaining background events are used for estimating the background event rate `BG rate` by dividing the total number of remaining events by the number of

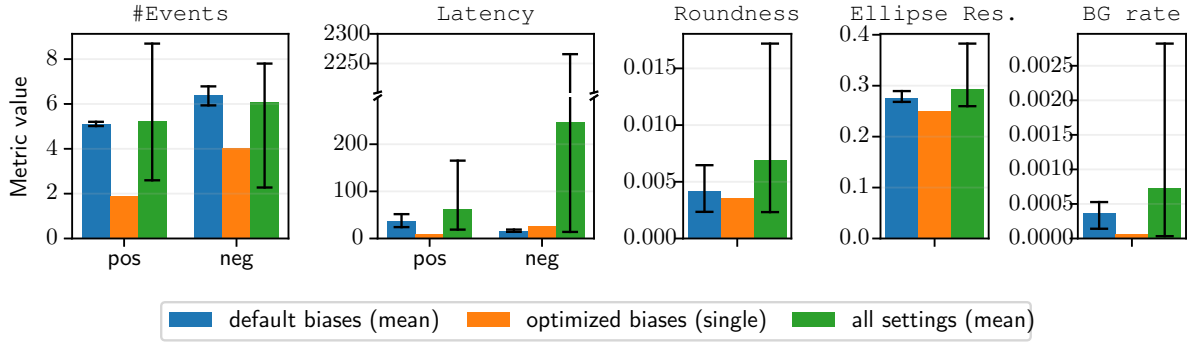


Figure 3: Metric fidelity for the High-f/6 hardware setting. For the default biases, minimum and maximum values of the 5 repeated experiments are shown as whiskers. For the variation across all bias settings, 5- and 95% quantile are shown. As a further reference, metric values for the empirically tuned operating point are also shown.

pixels and the total time across all timespans.

6 Bias Sweep Analysis

For each hardware configuration (illumination and aperture), we independently sweep each sensor bias while keeping all remaining biases at their default values. For every measurement, we extract the full metric set described in Sec. 5. We assess metric robustness, compare hardware regimes, and identify dominant sensitivities and trade-offs of individual parameters.

6.1 Metric Fidelity

Our measurement setup is subject to both random and systematic disturbances, including variations in rolling-sphere trajectories, illumination fluctuations, and sensor fixed-pattern noise. These disturbances are not a limitation of the proposed method, but reflect the variability expected in an online characterization and tuning scenario. Consequently, the proposed metrics must be robust enough to reveal meaningful and repeatable trends under realistic operating conditions, while remaining sensitive to actual changes in sensor settings.

To assess metric robustness, we repeat the identical experiment—five objects recorded at default bias settings—five times in succession and compare the resulting variation to the range observed across all bias settings. Fig. 3 shows the variation across repeated measurements relative to the variation induced by the bias sweeps.

The `#Events` (`pos/neg`) metric exhibits relatively low noise, with a clear separation between the default setting, the selected operating point, and the overall range across bias settings. The `Latency` (`pos/neg`) metrics show moderate variation, but the default setting remains clearly separated from the selected operating point. Moreover, the absolute magnitude of the noise is small, on the order of tens of μs , compared to the full range across bias settings.

The `Roundness` score exhibits the highest relative variation. Nevertheless, it remains useful for detecting large degradations in geometric consistency, which occur for a substantial fraction of the tested bias settings. The `Ellipse residual` shows very little variation across repeated measurements. No hot pixels were observed under the default bias settings; consequently, the `#Hot pixels` metric shows no variation across repeated measurements. The `BG rate` exhibits comparatively strong variation, likely due to changes in the static background scene and ambient illumination. Despite this, the observed differences remain significant relative to the selected operating point.

6.2 Hardware-Level Comparison

We compare different hardware regimes (background illumination and aperture) at a global level. For each recording in a given hardware configuration (as presented in Tab. 1), we extract the aforementioned metrics using our pipeline. Missing metric values (usually for extreme bias settings) are removed for all hardware settings to remain consistent. We then summarize each error component by taking the median across all available measurements in that hardware configuration, yielding a compact per-hardware “typical error profile”. The median is used to reduce the influence of outliers and to capture the typical behavior of each hardware setting across the various bias sweeps.

To make components comparable across different scales, each metric component is normalized across hardware settings using a per-metric z -score. Fig. 4 shows these normalized median error components. The plot highlights how the overall operating regime shifts with optics: low light increases noise-related and timing-related penalties, whereas very open apertures can degrade geometry-related components and lead to more events overall. Intermediate settings provide the most balanced trade-off across the task-driven metrics and therefore form the most

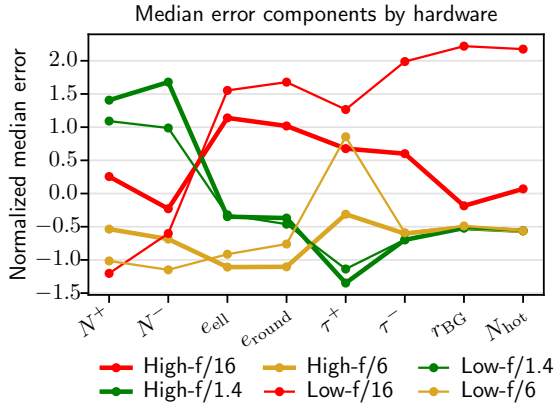


Figure 4: Hardware-level comparison across illumination and aperture settings. Each curve shows the normalized, per-metric z -score median error components aggregated over all measurements within a hardware configuration. Abbreviations: N^+/N^- are positive/negative event counts; e_{ell} is the ellipse residual; e_{round} is the roundness error; τ^+/τ^- are positive/negative trigger delays; r_{BG} is the background rate; and N_{hot} is the number of hot pixels.

promising region for subsequent bias-sweep interpretation and operating-point selection.

6.3 Per-Bias Analysis

The one-dimensional sweeps reveal interpretable sensitivities of individual metrics to specific control parameters and expose practical trade-offs for operating-point selection. We highlight three particularly interesting effects in the following.

6.3.1 Latency and ON/OFF asymmetry

We quantify latency using the LED trigger reference and evaluate ON and OFF responses separately. Fig. 5 compares representative contrast-threshold sweeps for the High-f/1.4 setting. Increasing `bias_diff_on` leads to a comparatively smooth and controllable increase in ON-related delays while leaving the overall behavior stable, which makes it a useful lever for reducing event load without strongly destabilizing timing. In contrast, increasing `bias_diff_off` induces a much stronger latency penalty and a pronounced ON/OFF asymmetry; at higher `bias_diff_off` values, the positive-event delays also increase, indicating that OFF-threshold tuning can affect the timing behavior more broadly than expected from a polarity-isolated view. For comparable event counts, large `bias_diff_off` settings produce substantially higher delays and are accompanied by a clear degradation in the roundness metric, consistent with systematically biased edge timing. Similar sensitivity of OFF-related bias parameters has been reported in benchmark-driven bias tuning studies [15].

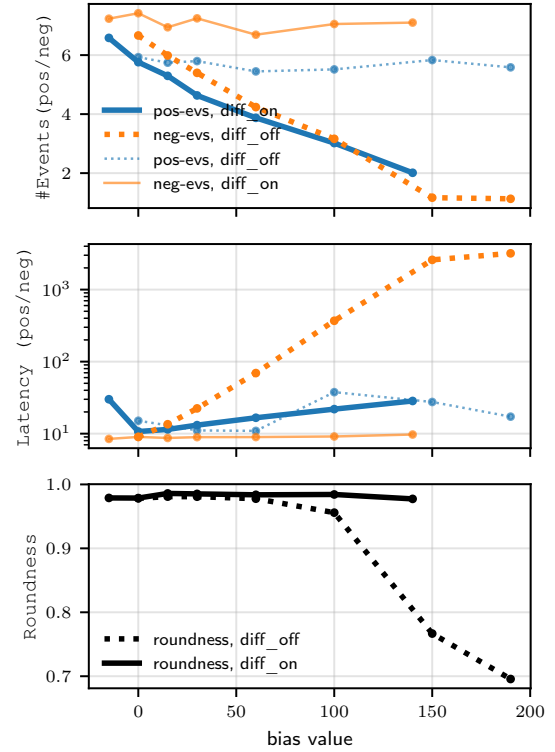


Figure 5: Contrast threshold sweeps for High-f/1.4 setting. The range of `bias_diff_on` is smaller than that of `bias_diff_off`.

6.3.2 Noise and Hot-pixel behavior

Beyond geometry and latency, sensor settings must also be judged by their tendency to generate *background activity* and *hot pixels*. Such pixels consume bandwidth without contributing task-relevant edge information and can degrade downstream tracking by injecting spurious events.

The sweeps in Fig. 6 show that hot-pixel behavior is strongly bias- and optics-dependent. The trends for `bias_diff_on` and `bias_fo` are qualitatively consistent with the expected noise-bandwidth trade-off: Especially at low illumination levels, shot noise has strong influence. This is amplified, if no low-pass filtering is applied (low `bias_fo`). Low contrast thresholds then lead to a large number of events being triggered (each time the noise exceeds the threshold).

Surprisingly, the sweep over `bias_hpf` exhibits a steep increase in the hot-pixel count. The manufacturer documentation suggests that increasing `bias_hpf` should make the sensor less sensitive to slow light changes, therefore decreasing bandwidth on the lower frequency end, but not influence high-frequency changes. Most of this activity is attributed to events with negative polarity.

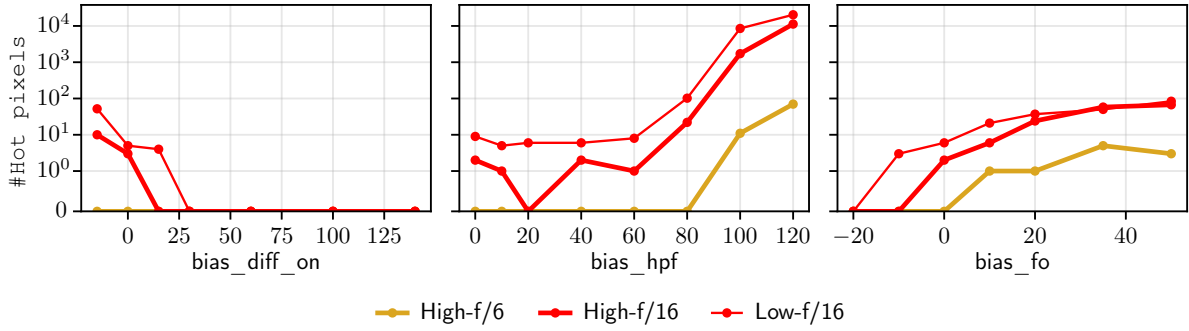


Figure 6: Hot-pixel behavior across bias sweeps and hardware settings. The plot shows the number of hot pixels for each bias setting, revealing strong sensitivities to certain parameters, particularly `bias_hpf`

6.3.3 Geometry versus refractory-period trade-off

The refractory period enforces a dead time where the pixel is idle after an event. This can lead to a suppression of event-bursts and is therefore an effective lever in reducing event rate. The upper panel of Fig. 7 shows this effect.

However, an extensive refractory period also means that if a pixel fires early due to texture induced intensity changes on the object, it can be inactive for the actual object boundary, which can lead to gaps in the event cloud and bias in the recovered contour. This is reflected in the roundness metric, which shows a clear deterioration for long refractory times and further can be qualitatively observed in the motion-compensated event clouds and fitted ellipses in the lower panels of Fig. 7.

7 Selection of Operating Points

The hardware comparison and bias sweeps provide practical guidance for selecting a suitable operating point without requiring black-box optimization. For the present sorting setup, the High-f/6 configuration offers a useful compromise between noise, and optical blur: compared to very small apertures it provides sufficient signal, while avoiding the stronger blur observed for very open apertures.

Based on the sweep results, we select an operating point with increased contrast thresholds (`bias_diff_on` and `bias_diff_off`) to reduce the event rate, while tuning the negative threshold more conservatively due to its instability at high bias values. The refractory bias `bias_refr` is adjusted to suppress event bursts without introducing visible gaps or degrading the fitted object shape. Finally, a higher `bias_fo` is used to improve temporal response; we found, that under the selected higher contrast thresholds, this does not lead to excessive background activity. The resulting operating point, shown in Fig. 2b), yields lower event multiplicity and latency while preserving the circular event contour. Its effect on the metrics

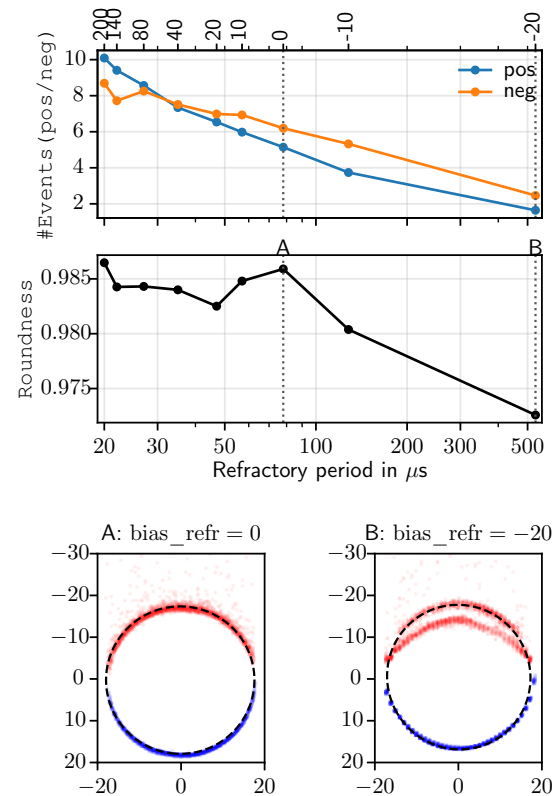


Figure 7: Refractory-period trade-off for the `bias_refr` sweep (at High-f/6). The tick marks on top of the upper axis indicate the numerical bias settings associated with the effective refractory times (lower axis labels). Bottom: two representative motion-compensated event clouds with fitted ellipse overlay, illustrating how overly long refractory times can create gaps and bias the recovered contour while reducing event throughput.

is highlighted in Fig. 3, where it is compared to the default setting.

8 Discussion and Conclusion

We introduce a task-driven framework for characterizing event-camera operating points using a controlled rolling-sphere experiment with known geometry and an LED-based timing reference. Using one-dimensional bias sweeps, we analyze how individual parameters affect the proposed output metrics across different illumination and aperture settings. The sweeps reveal distinct sensitivities of the metrics to contrast thresholds, refractory time, and filter-related biases, informing bias tuning. Based on these trends, we derive a valid operating point by selecting a balanced illumination-aperture regime and combining higher contrast thresholds, moderate refractory suppression, and faster temporal response. Future work will focus on extending the method to unknown objects and enabling online closed-loop parameter tuning. The presented methodology provides a basis for automated operating-point selection in event-based sorting and other event-camera applications where latency, data rate, noise, and geometric fidelity must be balanced explicitly.

References

- [1] G. Maier *et al.*, “A Survey of the State of the Art in Sensor-Based Sorting Technology and Research,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3350987.
- [2] G. Maier *et al.*, “Experimental Evaluation of a Novel Sensor-Based Sorting Approach Featuring Predictive Real-Time Multiobject Tracking,” *IEEE Transactions on Industrial Electronics*, 2021, doi: 10.1109/TIE.2020.2970643.
- [3] P. Lichtsteiner *et al.*, “A 128×128 120 dB 15 μ s Latency Asynchronous Temporal Contrast Vision Sensor,” *IEEE Journal of Solid-State Circuits*, 2008, doi: 10.1109/JSSC.2007.914337.
- [4] T. Delbrück, “Silicon retina with correlation-based, velocity-tuned pixels,” *IEEE Transactions on Neural Networks*, 1993, doi: 10.1109/72.217194.
- [5] G. Gallego *et al.*, “Event-Based Vision: A Survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, doi: 10.1109/TPAMI.2020.3008413.
- [6] Z. Yin *et al.*, “IRIS: Integrated Retinal Functionality in Image Sensors,” *bioRxiv*, 2022, doi: 10.1101/2022.08.14.503909.
- [7] P. Bäcker *et al.*, “Fusion between Event-Based and Line-Scan Camera for Sensor Based Sorting,” in *Conference “Sensor-Based Sorting & Control”* 2024, 2024.
- [8] F. Barranco *et al.*, “Real-Time Clustering and Multi-Target Tracking Using Event-Based Sensors,” *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, doi: 10.1109/IROS.2018.8593380.
- [9] P. Bäcker and G. Maier, “Motion Compensation of Event-Streams for Multitarget Tracking in Sensor-Based Sorting,” *Sensor and Measurement Science International*, pp. 115–116, 2025, doi: 10.5162/SMSI2025/B6.3.
- [10] J. Baßler and M. Heizmann, “Beyond the sparsity paradigm: efficient binary frames for event-based vision,” in *Automated Visual Inspection and Machine Vision VI*, 2025, doi: 10.1117/12.3062298.
- [11] J. Baßler *et al.*, “Ev2Gray: event-only intensity imaging – occlusion-referenced log-intensity reconstruction,” *tm - Technisches Messen*, 2026, doi: 10.1515/teme-2025-0124.
- [12] B. J. McReynolds *et al.*, “Experimental methods to predict dynamic vision sensor event camera performance,” *Optical Engineering*, 2022, doi: 10.1117/1.OE.61.7.074103.
- [13] A. Manakov *et al.*, “Characterisation metrics for performance comparison of area scan and event-based sensors,” *tm - Technisches Messen*, 2025, doi: 10.1515/teme-2025-0018.
- [14] B. McReynolds *et al.*, “Demystifying Event-based Vision Sensor Biasing to Optimize Signal to Noise for Space Domain Awareness,” 2023.
- [15] A. Ziegler *et al.*, “BiasBench: A reproducible benchmark for tuning the biases of event cameras.”
- [16] T. Delbruck *et al.*, “Feedback control of event cameras,” *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, doi: 10.1109/CVPRW53098.2021.00146.
- [17] G. B. Nair *et al.*, “Enhancing Visual Place Recognition via Fast and Slow Adaptive Biasing in Event Cameras,” *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, doi: 10.1109/IROS58592.2024.10802384.
- [18] “Spatter Tracking using Python — Metavision SDK Docs 4.6.2 documentation,” online. Available: https://docs.prophesee.ai/4.6.2/samples/modules/analytics/tracking_spatter_py.html?highlight=spatter%20tracking.

SNR Enhancement in Event-Based Grayscale Image Reconstruction via Pixel Sensitivity Characterization

Adrián Leonardo Delucchi¹; Johannes Baßler¹; Paul Bäcker²; Michael Heizmann¹ 

¹Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT), Karlsruhe

²Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB), Karlsruhe

Abstract This paper presents a robust method for enhancing the signal-to-noise ratio (SNR) in grayscale images reconstructed via the `Ev2Gray` method from event-based vision sensors. Reconstructed images often exhibit significant noise originating from the sensor's inherent fixed-pattern noise (FPN). We propose a two-phase technique involving the empirical extraction of a pixel sensitivity matrix followed by an algorithmic normalization. By adopting a logarithmic sensor response model, the multiplicative noise is transformed into a linear additive problem, allowing for computationally efficient correction. The contributions of this work are two-fold: First, we introduce a practical and straightforward method to quantify the FPN of a sensor without the requirement of laboratory equipment. Second, we demonstrate that using the knowledge of this quantified FPN, we can significantly improve image quality.

The proposed dynamic parameter estimation enables the algorithm to adapt to varying conditions, such as temperature and gain fluctuations, without requiring manual recalibration. Experimental results demonstrate a substantial reduction in structural correlation between the FPN and the reconstructed images, decreasing from $r = 0.257$ to -0.063 and $r = 0.138$ to -0.009 , respectively, while maintaining the global mean intensity. This approach effectively suppresses persistent sensor-level artifacts across diverse operating scenarios, providing a critical advancement for high-quality semantic and structural scene interpretation in event-based vision.

Keywords: Event-based vision, Image denoising, Fixed-pattern noise (FPN), Grayscale image reconstruction, Logarithmic sensor model

1 Introduction

Event-based vision is a novel change-based image acquisition method offering improved dynamic range, reduced data redundancy and latency relative to conventional frame-based vision [1]. With the help of the `Ev2Gray` method, grayscale images can be reconstructed from static scenes, using only information from an event-based camera [2]. This robust technique enables semantic and structural scene interpretation exclusively through event-based vision information [3]. However, reconstructed images exhibit pronounced sensor-specific noise. We propose a technique to decrease the noise in the reconstructed images through fixed-pattern noise (FPN) characterization of the sensor. This is done by stacking multiple recordings of the same homogeneous scene under identical conditions, which are then processed with the method `Ev2Gray`. By averaging these reconstructed images, statistical noise factors, such as shot noise or temporal dark-current noise, are diminished while the FPN is isolated. This allows us to determine a sensitivity matrix of the sensor, which can then be used to compensate and reduce the noise in newly reconstructed images. The proposed method not only improves the quality of reconstructed images, but quantifies the FPN of a sensor in an accessible way without expensive laboratory equipment, which can be later used for further processing.

2 Related Work

Event-based vision cameras exhibit several distinct noise sources. The dominant physical noise mechanisms include: photon-shot noise, leakage and dark-current noise, and FPN [4]. Most denoising approaches rely on the empirical observation that true events cluster in space and time, whereas noise events appear isolated [1].

2.1 Classical Methods for Filtering Noise

Early methods of noise reduction in event streams rely on filtering to remove the events caused by noise. A common method for this is the Background Activity Filter (BAF) (or Nearest-Neighbour-Filter). This filter removes events that have no other events occurring within their spatial neighborhood and a short temporal window. The underlying idea is that real brightness changes in the scene trigger multiple events in neighboring pixels and within close temporal proximity, whereas noise events often remain isolated [5]. Although the BAF is already efficient and can be implemented in hardware, it was further improved by Guo and Delbruck [6].

In addition to neighborhood filters, spatio-temporal smoothing methods have been proposed. Local event sequences can be evaluated by fitting a plane in space-time – if an event does not lie on a surface that can be explained by motion, it is classified as noise [7].

Statistical criteria have also been used for filtering out noise. A threshold for the event density within a small space-time window can be defined and events whose local density falls below this threshold are eliminated [8]. This removes random isolated events in a coarse filtering step, followed by fine filtering of hot-pixel activity.

Purely neighborhood based filters, while simple and fast, can reach their limits in the presence of complex noise patterns. Nevertheless, BAF and related methods remain the foundation of real-time noise suppression in event-based cameras to this day [8].

2.2 Learning-based Approaches

In recent years, machine-learning methods have been developed to detect noise events more reliably. Baldwin *et al.* (2020) proposed an approach in which an Event Probability Mask (EPM) is first computed, resulting in the probability for each pixel that a detected event is genuine. These probabilities are estimated from auxiliary information such as simultaneously recorded intensity images and IMU data. Using such labels, a CNN for noise removal could be trained (EDnCNN). This approach not only treats background activity like the filtering methods, but also false negatives, temporal jitter and random fluctuations in the event rate [9]. Further improvements were done by Fang *et al.* with AEDNet and Jiang *et al.* with EDFormer, which achieve state-of-the-art noise suppression through learned feature processing [10, 11].

While these networks provide high accuracy in separating signals from noise events, they are often computationally expensive. This creates a bottleneck, given the high data rates inherent to event-based vision cameras. Ruan *et al.* (2025) propose EDMamba, a framework based on state-space models. EDMamba encodes events as 4D point clouds and employs efficient spatio-temporal projections to remove noise events, allowing it to achieve state-of-the-art accuracy with significantly lower resource requirements [12].

2.3 Denoising in Image Reconstruction

The reconstruction of grayscale images using event-based vision data is significantly complicated by noise. For this reason, many reconstruction methods directly incorporate mechanisms for noise compensation.

Bardow *et al.* (2016) combined optical flow and intensity estimation and used a robust smoothing term to suppress noise and improve image contrast. This reduces image noise, but can cause a loss in detail [13]. New deep-learning approaches for event-to-video reconstruction (e.g. E2VID by Rebecq *et al.* (2019)) have learned how to handle noise during the training process, likely because simulated noise events were included during training. It has been observed that these approaches behave more

robustly in comparison to optical-flow-based methods [14]. Prefiltering methods, which include spatial smoothing via a time surface of temporal smoothing to remove isolated noise, are often applied to further increase the quality of the data [14].

To the best of our knowledge, no method currently exists to compensate for one of the biggest noise factors in event-based vision cameras: the FPN. This paper presents a robust, fast and efficient approach to compensate for the FPN in grayscale image reconstruction, significantly enhancing the quality of the resulting images.

3 Method Overview

The proposed method in this paper aims to improve the quality of reconstructed grayscale images from event-based vision data using Ev2Gray [3] by reducing the noise present in these images and thus increasing the signal-to-noise ratio (SNR). This is achieved by determining the particular FPN of the sensor and compensating for it in reconstructed images. The approach is structured into two main phases: the empirical extraction of the FPN and the subsequent correction of reconstructed images. In this work we treat the Ev2Gray method as a virtual sensor and only analyze its input and output.

3.1 Ev2Gray Method Description

Ev2Gray is a novel method for reconstructing pseudo-grayscale images from pure event-based vision data. By leveraging a moving opaque stripe to impose controlled occlusion of the scene, every pixel observes a uniform intensity anchor during the sweep. This approach enables the recovery of absolute brightness structures and fine details in static scenes without requiring external sensors or pre-trained models. The method effectively bridges the gap between asynchronous event streams and static scene interpretation, and further provides a computationally simple pathway for applying conventional machine vision algorithms developed for intensity images to event-based vision use cases. For further details about the Ev2Gray method the reader is referred to Baßler *et al.* (2025, 2026) [2, 3].

3.2 Determining the FPN

To quantify the FPN inherent to the used sensor, it must be isolated from transient noise sources. This is achieved by exploiting the different temporal characteristics of the noise components. While FPN is time-invariant due to the physical properties of the sensor (considering all parameters remain identical, e.g. temperature), all other noise sources (e.g., photon-shot noise, dark-current noise) are stochastic and time-variant.

Although these stochastic sources exhibit spatial variance in any single frame, their zero-mean distribution allows for effective suppression through temporal integration. By averaging over multiple

recordings, the variance of the stochastic spatial deviations is reduced, causing them to converge towards their expectation value. Consequently, the temporally averaged image $\bar{I}$ remains primarily affected by the time-invariant FPN, which is reinforced as the dominant spatial non-uniformity.

The practical implementation involves the following procedure: a homogeneous, static, evenly lit scene is recorded and reconstructed using the method `Ev2Gray` [3]. This is repeated N times while maintaining static scene conditions and ensuring that all camera parameters and `Ev2Gray` settings remain identical to guarantee experimental consistency. The resulting reconstructed grayscale images are then pixel-wise stacked to create a new grayscale image, which contains the average gray value for each pixel computed across all N images. This temporally averaged image is referred to as $\bar{I}$ and serves as the basis for deriving the sensitivity matrix S_{meas} , which provides an approximation of the relative sensitivity of each sensor pixel.

This approximation is achieved by determining the pixel-wise deviation of $\bar{I}$ from a reference value μ_{ref} . While a global average of $\bar{I}$ could serve as a theoretical baseline, achieving perfect homogeneity across the entire scene is difficult. Therefore, a moving average filter is used to determine μ_{ref} instead, resulting in a bounded positive value due to the nature of the output of `Ev2Gray`. This allows the reference value to adapt to local intensity variations while maintaining an ideal uniform response. We define the sensitivity matrix S_{meas} as:

$$S_{\text{meas}}(x, y) := \frac{\bar{I}(x, y)}{\mu_{\text{ref}}(x, y)} \quad (1)$$

The matrix S_{meas} represents the FPN of the sensor by stating the approximated empirically inferred sensitivity of each pixel. Each entry of S_{meas} expresses the deviation for that specific pixel, where values greater than 1 indicate increased sensitivity and values smaller than 1 indicate decreased sensitivity relative to the sensor's local average.

3.3 Algorithmic Correction

Once the FPN of the used event-based vision sensor is quantified with the sensitivity matrix S_{meas} , grayscale images which were recorded with the same sensor – thereby exhibiting the same FPN characteristics – can be corrected.

3.3.1 Modelling the Intensity Measurement

A logarithmic sensor response model is chosen due to the nature of event-based vision sensors [1]. The FPN present in these is predominantly caused by spatial variations in pixel sensitivity and contrast threshold. Additional additive offsets and background activity may contribute under low-signal conditions, but typically represent a secondary effect. For this reason, the mathematical model of

the sensor response used in this work will only take into consideration the FPN.

While noise in conventional frame-based imaging is often modeled as an additive component, event-based vision sensors require a different approach due to their inherent response to relative contrast changes rather than absolute intensity. In this context, spatial variations in pixel sensitivity act as a gain error. This has its cause in the fact that a change in the sensitivity of an event-based pixel produces an error which scales with the radiance. Under low-light conditions the absolute error is small, whereas in high-intensity scenarios, the absolute error becomes large [15].

Furthermore, the FPN is assumed to be of exponential nature to facilitate the calculation of the scaling factor α . This factor accounts for the varying magnitudes of FPN intensity resulting from the inconsistent normalization inherent in `Ev2Gray`.

We assume that the reconstructed pixel intensity I_{meas} is a function of I_{ideal} , the ideal scene radiance without FPN, modulated by S_{ideal} , the true spatial sensitivity pattern representing the exact FPN inherent to the sensor used, according to the following relationship:

$$I_{\text{meas}} := I_{\text{ideal}} \cdot \exp(\alpha \cdot S_{\text{ideal}}) \quad (2)$$

where α denotes a scalar weighting factor.

3.3.2 Logarithmic Linearization

To facilitate a computationally efficient correction, the multiplicative model is linearized by transforming the intensity values into the logarithmic domain. To ensure numerical stability and avoid undefined values at $I_{\text{meas}} = 0$, a small regularization constant $\epsilon = 10^{-9}$ is introduced:

$$\ln(I_{\text{meas}} + \epsilon) = \ln(I_{\text{ideal}}) + \alpha \cdot S_{\text{ideal}} \quad (3)$$

Defining $L := \ln(I)$, the relationship can be expressed as follows:

$$L_{\text{meas}} := L_{\text{ideal}} + \alpha \cdot S_{\text{ideal}} \quad (4)$$

3.3.3 Dynamic Parameter Estimation

The goal is to determine L_{ideal} by estimating the unknown coefficient α , as this would yield the FPN-free image. Since the intensity of FPN fluctuates based on environmental factors such as sensor temperature or gain settings, the scaling factor α must be estimated dynamically for each individual image [16]. Starting from the linear model defined in Eq. (4), we treat the estimation of α as an ordinary least-squares (OLS) problem and aim to minimize the residual difference between the observed signal (L_{meas}) and the noise model prediction ($\alpha \cdot S_{\text{ideal}}$). By centering both the measured log-intensity L_{meas} and the sensitivity matrix S_{ideal} around their respective global spatial means, $\bar{L}_{\text{meas}}$ and $\bar{S}_{\text{ideal}}$, we

isolate the structural fluctuations and eliminate the need for an explicit intercept term, focusing solely on the structural correlation.

The optimal scaling factor α is then found by minimizing the residual variance between the observed noise and our model:

$$J = \min_{\alpha} \sum_{\substack{x \in [0, W] \\ y \in [0, H]}} [(L_{\text{meas}}(x, y) - \bar{L}_{\text{meas}}) - \alpha(S_{\text{ideal}}(x, y) - \bar{S}_{\text{ideal}})]^2 \quad (5)$$

After centering L_{meas} and S_{ideal} by subtracting their respective spatial means, the estimation of α reduces to a least-squares regression without an intercept. The corresponding minimizer is:

$$\alpha = \frac{\text{cov}(L_{\text{meas}}, S_{\text{ideal}})}{\text{var}(S_{\text{ideal}})} \quad (6)$$

Hence, α captures the degree to which the measured log-intensity is structurally aligned with the FPN pattern and adapts the correction strength to the current image. A larger value of α indicates that the spatial deviations in L_{meas} more strongly follow the pattern encoded in S_{ideal} , and therefore that a stronger FPN correction is required.

3.3.4 Image Reconstruction and Validation

Once the optimal α is determined, the noise-compensated image in the log-domain can be calculated by solving Eq. (4) for L_{ideal} .

Because S_{meas} only approximates the true sensitivity matrix S_{ideal} , the corrected result is likewise an approximation of L_{ideal} , which we define as $L_{\text{corrected}}$:

$$L_{\text{corrected}} := L_{\text{meas}} - \alpha \cdot S_{\text{meas}} \quad (7)$$

$L_{\text{corrected}}$ represents an estimate for the log-image without FPN. The final image $I_{\text{corrected}}$ is reconstructed using the exponential function:

$$I_{\text{corrected}} := \exp(L_{\text{corrected}}) - \epsilon \quad (8)$$

To preserve numerical consistency for standard display pipelines, we apply min-max scaling to map the floating-point values onto the interval $[0, 255]$, followed by a conversion to 8-bit unsigned integer format.

4 Results

In the following chapter, the results obtained using the presented method are discussed. To evaluate the performance, a Prophesee EVK4 camera was utilized. Although the evaluation is limited to a single sensor type, FPN is a well-documented characteristic of event-based vision, and therefore this method remains applicable to other event-based architectures [1].

4.1 Quantitative Performance Analysis

To determine the sensitivity matrix S_{meas} , a kernel size of 250 was used for the moving average filter (see Section 3.2). The efficacy of the logarithmic FPN correction was evaluated by quantifying the structural correlation between the image and the sensitivity matrix S_{meas} . Ideally, a successful correction yields a correlation coefficient near zero, indicating a complete decoupling of sensor-induced patterns from the signal.

4.1.1 Test Case *Schloss Karlsruhe*

For the test case *Schloss Karlsruhe*, the algorithm estimated an optimal scaling factor $\alpha = 0.0033$.



Figure 1: Example *Schloss Karlsruhe*. Left: uncorrected image I_{meas} . Right: FPN corrected image $I_{\text{corrected}}$.

The test case image is presented in Fig. 1, where the left half displays the uncorrected reconstruction I_{meas} and the right half shows the result after FPN correction $I_{\text{corrected}}$.

As summarized in Table 1, the high initial correlation with the FPN structures ($r_{\text{uncorrected}} = 0.257$) – representing the visible FPN – was reduced to a minimal level ($r_{\text{corrected}} = -0.063$). The remaining residual correlation is attributed to the fact that the measured sensitivity matrix S_{meas} serves as an approximation of the underlying ideal sensitivity S_{ideal} , limited by experimental uncertainties. Nevertheless, this result indicates a successful suppression of the dominant noise patterns. Furthermore, the global mean intensity remained largely unchanged, confirming that the correction does not introduce unwanted brightness shifts.

Metric	Uncorrected	Corrected
α	–	0.0033
r	0.257	–0.063
$\bar{I}$	147.871	147.103

Table 1: Statistics of test case *Schloss Karlsruhe*

4.1.2 Test Case *Grayscale Chart*

For the test case *Grayscale Chart*, the algorithm estimated an optimal scaling factor $\alpha = 0.0019$. The test case image is presented in Fig. 2, displaying the uncorrected reconstruction I_{meas} in the left half, while the FPN corrected reconstruction $I_{\text{corrected}}$ is visible as the right half of the image.

As summarized in Table 2, the notable initial correlation between the image and the FPN structures ($r_{\text{uncorrected}} = 0.134$) was reduced to a negligible level ($r_{\text{corrected}} = -0.009$). Consistent with the aforementioned limitations of S_{meas} as an approx-

imation, this minor residual indicates an effective suppression of the dominant noise patterns. Furthermore, the global mean intensity also remained largely unchanged.

Metric	Uncorrected	Corrected
α	–	0.0019
r	0.134	–0.009
$\bar{I}$	139.505	139.167

Table 2: Statistics of test case *Grayscale Chart*

4.2 Visual Validation and Signal Quality

A qualitative comparison of the reconstructed grayscale images before and after FPN correction shows quite clearly that in the uncorrected state (Fig. 1 left half, Fig. 2 left half) pronounced high-frequency noise artifacts are visible, especially in homogeneous regions of the scene.

Following the application of the proposed method, these artifacts are severely reduced, resulting in a visually smoother intensity distribution with higher global SNR (Fig. 1 right half, Fig. 2 right half).

Unlike a Gaussian filter, our approach does not function as a simple low-pass filter. While it effectively suppresses noise, it preserves fine structural details and high-frequency components – such as the windows and fence in Fig. 1 – which would otherwise be lost to smoothing.

4.3 Correlation between Scene Illumination and Scaling Factor

To evaluate the adaptability of the proposed method, the relationship between the scene illumination E and the estimated coefficient α was evaluated. This investigation is of interest to determine whether the sensor's FPN scaling follows a

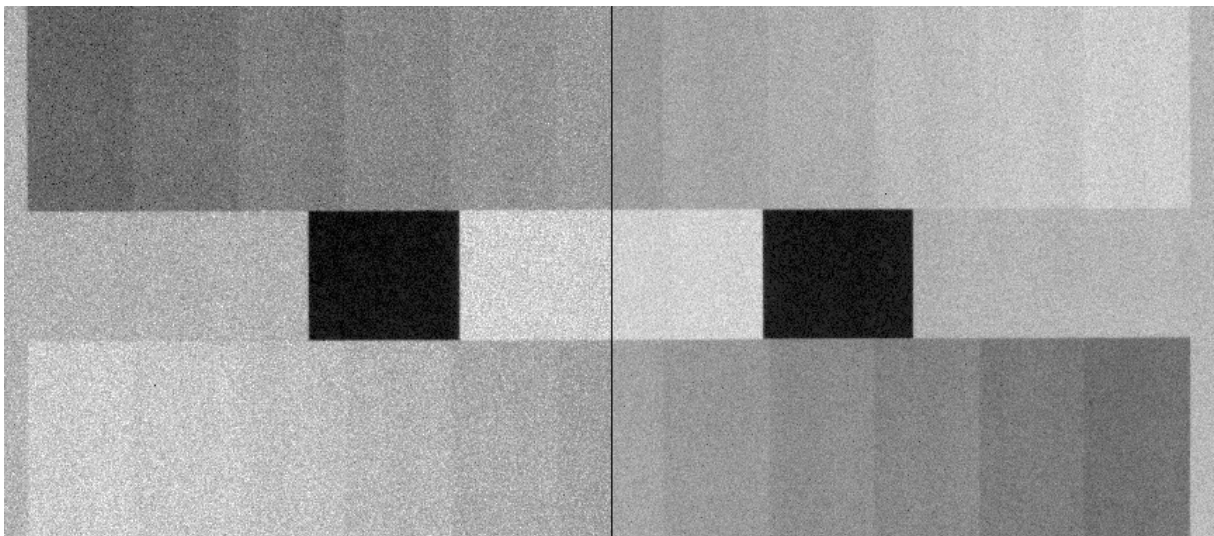


Figure 2: Example *Grayscale Chart*. Left: uncorrected image I_{meas} . Right: FPN corrected image $I_{\text{corrected}}$.

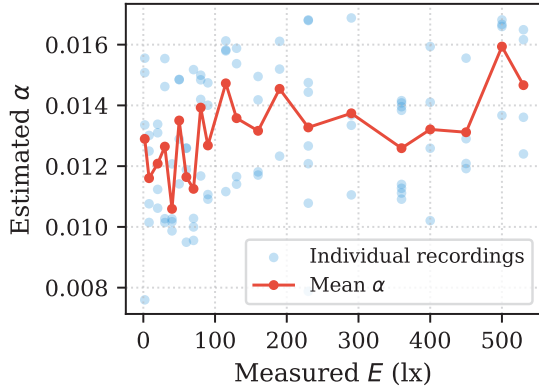


Figure 3: Correlation between scene illumination E and estimated scaling factor α . The scatter plot shows individual recordings, while the red line tracks the mean value per illumination level.

predictable, deterministic trend relative to illumination intensity, which could potentially allow for a static calibration bypass. Measurements of the same homogeneous, static, evenly lit scene were made, solely varying the scene illumination in 21 steps from 2 lux to 530 lux, capturing five recordings per illumination level.

Fig. 3 illustrates the distribution: blue dots represent individual scaling factors, while red points indicate the mean scaling factor for each illumination value. The Pearson correlation factor for this distribution is 0.2884, indicating a weak correlation between the scaling factor α and the lux value E of the scene. The high variance observed in the scaling factor suggests that while scene illumination influences the presence of FPN, it represents a secondary effect. Other factors appear to have a more significant impact on the scaling factor's distribution.

5 Conclusion

This paper presented a robust method for enhancing the SNR in event-based grayscale image reconstruction through characterization and compensation for the FPN of the sensor. By adopting a logarithmic sensor response model, we successfully transformed a multiplicative noise problem into an additive one. The proposed dynamic parameter estimation allows the algorithm to adapt to varying environmental conditions – such as gain and temperature fluctuations – without manual recalibration or additional measurements of the scene.

Experimental results on reconstructed images demonstrate that the method achieves a substantial reduction in structural correlation, decreasing from $r_{\text{uncorrected}} = 0.257$ to $r_{\text{corrected}} = -0.063$ for the test case *Schloss Karlsruhe* and from $r_{\text{uncorrected}} = 0.138$ to $r_{\text{corrected}} = -0.009$ for the test case *Grayscale Chart*. Furthermore, the stability of the

global mean intensity confirms that the correction preserves the original scene luminance. The presented method provides a critical step forward in making event-based vision more viable for high-quality semantic and structural scene interpretation by effectively suppressing persistent sensor-level artifacts.

6 Outlook

The findings in this work provide a foundation for several promising research avenues. First, while this study serves as a proof-of-concept using a single event-based vision sensor, future research will involve multi-camera setups to demonstrate the generalizability of the proposed method across different hardware architectures. Furthermore, the sensitivity matrix S_{meas} offers a unique diagnostic tool: by comparing matrices from various sensors, inherent manufacturing differences and hardware-specific noise characteristics can be quantified. This cross-sensor analysis will be crucial for developing universal denoising standards in event-based vision. Second, the direct integration of the presented FPN correction into the *Ev2Gray* reconstruction provides an immediate refinement of the method, yielding higher-fidelity images natively. This integration leads to a more streamlined and automated reconstruction pipeline.

Third, incorporating the sensitivity matrix S_{meas} directly into the event processing pipeline offers significant potential. By weighting or filtering individual events in the stream based on their specific sensitivity, it may be possible to suppress FPN in real-time. This approach could significantly enhance motion compensation and enable more efficient bulk sorting of event data, as events can be prioritized based on their reliability (weighted by S_{meas}). Such “event-level” denoising would enhance the data quality for all subsequent higher-level algorithms, such as object detection or tracking.

Finally, the core principles of this compensation method could be expanded to conventional frame-based CMOS cameras. Given that pixel-wise sensitivity variations are a universal phenomenon in CMOS technology, adapting this robust, empirical calibration approach could enhance the dynamic range and image quality of standard vision systems.

References

- [1] Guillermo Gallego, Tobi Delbrück, et al. "Event-based Vision: A Survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.1 (2022), pp. 154–180. DOI: 10.1109/TPAMI.2020.3008413. URL: <https://ieeexplore.ieee.org/document/9138762>.
- [2] J. Baßler and M. Heizmann. "Ev2Gray: generating intensity images from event streams". In: *Proc. Sensor and Measurement Science International (SMSI)*. Nürnberg, 2025, pp. 117–118. DOI: 10.5162/SMSI2025/B6.4.
- [3] J. Baßler, A. L. Delucchi, et al. "Ev2Gray: event-only intensity imaging – occlusion-referenced log-intensity reconstruction". In: *tm – Technisches Messen* (2026). DOI: 10.1515/teme-2025-0124.
- [4] Daxin Gu, Jia Li, et al. "Reliable Event Generation With Invertible Conditional Normalizing Flow". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46.2 (2024), pp. 927–943. DOI: 10.1109/TPAMI.2023.3326538.
- [5] Vandana Padala, Arindam Basu, and Garrick Orchard. "A Noise Filtering Algorithm for Event-Based Asynchronous Change Detection Image Sensors on TrueNorth and Its Implementation on TrueNorth". In: *Frontiers in Neuroscience* Volume 12 - 2018 (2018). ISSN: 1662-453X. DOI: 10.3389/fnins.2018.00118. URL: <https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2018.00118>.
- [6] Shasha Guo and Tobi Delbrück. "Low Cost and Latency Event Camera Background Activity Denoising". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45.1 (2023), pp. 785–795. DOI: 10.1109/TPAMI.2022.3152999.
- [7] Tobi Delbrück. "Frame-free dynamic digital vision". In: *Proceedings of the International Symposium on Secure-Life Electronics, Advanced Electronics for Quality Life and Society* (Mar. 2008). DOI: 10.5167/uzh-17620.
- [8] Yang Feng, Hengyi Lv, et al. "Event Density Based Denoising Method for Dynamic Vision Sensor". In: *Applied Sciences* 10 (Mar. 2020), p. 2024. DOI: 10.3390/app10062024.
- [9] R. Wes Baldwin, Mohammed Almatrafi, et al. "Event Probability Mask (EPM) and Event Denoising Convolutional Neural Network (EDnCNN) for Neuromorphic Cameras". In: *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 1698–1707. DOI: 10.1109/CVPR42600.2020.00177.
- [10] Huachen Fang, Wu Jinjian, et al. "AEDNet: Asynchronous Event Denoising with Spatial-Temporal Correlation among Irregular Data". In: Oct. 2022, pp. 1427–1435. DOI: 10.1145/3503161.3548048.
- [11] Bin Jiang, Bo Xiong, et al. "EDformer: Transformer-Based Event Denoising Across Varied Noise Levels". In: Nov. 2024, pp. 200–216. ISBN: 978-3-031-73382-6. DOI: 10.1007/978-3-031-73383-3_12.
- [12] Ciyu Ruan, Zihang Gong, et al. *EDmamba: A Simple yet Effective Event Denoising Method with State Space Model*. May 2025. DOI: 10.48550/arXiv.2505.05391.
- [13] Hanme Kim, Stefan Leutenegger, and Andrew Davison. "Real-Time 3D Reconstruction and 6-DoF Tracking with an Event Camera". In: vol. 9910. Oct. 2016, pp. 349–364. ISBN: 978-3-319-46465-7. DOI: 10.1007/978-3-319-46466-4_21.
- [14] Shintaro Shiba, Yoshimitsu Aoki, and Guillermo Gallego. *Simultaneous Motion And Noise Estimation with Event Cameras*. 2025. DOI: 10.48550/arXiv.2504.04029. arXiv: 2504.04029 [cs.CV]. URL: <https://arxiv.org/abs/2504.04029>.
- [15] Ziwei Wang, Yonhon Ng, et al. *Event Camera Calibration of Per-pixel Biased Contrast Threshold*. 2020. DOI: 10.48550/arxiv.2012.09378. arXiv: 2012.09378 [cs.CV]. URL: <https://arxiv.org/abs/2012.09378>.
- [16] Christian Brandli, Raphael Berner, et al. "A 240×180 130 dB 3 μ s Latency Global Shutter Spatiotemporal Vision Sensor". In: *IEEE Journal of Solid-State Circuits* 49.10 (2014), pp. 2333–2341. DOI: 10.1109/JSSC.2014.2342715.

Data-efficient optical 4D measurements of moving objects with stereo event cameras and tailored structured illumination

Christoph Freitag^{1,2}; Aliya Turelli Khusainova¹; Dr. Stefan Heist¹; Dr. Andreas Stark³; Dr. Peter Kühmstedt¹; Univ.-Prof. Dr. rer. nat. Gunther Notni^{1,4}; Jun. Prof. Dr. Christian Franke^{3,5,6}

1. Fraunhofer-Institut für Angewandte Optik und Feinmechanik IOF,
Albert-Einstein-Str. 7, 07745 Jena,

2. Friedrich-Schiller-Universität Jena, Institut für Angewandte Physik,
Albert-Einstein-Str. 15, 07745 Jena,

3. Friedrich-Schiller-Universität Jena, Institut für Angewandte Optik und Biophysik,
Max-Wien-Platz 1, 07743 Jena,

4. Technische Universität Ilmenau, Fachgebiet Qualitätssicherung und Industrielle Bildverarbeitung,
Gustav-Kirchhoff-Platz 2, 98693 Ilmenau,

5. Friedrich-Schiller-Universität Jena, Jena Center for Soft Matter,
Philosophenweg 7, 07743 Jena,

6. Friedrich-Schiller-Universität Jena, Abbe Center of Photonics,
Albert-Einstein-Str. 6, 07745 Jena

Contact: +49 3641 807-244, christoph.freitag@iof.fraunhofer.de

Abstract

We present an optical 4D (3D+time) measurement system for data-efficient reconstruction of shape and motion of rigid objects. In contrast to frame-based active stereo systems, which repeatedly resample surfaces at high rates, our approach exploits event cameras to break the tight dependence of scanning rate to object speed. The system combines a stereo pair of event cameras with a projector that generates a periodically rotating contrast edge to sparsely probe the measurement volume. In contrast to frame-based cameras, event cameras only report brightness changes and thus adapt their measurements to the dynamics of the scene. The moving contrast edge sparsely probes the measurement volume. Triangulation of temporally corresponding events in both cameras yields 3D surface points with individual timestamps. We specifically target sub-scan dynamics, i.e. significant object motion during a single slow scan of the contrast edge. Allowing significant sub-scan motion fundamentally breaks the assumptions of frame-based structured-light systems and introduces two key challenges: ambiguous temporal correspondences and motion-induced 3D rolling-shutter distortions. Our reconstruction framework addresses both effects. First, a geometric 4D model of the structured illumination is estimated and used to resolve ambiguity. Second, we estimate the object's motion and warp all 4D points to a common reference time, compensating the rolling-shutter effect. Experiments with a translating rigid object demonstrate that, despite low scan rates and significant sub-scan motion, the proposed method recovers the object's shape with sub-millimetre deviation from a static reference.

Keywords: Event cameras, Structured illumination, Event-based stereo, 4D shape measurement, Motion compensation

Introduction

Optical 3D metrology offers numerous techniques for measuring surfaces of technical objects across different sizes and materials [1,2]. A large class of systems uses active triangulation, for example fringe projection or GOBO projection [3]. In these systems a “scan” is a time interval in which the entire measurement volume is sampled, either sequentially or in parallel. For moving objects, most frame-based sensors have to keep motion within a scan negligible to maintain temporal

consistency of projected patterns and camera exposures. This leads to high scan rates, high-speed cameras and powerful illumination [3].

For rigid bodies, the shape is constant, and the primary interest is motion in space. Frame-based active stereo sensors yet acquire dense 3D snapshots at high rates, repeatedly resampling the same surface. From a metrology perspective, this is redundant and therefore data-inefficient when only rigid motion is present.

Event cameras offer a complementary paradigm by acquiring measurements only where and when intensity changes occur [4,5]. Each pixel independently monitors brightness and emits an event when a contrast threshold is exceeded. The output is an asynchronous stream of events (x, y, p, t) that is sparse in space and time and naturally concentrates on regions where changes occur. When combined with structured illumination, this enables a type of 4D sensor that sparsely samples the measurement volume but with high temporal precision.

In this contribution we extend our previous event-based stereo 3D method for static scenes [6] to 4D measurements of moving rigid objects. A rotating contrast edge generated by a projector probes the scene, and two event cameras observe the resulting brightness changes. We explicitly allow sub-scan dynamics, i.e. significant motion during a single slow scan, and compensate the resulting distortions in the reconstruction.

Related works

Event-based vision has been reviewed extensively by Gallego et al. [4], and recent sensor developments are summarized by [7]. For 3D reconstruction, many methods don't actively sample scenes with structured illumination and single or stereo event cameras. They rather focus on sparse 3D reconstruction and sparse SLAM. These approaches are not designed for dense, metrology-grade surface measurements.

To obtain dense depth information, several authors have combined event cameras with active illumination. Muglikar et al. introduced event-based structured light using a scanning laser point [8]. Fu et al. encoded both temporal and spatial information in an event-based structured-light system [9]. The named systems use a single event camera and, to the best of our knowledge none of them explicitly separate projector-induced events from events caused by moving object features.

Our previous work [6] used a stereo pair of event cameras and a moving contrast edge to reconstruct dense depth maps of static scenes. Raw events were binned into event frames and processed with a classical stereo correlation algorithm [10]. This implicitly assumed negligible motion during a scan and enforced a

one-to-one correspondence between (sub)pixels in both event cameras. It did not represent the ambiguity that occurs when objects move during pattern projection, and it worked with a fixed frame rate.

The present work departs from these approaches by keeping the full temporal resolution of events (no frame binning), explicitly representing one-to-many temporal correspondences, resolving them via a 4D model of the projector, and finally compensating motion-induced 3D rolling-shutter distortions. To the best of our knowledge, no existing approach explicitly models and resolves the ambiguity between illumination-induced and motion-induced events in a unified 4D reconstruction framework.

Methodology

Principle of event-based structured illumination

Our setup consists of a stereo rig of two event cameras observing a measurement volume illuminated by a projector. Inside the projector, a back-side-illuminated rotating mask wheel produces a sharp transition between bright and dark illumination. This projected contrast edge sweeps periodically through the measurement volume.

When the image of the contrast edge passes over an event camera's pixel, the brightness changes abruptly and the pixel emits a burst of events (see Fig. 1). Static regions without changes in illumination remain silent except for noise. Thus, each scan of the contrast edge samples the measurement volume in thin 3D slices over time. Using calibrated stereo geometry, temporally corresponding bursts in both cameras are triangulated to obtain 3D points. The matched stimuli's times provide the 4th coordinate.

An important consequence is the decoupling of scan duration and temporal resolution. The contrast edge can scan slowly across the field of view, while individual object points are sampled as quickly as the contrast edge's image passes from on pixel to its neighbour. We exploit this to allow significant object motion during a single scan (sub-scan dynamics) and correct the arising spatial distortion in post-processing.

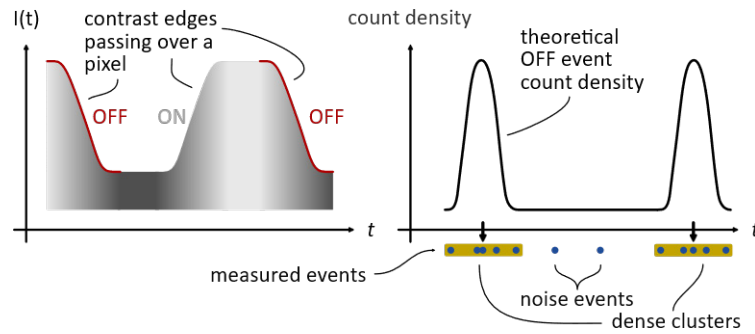


Fig. 1: How an event camera's pixel measures the passing of contrast edges. The left diagram shows the intensity above the pixel over time. Three contrast edges pass, of which two cause a decrease in brightness (OFF edges). Right: Measured events follow a density distribution. To distinguish signal events from noise events, they are clustered by density. Median times of clusters are defined as "stimuli" and pointed out by arrows on the timeline.

Event-generation model and detection of contrast-edge stimuli

The raw event streams contain bursts of events induced by the projector, events from moving object features, and electronic noise. To extract the temporal structure of the projected illumination without imposing an artificial frame rate, we detect contrast-edge stimuli for each pixel directly in the asynchronous event stream.

We model the event-generating process in a pixel by two categories. Burst-like stimuli appear when a contrast edge's image (from the projector or a moving object's feature) passes over a pixel. Between such bursts, there may be noise events. We use a clustering approach to find burst-like stimuli and separate from noise events. To achieve that, we compute the temporal differences between consecutive events. These differences form two classes: short intervals within bursts and longer intervals between bursts or noise events. A threshold duration that separates both classes can be found through different strategies. We use a two-class variance criterion based on Otsu's method [11]. Consecutive events separated by durations shorter than the threshold form a cluster, which is interpreted as a contrast-edge stimulus. The central (representative) stimulus time is defined as the median timestamp of the cluster. For each pixel, this yields a small set of stimulus times without defining frames or a fixed frame rate.

Stereo matching and 4D reconstruction

The next step is to find, for each stimulus in the primary camera, potential corresponding stimuli

in the secondary camera that might have originated from the same object point. We use calibrated intrinsic and extrinsic parameters (obtained with OpenCV [12]) to define epipolar curves in the original distorted image coordinates.

For each stimulus in a reference pixel of the primary camera: The corresponding epipolar curve in the secondary camera is computed. We slide a small three-pixel neighbourhood along the curve. In each of the three pixels we select the stimulus time closest to the reference time (if available). If all three pixels provide a stimulus, we fit a line $t(c) = v^{-1}c + t_0$ through the three (column c_i , time t_i) pairs. The slope v^{-1} represents the inverse image velocity of the contrast edge. If the temporal fit error is below a threshold and the back-projected column at the reference time lies within ± 0.5 pixel of the central column, we accept a match between the reference pixel and a sub-pixel coordinate in the secondary camera.

Because moving object edges can generate additional stimuli, a single stimulus in the primary camera can be matched to multiple candidate locations in the secondary camera. We keep this one-to-many ambiguity at this stage. Each accepted pair of corresponding pixels is undistorted and triangulated with the calibrated camera matrices to obtain a 3D point. The time coordinate of the resulting 4D point is the stimulus time in the primary camera. The raw 4D point cloud contains both true surface points and ghost points arising from false correspondences.

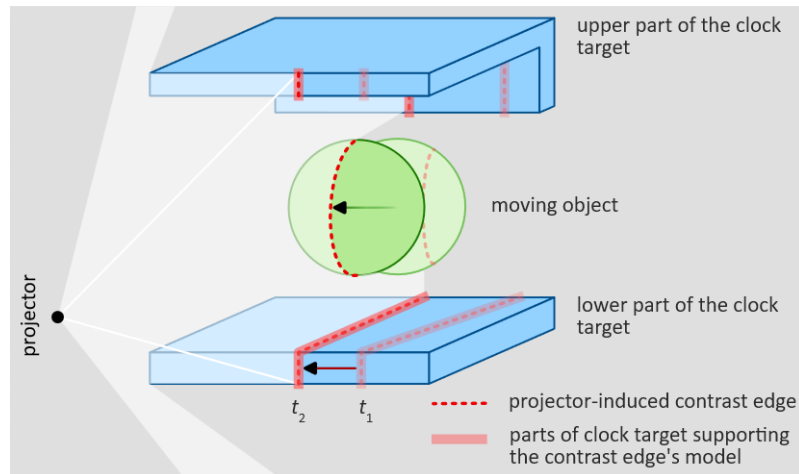


Fig. 2: Visualisation of the measurement volume to support the modelling of the structured illumination. The projector floods the measurement volume with two different levels of illumination, whose transition, a contrast plane, intersects with the objects' surfaces, creating a contrast edge (red). An object (green) is moving along the horizontal direction. Below and above the object, there is the static clock target (blue), whose unambiguous 4D reconstructed points are used to fit the model of the contrast plane which rotates around an axis at the projector's centre of projection. Two sources of events can be found in this scene: First, when the contrast edge's image passes over pixels of the event cameras, it triggers events. Second, images of the moving object's features with sufficient contrast (for instance, in this case the sphere's edge with different reflectivity than the background) trigger additional simultaneous events. Therefore, a horizontal temporal correspondence search on epipolar lines results in ambiguous matches.

Projector model and ghost-point filtering

To distinguish physically consistent 4D points from ambiguities, we introduce a geometric 4D model of the projector that enforces spatio-temporal consistency with the structured illumination (see Fig. 2). The plane rotates around an axis near the projector's centre of projection with approximately constant angular velocity. In the world coordinate system of the 4D points, the model has six parameters: two for the rotation axis' position on the xz plane, two for its orientation, one for angular speed, and one temporal offset.

A static "clock target" with known structure is placed in the measurement volume. During a scan, the rotating contrast plane intersects the clock target at several locations, providing 4D points that sample the plane. Since projector and cameras are not hardware-synchronised, we first detect scan periods using the temporal distribution of events on the clock target. Exploiting the periodic structure of the scans, the scan frequency and phase are estimated using a Rayleigh-type periodicity analysis [13].

For each scan (or scan segment), we then fit the projector model parameters to the 4D points belonging to the static clock target by

minimising their distances to the predicted contrast plane. The resulting model predicts which 3D locations are consistent with the projected contrast edge based on their time coordinates. All 4D points that deviate more than the fit's RMSD from the predicted plane at their timestamp are classified as ghost points and removed. This resolves most stereo ambiguities caused by moving object edges while also removing many noise 4D points.

Motion compensation

After ghost-point filtering, the remaining 4D points likely correspond to locations sampled by the projector's contrast edge. If the object moves during the scan, its surface points are sampled at different times, and the reconstructed shape exhibits a 3D rolling-shutter effect: it appears stretched along the motion direction.

We correct this distortion by explicitly modelling the temporal sampling process and estimating the object's motion to map all measurements to a common reference time t_0 . For a proof of concept, in our experiments we assume a single rigid object undergoing constant linear motion without rotation. Its displacement relative to its position at t_0 is described by

$$\Delta \mathbf{r}(t) = \mathbf{v}(t - t_0)$$

where $\mathbf{v}$ is the 3D velocity vector. Each 4D point $(\mathbf{r}_i, t_i)$ is mapped to

$$\mathbf{r}'_i = \mathbf{r}_i - \mathbf{v}(t_i - t_0)$$

and its time coordinate is set to t_0 . The resulting point cloud approximates the object shape at time t_0 and can be compared to a static reference.

To estimate $\mathbf{v}$ and t_0 , we first acquire a static reference point cloud with the object at rest. For a dynamic scan, we apply candidate motion parameters, warp the dynamic 4D points to t_0 , and compute, for each dynamic point, the distance to its nearest neighbour in the static point cloud. The root-mean-square of these distances serves as loss function. Since this loss function is not differentiable everywhere, we use a derivative-free optimisation method to minimise it and obtain $\mathbf{v}$ and t_0 . We took special care with non-overlapping parts between static and dynamic point clouds. Their associated large distances to points of the other point cloud can distort the loss function such that the optimisation produces erroneous estimates for $\mathbf{v}$ and t_0 .

Experimental setup

The experimental setup (Fig. 3 and Fig. 4) consists of a structured-light projector, a stereo rig of two event cameras, a static clock target, and a motorised linear stage in the measurement volume.

The projector is based on high-power LED illumination and a rotating mask wheel. The wheel comprises 80 masks. In the overlapping fields of view of the cameras only one mask image is visible, producing two contrast edges of opposite polarity. The cameras are configured to detect only one polarity (OFF events), effectively using a single contrast edge at a time. The contrast edge sweeps across the common field of view with an image velocity of approximately 2270 px s^{-1} , corresponding to about 265 mm s^{-1} at a mean object distance of 0.59 m . The scan frequency is about 1.6 Hz , so one scan takes roughly 600 ms .

The stereo rig holds two commercial event cameras with resolution of example 1280×720 pixels equipped with C-mount lenses of 25 mm focal length. The baseline and viewing geometry are chosen such that the projected contrast edge remains within the overlapping fields of view over the full scan. Camera biases

are tuned to ignore ON events and keep noise low in static regions.



Fig. 3: Measurement part of our experimental setup. Two event cameras observe a scene beyond the right edge. The mask wheel is illuminated from the light source within. Rotation causes the projected images of the mask to form moving contrast edges in the measurement volume. These sample the scene and trigger event measurements in the event cameras.

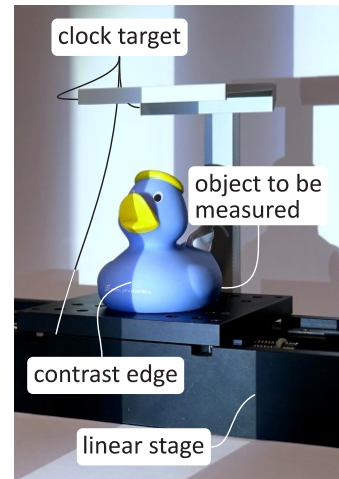


Fig. 4: A view into the measurement volume of our experimental setup with three elements: The duck, which is the object to be measured, a linear stage (black) and a clock target as reference for our contrast edge model.

A linear translation stage is used to move a rigid “duck” object horizontally through the measurement volume at controlled speeds. In the results shown below the object speed is approximately 10 mm s^{-1} , corresponding to about 85 px s^{-1} or a translation of about

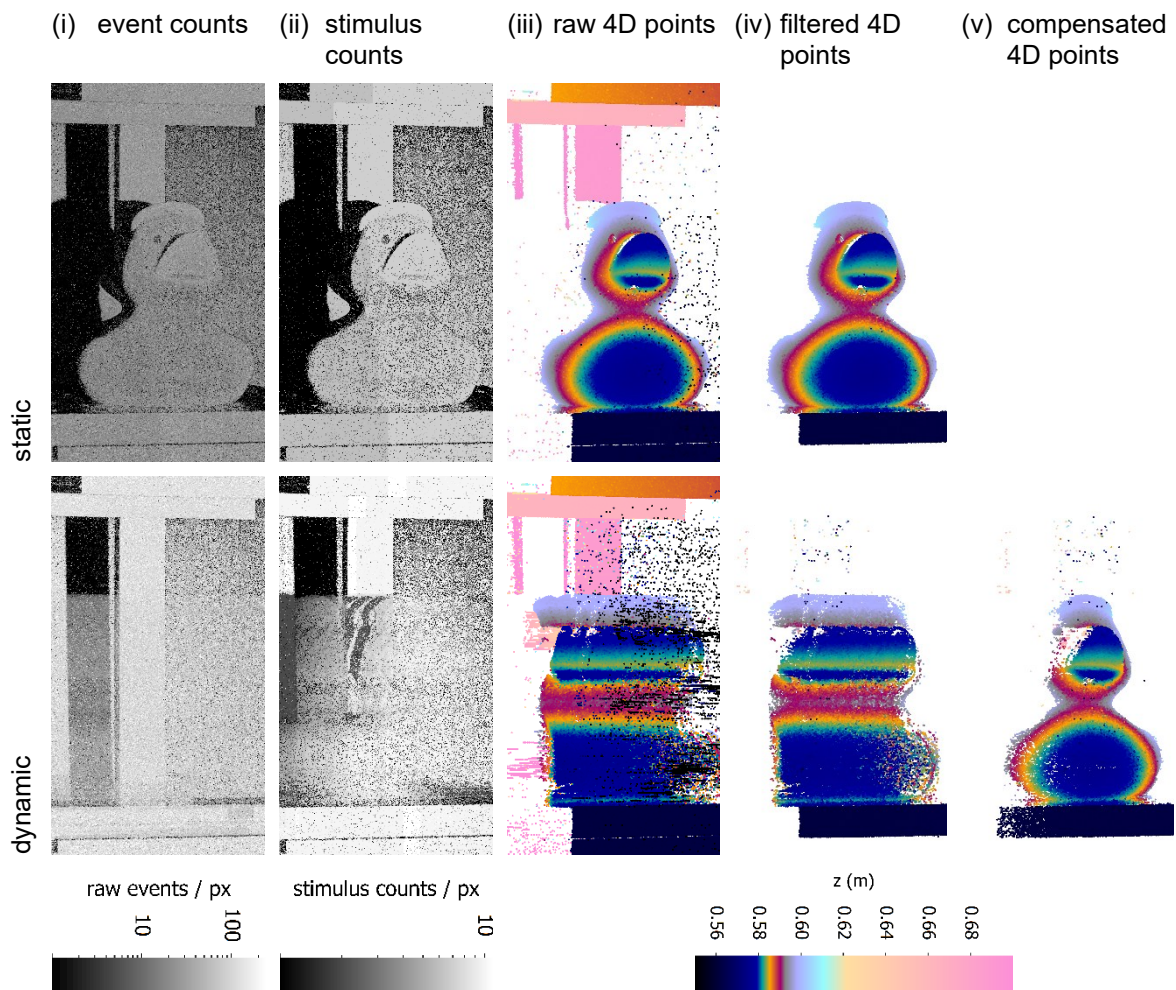


Fig. 5: Following the data from raw events (left) to the final motion-compensated point cloud (right). (i) The first column shows the total number (logarithmic) of events per pixel expressed in brightness. (ii) In the second row, raw events have been reduced to identified contrast edge stimuli per pixel. (iii) Ambiguously matched stimuli result in raw 4D points which consist of true object surface points and wrongly matched and triangulated ghost points. (iv) Using the model of the rotating contrast edge induced by the projector allows to remove ghost and noise points, resulting in a clean reference point cloud in the static case. The dynamic case shows multiple overlapping point clouds of the duck from multiple scans during motion. (v) The motion compensated point cloud is compared to the static reference.

53 pixels per scan in the camera's image plane. The static clock target spans the common field of view of both cameras. This is achieved by three horizontal planes, two at the top and one at the bottom of the measurement volume (see Fig. 2 and Fig. 4).

Results

Reconstruction pipeline

Fig. 5 illustrates the main processing stages for a static and a dynamic scene containing the duck object. For each case, the figure shows (i) event counts per pixel, (ii) the number of detected contrast-edge stimuli per pixel after clustering, (iii) raw 4D points after stereo triangulation, (iv) 4D points after projector-

based ghost-point filtering and removed clock target, and, (v) for the dynamic case, the motion-compensated 3D point cloud.

In the static case, projector-based filtering mainly removes a small number of wrongly matched 4D points. The shape of the duck is already largely correct in the raw 4D point cloud.

However, in the dynamic case, the raw 4D point cloud contains ghost points due to ambiguous temporal correspondences along the epipolar curves. These ghost points appear as spurious structures and violate the simple model of a rotating contrast plane. After fitting the projector model and rejecting inconsistent points, most ghost points are removed and a clean, but

motion-distorted 4D representation of the moving object remains.

Sub-scan dynamics and motion compensation

During a single scan of about 600 ms, the duck object moves roughly 6 mm, corresponding to more than 50 pixels in the image. This clearly violates the quasi-static assumption of classical active stereo systems. Nevertheless, the temporal resolution implied by the contrast-edge velocity on the sensor is on the order of 0.4 ms, many times finer than the scan duration.

If motion is not compensated, the filtered 4D point cloud of the dynamic scan exhibits a 3D rolling-shutter distortion: early points are measured when the object is on one side of the field of view, and later points when it has moved significantly, resulting in a stretched or contracted surface.

We acquired an additional static reference scan of the duck at rest. Using the procedure outlined in the Methodology section, we estimated the translational velocity vector and reference time t_0 that minimise the distance between the motion-compensated dynamic point cloud and the static reference. After compensation, the shape of the moving duck closely matches the static shape except for some sparsely recovered patches. Points are missed when Otsu's method does not yield an optimal separation threshold when finding stimuli in a pixel's raw events. The root-mean-square distance between static and compensated dynamic point clouds is about 0.3 mm when 1% of the largest residuals are discarded as outliers. Fig. 6 shows a spatial residual map over the duck's surface and the distribution of residuals. Larger residuals on parts of the translation platform are due to incomplete coverage of these areas in the static reference; these points were excluded from the optimisation.

From a metrology point of view, these results indicate that, despite low scan rates and substantial sub-scan motion, the event-based system yields shape deviations in the sub-millimetre range for the tested configuration.

Discussion

The presented method rethinks several stages of a structured-light 4D reconstruction pipeline for dynamic scenes based on actively probed event measurements. At the event level, a statistical model of the event-generating process motivates a clustering-based detection

of contrast-edge stimuli that preserves temporal precision without imposing a frame rate. Stereo matching is formulated as a one-to-many correspondence problem constrained by epipolar geometry and the local temporal behaviour of the moving contrast edge.

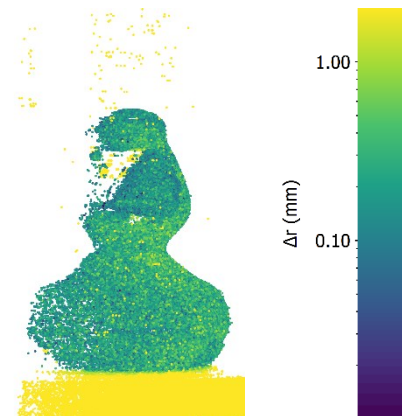


Fig. 6: 3D point cloud of the duck recovered from a dynamic measurement. Points are coloured as to reflect the distance to the nearest point of the static reference point cloud.

A key element is the 4D geometric model of the projector, which enables classification of 4D points as projector-consistent or inconsistent. This step removes ghost points originating from moving object edges and resolves ambiguities that cannot be eliminated in image space alone. Finally, a motion compensation stage treats the 3D rolling-shutter effect caused by sequential sampling of a moving surface.

In the current implementation, a static clock target is used as a calibration aid to identify contrast-edge events and fit the projector model. This component is not a fundamental requirement of the approach but a practical solution in the present proof-of-concept implementation. The clock target reduces the usable measurement volume and constrains the admissible motion directions. Moreover, the motion model is restricted here to a single rigid object undergoing constant linear motion, and an additional static scan is required for each object geometry.

Despite these limitations, the experimental results show that event-based structured light has the potential to make 4D measurements of moving rigid objects considerably more data-efficient than classical frame-based systems. We intend to test this with a thorough comparison in the future. Our sensor avoids high redundancy compared to frame-based 4D sensors when measuring rigid objects. This

makes the approach attractive for inline quality assurance and Industry 4.0 environments, where continuous 3D inspection of moving objects is required without interrupting the process.

Conclusions

We have presented a stereo event-based 4D measurement system with tailored structured illumination for data-efficient and physically consistent reconstruction of shape and motion of rigid objects. The system combines a slowly scanning contrast edge with the high temporal resolution of event cameras, detects contrast-edge stimuli by clustering raw events, resolves stereo ambiguities via a 4D projector model, and compensates the 3D rolling-shutter effect by estimating and inverting object motion.

References

- [1] T. Luhmann, S. Robson, S. Kyle, and J. Boehm, *Close-Range Photogrammetry and 3D Imaging* (De Gruyter, Berlin, Boston, 2020), <https://doi.org/10.1515/9783110607253>.
- [2] A. G. Marrugo, F. Gao, and S. Zhang, "State-of-the-art active optical techniques for three-dimensional surface metrology: a review," *J. Opt. Soc. Am. A* 37, B60–B77 (2020), <https://doi.org/10.1364/JOSAA.398644>.
- [3] S. Heist, P. Dietrich, M. Landmann, et al., "GOBO projection for 3D measurements at highest frame rates: a performance analysis," *Light: Science & Applications* 7, 71 (2018), <https://doi.org/10.1038/s41377-018-0072-3>.
- [4] G. Gallego, T. Delbrück, G. Orchard, et al., "Event-Based Vision: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.* 44, 154–180 (2022), <https://doi.org/10.1109/TPAMI.2020.3008413>.
- [5] T. Finateu, A. Niwa, D. Matolin, et al., "A 1280×720 Back-Illuminated Stacked Temporal Contrast Event-Based Vision Sensor with 4.86 μm Pixels, 1.066 GEPS Readout, Programmable Event-Rate Controller and Compressive Data-Formatting Pipeline," in *2020 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 112–114 (2020), <https://doi.org/10.1109/ISSCC19947.2020.9063149>.
- [6] C. Freitag, S. Heist, A. Stark, et al., "Photometrically optimized event-based stereo 3D measurements," *Opt. Express* 33, 2924–2939 (2025), <https://doi.org/10.1364/OE.541659>.
- [7] B. Chakravarthi, A. A. Verma, K. Daniilidis, et al., "Recent Event Camera Innovations: A Survey," in *Computer Vision – ECCV 2024 Workshops*, A. Del Bue, C. Canton, J. Pont-Tuset, and T. Tommasi, eds. (Springer Nature Switzerland, Cham, 2025), pp. 342–376, https://doi.org/10.1007/978-3-031-92460-6_21.
- [8] M. Muglikar, G. Gallego, and D. Scaramuzza, "ESL: Event-based Structured Light," in *2021 International Conference on 3D Vision (3DV)*, pp. 1165–1174 (2021), <https://doi.org/10.1109/3DV53792.2021.00124>.
- [9] J. Fu, Y. Zhang, Y. Li, et al., "Fast 3D reconstruction via event-based structured light with spatio-temporal coding," *Opt. Express* 31, 44588–44602 (2023), <https://doi.org/10.1364/OE.507688>.
- [10] P. Dietrich, S. Heist, M. Landmann, et al., "BICOS—An Algorithm for Fast Real-Time Correspondence Search for Statistical Pattern Projection-Based Active Stereo Sensors," *Appl. Sci.* 9 (2019), <https://doi.org/10.3390/app9163330>.
- [11] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* 9, 62–66 (1979), <https://doi.org/10.1109/TSMC.1979.4310076>.
- [12] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools* (2000).
- [13] T. Bai, "Methods of Periodicity Analysis: Relationship between the Rayleigh Analysis and a Maximum Likelihood Method," *The Astrophys. J.* 397, 584–590 (1992), <https://doi.org/10.1086/171816>.

On a translating test object, the method achieves sub-millimetre agreement between static and dynamic reconstructions while operating at a low scan rate and allowing significant motion within a scan. Future work will address more general motion models, reduce reliance on auxiliary calibration structures such as the clock target, and provide a quantitative comparison to high-speed frame-based systems.

Funding

This work was supported in part by the German Federal Ministry of Education and Research (BMBF, project 13N16890) and by the Thüringer Aufbaubank (project FGR 0060 KI-supER).

Entwicklung einer neuartigen photoakustischen Bodensonde für Kohlenstoffdioxid

Johannes Kapp^{1,3}, Christian Weber^{1,3}, Julian Brozon², Helmer Schack-Kirchner², Jürgen Wöllenstein^{1,3}

¹Universität Freiburg, Institut für Mikrosystemtechnik – IMTEK, 79110 Freiburg

²Universität Freiburg, Fakultät für Umwelt und Natürliche Ressourcen – UNR, 79104 Freiburg

³Fraunhofer-Institut für Physikalische Messtechnik IPM, Georges-Köhler-Allee 301, 79110 Freiburg

Zusammenfassung

Zur Erfassung quantitativer Daten über die CO₂-Konzentrationen im Waldboden wurden photoakustische CO₂-Sensoren entwickelt. Diese Sensoren wurden unter besonderem Focus auf schnelle Ansprechzeit, kompakte Bauweise und Robustheit konzipiert und nutzen die Zweikammer-Photoakustik-Technologie. Dazu wurden FEM-Simulationen in den Bereichen Raytracing, Akustik und Diffusion durchgeführt. Das Sensormodul ist in einer Sonde mit einem Durchmesser von 19 mm integriert, um minimalinvasive Messungen in verschiedenen Bodentiefen durchführen zu können. Die Sonde ist robust, gegen Flüssigkeiten und Partikel abgedichtet und bietet eine Auflösung von 1 ppm im Messbereich von 0 bis 50 000 ppm.

Keywords: Indirekte Zweikammer-Photoakustik, Kohlenstoffdioxid CO₂, Bodensonde, Messlanze, Waldboden, Bodenatmung

Motivation

Zersetzungsprozesse und Wurzelatmung verursachen Kohlenstoffdioxid-(CO₂)-Flüsse zwischen dem Waldboden und der Atmosphäre. Quantitative Daten zu unterirdischen CO₂-Konzentrationen und zur Diffusivität des Bodens ermöglichen die Untersuchung des Zusammenhangs zwischen der Bodenatmung und dem oberirdischen Stressniveau der Pflanzen. Die Bodenatmung R_{soil} ist der CO₂-Fluss zwischen Boden und Atmosphäre. Diese setzt sich zusammen aus der autotrophen Atmung R_{SA} , welche aus dem Pflanzenstoffwechsel von Blättern und Wurzeln herrührt. Und der heterotrophen Atmung R_{SH} aus der Zersetzung akkumulierter Biomasse (siehe Abb. 1). Die neuen Messdaten werden dazu beitragen, skalierbare empirische oder mechanistische Funktionen der unterirdischen CO₂-Produktion und Emission (durch Atmung) zu liefern, die deren Rolle im Kohlenstoffhaushalt des Ökosystems beschreiben [1]. Dazu soll die natürliche unterirdische CO₂-Konzentration über die Jahreszeiten hinweg gemessen werden. Zur Ermittlung der Diffusivität des Bodens erfolgt eine Injektion eines definierten Volumens von 100 % CO₂ in einem definierten Abstand zur Sonde. Die Diffusivität ergibt sich rechnerisch aus der Zeit, die vergeht, bis das Gas von der Sonde gemessen wird. Dafür

bedarf es einer hinreichend kurzen Ansprechzeit des Sensors.

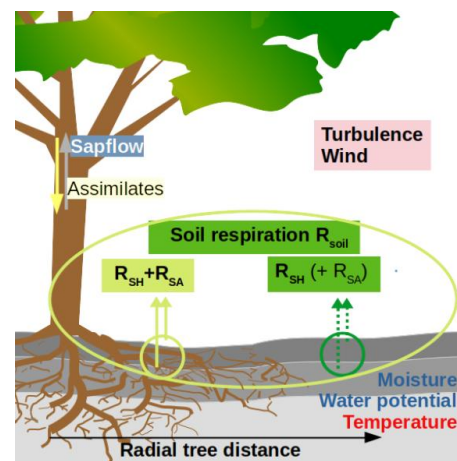


Abb. 1: Bodenatmung ist der CO₂-Austausch zwischen Boden und Atmosphäre. Sie besteht aus autotropher und heterotropher Atmung [1].

Die CO₂-Bodensonde Vaisala GMP343 gilt als Stand der Technik für Anwendungen im Boden. Sie ist für den Zweck zertifiziert und im Rahmen von ECOSENSE bereits im Einsatz – weist jedoch aufgrund ihrer langen Ansprechzeit und ihrer Größe Nachteile und Einschränkungen

auf. Für die Diffusivitätsabschätzung reagiert sie zu langsam und ihre Installation im Boden erfordert Aushubarbeiten, wodurch das zu untersuchende System zerstört wird. Deshalb werden schnelle, kleine und robuste CO₂-Sensoren auf Basis der Zweikammer-Photoakustik entwickelt (Abb. 2). Die Photoakustik ist ein weit verbreitetes, selektives infrarotspektroskopisches Messverfahren, das zur Messung von IR-aktiven Gasen in geringen Konzentrationen eingesetzt wird. Dies stellt einen neuen Ansatz in der Bodenmesstechnik dar und bietet potenziell hohe Empfindlichkeiten bei relativ geringem Energiebedarf und kompakten Abmessungen für minimalinvasive Installation [2, 3]. Photoakustische Sensoren sind zudem unempfindlich gegenüber Feuchtigkeitsschwankungen, und mögliche Störungen durch Temperaturänderungen lassen sich leicht mithilfe eines integrierten Temperatursensor als Referenz kompensieren.

Messprinzip

Der photoakustische (PA) Effekt wurde 1880 von Alexander Graham Bell entdeckt. Er beruht auf der Absorption von elektromagnetischer Strahlungsenergie durch das untersuchte Medium und deren Umwandlung in kinetische Energie der Moleküle. Diese führt zu einem Druckanstieg im absorbierenden Medium. Wird die anregende Strahlungsquelle nun moduliert, breitet sich im Medium eine thermische Druckwelle aus, die man als Schall mit einem Mikrofon messen kann. Die Schallfrequenz entspricht der Modulationsfrequenz der anregenden Strahlungsquelle. Die Amplitude des akustischen Signals ist direkt proportional zur Anzahl der absorbierenden Moleküle. Dadurch werden präzise Messungen von Gaskonzentrationen möglich [3]. Die durch ein Medium transmittierte Lichtintensität I wird durch das Lambert-Beer-Gesetz beschrieben. Sie nimmt entlang des Weges z exponentiell ab gemäß

$$I = I_0 \cdot \exp(-\sigma(\lambda) \cdot n \cdot z), \quad (1)$$

wobei I_0 die einfallende Strahlung, n die Teilchenanzahl des Zielgases und σ der Wirkungsquerschnitt des absorbierenden Moleküls ist [5]. Beim Zwei-Kammer-Prinzip, auch indirektes Prinzip genannt, bildet die erste Kammer den optischen-/ Absorptionspfad. Die erste Kammer ist nach außen hin geöffnet und durch eine Membran stellt sich diffusiv die Umgebungskonzentration ein. Je höher die CO₂-Konzentration ist, desto stärker wird gemäß Gl. (1) die Intensität der Strahlung abgeschwächt. Die zweite Kammer dient als photoakustischer Detektor. Sie ist durch ein aufgelötetes entspiegeltes Siliziumfenster (Si) hermetisch verschlossen und

zu 100 % mit dem Zielgas CO₂ gefüllt. Durch ein integriertes Mikrofon wird der PA-Effekt technisch nutzbar gemacht. Das Messprinzip ist somit invers und ähnelt dem eines Nichtdispersiven Infrarotsensors (NDIR), jedoch mit einer höheren Empfindlichkeit und intrinsischen Selektivität ohne zusätzlichen optischen Filter durch den photoakustischen Detektor.

Zur optischen Anregung von CO₂ bei 4,3 µm Wellenlänge wird die *nanoplus Mid-Infrared LED 4300* verwendet. Der schematische Aufbau ist in Abb. 2 dargestellt.

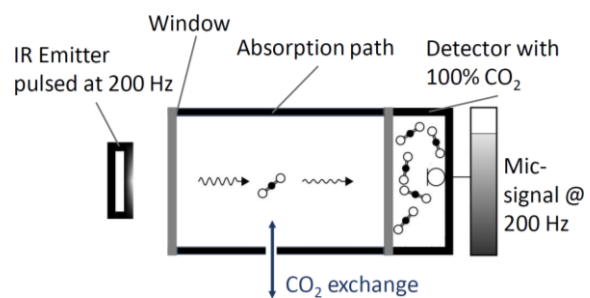


Abb. 2: Inverses Messprinzip bei der indirekten Zweikammer-Photoakustik. Die Messstrahlung des Emitters wird im Absorptionspfad proportional zur Umgebungskonzentration abgeschwächt. Die verbleibende Strahlungsintensität wird im Detektor photoakustisch gemessen. Das Signal ist invers-proportional zur CO₂-Konzentration.

Simulation

Die simulationsbasierte Kopplung von Raytracing, Strahlungsabsorption und Akustik in COMSOL Multiphysics bildet die gesamten photoakustischen Effekte ab. Ausgewertet wird die Höhe des akustischen Drucks (Schalldruck) im Detektor. Die *Thermoviscous Acoustic* Schnittstelle ist für diese Art von Anwendung gut geeignet,

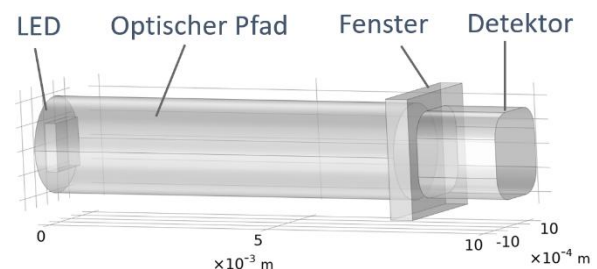


Abb. 3: Aufbau und Geometrie des COMSOL-Modells

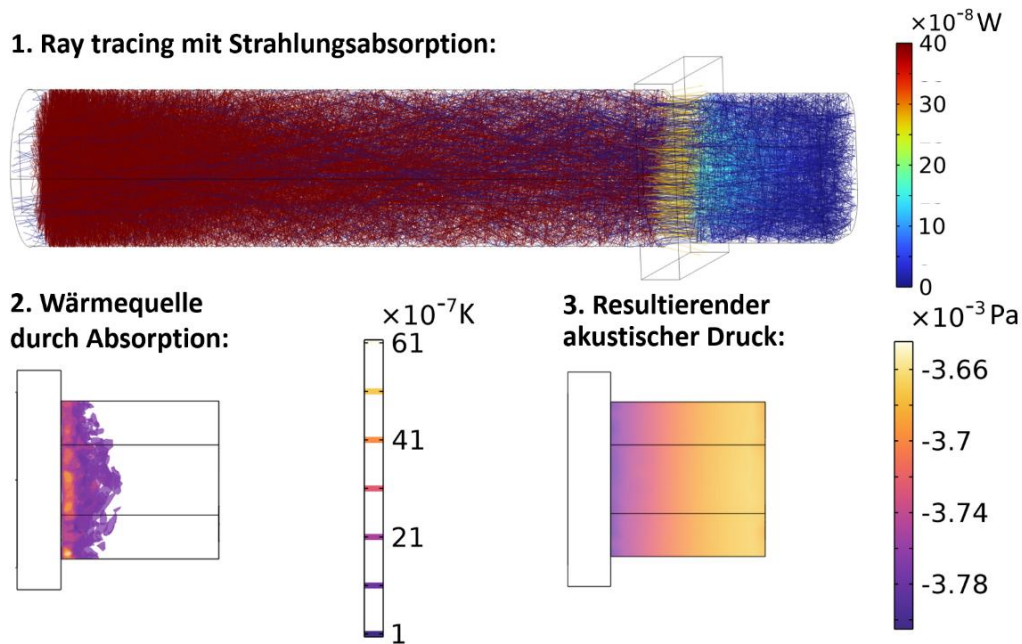


Abb. 4: Schritte der gekoppelten COMSOL-Simulation: (1) Ray Tracing mit Strahlungsabsorption und die daraus resultierende Wärmeverteilung (2) als Quelle für die akustische Anregung. (3) Der daraus resultierende akustische Druck bei 200 Hz als Maß für das Sensorsignal.

da sie die vollständigen linearisierten Navier-Stokes-Gleichungen für ein kompressibles Fluid einschließlich der Energiegleichung löst.

Abb. 3 zeigt, wie der Aufbau des Sensors in die Geometrie eines COMSOL-Modells überführt wurde. Modelliert wurden ausschließlich jene Komponenten, welche eine optische und / oder akustische Relevanz haben. Angefangen beim LED-Chip (1×1) mm, welcher in den verspiegelten optischen Pfad mit einer Länge von 8 mm hineinragt. Das 0,5 mm dicke Si-Fenster trennt die erste Kammer vom Detektorvolumen ($1,8 \times 1,8 \times 1,8$) mm. In letzterem wird ein Absorptionskoeffizient von 10^3 m^{-1} eingestellt, der dem von reinem CO_2 bei 1 bar und 20°C im Linienzentrum der $4,3 \mu\text{m}$ -Bande nachempfunden ist. Mithilfe der gekoppelten Simulation lassen sich verschiedenste optische oder geometrische Parameter variieren und der damit einhergehende Einfluss auf den akustischen Druck im Detektor, der als Maß für das Signal des realen Sensors steht, auswerten. Abb. 4 visualisiert die einzelnen Simulationsschritte. Angefangen mit Ray Tracing (1). Dabei treten 500 einzelne Strahlen mit einer optischen Gesamtleistung von 1 mW in zufälliger Richtung aus der Oberfläche des nachgebildeten LED-Chips aus. Die Farbe der einzelnen Strahlen ist ein Maß für die Intensität. Sie wandelt sich von initial rot zu blau, wenn Strahlungsenergie aufgrund von Absorption verloren geht. Beim Durchgang der Strahlen durch das Si-Fenster geht bereits ein Teil der Intensität verloren, was sich in einem Farbwechsel zu orange

wiederspielt. Im Detektorvolumen zeigt sich, dass die Intensität bereits nach Submillimeter- bis Millimeter-Weglängen stark zusammenbricht, also praktisch „sofort“ absorbiert wird und sich die Strahlen blau färben. Einige Strahlen werden in die erste Kammer rückreflektiert. Das daraus resultierende Wärmebild (2) bestätigt, dass die Energie direkt hinter dem Fenster dissipiert wird. Der resultierende akustische Druck (3) kann als Maß für Optimierung der Optik bzw. Geometrie herangezogen werden.

Detektor

Die Erkenntnisse aus den Simulationen haben zu einem Detektordesign beigetragen, bei dem die Tiefe der Kavität gemäß der „raschen“ Absorption bei Weglängen im Submillimeterbereich auf 0,7 mm festgelegt wurde.



Abb. 5: Aufgebauter photoakustischer Detektor, hier mit Kupferröhrchen zur Befüllung mit dem Zielgas CO_2

Abb. 5 zeigt den aufgebauten photoakustischen Detektor, bestehend aus einem vergoldeten Gehäuse mit Glasdurchführungen für die elektrische Kontaktierung des integrierten MEMS-Mikrofons *SPV1840LR5H* von *Knowles*. Oben ist ein am Rand metallisiertes Si-Fenster aufgelötet. Das Gehäuse ist mit angelöteten Kupferröhrchen ausgestattet. Über diese kann das Zielgas CO_2 eingefüllt werden. Das relativ weiche Kupfer kann mit einer Rundzange gequetscht und zusätzlich durch Verlöten der Enden hermetisch verschlossen werden. Alternativ ist zur Befüllung das Auflöten des Fensters direkt unter CO_2 -Atmosphäre möglich. Nach der Befüllung mit CO_2 wurde der Detektor spektral charakterisiert. Dazu wurde er in ein FTIR *Brüker Vertex V80* eingebaut. Abb. 6 zeigt das Fourier-transformierte Signal des Detektors in Abhängigkeit der Wellenlänge. Dabei zeigt sich die für CO_2 charakteristische Absorptionsbande um $4,3\ \mu\text{m}$ Wellenlänge und somit die intrinsische Selektivität des Detektors auf das Zielgas.

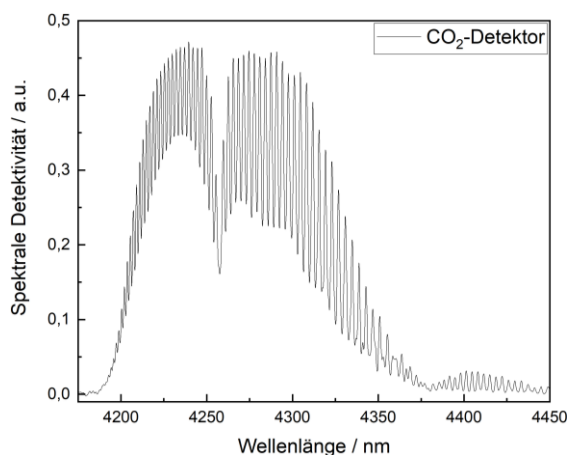


Abb. 6: Die spektrale Detektivität des photoakustischen Detektors gemessen mit FTIR-Spektrometrie. Die Schmalbandigkeit der Detektivität verspricht eine hohe Selektivität auf das Zielgas CO_2 .

Sensoraufbau

Abb. 7 zeigt den Aufbau des Sensormoduls, welches um weitere Module oder Rohre mit M17 Innen- bzw. Außengewinden zu einer lanzenförmigen Sonde erweitert werden kann. Besonderes Augenmerk bei der Entwicklung lag auf einer kurzen Ansprechzeit. Für eine schnelle diffusive Befüllung wurde das innere Gasvolumen klein gehalten. Gleichzeitig sollte die „aktive“ Membranfläche außen groß sein und das Gas über Drainagekanäle effektiv nach innen zur Messstrecke geleitet werden. Eine

Membran kann außen über den Drainagen aufgebracht werden. Abb. 8 veranschaulicht die innenliegende Geometrie, mit den sechs Entlüftungsöffnungen, die die Drainagen und den 8 mm langen optischen Pfad miteinander verbinden.

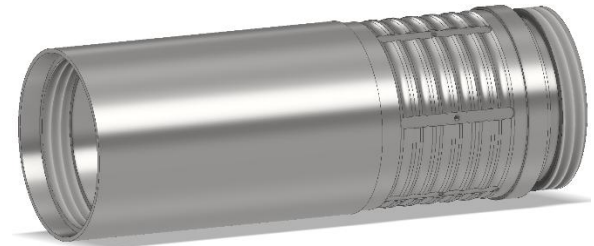


Abb. 7: Einzelnes Sensormodul. Radiale Drainagekanäle führen das Gas zu sechs Entlüftungsöffnungen, die mit der Messstrecke im Inneren verbunden sind.

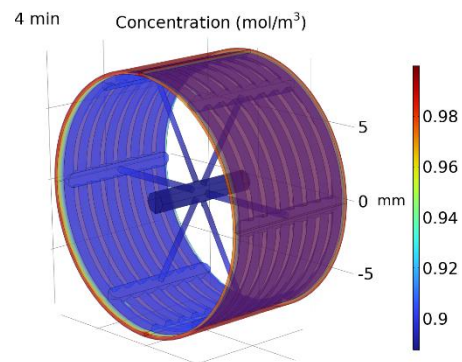


Abb. 8: Veranschaulichung der innenliegenden Geometrie. Die Diffusionssimulation zeigt die Konzentration in den Drainagen und Entlüftungsöffnungen zum optischen Pfad.

Zur Abschätzung der Ansprechzeit des Sensors wurde mittels COMSOL Multiphysics eine Simulation der Diffusion durchgeführt. An der Außenfläche der Sonde wurde eine initiale Konzentration von $1\ \text{mol/m}^3$ festgelegt und im inneren Volumen eine Konzentration von null angenommen. Bei der Materialauswahl für einen Filter bzw. eine Membran sollte nach Möglichkeit auf PTFE – weit verbreitetes Filtermaterial – verzichtet werden, da es sich um eine schwer abbaubare Chemikalie handelt. Deshalb wird zunächst Silikon (PDMS) evaluiert, da es im Vergleich zu anderen Polymeren eine relativ hohe selektive Durchlässigkeit für CO_2 aufweist. Abb. 9 zeigt das simulierte zeitliche Verhalten der relativen Konzentration auf der Messstrecke ohne Filter sowie mit einem 0,25 mm dicken Silikonfilter mit einer Diffusionskonstante von $D = 6,88 \times 10^{-6}\ \text{cm}^2/\text{s}$ (Literaturwert

[6]). Ohne Filter wird eine T_{90} -Ansprechzeit von einer halben Minute beobachtet. Mit dem Silikonfilter erhöht sich die Ansprechzeit auf ca. 4 Minuten. Dies wäre für eine langfristige Konzentrationsmessung über die Jahreszeiten hinweg ausreichend. Allerdings würde es zu lange dauern, die Diffusionsrate des Bodens anhand der Reaktionszeit nach einer externen Injektion von reinem CO_2 in den Boden zu schätzen, weshalb derzeit weitere Membranen untersucht werden.

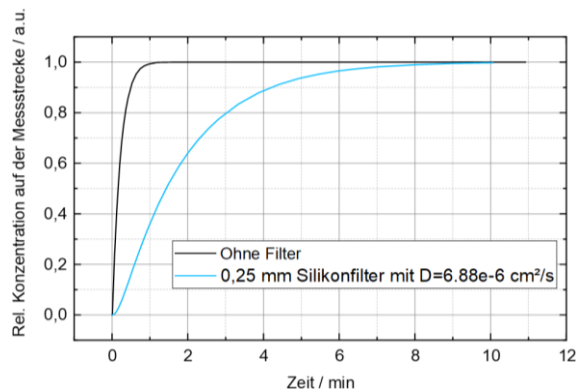


Abb. 9: Simulation der diffusiven Befüllung der ersten Kammer (optischer Pfad / Messstrecke) als Maß für die Ansprechzeit des Sensors.

Für eine Labormessung wurden LED und Detektor in ein Sensormodul aus Aluminium (wie in Abb. 7) integriert.

Messergebnis

Über Massenflussregler wurden unterschiedliche CO_2 -Konzentrationsstufen eingestellt, um den Sensor zu beaufschlagen. Reines CO_2 wurde entsprechend mit Stickstoff N verdünnt.

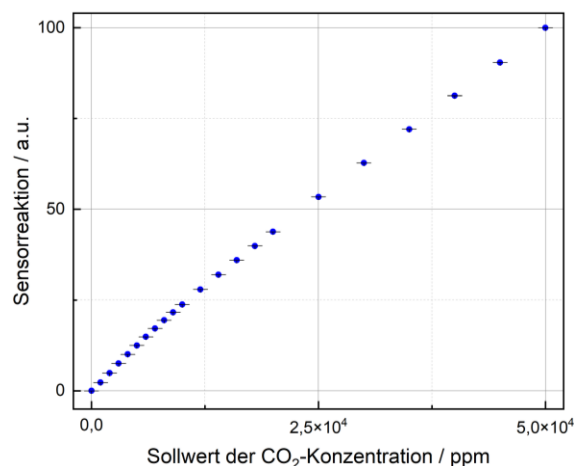


Abb. 10: Sensorreaktion auf unterschiedliche CO_2 -Konzentrationen von null bis 50 000 ppm.

Der Sensor wurde in einem geschlossenen Gehäuse vom Messgas bei einem Gesamtfluss von 1 L/min umströmt. Der Sensor wurde betrieben mit der Lock-in Zeitkonstante von 1 s, der LED-Modulationsfrequenz von 200 Hz und dem LED-Strom von 100 mA. Abb. 10 zeigt die Sensorreaktion für die beaufschlagten Konzentrationsstufen von null bis 50 000 ppm. Die Fehlerbalken zeigen die jeweilige Standardabweichung des Sensorsignals.

Fazit

Die Photoakustik verhilft zu einer robusten, kompakten und empfindlichen CO_2 -Boden-sonde. Mit einem Messbereich von null bis 50 000 ppm deckt der Sensor den relevanten Messbereich für die Anwendung ab. Die gesamte Prozesskette kann mit COMSOL Multiphysics simuliert und besser verstanden werden. Die Simulationen ermöglichen Optimierungen am Sensoraufbau und der Performance im Hinblick auf Empfindlichkeit und Ansprechzeit. Wir sind auf dem Weg zu einem Sensor, der zuverlässige Daten über die CO_2 -Konzentration im Erdboden liefert. Künftig können mit den Daten Zusammenhänge von oberirdischem Stress der Pflanzen und Bäume mit der Bodenatmung gezielter korreliert werden.

Danksagung

Wir danken der Deutschen Forschungsgemeinschaft für die Förderung des Sonderforschungsbereichs SFB 1537 „ECOSENSE“ (Projekt-ID: 459819582).

Literaturnachweis

- [1] Werner C, Wallrabe U, Christen A, Comella L, Dormann C, Göritz A, Grote R, Haberstroh S, Jouda M, Kiese R, Koch B, Korvink J, Kreuzwieser J, Lang F, Müller J, Prucker O, Reiterer A, Rühle J, Rupitsch S, Schack-Kirchner H, Schmitt K, Stobbe N, Weiler M, Woias P, Wöllenstein J (2024) ECOSENSE - Multi-scale quantification and modelling of spatio-temporal dynamics of ecosystem processes by smart autonomous sensor networks. Research Ideas and Outcomes 10: e129357. <https://doi.org/10.3897/rio.10.e129357>
- [2] Schjolberg-Henriksen, K., et al. „Sensitive and Selective Photoacoustic Gas Sensor Suitable for High-Volume Manufacturing“. *IEEE Sensors J.* 8 (9), pp. 1539–1545, 2008.
- [3] Bozóki, Z., et al. „Photoacoustic Instruments for Practical Applications: Present, Potentials, and Future Challenges“. *Applied Spectroscopy Reviews* 46 (1), pp. 1–37, 2011 [3] Reference 3
- [4] M. Wolff, B. Kost und B. Baumann, „Shape-Optimized Photoacoustic Cell: Numerical Consolidation and Experimental Confirmation,“

- International Journal of Thermophysics, Jg. 33, Nr. 10-11, S. 1953–1959, 2012, issn: 0195-928X. doi: 10.1007/s10765-012-1257-2
- [5] W. Demtröder, Laser Spectroscopy 1: Basic Principles, 5. 5th ed. 2014. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014, ISBN: 978-3-642-53858-2. Adresse: <http://nbn-resolving.org/urn:nbn:de:bsz:31-e-pflicht-1488598>
- [6] Thomas M. DeSutter, Thomas J. Sauer, Timothy B. Parkin, Porous tubing for use in monitoring soil CO₂ concentrations, Soil Biology and Biochemistry, Volume 38, Issue 9, 2006, Pages 2676-2681, ISSN 0038-0717, <https://doi.org/10.1016/j.soilbio.2006.04.022>.

Resonanter Photoakustik-Sensor für die Detektion von Isopren im mittleren Infrarotbereich

Sebastian Schnell¹, Carl Hild¹, Hojin Lee¹, Stefan Hugger², Katrin Schmitt^{1,3}, Jürgen Wöllenstein^{1,3}

¹ Institut für Mikrosystemtechnik – IMTEK, Universität Freiburg 79110 Freiburg i. Br.

² Fraunhofer Institut für angewandte Festkörperphysik IAF, 79110 Freiburg i. Br.

³ Fraunhofer Institut für physikalische Messtechnik IPM, 79110 Freiburg i. Br.

Zusammenfassung

In diesem Beitrag wird ein resonantes photoakustisches Laserspektrometer zur Detektion von Isopren im Spurenbereich und dessen Verhalten bei variabler Feuchte vorgestellt. Als Lichtquelle dient ein Quantenkaskadenlaser mit externem Resonator (External Cavity, EC-QCL) im Wellenzahlbereich von 820–980 cm⁻¹, während die akustische Detektion in einer zylindrischen Resonanzzelle mit einer Resonanzfrequenz von 5.32 kHz erfolgt. Die Analyse des Rauschverhaltens mittels Allan-Werle-Analyse zeigt, dass die Nachweisgrenze des Systems bis zu Integrationszeiten von etwa 120 s durch weißes Rauschen dominiert wird. Daraus ergibt sich bei einer Integrationszeit von 100 ms eine Nachweisgrenze (LOD) von 2.7 ppb. Zur Untersuchung des Einflusses von Wasserdampf wurden Isopren-Spektren im Bereich von 880–900 cm⁻¹ bei unterschiedlicher relativer Feuchte aufgenommen. Dabei wurden insbesondere die Absorptionsbanden ν_{27} und ν_{28} betrachtet. Während die ν_{28} -Bande bei 906.15 cm⁻¹ eine deutliche spektrale Überlagerung mit der Absorption von Wasser zeigt, ist die ν_{27} -Bande bei 893.72 cm⁻¹ weitgehend frei von Wasserinterferenzen. Kalibrationsmessungen bei unterschiedlichen Luftfeuchten zeigen, dass die Empfindlichkeit bei ν_{27} unabhängig von der Feuchte ist, während bei ν_{28} eine Abnahme der Sensitivität mit steigender Luftfeuchtigkeit beobachtet wird. Die Ergebnisse zeigen, dass die ν_{27} -Bande für den quantitativen Isopren-Nachweis unter realistischen, feuchtevariablen Bedingungen deutlich besser geeignet ist und dass eine aufwändige Trocknung der Gasproben vermieden werden kann, sofern der Wasserdampfgehalt bekannt ist. Das vorgestellte System stellt damit eine kompakte und kosteneffiziente Alternative zu etablierten Methoden wie GC-MS oder PTR-ToF-MS dar und eignet sich insbesondere für Anwendungen in der Atmosphärenchemie, Pflanzenphysiologie und Atemgasanalyse.

Keywords: Photoakustik, Isopren, C₅H₈, Spektroskopie, Atemgasanalyse, Atmosphärenchemie, Feuchte

Einleitung

Flüchtige organische Verbindungen (Volatile Organic Compounds, VOCs) bezeichnen eine signifikante Gruppe chemischer Verbindungen, die durch eine hohe Flüchtigkeit bei Raumtemperatur und eine hohe Reaktivität gekennzeichnet sind. Die wichtigsten VOCs gehören zu den Gruppen der Alkane, Alkene, Benzene, Terpene, organischen Säuren, Carbonyle und Alkohole [1]. Aufgrund ihrer chemischen Reaktivität spielen sie eine wichtige Rolle bei den chemischen Prozessen in der Atmosphäre. VOCs tragen zur Bildung von Ozon in der Troposphäre sowie zur Bildung sekundärer organischer Aerosole (SOA) bei [2], die als Kondensationskeime für die Wolkenbildung wirken und damit indirekt den Energiehaushalt von Ökosystemen beeinflussen. Da VOCs in der Atmosphäre mit Sauerstoff reagieren, können größere Mengen die gesamte Oxidationskapazität der

Atmosphäre reduzieren. Diese Reduktion führt zu längeren mittleren Lebensdauern des Treibhausgases Methan, welches dadurch länger in der Atmosphäre verbleibt. Über diesen Prozess tragen VOCs indirekt zur globalen Erwärmung bei [3].

Die jährlichen VOC-Emissionen weltweit betragen schätzungsweise 1150 Tg pro Jahr [4] und setzen sich aus anthropogenen Emissionen, hauptsächlich durch die Verbrennung fossiler Brennstoffe [5] sowie aus biogenen Quellen zusammen. Der Hauptteil stammt aus biogenen Quellen, auf die etwa 90 % der gesamten in die Atmosphäre emittierten VOCs entfallen [1, 4]. Unter den biogenen VOC-Quellen sind die Vegetation, insbesondere Wälder, von größerer Bedeutung.

Mit geschätzten 311 bis 637 Tg pro Jahr ist Isopren global das am häufigsten vorkommende und emittierte BVOC (Bio-VOC), mit etwa 50 % Anteil der gesamten BVOCs [1,

6, 7]. Durch Reaktionen mit NO_x -Verbindungen in der Atmosphäre trägt Isopren erheblich zur Bildung von troposphärischem Ozon bei, wodurch es zu Beeinträchtigung der Luftqualität auf regionaler Skala kommen kann [4, 7–9]. Die Quantifizierung der Isopren-Konzentrationen in der Luft ist daher für die Atmosphärenchemie von großem Interesse.

Isopren wird in großen Mengen von den Blättern vieler Pflanzenarten, abhängig von Beleuchtung und Temperatur, emittiert [4, 10]. Es entsteht in den Chloroplasten der Mesophyllzellen der Blätter über den 2-C-Methyl-D-erythritol-4-phosphat-(MEP)-Stoffwechselweg, der direkt mit dem Calvin-Zyklus gekoppelt ist und Vorläuferstoffe für die Isopren-Biosynthese liefert. Die Funktion der Isopren-Biosynthese und deren Emission werden seit Jahrzehnten untersucht. Es wurden mehrere Funktionen identifiziert, die hauptsächlich mit der Stressabwehr von Pflanzen zusammenhängen. Es wurde gezeigt, dass Isopren (i) die Toleranz der Photosynthese gegenüber hohen Temperaturen und Trockenheit erhöht, (ii) oxidativen Stress, beispielsweise durch Ozon hervorgerufen, mindert und (iii) als Botenstoff fungiert, beispielsweise in tritrophischen Interaktionen von Pflanzen mit Herbivoren und deren Parasiten [11]. Neuere Arbeiten zeigen, dass die Wirkung von Isopren möglicherweise nicht unmittelbar erfolgt, sondern auf seiner Funktion als Signalmolekül beruht. Isopren kann Genexpressionsmuster verändern und mehrere zelluläre Signalwege beeinflussen, welche an der Stressanpassung von Pflanzen beteiligt sind [11]. Vom Standpunkt der Ökologie und der pflanzlichen Stressphysiologie besteht ein großer Bedarf an genauen Bestimmungen der Isoprenemissionen aus Blättern und anderen

Pflanzenteilen.

Auch Atemgasanalyse von Isopren gewinnt als nichtinvasive Methode zur Untersuchung physiologischer und metabolischer Prozesse im menschlichen Körper zunehmend an Bedeutung. Isopren entsteht als Nebenprodukt bei der Verstoffwechselung von Cholesterin und wird über die Lunge ausgeatmet, wobei seine Konzentration im Atem typischerweise im ppb-Bereich liegt [12]. Charakteristisch ist eine deutliche Abhängigkeit von Faktoren wie körperlicher Aktivität, Durchblutung und Stoffwechselrate, was sich beispielsweise in einem starken Anstieg der Isopren-Konzentration nach körperlicher Belastung zeigt [13]. Aufgrund dieser Dynamik wird Isopren als potenzieller Biomarker für kardiovaskuläre Prozesse diskutiert [14]. Die Herausforderung liegt dabei in dem hochempfindlichen und zeitaufgelösten Nachweis, da die Konzentrationen gering sind und schnelle physiologische Änderungen abgebildet werden müssen.

Die Quantifizierung von Isopren in Pflanzenemissionen, in der Atmosphäre oder bei der Atemgasanalyse erfolgt in der Regel in Gasproben mit variierendem Feuchtegehalt. In dieser Arbeit wird ein photoakustisches Spektrometer vorgestellt, das über eine ausreichend hohe Empfindlichkeit für den Nachweis von Isopren im ppb-Bereich besitzt und zugleich eine zuverlässige Messung auch bei schwankender Feuchte der Gasproben ermöglicht.

Versuchsaufbau

Der Versuchsaufbau ist in Abb.1 schematisch dargestellt. Als Lichtquelle wurde ein

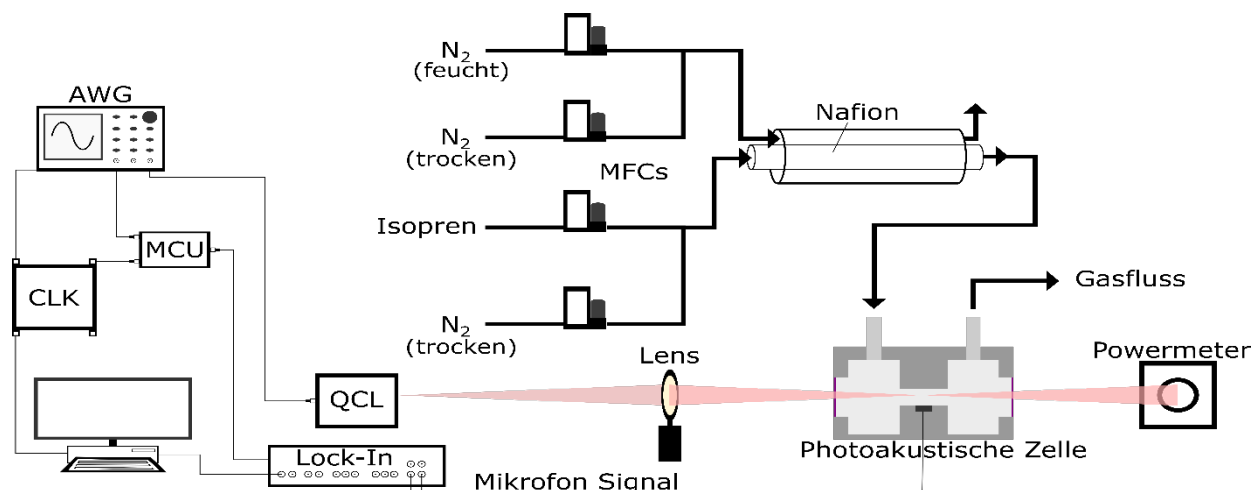


Abb. 1: Schematischer Aufbau des photoakustischen Spektrometers. Als Lichtquelle dient ein QCL mit externem Resonator, dessen Licht durch die Zelle fokussiert wird.

Quantenkaskadenlaser mit externem Resonator eingesetzt (EC-QCL). Dieser wird gepulst betrieben mit einem Tastverhältnis von 10% (1 MHz, 100 ns) und bietet einen Abstimm-bereich von 820 cm^{-1} bis 980 cm^{-1} bei einer Linienbreite von 0.3 cm^{-1} . Der QCL sowie die Kavität wurden am Fraunhofer Institut für angewandte Festkörperphysik entwickelt und aufgebaut. Der Resonator wurde in Litrow-Anordnung aufgebaut und erlaubt, durch ein drehbares Beugungsgitter, eine kontrollierte Rückkopplung. Das Beugungsgitter verfügt über 150 Linien pro Millimeter und ist auf einem Piezo-Slip-Stick betriebenen Drehtisch montiert, welcher eine Auflösung von $25\text{ }\mu^\circ$ besitzt. Das aus dem Medium rückseitig austretende Laserlicht wird am Gitter gebeugt und wieder in das Medium eingekoppelt. Die wieder eingekoppelten Photonen tragen verstärkend zur stimulierten Emission bei und verringern dadurch die Linienbreite. Da der Beugungswinkel von Photonen wellenlängen-abhängig ist, kann die emittierte Wellenlänge gesteuert werden, indem der entsprechende Winkel eingestellt wird.

Die akustische Resonanzzelle ist aus Aluminium gefertigt und in Form eines zylindrischen Resonators mit einer Länge von 30 mm und einem Durchmesser von 3 mm. An beiden Enden des Resonators sind zylindrische Puffervolumina mit einer Länge von 15 mm und einem Durchmesser von 12 mm angebracht. Die Puffervolumina sind durch Antireflex (AR) beschichtete Germanium-Fenster abgeschlossen. Das akustische Signal in der Mitte des Resonators von einem MEMS-Mikrofon (ICS-40720, TDK InvenSense) aufgenommen. Durch die mittige Positionierung des Mikrofons wird der Signalbeitrag der akustischen Grundmode maximiert.

Der QCL wird durch einen Signalgenerator (AWG, Tektronik AFG3022C) gepulst mit einer Frequenz von 1 MHz und einem duty-cycle von 10 %. Das Signal des AWG wird zusätzlich mit einer Rechteckfunktion moduliert, welche von der Mikrocontroller Einheit (MCU) erzeugt wird. Diese Modulation wird auf die Resonanzfrequenz der Zelle von 5.32 kHz abgestimmt, um eine akustische Signalverstärkung zu erreichen. Neben dem AWG wird diese Modulation dem Lock-In Verstärker (LIV) übergeben. Das Signal des LIV wird, per COM-Port, über ein Python Programm ausgelesen. Über ein Clockboard (Skyworks Si5332-6EX-EVB REV2.0) werden die Geräte miteinander synchronisiert.

Durch die Massendurchflussregler (MFC, Bronkhorst EL-FLOW Serie) kann 1 ppm Isopren (Westfalen AG) mit Stickstoff (Linde,

Reinheitsgrad 5.0) verdünnt werden. Dieses Gasgemisch wird durch einen Nafion-Schlauch in die Zelle eingeleitet. Das Äußere der Nafion-Membran kann mit einem Gemisch aus feuchtem und trockenem Stickstoff gespült werden. Dadurch kann das Isopren-Stickstoff-Gemisch im Inneren des Nafion-Schlauchs gezielt angefeuchtet werden. Das Feuchtelevel wird mittels Feuchtesensor (Sensor-Tec UFT75-ST2) bestimmt.

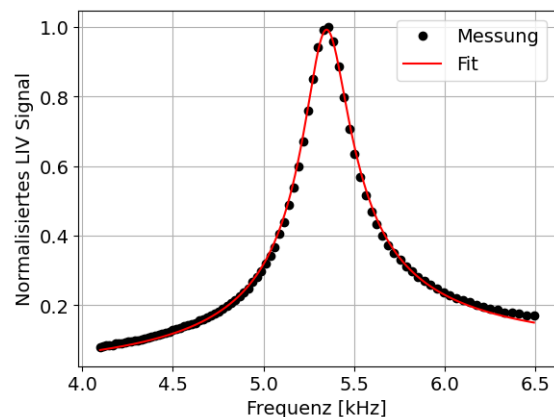


Abb.2: Messung zur Bestimmung der Resonanzfrequenz des akustischen Resonators. Das Signal wurde auf ein Maximum von 1 normiert. Die rote Linie zeigt einen Fit welcher auf Gl.(1) und Gl.(2) beruht.

Charakterisierung des Systems

Zunächst wurde die Resonanzfrequenz des akustischen Resonators bestimmt. Dazu wurde die Zelle mit 20 ppb Isopren bei 300 ml/min gespült und das LIV-Signal gemessen, während am MCU verschiedene Frequenzen eingestellt wurden. Die Messdaten sind in Abb. 2 dargestellt.

Zur Bestimmung der Resonanzfrequenz und der Güte Q der Messzelle betrachten wir die Amplitudenantwort eines angeregten, gedämpften harmonischen Oszillators [15]:

$$A = \frac{A_0}{\sqrt{\left(1 - \left(\frac{f}{f_0}\right)^2\right)^2 + \left(\frac{f}{f_0 Q}\right)^2}} \quad (1)$$

Hierbei ist f die Frequenz f_0 die Resonanzfrequenz der Grundmode.

Zusätzlich betrachten wir die einen Hochpassfilter, welcher als Instrumentenfunktion für das Mikrofon fungiert [16]:

$$|H_{hp}| = \frac{f/f_c}{\sqrt{1 + \left(\frac{f}{f_c}\right)^2}} \quad (2)$$

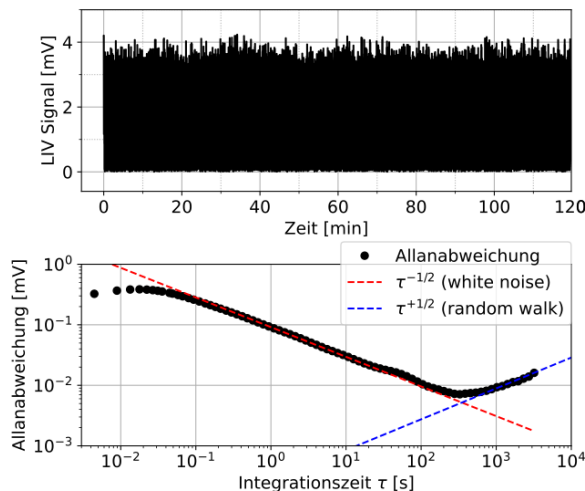


Abb. 3: Messung zur Bestimmung der Allan-Abweichung über eine Dauer von 120 min. Die rote Linie dient zum Vergleich mit weißem Rauschen, während die blaue Linie die Charakteristik von Random-Walk-Rauschen zeigt.

wobei f_c die Cut-Off-Frequenz des Mikrofons ist und laut Datenblatt 75 Hz beträgt. Zur Bestimmung der Güte Q und Resonanzfrequenz f_0 fitten wir eine Funktion der Form $A_{tot} = A_{Zelle} \cdot A_{Mikro} \cdot |H_{hp}|$ an die Messdaten an. Die Resonanzfrequenz des Mikrofons ist dabei laut Datenblatt 20 kHz. Wie Abb. 2 zeigt besteht eine große Übereinstimmung zwischen den Messdaten und dem Fit, was sich auch durch $R^2 = 0.999$ zeigt. Demnach ist der Gütefaktor $Q = 20.6$ und die Resonanzfrequenz $f_0 = 5.32$ kHz.

Ein weiterer Aspekt, der charakterisiert, wurde ist das Rauschverhalten und die Stabilität des Messsystems. Dazu wurde unter konstantem Stickstofffluss das LIV-Signal für 120 min gemessen. Dabei wurde die Integrationszeit des LIV auf 30 μ s reduziert. Ein weiterer Aspekt der charakterisiert wurde ist das Rauschverhalten und Stabilität des Messsystems. Dazu wurde unter konstantem Stickstofffluss das LIV-Signal für 120 min gemessen. Hierfür wurde die Integrationszeit

des LIV auf 30 μ s reduziert, um das Rauschen nicht durch Mittelungseffekte zu beeinflussen. Aus den Messdaten wurden Allan-Varianzen berechnet und gegen ihre Integrationszeit τ aufgetragen, siehe Abb. 3.

Die Allan-Abweichungen [17, 18] zeigen einen typischen Verlauf für weißes Rauschen von $\tau^{-1/2}$ bis zu einer Integrationszeit von über 100 s. Die Abweichung am Anfang zur roten Linie, und damit zu weißem Rauschen, ist typisch für Tiefpassfilter, welcher hier durch die endliche Mittelungszeit des LIV zum Tragen kommt. Bei etwa 120 s Integrationszeit befindet sich ein lokales Minimum in der Allan-Abweichung, die anschließend analog zu $\tau^{1/2}$ ansteigt. Dies deutet zwar auf ein Random-Walk-Rauschen hin, doch würde ein Beitrag durch lineare Drift zu einer deutlich steileren Steigung führen. Das Messsignal unterliegt also keiner signifikanten Drift

Isoprenspektrum und Wassereinfluss

Um den Einfluss von Wasserdampf auf das photoakustische Signal zu untersuchen, wurde eine Reihe von Spektren gemessen. Bei einem konstanten Gasfluss von 300 ml/min durch die Zelle wurde einmal ein reines Isopren-Spektrum und ein reines H₂O-Spektrum aufgezeichnet. Dabei wurde von 880 cm^{-1} bis 900 cm^{-1} in 0.01er Schritten die Wellenzahl durchgestimmt und mit einer Integrationszeit von 100 ms gemessen. Die Isopren-Messung erfolgte mit 167 ppb und die H₂O-Messung bei 57% rF. Anschließend wurde eine weitere Messung mit 167 ppb Isopren und 57% rF durchgeführt. Die Messungen sind in Abb. 4 dargestellt.

Der obere Plot zeigt die Einzelmessungen, wohingegen der untere Graph aus der feuchten Isopren-Messung besteht, von der die beiden Einzelmessungen subtrahiert werden. Im oberen Plot sind die beiden Absorptionsbanden ν_{27} und ν_{28} deutlich aufgelöst, sowie mehrere Signaturen, die von H₂O stammen. Die Absorbanzen stimmten in Form und Position mit der HITRAN Datenbank überein [19]. Dabei ist klar zu erkennen, dass die ν_{28} Bande bei 906.15 cm^{-1} von einer Wasserabsorbanz überlagert wird, während die ν_{27} Bande bei 893.72 cm^{-1} keinen signifikanten Überlapp mit dem Wasser vorweist.

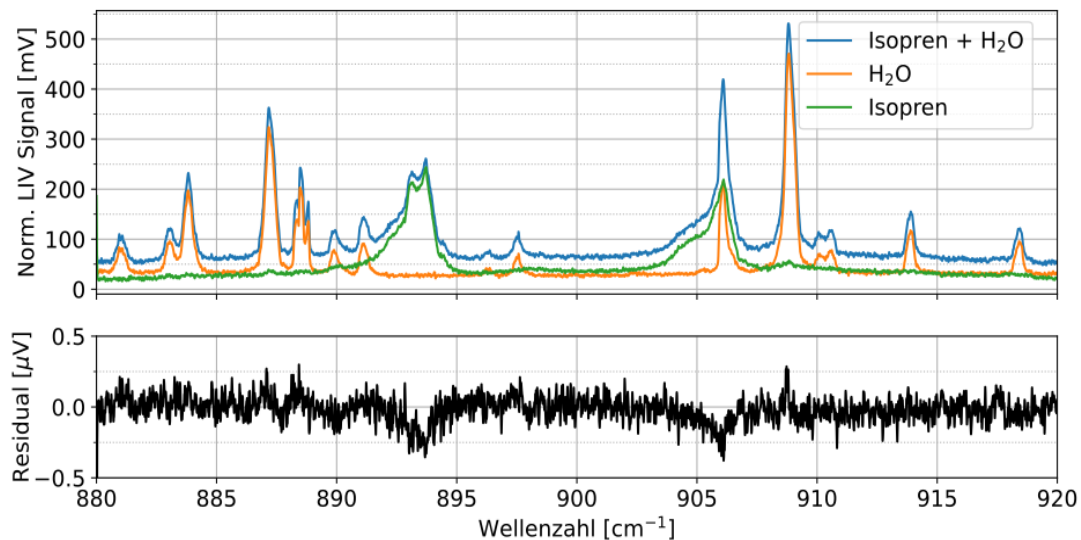


Abb. 4: Messungen eines Isopren-Spektrums bei 167 ppb Isopren (grün), einer Wasserdampfmessung bei 57 % rF bei Raumtemperatur (orange), sowie eine Messung mit 167 ppb Isopren bei 57 % rF (blau). Der untere Plot zeigt die kombinierte Messung (blau) abzüglich der Einzelmessungen (orange und grün).

In der Photoakustik kann die Befeuchtung von Gasproben je nach Material die Empfindlichkeit erhöhen, da Wassermoleküle als Kollisionspartner wirken und die Relaxationszeit des zu detektierenden Gases verkürzen können. Der Residualplot in Abb. 4 zeigt, dass dies hier nicht der Fall ist. Ansonsten würde eine stoßgetriebene Erhöhung der Empfindlichkeit zu deutlichen Maxima im Residualplot führen.

Dies lässt darauf schließen, dass Isopren-Messungen bei ν_{27} besser geeignet sind als Messungen bei ν_{28} aufgrund der deutlich geringeren Überlagerung mit der Wasserabsorbanz bei ν_{27} .

Kalibrierung und Nachweisgrenze

Wir führten für $\nu_{27} = 892.72 \text{ cm}^{-1}$ und $\nu_{27} = 906.15 \text{ cm}^{-1}$ eine Reihe von Kalibrationen durch, bei welchen die Isopren- und Wasserkonzentration der Gasmischung variiert wurden. Es wurden 5 Isopren-Konzentrationen bei jeweils 4 verschiedenen Wasserkonzentrationen in die Zelle gegeben. Bei fest eingestellter Wellenzahl wurde das LIV-Signal aufgezeichnet, wobei eine Integrationszeit von 100 ms eingestellt war. Die Messung ist in Abb. 5 dargestellt.

Für jede Isoprenstufe wird das letzte Drittel der Stufe gemittelt und zur Erstellung einer feuchtespezifischen Kalibration verwendet. Wir beschränkten uns dabei auf das letzte Drittel

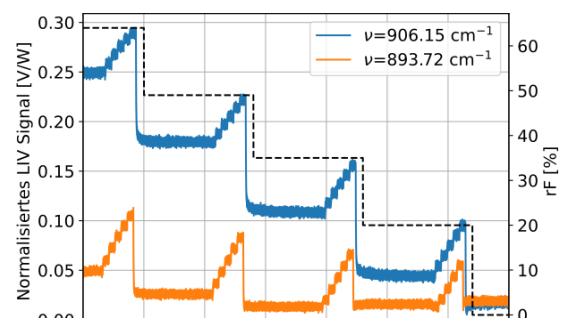


Abb.5: Der Graph zeigt eine Kalibrationsmessung, bei der das LIV-Signal für verschiedene Konzentrationen von Wasser und Isopren aufgezeichnet wurde. Das LIV-Spannungssignal wurde auf die gemessene Laserleistung normiert. Die beiden Messungen wurden nacheinander aufgenommen.

der Konzentrationsstufen, um Artefakte durch den Konzentrationswechsel in der Zelle zu vermeiden. Die Mittelwerte der Isopren-Stufen und deren Standardabweichung wurden gegen

ihre eingestellte Konzentration aufgetragen, siehe Abb. 6. Ein linearer Fit der Form $f(x) = ax + b$ wurde für jedes Feuchteniveau auf beiden Banden durchgeführt und dem Graphen hinzugefügt. Darüber hinaus ist eine Übersicht der gefitteten Parameter in Tab. 1 gelistet.

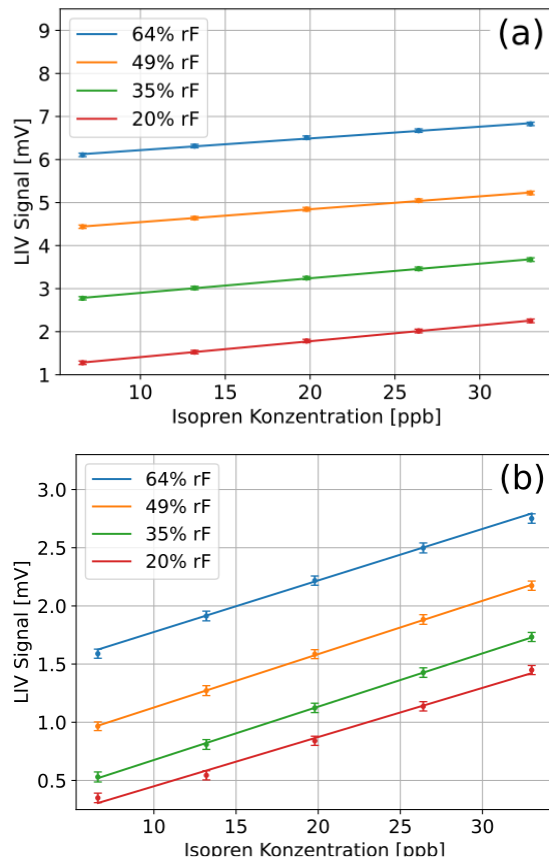


Abb.6: Die Graphen zeigen Mittelwerte die aus der Messung von Abb.5 erstellt wurden gegen die eingestellte Isoprenkonzentration. Es wurde für eine Kalibrationsgerade für jeden eingestellten Feuchtewert erstellt. Der Graph (a) zeigt die Kalibration die auf auf der v_{28} Bande gemessen während (b) die Kalibration der v_{27} Bande zeigt.

Es zeigt sich durch die Geradensteigungen a in Tab.1 dass sich bei der v_{27} die Steigungen der Kalibrationsgeraden sich kaum unterscheiden und innerhalb einer Standardabweichung zueinander liegen. Die Empfindlichkeit des Messsystems, die der Steigung der Kalibrationsgeraden entspricht, ist demnach unabhängig von der Wasserkonzentration. Der y-Achsenabschnitt b nimmt mit sinkender Feuchte ab, was bedeutet, dass das Wasser hier einen Signalbeitrag leistet. Das ist ebenso bei v_{28} der Fall, jedoch ist hier der Signalbeitrag deutlich höher. Hier zeigt sich außerdem, dass die Empfindlichkeit mit sinkender Feuchte steigt, was ein Hinweis auf veränderte Relaxationsdynamik durch Wassermoleküle als Stoßpartner ist. Demnach sind die Empfindlichkeit und Querempfindlichkeit von Isopren zu Wasser bei der v_{27} Bande deutlich

besser als bei v_{28} . Dies spiegelt sich auch in der Nachweisgrenze wider, die hier die Standardabweichung des Messsignals σ übersteigt:

$$LOD = 3 \cdot \sigma / a. \quad (3)$$

Aus Gl. (3) ergibt sich ein LOD von 2.65 ppb für v_{28} und ein LOD zwischen 4.35 ppb und 3.30 ppb für v_{28} abhängig von der Feuchte des Messgases. Da die Allan-Abweichung gezeigt

Tab.1: Auflistung der Parameter der in Abb.6 dargestellten Geraden-Fits.

	rF [%]	R ²	A [μ V/ppb W]	b [μ V/W]
v_{27}	64	0.999	44.3 ± 1.0	1332.8 ± 20
	49	0.999	45.8 ± 0.3	668.6 ± 8.2
	35	0.999	45.7 ± 0.5	218.2 ± 10.4
	20	0.994	42.3 ± 1.9	27.7 ± 42.4
v_{28}	64	0.996	27.2 ± 1.0	5946.1 ± 21.1
	49	0.999	30.0 ± 0.5	4246.6 ± 10.0
	35	0.999	34.0 ± 0.5	2561.6 ± 11.0
	20	0.999	36.9 ± 0.4	1040.8 ± 9.6

hat, dass das System bis 120 s Integrationszeit von weißem Rauschen dominiert wird, erwarten wir dort den geringsten möglichen LOD. Da weißes Rauschen mit $\tau^{-1/2}$ skaliert erwarten wir durch eine Integrationszeit von 120 s eine Rauschdeduktion um den Faktor 34. Auf den LOD der v_{27} Bande angewandt, bedeutet das einen geringsten möglichen LOD von 77 ppt bei einer Integrationszeit von 120 s.

Zusammenfassung und Ausblick

In diesem Beitrag wird ein resonantes photoakustisches Laserspektrometer zur Detektion von Isopren beschrieben. Das Resonanzverhalten des akustischen Resonators entspricht der Erwartung. Die resonante Signalverstärkung betrug 20.6, was durch die Güte Q gegeben ist. Die Untersuchung mit variablen Isopren-Konzentrationen und unterschiedlicher Feuchte ergab einen klaren favorisierten Messbereich für die Detektion von Isopren mit variierender Feuchte. Während bei einer Integrationszeit von 100 ms eine Nachweisgrenze von 2,7 ppb direkt gemessen werden konnte, deutet die Allan-Werle-Analyse darauf hin, dass das LOD bei einer Integrationszeit von 120 s auf 77 ppt sinkt.

Isoprenmessungen bei Pflanzen, im Atemgas oder in der Atmosphäre geschehen in der Regel mit feuchten Gasproben. Diese Gasproben werden – teilweise unter Einsatz komplexer

Verfahren – mithilfe von Kältefallen und Nafion oder Absorbern getrocknet. Dies kann jedoch zu Signalverlusten führen, wenn sich Isopren auf Oberflächen ablagert und führt zu längeren Austauschzeiten, wenn beispielsweise zusätzliche Volumina eingebracht werden, wie es bei Kältefallen der Fall ist. Diese Arbeit zeigt, dass auf das Trocknen des Gases verzichtet werden kann, sofern die Wasserkonzentration bestimmt wird.

Eine solche Bestimmung könnte durch eine spektrale Messung mit dem System selbst durchgeführt werden. Daher eignet sich das System für Messungen in allen genannten Bereichen geeignet und bietet speziell für den

Einsatz im Feld eine Alternative zu GC-MS und PTR-ToF Systemen. Bei gleicher Messgenauigkeit ist das photoakustische Spektrometer kompakter und kostengünstiger.

Danksagung:

Wir danken der Deutschen

Forschungsgemeinschaft

für die Förderung des

Sonderforschungsbereichs SFB 1537

„ECOSENSE“ (Projekt-ID: 459819582).

References

- [1] A. B. Guenther *et al.*, "The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): an extended and updated framework for modeling biogenic emissions," *Geosci. Model Dev.*, vol. 5, no. 6, pp. 1471–1492, 2012, doi: 10.5194/gmd-5-1471-2012.
- [2] R. Atkinson and J. Arey, "Atmospheric degradation of volatile organic compounds," *Chemical reviews*, vol. 103, no. 12, pp. 4605–4638, 2003, doi: 10.1021/cr0206420.
- [3] M. Boy *et al.*, "Positive feedback mechanism between biogenic volatile organic compounds and the methane lifetime in future climates," *npj Clim Atmos Sci*, vol. 5, no. 1, 2022, doi: 10.1038/s41612-022-00292-0.
- [4] A. Guenther *et al.*, "A global model of natural volatile organic compound emissions," *J. Geophys. Res.*, vol. 100, D5, pp. 8873–8892, 1995, doi: 10.1029/94JD02950.
- [5] B. C. McDonald *et al.*, "Volatile chemical products emerging as largest petrochemical source of urban organic emissions," *Science (New York, N.Y.)*, vol. 359, no. 6377, pp. 760–764, 2018, doi: 10.1126/science.aag0524.
- [6] S. Oumami, J. Arteta, V. Guidard, P. Tulet, and P. D. Hamer, "Evaluation of isoprene emissions from the coupled model SURFEX-MEGANv2.1," *Geosci. Model Dev.*, vol. 17, no. 8, pp. 3385–3408, 2024, doi: 10.5194/gmd-17-3385-2024.
- [7] A. Arneth, R. K. Monson, G. Schurgers, Ü. Niinemets, and P. I. Palmer, "Why are estimates of global terrestrial isoprene emissions so similar (and why is this not so for monoterpenes)?," *Atmos. Chem. Phys.*, vol. 8, no. 16, pp. 4605–4620, 2008, doi: 10.5194/acp-8-4605-2008.
- [8] C. N. Hewitt *et al.*, "Nitrogen management is essential to prevent tropical oil palm plantations from causing ground-level ozone pollution," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 44, pp. 18447–18451, 2009, doi: 10.1073/pnas.0907541106.
- [9] K. Ashworth, O. Wild, A. S. D. Eller, and C. N. Hewitt, "Impact of Biofuel Poplar Cultivation on Ground-Level Ozone and Premature Human Mortality Depends on Cultivar Selection and Planting Location," *Environmental science & technology*, vol. 49, no. 14, pp. 8566–8575, 2015, doi: 10.1021/acs.est.5b00266.
- [10] J. Kesselmeier and M. Staudt, "Biogenic Volatile Organic Compounds (VOC): An Overview on Emission, Physiology and Ecology," *Journal of Atmospheric Chemistry*, vol. 33, no. 1, pp. 23–88, 1999, doi: 10.1023/A:1006127516791.
- [11] R. K. Monson, S. M. Weraduwage, M. Rosenkranz, J.-P. Schnitzler, and T. D. Sharkey, "Leaf isoprene emission as a trait that mediates the growth-defense tradeoff in the face of climate stress," *Oecologia*, vol. 197, no. 4, pp. 885–

- 902, 2021, doi: 10.1007/s00442-020-04813-7.
- [12] A. J. Hills and P. R. Zimmerman, "Isoprene measurement by ozone-induced chemiluminescence," *Anal. Chem.*, vol. 62, no. 10, pp. 1055–1060, 1990, doi: 10.1021/ac00209a017.
- [13] T. Karl *et al.*, "Human breath isoprene and its relation to blood cholesterol levels: new measurements and modeling," *Journal of applied physiology (Bethesda, Md. : 1985)*, vol. 91, no. 2, pp. 762–770, 2001, doi: 10.1152/jappl.2001.91.2.762.
- [14] A. Amann, P. Mochalski, V. Ruzsanyi, Y. Y. Broza, and H. Haick, "Assessment of the exhalation kinetics of volatile cancer biomarkers based on their physicochemical properties," *Journal of breath research*, vol. 8, no. 1, p. 16003, 2014, doi: 10.1088/1752-7155/8/1/016003.
- [15] L. Kinsler, A. Frey, A. Coppens, and J. Sanders, *Fundamentals of acoustics*, 4th ed. New York, NY, Weinheim: Wiley, 2000. [Online]. Available: <http://www.loc.gov/catdir/description/wiley031/99049667.html>
- [16] D. T. Blackstock, *Fundamentals of physical acoustics*. New York, NY: Wiley, 2000. [Online]. Available: <http://www.loc.gov/catdir/description/wiley031/99042186.html>
- [17] P. Werle, "Accuracy and precision of laser spectrometers for trace gas sensing in the presence of optical fringes and atmospheric turbulence," *Appl Phys B*, vol. 102, no. 2, pp. 313–329, 2011, doi: 10.1007/s00340-010-4165-9.
- [18] P. Werle, R. Mcke, and F. Slemr, "The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS)," *Appl Phys B*, vol. 57, no. 2, pp. 131–139, 1993, doi: 10.1007/BF00425997.
- [19] S. W. Sharpe, T. J. Johnson, R. L. Sams, P. M. Chu, G. C. Rhoderick, and P. A. Johnson, "Gas-phase databases for quantitative infrared spectroscopy," *Applied spectroscopy*, vol. 58, no. 12, pp. 1452–1461, 2004, doi: 10.1366/00093702042641281.

System zum störungsreduzierten Messen von Strukturschwingungen mithilfe eines Laser-Doppler-Vibrometers und drohnengetragenen Spiegels

Robin Zimmermann¹, Dr.-Ing. Fangjian Wang², Prof. Dr.-Ing. Christian Rembe², Prof. Dr.-Ing. Umut Durak¹

¹Institut für Flugsystemtechnik, Deutsches Zentrum für Luft- und Raumfahrt (DLR), Lilienthalplatz 7, 38108 Braunschweig, Deutschland

²Institut für Elektrische Informationstechnik, Technische Universität Clausthal, Leibnizstr. 28, 38678 Clausthal-Zellerfeld, Deutschland

Zusammenfassung

Veränderungen von Schwingungseigenschaften können zur Früherkennung von Schäden in großen, von Menschen geschaffenen Bauwerken genutzt werden. Laser-Doppler-Vibrometer (LDV) eignen sich hierfür aufgrund ihrer einfachen Einrichtung, sind jedoch bei großen Einfallswinkeln auf die Zieloberfläche und verdeckten Messpunkten limitiert. Durch die Verwendung eines von einem unbemannten Luftfahrzeug (UAV) getragenen Spiegels kann der Messlaser flexibel auf die Zieloberfläche umgelenkt werden. Allerdings verfälschen die bei der fortlaufenden Ausrichtung und Umlenkung des Strahls auftretenden Vibrationen das LDV-Messsignal. In dieser Arbeit werden erstmals die aus Laborversuchen bekannten Filtermethoden für Spiegelschwingungen kombiniert und in einem Freiluft-Experiment unter realen Störbedingungen wie Windlasten und Bewegungen des UAVs angewandt. Hierzu wurden Micro-Electro-Mechanical Systems (MEMS) Beschleunigungssensoren an der LDV-Ausrichtungsplattform sowie am Umlenkspiegel montiert. Die Sensoren ermöglichen die Aufzeichnung der störenden Vibrationen und deren anschließende Filterung aus dem LDV-Signal. Für die verteilte Datenerfassung wurde ein drahtloses Synchronisationsverfahren auf Basis des globalen Navigationssatellitensystem (GNSS) basierten Puls pro Sekunde (PPS) Signals implementiert.

Erste Messungen an einer durch einen Shaker erregten Aluminium-Zielplatte zeigen, dass Störungen im Frequenzbereich unter 10 Hz effektiv reduziert werden können. Die Ergebnisse belegen das Potenzial des kombinierten Ansatzes für die kontaktlose Schwingungsüberwachung großer Strukturen, begründen zudem aber auch weitere Arbeiten, um eine Störreduzierung bei höheren Frequenzen zu erzielen.

Keywords: LDV, Drohne, Zustandsüberwachung

Einleitung und Motivation

Große, von Menschen errichtete Bauwerke sind ständig Korrosions- und Umweltbelastungen ausgesetzt. Veränderungen in ihrer strukturellen Integrität können durch die Überwachung ihrer Schwingungseigenschaften erkannt werden. LDV können Schwingungsamplituden im Pikometerbereich auflösen, sind hochgenau sowie einfach einzurichten und eignen sich daher für die Aufzeichnung der erforderlichen Schwingungsmessdaten [1]. Sie sind jedoch dadurch limitiert, dass sich ihre Signalqualität mit zunehmendem Einfallswinkel verschlechtert und die Zielmesspunkte durch Fremdkörper im Sichtfeld oder durch andere Teile der Struktur selbst verdeckt sein könnten. LDV können zwar zur Positionierung direkt von einer Drohne getragen werden [2], [3], [4], potenziell geeignete kleine

und leichte LDV haben jedoch noch einige Limitierungen [5]. Die für hochqualitative Messungen eingesetzten LDV wiegen meist mehrere Kilogramm. Drohnen die zum Tragen dieser Last geeignet wären, sind wiederum so groß, dass ihr Betrieb in der Nähe von vielen Baustrukturen mit Einschränkungen verbunden ist [6]. Darum wird hier der Ansatz verfolgt, den Laserstrahl des LDVs durch einen von einem UAV getragenen Spiegel auf die zu messende Oberfläche zu reflektieren.

Insbesondere durch den Einfluss von Wind ändern sich die Position der Drohne und des Spiegels ständig [7]. Beim kontinuierlichen Neuausrichten und Umleiten des Messlasers kommt es zur Aufzeichnung von Störungen im Messsignal, welche vorwiegend durch die dabei auftretenden Vibrationen der Ausrichtungsplattform und des Spiegels verursacht werden und von der

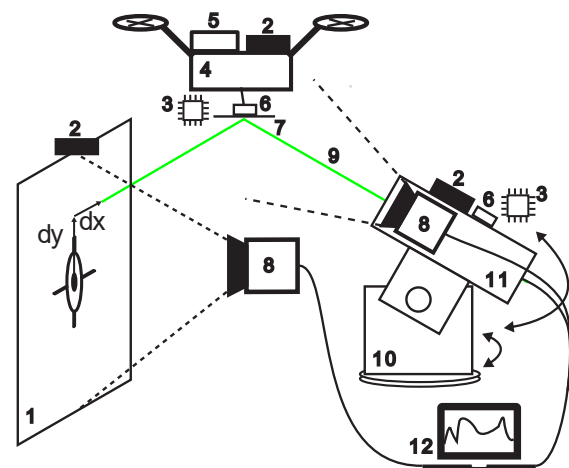
räumlichen Ausrichtung der einzelnen Systemkomponenten zueinander abhängen [8], [9]. In [10] wurde ein Differentielles Laser-Doppler-Vibrometer verwendet um einen Teil dieser Störungen zu kompensieren. Sie können jedoch auch mithilfe von geschickt angebrachten Beschleunigungssensoren bestimmt und somit aus dem LDV-Messsignal gefiltert werden. In [11] wird das Vorgehen zur Filterung von Schwingungen des LDVs selbst und die dafür optimalen Sensorpositionen beschrieben. In gleicher Weise wird in [12] die optimale Sensorposition sowie das Vorgehen zur Filterung von Schwingungen des Umlenkspiegels beschrieben. Beide Arbeiten haben gemeinsam, dass sie unter Laborbedingungen mit weitestgehend kontrollierten Randbedingungen durchgeführt wurden.

In dieser Arbeit werden die beiden Methoden zur Filterung der LDV- und Spiegelschwingungen zum ersten Mal kombiniert und auf einen Versuch im Freien mit realen Störungen durch Windlasten und Ausrichtungsbewegungen des Lasers angewandt. Für eine verteilte Aufzeichnung der Messsignale wird zudem eine Methode zur drahtlosen Synchronisation entwickelt und vorgestellt.

Messsystemaufbau

Die Komponenten des Messsystems sind in Abb. 1 zu sehen. Das sich am Boden befindliche VibroFlex QTec LDV von Polytec ist auf einer motorisierten Ausrichtungsplattform montiert. Die Umlenkung des LDV-Laserstrahls erfolgt mit einem Spiegel, welcher durch einen drohnengetragenen, motorisierten Gimbal ausgerichtet werden kann.

Zudem wird auf der Ausrichtungsplattform eine Kamera angebracht. Diese ist auf einer kinematischen Halterung montiert, mithilfe derer die Achse der Kamera parallel zur Achse des Laserstrahls ausgerichtet wird. Wie in Abb. 2 zu sehen, sind links und rechts vom Spiegel jeweils Aruco Marker zur optischen Erkennung angebracht. Mithilfe von Computer-Vision Algorithmen aus der Softwarebibliothek OpenCV werden diese Marker im Kamerabild erkannt. Dies ermöglicht die Ausrichtung der Ausrichtungsplattform auf einen bestimmten Punkt im Kamerabild, der so berechnet wird, dass der Achsabstand zwischen Kamera und LDV-Laserachse ausgeglichen ist, und der Laserstrahl das Zentrum der Spiegeloberfläche trifft.



Nr.	Bezeichnung
1	Zielfläche
2	GNSS-Empfänger
3	Mikrocontroller
4	Drohne
5	Onboard Computer
6	Beschleunigungssensor(en)
7	Spiegel
8	Kamera
9	Laserstrahl
10	Ausrichtungsplattform
11	LDV
12	Messdatenaufnahme- & Verarbeitungssystem

Abb. 1 Messsystemkonzept

Auf der Drohne ist ein RaspberryPi als sogenannter Onboard Computer angebracht, welcher die Sensordaten der Drohne aufzeichnet, den drahtlosen Datenaustausch mit der Bodenplattform sicherstellt und die Achswinkel des Gimbals kontrolliert.

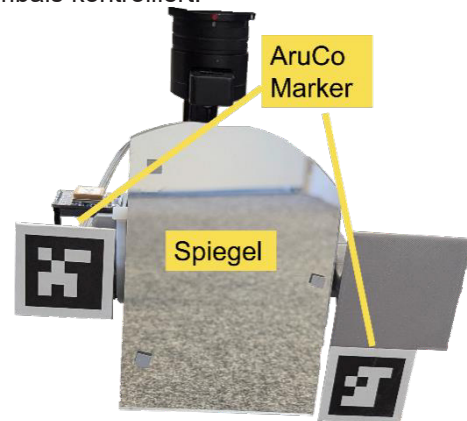


Abb. 2 Gimbal mit Nutzlast Vorderseite

Als Zielfläche kommt eine Aluminiumplatte zum Einsatz, die auch mit Aruco Markern ausgestattet wurde. Sie ist in Abb. 3 zu sehen. Die Bestimmung und Korrektur von Abweichungen des Laserauffreffpunkts vom Zielpunkt erfolgt ebenfalls mithilfe einer Kamera und klassischen

Bildverarbeitungsmethoden, wobei die Aruco Marker hier als räumliche Referenzpunkte für das lokale Koordinatensystem des Auftreffpunktes dienen. Um die inkrementalen Winkelkorrekturen der Spiegelausrichtung zu korrigieren, sind zudem die Positionen der drei Komponenten Drohne, Zielfläche und Ausrichtungsplattform vonnöten. Die statischen Positionen der Zielfläche und Ausrichtungsplattform werden mithilfe von Real-Time Kinematic positioning (RTK) korrigierten Positionsdaten von mobilen GNSS-Empfängern einmalig bestimmt, während die sich zeitlich verändernde Position der Drohne von den darin verbauten GNSS-Empfängern während der Messung kontinuierlich ermittelt wird. Durch die RTK-Korrektur ist die Genauigkeit der Positionsdaten üblicherweise < 2 cm.

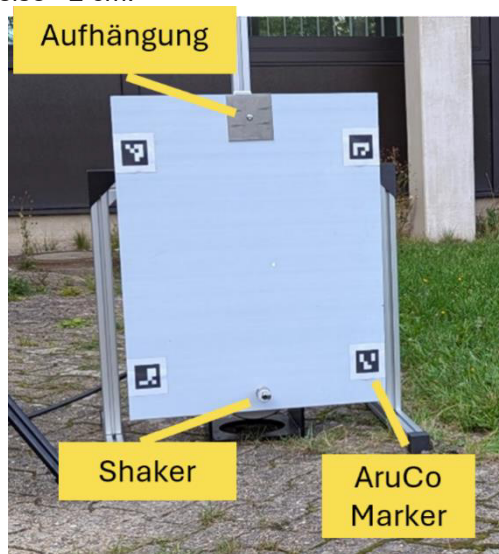


Abb. 3 Zielplatte mit Shaker und Aruco Markern

Um eine möglichst hohe Bandbreite des Messsignals nutzen zu können, muss jegliche Art von zeitabhängiger Weglängenänderung des Laserstrahls bestimmt werden können. Dafür sind mehrere Messsignale nötig, die insbesondere, wenn sie höherfrequente Anteile enthalten, hochgenau zeitlich synchronisierbar sein müssen. Anders als beim Aufbau des Messsystems im Labor, wäre eine Anbindung aller Korrektursensoren an ein zentrales Messdatenerfassungssystem per Kabel ein limitierender Faktor für den Einsatz eines UAV im Freien. Um trotzdem eine zeitliche Synchronisation der Messdaten zu erreichen, ist für den Messsystemaufbau ein Konzept mittels GNSS vorgesehen. Störschwingungen des LDVs und des Umlenkspiegels werden mithilfe von an diesen Komponenten angebrachten Beschleunigungssensoren aufgezeichnet. Um die Vibrationen des Spiegels aufzuzeichnen, werden drei MEMS Beschleunigungssensoren mit der Bezeichnung ADXL355 von Analog Devices an dessen

Rückseite angebracht. Sie sind auf einem Kreis angeordnet, sodass in der Zukunft auch eine Bestimmung der Winkel-Vibrationen in den Drehachsen des Spiegels erfolgen kann. Die Anbringung ist in Abb. 4 zu sehen. Die Sensoren besitzen eine digitale Schnittstelle zum Auslesen der Messdaten. Dass die Beschleunigungsdaten auf dem Spiegel zur gleichen Zeit aufgenommen werden, wird durch die Vorgabe eines Takts durch einen Mikrokontroller ermöglicht. Das vom Hersteller spezifizierte DRDY-Signal der Beschleunigungssensoren zeigt an, zu welchem Zeitpunkt neue Messdaten vorliegen. Dieser Zeitpunkt wird durch den Mikrokontroller mit einem Hardware Interrupt erkannt und mit einem von seinem freilaufenden Taktgeber erstellten Zeitstempel versehen. Dieser Signalfluss ist in Abb. 5 dargestellt.

Auf die gleiche Art bekommen auch die Zeitpunkte einer steigenden Flanke des GNSS PPS Signals einen Zeitstempel. Damit entsteht ein relativer Bezug zwischen den Zeitstempel der Beschleunigungsdaten und der durch das GNSS-Modul ermittelten koordinierten Weltzeit (UTC).

Analog werden auch die Vibrationen der LDV-Längsachse aufgezeichnet, welche durch die Ausrichtungsplattform bei der Ausrichtung erzeugt werden und mit einem Zeitstempel versehen. Die Anbringung dieses Sensorsystems an das LDV ist in Abb. 6 dargestellt.

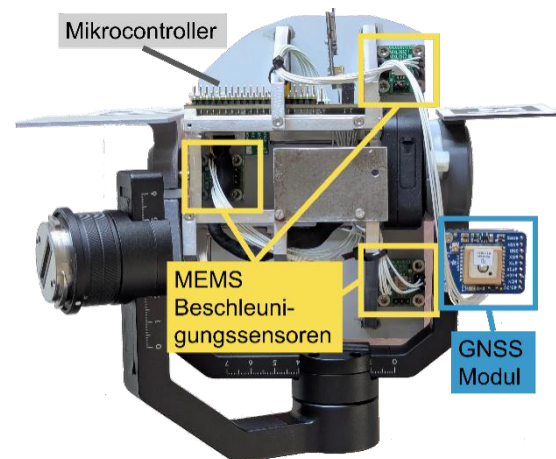


Abb. 4 Gimbal mit Nutzlast Rückseite

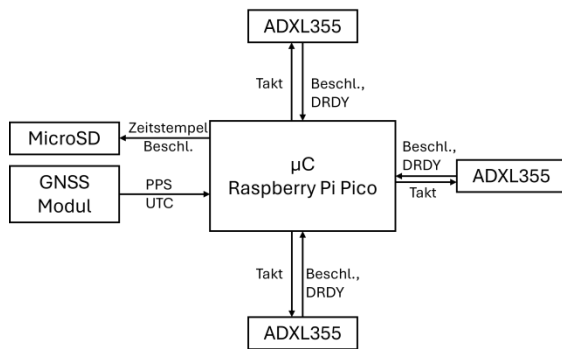


Abb. 5 Signalfluss Messsystem Gimbal

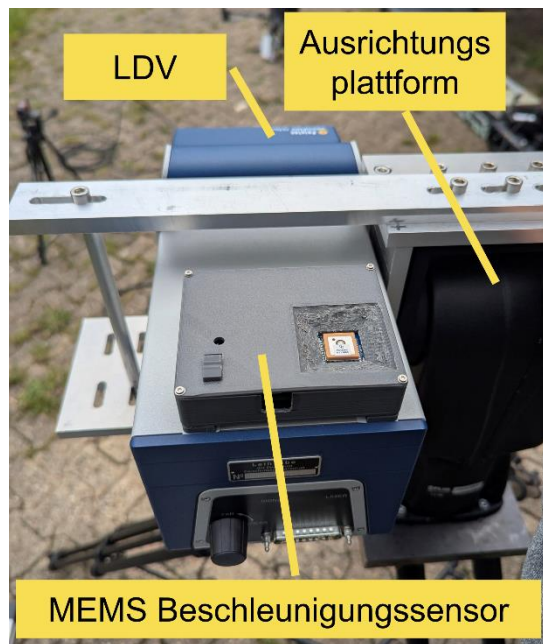


Abb. 6 Korrektursensor auf Ausrichtungsplattform

Messdatenverarbeitung

Die verteilt aufgenommenen Messdaten werden nach einer Messung gesammelt und zentral ausgewertet.

Im ersten Schritt wird das Geschwindigkeitssignal des LDVs abgeleitet, um ein Beschleunigungssignal zu erhalten. Daraufhin erfolgt eine Anpassung der Signalabtastrate an die der Beschleunigungssensoren.

Die Messdaten müssen dann noch zeitlich miteinander synchronisiert werden. Dafür wird eine Kreuzkorrelation durchgeführt. Um dabei ein möglichst genaues Ergebnis erzielen zu können, werden weitere Schritte zur Vorbereitung der Signale durchgeführt. Da es sich beim ADXL355 um einen digitalen Beschleunigungssensor handelt, werden die Daten im Sensor vor Ausgabe mit einem Tiefpass Filter mit einer 3 dB Grenzfrequenz bei ca. 1 kHz gefiltert. Die

Bandbreite des LDV-Signals ist weitaus größer und wird daher vor der Kreuzkorrelation auch zuerst mit einem Tiefpassfilter mit einer Grenzfrequenz von 1 kHz versehen. Als charakteristischer Abschnitt bei welchem eine deutliche Korrelation der zwei Signale zu erkennen ist, wird der Start der Drohnenmotoren vor Abflug identifiziert. Anhand dieses kurzen Zeitabschnittes wird der Zeitversatz der zwei Signale approximiert. Die abschließende Zeitsynchronisation erfolgt anhand der PPS-Zeitstempel, sodass die Signale für die nachfolgende Störreduktion zeitlich weitestgehend miteinander synchronisiert sind.

Der Doppler-Effekt tritt beim Reflektieren des Laserstrahls am drohnengetragenen Umlenkspiegel jeweils auf dem Hin- und Rückweg auf. Wie stark die Vibrationen des Umlenkspiegels dabei in das Messsignal mit eingehen, hängt maßgeblich vom Einfallswinkel ab. Der damit verbundene Skalierungsfaktor s wird angelehnt an [12] mithilfe der Berechnungsvorschrift

$$s = -2(\hat{a}_{pd} \cdot \hat{n}_m) \quad (1)$$

bestimmt. Hier ist $\hat{a}_{pd}$ der normierte Vektor zwischen der Position der Ausrichtungsplattform und der Drohne und $\hat{n}_m$ ist der Normalvektor des Spiegels. Der Normalvektor des Spiegels wird mit

$$\hat{n}_m = \frac{\hat{a}_{dt} - \hat{a}_{pd}}{\|\hat{a}_{dt} - \hat{a}_{pd}\|} \quad (2)$$

berechnet. $\hat{a}_{dt}$ ist der normierte Vektor zwischen der Position des Spiegels und dem Auftreffpunkt des Lasers.

Abschließend werden die Beschleunigungswerte der Korrektursensoren gemittelt, mit dem Faktor s skaliert und im Zeitbereich von den Beschleunigungsmesswerten des LDVs subtrahiert, um ein korrigiertes LDV-Signal zu erhalten.

Versuchsaufbau

Die Effektivität des Messsystemkonzepts wird anhand von in einem Versuch aufgenommenen Messdaten bewertet. Dafür wird eine Zielscheibe mit einer Sinusschwingung mit bekannter Frequenz angeregt und diese mit dem durch den drohnengetragenen Spiegel umgelenkten LDV-Laserstrahl gemessen. Anschließend erfolgt ein Vergleich der Messung vor und nach der Störreduktion. Der Versuchsaufbau ist in Abb. 7 dargestellt. Der Abstand sowohl zwischen Drohne und Zielplatte als auch zwischen Ausrichtungsplattform und Drohne beträgt 5-6 m.

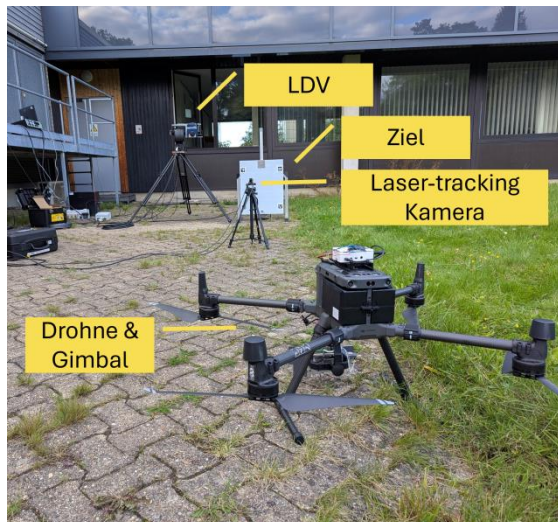
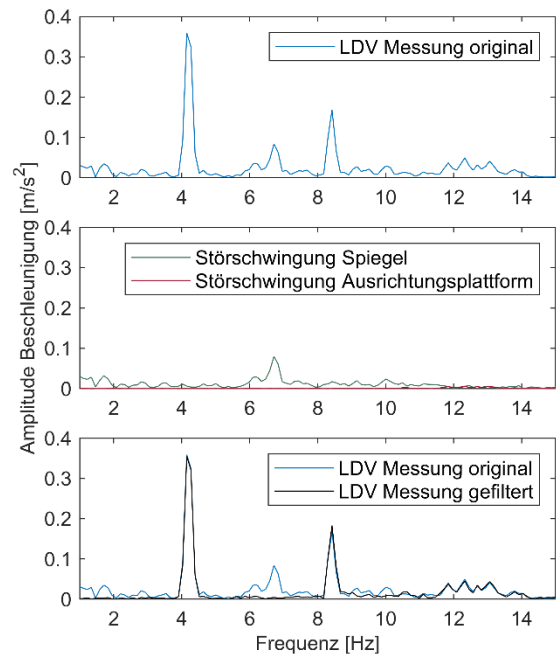


Abb. 7 Versuchsaufbau

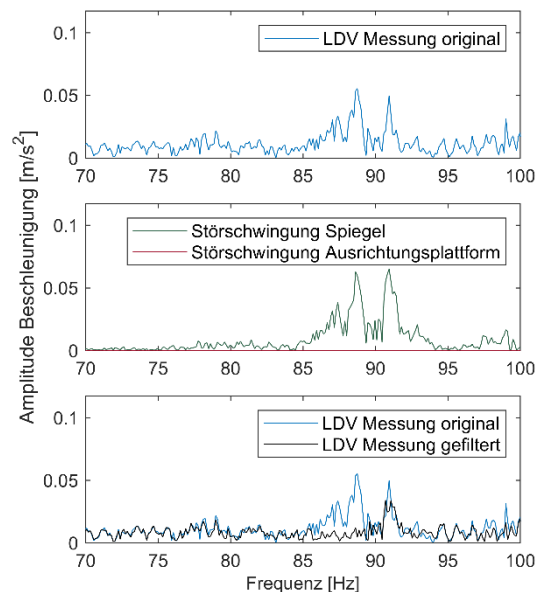
Ergebnisse

Während der Messreihe, die in Abb. 8 und Abb. 9 zu sehen ist, wurde die Zielplatte mit ca. 4,2 Hz angeregt. Es ist aus Abb. 8 ersichtlich, dass die Amplituden der Störschwingungen, welche durch die Ausrichtungsplattform verursacht werden, im Vergleich zu denen des Spiegels klein sind. Wie im Bildausschnitt (a) zu erkennen, weist das mithilfe der Korrektursensoren gefilterte Signal bei den Frequenzen unter 10 Hz außerhalb von Vielfachen der Anregungsfrequenz des Shakers geringere Amplituden auf. In Bildausschnitt (b) lässt sich zudem eine Reduktion der Amplituden im Bereich der doppelten Rotorfrequenz der Drohne bei ca. 85-95 Hz feststellen. Zwischen den dargestellten Frequenzbereichen treten breitbandige Störungen auf, bei denen sich keine signifikante Abweichung der Amplituden erkennen lässt. Die Amplituden der Frequenzen bei denen eine Anregung der Zielplatte durch den Shaker zu erwarten ist, wie z.B. 4,2 Hz, 8,4 Hz, 12,6 Hz etc., bleiben größtenteils unverändert.

In Abb. 9 sind im Zeitsignal der LDV Messdaten Beschleunigungsspitzen zu sehen, deren Betrag weit über den umliegenden Werten liegen.



(a)



(b)

Abb. 8 Störschwingungsreduzierung im LDV-Signal. Anregungsfrequenz 4,2 Hz

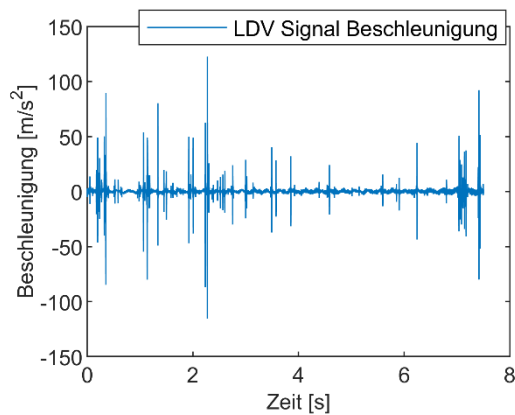


Abb. 9 Mit LDV gemessene Beschleunigung vor Filterung. Anregungsfrequenz 4,2 Hz

Bei einer zweiten Messung wurde der Shaker bei 68 Hz betrieben. Die Ergebnisse sind in Abb. 10 und Abb. 11 zu sehen. Im Zeitsignal der Abb. 10 heben sich die Beschleunigungsspitzen nicht mehr so sehr vom Rest des Signals ab, da die Schwingungsamplituden insgesamt höher sind. In Abb. 11 kann wieder eine Reduktion der Schwingungsamplituden bei <10 Hz beobachtet werden. Auch hier sind die Amplituden im Messsignal des Sensors auf der Ausrichtungsplattform gerade im unteren Frequenzbereich um mehrere Größenordnungen niedriger und haben somit kaum einen Effekt bei der Störreduzierung.

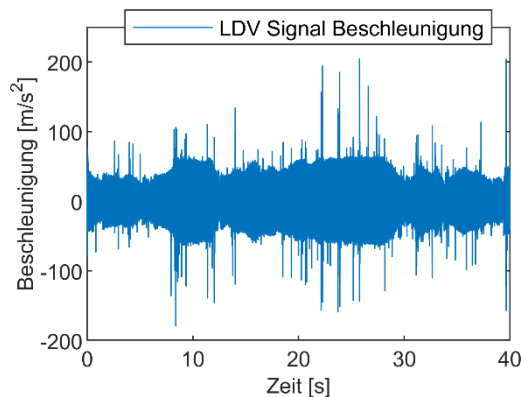


Abb. 10 Abgeleitetes LDV-Signal vor Filterung. Anregungsfrequenz 68 Hz

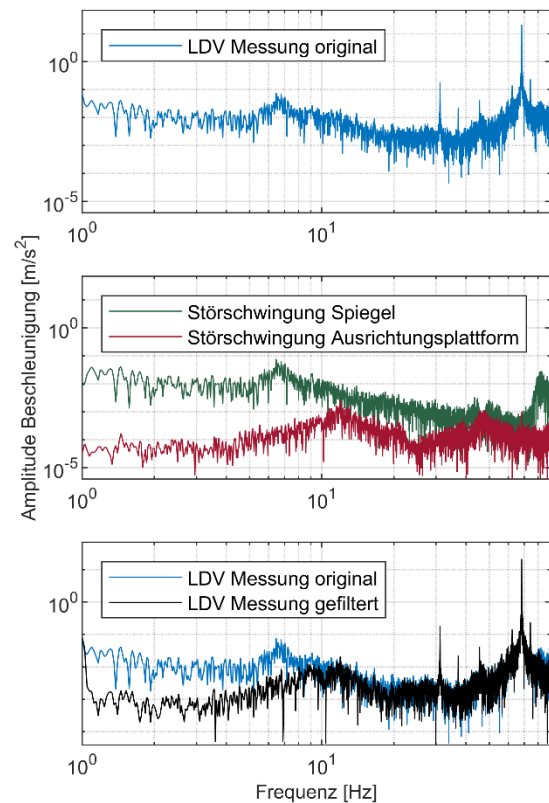


Abb. 11 Störschwingungsreduzierung im LDV-Signal. Anregungsfrequenz 68 Hz

Diskussion

Im Frequenzbereich unter 10 Hz erweist sich die Methode zur Störreduktion als effektiv. Es wurde zwar keine simultane Referenzmessung durchgeführt, jedoch sind die Vibrationen des Spiegels und der Zielplatte nicht gekoppelt, weshalb bei einer Störreduktion über mehrere Frequenzen mit hoher Sicherheit davon ausgegangen werden kann, dass tatsächlich die Schwingungen des Spiegels und nicht andere zufällig korrelierten Signalanteile unterdrückt werden.

Erwartungsgemäß nimmt die Effektivität der Störreduktion über 10 Hz signifikant ab. Dies kann mehrere Ursachen haben. Zum einen wurde keine Kalibrierung der Beschleunigungssensoren vorgenommen. Da die Sensitivität der MEMS-Sensoren bei höheren Frequenzen abnimmt, ist hier von einer verminderten Effektivität bei der Störungsreduzierung auszugehen. Bei höheren Frequenzen spielen auch kleine Zeitversätze zwischen den Signalen eine immer größere Rolle. Das Datenerfassungssystem des LDVs sowie auch die MEMS-Sensoren haben beide eine unterschiedliche Gruppenlaufzeit, die bei der Synchronisation mithilfe des GNSS PPS Signals noch nicht berücksichtigt

wurde. Dies führt zu einer mit der Frequenz ansteigenden Phasenverschiebung zwischen den Signalen.

Auch eine zu niedrige Abtastrate der Signale kann bei höheren Frequenzen zu einer Phasenverschiebung führen, da die Signale nur anhand der diskreten Abtastpunkten synchronisiert werden können. Sind diese weit voneinander entfernt, führt dies zu einer signifikanten Phasenverschiebung.

Signaleinträge, welche durch eine Ausrichtungsänderungen des Spiegels verursacht werden, sind hier nicht berücksichtigt worden. Je nach Auftreffpunkt des Lasers auf dem Spiegel, haben Winkelschwingungen des Spiegels einen direkten Einfluss auf das Messsignal. Wichtiger ist wohl aber, dass sie auch zu einer Verschiebung des Laserauftreffpunkts auf der Zielplatte führen. Trifft der Laserstrahl nicht nahezu rechtwinklig auf das Ziel auf, führen auch kleine Verschiebungen demzufolge zu einer Aufzeichnung einer nicht vernachlässigbaren Geschwindigkeit.

In den Zeitsignalen der abgeleiteten LDV Messsignale in Abb. 9 und Abb. 10 treten zeitlich hochgradig konzentrierte Beschleunigungsspitzen auf, welche mit hoher Wahrscheinlichkeit auf den Laser-Speckle Effekt zurückzuführen sind. Hier kommt es auf einer rauen Oberfläche zu einer Interferenz des Lasers mit sich selbst und zu einem kurzen Signalverlust, welcher wegen der teilweise schnellen Auslenkung des Laserauftreffpunkts auf der Zielplatte nicht vom QTec LDV kompensiert werden kann. Mit einer anderen Implementierung könnte die Speckle-Unterdrückung jedoch verbessert werden, siehe [13]. Da die Beschleunigungsspitzen mit einem großen Betrag auftreten, führen sie in der FFT zu einem breitbandigen Energieeintrag, welcher möglicherweise eine Störreduktion bei höheren Frequenzen überdeckt.

Der Beschleunigungssensor auf der Ausrichtungsplattform verzeichnete Schwingungen mit nur sehr geringer Amplitude. Ein möglicher Grund dafür ist, dass der Sensor sich auf einer Platine aufgelötet in einem Plastikgehäuse befand, welches per doppelseitigem Klebeband auf dem LDV befestigt war. Dadurch war die Anbindung möglicherweise nicht steif genug, um auch höhere Frequenzanteile mit niedriger Dämpfung zu übertragen. In zukünftigen Messungen wird der Sensor direkt mit einem Adapter aus Aluminium auf das LDV angebracht.

Zusammenfassung

In dieser Arbeit werden erste Ergebnisse eines neu entwickelten Messsystems zur Früherkennung von Schäden an Baustrukturen vorgestellt. Dieses setzt sich aus einem bodenbasier-

ten LDV und einem drohnengetragenen Spiegel, welcher den Laserstrahl des LDVs auf die zu untersuchende Oberfläche umleitet, zusammen. Beim Ausrichten und Umlenken des Laserstrahls auf das Zielobjekt treten Vibrationen auf, die im Messsignal des LDVs verzeichnet werden. Um diese Vibrationen aufzeichnen und aus dem Messsignal des LDVs zu filtern, werden MEMS-Beschleunigungssensor basierte Messsysteme für die LDV-Ausrichtungsplattform und den Spiegel entwickelt. Diese erlauben eine nachträgliche Synchronisation der Messdaten über GNSS PPS Signale. Erste Messsignale werden mit einer durch einen Shaker angeregten Zielplatte aus Aluminium aufgenommen. Eine Analyse der Messdaten zeigt, dass Störungen im für große Baustrukturen relevanten Frequenzbereich unter 10 Hz effektiv reduziert werden können. Bei höheren Frequenzen scheint die Störreduktion an Effektivität zu verlieren. Dies könnte zum einen in einem Speckle-Effekt bedingten Signalverlust begründet sein, durch welchen es zu Messwertspitzen im abgeleiteten LDV-Signal kommt. Zum anderen wurde der Effekt von Winkelvibrationen des Umlenkspiegels auf das Messsignal bisher noch nicht berücksichtigt. Neben der Umsetzung einer Sensorkalibrierung der Beschleunigungssensoren und einer zeitgleichen Referenzmessung der Zieloberfläche, sollten diese Effekte in Zukunft näher untersucht werden, um die Effektivität der Störreduzierung zu verbessern.

Förderung

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - 496079937

Literatur

- [1] C. Rembe, B. J. Halkon, und M. A. A. Ismail, „Measuring Vibrations in Large Structures with LDV and UAS: A Review and Outlook“, *Adv Devices Instrum.*, S. adi.0103, Juli 2025, doi: 10.34133/adi.0103.
- [2] P. Garg u. a., „Measuring Transverse Displacements Using Unmanned Aerial Systems Laser Doppler Vibrometer (UAS-LDV): Development and Field Validation“, *Sensors*, Bd. 20, Nr. 21, Art. Nr. 21, Jan. 2020, doi: 10.3390/s20216051.
- [3] H. Zheng, T. Guo, G. Zhi, und Z. Hu, „A Novel Debonding Damage Identification Approach of Hidden Frame-Supported Glass Curtain Walls Based on UAV-LDV System“, *Applied Sciences*, Bd. 14, Nr. 13, Art. Nr. 13, Jan. 2024, doi: 10.3390/app14135412.
- [4] P. Garg, F. Moreu, A. Ozdagli, M. R. Taha, und D. Mascareñas, „Noncontact Dynamic Displacement Measurement of Structures Using a Moving Laser Doppler Vibrometer“, *Journal of Bridge Engineering*, Bd. 24, Nr. 9, S. 04019089, Sep. 2019, doi: 10.1061/(ASCE)BE.1943-5592.0001472.

- [5] Y. Li, E. Dieussaert, und R. Baets, „Miniaturization of Laser Doppler Vibrometers—A Review“, *Sensors*, Bd. 22, Nr. 13, Art. Nr. 13, Jan. 2022, doi: 10.3390/s22134735.
- [6] *Durchführungsverordnung (EU) 2019/947 der Kommission vom 24. Mai 2019 über die Vorschriften und Verfahren für den Betrieb unbemannter Luftfahrzeuge (Text von Bedeutung für den EWR.)*, Bd. 152. 2019. Zugriffen: 17. August 2024. [Online]. Verfügbar unter: http://data.europa.eu/eli/reg_impl/2019/947/oj/deu
- [7] R. Zimmermann, U. Durak, und M. A. A. Ismail, „Towards laser-doppler-vibrometry with UAVs - the effect of wind disturbances on the position of a mirror attached to a drone“, *CEAS Aeronaut J*, Feb. 2026, doi: 10.1007/s13272-026-00955-1.
- [8] M. Schewe, M. A. A. Ismail, und C. Rembe, „Differentielle Laser-Doppler-Vibrometrie für luftgestützte Vibrationsmessungen mit Drohnen“, *tm - Technisches Messen*, Bd. 88, Nr. s1, S. s42–s47, Sep. 2021, doi: 10.1515/teme-2021-0071.
- [9] B. J. Halkon und S. J. Rothberg, „Towards laser Doppler vibrometry from unmanned aerial vehicles“, *J. Phys.: Conf. Ser.*, Bd. 1149, S. 012022, Dez. 2018, doi: 10.1088/1742-6596/1149/1/012022.
- [10] B. J. Halkon, C. Gao, R. Zimmerman, und C. Rembe, „Differential Laser Doppler Vibrometer Measurements Using ``Flyable Mirrors`` for the Vibration Assessment of Remote Structures“, 22. März 2025, *Social Science Research Network, Rochester, NY*: 5189527. doi: 10.2139/ssrn.5189527.
- [11] B. J. Halkon und S. J. Rothberg, „Taking laser Doppler vibrometry off the tripod: correction of measurements affected by instrument vibration“, *Optics and Lasers in Engineering*, Bd. 91, S. 16–23, Apr. 2017, doi: 10.1016/j.optlas-eng.2016.11.006.
- [12] B. J. Halkon und S. J. Rothberg, „Restoring high accuracy to laser Doppler vibrometry measurements affected by vibration of beam steering optics“, *Journal of Sound and Vibration*, Bd. 405, S. 144–157, Sep. 2017, doi: 10.1016/j.jsv.2017.05.014.
- [13] F. Wang, S. Krause, und C. Rembe, „Signal diversity for the reduction of signal dropouts and speckle noise in a laser-Doppler extensometer“, *Measurement: Sensors*, Bd. 22, S. 100377, Aug. 2022, doi: 10.1016/j.measen.2022.100377.

Voruntersuchung zum Einsatz von kostengünstigen MEMS-Beschleunigungssensoren für die kontinuierliche Bestimmung der Lageänderung von Brückenbauwerken

Dominik Thiem¹, Andreas Hennig¹, Gregor Schiele², Martin Krückhans³

¹Hochschule Ruhr West, Institut Mess- und Sensortechnik, 45479 Mülheim an der Ruhr, Deutschland

²Universität Duisburg-Essen, Eingebettete Systeme der Informatik, 47057 Duisburg, Deutschland

³Stadt Essen, Amt für Geoinformation, Vermessung und Kataster, 45127 Essen, Deutschland

Zusammenfassung

Die kontinuierliche Überwachung von Brückenbauwerken gewinnt vor dem Hintergrund einer alternden Infrastruktur und steigender Verkehrsbelastungen zunehmend an Bedeutung. Für eine flächendeckende Anwendung sind kostengünstige, langlebige und zugleich hinreichend genaue Sensorsysteme erforderlich. Der vorliegende Beitrag untersucht, inwieweit sich preisgünstige MEMS-Beschleunigungssensoren für die kontinuierliche Erfassung von Lageänderungen an Brücken eignen. Im Mittelpunkt stehen dabei zwei wesentliche Störgrößen: verkehrsbedingte dynamische Einflüsse an der Messstelle sowie temperaturabhängige Änderungen der Sensoreigenschaften.

Zur Bewertung des Verkehrseinflusses wurde ein Messsystem in einer stark frequentierten Straßenbrücke installiert und über mehrere Monate kontinuierlich betrieben. Ergänzend wurden die Genauigkeit der Sensorik unter Laborbedingungen sowie parasitäre Effekte in einem Klimaschrank über einen weiten Temperaturbereich untersucht. Die Ergebnisse zeigen, dass sich verkehrsbedingte Störungen aufgrund ihrer höheren Frequenzanteile durch eine ressourcenschonende Filterung wirksam um etwa 55 dB reduzieren lassen. Darüber hinaus lassen sich temperaturbedingte Abweichungen einzelner Sensorachsen durch sensor- und ausrichtungsabhängige Referenzmessungen sowie eine polynomiale Kompensation deutlich verringern, die Streuung sinkt um etwa den Faktor sechs.

Die Untersuchungen belegen, dass die eingesetzte Sensorik die geforderte Winkelauflösung von $0,0057^\circ$ grundsätzlich erreichen kann. Gleichzeitig zeigt sich die Relevanz einer individuellen Charakterisierung jedes Sensorelements, da jede Sensorachse in unterschiedlichen Einbaulagen markante temperaturbedingte Störungen aufweisen kann. Für quasistationäre Größen dieses Anwendungsfalls mit einer gleichbleibenden Ausrichtung lässt sich der Aufwand der erforderlichen Referenzmessung auf Basis dieser Voruntersuchung jedoch erheblich reduzieren.

Keywords:

kapazitive Beschleunigungssensoren, Inklinometer, Temperaturkompensation, Brücke, Structural Health Monitoring,

Einleitung

Nachbergbauliche Bodenbewegungen sowie eine alternde Infrastruktur, insbesondere im Brückenbestand, führen zu einem wachsenden Bedarf an flächendeckender Überwachung. In Deutschland existieren rund 130000 Brücken, darunter etwa 39500 Brücken an Autobahnen und Bundesstraßen sowie rund 64000 kommunale Straßen-, Fußgänger- und Radwegbrücken. Besonders kritisch ist das Alter vieler Bauwerke: Mehr als die Hälfte wurde vor 1985 errichtet und muss heute deutlich höhere Verkehrslasten aufnehmen als ursprünglich vorgesehen [1, 2].

Die Überwachung erfolgt derzeit überwiegend durch gesetzlich vorgeschriebene Regelprüfungen im Abstand von drei bis sechs

Jahren [3]. Diese Prüfungen liefern jedoch lediglich Momentaufnahmen des Bauwerkszustands. Kontinuierliches Structural Health Monitoring mit fest installierter Sensorik ist bislang die Ausnahme und findet meist nur im Rahmen einzelner Pilotprojekte Anwendung [4–7]. Der Fokus liegt dabei häufig auf bekannten Schadstellen oder zeitlich begrenzten Belastungssituationen. Die geringe Verbreitung solcher Systeme ist vor allem auf hohe Kosten, fehlende Standardisierung sowie den hohen Implementierungsaufwand in Bestandsbauwerken zurückzuführen [8, 9].

Problemstellung

Für eine flächendeckende Anwendung sind kostengünstige, langlebige und zuverlässige Sensoren essenziell. Eine wesentliche

Herausforderung besteht in der Überlagerung der eigentlichen Messgröße durch Störeinflüsse, insbesondere durch Verkehr, Temperaturschwankungen und Langzeitdrift der Sensorelemente. Kostengünstige MEMS-Beschleunigungssensoren haben ihre Eignung zur Erfassung dynamischer Effekte, etwa von Eigenschwingungen, bereits gezeigt [10, 11]. Für die Bestimmung quasistatischer Größen wie Lageänderungen können sie jedoch nicht ohne Weiteres eingesetzt werden.

Die Ursache liegt insbesondere in ihrer ausgeprägten Temperaturabhängigkeit. Diese kann unter anderem durch thermische Spannungen infolge unterschiedlicher Ausdehnungskoeffizienten der Materialien sowie durch Drift der Ausleseelektronik hervorgerufen werden [12–14]. Bereits Temperaturänderungen von wenigen Grad Celsius können dadurch Messfehler in der Größenordnung der zu erfassenden Neigungsänderung verursachen. Da bei jedem Sensor die einzelnen Achsen individuelle Temperaturabhängigkeiten aufweisen, die zusätzlich auch von der Einbaulage und gegebenenfalls dem Alter abhängig sind, ist ihre Vorabcharakterisierung aufwändig. Genau dies erschwert bislang eine breite Massenanwendung solcher Sensoren. Der vorliegende Beitrag untersucht daher zwei zentrale Einflussgrößen: Erstens den Einfluss von Verkehr und umgebungsbedingten Störungen an einer realen Messstelle und zweitens die Temperaturabhängigkeit sowie die allgemeine Genauigkeit der Sensorik unter Laborbedingungen. Auf Basis der gewonnenen Teilergebnisse wird ein reduzierter Ansatz zur Kompensation dieser Einflüsse vorgestellt.

Grenzen der Sensortechnologie

Als erster Indikator für ein kontinuierliches Monitoring wird die Änderung der Lage eines Bauwerks betrachtet. Dabei ist nicht die absolute Lage des Sensors entscheidend, sondern ausschließlich deren zeitliche Änderung. Die eingesetzte Sensorik basiert auf kostengünstigen MEMS-Beschleunigungssensoren, die die Beschleunigung in zwei orthogonalen Achsen erfassen. In der vorliegenden Anwendung soll der Sensor keiner relevanten dynamischen Beschleunigung ausgesetzt sein, sondern im Wesentlichen nur die Erdbeschleunigung, als Grundlage für die Berechnung der Lage, detektieren.

Durch die Auswertung zweier Achsen lassen sich sowohl Verdrehungen mit dem Winkel φ als auch Verkippungen um die y-Achse des Bezugssystems bestimmen (siehe Abb. 1), ohne dass das Ergebnis unmittelbar von Ausrichtungsfehlern oder vom exakten Betrag der Erdbeschleunigung abhängt [15].

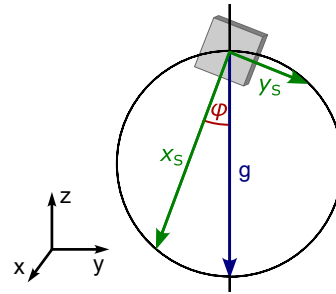


Abb. 1: Ausrichtung eines Sensorelements mit zugehörigen Beschleunigungskomponenten und Bezugssystem.

Als Zielauflösung wird eine Winkeländerung von $0,0057^\circ$ angesetzt, für eine direkte Vergleichbarkeit mit einem an der Messstelle genutztem Messverfahren. Dies kann eine Höhenänderung von 1 mm auf einer Messstrecke von 10 m detektieren, was gemäß dem geometrischen Zusammenhang den genannten Winkel entspricht. Voruntersuchungen zeigten, dass neben der Auflösung insbesondere Skalierungsfehler und Temperaturabhängigkeiten einzelner Achsen die größten Einflüsse auf das Messergebnis haben. Der IIS2ICLX der Firma STMicroelectronics wurde als Sensor gewählt, da er den Stand der Technik für hoch präzise Beschleunigungssensoren repräsentiert. Dieser zeichnet sich durch geringes Rauschen, hohe Stabilität und eine hohe Wiederholgenauigkeit aus und wird daher auch für Inklinometeranwendungen empfohlen. Unter Berücksichtigung der Herstellerangaben ist bei einer Sensitivität von $0,061 \text{ mg/LSB}$ je nach Ausrichtung ein maximaler Quantisierungsfehler von weniger als $0,005^\circ$ zu erwarten (siehe Abb. 2) [15].

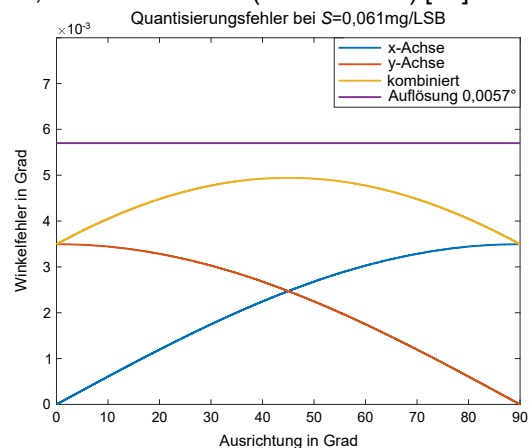


Abb. 2: Maximaler Winkelfehler durch Quantisierungsfehler einer Achse (blau, rot) in Abhängigkeit von der Ausrichtung und bei Kombination beider Achsen (gelb).

Gleichzeitig ist die temperaturabhängige Änderung der Sensitivität mit bis zu $\pm 0,012\%/^{\circ}\text{C}$ angegeben. Daraus ergibt sich, dass bei unveränderter Sensorausrichtung bereits deutliche scheinbare Winkeländerungen auftreten können. Im ungünstigsten Fall kann bei einer Ausrichtung von 45° (gemäß der Definition in Abb. 1) und entgegengesetzten Temperaturkoeffizienten beider Achsen bei einer Temperaturspanne von 50°C eine scheinbare Winkeländerung von etwa $0,34^{\circ}$ entstehen.

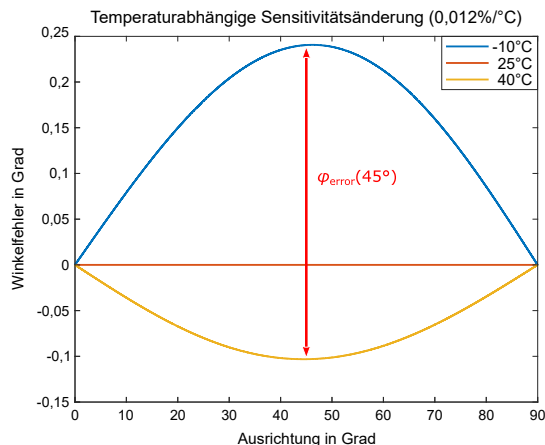


Abb. 3: *Maximaler Winkelfehler durch temperaturabhängige Sensitivitätsänderung in Abhängigkeit von der Ausrichtung*

Im Folgenden wird deshalb untersucht, wie stark die tatsächlichen Störungen an einer repräsentativen Messstelle ausfallen (*Testinstallation im Brückenbauwerk*) und welche praktische Genauigkeit unter idealisierten Bedingungen erreicht werden kann (*Genauigkeit der Sensorik*).

Testinstallation im Brückenbauwerk

Zur Untersuchung realer Störeinflüsse wurde in einer etwa 60 Jahre alten Straßenbrücke, die eine Bundesstraße im Ruhrgebiet trägt, ein Messsystem mit zwei Sensoren installiert. Beide Sensoren zeichnen kontinuierlich mit einer Abtastrate von 5 Hz auf. Die x-Achse ist jeweils in Fahrtrichtung, also näherungsweise horizontal ausgerichtet, die y-Achse zeigt näherungsweise in Richtung der Erdbeschleunigung, also vertikal.

Daraus ergibt sich, dass die x-Achse vor allem sensorinterne Störungen und Bewegungen des Bauwerks abbildet, während die y-Achse zusätzlich den größtmöglichen Einfluss des Straßenverkehrs enthält. Die ungefilterten Messwerte (siehe Abb. 4) bestätigen diese Erwartung. Im Verlauf der x-Komponente ist eine deutliche Ähnlichkeit zum Temperaturverlauf erkennbar. Diese kann sowohl auf temperaturinduzierte Bauwerksbewegungen

als auch auf die Temperaturabhängigkeit des Sensors zurückzuführen sein; eine nachträgliche Trennung der Ursachen ist an dieser Messstelle jedoch nicht möglich. In der y-Komponente ist dieser Zusammenhang aufgrund der Skalierung und der stärkeren Überlagerung durch verkehrsbedingte Störungen weniger direkt sichtbar.

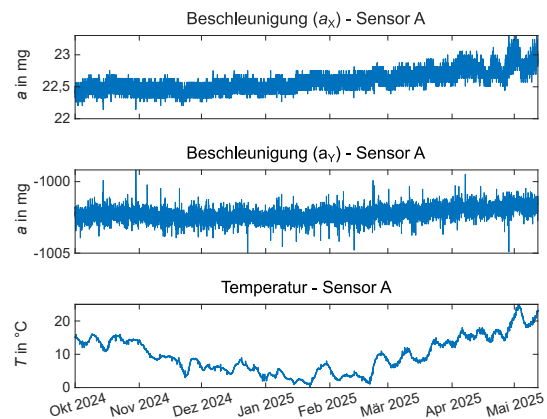


Abb. 4: *Ungefilterte Messwerte der Beschleunigungskomponenten und Sensortemperatur.*

Werden beide Beschleunigungskomponenten geometrisch kombiniert (siehe Abb. 5), treten die verkehrsbedingten Störungen deutlicher hervor. Im Mittel sollte näherungsweise 1 g detektiert werden, die beobachteten Auslenkungen sind primär dem Verkehr zuzuschreiben. Temperaturbedingte Einflüsse können in kurzen Betrachtungszeiträumen aufgrund ihrer deutlich geringeren Änderungsrate vernachlässigt werden. Im Zeitverlauf zeigt sich, dass die Störungen vor allem werktags und tagsüber ausgeprägt sind, während sie an Wochenenden sowie in den Randzeiten des Tages abnehmen. Dieses Verhalten stimmt mit dem zu erwartenden Verkehrsaufkommen überein. Im Spektrogramm konzentrieren sich die dominanten Anteile der Störungen insbesondere auf Frequenzbereiche um 1,1 Hz und 2,2 Hz.

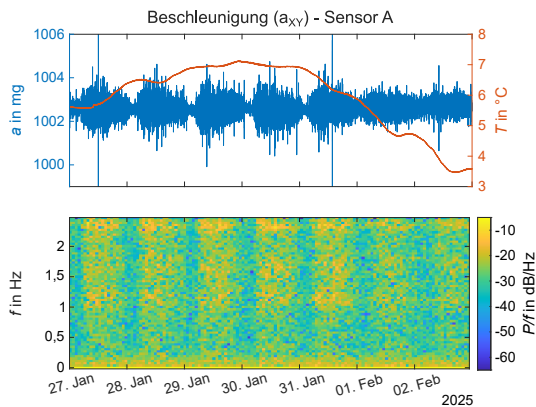


Abb. 5: Oben: Beschleunigung und Temperatur eines Sensors innerhalb einer Woche. Unten: Spektrogramm der gemessenen Beschleunigung.

Für eine detailliertere Analyse werden zwei Zeitfenster im Frequenzbereich verglichen (siehe Abb. 6): Montag von 6:00 bis 10:00 Uhr als Zeitraum hohen Verkehrsaufkommens und Sonntag von 0:00 bis 4:00 Uhr als Zeitraum geringen Verkehrsaufkommens. Im Spektrum zeigt sich ab etwa 20 mHz ein klarer Unterschied zwischen beiden Zuständen. Diese Differenz kann dem Einfluss des Verkehrs zugeschrieben werden.

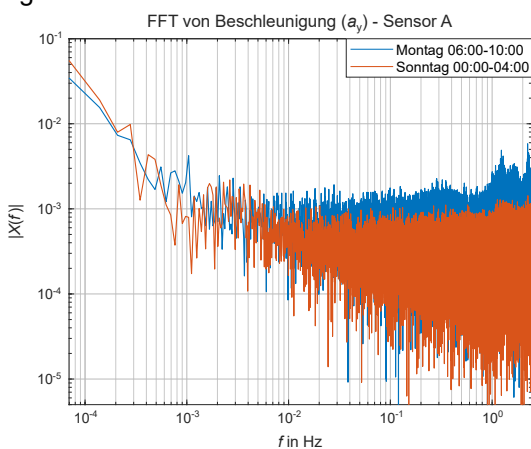


Abb. 6: FFT für verschiedene Zeitbereiche, Einfluss des Verkehrsaufkommens.

Bereits durch den Einsatz einer Tiefpassfilterung mit einem Moving-Average-Filter lässt sich das Rauschen deutlich reduzieren [16]. Die zuvor beschriebenen verkehrsbedingten Störungen sind danach im Spektrogramm deutlich weniger ausgeprägt.

Genauigkeit der Sensorik

Unabhängig von den an der Messstelle beobachteten Störungen ist die tatsächliche Genauigkeit der verwendeten Sensoren unter idealen Umgebungsbedingungen zu bewerten.

Dazu wird betrachtet wie nah der Messwert am wahren Wert liegt. Hierzu wird ein Sensorpaar auf einem drehbar gelagerten Balken montiert. Dieser Balken kann über einen motorisierten Linearpositionierer mit mikrometergenauer Verstellung (*ThorLabs LNR50SE* und *ThorLabs BSC103*) schrittweise bewegt werden. In Abhängigkeit vom Abstand zwischen Drehpunkt und Positionierer lässt sich so durch eine lineare Bewegung eine definierte Rotation erzeugen. Im vorliegenden Laboraufbau entspricht eine translatorische Bewegung von 45 μm einer Rotation von 0,0057 $^\circ$ und damit der angestrebten Zielauflösung.

Die resultierenden Signalverläufe (siehe Abb. 7) zeigen die diskreten Verstellsschritte deutlich und stimmen nach kurzer Einschwingzeit gut mit dem theoretischen Verlauf überein. Damit kann der ausgewählten Sensorik grundsätzlich die geforderte Auflösung und Genauigkeit für die betrachtete Anwendung zugesprochen werden.

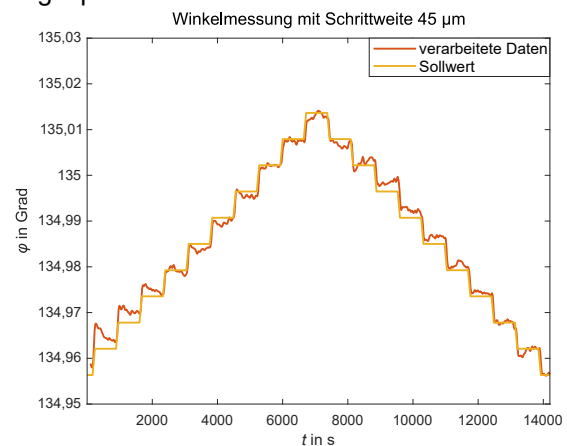


Abb. 7: Verarbeitete Sensordaten (rot) und theoretische Werte (gelb) einer Winkelmessung bei stufenweiser Lageänderung.

Temperaturabhängigkeit

Gemäß den auf Herstellerangaben basierenden Vorüberlegungen stellt die temperaturabhängige Sensitivitätsänderung die wichtigste Ursache für die zu erwartende Messungenauigkeit dar und liegt oberhalb der angestrebten Zielgenauigkeit. Daher wurden mehrere Sensoren in unterschiedlichen Ausrichtungen über einen Temperaturbereich von -30 $^\circ\text{C}$ bis 50 $^\circ\text{C}$ untersucht [14]. Hierzu durchliefen die Sensoren jeweils zwölf Zyklen in einem Klimaschrank des Typs *Weiss WK3* bei konstanter Luftfeuchtigkeit und einer rampenförmigen Temperaturänderungsrate von 0,25 $^\circ\text{C}/\text{min}$. Die geringe Änderungsrate wurde gewählt, um thermische Ausgleichsvorgänge weitgehend zu ermöglichen. Da die Lage der

Sensoren während der Messung konstant gehalten wurde, wird jede Änderung der Messgröße als Störung interpretiert.

Bereits bei der geringen Stichprobengröße (siehe Abb. 8) mit vier Sensoren (Zeilen) in vier Ausrichtungen (Spalten) werden vier wesentliche Eigenschaften sichtbar.

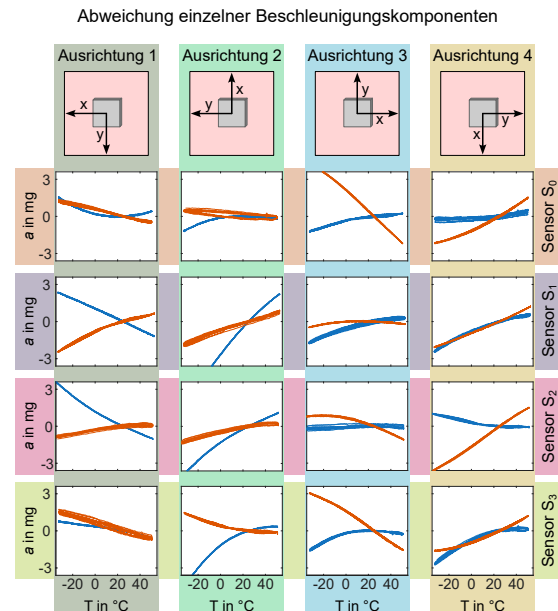


Abb. 8: Abweichung der Beschleunigungskomponenten verschiedener Sensoren (Zeilen) bei verschiedenen Ausrichtungen (Spalten) im Bereich von -30 °C bis 50 °C bei $0,25\text{ °C/min}$; blau: x-Achse, rot: y-Achse.

Erstens weisen einzelne Achsen unterschiedlich starke Temperaturabhängigkeiten auf. Für die in Richtung der Erdbeschleunigung ausgerichtete Achse wurden Abweichungen von bis zu $\pm 0,08\text{ mg/°C}$ beobachtet, diese entsprechen einer relativen Abhängigkeit von $\pm 0,008\%/°\text{C}$ und damit leicht unterhalb der Herstellerangabe. Die horizontale Komponente erreicht Werte bis zu $\pm 0,04\text{ mg/°C}$. Aufgrund des geringeren Absolutwerts entspricht dies jedoch einer vergleichsweise hohen relativen Abweichung. In diesem Fall ist wahrscheinlich der temperaturabhängige Nullpunkt dominanter als die eigentliche Sensitivitätsänderung.

Zweitens führt eine Spiegelung der Achse nicht zwangsläufig zu einer spiegelbildlichen Signalantwort. Zwar wechselt bei der vertikalen Achse das Vorzeichen der Temperaturabhängigkeit häufig mit der Spiegelung der Einbaulage, die Stärke der Abhängigkeit kann jedoch deutlich variieren. Bei der horizontalen Achse zeigt sich ein solcher Vorzeichenwechsel nicht in allen Fällen. Diese Betrachtung ist unter anderem relevant für eine Beurteilung zum Einsatz von zwei Sensoren in einer differentiellen

Auswertung. Bei gegenphasigen Signalen können Störungen die auf beiden Signalen wirken besonders effektiv kompensiert werden. Drittens sind die Verläufe teilweise nur eingeschränkt linear und können vereinzelt sogar einen Scheitelpunkt innerhalb des untersuchten Temperaturbereichs aufweisen.

Viertens lässt sich weder eine bestimmte Ausrichtung noch ein bestimmter Sensor als grundsätzlich besonders vorteilhaft identifizieren.

Insgesamt zeigt sich somit, dass jede Sensorachse in jeder Ausrichtung eine individuelle Temperaturabhängigkeit besitzt. Ein Sensor kann in einer Einbaulage eine sehr geringe Streuung aufweisen und in einer anderen Lage gleichzeitig zu den schwächsten Exemplaren der Stichprobe gehören. Soll die geforderte Genauigkeit über den zu erwartenden Temperaturbereich in der realen Anwendung erreicht werden, muss daher jeder Sensor in seiner späteren Installationsausrichtung bei verschiedenen Temperaturwerten charakterisiert werden.

Als Lösungsansatz für dieses Teilproblem wird auf Basis der vorangegangenen Referenzmessungen die Temperaturabhängigkeit mithilfe eines polynomialen Modells beschrieben [17]. Der temperaturabhängige Anteil wird anschließend für jeden Messpunkt aus dem Sensorsignal subtrahiert.

Die Auswertung zeigt eine deutliche Reduktion der Streuung nach der Kompensation (siehe Abb. 9). Besonders effizient ist der gewählte Ansatz bei der vertikalen Achse, die zuvor eine höhere Varianz aufwies. Bei der horizontalen Achse ist die Kompensation hingegen weniger wirksam. Nach der Stauchung der Daten wird dort sichtbar, dass sich teilweise Bahnkurven ausbilden, die an Lissajous-Figuren erinnern. Dies deutet auf eine verzögerte Systemantwort beziehungsweise auf ein Hystereseverhalten hin [18, 19]. Offenbar hängt das Verhalten in diesen Fällen zusätzlich davon ab, ob ein Aufheiz- oder ein Abkühlvorgang vorliegt. Aufgrund der geringen verbleibenden Abweichungen wird ihr Einfluss auf den berechneten Winkel jedoch im Hinblick auf die geforderte Auflösung als vernachlässigbar bewertet.

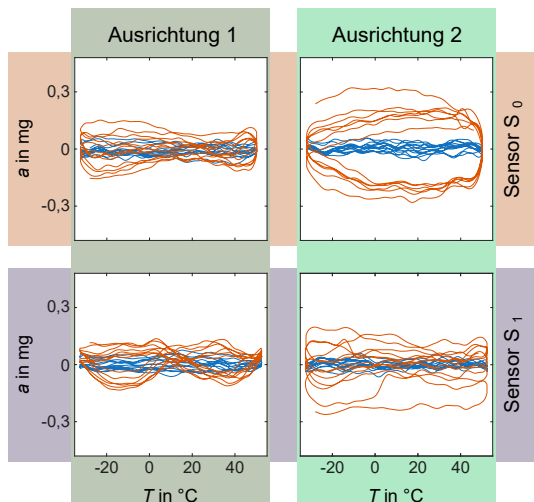


Abb. 9: Abweichung von kompensierten Beschleunigungskomponenten bei verschiedenen Ausrichtungen im Bereich von -30 °C bis 50 °C bei $0,25\text{ °C/min}$; blau: x-Achse, rot: y-Achse.

Gesamtkonzept

Gemäß den ermittelten Einflussfaktoren und darauf abgestimmten einzelnen Lösungsansätzen ergibt sich folgende Signalkette mit mehreren Verarbeitungsschritten, von der Aufnahme der Rohdaten bis zur berechneten Neigung (siehe Abb. 10). Zunächst wird auf die gemessenen Beschleunigungskomponenten eine Tiefpassfilterung angewendet, um allgemeine Störungen und insbesondere den Einfluss des Verkehrs zu reduzieren. Anschließend erfolgt die Temperaturkompensation auf Basis des polynomialen Modells und einer Referenzmessung in identischer Ausrichtung. Im Rahmen dieser Referenzmessung kann zusätzlich der integrierte Temperatursensor des Beschleunigungssensors kalibriert werden. Dies ist für die reine Messgenauigkeit nicht zwingend erforderlich, verbessert jedoch die Vergleichbarkeit der kompensierten Messwerte. Abschließend werden aus den kompensierten Beschleunigungskomponenten die Neigungswerte berechnet und im letzten Schritt mehrere Sensoren zur weiteren Erhöhung der Genauigkeit kombiniert. Gemäß der Anzahl an genutzten Sensoren sinkt die theoretische Empfindlichkeit des Gesamtsystems und Störungen einzelner Sensoren weisen einen geringeren Beitrag zum ausgegebenen Wert des Sensorsystems. Es werden sowohl Kombinationen aus zwei und vier Sensoren betrachtet.

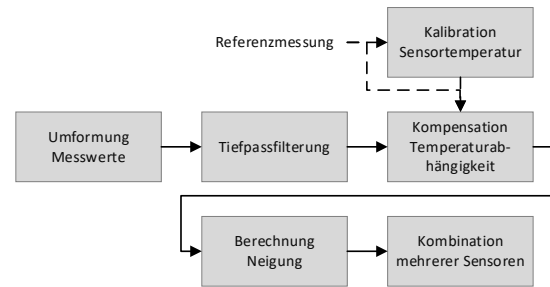


Abb. 10: Vollständige Signalkette von der Aufnahme der Beschleunigungswerte bis zur berechneten Neigung.

Ergebnisse

Das Sensorsystem liefert unter realen Umgebungsbedingungen zuverlässig kontinuierliche Daten. Die durch den Verkehr hervorgerufenen Störungen sind im Frequenzspektrum klar identifizierbar und erfordern zusätzliche Maßnahmen. Da diese deutlich höherfrequent sind als die zu beobachtenden Lageänderungen, lassen sie sich durch eine ressourcenschonende Tiefpassfilterung wirksam mindern (siehe Abb. 11). Beim Vergleich zwischen den Rohdaten und den gefilterten Daten beträgt das Signal-Rausch-Verhältnis (SNR) bei etwa 55 dB.

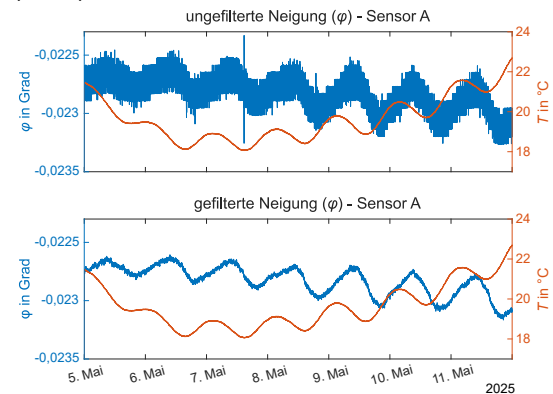


Abb. 11: Berechnete Neigung auf Basis ungefilterter und gefilterter Beschleunigungen von Messdaten innerhalb des Brückenbauwerks.

Die ausgewählten Sensoren erreichen in einer kontrollierten Laborumgebung bei einer Empfindlichkeit von $0,061\text{ mg/LSB}$ nach kurzer Einschwingzeit und bei konstanter Temperatur eine Winkelauflösung unterhalb der geforderten Auflösung von $0,0057^\circ$. Damit ist ihre grundsätzliche Eignung für die vorgesehene Anwendung belegt. Durch eine Referenzmessung bei konstanter Ausrichtung lässt sich außerdem die Temperaturabhängigkeit im erweiterten Bereich von -30 °C bis 50 °C deutlich reduzieren. Die folgende Tabelle zeigt die mittlere, minimale und maximale

Standardabweichung der zuvor präsentierten Messreihe mit vier verschiedenen Sensoren in jeweils vier Ausrichtungen nach den einzelnen Schritten der Verarbeitungskette. Dies bestätigt die Effizienz der Signalkette. Bei der Kombination mehrerer Sensoren sollten Sensoren mit gleicher oder entgegengesetzter Ausrichtung verwendet werden; in anderen Fällen führt dies in der Regel zu keiner weiteren Verbesserung der Auflösung.

Tab. 1: Übersicht der Standardabweichung s (Mittelwert, Minimum, Maximum) der berechneten Neigungen nach den einzelnen Verarbeitungsschritten.

	$\bar{s}_\varphi$	$s_{\min,\varphi}$	$s_{\max,\varphi}$
ungefiltert	2,12E-01	9,23E-03	1,45E+00
gefiltert (TP)	3,13E-02	7,13E-03	5,34E-02
Temperatur kompensiert	5,30E-03	3,21E-03	1,02E-02
Kombination zweier Sensoren	3,83E-03	2,50E-03	6,97E-03

Diskussion

Die Ergebnisse bestätigen, dass die gewählte Sensorik nach geeigneter Kalibrierung die geforderte Genauigkeit grundsätzlich erreichen kann. Zugleich wird deutlich, dass die Kalibrierung für jeden Sensor und dessen konkrete spätere Einbaulage individuell erfolgen muss. Dieser zusätzliche Aufwand erscheint jedoch beherrschbar, sofern ein standardisierter Referenzprozess etabliert wird. Bei geringen Beschleunigungswerten zeigt die verwendete Kompensation eine geringere Wirksamkeit. Ursache hierfür ist vermutlich, dass in diesem Bereich der temperaturabhängige Nullpunkt (Zero-g level change vs. temperature) stärker ins Gewicht fällt als die primär betrachtete Sensitivitätsänderung. Darüber hinaus wurde sichtbar, dass auch Hysterese-Effekte eine Rolle spielen können. Dies sollte in zukünftigen Modellen stärker berücksichtigt werden. Der zunächst stärkere Einfluss des vorherrschenden Verkehrs an der Messstelle lässt sich dagegen bereits mit einfachen Filterverfahren wirksam auf ein Minimum reduzieren.

Ausblick

Die vorliegenden Untersuchungen zeigen innerhalb einer noch kleinen Stichprobe eine deutliche Variabilität zwischen einzelnen Sensoren. Da diese Unterschiede zwar reproduzierbar, aber individuell ausgeprägt sind, sollte der Ansatz künftig an einer größeren Anzahl von Sensoren validiert werden, um seine allgemeine Übertragbarkeit zu belegen. In

diesem Zusammenhang sind auch Untersuchungen mit alternativen Sensortypen sinnvoll.

Darüber hinaus ergeben sich mehrere weiterführende Fragestellungen. Zu klären ist insbesondere, wie stark sich der zeitliche Umfang einer Referenzmessung in kontrollierter Umgebung reduzieren lässt, ohne die erreichbare Genauigkeit wesentlich zu verschlechtern. Ebenso ist zu untersuchen, ob eine praxisnähere Referenzmessung mit geringerem apparativem Aufwand möglich ist, beispielsweise durch gezielte Erwärmung des Sensorelements. In diesem Zusammenhang wäre auch zu prüfen, ob auf Temperaturphasen im niedrigen oder negativen Bereich verzichtet werden kann. Schließlich sollten die beobachteten Hystereseeffekte systematisch untersucht und nach Möglichkeit in zukünftige Kompensationsmodelle integriert werden.

Literaturnachweis

- [1] B. Novák *et al.*, „Erfahrungssammlung Monitoring für Brückenbauwerke“ 2024, doi: 10.60850/bericht-b197.
- [2] I. Hindersmann, „Monitoring von Straßenbrücken - aktueller Einsatz und zukünftige Anwendung“ *e-Journal of Nondestructive Testing*, Jg. 29, Nr. 04, 2024.
- [3] F. Wedel, S. Pitters, R. Herrmann, R. Schneider, F. Hille und I. Hindersmann, „Guideline for the strategic application of monitoring of road bridges in Germany“ *eJNDT*, Jg. 29, Nr. 7, 2024, doi: 10.58286/29582.
- [4] S. Selvakumaran *et al.*, „Combined InSAR and Terrestrial Structural Monitoring of Bridges“ *IEEE Trans. Geosci. Remote Sensing*, Jg. 58, Nr. 10, S. 7141–7153, 2020, doi: 10.1109/TGRS.2020.2979961.
- [5] M. Reiterer, L. Praxmarer und A. Kammersberger, „Maintenance und Monitoring. Dauermonitoring zur Optimierung der Nutzungsdauer einer Brücke“ *Beton und Stahlbetonbau*, Jg. 103, Nr. 2, S. 103–111, 2008, doi: 10.1002/best.200700599.
- [6] R. Geier, T. Mack und E. Krebes, „Monitoring der Seitenhafenbrücke in Wien“ *Beton und Stahlbetonbau*, Jg. 109, Nr. 7, S. 486–495, 2014, doi: 10.1002/best.201400023.
- [7] F. Wedel, S. Pitters, F. Hille, R. Herrmann und R. Schneider, „Leitfaden – Strategischer Einsatz von Monitoring für Ingenieurbauwerke“ 2025, doi: 10.60850/bericht-b210.
- [8] I. Hindersmann, „Monitoring bei Straßenbrücken – Eine Kosten-Nutzen-Betrachtung“ *e-Journal of Nondestructive Testing*, Jg. 31, Nr. 03, 2026, doi: 10.58286/32740.
- [9] W. Ostachowicz, R. Soman und P. Malinowski, „Optimization of sensor placement for structural health monitoring: a review“ *Structural Health Monitoring*, Jg. 18, Nr. 3, S.

- 963–988, 2019, doi:
10.1177/1475921719825601.
- [10] C. Bedon, E. Bergamo, M. Izzi und S. Noè, „Prototyping and Validation of MEMS Accelerometers for Structural Health Monitoring—The Case Study of the Pietratagliata Cable-Stayed Bridge“ *JSAN*, Jg. 7, Nr. 3, S. 30, 2018, doi: 10.3390/jsan7030030.
 - [11] S. Morichika, H. Sekiya, Y. Zhu, S. Hirano und O. Maruyama, „Estimation of Displacement Response in Steel Plate Girder Bridge Using a Single MEMS Accelerometer“ *IEEE Sensors J.*, Jg. 21, Nr. 6, S. 8204–8208, 2021, doi: 10.1109/JSEN.2021.3051697.
 - [12] X. Ru, N. Gu, H. Shang und H. Zhang, „MEMS Inertial Sensor Calibration Technology: Current Status and Future Trends“ *Micromachines*. doi: 10.3390/mi13060879.
 - [13] C. Nagel, F. Ante, M. Putnik, J. Classen und J. Mehner, „Characterization of Temperature Gradients on MEMS Acceleration Sensors“ *Procedia Engineering*, Jg. 168, S. 888–891, 2016, doi: 10.1016/j.proeng.2016.11.298.
 - [14] M. Catelani *et al.*, „3-Axis accelerometer characterization: A dynamic analysis varying the operating temperature“ *Measurement: Sensors*, Jg. 38, S. 101599, 2025, doi: 10.1016/j.measen.2024.101599.
 - [15] S. Łuczak, „Guidelines for tilt measurements realized by MEMS accelerometers“ *Int. J. Precis. Eng. Manuf.*, Jg. 15, Nr. 3, S. 489–496, 2014, doi: 10.1007/s12541-014-0362-5.
 - [16] A. I. Ozdagli, F. Moreu, J. Gomez, S. Vemuganti und P. Garg, „Data Fusion of Accelerometers with Inclinometers for Reference-free High Fidelity Displacement Estimation“ *e-Journal of Nondestructive Testing*, Jg. 21, Nr. 08, 2016.
 - [17] Y. Gunhan und D. Unsal, „Polynomial degree determination for temperature dependent error compensation of inertial sensors“ in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, 2014, S. 1209–1212, doi: 10.1109/PLANS.2014.6851494.
 - [18] H. J. Kim und H. K. Jung, „Temperature Hysteresis Calibration Method of MEMS Accelerometer“ *Sensors*. doi: 10.3390/s25196131.
 - [19] F. Gulmammadov, „Analysis, modeling and compensation of bias drift in MEMS inertial sensors“ in *2009 4th International Conference on Recent Advances in Space Technologies*, Istanbul, Jun. 2009 - Jun. 2009, S. 591–596, doi: 10.1109/RAST.2009.5158260.

Miniaturisierter Fluoreszenzsensor auf Ru(dpp)_3^{2+} -Basis zur Echtzeitbestimmung der Atemsauerstoffkonzentration

Jonas Steinebrunner¹, Christian Weber^{1,2}, Maximilian Willaredt¹, Dr. Katrin Schmitt^{1,2}, 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.

Zusammenfassung

Im vorliegenden Beitrag wird ein fluoreszenzbasierter Sauerstoffsensor vorgestellt, der sich in einen standardisierten T-Verbinder (Innendurchmesser 15 mm) für Atemmasken und Beatmungssysteme integriert werden kann. Dies ermöglicht eine einfache Einbindung in bestehende Monitoringsysteme möglich. Das Messprinzip beruht auf dem Effekt der Fluoreszenzlöschung (Quenching) durch molekularen Sauerstoff. Als Farbstoff wird ein Rutheniumkomplex (Ru(dpp)_3^{2+}) verwendet, der in eine Sol-Gel-Matrix eingebettet ist. Zur Anregung des Farbstoffes wird eine blaue LED mit einer zentralen Emissionswellenlänge von 450 nm verwendet. Der Farbstoff Ru(dpp)_3^{2+} absorbiert das Anregungslicht und emittiert eine charakteristische Fluoreszenz mit einem Intensitätsmaximum bei ca. 610 nm. Diese Verschiebung des Emissionsspektrums gegenüber dem Anregungslicht (Stokes-Shift) bildet die Grundlage für das angewendete Quenching-Messprinzip. Ein Langpassfilter mit einer Grenzwellenlänge von 550 nm blockiert das reflektierte, energiereichere Anregungslicht, sodass nur das von der Probe emittierte Fluoreszenzlicht von einem Photodetektor erfasst wird. Für eine effiziente Lichtführung sind zwei plankonvexe Linsen mit einem Durchmesser von 6 mm in den Sensorkopf integriert. Eine befindet sich vor der LED und eine vor dem Detektor, um sowohl das Anregungs- als auch das Fluoreszenzlicht möglichst effektiv zu fokussieren. Die gesamte optische Anordnung einschließlich des Langpassfilters ist kompakt im Sensorkopf eingebaut, der mittels additiver Fertigung (3D-Druck) hergestellt wurde.

Keywords: Fluoreszenzlöschung, Sauerstoffsensor, Atemgasanalyse, Sol-Gel, Ru(dpp)_3^{2+}

Einleitung

Der menschliche Organismus kann nur wenige Minuten ohne Sauerstoff überleben. Entsprechend zentral ist die Beurteilung des O_2 -Status in nahezu allen Bereichen der Medizin von der Notaufnahme über den Rettungsdienst bis hin zur hausärztlichen Versorgung. Eine ausreichende O_2 -Versorgung gewährleistet die kontinuierliche zelluläre Energieproduktion, insbesondere in hochmetabolisch aktiven Organen wie Herz, Gehirn und Leber. Die Sauerstoffaufnahme (VO_2) ist ein direkter Indikator für den Energiebedarf und die metabolische Aktivität von Geweben [1]. Eine effiziente VO_2 setzt einen intakten Gasaustausch in der Lunge, eine ausreichende kardiovaskuläre Versorgung und eine funktionierende zelluläre Sauerstoffverwertung voraus. Störungen in einem dieser Systeme führen zu einer verminderten Energieproduktion und können frühzeitig auf pathophysiologische Veränderungen hinweisen, was die Notwendigkeit einer präzisen und kontinuierlichen Überwachung der O_2 -Versorgung unterstreicht [2].

Derzeit erfolgt die klinische Beurteilung der Sauerstoffversorgung überwiegend über die Pulsoxymetrie, welche die arterielle Sauerstoffsättigung (SpO_2) am Finger misst. Diese Methode ist einfach zu handhaben, liefert Echtzeitwerte, ist nicht-invasiv und kostengünstig [3]. Daher gehört der Pulsoxymetrie-Sensor von der Notfallmedizin bis zur Hausarztpraxis zum medizinischen Standard-equipment. Allerdings liefert der Sensor nur indirekte Informationen über den Gasaustausch in der Lunge und erlaubt keine Aussage über die tatsächliche VO_2 . Außerdem ist die Messung anfällig für verschiedenste Störfaktoren, wie beispielsweise eine eingeschränkte periphere Durchblutung, starke Hautpigmentierung, Bewegungsartefakte oder Nagellack [4]. Aufgrund dessen ist bei unklaren Messungen häufig eine Blutanalyse notwendig. Diese gelten als klinischer Goldstandard zur Bewertung des metabolischen Status. Nachteile ergeben sich jedoch aus der invasiven Entnahme, der zeitaufwendigen und kostenintensiven Laboranalyse sowie dem Verlust der Echtzeitmessung gegenüber einer

Momentaufnahme [5]. Bislang existieren keine portablen Echtzeit Monitoring-Systeme, die die Pulsoxymetrie durch eine gleichzeitige Messung des Sauerstoffgehalts in der Atemluft ergänzen.

Im vorliegenden Beitrag wird ein miniaturisierter Fluoreszenzsensor auf Basis des Ruthenium-komplexes Ru(dpp)_3^{2+} vorgestellt, der in einen standardisierten T-Verbinder (Innendurchmesser 15 mm) für Beatmungsmasken und Beatmungssysteme integrierbar ist. Der Schwerpunkt des vorliegenden Beitrags liegt auf dem physikalischen Messprinzip, dem Herstellungsprozess der Sensormembran sowie dem optischen und mechanischen Aufbau des Systems. Außerdem werden erste experimentelle Labormessungen einer Machbarkeitsstudie vorgestellt.

Messprinzip und theoretische Grundlagen

Fluorophore besitzen die Charakteristik Photonen nach Absorption der Strahlungsenergie, zum Teil strahlend zu emittieren. Das emittierte Licht wird Lumineszenz genannt und ist durch die Stokes-Verschiebung (längere Wellenlänge) vom energiereicheren Anregungslicht unterscheidbar. Ein Spektralfilter sorgt in der Praxis dafür, dass ausschließlich Fluoreszenzlicht von einem Detektor gemessen wird. Die Bestimmung der Sauerstoffkonzentration erfolgt über das Prinzip des Fluoreszenzquenchings, bei dem molekularer Sauerstoff als dynamischer Quencher die Emission eines geeigneten Indikatorfarbstoffs abschwächt [6]. Beim Ru(dpp)_3^{2+} kommt es durch den Heavy-Atom-Effekt des Rutheniums zu ultraschnellem Intersystem Crossing (ISC) vom initial angeregten Singulettzustand zum langlebigen Triplettzustand, der Phosphoreszenz (~ 610 nm, $\tau \approx 5$ μs) (vgl. Abb. 1). Molekularer Sauerstoff interagiert durch seinen Triplett-Grundzustand ($^3\Sigma$) strahlungslos mit diesem Triplettzustand des Fluorophors, wodurch Phosphoreszenzintensität -und lebensdauer konzentrationsabhängig reduziert werden [7].

Das maximale Messsignal bei der Fluoreszenzquenching-Messtechnik tritt somit bei 0 Vol.-% O_2 auf. Der Zusammenhang zwischen Fluoreszenzintensität und Sauerstoffkonzentration wird durch die Stern-Volmer-Gleichung beschrieben:

$$I_0/I = 1 + K_{SV}[\text{O}_2], \quad (1)$$

wobei I_0 die Fluoreszenzintensität in Abwesenheit von Sauerstoff, I die gemessene Intensität bei einer gegebenen O_2 -Konzentration und K_{SV} die Stern-Volmer-

Konstante bezeichnet. Diese ergibt sich aus dem Produkt der bimolekularen Löschratenkonstante k_q und der natürlichen Lebensdauer T_0 :

$$K_{SV} = k_q + T_0. \quad (2)$$

Eine hohe natürliche Lebensdauer T_0 führt gemäß Gl. (2) zu einer hohen Stern-Volmer-Konstante und damit zu hoher Sauerstoffempfindlichkeit.

Die Sauerstoffabhängigkeit lässt sich sowohl über die Fluoreszenzintensität als auch über die Fluoreszenzlebensdauer beschreiben. Im vorliegenden Sensorsystem erfolgt die Auswertung über die Änderung der Fluoreszenzintensität.

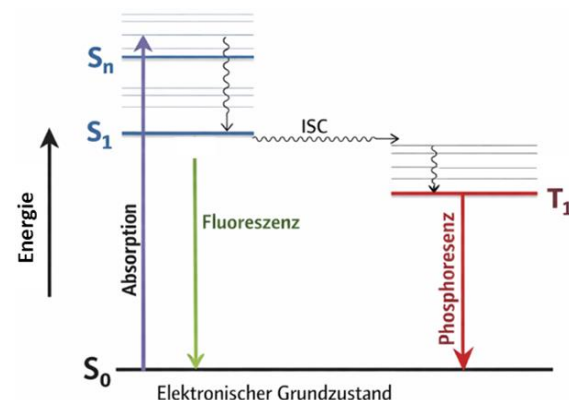


Abb. 1: Darstellung der physikalischen Prozesse nach optischer Anregung in einem Jablonski-Diagramm. Gezeigt sind Absorption aus dem elektronischen Grundzustand S_0 in angeregte Singulettzustände S_n , strahlungslose Relaxation zum Zustand S_1 , Fluoreszenz, Intersystem Crossing (ISC) in den Triplettzustand T_1 sowie Phosphoreszenz.

Farbstoff Ru(dpp)_3^{2+} und Sol-Gel Einbettung

Ru(dpp)_3^{2+} ist ein bereits vielfach eingesetzter Fluoreszenzfarbstoff für die optische Sauerstoffmessung. Für den Sensoreinsatz muss der Farbstoff auf einem Festkörpersubstrat immobilisiert werden, wobei die Zugänglichkeit für Sauerstoffmoleküle erhalten bleiben muss. Die Sol-Gel-Technologie auf Basis von n-Propyltriethoxysilan (PTEOS) stellt hierfür eine bewährte Methode dar. Im Sol-Gel-Prozess entsteht durch Hydrolyse und anschließende Kondensation von PTEOS ein nanoporöses SiO_2 -Netzwerk. Das Ru(dpp)_3^{2+} wird bereits vor der Gelierung der Ausgangslösung zugesetzt und während der Netzbildung in die entstehende Sol-Gel-Matrix eingebettet. Anschließend wird das Gel bei erhöhter Temperatur ausgehärtet. Die

entstehenden Nanoporen ermöglichen eine schnelle Diffusion von Sauerstoff zu den immobilisierten Farbstoffmolekülen. Schichtdicke, PTEOS-Konzentration, Katalysator-pH und Farbstoff-zu-Precursor-Verhältnis beeinflussen dabei die optische Dichte sowie Empfindlichkeit und Ansprechzeit der Sensor-membran.

Sensordesign

Das Sensordesign wurde mittels additiver Fertigung (FDM) aus Polyactide (PLA) realisiert (Vgl. Abb. 2). Es wurde ein modularer Aufbau entwickelt, der aus mehreren Baugruppen besteht. Dazu gehören der Sensorkopf mit integrierter LED und Photodetektor sowie zugehörige Steuerelektronik (PCB). Ein Probenhalter aus Acrylnitril-Butadien-Styrol ABS fixiert unterschiedlichste Substratträger durch eine passgenaue Vertiefung.

Bereits im ersten Demonstrator wurde das Konzept so umgesetzt, dass der Sensor in einen standardisierten medizinischen T-Verbinder mit einem Innendurchmesser von 15 mm für Atemmasken und Monitoringsysteme integriert werden kann (vgl. Abb. 3 (a)). Zur Anregung der Probe wird eine blaue LED mit einer Emissionswellenlänge von 450 nm verwendet. Der Farbstoff Ru(dpp)_3^{2+} absorbiert das Anregungslicht und emittiert eine charakteristische Fluoreszenz mit einem Intensitätsmaximum bei ca. 610 nm. Ein Langpassfilter (15-198 Edmund Optics, USA) mit einer Grenzwellenlänge von 550 nm unterdrückt das reflektierte, energiereichere Anregungslicht, sodass nur das von der Probe emittierte Fluoreszenzlicht von einem Photodetektor (SFH 2700 A01 AMS Osram, Austria) erfasst wird. Für eine effiziente Lichtführung sind zwei plankonvexe Linsen (LA4280 Thorlabs, USA) mit einem Durchmesser von 6 mm in den Sensorkopf integriert, eine vor der LED und eine vor dem Detektor, um sowohl das Anregungs- als auch das Fluoreszenzlicht möglichst effektiv zu bündeln.

Experimentelle Durchführung

Das Sol-Gel wurde gemäß [8] unter Verwendung von n-Propyltriethoxysilan (PTEOS), Ethanol, Wasser und Salzsäure im molaren Verhältnis von 1: 6,25: 4: 0,007 hergestellt. Der Farbstoff Ru(dpp)_3^{2+} wurde in einer Konzentration von 2,5 g/l in Bezug auf das Gesamtvolumen des Sols beigemischt.

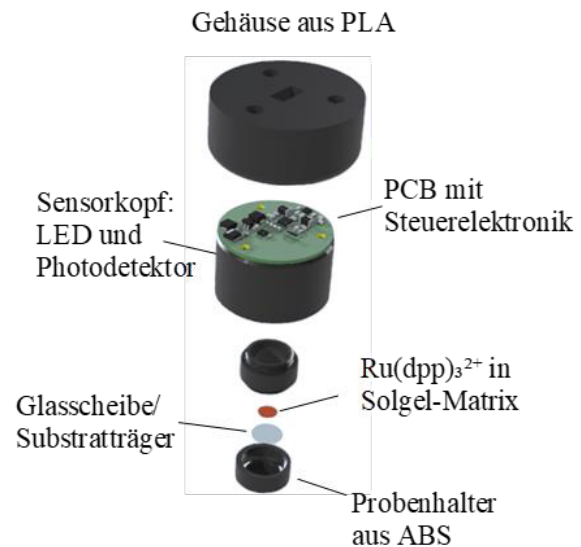


Abb. 2: Schematische Explosionszeichnung des Demonstrators. Dieser besteht aus mehreren Einzelkomponenten. Hervorzuheben sind der Sensorkopf mit LED und Photodetektor sowie die Steuerelektronik auf einer passenden Leiterplatte. Der Probenhalter aus ABS ist mit einem integrierten Substratträger aus Glas ausgestattet.

Anschließend wurden 2 μL der fertigen Farbstoff-Matrix mit einer Mikropipette in die Mitte des Substrats aufgebracht und 24 Stunden bei 110 °C ausgehärtet. Es wurden hierfür vier verschiedene Substratträger untersucht: Aluminium, Glas, PLA und ABS. Abb. 4 zeigt eine mikroskopische Aufnahme der ausgehärteten Ru(dpp)_3^{2+} -Sol-Gel-Mischung für jedes Substratmaterial.

Messparameter

Die LED wurde mit einer Modulationsfrequenz von 5 kHz und einer Stromamplitude von 20 mA betrieben. Die Fluoreszenzsignale wurden mittels eines Lock-In-Algorithmus mit einer Zeitkonstante von 10 ms digital verarbeitet. Der Sensor wurde für alle Messreihen in eine additiv hergestellte Messkammer integriert (Vgl. Abb. 4 (b)).

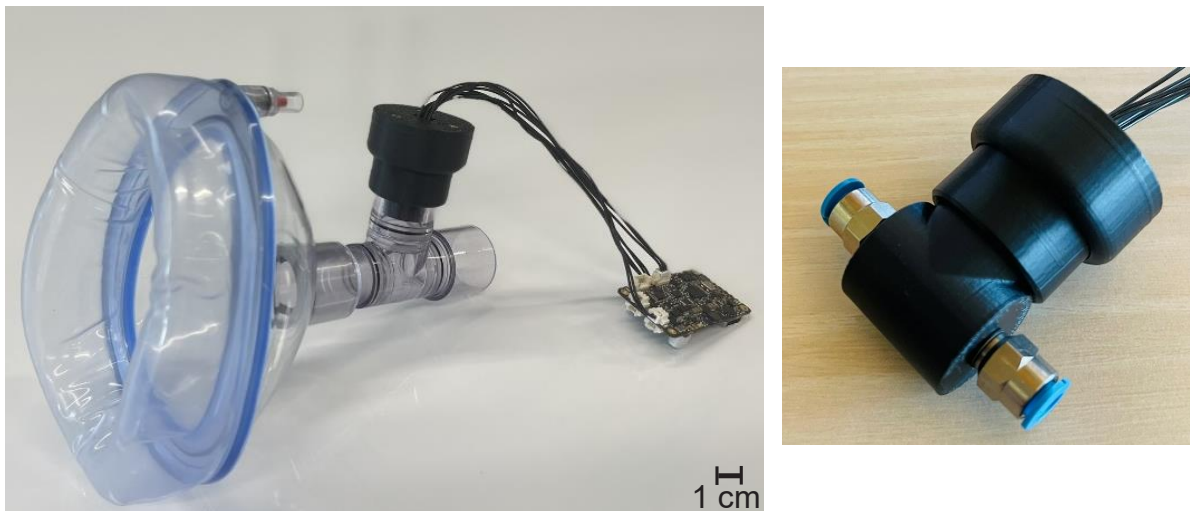


Abb. 3 (a) Foto des Sensors, der in einen medizinischen T-Verbinder integriert und an eine Beatmungsmaske angeschlossen ist. (b) Messaufbau für die ersten Funktionstests.

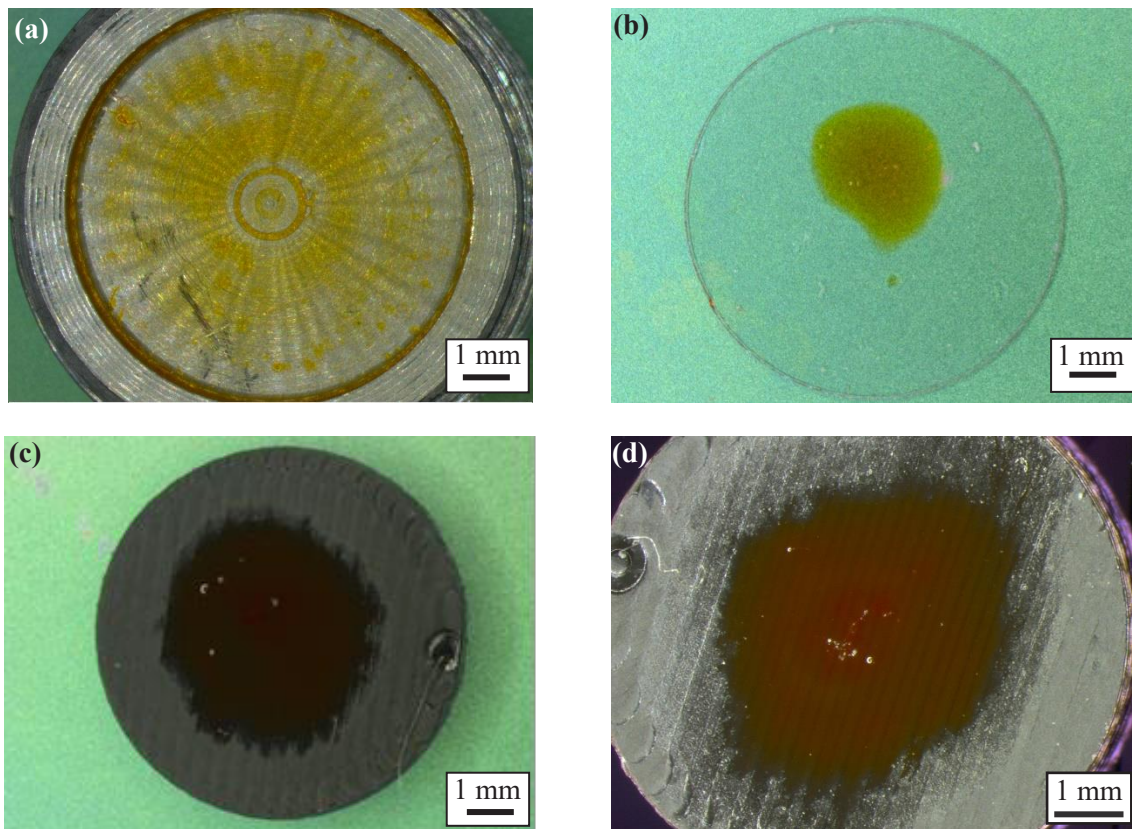


Abb. 4: Lichtmikroskopische Aufnahmen der Ru(dpp)_3^{2+} -Sol-Gel-Schicht auf verschiedenen Substratträgern: (a) Aluminium, (b) Glas, (c) PLA und (d) ABS. Die Aufnahmen wurden mit einem ZEISS Smartzoom 5 Mikroskop aufgenommen.

Ergebnisse und Diskussion

Im Rahmen eines Funktionstest wurde die Sauerstoffkonzentration im Bereich von 0 bis 20 Vol.-% O₂ in N₂ variiert. Hierfür wurden Massendurchflussregler (MFC, Bronkhorst, Deutschland) und zwei zertifizierte Gasflaschen (Westfalen, Deutschland) zum verwendet: eine mit reinem Sauerstoff und eine mit reinem Stickstoff. Während der Messung wurde der Gasfluss konstant auf 2 l/min eingestellt. Die Sauerstoffkonzentration wurde schrittweise auf 0, 5, 10 und 20 Vol.-% eingestellt, während die Intensität der Fluoreszenzlöschung kontinuierlich gemessen wurde. Das Fluoreszenzsignal wurde mit einer Messrate von 1 Hz aufgezeichnet (Abb. 5).

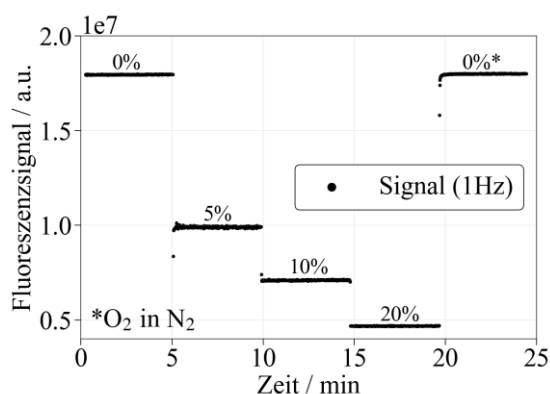


Abb. 5: Fluoreszenzsignal bei einer Messrate von 1 Hz aufgetragen über die Zeit bei unterschiedlichen O₂-Konzentration (0,5,10,20 Vol.-% O₂ in N₂) bei Raumtemperatur.

Der Funktionstest zeigt neben der grundsätzlichen Sauerstoffempfindlichkeit auch ein schnelles Ansprechverhalten des Sensors. Es zeigt sich, dass bei einer Flussrate von 2 l/min kein einziger Datenpunkt zwischen den variierenden O₂-Konzentrationsstufen liegt. Aufgrund der relativ niedrigen Messrate (1 Hz) lässt sich daraus lediglich schließen, dass die Antwortzeit unter 1 s liegt. Trotzdem deutet es daraufhin, dass die Sensorreaktion kein limitierender Faktor für die Echtzeitmessung im Atem darstellt. Die bei der Atmung auftretenden Volumenströme liegen über dem im Experiment eingestellten Wert von 2 l/min. Das Atemvolumen eines Erwachsenen beträgt typischerweise zwischen 6 bis 40 l/min [9]. Es ist daher zu erwarten, dass der Sensor in der praktischen Anwendung eine Antwortzeit deutlich unter 1 s aufweist, da ein höherer Gasfluss den Austausch in der Messkammer zusätzlich beschleunigt.

Zur Untersuchung des Einflusses des Substratträgers auf die Sensorantwort wurden Glas, Aluminium sowie die additiv gefertigten Kunststoffe ABS und PLA als Trägermaterialien für die Ru(dpp)₃²⁺-Sol-Gel-Kombination eingesetzt. Die Kunststoffe wurden dabei mittels 3D-Druck hergestellt. Für diese Messreihe wurde die Sauerstoffkonzentration zwischen 16 und 21 Vol.-% O₂ in N₂ variiert. Diese Konzentrationswerte wurden gewählt, da 21 Vol.-% in etwa dem Sauerstoffgehalt der Umgebungsluft entspricht, während 16 Vol.-% O₂ einen typischen Sauerstoffgehalt der Ausatemluft darstellt [9]. Die Differenz zwischen beiden Werten entspricht näherungsweise dem vom Körper aufgenommenen Sauerstoffanteil. Zur Bewertung der Sensorantwort auf den jeweiligen Substratträger wurde das Verhältnis von Signaländerung zu Rauschen herangezogen (Gl. 3).

$$SNR_{\Delta} = (S_0 - S_1) / \sigma_1 \quad (3)$$

Dabei ist S₀ der gemittelt Fluoreszenzwert über 100 Messpunkte (100 s) bei einer Sauerstoffkonzentration von 16 Vol.-%, S₁ der Fluoreszenzwert (Mittelwert, 100 s) bei einer Sauerstoffkonzentration von 21 Vol.-% und σ_1 die Standardabweichung des Mittelwerts bei der 21 Vol.-% O₂-Stufe. Die Ergebnisse der Berechnungen sind in Tabelle 1 aufgelistet.

Tab. 1: Verhältnis zwischen Signaländerung (16 zu 21 Vol.-% O₂ in N₂) und Rauschen für vier verschiedene Substratträger.

Substrat	SNR _Δ
Alu	204
Glas	131
PLA	75
ABS	202

Es zeigt sich, dass Alu und ABS mit einem SNR_Δ von ca. 200 das beste Verhältnis liefert und damit für die beste Auflösung sorgt. Besonders ABS erweist sich als interessanter Substratträger, da sowohl die Material- als auch die Fertigungskosten im Vergleich zu Aluminium und Glas deutlich geringer sind. PLA hingegen liefert als Substratträger die schlechtesten Ergebnisse, das Signal erscheint bereits beim „Eye-Test“ deutlich instabiler (Vgl. Abb. 6 (c)). Dies ist darauf zurückzuführen, dass der Farbstoff auf PLA nicht unter den erforderlichen Aushärtungsbedingungen verarbeitet werden konnte, da PLA lediglich eine Wärmeformbeständigkeit von etwa 50 bis 60 °C

aufweist und somit keine Aushärtung bei Temperaturen über 100 °C ermöglicht.

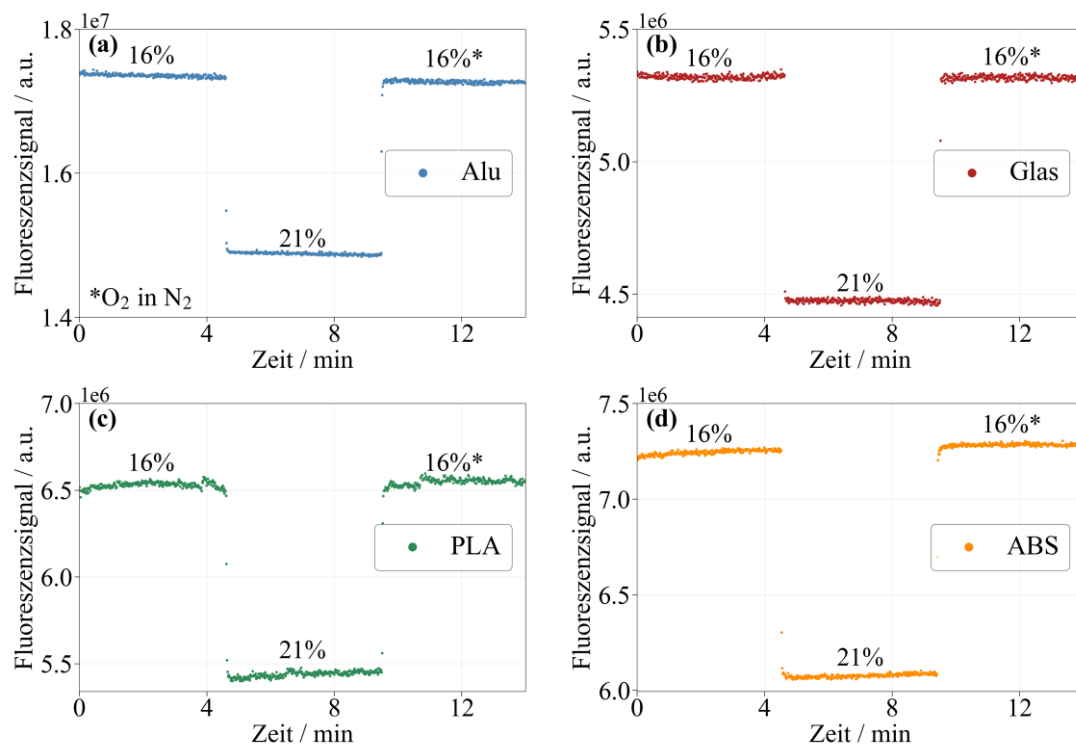


Abb. 6: Fluoreszenzsignal aufgetragen über die Zeit für verschiedene Substratträger: Alu (a), Glas (b), PLA (c) und ABS (d). Bei jeder Messreihe (a)-(d) wurde die Sauerstoffkonzentration bei Raumtemperatur zwischen 16 und 21 Vol. % O₂ in N₂ variiert.

Literatur

- [1] J. E. Hall und A. C. Guyton, *Guyton and Hall textbook of medical physiology*, 13. Aufl. (Student consult). Philadelphia, Pa.: Elsevier, 2016.
- [2] D. C. Poole und A. M. Jones, "Oxygen uptake kinetics," *Comprehensive Physiology*, Jg. 2, Nr. 2, S. 933–996, 2012, doi: 10.1002/cphy.c100072.
- [3] J. J. Pretto, T. Roebuck, L. Beckert und G. Hamilton, "Clinical use of pulse oximetry: official guidelines from the Thoracic Society of Australia and New Zealand," *Respirology (Carlton, Vic.)*, Early Access. doi: 10.1111/resp.12204.
- [4] D. León-Valladares *et al.*, "Determining factors of pulse oximetry accuracy: a literature review," *Revista clinica espanola*, Early Access. doi: 10.1016/j.rceng.2024.04.005.
- [5] H. Prasad *et al.*, "Correlation and agreement between arterial and venous blood gas analysis in patients with hypotension-an emergency department-based cross-sectional study," *International journal of emergency medicine*, Early Access. doi: 10.1186/s12245-023-00486-0.
- [6] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*. Boston, MA: Springer US, 2006.
- [7] V. Balzani und S. Campagna, *Photochemistry and Photophysics of Coordination Compounds I* (280). Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
- [8] C. Higgins, D. Wencel, C. S. Burke, B. D. MacCraith und C. McDonagh, "Novel hybrid optical sensor materials for in-breath O₂ analysis," *The Analyst*, Early Access. doi: 10.1039/B716197B.
- [9] J. D. Pleil, M. Ariel Geer Wallace, M. D. Davis und C. M. Matty, "The physics of human breathing: flow, timing, volume, and pressure parameters for normal, on-demand, and ventilator respiration," *Journal of breath research*, Early Access. doi: 10.1088/1752-7163/ac2589.

Differential Torque Measurement Concept for Condition Monitoring of a Cycloidal Gearbox in a Flight Control Actuator

M. Schmidt¹

¹DLR, Institute of Flight Systems, Lilienthalplatz 7, 38108 Braunschweig, Germany

Abstract

Electromechanical actuators for primary flight control require reliable monitoring of transmission-related loads because the gearbox is a critical subsystem with wear-prone components and potential single-point failures. This paper presents a differential torque measurement concept for a Sumitomo cycloidal gearbox used in a flight control actuator. The approach is based on the simultaneous measurement of torque and rotational speed at the gearbox input and output in order to derive load-dependent loss characteristics and operating signatures of the drivetrain. For this purpose, a compact contactless torque measurement shaft was integrated between the motor and gearbox, while complementary sensor systems were used to acquire torque and position data at both measurement locations. The paper focuses on the measurement architecture, sensor integration, and experimental implementation on an EMA test bench. In addition, the acquired measurement data were combined with model-based fault models and simulation data to perform an initial condition-monitoring-oriented evaluation. Time- and frequency-domain analyses, including FFT and order tracking, were used to assess the sensitivity of the measured quantities to fault-related effects such as increased friction and backlash variation. The results indicate that the proposed measurement concept provides physically meaningful features for distinguishing relevant gearbox fault patterns and forms a promising basis for further condition-monitoring investigations of cycloidal gearboxes.

Keywords: condition monitoring, cycloidal gearbox, differential torque measurement, electromechanical actuator, anomaly detection

1. Introduction

Electromechanical actuators are a key enabling technology for More Electric Aircraft concepts because they can replace conventional hydraulic actuation systems and thereby support weight reduction, improved efficiency, and simplified system integration [1]. In primary flight control applications, however, electromechanical actuators (EMAs) must also satisfy stringent requirements regarding reliability, fault detection, and maintainability, which has increased interest in prognostic and health management methods for safety-critical aerospace systems [2, 11, 13].

Within an EMA, the transmission is a particularly critical subsystem containing several wear-prone mechanical components and strongly influences the mechanical power flow of the actuator [2, 3]. For cycloidal gearboxes, this is especially relevant because their compact, low-backlash design and complex internal kinematics lead to characteristic dynamic load behavior that is difficult to observe using only conventional monitoring approaches alone [4, 5]. Reliable

access to torque- and speed-related quantities at the gearbox input and output is therefore an important prerequisite for understanding drivetrain losses, identifying abnormal operating behavior, and deriving physically interpretable health indicators [8, 9].

Against this background, the present work focuses primarily on the measurement system itself. It presents a differential torque measurement concept for a cycloidal gearbox in a flight control actuator, including sensor selection, mechanical integration, signal acquisition, and implementation on an EMA test bench. Chapter 2 introduces the differential measurement concept, and Chapter 3 describes its implementation on the EMA test bench. Chapter 4 presents the fault modelling and data basis, and Chapter 5 provides an exemplary feature analysis. In this way, the work demonstrates that the proposed measurement concept not only enables detailed observation of drivetrain behavior, but also provides a suitable experimental basis for fault-sensitive monitoring investigations of cycloidal gearboxes [2,3]. Chapters 6 and 7 discuss the

results and conclude with an outlook on future work.

2. Measurement concept

The proposed measurement concept aims to describe the mechanical behavior of the cycloidal gearbox by comparing the drivetrain state at the gearbox input and output. The proposed concept describes gearbox behavior by comparing drivetrain states at the gearbox input and output. For this purpose, torque M , angular position θ , angular velocity ω , and mechanical power are considered at both sides of the gearbox. The measured quantities are transformed to a common reference side in order to derive differential quantities that reflect transmission-internal effects more directly than absolute signals alone [2, 5, 9].

The general measurement principle is illustrated in Fig. 3.1 of the experimental setup. The gearbox is treated as a transmission element with an input-side mechanical state (M_1, θ_1, ω_1) and an output-side mechanical state (M_2, θ_2, ω_2). In order to compare both states consistently, the output quantities are referred to the input side by means of the transmission ratio $i = 29$ for the investigated Sumitomo cycloidal gearbox. The first step is the transformation of the measured output quantities to the input-side reference frame. With this, the referenced output quantities are defined as

$$M_{2,ref} = \frac{M_2}{i}, \theta_{2,ref} = i\theta_2, \omega_{2,ref} = i\omega_2 \quad (1)$$

Based on this reference transformation, differential quantities can now be formulated in a compact and physically interpretable way. A first derived quantity is the torque difference

$$\Delta M = M_1 - M_{2,ref} = M_1 - \frac{M_2}{i} \quad (2)$$

which describes deviations in load transfer across the gearbox. This quantity is less influenced by the absolute actuator load level than a single torque signal and emphasizes changes that originate within the transmission itself.

A second quantity is the relative angular deviation

$$\Delta\theta = \theta_1 - \theta_{2,ref} = \theta_1 - i\theta_2 \quad (3)$$

which represents the differential motion between input and output. This term can include elastic deformation, backlash-related effects, and general transmission irregularities. In this way, the position measurement does not only provide kinematic information, but also contributes to a more direct description of the internal mechanical coupling.

The corresponding speed deviation follows as

$$\Delta\omega = \omega_1 - \omega_{2,ref} = \omega_1 - i\omega_2 \quad (4)$$

and extends the differential comparison into the dynamic domain. It is particularly useful for describing transient effects and oscillatory deviations in the motion transfer.

Based on the same reference-side formulation, a differential power quantity can be introduced as

$$P_{Diff} = M_1\omega_1 - M_{2,ref}\omega_{2,ref} \quad (5)$$

which yields

$$P_{Diff} = M_1\omega_1 - \frac{M_2}{i}(i\omega_2) \quad (6)$$

and thus

$$P_{Diff} = M_1\omega_1 - M_2\omega_2 \quad (7)$$

Although the transmission ratio cancels out algebraically in the final expression, the reference transformation remains an important conceptual part of the differential method because it places all compared quantities in one common mechanical frame. The resulting power difference represents the energetic deviation across the gearbox. In stationary operation it is dominated mainly by internal losses, whereas in dynamic operation it may additionally contain elastic and inertial drivetrain effects [7, 8, 9].

Finally, the differential load and motion quantities can also be combined into an equivalent stiffness-related measure,

$$k_{eq} = \frac{\Delta M}{\Delta\theta} \quad (8)$$

provided that $\Delta\theta \neq 0$. This term is not used here as a full physical gearbox model, but as an additional compact descriptor of the mechanical coupling behavior.

Overall, the differential measurement concept is based on the comparison of mechanically related quantities at both sides of the gearbox. Its main advantage is that it shifts the focus from absolute operating levels to transmission-internal effects between the two measurement points. The resulting quantities therefore provide a physically interpretable basis for subsequent feature generation in time and frequency domain, while the concrete sensor implementation is addressed in the following chapter [4].

3. Experimental setup

This chapter describes the practical realization of the differential measurement architecture for the evaluation of the Sumitomo cycloidal gearbox drivetrain. It focuses on the sensor integration at gearbox input and output such that the mechanical quantities M_1 , θ_1 , M_2 , and θ_2 are simultaneously available for evaluation.

The investigated system consists of an electric motor, the cycloidal gearbox, and an output stage connected to an external load machine via a rigid lever-based kinematic chain. In this setup, the load machine is used to apply reproducible output loads and operating conditions, while the main contribution lies in the sensor integration inside and directly around the actuator drivetrain.

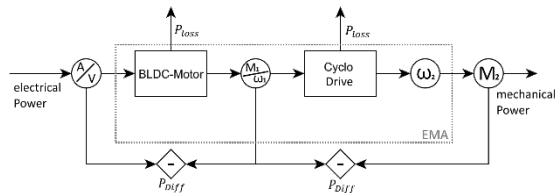


Fig. 3.1: Schematic overview of the measurement concept and sensor positions

The measurement architecture is based on the concept introduced in the previous chapter. On the input side, the transmitted torque M_1 is measured by a compact contactless torque measurement shaft integrated between motor and gearbox, while the input angle θ_1 is provided by a resolver. On the output side, the load torque M_2 is measured by an external torque flange installed between actuator output and load machine, and the output angle θ_2 is acquired by an encoder. Angular velocities ω_1 and ω_2 are not measured directly but derived from the respective position signals. Together, these four measured quantities provide the complete experimental basis for computing the differential variables $\Delta M, \Delta \theta, \Delta \omega$ and P_{diff} as defined in Chapter 2.

The core element of the setup is the input-side torque measurement shaft from MANNER, a compact custom design combining a strain-gauge sensing element with an integrated contactless telemetry system. The shaft has a torque range of 0.1 kNm and a measurement bandwidth of up to 1 kHz . The output-side torque flange covers load levels up to 1 kNm and thus characterizes the full mechanical load at the actuator output. Both sensors are connected to dedicated signal conditioning units for amplification and preprocessing before the signals are passed to the central acquisition system.

The mechanical integration of the torque measurement shaft required modifications to the original actuator. In the original design, the extended motor shaft is connected to the gearbox input via a spline connection. The same spline geometry was transferred to the custom measurement shaft, which was inserted between motor shaft and gearbox input. The required additional axial space was created by

an intermediate flange between motor and gearbox housing. This flange also carries the stationary pickup for contactless signal transmission and provides cable routing to the outside of the actuator. To minimize the axial extension, the measurement shaft was realized as a compact custom design with integrated antenna and minimal shaft length.

All signals were acquired by a dSPACE system, which also handles test bench control and regulation. Torque signals were recorded at 1 kHz , providing a common temporal basis for the differential evaluation. Using a single acquisition platform ensures consistent time alignment between input- and output-side measurements, which is essential for physically meaningful differential quantities.

The experiments were conducted under dynamic conditions derived from representative primary rotor blade control trajectories of a helicopter flight-control application. The reference trajectories are approximately sinusoidal with two higher-harmonic components, covering excitation frequencies between roughly 1 Hz and 6 Hz and output amplitudes up to 10° . These conditions excite relevant drivetrain dynamics while remaining sufficiently structured for systematic signal analysis and comparison in the time and frequency domain [2, 3].

4. Fault modelling and data basis

To provide a basis for the condition-monitoring-oriented evaluation, the failure modes relevant to the investigated Sumitomo cycloidal gearbox were identified and modelled analytically. Rather than deriving features directly from experimental fault states, the present approach uses a physics-based actuator model to generate fault-affected signal data and compares these data with nominal measurements from the EMA test bench.

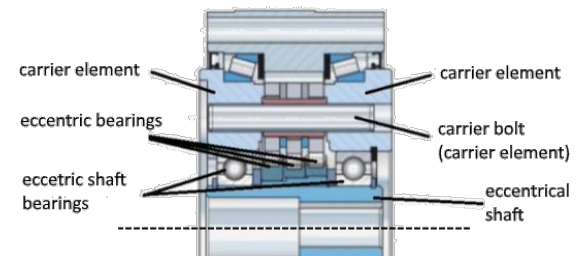


Fig. 4.1: Relevant components of the investigated cycloidal gearbox.

Figure 4.1 shows the main gearbox components relevant to the subsequent fault analysis, including the eccentric shaft, the associated bearings, the carrier elements, and the carrier bolt. This component-level view

provides the structural basis for identifying the considered fault mechanisms.

The failure modes considered in this work are based on a systematic analysis of critical damage mechanisms in EMA transmissions [3, 13]. The cycloidal gearbox contains several wear-prone components whose degradation can produce characteristic changes in the mechanical quantities accessible through the differential measurement concept.

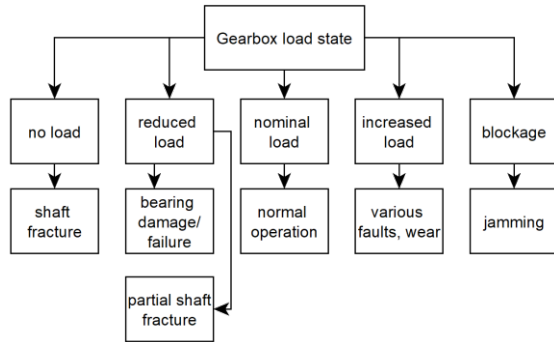


Fig. 4.2: Classification of gearbox load states and exemplary associated failure cases.

As illustrated in Fig. 4.2, the spectrum of possible load-path deviations ranges from a fully open load path to complete blocking of the drivetrain through bearing jamming. Between these extremes, progressive damage mechanisms produce more gradual changes in transmission behavior that are of particular interest for condition monitoring.

Tab. 4.1: Overview of considered failure modes, including component, fault class, and description

Component	Failure	Description
Rolling bearing / tribological contact	Increased friction	Additional resistive torque caused by wear, degradation, or increased contact friction
Transmission kinematic coupling	Increased backlash	Enlargement of the angular dead zone in the drivetrain
Carrier bolt	Carrier bolt wear or deformation	Local wear or deformation at carrier bolt in contact with the cycloidal discs

The three fault classes addressed in this work are summarized in Tab. 4.1. Particular focus is placed on carrier bolt wear or deformation, since the carrier bolt is a primary wear location in this gearbox type [1]. This failure mode is expected to manifest as increased internal friction and increased backlash, both of which produce observable changes in the differential

quantities ΔM , $\Delta\theta$, and P_{Diff} defined in Chapter 2. Increased friction more broadly captures degradation of rolling bearings and lubrication-related effects, while backlash changes mainly affect the position-domain differential quantity $\Delta\theta$ and the equivalent stiffness descriptor k_{eq} introduced in Eq. (8).

The fault modelling is based on an existing physics-based actuator model extended with parametric fault representations. The model describes the mechanical state of the drivetrain and captures the relevant input–output quantities M_1 , θ_1 , M_2 , and θ_2 as functions of the applied load trajectory and internal model parameters. Increased friction is introduced into the model as an additional resistive torque term acting on the input side of the gearbox. The friction torque is described by a combined Coulomb and viscous model

$$M_f(\omega_1) = \text{sign}(\omega_1) (k_v |\omega_1| + M_c) \quad (9)$$

where k_v denotes the viscous friction coefficient and M_c the Coulomb friction offset. Under nominal conditions, both parameters are set to zero. For the fault case FM 211, the Coulomb offset is increased to a small positive value representing the additional resistive torque caused by carrier bolt wear. This additional torque directly increases M_1 at the input while leaving M_2 largely unchanged for the same output load, thereby increasing the torque difference ΔM and the differential power P_{Diff} . Backlash is introduced as a dead zone element in the kinematic coupling between input and output, as illustrated in Fig. 4.3. Under the fault condition, this dead zone is widened parametrically, which produces a loss of mechanical coupling in the $\Delta\theta$ signal during direction reversals and is directly visible in the equivalent stiffness measure k_{eq} .

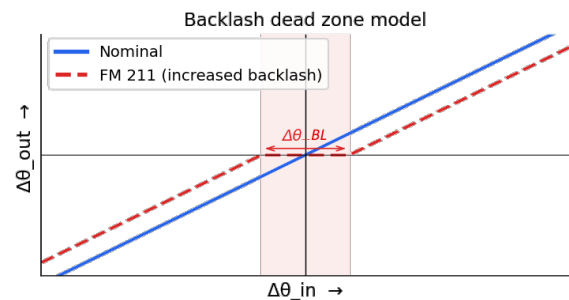


Fig. 4.3: Qualitative backlash dead zone characteristic of the gearbox model

The simulation data generated with the fault model cover the same dynamic operating conditions used during experimental measurements on the EMA test bench, as described in Chapter 3. The trajectories are approximately sinusoidal with higher-harmonic content at excitation frequencies between

roughly 1 Hz and 6 Hz, consistent across simulation and experiment. This allows direct comparison of differential signal characteristics between the nominal simulation, the fault-injected simulation, and the measured test bench data.

The available experimental data comprise a set of representative measurements under nominal operating conditions at different trajectory amplitudes and load levels. These data are intended to serve as a reference baseline for future validation of the model behavior against measured drivetrain signatures. The exemplary evaluation in Chapter 5 is entirely based on simulated signals, comparing nominal simulation output against fault-injected simulation data for the selected fault case, with the purpose of illustrating the proposed condition monitoring approach [2,3].

5. Exemplary feature analysis

Based on the differential quantities introduced in Chapter 2 and the fault modelling described in Chapter 4, an exemplary feature analysis was performed to illustrate how fault-related effects can be extracted from the simulated signals [2,3,6]. The purpose of this chapter is to demonstrate a compact and practical evaluation workflow that links the differential measurement concept to simple fault-sensitive signal characteristics.

As a representative example, the fault case carrier bolt wear or deformation is considered in the following. This fault is particularly suitable because it combines two effects that are relevant for the proposed method, namely increased internal friction and increased backlash. Both effects are expected to influence the dynamic load transfer of the gearbox and should therefore become visible in the differential quantities [1,5].

For the present study, the analysis is performed on short, synchronized observation windows of the simulated trajectories. This keeps the procedure transparent and allows a direct comparison between nominal and fault-affected operation [2,3].

A compact first representation is obtained in the time domain by evaluating the differential power derived in Eq. 7. Since this quantity directly describes the energetic deviation across the gearbox, it combines the information contained in torque and speed into a single physically interpretable signal [4]. A visual comparison of the nominal and fault-affected cases already indicates whether the fault leads to a changed loss level or to increased dynamic irregularities.

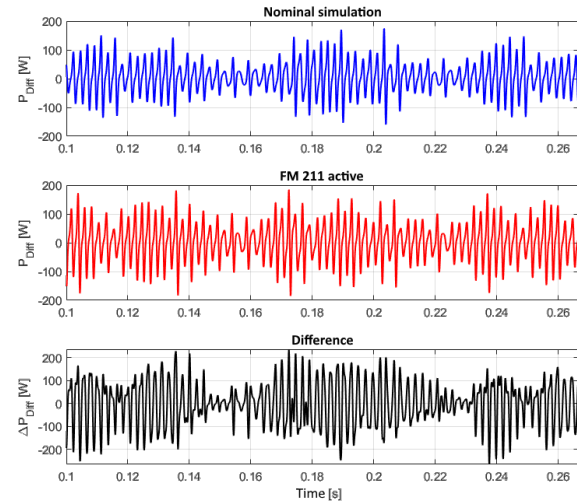


Fig. 5.1: Time-domain comparison of the differential power P_{Diff} for nominal operation and the simulated fault case carrier bolt wear or deformation.

For a simple quantitative description, basic statistical features can be extracted from each observation window. Suitable examples are the mean value, the root-mean-square value, and the peak-to-peak value of P_{Diff} [2,3]. In the considered fault case, an increase in the mean value can indicate additional dissipative losses, while an increase in the peak-to-peak value can point to stronger dynamic variations caused by the altered transmission behavior.

To complement the time-domain view, the same windows can be transformed into the frequency domain by applying a fast Fourier transform to P_{Diff} [2,3,5].

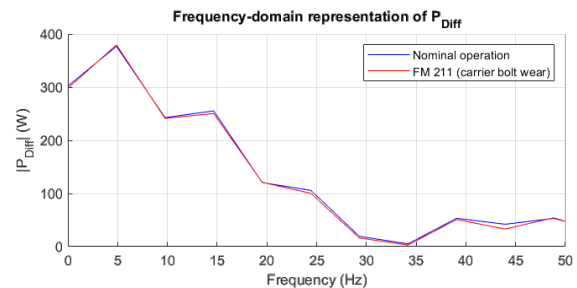


Fig. 5.2: Frequency-domain representation of the differential power P_{Diff}

As shown in Fig. 5.2, the FFT-based representation reveals that the overall spectral shape of nominal operation and the fault-affected case remain largely similar, with no clearly dominant new frequency components appearing under the fault condition. This indicates that the fault-related effect of carrier bolt wear or deformation does not produce a strong isolated spectral signature in the fixed-frequency domain under the investigated operating conditions.

A more sensitive representation is obtained by means of order tracking [2,3,10]. In contrast to the FFT, order tracking relates the spectral content to the instantaneous rotational speed and thereby compensates for speed variations in the trajectory. As shown in Fig. 5.3, this representation reveals a clearly elevated spectral content across multiple orders in the fault-affected case compared to nominal operation, which is not visible in the fixed-frequency FFT. This demonstrates that the fault-related changes in P_{Diff} are more effectively captured in the rotational-reference domain, and that order tracking provides a more discriminative basis for condition-monitoring-oriented feature extraction under dynamic operating conditions [3, 5, 10].

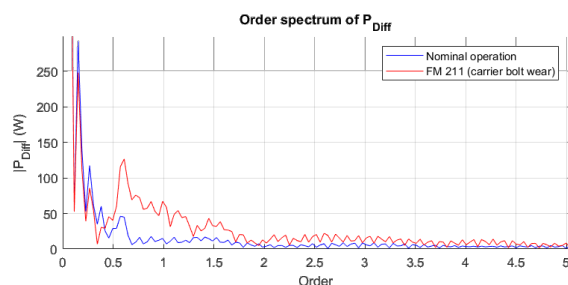


Fig. 5.3: Order spectrum of the differential power P_{Diff}

The three representations shown here, namely time signal, FFT, and order spectrum, provide an initial basis for illustrating the proposed workflow. They show how the differential measurement concept can be transferred into simple and physically motivated features that are suitable for distinguishing fault-related behavior [2,3,6]. A neural-network-based classification could be used in a subsequent step to extend the present feature analysis and potentially improve the separation of fault-related operating states based on the differential signal representations [2,3,6].

6. Discussion of the Results

The results presented in the previous chapters show that the proposed measurement concept provides meaningful access to gearbox-related behavior that cannot be obtained from a single measurement location alone. By combining torque and position information at gearbox input and output, the drivetrain can be evaluated in terms of differential load transfer, relative motion, and energetic deviation across the transmission. This makes internal gearbox effects more directly observable and provides a physically interpretable basis for feature extraction and condition-monitoring-oriented assessment [2, 3, 5].

The exemplary analysis of the simulated fault case carrier bolt wear or deformation indicates that the derived quantities are sensitive to fault-related changes in the gearbox. In particular, the differential power signal provides a compact representation of dissipative losses and dynamic irregularities. Under the investigated conditions, order tracking proved more sensitive than fixed-frequency FFT analysis.

Besides differential power, the other derived quantities also contribute to the interpretation of drivetrain behavior. Differential torque emphasizes deviations in load transfer across the gearbox and is therefore closely linked to changes in internal friction and transmission losses. Differential angle captures relative motion effects between input and output and is therefore particularly relevant for backlash-related behavior, elastic deformation, and local irregularities in the mechanical coupling. In summary, these quantities help to distinguish loss-related effects from backlash-related motion deviations and thereby support the interpretation of gearbox behavior.

A central result of the work is the successful experimental realization of the measurement architecture on the EMA test bench. The integration of the compact contactless torque measurement shaft into the restricted installation space of the actuator, together with synchronized acquisition of input- and output-side quantities, enables the computation of differential variables directly from measured drivetrain states. Compared with approaches that rely on isolated local measurements or purely external load information, this improves the observability of transmission-related effects and extends the experimental basis for gearbox monitoring [4].

The results also underline the benefit of comparing both mechanical interfaces of the gearbox rather than considering absolute quantities at only one side. In the differential representation, the dominance of the overall operating level is reduced and the focus is shifted more strongly toward processes within the gearbox itself. This is particularly useful for cycloidal transmissions, where compact design, low backlash, and complex internal kinematics can make the interpretation of conventional monitoring signals more difficult [2, 3, 5]. Under these conditions, the proposed approach provides a clearer relation between measured quantities and internal effects such as friction increase, altered load transfer, and backlash-related deviations.

The combination of experimental nominal data and model-based fault representation further supports the benefits of the concept. The experimental measurements provide a realistic

reference for drivetrain behavior on the test bench, while the simulation-based fault cases allow expected effects of relevant damage mechanisms to be studied in a controlled manner. In this form, the available data provide a basis for an initial fault-sensitive evaluation of the selected differential quantities and signal representations, while the exemplary feature analysis presented here is based on simulated data [3, 11, 12].

At the same time, some limitations remain in the current state of the study. The fault-related effects discussed here are based on model-based signal variation and the exemplary evaluation of one representative fault case. The present results therefore provide an initial indication of the potential of the proposed differential quantities and signal representations. These considerations provide the basis for the concluding assessment of the proposed measurement concept in the following chapter.

7. Conclusion and outlook

This paper presents a differential torque measurement concept tailored to the condition-monitoring-oriented investigation of a cycloidal gearbox in a flight control actuator. The contribution of the work is not the introduction of differential quantities as a general principle, but their application to this specific drivetrain subsystem, together with the corresponding measurement architecture, sensor integration, and experimental realization on an EMA test bench. In this way, the study provides a practically implemented measurement setup that makes gearbox-related behavior accessible through the comparison of mechanical states at the input and output of the transmission.

The results indicate that the selected differential quantities provide a suitable basis for deriving physically interpretable features from drivetrain measurements. In particular, the differential power quantity P_{Diff} proved to be a compact and informative representation because it combines torque and speed information and reflects energetic deviations across the gearbox. Together with the time- and order-domain representations considered in this work, the presented approach enables an initial fault-sensitive assessment of characteristic effects such as increased friction and backlash-related changes [2, 3, 5]. While a fixed-frequency FFT representation showed limited discriminability for the investigated fault case, order tracking proved effective in revealing fault-related spectral changes under dynamic operating conditions.

A central outcome of the work is the implemented combination of input- and output-side measurement, differential signal representation, and first condition-monitoring-oriented evaluation for the investigated EMA gearbox application. The presented example shows how measurement implementation, physically motivated signal representation, and condition-monitoring-oriented evaluation can be linked at component level. Beyond the investigated gearbox application, the same approach may also be transferred to other drivetrain components, such as the motor or additional transmission stages, provided that suitable measurement access is available.

Future work should focus on a broader validation of the proposed approach using additional measurement data and extended operating conditions. This includes the comparison of model-based and experimentally observed fault signatures, the consideration of further relevant fault cases, and the assessment of robustness under varying loads, trajectories, and boundary conditions. In addition, the differential quantities introduced here can be extended toward further feature generation and more advanced monitoring approaches, including model-based and data-driven methods. Quantities such as differential torque, differential angle, and equivalent stiffness-related measures may provide additional information for this purpose. The limited sensitivity of fixed-frequency spectral analysis further motivates the use of rotation-synchronous signal processing methods for fault-sensitive feature extraction [10]. In this sense, the presented measurement concept provides a suitable experimental basis for future condition-monitoring studies in electromechanical actuator drivetrains [3, 12, 13].

Acknowledgements

This work was carried out within the joint research project smartHUMS within the framework of the German Aviation Research Programme LuFo VI-1 under grant number 20D1918E. The presented paper was prepared in the context of this project

Literature

- [1] G. Qiao, G. Liu, Z. Shi, Y. Wang, S. Ma, and T. C. Lim, "A review of electromechanical actuators for More/All Electric aircraft systems", Proceedings of the Institution of Mechanical Engineers, Part C: *Journal of Mechanical Engineering Science*, vol. 232, no. 22, pp. 4128–4151, 2018, doi: 10.1177/0954406217749869.

- [2] S. Fu and N. P. Avdelidis, "Prognostic and Health Management of Critical Aircraft Systems and Components: An Overview", *Sensors*, vol. 23, no. 19, art. 8124, 2023, doi: 10.3390/s23198124.
- [3] Z. Yin, N. Hu, J. Chen, et al., "A review of fault diagnosis, prognosis and health management for aircraft electromechanical actuators", *IET Electrical Power Applications*, vol. 16, no. 11, pp. 1249–1272, 2022, doi: 10.1049/elp2.12225.
- [4] A. Schütze, N. Helwig, and T. Schneider, "Sensors 4.0 – smart sensors and measurement technology enable Industry 4.0," *Journal of Sensors and Sensor Systems*, vol. 7, pp. 359–371, 2018, doi: 10.5194/jsss-7-359-2018.
- [5] I. Komorska, "Fault Diagnosing of Cycloidal Gear Reducer Using Statistical Features of Wavelet Leader Multifractal Spectra", *Sensors*, vol. 23, no. 3, art. 1645, 2023, doi: 10.3390/s23031645.
- [6] G. Balaban, A. Saxena, P. Bansal, et al., "Prognostic Health-Management System Development for Electromechanical Actuators", *Journal of Aerospace Information Systems*, vol. 12, no. 3, pp. 249–263, 2015, doi: 10.2514/1.1010171.
- [7] K. Michaelis, B.-R. Höhn, and M. Hinterstoißer, "Influence factors on gearbox power loss", *Industrial Lubrication and Tribology*, vol. 63, no. 1, pp. 46–55, 2011, doi: 10.1108/00368791111101830
- [8] C. M. C. G. Fernandes, P. M. T. Marques, R. C. Martins, and J. H. O. Seabra, "Gearbox power loss. Part I: Losses in rolling bearings", *Tribology International*, 2015, doi: 10.1016/j.triboint.2014.10.017
- [9] C. M. C. G. Fernandes, P. M. T. Marques, R. C. Martins, and J. H. O. Seabra, "Gearbox power loss. Part II: Friction losses in gears", *Tribology International*, 2015, doi: 10.1016/j.triboint.2014.10.019
- [10] R. B. Randall, "Analysis of computed order tracking," *Mechanical Systems and Signal Processing*, vol. 11, 1997, doi: 10.1006/mssp.1996.0056
- [11] R. Dixon and A. W. Pike, "Application of condition monitoring to an electromechanical actuator: a parameter estimation-based approach," *Computing & Control Engineering Journal*, 2002, doi: 10.1049/cce:20020203
- [12] M. Mazzoleni, M. Scandella, Y. Maccarana, F. Previdi, G. Pispola, and N. Porzi, "Condition assessment of electro-mechanical actuators for aerospace using relative density-ratio estimation", *IFAC-PapersOnLine*, vol. 51, no. 15, pp. 957–962, 2018, doi: 10.1016/j.ifacol.2018.09.391
- [13] M. Ismail, E. Balaban, and J. Windelberg, "Spall Fault Quantification Method for Flight Control Electromechanical Actuator", *Actuators*, vol. 11, 2022, doi: 10.3390/act11020029, doi: 10.1108/00368791111101830

Joint Simulation–Measurement Framework for Estimation and Compensation of Distributed Parasitics in a Flexible PCB Sensor Matrix

Xiaohu Ma¹, Guozheng Wu², Olfa Kanoun², Alice Fischerauer¹, Gerhard Fischerauer¹

¹University of Bayreuth, Chair of Measurement and Control Systems, Bayreuth, Germany

²Technische University of Chemnitz, Professur Mess- und Sensortechnik, Chemnitz, Germany

Correspondence: mrt@uni-bayreuth.de

Abstract

Distributed sensor networks based on impedance-based pressure sensors can be used for pressure field measurements. However, accurate impedance characterization of such sensor matrices is often compromised by distributed parasitics arising from transmission lines, connectors, and multiplexer switching paths. In this work, a hybrid simulation-measurement approach is presented to compensate for these parasitics in the transmission lines of a nanocomposite pressure sensor matrix fabricated on a flexible printed circuit board (FPCB) with 2×7 nodes for pressure distribution measurement. Impedances between all node pairs were measured from 50 kHz to 1 MHz using an Analog Discovery 2 impedance analyzer. After compensation, the proposed method was validated under multiple loading conditions of the sensor matrix, demonstrating good agreement between the corrected and expected impedance responses. Across the investigated frequency range, the RMS relative error in impedance magnitude across all measurements remained below 8%, with a forward-model RMS relative error of approximately 2.5% and an inverse reconstruction error below 10%, demonstrating the effectiveness of the parasitic compensation framework. This approach enables reliable extraction of intrinsic impedance, providing a foundation for subsequent analysis and applications such as machine-learning-based interpretation of sensor data.

Keywords: Flexible PCB, simulation-measurement framework, compensation of distributed parasitics, impedance network

1 Introduction

Distributed sensor networks based on impedance-based pressure sensors are used in flexible systems for measuring pressure fields, tactile sensing, and human–machine interfaces [1–3]. The accurate characterization of such sensor arrays requires the accurate measurement of the electrical impedance between sensor nodes. However, measurements are often distorted by distributed parasitics arising from transmission lines, connectors, and multiplexers used for sequential node interrogation [4, 5]. These parasitics introduce resistive, inductive, and capacitive coupling that obscures the intrinsic behavior of the actual (= de-embedded) sensor network and can compromise subsequent data-driven analysis or control applications [2, 6].

Various approaches have been proposed to model parasitic effects in electronic interconnects and sensor networks and to remove artifacts due to their presence. Data-driven extraction techniques and curve-fitting me-

thods based on measured responses provide valuable insight into parasitic contributions but often rely on simplified equivalent circuits or specific test structures, which may limit their accuracy and generalizability [4, 7].

Hybrid estimation methods, which combine physics-based simulations with experimental measurements, provide a flexible and practical framework for identifying parasitic elements while preserving the intrinsic response of the network, particularly in impedance networks implemented on flexible substrates where distributed parasitics along densely routed transmission lines are complex and difficult to model accurately [4, 8].

In this work, we present a joint simulation–measurement framework for compensating distributed parasitics in a flexible nanocomposite pressure sensor matrix. The method leverages measured impedance matrices and physics-informed circuit simulation to reconstruct the intrinsic sensor behavior, enabling more accurate sensor characterization and downstream data analysis.

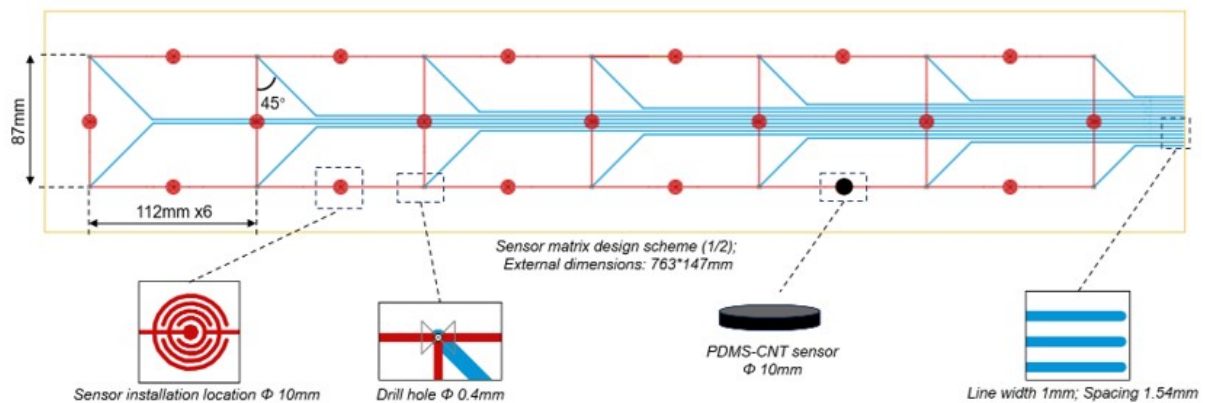


Fig. 1: Layout of the flexible PCB-based pressure sensor matrix.

2 Sensor matrix and measuring setup

Our sensor matrix was developed as part of a versatile rehabilitation and care-assistance system aimed at supporting both macro- and micro-positioning of patients in outpatient care. It comprises 19 sensors arranged in a 2×7 node grid and is fabricated from a composite of polydimethylsiloxane (PDMS) and carbon nanotubes (CNTs), as shown in Fig. 1. The PDMS–CNT pressure sensors operate via a piezoresistive mechanism: conductive CNTs are dispersed within the elastomeric PDMS matrix, which provides mechanical flexibility. When external pressure is applied, the PDMS deforms elastically, reducing the distance between CNTs and increasing both contact points and tunneling pathways, thereby decreasing the overall electrical impedance. Upon release of pressure, the PDMS returns to its original shape, restoring the CNT spacing and electrical resistance. The matrix enables the high-resolution measurement of pressure distributions, which is essential for monitoring and assisting patient positioning, providing real-time feedback for safe and effective rehabilitation and care interventions [9].

The sensor matrix layout is specifically designed for electrical impedance tomography (EIT) to estimate the pressure distribution. Unlike direct sensor readout, EIT exploits the relationships between node impedances to reconstruct a high-resolution pressure map, allowing more “pixels” of information than the number of impedance measurements. This approach reduces the number of direct measurements required, but it also prevents access to individual sensor signals, making the determination of parasitic elements in the transmission lines particularly challenging. To

accurately estimate these parasitics, all possible node-pair combinations must be characterized. For the 14 nodes considered here, this corresponds to $14 \times 13/2 = 91$ impedance measurements. These measurements were conducted over a frequency range from 50 kHz to 1 MHz at 5 discrete frequencies using an Analog Discovery 2 impedance analyzer (AD2), providing representative coverage of the electrical behavior of the sensor network. The AD2 was interfaced with a personal computer via USB, and data acquisition was automated through a custom Python-based measurement script. A broad frequency range is generally desirable to capture the frequency-dependent response of the sensor system. However, the AD2 employs a reference resistor whose value must be comparable to that of the device under test (DUT) to ensure accurate measurements [10].

At lower frequencies, the impedance magnitude of the sensor network increases to the megaohm range, while the available reference resistor is limited to approximately 10 kΩ. This mismatch leads to reduced measurement accuracy and increased uncertainty in the low-frequency regime. Consequently, the usable frequency range is constrained by the measurement system, and frequencies below 50 kHz are excluded from the analysis.

The measurement setup is illustrated in Fig. 2.

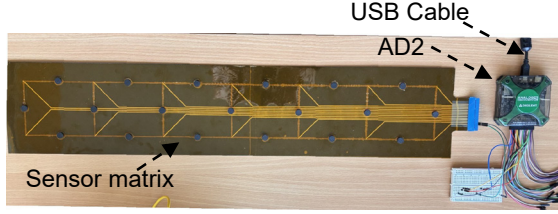


Fig. 2: Impedance measurement setup.

3 Hybrid Estimation of Intrinsic Sensor Impedance

To determine the distributed parasitic impedances in the transmission lines, the intrinsic (ideal) impedances between node pairs are required. Since they are not directly accessible from measurements, they are obtained through numerical simulation based on impedance network theory [11].

As the individual sensor impedances cannot be measured directly due to the multiplexed network configuration, an indirect estimation approach is adopted. Impedances between node pairs are measured under two distinct loading conditions: (i) a no-load condition and (ii) a loaded condition with an applied pressure of 1.25 bar.

Under the no-load condition, impedances are measured for all 91 node pair combinations at 5 discrete frequencies, resulting in a total of 455 complex values, denoted as $Z_{meas}(\omega)$. Under the loaded condition, each sensor is sequentially subjected to a pressure of 1.25 bar, and only the impedance of the node pair directly connected to the activated sensor is measured. This results in 19 impedance measurements per frequency, corresponding to 95 impedance values in total, denoted as $Z_1(\omega)$.

Between these two measurement conditions, only the sensor impedances vary whereas the distributed parasitic impedances associated with the transmission lines and measurement setup remain unchanged. This enables the formulation of an inverse estimation problem, where the intrinsic sensor parameters (modeled as frequency-dependent RC elements) are identified by minimizing the difference between measured and simulated impedance responses across both loading conditions.

Each sensor is modeled as a parallel RC element:

$$Z_i^{(k)}(\omega) = \frac{R_i}{1 + j\omega R_i C_i} \quad (1)$$

where $k=0$ and $k=1$ respectively denote the unloaded and loaded case and i runs from 1 to 19.

Let

$$Z_{simu}^{(0)}(\omega) = f(Z_i^{(0)}(\omega)) \quad (2)$$

denote the simulated pairwise impedance vector (size $M=91$) under no-load condition obtained from the impedance network mode [11]. Then the parasitic admittance is estimated by

$$Y_{CP}(\omega) = \frac{1}{Z_{meas}(\omega)} - \frac{1}{Z_{simu}(\omega)} \quad (3)$$

where $Z_{meas}(\omega)$ is associated with impedance measurements under no-load condition. For each sensor i , a hybrid state is constructed:

$$Z_i^{(k)}(\omega) = \begin{cases} Z_i^{(1)}(\omega), & k = i \\ Z_i^{(0)}(\omega), & k \neq i \end{cases} \quad (4)$$

When one sensor is pressed, the corresponding intrinsic impedances between node pairs can be simulated:

$$Z_{simu}^{(k)}(\omega) = f(Z_i^{(k)}(\omega)). \quad (5)$$

Let $p(i)$ denote the node pair corresponding to sensor i which is pressed. The residual is defined in the admittance domain:

$$r_i(\omega) = \frac{1}{Z_{1,i}(\omega)} - \frac{1}{Z_{simu,p(i)}^k(\omega)} - Y_{CP,p(i)}(\omega). \quad (6)$$

As the impedance magnitude varies by decades between different frequencies the relative residual is used:

$$\tilde{r}_i(\omega) = \frac{r_i(\omega)}{|Z_{1,i}(\omega)|}. \quad (7)$$

The unknown parameters are:

$$\theta = \{R_i^0, C_i^0, R_i^1, C_i^1\}_{i=1}^{19}. \quad (8)$$

The objective is to minimize the total residual over all sensors and frequencies:

$$\min_{\theta} \sum_{\omega, i} (\text{Re}(\tilde{r}_i(\omega))^2 + \text{Im}(\tilde{r}_i(\omega))^2). \quad (9)$$

The dataset used for parameter estimation comprises 455 impedance measurements under no-load conditions and 95 values under loaded conditions, yielding a total of 550 observations across multiple frequencies. These measurements are employed to estimate 76 parameters corresponding to the sensor work, resulting in an overdetermined optimization problem. It is noteworthy that under a load of 1.25 bar, the impedance of the pressure-sensitive elements decreases significantly, becoming approximately one order of magnitude smaller than both the parasitic

impedance and the corresponding no-load sensor impedance [9]. In this regime, the measured impedance is predominantly determined by the sensor response and can therefore serve as an approximation of the intrinsic sensor impedance. In the proposed framework, this approximation is not directly used for parameter identification; rather, it provides physically meaningful initial values for the optimization. This initialization improves convergence and mitigates the risk of non-physical local minima in the subsequent nonlinear estimation process.

The reconstruction accuracy of the proposed model is evaluated by comparing measured and reconstructed pairwise impedances across all frequencies. As shown in Fig. 3, the relative errors in impedance magnitude remain below 10% and phase error below 10 degrees over the entire frequency range. These results demonstrate that the proposed hybrid estimation framework provides an accurate reconstruction of the impedance magnitude and a reasonable approximation of the phase response.

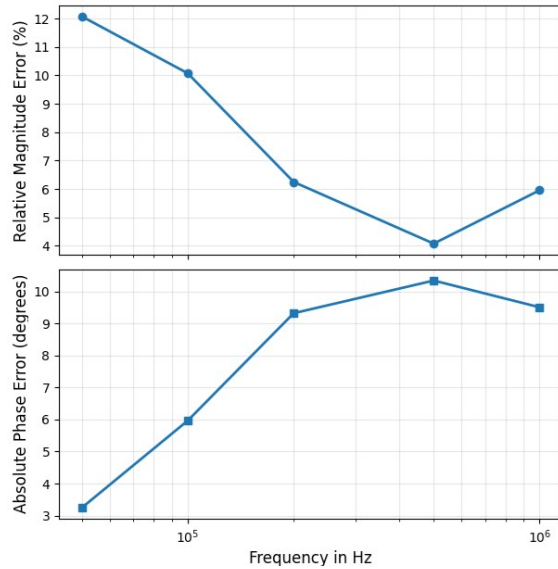


Fig. 3: Relative impedance magnitude error (top) and phase error (bottom) across all frequencies.

4 Extraction and Validation of Distributed Parasitic Admittances

With the pairwise impedances obtained from simulation, the corresponding parasitic admittance $Y_{CP}(\omega)$ can be calculated by Eqn. (3). In general, distributed parasitics in transmission lines may include resistive, inductive, and capacitive components. However, in the present setup, the lines are shorter than 1 m, and the operating frequency range is limited to 1 MHz. Under these conditions, the inductive and resistive contributions are relatively small

compared to the capacitive coupling effects and can therefore be neglected in a first-order approximation.

Consequently, the distributed parasitics are modeled using a leaky capacitive representation that accounts for both capacitive coupling and conductive leakage effects. Specifically, each parasitic path is described by a shunt combination of a coupling capacitance $C_p(k)$ and a leakage conductance $G_p(k)$, resulting in a frequency-dependent admittance of the form

$$Y_{CP}^{(k)}(\omega) = G_p^{(k)} + j\omega C_p^{(k)}, \quad k = 1, \dots, 19. \quad (10)$$

Since each pairwise impedance measurement corresponds to an independent signal path within the multiplexed network, the associated parasitic effects are assumed to be independent for each node pair. Accordingly, the distributed parasitic model is parameterized by 91 independent pairs of parameters ($C_p(k)$, $G_p(k)$), each corresponding to one node pair.

The estimation of these parameters is formulated as a nonlinear least-squares optimization problem in the complex domain. For all frequencies ω and node pairs k , the model parameters are obtained by minimizing the normalized difference between the modeled parasitic admittance and the measured admittance:

$$\min_{C_p, G_p} \sum_{\omega, k} \left\| \frac{Y_{model}^{(k)}(\omega) - Y_{meas}^{(k)}(\omega)}{Y_{meas}^{(k)}(\omega)} \right\|^2. \quad (11)$$

Here, $Y_{meas}(\omega)$ is obtained from the measured impedance by subtracting the simulated intrinsic network response, and $Y_{model}(\omega)$ is given by the parametric leaky capacitive model. The normalization by $Y_{meas}(\omega)$ ensures balanced weighting across frequencies and measurement channels.

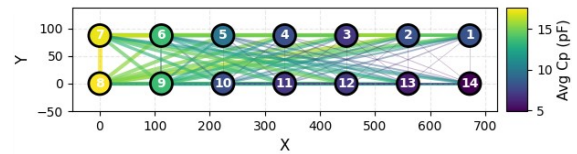


Fig. 4: Physical layout of coupling capacitance in the transmission lines (line thickness & color = coupling strength)

The extracted coupling capacitances $C_p(k)$ of the transmission lines are mapped onto the sensor matrix, as illustrated in Fig. 4. It can be observed that $C_p(k)$ strongly depends on the length of the transmission lines. In particular, the connections associated with nodes 7 and 8, which define the longest transmission paths, show the highest coupling capacitance values.

This trend is consistent with the physical expectation that longer transmission lines introduce larger parasitic capacitances. Therefore, the extracted results not only provide a good numerical fit but also demonstrate clear physical consistency with the underlying system geometry. In contrast, the extracted conductance $G_p(k)$ does not show a clear correlation with the transmission line length. As a result, it is not included in the spatial visualization.

With the extracted coupling capacitances $C_p(k)$, a measured pairwise impedance can be accurately reconstructed from the simulated intrinsic impedance. Conversely, the intrinsic impedance can be recovered from the parasitic-affected measurements through the proposed correction method. Over the entire frequency range, the RMS relative error in impedance magnitude across all measurements was 2.5% for the forward reconstruction of $Z_{meas}(\omega)$ and 7.7% for the inverse reconstruction of the intrinsic impedance $Z_{simu}(\omega)$.

To further validate the estimated parasitic parameters, an additional load condition was introduced in which two sensors were simultaneously pressed. Under this condition, all pairwise impedances were measured and reconstructed across the full frequency range. The reconstructed results strongly agree with the measurements, confirming the validity and robustness of the proposed model. Results at 100 kHz are presented in Figs. 5–7.

5 Conclusion

The proposed inversion, or de-embedding, framework provides a robust and generalizable method for the extraction of distributed parasitics in flexible PCB sensor matrices. By suppressing the parasitic influence, the framework enables the recovery of the intrinsic impedance response that can be directly utilized for subsequent analysis and machine-learning-based modeling of distributed impedance sensors and flexible electronic systems.

The normalization strategy in the admittance domain ensures balanced weighting across frequencies and measurement channels, leading to stable and physically meaningful parameter estimation.

The achieved reconstruction accuracy demonstrates the effectiveness of the method, with RMS relative error of 2.5% in impedance magnitude across all measurements for the forward reconstruction and 7.7% for the inverse reconstruction of the intrinsic impedance.

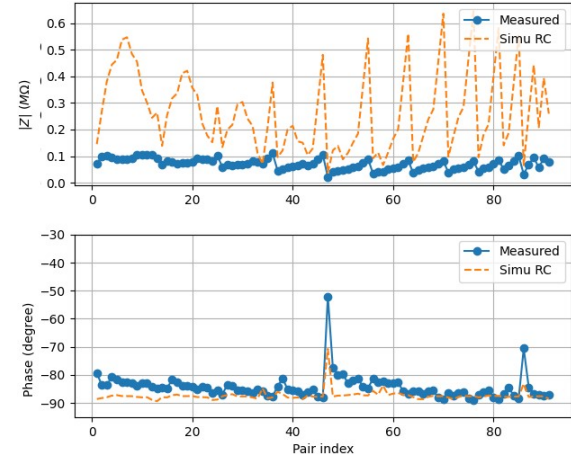


Fig. 5: Measured and simulated impedance magnitudes (top) and impedance phases (bottom) between selected nodes in the distributed-sensor matrix at 100 kHz without taking distributed parasitics into account.

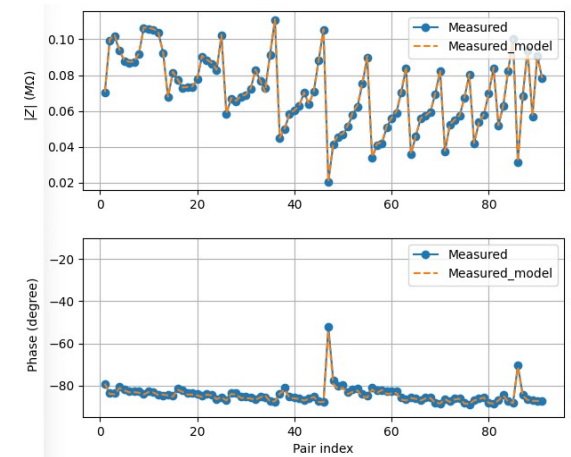


Fig. 6: Comparison of measured and modeled impedance magnitudes (top) and phases (bottom) between all node pairs in the distributed sensor matrix at 100 kHz.

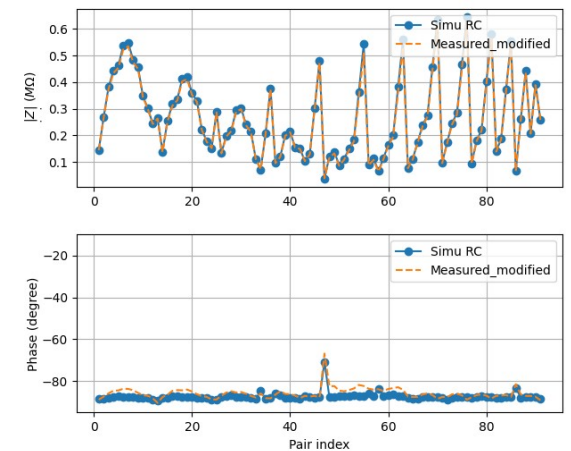


Fig. 7: Comparison of de-embedded measured impedances (= measured impedances with parasitic influences removed) with simulated impedances. Top: impedance magnitudes. Bottom: phases.

Furthermore, validation under a new load condition confirms the generalizability of the estimated parameters and highlights the effectiveness of the parasitic compensation in recovering the intrinsic impedance response.

Overall, the proposed de-embedding framework provides a robust and scalable solution for modeling distributed parasitics in flexible electronic systems. By enabling accurate extraction of intrinsic impedance characteristics, it establishes a reliable foundation for subsequent analysis and machine-learning-based signal processing for distributed impedance sensors.

Acknowledgment

This work was partially supported by the ZIM Program (Grant KK5155006LU3) by the German Federal Ministry for Economic Affairs and Energy.

Declaration of Generative AI Use

While preparing this work, the authors used ChatGPT to check for possible spelling and grammatical errors. In addition, prompts such as “Improve it” were used to ensure the quality of the text. After using this service, the authors reviewed and edited the content as needed. They assume full responsibility for the content of the manuscript.

Literature

- [1] R. Ramalingame, Z. Hu, C. Gerlach, D. Rajendran, T. Zubkova, R. Baumann, O. Kanoun, *J. Sens. Sens. Syst.* 8, 1–7 (2019); doi: 10.5194/jsss-8-1-2019
- [2] N. Carbonaro, M. Laurino, L. Arcarisi, D. Menicucci, A. Gemignani, A. Tognetti, *Appl. Sci.* 11, 2552 (2021); doi: 10.3390/app11062552
- [3] C. Xu, J. Chen, Z. Zhu, M. Liu, R. Lan, X. Chen, W. Tang, Y. Zhang, H. Li, *Small* 20, 2306655 (2024); doi: 10.1002/smll.202306655
- [4] R. A. Khan, M. M. Muhaisin, G. W. Roberts, *IEEE Trans. Circuits Syst. I* 68, 4025–4037 (2021); doi: 10.1109/TCSI.2021.3102103
- [5] G. Bhatti, Y. Agrawal, V. Palaparthi, *Proc. Int. Symp. VLSI Design and Test (VDAT)*, 1–6 (2024); doi: 10.1109/VDAT63601.2024.10705698
- [6] P. Ben Ishai, E. Talary, A. Caduff, E. Levy, Y. Feldman, *Meas. Sci. Technol.* 24, 102001 (2013); doi: 10.1088/0957-0233/24/10/102001
- [7] A. Llamazares, M. García-Gracia, S. Martín-Arroyo, *IEEE Access* 9, 40695–40705 (2021); doi: 10.1109/ACCESS.2021.3064190
- [8] C. Venkataiah, K. S. Prasad, T. J. C. Prasad, *Proc. Int. Conf. Communication and Electronics Systems (ICCES)*, 1–4 (2016); doi: 10.1109/CESYS.2016.7889958
- [9] D. Rajendran, R. Ramalingame, S. Palaniyappan, G. Wagner, O. Kanoun, *Sensors* 21, 6082 (2021); doi: 10.3390/s21186082
- [10] Digilent Inc.: Using the Impedance Analyzer (WaveForms Guide). Available: <https://digilent.com/reference/test-and-measurement/guides/waveforms-impedance-analyzer> (accessed Apr. 14, 2026)
- [11] W. J. Tzeng, F. Y. Wu, *J. Phys. A: Math. Gen.* 39, 8579–8591 (2006); doi: 10.1088/0305-4470/39/27/011

Integration siliziumbasierter Dehnmessstreifen in Keramik zur Realisierung von Nass-Nass Differenzdrucksensoren

*Cathleen Kleinholz¹, Thomas Frank¹, Andrea Cyriax¹, Stefan Jagomast¹, Christian Maier¹,
Michael Hintz¹, Annett Schroeter², Uwe Krieger², Thomas Ortlepp¹*

¹ CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Germany

² VIA electronic GmbH, Robert-Friese-Straße 3, 07629 Hermsdorf, Germany

Zusammenfassung

In diesem Beitrag wird die Herstellung von vollständig integrierten Silizium-Dehnmessstreifen (Si-DMS) in ein Keramikmaterial zur Erzeugung von Differenzdrucksensoren für Anwendungen mit Medienkontakt vorgestellt. Die technologischen Herausforderungen im Herstellungsprozess werden aufgezeigt, Fertigungsprozesse etabliert und das Design optimiert. Ziel dieser Untersuchung ist die Herstellung eines membranfreien, medienbeständigen Differenzdrucksensors für Nass-Nass-Anwendungen.

Es werden zwei verschiedene Keramikmaterialien zur Einbettung von Si-DMS untersucht: Micromax™ GreenTape™ 951PX (Celanese) und bondbares Keramiktape (*engl. bondable ceramic tape* – kurz: BCT, Fraunhofer IKTS). Die Dehnmessstreifen werden auf die noch ungesinterte Keramik aufgebracht und gemeinsam mit der Keramik bei Temperaturen zwischen 850-900°C gesintert. Die Charakterisierung der eingebetteten Si-DMS erfolgt anhand verschiedener Kriterien: optisch, elektrisch und mechanisch.

Keywords: Biegeplatte, Dehnmessstreifen, Differenzdrucksensor, Keramik, Silizium.

Einleitung

In der Mikrosystemtechnik basieren Kraftmessungen in der Regel auf dehnungsempfindlichen Widerständen, sogenannten Dehnmessstreifen (DMS), die zur elektrischen Messung mechanischer Beanspruchung eingesetzt werden. Die Dehnmessstreifen können aus verschiedenen Materialien hergestellt werden, wie beispielsweise aus Metall, Silizium oder Glasfaser [1, 2]. Dabei beeinflusst das verwendete Material die Empfindlichkeit, Temperaturstabilität, Kosten und das Anwendungsgebiet. Im Gegensatz zu metallischen DMS bieten Silizium-DMS eine bis zu 40-mal höhere Empfindlichkeit, weisen nahezu kein Kriechverhalten auf und ermöglichen eine höhere Messgenauigkeit. Bei Ausübung mechanischer Beanspruchung auf die dotierten Bereiche ändert sich der Leitungsmechanismus, was zu einer Änderung des ohmschen Widerstandes führt. Der elektrische Kontakt wird durch Drahtbonden erreicht, weshalb die Anschlussseite nicht für einen physischen Kontakt mit Medien vorgesehen ist. Aus diesem Grund lassen sich Differenzdrucksensoren mit beidseitigem Medienkontakt (Nass-Nass Anwendungen) mit dieser Methode nicht realisieren. Ein bislang ungelöstes Problem war der Schutz des Si-DMS vor dem Kontakt mit verschiedenen

aggressiven Medien und mechanischen Beschädigungen, sowie der gleichzeitige Erhalt der Empfindlichkeit des Sensors. Kontaktpads und Bonddrähte lassen sich beispielsweise durch Silikon vor Feuchtigkeit schützen, sind jedoch mechanischen Einflüssen ausgesetzt, die zu Beschädigungen führen können.

Um den Si-DMS zuverlässig vor mechanischer Manipulation wie auch vor dem Medium zu schützen, soll dieser in eine medienbeständige Biegeplatte integriert werden. Aus diesem Grund wird eine vollständige Integration von Si-DMS in einer Niedertemperatur-Einbrand-Keramik (*engl. Low Temperature Cofired Ceramic* – kurz: LTCC) vor dem Sinterprozess in Betracht gezogen. Bei keramischen Substraten können die Leitbahnen innerhalb des Substrates (vergraben) realisiert werden, sodass die Kontaktanschlüsse vor mechanischen und chemischen Einflüssen geschützt sind [3]. Somit berührt lediglich die LTCC das Medium und eine medienbeständige Biegeplatte für Nass-Nass-Anwendungen kann erzielt werden. Frühere Untersuchungen haben gezeigt, dass Si-DMS erfolgreich in eine Keramik mithilfe der bekannten LTCC-Mehrlagentechnologie partiell eingebettet werden kann [4].

In diesem Beitrag werden Si-DMS mit verschiedenen kommerziell erhältlichen Keramiken hergestellt und hinsichtlich ihrer

Kompatibilität zum jeweiligen Metallisierungssystem untersucht. Der Fokus dieser Untersuchung liegt auf der vollständigen Integration des Si-DMS in ein keramisches Substrat sowie auf der Herstellung einer zuverlässigen elektrischen Kontaktierung.

Silizium-Dehnmessstreifen

Für die Untersuchung wurde ein Si-DMS mit einer Vollbrücke verwendet – bestehend aus jeweils zwei transversal und zwei longitudinal wirkenden Messwiderständen, die als Wheatstone'sche Messbrücke auf der Biegeplatte angeordnet sind (vereinfachte Darstellung in Abb. 1). Diese Anordnung gewährleistet, dass alle Widerstände gleichermaßen beeinflusst werden und keine Verfälschung des Ergebnisses auftritt.

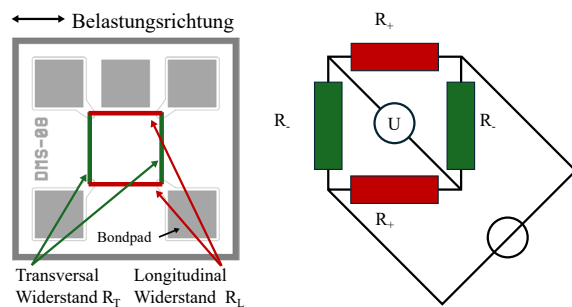


Abb. 1: Si-DMS aus <100> p-Silizium mit vollständiger Messbrücke (je zwei Widerstände für Längs- und Quereffekte), adaptiert und modifiziert von [5].

Integration von Si-DMS in LTCC

Die Herstellung des LTCC-Mehrlagensystems beginnt mit der Präparation der ungesinterten Keramiktapes, auch als „green tapes“ bekannt. Es erfolgt der Zuschnitt auf die geforderte Größe und anschließende thermische Vorbehandlung. Designspezifische Strukturierung der einzelnen Tapelagen wird durch Stanzen der Löcher für Durchkontaktierungen erzielt. Im Anschluss erfolgt das Füllen der Durchkontaktierungen, sogenannter „Vias“, und der Siebdruck der Leiterbahnen. Nach abgeschlossener individueller Tape-Prozessierung erfolgt der Aufbau und die Bestückung des Si-DMS zwischen den ungesinterten Tapes, bevor der LTCC-Stapel laminiert wird. Die Keramik wird druckunterstützt gesintert bei Spitzensinter-temperaturen bei ca. 850°C bis 900°C, entsprechend des jeweiligen Keramiksystems. Eine schematische Darstellung der Fertigungsschritte ist in Abb. 2 gezeigt. [3, 5, 6]

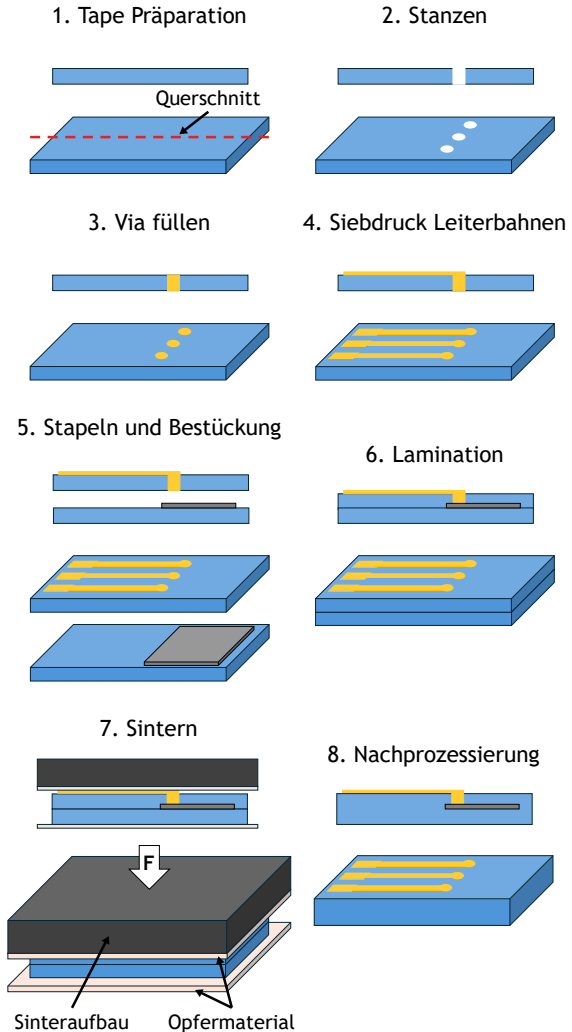


Abb. 2: Prozessablauf zur Herstellung vollständig eingebetteter Si-DMS in Keramik.

Keramikmaterial und Aufbauvarianten

Für die Einbettung des Si-DMS wurden zwei Keramikmaterialien untersucht: Micromax™ GreenTape™ 951PX (Celanese, im Folgenden 951PX; [6]) und ein bondbares Keramiktape (engl. *bondable ceramic tape* – kurz: BCT; Fraunhofer IKTS). Das 951PX ist ein etabliertes LTCC-Material für mehrlagige, cofired Keramiksubstrate, mit optimierten Metallisierungssystemen, verbesserter Stanzbarkeit/Handling und standardisierten Durchkontaktierungsprozessen. BCT ist ein speziell adaptiertes Keramiktape, mit einem an Silizium angepassten Wärmeausdehnungskoeffizienten mit besonderem Fokus auf hybride Integration. Die Fertigungsabläufe sind bei beiden Materialien identisch, Abweichungen finden sich in der Lagenanzahl und vereinzelt in Prozessparametern wie beispielsweise Laminierverfahren, -druck und -temperatur, sowie im Sinterprofil.

Für die konzeptionelle Betrachtung werden zwei Layoutvarianten für vollständig integrierte Si-DMS in Keramik vorgeschlagen, siehe Abb. 3. Zunächst erfolgt die Validierung der siebgedruckten Metallisierung auf der Außenlage (Abb. 3a) und im nächsten Schritt die Verlagerung der Leiterbahnen in eine innenliegende Lage mit Ausnahme der Kontaktierungsflächen am Substratrand (Abb. 3b), um direkten Kontakt mit aggressiven Medien zu vermeiden. Die Anzahl der Tapelagen wurde an die jeweilige Materialdicke angepasst. Für den Probenaufbau wurde eine Gesamtdicke von ca. 300–350 µm nach dem Sintern angestrebt.

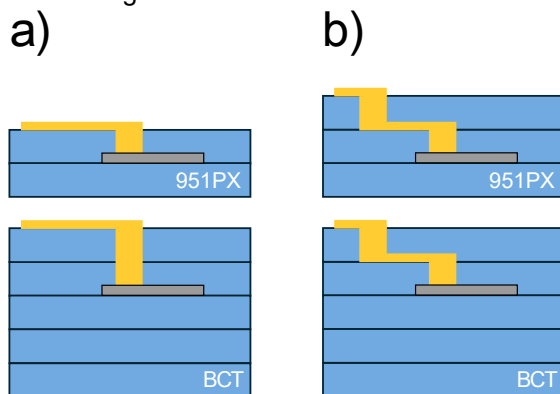


Abb. 3: Schematische Darstellung des Keramikdesigns mit unterschiedlicher Positionierung der Metallisierung und des Si-DMS: a) Leiterbahn auf der Toplage und b) vergrabene Leiterbahn.

Methoden zur Evaluierung

Vor Beginn der Untersuchung wurden essentielle Kriterien für die Bewertung der Ergebnisse definiert:

- Kompatibilität der Metallisierungen von Si-DMS und Keramik
- Messungen der elektrischen Eigenschaften
- Vollständige Integration des Si-DMS in die Keramik.

Zur Bewertung der späteren Sensorzuverlässigkeit werden die hergestellten Substrate zunächst einer elektrischen Kontaktprüfung unterzogen und anschließend einem Temperaturzyklus (Raumtemperatur bis 4 K) ausgesetzt. Oberflächennahe Schäden wie Risse, Delaminationen sowie unvollständige oder beschädigte Leiterbahnen werden mikroskopisch untersucht. Ergänzende Röntgenanalyse dienen der Identifikation innenliegender Defekte wie beispielsweise Metallisierungsunterbrechungen oder Positioniergenauigkeiten. Die Grenzfläche

zwischen Si-DMS und Keramik wird anhand von Querschnitten bewertet.

Elektrische Evaluierung

Die nachfolgend beschriebenen Ergebnisse beziehen sich ausschließlich auf die erste Aufbauvariante (Abb. 3a) für beide Keramiksysteme.

Der Brückenwiderstand der eingebetteten Si-DMS in beiden Keramikmaterialien wurde vor und nach dem thermischen Zyklus gemessen, $R_{951PX} = 4,6 \text{ k}\Omega$ und $R_{BCT} = 4,9 \text{ k}\Omega$; ein nicht eingebetteter Si-DMS liegt bei $R_{Si-DMS} = 5 \text{ k}\Omega$. Die abweichenden Brückenwiderstände lassen sich durch unterschiedliche Materialien und daraus resultierendes Materialverhalten während des Sinterns erklären, da die Schrumpfung einen direkten Einfluss auf die induzierten Materialspannungen und damit auf den Si-DMS Widerstand hat. In Abb. 4 sind die Messverläufe von 951PX und BCT über den Temperaturbereich $RT \rightarrow 4 \text{ K} \rightarrow RT$ dargestellt.

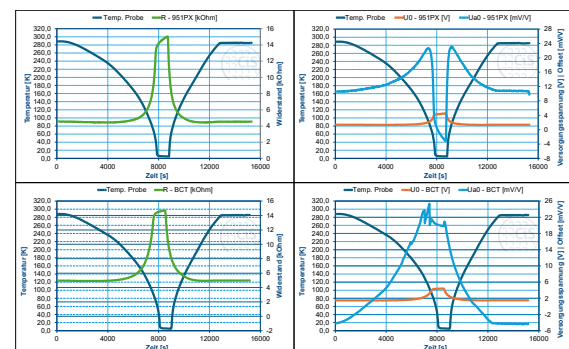


Abb. 4: Temperaturzyklusmessung eines in Keramik eingebetteten Silizium-Dehnmessstreifens, oben: 951PX und unten: BCT. Links ist jeweils die Temperaturabhängigkeit des Widerstandes dargestellt und rechts die Temperaturabhängigkeit der Versorgungsspannung (U_0) und des Offsets/ normierte Brückenausgangsspannung (U_{a0}).

Im Temperaturbereich von Raumtemperatur bis 4 K stiegen die Brückenwiderstände R_{951PX} und R_{BCT} reversibel an und kehrte beim Wiedererwärmen auf den Ausgangswert zurück. Dieses Verhalten ist charakteristisch für Carrier-Freeze-out sowie für temperaturabhängige Änderungen der Kontaktwiderstände und wird auf piezoresistive Effekte und thermisch induzierte Eigenverspannungen im Verbund zurückgeführt [7]. Die relative Änderung des Brückenwiderstands betrug bei 951PX

-0,46%FS, während für BCT eine Änderung von +0,46%FS gemessen wurde. Die Versorgungsspannung U_0 veränderte sich proportional Brückenwiderstand. Für den Offset U_{a0} wurde ein substratabhängiges Verhalten beobachtet, bei 951PX trat ein Vorzeichenwechsel nahe 4 K auf, während bei BCT ein kontinuierlicher Anstieg verzeichnet wurde. Die unterschiedlichen Verläufe von U_{a0} lassen sich durch eine Kombination aus spannungsinduzierten Effekten infolge unterschiedlicher thermischer Ausdehnungskoeffizienten (CTE-Mismatch) sowie variierendem Kontaktverhalten der verwendeten Silberpasten erklären. Basierend auf den durchgeführten Untersuchungen bestätigt sich eine hohe Stabilität der eingebetteten Si-DMS selbst unter extremen thermischen Bedingungen und zeigt gleichzeitig eine Substratabhängigkeit. Der Vorzeichenwechsel ist Gegenstand laufender Untersuchungen. Zusätzliche reproduzierbare Messreihen und Experimente sind notwendig zur abschließenden Klärung.

Optische Evaluation

Die Ergebnisse der optischen Untersuchungen beider Keramiken sind in Abbildung 5 dargestellt.

Bei der mikroskopischen Auswertung der siebgedruckten Leiterbahnen (Abb. 5 a & d) lag der Fokus auf der Vollständigkeit und Unversehrtheit der Leiterbahnen, der Adhäsion zum LTCC-Substrat sowie dem Vorhandensein von Rissen und Delaminationen. Für beide untersuchten LTCC-Keramiken zeigten sich in der mikroskopischen Auswertung keine Auffälligkeiten.

Im Anschluss erfolgten Röntgenaufnahmen (b & e), um Informationen über innere Defekte, Via-Integrität und Positioniergenauigkeit zu erhalten. Hierbei zeigte sich, dass bei beiden Keramiken um die Vias ringförmige Kontrastunterschiede sichtbar sind, was auf ein partielles Ablösen bzw. Wegschrumpfen der Keramik in diesen Bereichen hindeutet. An den untersuchten Substraten ließen sich für beide LTCC-Materialien keine Fehler in der Positionierung des Si-DMS nachweisen.

Zum Abschluss der Untersuchungen wurden Querschnitte angefertigt (c & f). Die Einbettung des Si-DMS in beide Keramiken gelang insgesamt zufriedenstellend, seitlich waren jedoch noch kleine Hohlräume erkennbar. Der Querschnitt des 951PX (c) bestätigte die Vermutung von Hohlräumen um die Durchkontaktierungen (Vias) nicht. Beim BCT zeigte der Querschnitt hingegen deutliche sinterungsbedingte Hohlräume zwischen

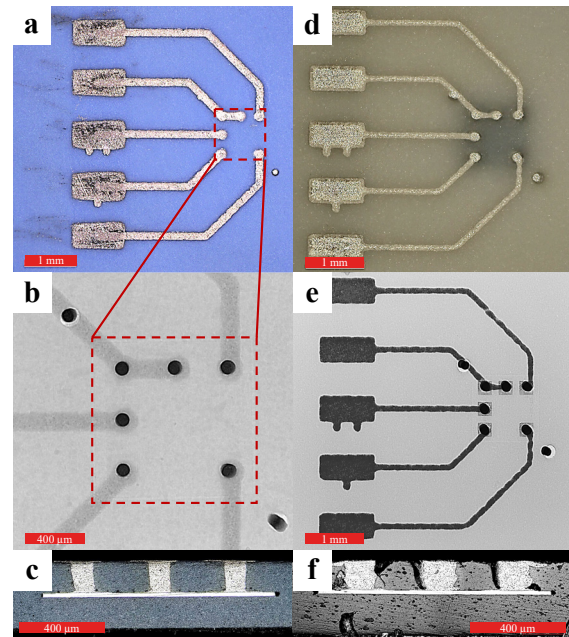


Abb. 5: Vergleich der gesinterten Keramiken 951PX (a - c) und BCT (d - f): mikroskopische Untersuchung der Leiterbahnen (a & d), Röntgenuntersuchung der gefüllten Durchkontaktierungen und der Ausrichtungsgenauigkeit (b & e) und Querschnitt der vollständig eingebetteten Si-DMS in Keramik (c & f). Adaptiert und modifiziert von [5].

Keramik und Via. Diese sind durch das Schwindungsverhalten des Keramiktafes während des Sinterprozesses bedingt. Zu berücksichtigen ist, dass für BCT bislang kein vom Hersteller validiertes Metallisierungssystem verfügbar war. Daher wurde auf ein handelsübliches Metallisierungssystem eingesetzt, dessen Siebdruckpaste bereits für BCT validiert wurde, die für die Via-Metallisierung verwendete Paste bislang nur teilweise validiert war. Die vorliegenden Ergebnisse deuten darauf hin, dass bei der Auswahl bzw. Anpassung des Metallisierungssystems für BCT weiterer Optimierungsbedarf besteht.

Aufgrund der sinterungsbedingten Schrumpfung des Keramiktafes nahe der Vias sowie des vermuteten CTE Mismatches wurden ergänzende Analysen mittels Rasterelektronenmikroskopie (REM) durchgeführt. Die REM-Aufnahmen dokumentierten Mikrorisse in der Keramik um die Vias (siehe Abb. 6, rote Markierungen).

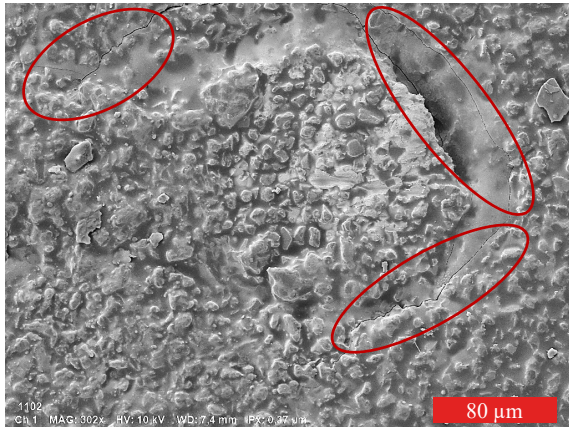


Abb. 6: Detailaufnahme mittels REM der BCT-Keramik in Via-Nähe.

Zusammenfassung und Ausblick

Ein Lösungsansatz für direkten Medienkontakt, sogenannte Nass-Nass-Sensoranwendungen auf Grundlage von Silizium und LTCC wurde untersucht. Silizium-Dehnungsmessstreifen konnten erfolgreich in zwei Keramikmaterialien, integriert, gesintert und ein elektrischer Kontakt nachgewiesen werden.

Die Brückenwiderstände R_{951PX} und R_{BCT} der vollständig eingebetteten Si-DMS kehren nach dem Temperaturwechselzyklus (Raumtemperatur – 4 K – Raumtemperatur) zu ihren ursprünglichen Werten zurück. Somit ist eine dauerhafte Schädigung durch die thermische Untersuchung für beide Keramiken auszuschließen. Beide Keramiken sind auf Grundlage der durchgeführten Experimente kryogen- und hochtemperaturbeständig, wie in dieser Studie gezeigt wurde. BCT erwies sich als vorteilhaft, da es nach dem Sintern semi-transparent erscheint und somit eine vorläufig einfache Überprüfung der Positioniergenauigkeit des Si-DMS ohne zusätzliche Werkzeuge ermöglicht.

Trotz des bei BCT nachgewiesenen sinterungsbedingten Schrumpfungsverhaltens der Keramik um die Vias traten keine elektrischen Ausfälle auf.

Als nächster Arbeitsschritt werden mit beiden Keramik-Materialsystemen die zweite Aufbauvariante mit innenliegenden Leiterbahnen gefertigt und charakterisiert. Anschließend werden an diesen Proben Temperaturwechselzyklen durchgeführt. Parallel dazu wird für das BCT eine gezielte Optimierung der eingesetzten Via-Metallisierung untersucht. Abschließend soll ein erster Prototyp eines membranfreien Differenzdrucksensors aufgebaut werden.

Danksagung

Die Autoren bedanken sich für die Förderung durch das Bundesministerium für Wirtschaft und Klimaschutz/Energie (BMWK/BMW) unter den Förderkennzeichen 49MF230066 (DMS in Keramik) und 20M2216E (H2ELZET). Wir danken B. Müller, M. Fischer und J. Müller (Gruppe Elektroniktechnik, Technische Universität Ilmenau) für die Bereitstellung der Geräte, die Unterstützung während des Herstellungsprozesses mit dem BCT-Material und bei der Durchführung der Röntgenuntersuchung. Wir bedanken uns bei den Kolleginnen und Kollegen A. Böttger für die Querschnitt- und REM-Analysen sowie K. Lauer und S. Hermann für die Unterstützung und Durchführung der Temperaturwechselzyklen.

Literaturnachweis

- [1] A. A. Velichko et al., "The Strain Gauges on the SOI Structures," 2021 XV International Scientific-Technical Conference on Actual Problems Of Electronic Instrument Engineering (APEIE), Novosibirsk, Russian Federation, 2021, pp. 53-57, doi: [10.1109/APEIE52976.2021.9647644](https://doi.org/10.1109/APEIE52976.2021.9647644).
- [2] R. Tamura et al., "High Dynamic Range 6-Axis Force Sensor Employing a Semiconductor–Metallic Foil Strain Gauge Combination," in IEEE Robotics and Automation Letters, vol. 6, no. 4, pp. 6243-6249, 2021, doi: [10.1109/LRA.2021.3093008](https://doi.org/10.1109/LRA.2021.3093008).
- [3] Y. Imanaka, "Multilayered Low Temperature Cofired Ceramics (LTCC) Technology", Springer (2004).
- [4] T. Frank et al., "Siliziumdehnungssensoren zur Spannungsanalyse," MikroSystemTechnik Kongress, Stuttgart-Ludwigsburg, Germany, 2021.
- [5] C. Kleinholt et al., "Fabrication and Evaluation of Fully Embedded Silicon Strain Gauges in Ceramic Material for Wet-Wet Applications," International Conference on Electronics Packaging (ICEP), 2026.
- [6] Celanese Corporation, Technical Datasheet Celanese Micromax™ GreenTape™ 951PX, 2024, available at: <https://www.celanese.com/en/products/micromax>.
- [7] P. Verma, D. Punetha and S.K. Pandey, "Sensitivity Optimization of MEMS Based Piezoresistive Pressure Sensor for Harsh Environment," Silicon [12](#), 2020, 2663–2671.

Temperature-Compensated Hydrogen Sensing with In-Fiber Dual-FBG Fabry-Pérot Structures and Palladium-Based Thin Films

Anja Braunmüller¹, Fabian Buchfellner¹, Markus Becherer², Johannes Roths¹

¹Photonics Lab, Munich University of Applied Sciences, Munich, 80335, Germany

²School of Computation, Information and Technology (CIT), Technical University of Munich, Munich, 80333, Germany

anja.braunmueller@hm.edu (Anja Braunmüller), fbuchfel@hm.edu (Fabian Buchfellner), markus.becherer@tum.de (Markus Becherer), roths@hm.edu (Johannes Roths)

Abstract

Hydrogen sensors based on in-fiber dual-FBG Fabry-Pérot (FP) structures present an advanced sensor platform with the ability to achieve a highly precise detection of hydrogen and temperature. By selectively covering the cavity formed between two UV imprinted Bragg reflectors with a Pd80:Au20 nanofilm provides the sensor with the potential for simultaneous measurements of temperature and hydrogen concentration in a single optical fiber and sensor element. For this purpose, current efforts aim for decoupling the sensor response in regard of temperature and hydrogen-induced strain by applying a fully zero-point referenced sensor calibration scheme. This work addresses the fabrication of such a dual-FBG FP sensor under consideration of suitable design parameters and its subsequent calibration. For the latter, the determination and compensation of the intrinsic temperature response of the sensor will be demonstrated on the example of measurements of H₂ concentrations between 100 and 20,000 ppm at temperatures of 30°C and 50°C. This compensation substantially reduces the cross-sensitivity of the sensor regarding temperature- and hydrogen-induced wavelength shifts. Above all, however, the results obtained with this sensor configuration show hydrogen measurements free of hysteresis with the capability to extend for multi-point sensor probes by wavelength division multiplexing.

Keywords: Hydrogen sensing, optical fiber sensing, temperature compensation, palladium-based metal-hydride, fiber Bragg grating, Fabry-Pérot

Introduction

In recent years, hydrogen (H₂) and hence the research into suitable sensor technologies has significantly gained importance [1-3]. Considering the high flammability of hydrogen in particular [4], it is crucial that sensors are capable of reliably detecting the present H₂ concentration at high precision. Such sensors are required not only for detecting leakages, but for example also the hydrogen concentration in inert gases [5,6]. For industrial applications, however, they must comply with ISO 26142 [7], which specifies the technical requirements summarized in Tab. 1. Fiber-based sensor systems have proven to be particularly promising in this context, since these offer a number of advantages, including the possibility of multiplexing, immunity to electromagnetic interference and their usability in explosive environments due to galvanic insulation. This makes fiber Bragg gratings (FBGs) an especially powerful sensor technology, which is usually applied to enable the detection of parameters such as temperature and strain [8,9]. The use of functionalizing coatings makes

it possible to extend their range of applications for measuring further parameters such as the concentration of gaseous H₂ through Palladium (Pd) based thin films [10-13].

Tab 1: Requirements for H₂ sensors according to ISO 26142 [7]

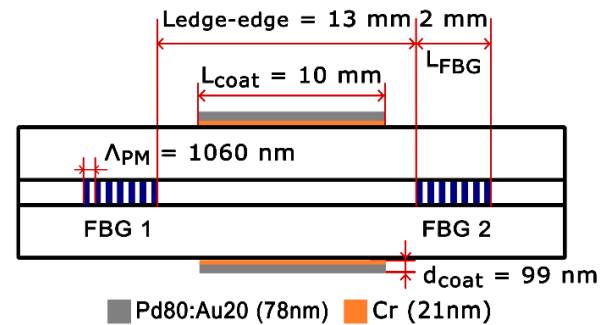
Parameter	Specification
Measurement Range	Up to 4 vol-%
Detection Limit	100 ppm
Accuracy	±50% (≤ 500 ppm) ±25% (≥ 500 ppm)
Repeatability	±10%
Time of Response	≤ 30 s
Time of Recovery	≤ 60 s
Temperature	- 20 °C to 50 °C
Pressure	80 kPa to 108 kPa
Humidity	20 % to 80 %

In 1991, Butler [14] first presented a hydrogen sensor based on a Pd-coated, fiber-optical Mach-Zehnder interferometer. Sutapun et al. [10] later transferred this concept to fiber Bragg gratings by coating a single FBG entirely with a thin layer of pure palladium. Therefore, the mechanical strain arising from the H₂ absorption in the Pd film created a measurable shift of the Bragg wavelength and hence a suitable sensor signal. Ever since, significant progress has been made in optimizing this sensor concept, for example in terms of sensitivity, response time, durability and fabrication feasibility [15].

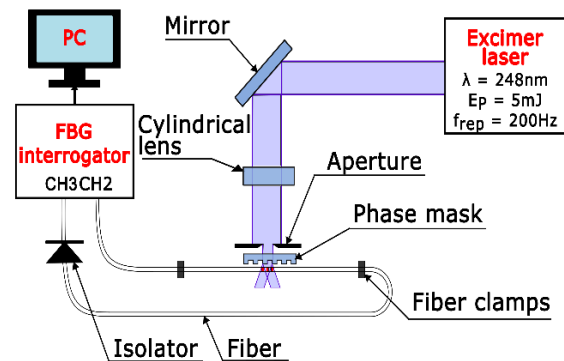
The in-fiber hydrogen sensor presented in this work offers a further improvement by employing an interferometric approach based on a Fabry-Pérot etalon. Applying Pd-coated dual-FBG Fabry-Pérot cavities, allows the suggested sensor platform to provide an increased sensor response and hydrogen measurements at higher resolution, while keeping the spectral selectivity provided by the FBGs. Apart from its potential regarding wavelength multiplexing, the key advantage of this arrangement lies in the possibility of dual parameter sensing [16,17]. An essential step to achieve this is the compensation of the intrinsic temperature sensitivity of the sensor as part of an appropriate sensor calibration, which is addressed in this work.

Sensor Fabrication

The sensor architecture of the fiber FP structures presented in this report is schematically shown in Fig. 1a including the manufacturing parameters. The sensor consists of two FBGs, which are created by photo imprinting a refractive index modulation into the core of a Germanium doped fused silica fiber according to the procedure described by Hill et al. [18]. Therefore, a single mode fiber with a diameter of 80 μm is positioned within the horizontal focal line of a KrF laser (248 nm) created with a cylindrical lens as shown in Fig. 1b. In addition, an adjustable aperture slit for setting the length of the inscribed FBG and a phase mask are placed in the beam path between cylindrical lens and photosensitive fiber. In this case, a phase mask with a pitch of $\Lambda_{\text{PM}} = 1060 \text{ nm}$ was used. The mask is the central component of the FBG fabrication setup. It creates a spatially modulated intensity pattern in the region, where the partial beams of the +1st and -1st diffraction order overlap. This leads to locations of high and low UV radiation, where the corresponding increase of the refractive index (RI) leads to the desired periodic RI modulation in the fiber's core.



(a) Schematic of the dual-FBG Fabry-Pérot based H₂ sensor



(b) Experimental setup for the FBG inscription with UV exposure through a phase mask

Fig 1: Fabrication of a dual-FBG Fabry-Pérot based H₂ sensor

After the inscription of the first FBG, a second grating is created at an edge-to-edge distance of 13 mm by applying identical parameters. Both gratings act as wavelength selective mirrors forming a cavity according to an in-fiber FP etalon. The reflection spectrum of the created etalon measured with an optical 4-channel I4G FBG interrogator (Optics 11, Netherlands) is shown in Fig. 2. The resulting signal consists of a high frequency signal with multiple narrow spikes (yellow) enclosed by a wider envelope (red), whose shape is defined by the line-shape of the two FBGs. Whereas the shift of the cavity-induced high frequency signal is tracked by the shift of the most central minimum, the so-called fringe, the FBG-induced envelope is represented by the vertex of a parabolic fit through the peaks of the FP spikes.

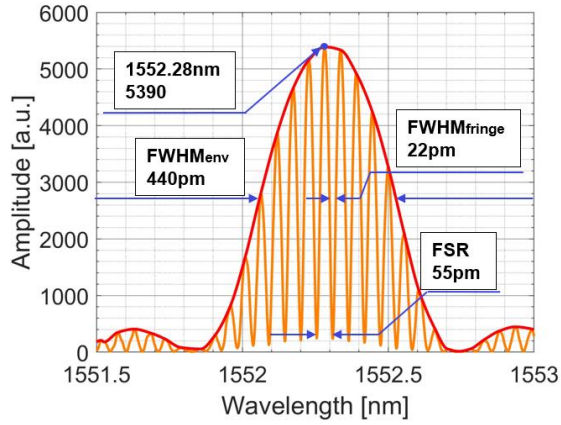


Fig 2: Spectrum of the FBG-induced FP etalon with envelope and comb-like modulation

Parameters such as the fiber diameter, FBG reflectivity and cavity length were intentionally selected to increase the sensor's strain sensitivity and the precision of both fringe and vertex signal. Accordingly, the cavity length must suffice to obtain enough support points for the vertex fit, while a larger FBG length results in a higher reflectivity, a narrower linewidth of the reflection spectrum and a higher precision of the fringe signal due to steeper transitions between maxima and minima. Further details to the fabrication of this sensor can also be found in [12].

In order to use the introduced sensor for the detection of hydrogen the optical cavity of the fabricated FP etalon is modified by partially coating its cavity with a Pd₈₀:Au₂₀ thin film via magnetron sputtering, while the FBGs remain unmodified (see Fig. 1a). Since Pd alone shows a tendency to peel off due to its poor adhesion towards optical fibers [10], the actual coating includes an intermediate chromium (Cr) layer before the gas-sensitive Pd alloy is deposited by accessing the Pd and Au targets simultaneously with different sputter powers. Due to the pronounced hysteresis behavior of Pd at its miscibility gap, gold is added to stabilize the phase transition of PdH_x by expanding the host lattice and increasing the H₂ solubility of the metal alloy [19]. For monitoring purposes, a fragment of a silicon (Si) wafer was coated in the same run as the fiber sensor. This enables the determination of the exact coating composition after the sputter process. The result of the performed energy-dispersive X-ray (EDX) measurement is shown in Tab. 2. From the scanning electron microscopy (SEM) image in Fig. 3, a thickness of approximately 21 nm is deduced for the Cr layer, while the Pd film measures roughly 78 nm.

Tab 2: Elemental composition of the H₂-sensitive coating determined by EDX

Element	Atom-%
Cr	3.05
Pd	78.99
Au	17.96

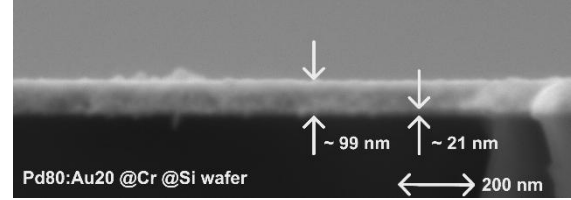


Fig 3: SEM image of the coating deposited via magnetron sputtering on a fraction of an Si wafer

Sensor Principle

Palladium (Pd) is known for its high H₂ solubility. In case of exposure, H₂ dissociates and diffuses into the coating under the formation of the metal hydride PdH_x. The hydrogen load leads to a volumetric expansion of the Pd lattice [20,21] and hence transfers a strain on the fiber along the FP cavity, supported by the reduced fiber diameter of only 80 μm compared to the 125 μm of a standard SM fiber [22]. The provoked strain again results in an increase of the optical path within the cavity and therefore in a measurable wavelength shift of the comb-like interference fringes. While the coated cavity reacts to both - a change in hydrogen concentration and temperature - with a shift in fringe wavelength, the FBGs remain uncoated. This means that a change in the spectral position of the envelope's vertex can be clearly attributed to a change in temperature, enabling compensation for temperature-related wavelength changes within a single spectrum. The compensated wavelength shift $\delta\lambda$ can be calculated via

$$\begin{aligned}
 \delta\lambda(c_{H_2}, \vartheta) &= \Delta\lambda_f - \Delta\lambda_{env} \quad (1) \\
 &= \lambda_f(c_{H_2}, \vartheta) - \lambda_f(c_{H_2} = 0 \text{ ppm}, \vartheta = 0^\circ\text{C}) \\
 &\quad - [\lambda_{env}(\vartheta) - \lambda_{env}(\vartheta = 0^\circ\text{C})] \\
 &= \lambda_f(c_{H_2}, \vartheta) - \lambda_{env}(\vartheta) - \delta\lambda_0
 \end{aligned}$$

with

$$\begin{aligned}
 \delta\lambda_0 &= \lambda_f(c_{H_2} = 0 \text{ ppm}, \vartheta = 0^\circ\text{C}) \\
 &\quad - \lambda_{env}(\vartheta = 0^\circ\text{C}) \quad (2)
 \end{aligned}$$

where $\Delta\lambda_{env}$ denotes the change in vertex wavelength, while $\Delta\lambda_f$ indicates the shift in fringe wavelength and $\delta\lambda_0$ denotes their joint zero reference, which is obtained as part of the subsequent temperature compensation procedure, described in detail in [12].

Temperature compensated H2 measurements

In order to determine the sensor's H2 sensitivity, it was integrated in the setup shown in Fig. 4.

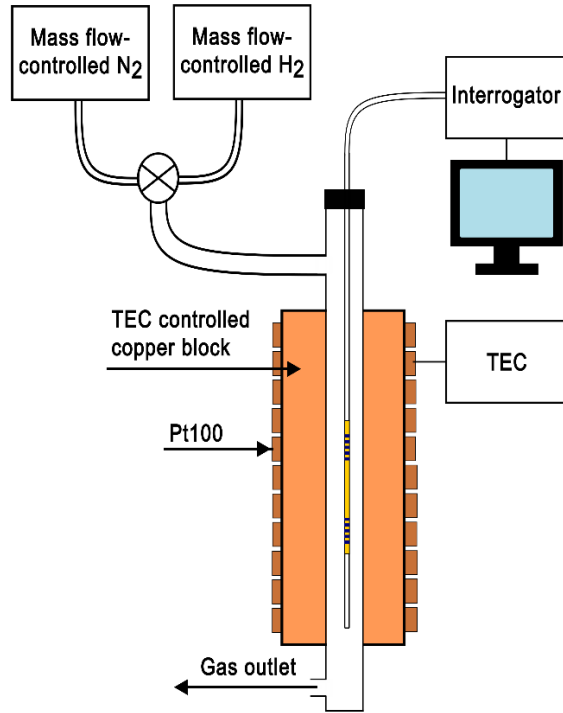


Fig 4: Schematic of the H2 calibration system

The calibration system consists of a glue-sealed tube holding the fiber sensor. This tube is temperature-controlled with an array of 12 Peltier elements and is connected to a mass flow-controlled gas supply making it possible to set a specific concentration of gaseous H2 surrounding the sensor by controlling the ratio between H2 and N2 flow. Within the scope of these measurements, the sensor was exposed to H2 concentrations in the range of 100 ppm to 20,000 ppm at temperatures of 30°C and 50°C. These measurements were performed at cycles of increasing and decreasing H2 concentration to assess the hysteresis behavior of the sensor. Fig. 5 shows the signals arising for both, the envelope vertex (violet) and fringe (yellow). As expected, the fringe signal shows a distinct modulation due to the change in H2 concentration, while the vertex signal remains almost constant during an isothermal hydrogen cycle.

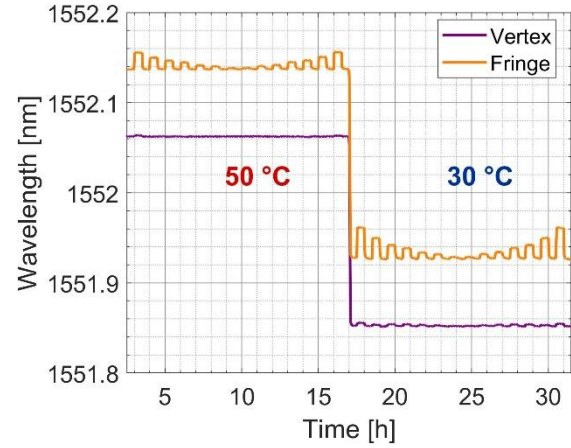


Fig 5: Fringe and vertex signal recorded at 30°C and 50°C for H2 concentrations between 100 and 20,000 ppm

The first step to a successful temperature compensation requires the determination of the reference wavelengths $\lambda_f(cH_2 = 0\text{ppm}, \vartheta = 0^\circ\text{C})$ and $\lambda_{env}(\vartheta = 0^\circ\text{C})$. Since the temperature response of FBG based sensors follows a quadratic function [23], these references emerge from the y-axis intercept of a second order polynomial fitted to the measurement signals obtained for fringe and vertex. The fit function

$$\lambda_{q, fit}(cH_2, \vartheta) = \alpha_1 \vartheta^2 + \alpha_2 \vartheta + \alpha_3 \quad (3)$$

(q = 'f' or 'env') yields the fit parameters denoted in Tab. 3 with α_3 corresponding to the zero-reference of the according signal. Using these parameters, $\delta\lambda_0$ can be calculated.

Tab 3: Fit coefficients resulting for the intrinsic temperature response of fringe and vertex signal of the sensor

λ_i	α_1 [pm/°C ²]	α_2 [pm/°C]	α_3 [nm]
$\lambda_f(cH_2, \vartheta)$	-0.098	15.998	1551.582
$\lambda_{env}(\vartheta)$	0.017	9.150	1551.562

Based on measurements of the vertex $\lambda_{env}(\vartheta)$ and fringe wavelengths $\lambda_f(cH_2, \vartheta)$, the difference results in the temperature-compensated wavelength shift $\delta\lambda$, because fringe and vertex wavelength show equal temperature dependencies. This is evident from the shared baseline for cycles at different temperatures in Fig. 6. The sensor clearly resolves H2 concentrations to a level as low as 100 ppm at both temperatures.

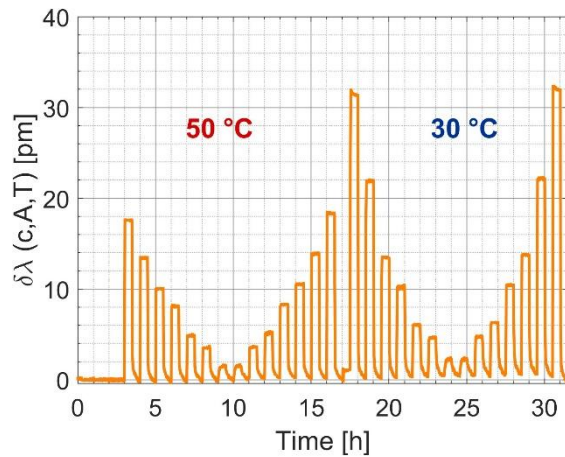


Fig 6: Time-resolved zero-referenced signal $\delta\lambda$ for H_2 concentrations ranging from 100 to 20,000 ppm at 30°C and 50°C

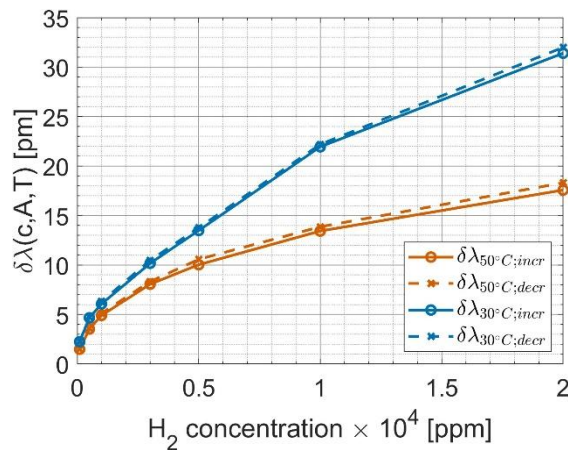


Fig 7: Temperature-compensated sensor response $\delta\lambda$ as a function of the H_2 concentration in the surrounding gas atmosphere in case of increasing and decreasing hydrogen load at 30°C and 50°C

The curves in Fig. 7 show the characteristic hydrogen curves at constant temperatures with a continuous, strongly nonlinear increase of $\delta\lambda(c_{H_2}, \vartheta)$ towards higher H_2 concentrations representing a particularly high sensitivity of the sensor at small concentrations. At the same time, the increasing and decreasing H_2 loadings are consistent, indicating a reversible, hysteresis-free sensor response. Even after the compensation of intrinsic temperature effects, a remaining temperature dependence was observed, being a result of the temperature dependent solubility of H_2 in the Pd thin film. This indicates the need of a temperature-dependent calibration procedure. Although the overall signal is highly nonlinear, Eq. (4) gives an adequate description of the sensor behavior with respect to H_2 concentration and temperature, given by

$$\delta\lambda(c_{H_2}, \vartheta) = \beta_1(\vartheta)\sqrt{c_{H_2}} + \beta_2(\vartheta)c_{H_2} + \beta_3(\vartheta)c_{H_2}^2 \quad (4)$$

The fit parameters β_i corresponding to the characteristic curves for the considered temperatures of 30°C and 50°C shown in Fig. 7 are summarized in Tab. 4.

Tab 4: Fit coefficients resulting for the H_2 dependent behavior of the hydrogen induced wavelength shift $\delta\lambda$

ϑ [°C]	β_1 [pm/ $\sqrt{\text{ppm}}$]	$\beta_2 \times 10^{-4}$ [pm/ppm]	$\beta_3 \times 10^{-9}$ [pm/ppm ²]
30	0.178	3.913	-2.366
50	0.170	3.254	1.876

While a straight forward compensation of the temperature dependent H_2 solubility is not possible, Buchfellner et al. [12] presented an algorithm based on iterative local linearization that allows a full calibration of the sensor under consideration of the sensitivities of fringe and vertex determined during the described compensation process.

Conclusion

The presented dual-FBG Fabry-Pérot sensor offers a highly precise detection of small H_2 concentrations. This sensor shows a reversible sensor response providing one of the highest sensitivities yet reported for unmodified, intact optical fibers (no etching or tapering required). Its suitability for multiplexing makes this sensor concept especially appealing. Additionally, it allows the compensation for internal temperature effects. Future efforts aim to characterize this sensor for broader H_2 and temperature ranges in various background compositions and for multi-point, multi-parameter sensor probes.

Funding

German Research Foundation (DFG, 448330062).

Literature

- [1] J.Egerer, N. Farhang-Damghani, V. Grimm, et al., The industry transformation from fossil fuels to hydrogen will reorganize value chains: Big picture and case studies for Germany, *Applied Energy* 358, 122485 (2024); doi: 10.1016/j.apenergy.2023.122485
- [2] I. Staffell, D. Scamman, A. Abad, P. Balcombe, P. Dodds, P. Ekins, N. Shah, K. Ward, The role of hydrogen and fuel cells in the global energy system, *Energy & Environmental Science* 12, 463-491 (2019); doi: 10.1039/C8EE01157E
- [3] M. Bhuiyan, Z. Siddique, Hydrogen as an alternative fuel: A comprehensive review of

- challenges and opportunities in production, storage, and transportation, *International Journal of Hydrogen Energy* 102, 1026-1044 (2025); doi: 10.1016/j.ijhydene.2025.01.033
- [4] A. Sakib, T. Islam, P. Resnick, A. Habib, S. Chowdhury, Comprehensive Safety Assessment of Hydrogen: From Production to Application in Energy Systems, *International Journal of Energy Research*, 8857513 (2025); doi: 10.1155/er/8857513
- [5] T. Hübert, L. Boon-Brett, G. Black, et al., Hydrogen sensors – A review, *Sensors and Actuators B: Chemical* 157, 329-352 (2011); doi:10.1016/j.snb.2011.04.070
- [6] S. Menon, A. Kumar, S. Mondal, Advancements in Hydrogen Gas Leakage Detection Sensor Technologies and Safety Measures, *Clean Energy* 9, 263–277 (2025); doi: 10.1093/ce/zkae122
- [7] ISO 26142, Hydrogen detection apparatus – Stationary applications, (2010)
- [8] A. Kersey et al., Fiber grating sensors, *Journal of Lightwave Technology* 15, 1442-1463 (1997); doi: 10.1109/50.618377
- [9] A. Alhussein, M. Qaid, T. Agliullin, B. Valeev, O. Morozov, A. Sakhabutdinov, Fiber Bragg Grating Sensors: Design, Applications, and Comparison with Other Sensing Technologies, *Sensors* 25, 2289 (2025); doi: 10.3390/s25072289
- [10] B. Sutapun, M. Tabib-Azar, A. Kazemi, Pd-coated elastooptic fiber optic Bragg grating sensors for multiplexed hydrogen sensing, *Sensors and Actuators B: Chemical* 60, 27-34 (1999); doi:10.1016/S0925-4005(99)00240-3
- [11] M. Alkhabet, S. Girei, Z. Al-Isawi, O. Shareef, A. Farhan, O. Altalebi, A. Khalaf, J. Jaafar, M. Yaacob, Palladium (Pd) coated fiber optic hydrogen sensors: A review, *Materials Science in Semiconductor Processing* 188, 109204 (2025); doi: 10.1016/j.mssp.2024.109204
- [12] F. Buchfellner, A. Röhr, M. Yang et al., Temperature-compensated fiber optic hydrogen sensing with an in-fiber Fabry-Perot etalon and a Pd:Ni-coated cavity. *Optics Express* 32, 46353–46365 (2024); doi: 10.1364/OE.543484
- [13] M. Wienecke, M. Lengert, M. Weidner, R. Ciudin, J. Heeg, T. Frank, P. Kienke, Pd based MEMS Hydrogen Sensors, *IEEE 36th International Conference on Micro Electro Mechanical Systems (MEMS)*, 822-825 (2023); doi: 10.1109/MEMS49605.2023.10052177
- [14] M. Butler, Micromirror optical-fiber hydrogen sensor, *Sensors and Actuators B: Chemical* 22, 155-163 (1994); doi: 10.1016/0925-4005(94)87015-2
- [15] K. Sharma, Y. Singh, R. Selvaraj, Hydrogen Sensors: A Review, *Biomedical Engineering Science and Technology* 2635, 241-256 (2026); doi: 10.1007/978-3-032-14908-4_20
- [16] F. Buchfellner, Q. Bian, W. Hu, X. Hu, M. Yang, A. Koch, J. Roths, Temperature-decoupled hydrogen sensing with pi-shifted fiber bragg gratings and a partial palladium coating, *Optics Letters* 48, 73-76 (2023); doi:10.1364/OL.474083
- [17] I. Yulianti, A. Supa'at, S. Idrus, et al., Design of fiber bragg grating-based fabry-perot sensor for simultaneous measurement of humidity and temperature, *Optik* 124, 3919-3923 (2013); doi: 10.1016/j.ijleo.2012.11.043
- [18] K. Hill, et al., Bragg gratings fabricated in monomode photosensitive optical fiber by UV exposure through a phase mask, *Applied Physics Letters* 62, 1035–1037 (1993); doi: 10.1063/1.108786
- [19] C. Wadell, F. Nugroho, E. Lidström, B. Iandolo, J. Wagner, C. Langhammer, Hysteresis-free nanoplasmonic pd-au alloy hydrogen sensors, *Nano letters* 15, 3563–3570 (2015); doi:10.1021/acs.nanolett.5b01053
- [20] H. Wipf, Solubility and diffusion of hydrogen in pure metals and alloys, *Physica Scripta* 43 (2001); doi: 10.1238/Physica.Topical.094a00043
- [21] H. Peisl, Lattice strains due to hydrogen in metals, *Topics in Applied Physics* 28, 53–74 (1978); doi: 10.1007/3540087052_42
- [22] F. Buchfellner, A. Roehrl, S. Heissel, C. Eulenkamp, A. W. Koch, J. Roths, Optical, thermal and mechanical characterization of standard and thin-clad germanosilicate single-mode fibers for sensitivity-enhanced force sensing with fiber bragg gratings: (under review), *Journal of Light wave Technology* (2025)
- [23] G. Flockhart, R. Maier, J. Barton, W. MacPherson, J. Jones, K. Chisholm, L. Zhang, I. Bennion, I. Read, and P. Foote, Quadratic behavior of fiber Bragg grating temperature coefficients, *Applied Optics* 43, 2744-2751 (2004); doi: 10.1364/AO.43.002744

Qualifizierung eines Gas-Sensors für den Einsatz von VOC-Messungen bei Abluftreinigungen

Annika Braun^{1,2}, Hartmut Sauter², Ralf Moos¹, Gunter Hagen¹

¹Universität Bayreuth, Zentrum für Energietechnik (ZET), Lehrstuhl für Funktionsmaterialien, Universitätsstraße 30, 95447 Bayreuth, Deutschland

²PlasmaAir AG, Am Lindenberg 8, 71263 Weil der Stadt – Hausen

Zusammenfassung

Die verlässliche Messung flüchtiger organischer Verbindungen ist in der industriellen Abgasreinigung unerlässlich. Für Inbetriebnahme und Überwachung von Düsenbodenwäschern ist der Einsatz von Flammenionisationsdetektoren Stand der Technik. Allerdings ist damit ein sehr hoher Aufwand verbunden. Im vorliegenden Beitrag wird gezeigt, dass ein kostengünstiger Hochtemperatur-Gassensor mit einfacher Elektronik vergleichbare Ergebnisse liefert. Weiterentwicklungen könnten auf ein Messgerät zur direkten Umsatzmessung abzielen.

Keywords: Gassensor, industrielle Abgasreinigung, VOC-Messung, Ethanol, Isopropanol.

Motivation und Zielsetzung

Nach der TA Luft liegt der Grenzwert für flüchtige organische Verbindungen (VOC) bei 50 mgC/m^3 [1], dessen Einhaltung regelmäßige Emissionsmessungen erfordert. In Anlagen werden häufig Analytik-Geräte wie Flammenionisationsdetektoren (FID) oder Photoionisationsdetektoren (PID) verwendet. Hohe Anschaffungskosten sowie vor Ort notwendige Peripheriegeräte und Wartungsaufwand sind Gründe für die Erforschung und Entwicklung handlicher und kostengünstiger Alternativen. Im vorliegenden Beitrag wird der Einsatz eines kostengünstigen in keramischer Dickschicht-technik hergestellten Sensors an einer Düsenbodenwäscher-Technikumsanlage untersucht. Es wird die Frage geklärt, ob sich der Sensor zur Detektion der Schadstoffe Ethanol und Isopropanol im Bereich des Roh- und des Reingases eignet, was die kontinuierliche Überwachung von VOCs in industriellen Abluftströmen oder sogar die sensorbasierte Regelung einer Abgasreinigungsanlage ermöglichen könnte.

Methodik und Experimentelles

Der Sensor arbeitet nach dem Exothermieprinzip. Die Temperaturerhöhung an einer katalytisch aktiven Schicht durch Oxidation der Testgase wird mittels einer speziellen Thermoelementstruktur im Vergleich zu einem inerten Bereich auf der Sensoroberfläche gemessen. Details zum Sensoraufbau finden sich in [2]. Vorteil dieses Sensorprinzips ist die Robustheit und zu erwartende Langzeitstabilität mit einer

linearen Kennlinie. Für die Untersuchungen wurde der Sensor unter realitätsnahen Betriebsbedingungen eingesetzt. Ein FID saugt kontinuierlich Gas aus dem Prozess an, welches zuerst am Sensor vorbeigeleitet wird. Durch den direkten Signalvergleich können die Linearität, die Sensitivität und das dynamische Ansprechverhalten des Sensors analysiert, sowie sein Potenzial als kostengünstige und robuste Alternative bewertet werden. Es wurden Positionierungen und Applikationsparameter variiert und deren Einfluss auf das Sensorsignal hinsichtlich Signalgüte und -stabilität untersucht. Als Testgase kamen Lösemittel mit Konzentrationen im Bereich von $2 - 4 \text{ gC/m}^3$ Ethanol und Isopropanol, sowie deren Mischungen (1:1) zum Einsatz, die in den Abluftstrom vor dem Wäscher zudosiert wurden.

Ergebnisse und Diskussion

Der Sensor zeigt im untersuchten Bereich ein dynamisches Ansprechverhalten mit schneller Reaktionszeit bzgl. Konzentrationsänderungen. Abbildung 1 zeigt exemplarischen einen zeitlichen Verlauf einer Rohgasmessung, in der die FID-Daten (in ppm, rot) und das Rohsignal des Sensors (in mV, korrigiert um die Offset-Spannung zu Beginn der Messung, schwarz) gemeinsam dargestellt sind. Die direkte Korrelation der Messdaten (Abbildung 2) zeigt die sehr gute Übereinstimmung von Sensormessung und Analytik und die lineare Abhängigkeit des Sensorsignals von der VOC-Konzentration mit einer Empfindlichkeit von

164 μV je 1000 ppm Ethanol. Ähnliche Ergebnisse konnten bei weiteren Messungen mit Isopropanol als Testgas erzielt werden, wobei die Empfindlichkeit hier auf Grund des 1,5 fachen C-Anteils erwartungsgemäß um ca. den Faktor 1,5 erhöht ist (240 μV je 1000 ppm Isopropanol). Die Messung von Mischungen beider Lösemittel (Verhältnis 1:1) zeigte eine Empfindlichkeit von 197 μV je 1000 ppm. Messungen auf der Reingasseite des Wäschers im Konzentrationsbereich bis 50 ppm Lösemittel konnten ebenfalls erfolgreich durchgeführt werden. Damit wurde nachgewiesen, dass das Sensorkonzept für diesen Anwendungsfall sehr gut geeignet ist.

Literaturnachweis

- [1] Neufassung der Ersten Allgemeinen Verwaltungsvorschrift zum Bundes-Immissionsschutzgesetz (Technische Anleitung zur Reinhaltung der Luft – TA Luft), 2025, https://www.verwaltungsvorschriften-im-internet.de/bsvwbund_18082021_IGI25025005.htm (Zugriff am: 10.09.2025).
- [2] Hagen, G.; Leupold, N.; Wiegärtner, S.; Moos, R.: Sensor Tool for Fast Catalyst Material Characterization, Topics in Catalysis 60 (2017), 3, p. 312-317. <https://doi.org/10.1007/s11244-016-0617-8>.

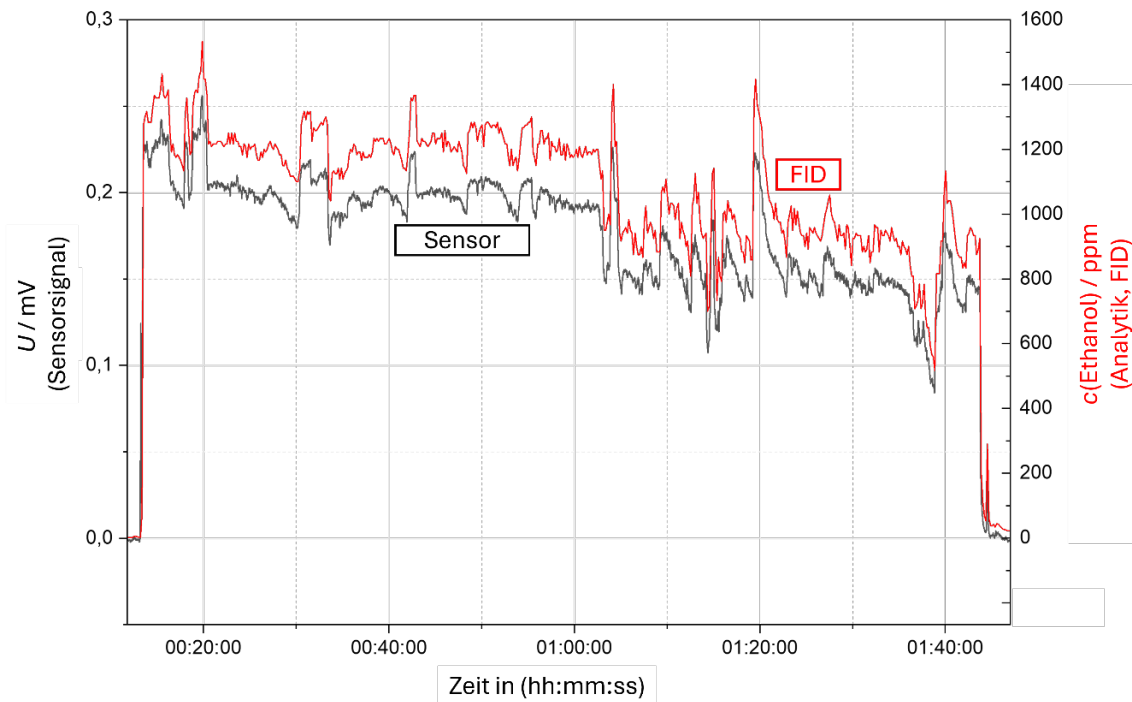


Abb. 1: Zeitlicher Verlauf einer Messung auf der Rohgasseite eines Düsenbodenwäschers mit Ethanol als Testgas-Komponente: Vergleich von Analytik (FID, rote Kurve, Achse rechts) und Sensorsignal (Rohdaten, korrigiert einmal um die Offset-Spannung zu Beginn der Messung, schwarz, Achse links)

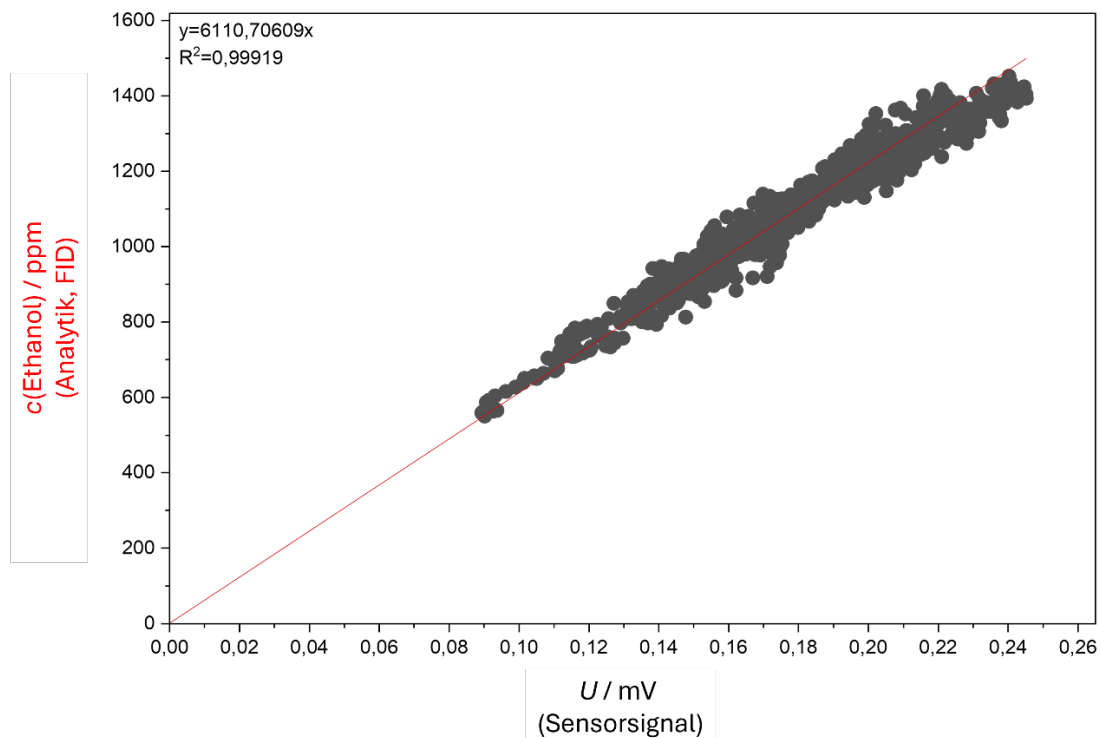


Abb. 2: Direkte Korrelation der Messdaten von Analytik und Sensor aus Abbildung 1. Ermittelte Steigung aus der linearen Sensorkennlinie als Empfindlichkeit des Sensors auf Ethanol: 164 μV / 1000 ppm.

Charakterisierung und Anwendung von additiv gefertigten Metallkomponenten in der Gasmesstechnik

Michael Stopp¹, Oliver Maurer², Dirk Bähre², Christian Bur¹

¹*Lehrstuhl für Messtechnik, Universität des Saarlandes, 66123 Saarbrücken, Deutschland*

²*Lehrstuhl für Fertigungstechnik, Universität des Saarlandes, 66123 Saarbrücken, Deutschland*

+49(0)681-302-3904, (m.stopp), (c.bur) @imt.uni-saarland.de

oliver.maurer@uni-saarland.de, lft@mx.uni-saarland.de

Zusammenfassung

In dieser Studie wurden Metallkomponenten additiv mittels Laser-Powder-Bed-Fusion (L-PBF) Verfahren gefertigt und auf Eignung als funktionelle Komponenten in der Gasmesstechnik untersucht. Als untersuchte Materialien wurden eine Aluminium- und eine Edelstahllegierung verwendet. Wasserstoff zur Prüfung auf Gasdichtheit und Propofol (2,6-Diisopropylphenol) zur Untersuchung von Verschleppungseffekten wurden mit einer Gasmischanlage beaufschlagt und Metalloxid Halbleitersensoren wurden als Detektor benutzt. Die additivgefertigten Komponenten wurden mit etablierten Werkstoffen (Metalle und Kunststoffe), in der Gasführungstechnik verglichen, welche als Normbauteile verfügbar sind. Bei beiden Versuchsreihen zeigten sich die gefertigten Testkörper sowohl als gasdicht als auch als unkritisch hinsichtlich der Verschleppung von chemischen Stoffen. von der Nutzbarkeit der Technologie, um mit kurzer Vorlaufzeit, hoher Gestaltungsfreiheit und geringem Nachbearbeitungsbedarf, metallische Komponenten für speziell maßgeschneiderte Messaufgaben herzustellen.

Keywords: Gassensorik, Additive Fertigung, MOS-Sensoren, Mobile Messsysteme, Medizintechnik

Motivation und Einleitung

Neben der eigentlichen Sensorik, Elektronik und Signalverarbeitung ist in der Gasmesstechnik auch die Gasführung der Misch- und Messstrecke zum Sensor ein wichtiger Teil des Gesamtsystems, welcher das Messergebnis beeinflusst und Raum zur Optimierung aufzeigt [1].

Hier bietet die Fertigungstechnologie des Laser-Powder Bed Fusion (L-PBF) eine vielversprechende Möglichkeit, Komponenten für die gasführenden Teile eines Messsystems zu generieren. Es bietet die geometrischen Freiheiten, wie sie vom klassischen 3D-Druck her bekannt sind und gleichzeitig können chemisch inerte und beständige Materialien, wie Edelstähle, verarbeitet werden.

Klassische additive Fertigungsverfahren (oft als „3D-Druck“ bezeichnet) nutzen hauptsächlich Polymere als Werkstoff, Nachteilig ist hier insb. die Offenporigkeit und die damit mangelnde Gasdichtheit. Weiterhin kritisch sind Ausgasungen von organischen Verbindungen, die damit die Gaszusammensetzung ändern und das Messergebnis verfälschen [2].

Durch die Herstellungsweise des schichtweisen aufschmelzen aus dem Pulverbett in Argon-Atmosphäre, (Abb. 1) haben die Oberflächen jedoch immer eine deutlich raue Oberfläche,

vergleichbar mit Metallguss. In innenliegenden Kanalstrukturen nur mit aufwändigen Verfahren, wie Gleitfließlappen, bearbeiten lassen. Weiterhin ist durch Laserstrahldurchmesser und Korngröße des verwendeten Metallpulvers die Auflösung der fertigen Strukturen nach unten begrenzt. Damit ergibt sich eine minimal fertigbare Geometrie und, insbesondere für die Gasmesstechnik wichtig, eine minimale Wandungsstärke gasführender Kanäle. Hier muss sichergestellt sein, dass sich keine Porenstrukturen in den Wandbereichen bilden, welche zu einer Gasundichtigkeit des Bauteils führen.

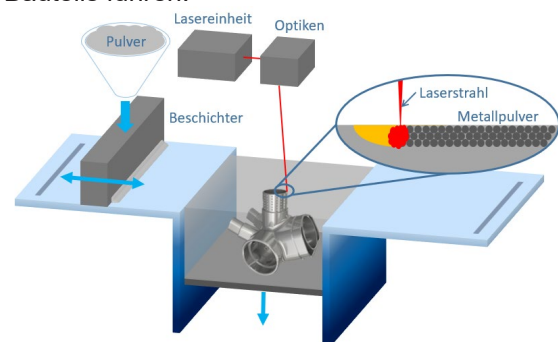


Abb. 1: Vereinfachte Darstellung des L-PBF Bauprozesses, nach [3].

Ziel dieser Arbeit ist es daher, anhand von Testgeometrien, L-PBF-Bauteile auf ihre Gasdichtheit und auf mögliche Verschleppung bei stark adsorbierenden Analyten in Prüfgasgemischen hin zu untersuchen. Für schwer- und teillüchtige Substanzen besteht die Gefahr von Verschleppungseffekten durch Anhaftung an inneren Oberflächen. Durch die höhere Oberflächenrauheit - und damit größere Oberfläche - von additiv gefertigten Bauteilen wird dieser Effekt noch begünstigt. Als weitere Ziel dieser Arbeit werden daher Verschleppungseffekte mit Propofol untersucht.

Design und Fertigung der Versuchskörper

Zur Fertigung aller Testkörper und Anwendungsbauteile in diesem Projekt wurde eine L-PBF Anlage der Baureihe SLM 125 (Nikon SLM Solutions AG, Lübeck, Deutschland) verwendet, welche über ein kubisches Bauvolumen von 125 mm x 125 mm x 125 mm verfügt. Bauteile können zwar frei in diesem Volumen angeordnet werden, benötigen jedoch trotzdem zusätzliche Stützstrukturen, unter anderem zur Wärmeabfuhr, um Verformung und Rissbildung vorzubeugen [4]. Dadurch ergeben sich, trotz hoher Designfreiheit, gewisse Beschränkungen in dem Entwurf der additiv gefertigten Bauteile und ihrer Positionierung im Bauraum. Weiterhin beeinflussen das Pulver (Material und Körnung) sowie die Ausrichtung des Bauteils den Bauprozess.

Für die Teststrukturen in den Versuchsreihen wurden zwei verschiedene Materialien untersucht: eine Aluminiumlegierung (AlSi10Mg) und eine Edelstahllegierung (316L; 1.4404) (beide von m4p Material Solutions GmbH, Magdeburg, Deutschland), wobei alle, nachfolgend an die Versuchsreihen gefertigten, anwendungsbezogenen Bauteile nur aus 316L gefertigt wurden.

Zum Test auf Druckfestigkeit wurde ein tonnenförmiges Probendesign entworfen (Abb. 2 links). Untersuchungen zu Verschleppungen und Memory Effekten sowie zur Diffusionsfestigkeit von dünnen Wandungen wurde mit einer Doppelkanalprobe (siehe Abb. 2 rechts und Abb. 3) durchgeführt.

Hierbei nähern sich zwei gasführende Kanäle im Bauteil auf einer definierten Strecke aneinander an und sind nur über eine definierte Wandstärke voneinander getrennt. Die Wandstärke wurde variiert (siehe Tab. 1) und soll ein Grenzwert aufzeigen, ab dem eine Gasdichtheit nicht mehr gewährleistet ist.

Diese Bauteile (Abb. 2) wurden sowohl in 316L als auch in AlSi10Mg hergestellt, minimal spanend nachbearbeitet (1/8"-Gewinde und bei Bedarf Planschliff der Dichtfläche) und gereinigt, bevor sie mit Swagelok-Anschlüssen zur

Integration in den Versuchsaufbau versehen wurden.

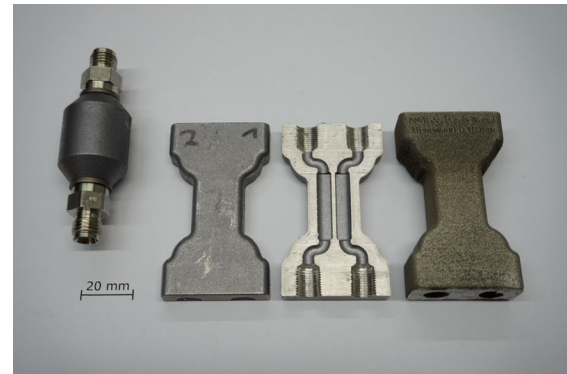


Abb. 2: Mittels L-PBF gefertigte Probengeometrien. Links: Tonnenproben, mit Swagelok-Anschlüssen. Mitte: Aufgeschnittene Doppelkanalprobe aus AlSi10Mg. Rechts: Doppelkanalprobe aus 316L.

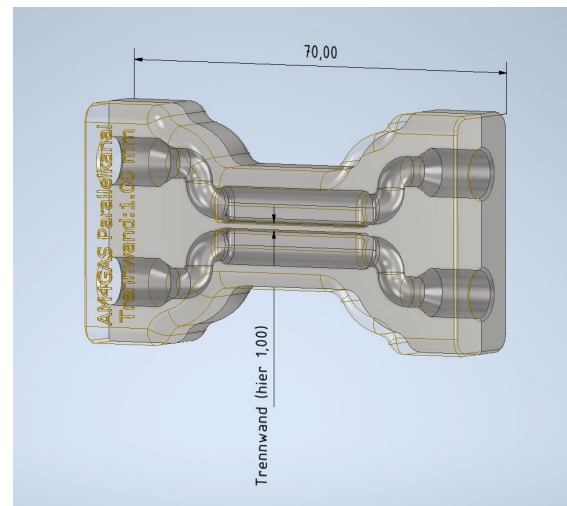


Abb. 3: CAD-Modell der Doppelkanalprobe, zur Visualisierung der Kanalgeometrie.

Tab. 1: Hergestellte Wandstärkenvariationen der Prüfkörpergeometrien.

Art der Geometrie (Material)	Wandstärke [mm]	Anzahl gefertigter Teile
Tonnenprobe (AlSi10Mg)	0,75	3
	1	3
	1,5	3
Doppelkanal (AlSi10Mg)	0,75	2
	1	2
	1,5	2
Doppelkanal (316L)	0,3	1
	0,5	1
	0,75	1

Methodik des Messaufbaus für Gasdichte- heit und Verschleppung

Für die Untersuchung zur „Gasdichtheit“ und „Verschleppung“ wurden Gasmischanlagen (GMA) [5], [6] eingesetzt, die gezielt Prüfgasgemische erzeugen können.

Flüchtige Verbindungen werden üblicherweise über Prüfgasflaschen bereitgestellt, wohingegen schwerflüchtige Verbindungen wie Propofol über Permeationsröhrchen und zugehörige Permeationsöfen (Dynacalibrator Model 150, VICI Metronics Inc., Twelve Trees, USA) bereitgestellt. Die relative Feuchte des Trägergasstromes wird durch die Anfeuchtung von Nullluft über Waschflaschen und das anschließende zumischen in den Hauptstrom realisiert. Der Trägergasstrom wird anschließend durch die Messstrecke mit der Doppelkanalprobe und den Sensoren und Messgeräten geleitet.

Bei den Dichtheitsversuchen wurde immer jeweils ein Kanal mit dem Prüfgas statisch beaufschlagt, während im Messkanal ein konstanter Fluss Nullluft von 500 mL/min bei 50% relativer Feuchte (r. F.) eingestellt bleibt. Das vereinfachte Schema des Messaufbaus ist in Abb. 4 oben dargestellt.

Durch die geänderten Anforderungen bei den Verschleppungsmessungen wurde eine andere GMA verwendet, welche Propofol über einen

Permeationsofen bereitstellen kann. Hierzu wird der Permeationsofen mit Nullluft durchströmt, wodurch sich eine Sättigungskonzentration abhängig von der Permeationsrate des verwendeten Permeationsröhrchens einstellt. Diese wird über Massenflussregler mit vorberechnetem Fluss in den Hauptstrom der GMA beigemischt, sodass sich die gewünschte Konzentration von 18 ppb Propofol im Prüfgas befindet. Weiterhin ist die Anordnung der Sensoren verändert, da über Sensoren vor der Probe und Sensoren nach der Probe die Rückkehr des Sensorsignals auf den Wert vor dem Gasangebot beobachtet werden soll (Abb. 4 unten).

Überprüfung der Gasdichtheit

Zum Nachweis der Gasdichtheit wird Wasserstoff mit einer Konzentration von 4 vol-% verwendet, da es aufgrund seiner kleinen Moleküle leicht durch Materialien hindurchwandern kann. Um eine Gefährdung durch Explosionsgefahr während der Versuchsreihen auszuschließen, wurde ausschließlich Wasserstoff in Stickstoff genutzt, dessen Konzentration von 4 vol.-% H_2 unterhalb der unteren Explosionsgrenze (UEG) liegt [7]. Dieser wurde unverdünnt auf den zu prüfenden Kanal beaufschlagt.

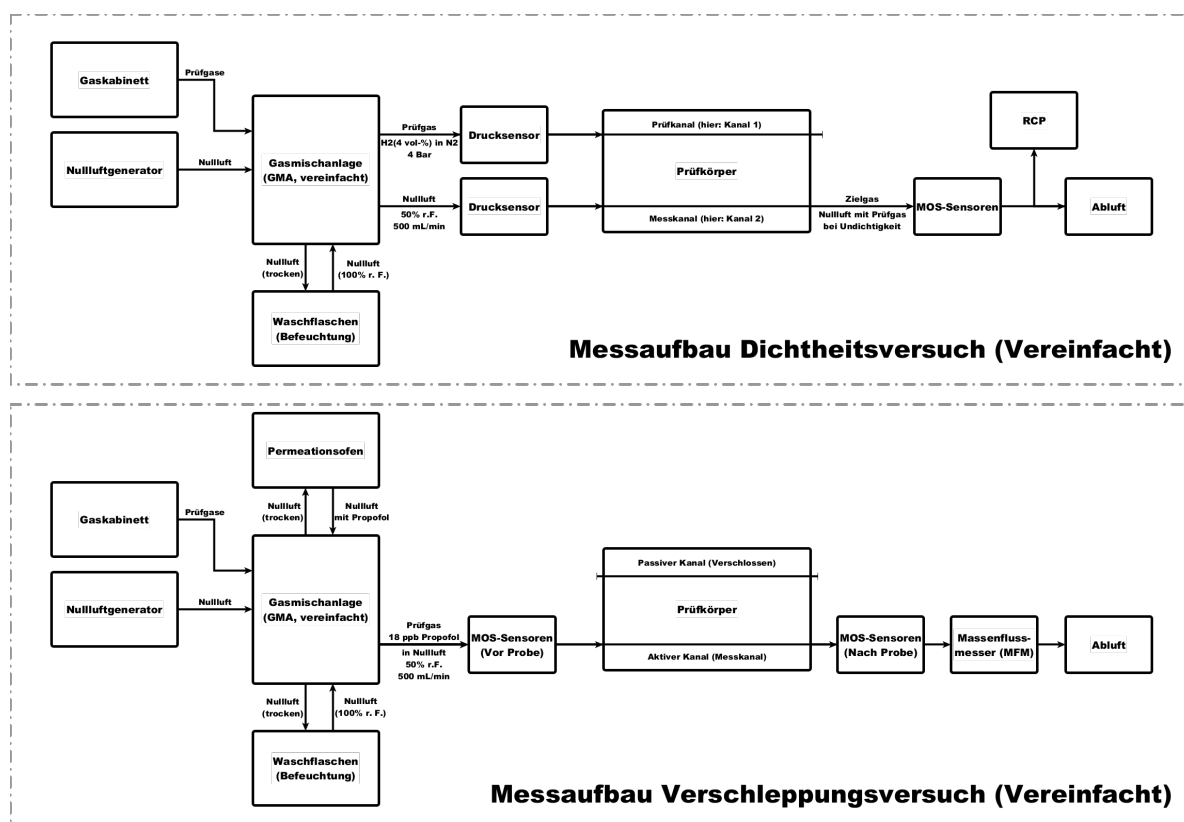


Abb. 4: Vereinfachte Schemata der Messaufbauten mit den Parallelkanalproben. Oben: Messaufbau für den Dichtheitsversuch. Unten: Messaufbau für die Verschleppungsversuche.

Zur Detektion des Wasserstoffes im Falle einer Undichtigkeit werden sowohl Metalloxidhalbleiter-Gassensoren (MOS-Sensoren) verwendet (Abb. 5) und in diesem Aufbau zusätzlich ein selektiv auf H_2/CO messendes Reducing Compund Photometer (RCP, Peak Performer 1 RCP von Peak Laboratories LLC, Mountain View, USA).

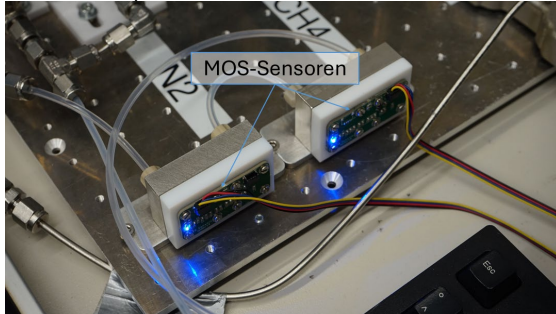


Abb. 5: MOS-Gassensor Boards, nach dem Messkanal der Probe, in der Messstrecke eingebunden.

Somit kann eine Sensorausschlag im Leitwert der MOS-Gassensoren (z.B. durch Feuchteschwankungen) bei gleichzeitigem Ausbleiben eines gestiegenen Messwertes am RCP als Undichtigkeit der Kanaltrennwand ausgeschlossen werden. In diesem Versuchsaufbau wurden zwei Multi-Sensor-Boards mit fünf MOS-Gassensoren (ein Board temperaturkonstant, ein Board temperaturzyklisch) und einem Temperatur- und Feuchtesensor verwendet (siehe Tab. 2).

Tab. 2: In der Gasdichtheitsprüfung und Verschleppungsversuch verwendete Sensoren

Sensorname (Hersteller)	Sensortyp
BME680 (Bosch Sensortec GmbH, Reutlingen, Deutschland)	Temperatur, Feuchte, Druck, MOS-Gassensor (nur Gassensor genutzt)
ENS160 (ScioSense B.V., Eindhoven, Niederlande)	MOS-Gassensor (4 Schichten)
SGP40 (Sensirion AG, Stäfa, Schweiz)	MOS-Gassensor (4 Schichten)
ZMOD4410 (Renesas Electronics Corporation, Japan)	MOS-Gassensor
ZMOD4510 (Renesas Electronics Corporation, Japan)	MOS-Gassensor
SHT35 (Sensirion AG, Stäfa, Schweiz)	Temperatur und Feuchte

Die zu erwartende Menge an H_2 im Messkanal bei einer Undichtigkeit wurde hier durch eine vereinfachte Finite-Elemente-Simulation mit Hilfe der Software COMSOL Multiphysics 5.4 abgeschätzt (siehe Abb. 6) und würde bei ei-

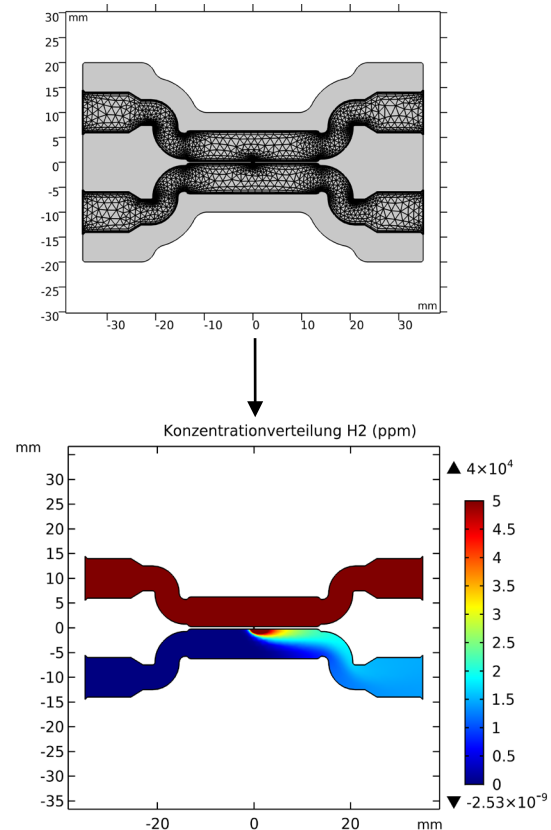


Abb. 6: Vereinfachtes Finite-Elemente Modell des Doppelkanals (oben) und Ergebnis der Simulation zur erwarteten Konzentration (unten).

ner Undichtigkeit mit einer modellierten Risslänge von $100\ \mu\text{m}$ am Ende des Messkanales bei $1435\ \text{ppb}$ liegen. Sowohl die MOS-Gassensoren als auch das RCP können deutlich geringere Konzentrationen nachweisen [8]. Jeweils ein Kanal der Probe wird dabei mit einem Blindstopfen verschlossen und an die Leitung mit dem Wasserstoffprüfgas angeschlossen. Dies wird in einem definierten Zeitraum (Prüfphase) von vier Stunden über ein Ventil mit 4 Bar Überdruck Wasserstoffprüfgas mit 4 vol.-% H_2 in N_2 beaufschlagt (vgl. Abb. 7 Kanal „1“). Der Druck wurde mit Drucksensoren kontinuierlich geprüft (Abb. 7). Im zweiten Kanal (vgl. Abb. 7 Kanal „2“) wird ein kontinuierlicher Fluss von $500\ \text{mL/min}$ Nullluft mit 50% relativer Feuchte bei Raumtemperatur (20°C) eingestellt und nach der Probe sowohl an den MOS-Sensoren vorbeigeleitet als auch an einer Zweigstelle vom RCP abgesaugt und von diesem zyklisch gemessen.

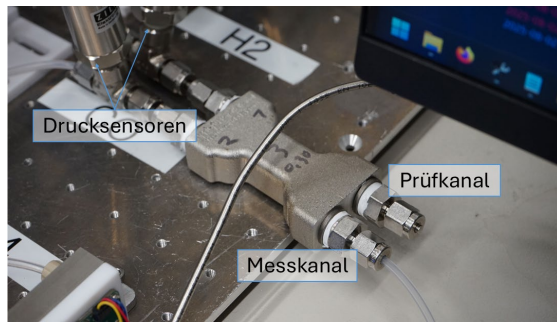


Abb. 7: Prüfung der Doppelkanalprobe mit Prüfkanal (Kanal 1) und Messkanal (Kanal 2) sowie Drucksensoren.

Eine Messreihe dauert dabei insgesamt zwölf Stunden und gliedert sich in drei Teilabschnitte:

1. Einlaufen und Stabilisierung der Sensoren auf dem Messkanal bei 500 mL/min, keine Gasbeaufschlagung auf dem Prüfkanal für sechs Stunden.
2. Bedrücken des Prüfkanals mit 4 Bar Überdruck Wasserstoffprüfgas bei gleichbleibendem Fluss auf Prüfkanal für vier Stunden.
3. Prüfkanal drucklos schalten und Auslaufphase des Messvorganges bei unverändertem Fluss auf Messkanal für zwei Stunden.

Ein Ausschlag der MOS-Sensoren während Phase 2, bei gleichzeitigem Anstieg des H₂-Messwertes am RCP, definiert dabei eine Undichtigkeit des jeweils getesteten Prüfkörpers. Zwecks Validierung und zum Ausschluss systematischer Fehler wurden die Kanäle in einer zweiten Messreihe vertauscht. Dadurch wurde jeder Prüfkörper effektiv zweimal auf Undichtigkeit geprüft.

Die Abb. 8 zeigt exemplarisch eine Messung des Edelstahl Prüfkörpers mit 0,3 mm Wandstärke zwischen den beiden Kammern (dünnste Wandstärke). Die ersten vier Diagramme zeigen das logarithmierte Sensorsignal (Leitwert) der vier gassensitiven Schichten des SGP40 MOS-Sensoren bei konstanter Temperatur von 400°C. Im untersten Diagramm ist der Messwert des RCP dargestellt. Der rote Bereich markiert die Phase der Wasserstoffbeaufschlagung mit 4 Bar Überdruck im zu prüfenden Kanal. In Blau ist das Sensorsignal von Kanal 1 und in Rot von Kanal 2 als Messkanal gezeigt. Somit wurde beim blauen Grafen Kanal 2 und beim roten Grafen Kanal 1 mit dem Prüfgas beaufschlagt.

Weder eine der Sensorschichten noch das RCP zeigten einen signifikanten Anstieg während des Zeitraums der Wasserstoffbeaufschlagung. Dies belegt die Gasdichtheit des additiv gefertigten Bauteils. Vergleichbare Er-

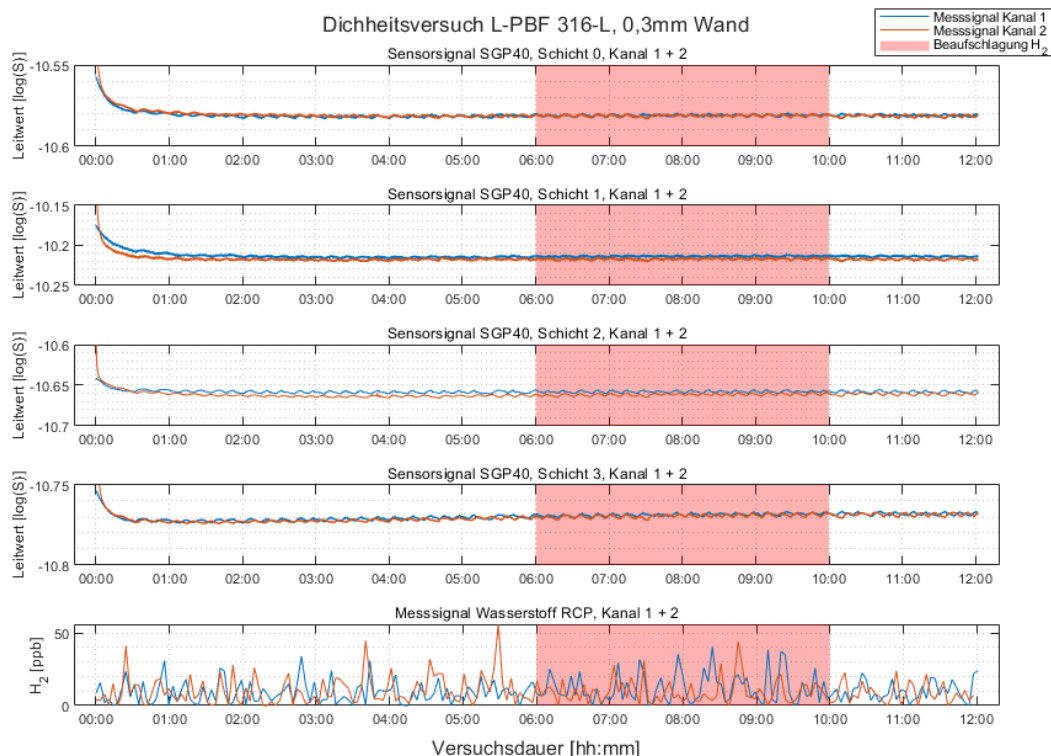


Abb. 8: Dichtheitsversuch mit Wasserstoffbedruckung einer Doppelkanalprobe. Gezeigt sind exemplarisch die vier gassensitiven Schichten des SGP40 Sensors für den Prüfkörper aus 316 Edelstahl mit einer Wandstärke von 0,3 mm. Zusätzlich ist unten die Überwachungsmessung des RCP dargestellt. In Blau ist Kanal 1 und in Rot Kanal 2 dargestellt.

gebnisse wurden mit allen getesteten Wandstärken erzielt sowie mit der Aluminiumlegierung als auch in Edelstahl. Alle Prüfkammern wurden, durch den Kanaltausch, mehrfach gemessen.

Dieses Ergebnis lässt bereits eine Verwendung für in L-PBF gefertigte Bauteile mit integrierten Gasführungen mit minimalen Wandstärken bis 0,3 mm zu.

Untersuchungen von Verschleppungseffekten

Zur Untersuchung der Verschleppung wurde ebenfalls die Doppelkanalkammer genutzt, wobei nur einer der beiden Kanäle betrachtet wird. Allerdings wurde ein MOS-Gassensorarray, mit Sensorbestückung nach Tab. 2, sowohl vor und ein weiteres Array nach der Doppelkanalprobe im Gasstrom vorgesehen. Diese Anordnung (Abb. 10) ermöglicht die Untersuchung von Verschleppungseffekten durch Adsorption und Desorption von teils flüchtigen Substanzen an inneren Wandungen. Der unbenutzte Kanal wird dennoch mit Blindstopfen verschlossen, um eine Anlagerung von Stoffen aus der Laborumgebung während der Mess- und Lagerzeit auszuschließen.

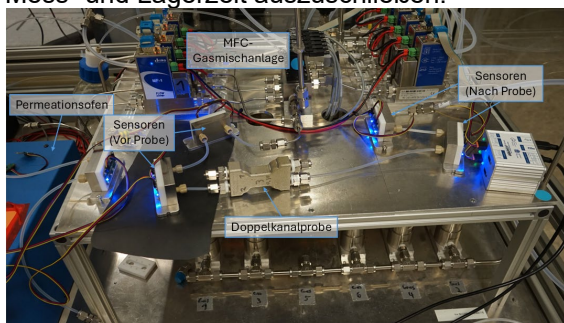


Abb. 9: Messaufbau an GMA mit Permeationsofen, zur Überprüfung der Verschleppung.

Der Messablauf dauert insgesamt 24 Stunden und gliedert sich in folgende Phasen:

1. Einlaufen der Sensoren und Herstellen einer Grundlinie der MOS-Sensoren bei 500 mL/min mit 50% relativer Feuchte bei Raumtemperatur (20°C) für acht Stunden.
2. Aufschalten des Prüfstoffes (18 ppb 2,6-Diisopropylphenol) für acht Stunden.
3. Beenden des Gasangebotes mit Propofol und Spülen des Kanales für weitere acht Stunden mit Nullluft zum Ermitteln der Verschleppung (Sensorrücklauf auf die Grundlinie).

Anschließend wurde die Messung mit dem zweiten Kanal derselben Probe wiederholt, wodurch auch hier jede Probe zweimal vermessen wurde. Zur Bewertung der Anhaftung von Propofol an inneren Wandungen und da-

mit von Verschleppungseffekten, wurde die Abklingzeit nach Propofol-Aufgabe ausgewertet (siehe Abb. 11 für Edelstahl und Abb. 12 für Aluminium). Hier sind in Blau (vor Probe) und Rot (nach Probe) die Sensorsignale von Kanal 1 dargestellt und in Gelb (vor Probe) und Lila (nach Probe) das Sensorsignale von Kanal 2. Der Zeitraum des Gasangebotes mit 18 ppb Propofol ist in Hellgelb farbig hinterlegt.

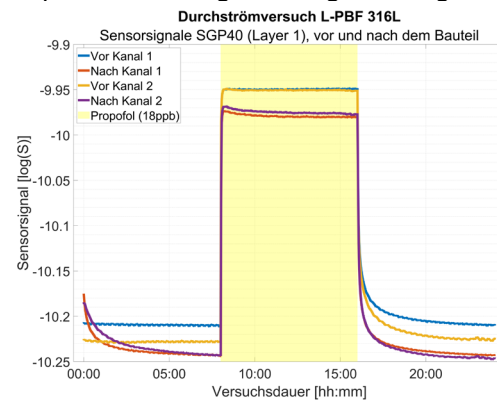


Abb. 10: Verschleppungsversuch an L-PBF Doppelkanal aus 316L. Es wurden jeweils dieselben Sensoren bei Kanal 1 und 2 verwendet. Zeitraum der Propofolangebots ist in Gelb hinterlegt.

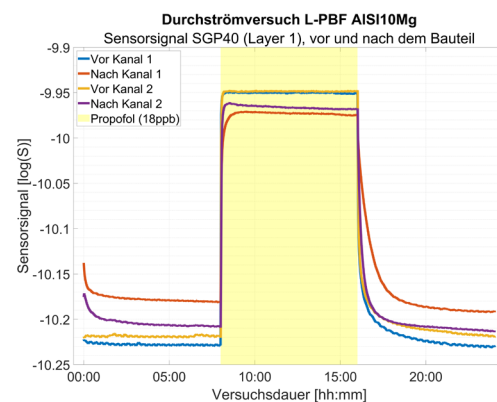


Abb. 11: Verschleppungsversuch an L-PBF Doppelkanal aus AISi10Mg. Darstellung analog zu Abb. 11.

Hierbei zeigt sich im direkten Vergleich zwischen den beiden Materialien, mit einem SGP40 mit entfernter Schutzmembran, dass die Doppelkanalproben aus 316L (Abb. 11) ein schnelleres Rücklaufen auf die Grundlinie besitzen als die Proben aus AISi10Mg (Abb. 12). Insbesondere an Kanal 1 (rote Kurve) in Abb. 12 ist ein längeres Rücklaufen auf die Grundlinie zu beobachten.

Dieses Verhalten ist bei den Doppelkanalproben aus AISi10Mg ausgeprägter, wodurch sich das Material 316L besser für eine Anwendung mit Medienberührung eignet.

Vergleich der Verschleppung bei L-PBF Teilen mit konventionellen Leitungen

Neben dem Vergleich von additiv gefertigten Bauteilen mit einer Aluminiumlegierung und aus Edelstahl, wird im Folgenden der Edelstahlprüfkörper mit Rohrstücken und Schläuchen, die standardmäßig zur Gasführung eingesetzt werden (vgl. Tab. 3), verglichen.

Tab. 3: In den Verschleppungsversuchen verwendete Materialien.

Werkstoff	Ausführung
Edelstahl (316L)	1/4"-Rohr
	L-PBF
Aluminium (unlegiert)	6 mm Rohr
Aluminium (AISI10Mg)	L-PBF
Perfluorakloxy-Polymer (PFA)	1/4"-Rohr
Polyurethan (Lebensmittel-echt) (PUR-L)	6 mm Schlauch

Die Länge der Rohr- und Schlauchstücke wurde abhängig von dem Innendurchmesser so angepasst, dass die nominelle (d.h. ohne Beachtung der Rauheit) Oberfläche, der der additiv gefertigten Prüfkörper entspricht. Zusätzlich wurde eine Messung ohne Bauteil (nur Swagelok-Verbinder) durchgeführt.

Das Bauteil aus 316L zeigt hier im Vergleich (Abb. 13) einen Verlauf, welcher mit dem von etablierten Materialien wie PFA (lila Kurve) [9] oder gezogenen Edelstahlrohren (gelbe Kurve) vergleichbar sind. Dem gegenüber ist für Schlauch aus PUR-L (grüne Kurve) selbst innerhalb der Abklingzeit nicht mehr auf den Wert vor dem Gasangebot abgefallen.

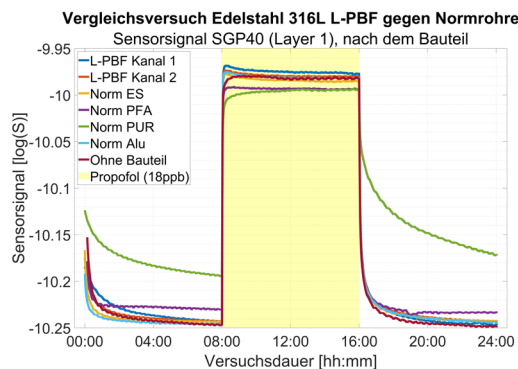


Abb. 12: Vergleich des Verschleppungsversuch von 316L mit Vergleichsmaterialien und ohne Messstrecke (dunkelrote Kurve).

Beispielhafte L-PBF Bauteile mit Anwendungsbezug

Neben den Prüfkörpern wurden auch anwendungsspezifische Körper durch additive Fertigung realisiert.

Eine erste Anwendung ist eine Sensoraufnahme (Abb. 14) für die Sonde eines Sauerstoffsensors sowie der Gasanschluss für ein Kapnometer. Die Aufnahme ist Teil eines modifizierten Oxygenators einer Herz-Lungen-Maschine, die in einem ex-vivo Schweinelungenmodell [10] eingesetzt wird (Abb. 15).

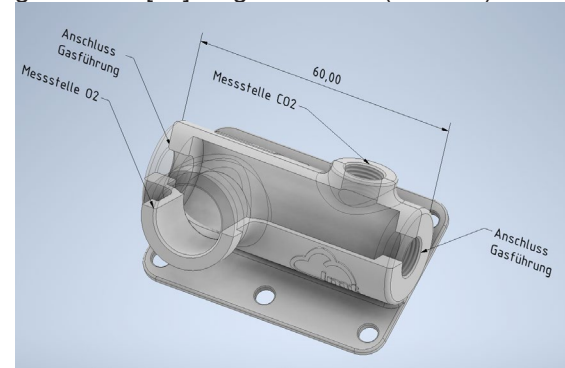


Abb. 13: CAD-Modell des Sensoradapters für die Oxygenatorsteuerung, mit halbtransparentem Dreiviertelschnitt und Anschlussbezeichnung.

Bei diesem wird an zwei Stellen im Gasfluss O₂ und CO₂-Gehalt der Luft gemessen.

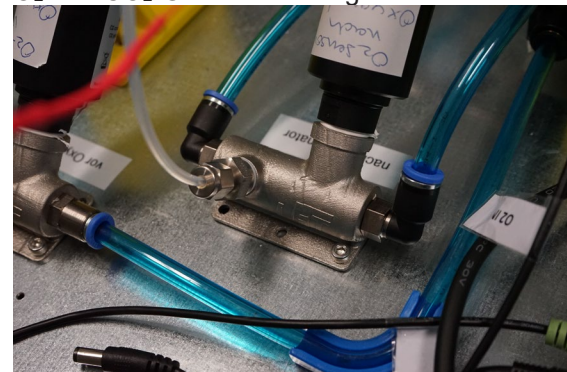


Abb. 14: Messstelle im Oxygenator mit gefertigtem Sensoranschluss aus 316L.

In einem zweiten Beispiel wurde integrierte fluidische Verbindungen innerhalb eines miniaturisierten Ventilblocks (Abb. 16) für ein mobiles Gasmesssystem realisiert.

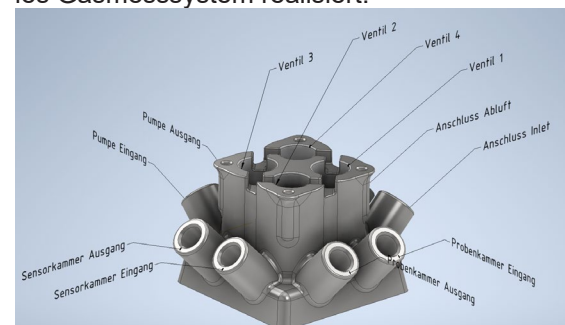


Abb. 15: CAD-Modell des Ventilblockes mit Anschlussbezeichnungen.

Hier wird die Verschaltung von vier 3/2-Wege-Kartuschen-Ventilen (Typ Bürkert 6164) untereinander komplett durch integrierte Kanäle mit durchschnittlich 1,4 mm Kanaldurchmesser abgebildet. Hierdurch wurde das Totvolumen der Ventilvorrichtung minimiert, was zu kürzeren Gaswechselzeiten führt.

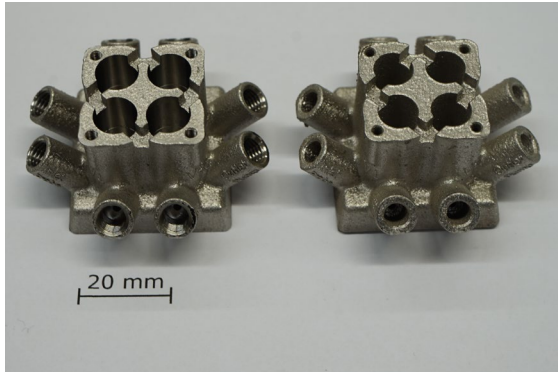


Abb. 16: L-PBF Bauteil „Ventilblock“, für vier 3/2-Wege-Kartuschenventile. Links: spanend nachbearbeitet. Rechts: L-PBF Rohteil.

Fazit und Ausblick

In dieser Arbeit konnte gezeigt werden, dass durch per L-PBF Verfahren additiv hergestellte Bauteile sich prinzipiell für den Einsatz in der Gasmesstechnik eignen. Hierbei ist von den beiden untersuchten Materialien, AlSi10Mg und 316L, das Letztere durch seine weniger stark ausgeprägten Verschleppungseffekte besser für medienberührende Bauteile mit Prüfgasen, geeignet. Die Aluminiumlegierung ermöglicht allerdings ebenfalls gasdichte Bauteile und bietet sich durch ihre deutlich bessere Wärmeleitfähigkeit für temperierte Bauteile und unkritische Gase an.

Das Fertigungsverfahren selbst ermöglicht die Herstellung von metallischen Bauteilen mit kurzer Vorlaufzeit bei gleichzeitig hoher Designfreiheit und geringem Nachbearbeitungsbedarf. Gezeigt wurde das für die Sensoraufnahme zur Anwendung in der medizinischen Forschung. Weiterhin ermöglicht das Fertigungsverfahren hochintegrierte Bauteile mit Fluidik-Komponenten, da Gaskanäle hergestellt werden können, welche mit klassischen, spanenden Fertigungsverfahren nicht realisierbar sind.

In zukünftigen Anwendungen werden L-PBF Komponenten eine weitere Verringerung des Totvolumens bei gleichzeitig verbesserter Integration weiterer Funktionen, wie integrierten Ventilen, ermöglichen.

Förderhinweis

Die hier vorgestellten Arbeiten wurden im Rahmen des Landesforschungsförderungspro-

gramm – LFFP Open Call „Drittmitel.Transformation.Saarland“ unter dem Vorhaben „Untersuchung des Potentials von Additiver Fertigung (3D-Druck) von Metallen für die Gasmesstechnik“ vom Ministerium der Finanzen und für Wissenschaft des Saarlandes gefördert.

Literaturnachweis

- [1] O. Brieger, C. Bur, A. Schütze, und T. Sauerwald, „Setup for characterization of MOS gas sensors as GC detector“, *Sensors and Actuators B: Chemical*, Bd. 444, S. 138396, Dez. 2025, doi: 10.1016/j.snb.2025.138396.
- [2] B. J. Lotesoriere u. a., „Experimental setup to study poisoning effects of different materials on chemical sensors used in E-nose systems“, *Journal of Sensors and Sensor Systems*, Bd. 14, Nr. 2, S. 237–247, Okt. 2025, doi: 10.5194/jsss-14-237-2025.
- [3] O. Maurer, „Sono-L-PBF: Entwicklung und Qualifikation von schallunterstütztem Laser Powder-Bed Fusion von AlSi10Mg“, Dissertation, Universität des Saarlandes, Saarbrücken, 2024. doi: 10.22028/D291-43539.
- [4] O. Maurer, M. Stopp, C. Bur, und D. Bähre, „Influence of the supported part surface area on part properties in Laser Powder-Bed Fusion of 316L for gas measurement accessory“, *Applications in Engineering Science*, Bd. 25, S. 100309, März 2026, doi: 10.1016/j.apples.2026.100309.
- [5] M. Leidinger, C. Schultealbert, J. Neu, A. Schütze, und T. Sauerwald, „Characterization and calibration of gas sensor systems at ppb level—a versatile test gas generation system“, *Meas. Sci. Technol.*, Bd. 29, Nr. 1, S. 015901, Jan. 2018, doi: 10.1088/1361-6501/aa91da.
- [6] D. Arendes, J. Amann, C. Tessier, O. Brieger, A. Schütze, und C. Bur, „Qualification and optimisation of a gas mixing apparatus for complex trace gas mixtures“, *tm - Technisches Messen*, Bd. 90, Nr. 12, S. 822–834, Dez. 2023, doi: 10.1515/teme-2023-0075.
- [7] „Wasserstoff“, GESTIS-Stoffdatenbank. [Online, letzter Zugriff am 14.04.2026]. Verfügbar unter: <https://gestis.dguv.de/data?name=007010>
- [8] T. Baur, J. Amann, C. Schultealbert, und A. Schütze, „Field Study of Metal Oxide Semiconductor Gas Sensors in Temperature Cycled Operation for Selective VOC Monitoring in Indoor Air“, *Atmosphere*, Bd. 12, Nr. 5, S. 647, Mai 2021, doi: 10.3390/atmos12050647.
- [9] F. Maurer u. a., „Adherence of volatile propofol to various types of plastic tubing“, *J. Breath Res.*, Bd. 11, Nr. 1, S. 016009, Jan. 2017, doi: 10.1088/1752-7163/aa567e.
- [10] S. Radermacher u. a., „A versatile ex-vivo model platform for detection of exhaled drugs using isolated perfused porcine lungs“, *J. Breath Res.*, Bd. 20, Nr. 2, S. 026004, Juni 2026, doi: 10.1088/1752-7163/ae4bfe.

Ultraschall-Messmethode zur Analyse des Heilungsfortschritts von Knochenfrakturen

Link, Eric N.¹; Lützelberger, Jan¹; Drese, Klaus Stefan¹

¹ISAT - Institut für Sensor- und Aktortechnik, Hochschule für angewandte Wissenschaften Coburg, Am Hofbräuhaus 1b, 96450 Coburg, Deutschland.

Eric-Nepomuk.Link@hs-coburg.de; Jan.Luetzelberger@hs-coburg.de; Klaus.Drese@hs-coburg.de

Zusammenfassung

Die Überwachung des Heilungsverlaufs von Knochenfrakturen stellt in der klinischen Praxis eine zentrale Herausforderung dar. Die Beurteilung der Frakturheilung erfolgt derzeit überwiegend mittels Röntgendiagnostik. Frühzeitige mechanische Veränderungen im Frakturspalt können damit jedoch nur eingeschränkt erfasst werden, zudem entsteht eine Strahlenbelastung. Daher besteht Bedarf an einer nicht-invasiven und strahlungsfreien Alternative. In diesem Beitrag wird eine ultraschallbasierte Messmethode vorgestellt, welche die Analyse des Heilungsverlaufs von Knochenfrakturen ermöglichen soll. Grundlage ist die Annahme, dass Änderungen der Steifigkeit und der akustischen Impedanz im Frakturspalt die Transmission geführter Ultraschallwellen beeinflussen. Zur Untersuchung wurden Frakturmodelle mit aushärtendem Epoxidharz verwendet und die Empfangssignale mithilfe der Hilbert-Einhüllenden ausgewertet. Dabei zeigten Messungen an ebenen und hohlzylindrischen Proben während der Aushärtung einen kontinuierlichen Anstieg der Empfangsamplitude. Dieses Verhalten stimmt mit der erwarteten Zunahme von Steifigkeit und akustischer Impedanz im Frakturspalt überein. Zudem erwies sich die Methode als robust gegenüber geometrischen Variationen der Proben. Die Ergebnisse zeigen das Potenzial des Verfahrens zur nicht-invasiven und strahlungsfreien Überwachung von Konsolidierungsprozessen.

Keywords: Ultraschall, Frakturheilung, GAW, Hilbert-Peak-Analyse, Knochenphantom

Einleitung

Knochenfrakturen zählen zu den häufigsten Verletzungen im klinischen Alltag und stellen insbesondere bei verzögerter oder ausbleibender Heilung (Nonunion) ein erhebliches medizinisches und sozioökonomisches Problem dar [1,2]. Eine frühzeitige und zuverlässige Beurteilung des Heilungsverlaufs ist daher entscheidend, um therapeutische Maßnahmen rechtzeitig anpassen zu können [3].

Die Knochenheilung verläuft in mehreren Phasen: nach der Fraktur bildet sich zunächst ein Hämatom, das eine Entzündungsreaktion initiiert [4]. In den folgenden Tagen entsteht Granulationsgewebe, aus dem sich ein weicher Kallus mit zunächst geringer mechanischer Stabilität entwickelt [5]. Im weiteren Verlauf kommt es zur Mineralisation und zur Bildung von geflechtartigem Knochen, wodurch die Steifigkeit deutlich zunimmt, bevor über mehrere Monate ein Umbau zu lamellärem Knochen erfolgt [4,5,6]. Der Heilungsverlauf ist schematisch in Abbildung 1 dargestellt.

Der aktuelle Stand der klinischen Diagnostik basiert überwiegend auf bildgebenden Verfahren wie der Röntgendiagnostik [3,7]. Diese erlaubt die Beurteilung fortgeschrittener Heilungsstadien, weist jedoch in frühen Phasen eine geringe Sensitivität gegenüber mechanischen Veränderungen im Frakturspalt auf und ist zudem mit Strahlenbelastung verbunden [3,7]. Außerdem werden ultraschallbasierte Sensorensysteme zwar seit über 20 Jahren an Knochen erforscht [8], jedoch wurde bislang keine Methodik für die Untersuchung der Gesamtknochenfraktur etabliert [9].

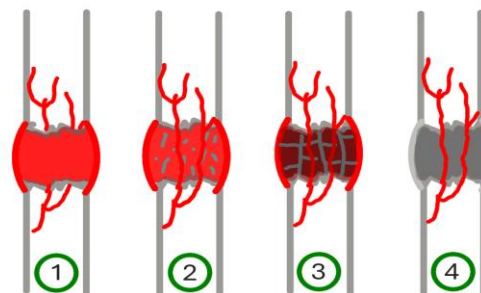


Abb. 1: Phasen der Knochenheilung (1 Verletzungsphase; 2 Granulationsphase; 3 Kallushärtungsphase; 4 Remodellierungsphase)

Damit besteht das zentrale Problem darin, dass die für den Heilungsverlauf entscheidende Zunahme der mechanischen Steifigkeit im Frakturspalt mit etablierten Methoden nur eingeschränkt und oft erst in späten Stadien erfasst werden kann, sodass eine frühzeitige Detektion einer verzögerten Heilung oder Nonunion bislang nur begrenzt möglich ist [1].

Vor diesem Hintergrund wurde in der vorliegenden Arbeit ein ultraschallbasierter Messansatz zur nicht-invasiven Überwachung der Frakturheilung untersucht. Ziel war es, die zeitliche Entwicklung der mechanischen Eigenschaften im Frakturspalt mit einem geeigneten Ersatzmodell experimentell abzubilden und eine Auswertungsmethodik zu entwickeln, welche es ermöglicht, die aus der Veränderung der mechanischen Eigenschaften resultierende Veränderungen in der Ultraschallausbreitung zeitlich zu verfolgen.

Methoden

Zur Durchführung der Untersuchungen werden im Folgenden die relevanten methodischen Grundlagen dargestellt. Hierzu zählen zunächst die physikalischen Mechanismen der Ultraschallausbreitung sowie deren Wechselwirkung mit dem Frakturmodell. Darauf aufbauend wird die Auswertungsmethodik zur Bestimmung charakteristischer Signalparameter beschrieben. Abschließend wird der verwendete experimentelle Versuchsaufbau erläutert.

Physikalisches Prinzip

Röhrenförmige Knochen wie beispielsweise das Schienbein können aus physikalischer Sicht näherungsweise, als hohlzylindrische Strukturen beschrieben werden. Eine Fraktur lässt sich dabei vereinfacht als Unterbrechung beziehungsweise Spalt im Hohlzylinder modellieren, dessen Elastizitätsmodul deutlich geringer ist als das des umgebenden intakten Knochengewebes. Im Verlauf der Frakturheilung steigt das Elastizitätsmodul des Kallusgewebes kontinuierlich an, wodurch sich die mechanischen Eigenschaften des Frakturbereichs zunehmend denen des gesunden Knochens annähern.

Die Wellenausbreitung in einem isotropen elastischen Festkörper wird durch die elastische Wellengleichung beschrieben

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} \quad (1)$$

dabei beschreibt $\mathbf{u}$ den Verschiebungsvektor, während λ und μ die Lamé-Konstanten sind [10].

In den vorliegenden Platten- und Hohlzylinder-Modellen führen die Randbedingungen an den freien Oberflächen dazu, dass longitudinale und transversale Wellen gekoppelt werden. In diesem Fall spricht man von sogenannten geführten akustischen Wellen (GAW).

Im Frequenzbereich bis 120 kHz treten hauptsächlich die beiden Fundamentalmoden L_0 und F_0 auf [10]. Die transversale Grundmode F_0 reagiert besonders empfindlich auf lokale Änderungen von Steifigkeit und Dicke und eignet sich daher besonders für die Untersuchung von Veränderungen im Bereich des Frakturspalts [10].

Die Phasengeschwindigkeit c_{ph} einer geführten akustischen Welle (GAW) ist definiert als

$$c_{ph} = \frac{\omega}{k} \quad (2)$$

und die Gruppengeschwindigkeit c_{gr} als

$$c_{gr} = \frac{d\omega}{dk} \quad (3)$$

mit der Kreisfrequenz ω und der Kreiswellenzahl k [10]. Im Inneren einer Knochenfraktur, steigt das Elastizitätsmodul E während der Ausheilung um mehrere Größenordnungen an [6]. Der Anstieg im Elastizitätsmodul führt wiederum zu einer ansteigenden Schallgeschwindigkeit im Frakturspalt. Die akustische Impedanz Z eines Mediums ist definiert als

$$Z = \rho c \quad (4)$$

mit der Dichte ρ und der Schallgeschwindigkeit c [10]. Dadurch führt eine Erhöhung des Elastizitätsmoduls ebenfalls zu einer höheren akustischen Impedanz.

Trifft eine Ultraschallwelle senkrecht auf die Grenzfläche zweier Medien, so wird ein Teil der Welle reflektiert und ein Teil transmittiert. Der Amplituden-Reflexionskoeffizient R ergibt sich zu

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1} \quad (5)$$

und der Transmissionskoeffizient T zu

$$T = \frac{2Z_2}{Z_2 + Z_1} \quad (6)$$

wobei Z_1 die Impedanz des Ursprungsmediums und Z_2 die Impedanz des Mediums auf, welches die Welle übertragen werden soll, darstellt [10].

Diese Veränderungen beeinflussen die Ausbreitung von Ultraschallwellen. Insbesondere die Transmission zwischen zwei Bereichen des Knochens wird durch den Impedanzunterschied zwischen Knochen und Frakturspalt bestimmt. Die zeitliche Entwicklung der empfangenen Ultraschallamplitude kann daher als Maß für den Fortschritt der Frakturheilung genutzt werden.

Je größer der Unterschied zwischen den Impedanzen beider Medien ist, desto größer ist die Reflexion und desto geringer ist die Transmission. Im frühen Stadium der Frakturheilung besitzt der Frakturspalt nur eine geringe akustische Impedanz, hierbei wurde ein Wert ähnlich Weichgewebe und Blut von 1.5 MRayl [11] angenommen für das Material im Inneren der Knochenfraktur, für den Knochen selbst wurde eine typische Impedanz von 7.8 MRayl [11] zur Abschätzung verwendet. Der Impedanzsprung wird im Heilungsverlauf kleiner und ein größerer Anteil der Ultraschallenergie wird transmittiert. Daraus folgt eine zunehmende Empfangsamplitude.

Versuchsaufbau und Probenauswahl

Für die Modellierung wurde eine typische Tibia-Wand-Stärke im Bereich von 2 mm bis 6 mm angenommen [12]. Insgesamt kamen zwei Epoxidharz-Platten-Modelle sowie drei Hohlzylinderphantome mit unterschiedlichen Wandstärken zum Einsatz. Die experimentell bestimmten Plattendicken betragen (3 ± 0.35) mm und (4 ± 0.35) mm, die Wandstärken der Hohlzylinderphantome betragen (2.85 ± 0.1) mm, (6.85 ± 0.1) mm und (11 ± 0.1) mm. Letzterer wurde bewusst oberhalb des physiologischen Bereichs der Tibia gewählt, um auch größere Wandstärken, wie sie beispielsweise bei Knochen des Oberschenkels auftreten, systematisch zu untersuchen und den Einfluss der Geometrie über den Bereich der Tibia hinaus analysieren zu können.

Zur experimentellen Umsetzung des Materials im Frakturspalt wurden vereinfachte Frakturmodelle verwendet, bei dem ein definierter Spalt mit einem Epoxidharz (CW 1302 GB mit Härter HY 1300 GB, Araldite) gefüllt wurde. Dieses weist während der Aushärtung eine kontinuierliche Zunahme der Steifigkeit sowie im ausgehärteten Zustand akustische Eigenschaften auf, die eine geeignete Modellierung des Frakturspalts ermöglichen. Das Epoxidharz wurde unmittelbar vor Beginn jeder Messung in den künstlich erzeugten Frakturspalt eingebracht.

Für die Ultraschallanregung wurden piezoelektrischer Wandler vom Typ PIC255 (Bezeichnung: PIC255 77362 x10 y1.9 t4) im

Transmissionsbetrieb verwendet, wobei das Signal auf der gegenüberliegenden Seite des Spaltes mit aushärtendem Epoxidharz detektiert wurde (vgl. Abb. 2).

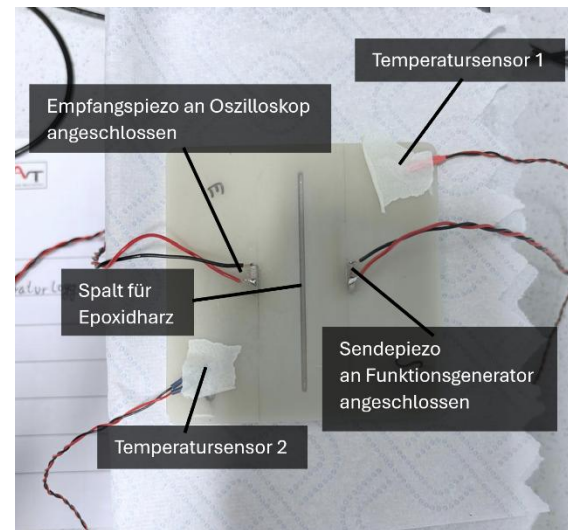


Abb. 2: Proben für Plattenversuche

Zur Ansteuerung des Sendepiezoes wurde ein Funktionsgenerator (Typ 33521A, Agilent Technologies) verwendet. Die Anregung erfolgte im Puls-Betrieb mit einer Spitzenspannung von 10 V und einer Puls-Periode von 10 ms. Sowohl das Trigger- als auch das Sendesignal wurden als Referenz auf ein Digitaloszilloskop (WaveRunner 604Zi, Teledyne LeCroy) geführt. Zur Messung des Sendesignals kam ein 10:1-Tastkopf zum Einsatz.

Aus den bereits beschriebenen Hohlzylinderphantomen wurden jeweils 5 cm lange Probenstücke gefertigt und paarweise in einer Halterung mit einem definierten Spalt von 5 mm fixiert. Der innere Hohlraum wurde mittels eines Luftballons abgedichtet, um ein Austreten des Epoxidharzes zu verhindern.

Der äußere Spaltbereich wurde mit Klebeband versiegelt. Anschließend wurde das Harz über eine Öffnung eingebracht, die danach ebenfalls verschlossen wurde. Während der Aushärtung wurde über mehrere Stunden die Empfangsamplitude der transmittierten Ultraschallwelle (vgl. Abb. 4) aufgezeichnet.

Der gesamte Versuchsaufbau wurde in einer Temperaturkammer bei 40 °C betrieben (vgl. Abb. 3).

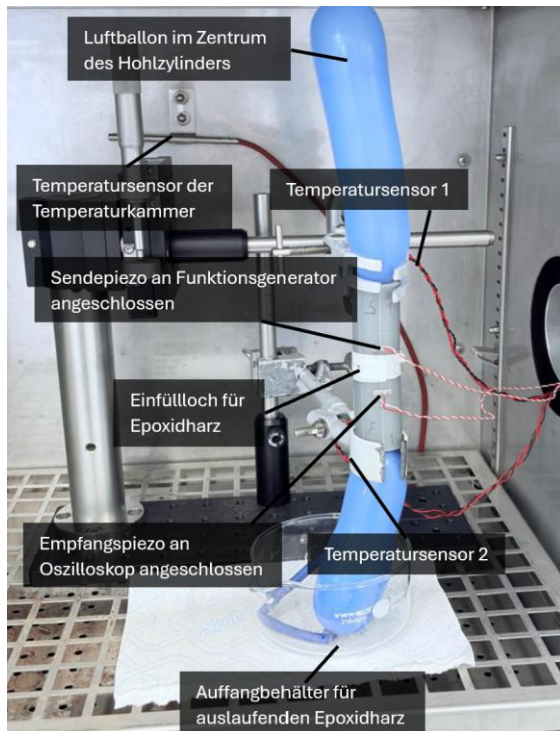


Abb. 3: Versuchsaufbau mit Hohlzylinder

Prinzip der Signalauswertung

Zur Anregung wurde ein zweifacher sinusförmiger Puls mit Hanning-Fenster verwendet. Die Mittenfrequenz betrug 100 kHz. Ein Schema dieses Signals ist in Abbildung 4 dargestellt. Diese Frequenz wurde auf Grundlage von Dispersionsrechnungen mit der Software *Dispersion Calculator* Version 3.1 des Deutschen Zentrums für Luft- und Raumfahrt (DLR) gewählt, da aus den Dispersionsdiagrammen hervorgeht, dass in diesem Frequenzbereich überwiegend fundamentale Moden auftreten und Überlagerungen höherer Moden weitgehend vermieden werden können.

Das ausgesendete Signal lässt sich durch

$$s(t) = \sin(2\pi f_0 t) \cdot w(t) \quad (7)$$

beschreiben, wobei f_0 die Mittenfrequenz und $w(t)$ die Hanning-Fensterfunktion ist.

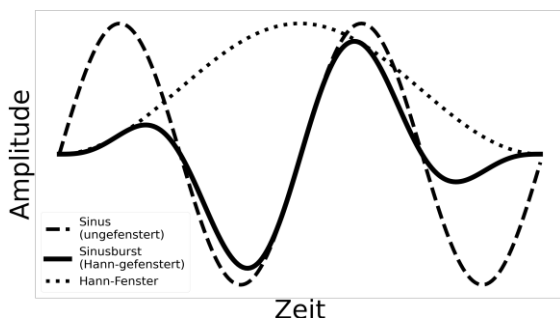


Abb. 4: Anregesignal, 2-fach-Sinus mit Hanning-Fenster

Die Empfangssignale wurden in Zeitabständen von 30 Sekunden für den mehrere Stunden andauernden Aushärtungsprozess des Epoxidharzes aufgezeichnet. Dadurch konnte die zeitliche Entwicklung des Ultraschallsignals während der Aushärtung verfolgt werden.

Zur Auswertung wurde zunächst für jedes Signal die Hilbert-Transformation berechnet. Daraus ergibt sich das analytische Signal

$$\hat{y}(t) = y(t) + i\mathcal{H}\{y(t)\} \quad (8)$$

mit dem gemessenen Signal $y(t)$ und dessen Hilbert-Transformation $\mathcal{H}\{y(t)\}$. Die Hüllkurve [13] ergibt sich aus

$$A(t) = |\hat{y}(t)| \quad (9)$$

und wurde als Grundlage für die Merkmalsextraktion verwendet. Das Verfahren wurde in Abbildung 5 bildlich dargestellt.

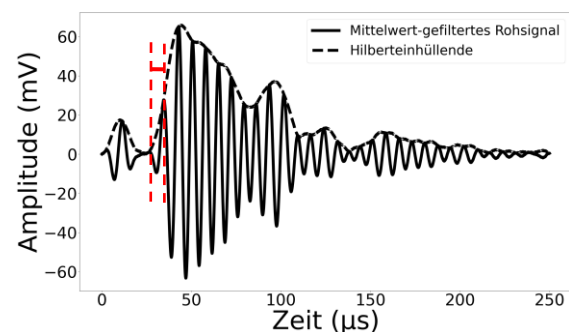


Abb. 5: Schematische Darstellung der Auswertung anhand der Rohsignale (der markierte Bereich zeigt das Auswertungsfenster der Hilberteinhüllenden)

Anschließend wurde innerhalb eines definierten Zeitfensters der Mittelwert der Hüllkurve bestimmt. Dieser Mittelwert wurde als Kennwert für die empfangene Ultraschallamplitude verwendet. Die Lage des Zeitfensters wurde so gewählt, dass ausschließlich der relevante Anteil der transmittierten Welle erfasst wurde. Störsignale sowie spätere Reflexionen, die an den Probenrändern entstanden sein könnten, wurden dadurch weitgehend unterdrückt bzw. ausgeschlossen.

Da einzelne Messungen durch Störungen oder kleine Bewegungen der Sensoren beeinflusst werden konnten, wurden die Daten zusätzlich geglättet und auf Ausreißer untersucht. Hierzu wurde ein gleitender Mittelwert sowie ein Hampel-Filter verwendet [14].

Ergebnisse

Nach der Auswertung der Messsignale über den gesamten Messzeitraum mithilfe des zuvor beschriebenen Verfahrens wurden die zeitlichen Verläufe der Signalamplituden bestimmt und analysiert. Im Folgenden werden die Ergebnisse für die verschiedenen Probengeometrien dargestellt und miteinander verglichen.

Versuche an Epoxidharz-Platten

Unter temperaturgeregelten Bedingungen zeigte sich für beide Plattendicken ein stabilerer Verlauf der Empfangsamplitude. Nach der Befüllung des Spalts nahm die Empfangsamplitude über mehrere Stunden kontinuierlich zu. Der Verlauf war dabei bei beiden Proben qualitativ sehr ähnlich. Zu Beginn der Messung wurden nur geringe Signalamplituden registriert. Im weiteren Verlauf stieg die Amplitude zunehmend an und erreichte schließlich ein Maximalniveau, der gesamte Verlauf wurde in Abbildung 6 dargestellt.

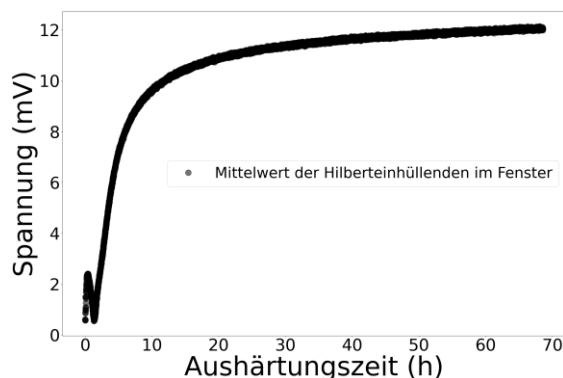


Abb. 6: Verlauf der Hilbert-Einhüllenden im relevanten Zeitbereich des Empfangssignals während der Epoxidharz-Aushärtung an einem Platten-Modell mit 4 mm Dicke

Versuche an Epoxidharz-Hohlzylindern

Für die Hohlzylinderproben zeigte sich ein Signalverlauf, welcher dem der Plattenversuchen ähnelte. Ein Beispiel eines solchen Verlaufes wurde in Abbildung 7 dargestellt. Die beobachteten Verläufe waren zwischen den einzelnen Proben sehr ähnlich.

Trotz unterschiedlicher Wanddicken und geringfügiger Unterschiede im Versuchsaufbau trat in allen Messreihen derselbe grundsätzliche Trend auf, die Amplitude des Empfangssignals stieg an und näherte sich zum Ende des Versuchszeitraumes einem Maximum an.

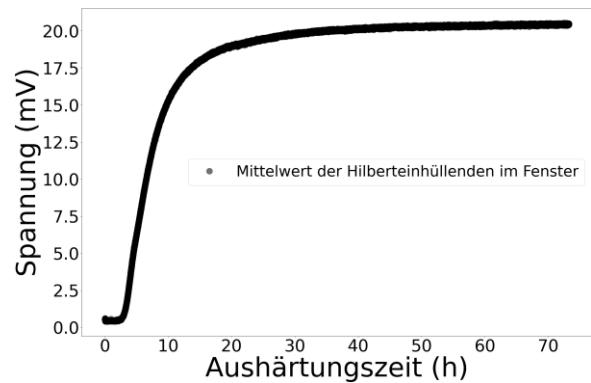


Abb. 7: Verlauf der Hilbert-Einhüllenden im relevanten Zeitbereich des Empfangssignals während der Epoxidharz-Aushärtung an einem Hohlzylinder-Modell mit 2.85mm Wandstärke

Die absolute Höhe der Empfangsamplitude unterschied sich jedoch zwischen den einzelnen Hohlzylindern. Proben mit geringerer Wanddicke zeigten in der Regel höhere Signalamplituden als dickwandige Proben.

Neben der reinen Amplitudenentwicklung wurde auch der Zusammenhang zwischen dem Signalverlauf und der Qualität der ausgehärteten Klebeschicht untersucht. Hierzu wurden die Proben nach Abschluss der Messung visuell bewertet. Es zeigte sich, dass Messreihen mit einem besonders starken und gleichmäßigen Amplitudenanstieg meist mit einer guten und vollständigen Benetzung des Spalts, sowie die lotrechte Ausrichtung der Hohlzylinderstücke zueinander einhergingen. Ein Beispiel hierfür ist in Abbildung 8b dargestellt. Dagegen führten Luft einschüsse oder unvollständig gefüllte Bereiche, wie sie in Abbildung 8a zu erkennen sind, zu geringeren Empfangsamplituden sowie teilweise zu verrauschten und wenig reproduzierbaren Verläufen.



Abb. 8: Darstellung des Klebespaltes von zwei verschiedenen Hohlzylindern, a) Luft einschüsse im Klebespalt, b) ohne Luft einschüsse

Diskussion

Die Versuche an ebenen Platten bestätigten zunächst die grundsätzliche Eignung des Messprinzips zum Erfassen des Aushärtungsprozesses. Anschließend lieferten die hohlzylindrischen Proben im Vergleich zu den Platten eine realitätsnähere Darstellung der Schallausbreitung, wie sie in medizinischen Anwendungen am Knochen der Fall wäre. Ihre Geometrie entspricht eher der Form eines Röhrenknochens und stellt damit eine realistischere Annäherung an die klinische Situation dar. Trotz unterschiedlicher Wandstärken wurde in nahezu allen Versuchen der gleiche grundsätzliche Verlauf beobachtet. Dies spricht dafür, dass die entwickelte Methode nicht auf einen einzelnen Versuchsaufbau beschränkt ist, sondern auch unter veränderten geometrischen Randbedingungen funktioniert.

Bei den durchgeführten Messungen war zu erkennen, dass sich der Aushärtungsprozess des Epoxidharzes im modellhaften Frakturspalt anhand der Empfangsamplitude geführter Ultraschallwellen verfolgen ließ. Hierbei war vor allem die Beobachtung, dass dünnwandigere Proben in der Regel eine höhere Signalamplitude als dickwandigere Proben besaßen, plausibel, da bei größerer Wanddicke eine stärkere Dämpfung und eine komplexere Modenstruktur auftreten. Dennoch blieb der zeitliche Verlauf auch bei den dickeren Proben deutlich erkennbar. Auch die Beobachtung eines kontinuierlichen und überwiegend monotonen Anstiegs der Hilbert-Peaks über mehrere Stunden bei den hohlzylindrischen Proben spricht für die Eignung des Messprinzips. Dieses Verhalten stimmt mit der zugrunde liegenden Hypothese überein, dass die fortschreitende Aushärtung des Epoxidharzes zu einer Zunahme der mechanischen Steifigkeit und der akustischen Impedanz im Frakturspalt führt. Dadurch verringert sich der Impedanzsprung zwischen Spalt und umgebendem Material, sodass ein größerer Anteil der Ultraschallwelle durch den Spalt transmittiert wird.

Des Weiteren reagierte die Ultraschallmessung nicht nur auf die zeitliche Veränderung der Materialeigenschaften während der Aushärtung, sondern auch auf die Qualität der Verbindung im Epoxidharzspalt. Der Klebespalt wurde nach der Aushärtung hinsichtlich seiner Qualität geprüft; hierbei wurde auf Lufteinschlüsse und die Lotrechtigkeit der einzelnen Rohrstücke zueinander geachtet. Dabei fiel auf, dass Klebeverbindungen mit wenigen Lufteinschlüssen und guter Ausrichtung der Hohlzylinderstücke eine höhere Empfangsamplitude besaßen als Proben, bei denen dies nicht der Fall war. Dennoch

konnte gezeigt werden, dass auch Proben mit schlechter ausgebildeter Verbindung im Verlauf der Zeit eine Zunahme der Festigkeit aufweisen. Der Unterschied zu gut verbundenen Proben bleibt jedoch in der Auswertung deutlich sichtbar.

Die Messdaten deuten darauf hin, dass die Empfangsamplitude eine sensitive Größe zur Beschreibung des Aushärtungsprozesses darstellt. Bereits geringe Änderungen der Materialeigenschaften im Spalt führen zu messbaren Unterschieden in der transmittierten Amplitude.

Trotz der vielversprechenden Ergebnisse ist die Aussagekraft der vorliegenden Untersuchungen noch eingeschränkt. Die verwendeten Proben bilden die reale Knochenheilung nur vereinfacht ab und berücksichtigen weder biologische Prozesse noch die komplexe Geometrie eines echten Knochens. Auch die Einflüsse von Weichgewebe zwischen Messsystem und Knochen konnten hierdurch nicht abgebildet werden.

Dennoch zeigen diese Vereinfachungen der Untersuchungen an idealisierten Knochen- und Frakturmodellen, dass das verwendete Messprinzip grundsätzlich geeignet ist, Veränderungen der mechanischen Eigenschaften im Frakturbereich mittels Ultraschalls zu erfassen. Die beobachteten Zusammenhänge zwischen den simulierten Heilungszuständen und den gemessenen Signalparametern verdeutlichen das Potenzial der Methode für eine nichtinvasive und strahlungsfreie Analyse des Frakturheilungsverlaufs. Die vorliegenden Ergebnisse stellen somit einen ersten Schritt auf dem Weg zu einem ultraschallbasierten Sensorsystem zur objektiven Beurteilung der Knochenheilung dar.

Für zukünftige Arbeiten sollten die verwendeten Modelle schrittweise an die physiologischen Bedingungen angepasst werden. Hierzu zählen insbesondere die Verwendung geometrisch realitätsnäherer Knochenmodelle sowie die Einbeziehung geeigneter Weichgewebe-Ersatzmaterialien zur Berücksichtigung der Schallausbreitung durch das umliegende Gewebe. Darüber hinaus ist die Integration von Sender, Empfänger sowie der Ansteuer- und Auswerteelektronik in ein kompaktes Sensorsystem anzustreben, um die praktische Anwendbarkeit der Methode weiter zu erhöhen und eine spätere Übertragung in klinische Anwendungen zu ermöglichen.

Literaturverzeichnis

- [1] Cunningham B.P., Brazina S., Morshed S., Miclau T.; Fracture healing: A review of clinical, imaging and laboratory diagnostic options, *Injury*. 2017; 48(1):69-75.
<https://doi.org/10.1016/j.injury.2017.04.020>
- [2] Antonova, E., et al.; Tibia shaft fractures: costly burden of nonunions. *BMC musculoskeletal disorders*. 2013; 14(1):42.
<https://doi.org/10.1186/1471-2474-14-42>
- [3] Bowers K.M., Anderson D.E.; Delayed Union and Nonunion: Current Concepts, Prevention, and Correction: A Review. *Bioengineering*. 2024; 11(6):525.
<https://doi.org/10.3390/bioengineering11060525>
- [4] LaStayo, Paul C., Kerri M., Winters, Hardy M.; Fracture healing: bone healing, fracture management, and current concepts related to the hand. *Journal of Hand Therapy*. 2003; 16(2):81-93.
[https://doi.org/10.1016/S0894-1130\(03\)80003-0](https://doi.org/10.1016/S0894-1130(03)80003-0)
- [5] Sathyendra, V., Darowish, M.; Basic Science of Bone Healing, *Hand Clinics*. 2013; 29(4):473-481.
<https://doi.org/10.1016/j.hcl.2013.08.002>
- [6] Schell, H., Duda, G.N., Peters, A. et al.; The haematoma and its role in bone healing. *Journal of Experimental Orthopaedics*. 2017; 4(1):5.
<https://doi.org/10.1186/s40634-017-0079-3>
- [7] Hak D.J., Fitzpatrick D., Bishop J.A., Marsh J.L., Tilp S., Schnettler R., Simpson H., Alt V.; Delayed union and nonunions: Epidemiology, clinical issues, and financial aspects, *Injury*. 2014; 45(2).
<https://doi.org/10.1016/j.injury.2014.04.002>
- [8] Tran T.N.H.T., Le L.H., Ta D.; Ultrasonic Guided Waves in Bone: A Decade of Advancement in Review. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2022 Oct; 69(10):2875-2895.
<https://doi.org/10.1109/TUFFC.2022.3197095>
- [9] Protopappas, V.C., Vavva M.G., Fotiadis D.I., Malizos K.N.; Ultrasonic monitoring of bone fracture healing. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*. 2008 Jun; 55(6):1243-1255.
<https://doi.org/10.1109/TUFFC.2008.787>
- [10] Rose J.L.; Ultrasonic Guided Waves in Solid Media. *Cambridge University Press*. 2014;
<https://doi.org/10.1017/CBO9781107273610>
- [11] Ludwig G.D.; The velocity of sound through tissues and the acoustic impedance of tissues. *The journal of the acoustical society of America*. 1950; 22(6):862-866.
<https://doi.org/10.1121/1.1906706>
- [12] Zadeh, A.F., Hanafi, M.G., Kiasat, A., Mousavi, M.; Evaluation of the tibial cortical thickness accuracy in osteoporosis diagnosis in comparison with dual energy X-ray absorptiometry. *Journal of Family Medicine and Primary Care*. 2019 Feb; 8(2):523-527.
https://doi.org/10.4103/jfmpc.jfmpc_456_18
- [13] Feldman M.; Hilbert transform in vibration analysis, *Mechanical Systems and Signal Processing*. 2011; 25(3):735-802,
<https://doi.org/10.1016/j.ymssp.2010.07.018>
- [14] Pearson R.K., Neuvo Y., Astola J. et al.; Generalized Hampel Filters. *EURASIP Journal on Advances in Signal Processing*. 2016, 87
<https://doi.org/10.1186/s13634-016-0383-6>

Ultraschall-Durchflussmessung mit geführten akustischen Wellen: Der Einsatz von Zylindermoden in Rohrsystemen

Alexander Backer¹, Klaus Stefan Drese¹

*¹Institut für Sensor- und Aktortechnik (ISAT), Hochschule für angewandte Wissenschaften Coburg
Am Hofbräuhaus 1b, 96450 Coburg, Deutschland
Alexander.Backer@hs-coburg.de, Klaus.Drese@hs-coburg.de*

Zusammenfassung

In dieser Arbeit wird ein Ansatz zur ultraschallbasierten Durchflussmessung unter Verwendung von Zylindermoden vorgestellt. Im Gegensatz zu etablierten Verfahren wie der diagonalen Durchschallung oder Lambwellen wird hierbei das gesamte Rohr samt Fluid zum Schwingen angeregt, wodurch eine nicht-invasive Messung über die Rohrwand ermöglicht wird. Zur gezielten Sensorauslegung wurde ein Berechnungsansatz entwickelt, der auf der Analyse von Dispersionsdiagrammen mit und ohne Hintergrundströmung basiert. Dadurch lässt sich der zu erwartende Messeffekt aus der Änderung der Phasen- und Gruppengeschwindigkeit bestimmen.

Die Simulationen wurden mit COMSOL Multiphysics unter Berücksichtigung turbulenter Strömungsprofile durchgeführt und dienten der Identifikation eines optimalen Arbeitspunktes, der bei 280 kHz gefunden wurde. Darauf aufbauend wurde ein Sensor mit ringförmig angeordneten Piezokeramiken entwickelt und experimentell an einem Wasserprüfstand validiert.

Die Messergebnisse zeigen eine gute Übereinstimmung der berechneten und experimentell bestimmten Signallaufzeitverschiebungen. Insgesamt demonstriert die Arbeit, dass die Berechnung von Dispersionsdiagrammen mit Hintergrundströmung eine effiziente und zielgerichtete Entwicklung von Ultraschalldurchflusssensoren ermöglicht und großes Potenzial für zukünftige Anwendungen bietet.

Keywords: Nicht-Invasive Durchflussmessung, Ultraschalldurchflussmessung, Geführte Akustische Wellen, Zylindermoden

Einleitung

Der Durchfluss von Flüssigkeiten und Gasen ist in der industriellen Prozessführung, der Herstellung von flüssigen Produkten in der Lebensmittel- und Chemieindustrie sowie dem Vertrieb von z. B. fossilen Energieträgern eine zentrale Messgröße und wird selbst im privaten Haushalt für die Abrechnung des Brauchwassers benötigt [1]. Durch die hohe Anzahl an Einsatzgebieten und den unterschiedlichsten Einsatzbedingungen wurden in den vergangenen Jahrzehnten eine Vielzahl von Durchflusssensoren, basierend auf unterschiedlichsten physikalischen Prinzipien entwickelt [1]. Neben thermischen, Turbinen- und Coriolis-Sensoren sind vor allem Ultraschallbasierte Durchflusssensoren durch die Möglichkeit einer nicht invasiven Messung weit verbreitet [1]. Neben der klassischen, diagonalen Durchschallung des Fluids gibt es auch Varianten basierend auf Lambwellen [2]. Lambwellen sind geführte akustische Wellen, die unter einem bestimmten Winkel in das Fluid einstrahlen und dadurch eine Durchflussmessung, ähnlich der diagonalen Durchschallung, ermöglichen. In diesem

Beitrag wird eine dritte Variante der Ultraschallbasierten Durchflussmessung unter der Verwendung von Zylindermoden beschrieben. Zylindermoden sind eine Art von geführten akustischen Wellen, die sich entlang von zylindrischen Strukturen, wie es ein flüssigkeitsgefülltes Rohr ist, ausbreiten können [3]. Im Vergleich zu der klassischen, diagonalen Durchschallung bzw. der Einstrahlung der Lambwellen unter einem bestimmten Winkel wird bei dieser Wellenart das gesamte Rohr inklusive Fluid zum Schwingen angeregt und es findet eine axial gerichtete Wellenausbreitung statt. Durch die Ausbreitung der Welle sowohl in der Rohrwand als auch in dem Fluid entsteht eine Sensitivität gegenüber Fluidströmungen, die durch die Interaktion zwischen Fluid und Rohrwand über die Außenseite des Rohres zugänglich ist und für die Durchflussmessung verwendet werden kann. Diese Methode wurde bereits für die Durchflussmessung im Milliliter- und Mikroliter-Bereich eingesetzt und wurde jetzt auf größere Rohrdurchmesser und höhere Flussraten übertragen [4,5,6].

Berechnung des Messeffekts

Für die Entwicklung und Auslegung eines Ultraschalldurchflusssensors ist die Berechnung des zu erwartenden Messeffekts in Abhängigkeit der Rohr- und Fluideigenschaften von entscheidender Bedeutung. Bei der diagonalen Durchschallung bzw. der Verwendung von Lambwellen lässt sich der Messeffekt über den Einstrahlwinkels in das Fluid berechnen. Bei Zylindermoden hingegen existiert dieser Winkel nicht, wodurch ein anderer Ansatz basierend auf Dispersionsdiagrammen verwendet wurde. Der Grundgedanke besteht darin ein Dispersionsdiagramm ohne und mit Rohrströmung zu berechnen und anhand der Veränderung der Ausbreitungsgeschwindigkeiten der Zylindermoden den zu erwartenden Messeffekt zu berechnen.

Die Berechnung der Dispersionsdiagramme wurde in COMSOL Multiphysics 6.1 mit der von Landskron et al. [7] vorgestellten und auf Zylindermoden übertragenen Methode durchgeführt. Als Physik wurde im Bereich des Rohres Solid Mechanics und im Fluid Linearized Navier-Stokes, Frequency Domain gewählt. Im Vergleich zur Pressure Acoustics ist es so möglich sowohl eine Hintergrundströmung als auch die Aeroacoustic-Structure Boundary als Multiphysikkopplung zwischen Fluid und Rohr zu modellieren.

Die verwendete Rohrgeometrie und Materialeigenschaften sind in Tab. 1 aufgelistet.

Tab. 1: Rohrgeometrie und Materialeigenschaften des COMSOL Modells

Parameter	Wert
Rohr	
Innenradius	7,5 mm
Außenradius	9 mm
E-Modul	193,79 GPa
Poissonzahl	0,294
Dichte	7969 kg/m ³
Fluid	
Schallgeschwindigkeit	1481,4 m/s
Dichte	998,20 kg/m ³
Dynamische Viskosität	1,0093 µPa·s
Volumenviskosität	2,8161 µPa·s
Adiabatexponent	1,0066
Wärmekapazität bei konstantem Druck	4186,9 J/(kg·K)

Als Hintergrundströmung wurde ein über das Modified Power Law beschriebenes, turbulentes Strömungsprofil verwendet [8]. Die

Geschwindigkeitsverteilung im Rohr lässt sich durch Gl. (1) beschreiben. Die Variable m wird als konstant mit dem Wert 2 angenommen und n lässt sich mit Gl. (2) berechnen.

$$v(r) = v_{max} \cdot \left[1 - \left(\frac{r}{R_i}\right)^m\right]^{\frac{1}{n}} \quad (1)$$

$$n = 0,77 \cdot \ln(Re) - 3,47 \quad (2)$$

Als Reynoldszahl Re wurde 50.000 verwendet, wodurch sich eine maximale Geschwindigkeit v_{max} von 4,0640 m/s ergibt.

Durch die Wellenausbreitung auf einem rotationssymmetrischen System können sich höhere Moden – die sogenannten Azimutalschwingungen – ausbilden, die einer Resonanz entlang des Rohrfangs entsprechen. Welche Moden angeregt werden können und letztendlich angeregt werden hängt hauptsächlich von der Frequenz und dem verwendeten Sensoraufbau ab.

In Abb. 1 ist das berechnete Phasen- und in Abb. 2 das Gruppengeschwindigkeitsdiagramm des mit Wasser gefüllten Edelstahlrohres für die 0-te Azimutalschwingung dargestellt.

Der hierbei farblich visualisierte Messeffekt lässt sich durch die Gl. (3) berechnen.

$$\Delta t_{vph} = \frac{L}{v_{ph+}} - \frac{L}{v_{ph-}} \quad (3)$$

v_{ph+} und v_{ph-} steht hierbei für die Phasengeschwindigkeit der sich ausbreitenden Welle einmal in und einmal entgegen der Strömungsrichtung. In Kombination mit dem Sensorabstand L lässt sich die resultierende Zeitverschiebung, was dem Messeffekt entspricht, berechnen.

In den Abbildungen ist der Messeffekt auf eine Länge von 1 m und eine mittlere Fluidgeschwindigkeit von 1 m/s normiert.

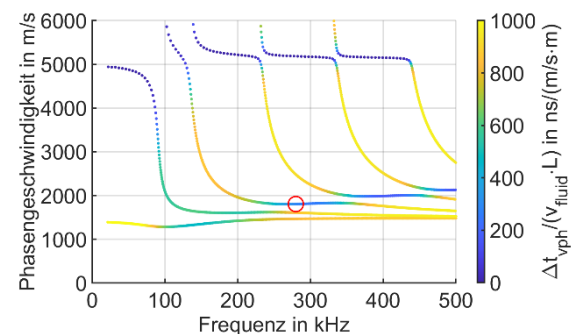


Abb. 1: Phasengeschwindigkeitsdiagramm mit farblich gekennzeichnetem Messeffekt eines mit Wasser gefüllten Rohres. 0-te Azimutalschwingung.

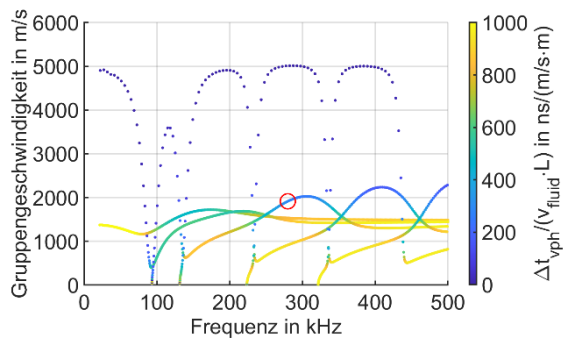


Abb. 2: Gruppengeschwindigkeitsdiagramm mit farblich gekennzeichnetem Messeffekt eines mit Wasser gefüllten Rohres. 0-te Azimutalschwingung.

In Abb. 3 ist der zuvor farblich dargestellt Messeffekt ergänzend als einzelner Graph dargestellt. Die für die Messungen verwendete Mode ist hierbei gestrichelt eingezeichnet und der verwendete Arbeitspunkt ist in den drei Abbildungen mit einem roten Kreis markiert.

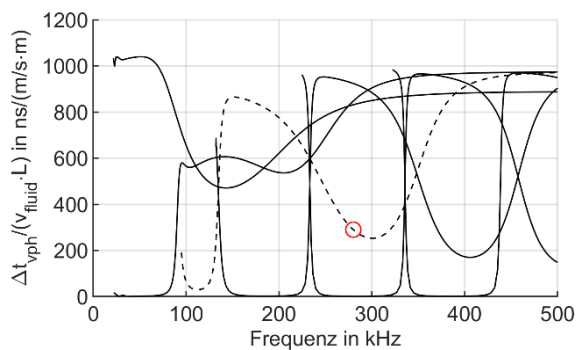


Abb. 3: Berechneter Messeffekt eines mit Wasser durchflossenen Rohres. 0-te Azimutalschwingung.

Obwohl nur die 0-te Azimutalschwingung betrachtet wird, bilden sich in dem simulierten Frequenzbereich eine Vielzahl von Moden aus. Neben den Grundmoden, die bei 0 Hz beginnen, kommen bei steigender Frequenz weitere Moden hinzu, die zu höheren Resonanzen im Fluidraum gehören.

Für die Auswahl eines geeigneten Arbeitspunktes müssen mehrere Faktoren berücksichtigt werden. Hierzu zählt die Separation von anderen Moden, die durch eine sich deutlich unterscheidende Gruppengeschwindigkeit – siehe Abb. 2 – erreicht werden kann. Zudem muss der Sensor in der Lage sein die angestrebte Mode bei dieser Frequenz anzuregen und die Mode sollte in diesem Bereich auch einen ausreichend hohen Messeffekt – siehe Abb. 3 – aufweisen. Letztendlich müssen diese Faktoren bei der Suche nach einem geeigneten Arbeitspunkt gegeneinander abgewogen werden. Bei dem

hier vorgestellten Sensor hat sich dabei der Arbeitspunkt von 280 kHz ergeben.

Sensoraufbau und Durchflussprüfstand

Für die Anregung der 0-ten Azimutalschwingung und damit einer rotationssymmetrischen Mode bietet sich ein zum Schwingungsverhalten der Mode passender Sensoraufbau an. Der in Abb. 4 und Abb. 5 dargestellte Sensor besteht aus 22 quaderförmigen Piezokeramiken. Die von PI Ceramic GmbH aus dem Material PIC255 hergestellten Piezokeramiken haben eine Kantenlänge von 2,54 mm. Zwischen den Piezokeramiken und dem Rohr befindet sich eine 100 µm dicke Indiumfolie. Diese verformt sich durch das feste Umwickeln und Festspannen der Sensoren mit Zahnseide und gewährleistet eine gute Schallübertragung zwischen den Sensoren und dem Rohr. Die elektrische Kontaktierung erfolgte über das Rohr sowie einem unter der Zahnseide eingewickelten Kupferdraht.

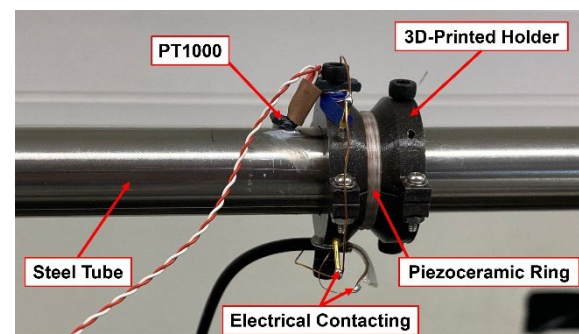


Abb. 4: Eine Seite des Ultraschalldurchflusssensors mit elektronischer Kontaktierung und PT1000 Temperatursensor.

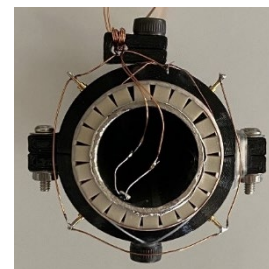


Abb. 5: Querschnitt des an einem Testrohr aufgebauten Ultraschalldurchflusssensors mit 22 Piezokeramiken

Zwei dieser Sensoren wurden an dem 1,6 m langen DN15 Rohr in einem Abstand von 200 mm positioniert. Die Vorlaufstrecke zwischen dem ersten Sensor und der nächstgelegenen Rohrkupplung betrug 1 m. Der so aufgebaute Sensor wurde, in dem in

Abb. 6 dargestellten Wasserprüfstand vermessen. Die Durchflussrate wurde über die Drehzahl einer Kreislumppe gesteuert. Ein Promass Q Coriolissensor der Firma Endress+Hauser diente als Referenzsensor. Die zeitgleiche Anregung erfolgte mit einem Keysight 33500B Funktionsgenerator und die Messdatenaufnahme mit einem LeCroy HDO6054B Oszilloskop. Das Umschalten zwischen Senden und Empfangen wurde durch einen Zweikanal Multiplexer realisiert. Als Anregesignal wurde ein 20 Vpp, 280 kHz, 10-fach Sinus Burst mit Hanning-Fensterung verwendet.

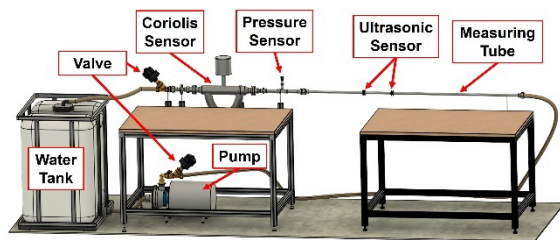


Abb. 6: Durchflussprüfstand

Ergebnisse und Diskussion

Die von beiden Sensoren parallel aufgenommenen Empfangssignale wurden miteinander kreuzkorreliert und die Signalverschiebung bestimmt.

In Abbildung 7 ist oben die gemessene Signalverschiebung inklusive einer linearen Regression durch ein Polynom 2. Ordnung dargestellt. Die Durchflussrate wurde hierbei über die Pumpendrehzahl rampenförmig erhöht.

Der lineare Term spiegelt die Empfindlichkeit der angeregten Mode gegenüber einer Fluidströmung wider. Bei einer Strömungsgeschwindigkeit von 1 m/s ergibt sich bei dem Sensorabstand von 200 mm eine differentiell ausgewertete Signallaufzeitverschiebung von 54,2 ns. Durch Multiplikation mit dem Faktor 5 – Normierung auf 1 m – lässt sich dieser Wert auf den berechneten Messeffekt in Abb. 3 umrechnen. Die sich so ergebende Signallaufzeitverschiebung von 270,8 ns liegt dicht an dem berechneten Wert von 290 ns.

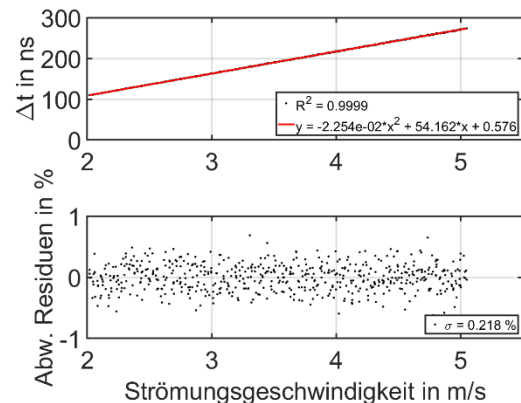


Abb. 7: Durchflussabhängige Signalverschiebung

Für die Abweichungen gibt es eine Vielzahl von möglichen Ursachen. Hierzu zählt die Verwendung von Literaturwerten für die Eigenschaften von Wasser sowie Herstellerangaben bezüglich der Materialeigenschaften und Abmessungen der verwendeten Edelstahlrohre. Zudem ist bekannt, dass der Piezo eine Art Übertragungsfunktion besitzt und es somit Abweichungen zwischen dem Anregesignal und der vom Sensor angeregten geführten akustischen Welle gibt.

Zudem hängt die bestimmte Signallaufzeitverschiebung von dem betrachteten Bereich des Empfangssignals sowie der Auswertemethode ab. Die verwendete Kreuzkorrelation ist wie z. B. auch eine Nullstellenauswertung überwiegend sensitiv auf die Phasengeschwindigkeitsänderung, erfasst jedoch auch zu geringen Teilen die Gruppengeschwindigkeitsänderung.

Wohingegen das Auswerten des Maximums der Einhüllenden empfindlich auf die Gruppengeschwindigkeitsänderung ist. In Abhängigkeit der Auswertemethode ist somit der Messeffekt aus der Änderung der Phasen- bzw. Gruppengeschwindigkeit zu berechnen.

Zusammenfassend lässt sich sagen, dass die Berechnung von Dispersionsdiagrammen mit Hintergrundströmung einen wichtigen Beitrag bei der Entwicklung von Ultraschalldurchflusssensoren basierend auf Zylindermoden liefert und eine gezielte Auslegung und Optimierung ermöglicht. Dies ist besonders durch die hohe Anzahl von Moden sowie dem Auftreten von höheren Azimutalschwingungen, die alle unterschiedlich auf eine Fluidströmung reagieren, hilfreich. Ohne die Berechnung des Messeffekt gäbe es sonst eine quasi nicht zu bewältigende Anzahl an experimentell auf ihre Eignung für die Durchflussmessung zu überprüfenden Moden. Nachdem der Grundstein gelegt wurde, gilt es

im nächsten Schritt das Potential hinter der Verwendung von Zylindermoden für die Durchflussmessung genauer zu untersuchen.

Förderung

Diese Forschungsarbeit wurde vom Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Rahmen des Förderprogramms „Forschung an Fachhochschulen“ unter dem Projekttitel „HydrAmess“ und der Fördernummer 13FH094KX1 gefördert.

Literaturnachweis

- [1] R. C. Baker, *Flow Measurement Handbook: Industrial Designs, Operating Principles, Performance, and Applications*, 2. Aufl. Cambridge University Press, 2015. doi: [10.1017/cbo9781107054141](https://doi.org/10.1017/cbo9781107054141).
- [2] D. A. Kiefer, A. Benkert, und S. J. Rupitsch, „Transit Time of Lamb Wave-Based Ultrasonic Flow Meters and the Effect of Temperature“, *IEEE Trans Ultrason Ferroelectr Freq Control*, Bd. 69, Nr. 10, S. 2975–2983, Sep. 2022, doi: [10.1109/TUFFC.2022.3201106](https://doi.org/10.1109/TUFFC.2022.3201106).
- [3] J. L. Rose, *Ultrasonic Guided Waves in Solid Media*. Cambridge: Cambridge University Press, 2014. doi: [10.1017/CBO9781107273610](https://doi.org/10.1017/CBO9781107273610).
- [4] A. Backer, P. Arneth, P. Linke, und K. S. Drese, „NON-INVASIVE ULTRASONIC FLOW MEASUREMENT METHOD FOR THE DETECTION OF FLOW RATES IN THE MICROLITER RANGE“, in *The 5th Conference on MicroFluidic Handling Systems (MFHS 2024)*, Munich, Germany, Feb. 2024, S. 29–32. [Online]. Verfügbar unter: <https://www.utwente.nl/en/eemcs/mfhs2024/mfhs2024-proceedings.pdf>
- [5] A. Backer und K. S. Drese, „Ultrasonic Flow Measurement Using Guided Acoustic Waves: Application of Cylindrical Modes in Pipe Systems“, in *Proceedings, 2025 ICU PADERBORN, 9th International Congress on Ultrasonics – ICU*, Paderborn: AMA Service GmbH, Sep. 2025, S. 85–88. doi: [10.5162/Ultrasonic2025/A12-b3](https://doi.org/10.5162/Ultrasonic2025/A12-b3).
- [6] M. Takamoto, H. Ishikawa, K. Shimizu, H. Monji, und G. Matsui, „New measurement method for very low liquid flow rates using ultrasound“, *Flow Measurement and Instrumentation*, Bd. 12, Nr. 4, S. 267–273, Aug. 2001, doi: [10.1016/S0955-5986\(01\)00023-1](https://doi.org/10.1016/S0955-5986(01)00023-1).
- [7] [1] J. Landskron, A. Backer, C. R. Wolf, G. Fischerauer, und K. S. Drese, „FEM-Based Modeling of Guided Acoustic Waves on Free and Fluid-Loaded Plates“, *Applied Sciences*, Bd. 15, Nr. 16, S. 9116, Aug. 2025, doi: [10.3390/app15169116](https://doi.org/10.3390/app15169116).
- [8] A. Salama, „Velocity Profile Representation for Fully Developed Turbulent Flows in Pipes: A Modified Power Law“, *Fluids*, Bd. 6, Nr. 10, S. 369, Okt. 2021, doi: [10.3390/fluids6100369](https://doi.org/10.3390/fluids6100369).

Einzel- und Doppelmikrofonmessungen in einem Helmholtz-Resonator zur Bestimmung der CO₂-Konzentration in trockener Stickstoffatmosphäre

*Sebastian Kindorf^{1,2}, Christian Weber^{1,2}, Mike Benkendorf¹, Benedikt Bierer^{1,2}, Jürgen Wöllenstein^{1,2},
Franz Jost¹, Matthias Eberl³*

¹Fraunhofer-Institut für Physikalische Messtechnik IPM, Georges-Köhler-Allee 301, 79110 Freiburg

²Institut für Mikrosystemtechnik – IMTEK, Professur für Gassensoren, Universität Freiburg,
Georges-Köhler-Allee 102, 79110 Freiburg

³Infineon AG, Am Campeon 1-15, 85579 Neubiberg, Germany

Zusammenfassung

Akustische Resonanzverfahren ermöglichen eine kostengünstige und schnell ansprechende CO₂-Detektion auf der Basis der Schallgeschwindigkeitsänderung in Gasgemischen. Motiviert durch die Anforderungen an die CO₂-Leckagedetektion in Maschinenräumen, die mit R744-Kälteanlagen ausgestattet sind und in denen Dichtungsversagen zu einer raschen Kohlendioxidanreicherung führen kann, wurden zwei auf MEMS-Dual-Backplate-Mikrofonen basierende Helmholtz-Resonator-Sensorkonzepte untersucht: ein Einzelmikrofon- sowie ein Doppelmikrofonsystem. Die innere Geometrie der MEMS-Mikrofone bildet dabei jeweils einen Helmholtz-Resonator, dessen Resonanzfrequenz direkt von der Schallgeschwindigkeit des eingeschlossenen Messgases abhängt. Beide Systeme wurden als Aufsteckplatinen für ein PSoC5-Entwicklungsboard implementiert und seriell in einem gemeinsamen Strömungsweg in trockener Stickstoffatmosphäre betrieben. Die CO₂-Konzentration wurde schrittweise von 0 bis 50.000 ppm variiert. Beide Konzepte zeigen eine deutliche und reproduzierbare Signaländerung, wobei das Doppelmikrofonsystem eine höhere Signalstabilität aufweist. Die erzielte Nachweisgrenze beträgt $3\sigma = 2.732$ ppm (Einzelmikrofon) bzw. $3\sigma = 1.040$ ppm (Doppelmikrofon).

Keywords: Gassensorik, Ultraschall, Helmholtz-Resonator, Dual-Backplate Mikrofon, Leckdetektion

Motivation

Die Erkennung von CO₂-Leckagen in Industrieanlagen ist für die Arbeitssicherheit, den Umweltschutz und die Anlagenverfügbarkeit unerlässlich. In Maschinenräumen mit R744-Kälteanlagen kann Dichtungsversagen zu einer raschen Kohlenstoffdioxidanreicherung führen. Benötigt wird ein kostengünstiges Sensorsystem mit schnellen Ansprechzeiten, hoher Zuverlässigkeit und der Fähigkeit, auch unter wechselnden Umgebungsbedingungen eine robuste Erfassung zu gewährleisten [1]. Akustische Verfahren, die auf resonanzbasierten Messprinzipien beruhen, adressieren diese Anforderungen. Die Schallgeschwindigkeit in Gasgemischen ist eine Funktion der Zusammensetzung sowie der Zustandsgrößen Temperatur, Druck und Feuchte. Die daraus resultierenden Änderungen verschieben die Resonanzfrequenz messbar und lassen Rückschlüsse auf die CO₂-Konzentration zu [2].

Messprinzip und Ansatz

Mikrofonbasierte Resonatorsysteme zur CO₂-Messung nutzen den physikalischen

Zusammenhang zwischen der Zusammensetzung eines Gasgemisches und der darin herrschenden Schallgeschwindigkeit. Die Schallgeschwindigkeit c in einem idealen Gas ist durch die thermodynamische Zustandsgleichung in Gl. (1) definiert:

$$c = \sqrt{\frac{\gamma RT}{M}} \quad (1)$$

wobei γ der adiabatische Exponent, R die universelle Gaskonstante, T die absolute Temperatur und M die mittlere molare Masse des Gases ist [3]. Da CO₂ eine deutlich höhere molare Masse als die anderen Bestandteile von Luft besitzt und zudem einen kleineren adiabatischen Exponenten aufweist, nimmt die Schallgeschwindigkeit mit zunehmendem CO₂-Anteil im Gasgemisch messbar ab. Die molare Masse von CO₂ beträgt 44,01 g/mol gegenüber 28,97 g/mol für trockene Luft, während der adiabatische Exponent mit 1,289 ebenfalls unter dem von Luft mit 1,400 liegt [4,5].

Zur Messung der Schallgeschwindigkeit wird ein akustischer Resonator eingesetzt. Ein Helmholtz-Resonator besteht aus einem

geschlossenen Gasvolumen V , das über eine Öffnung der Querschnittsfläche A und der effektiven Länge L_{eff} mit der Umgebung verbunden ist. Das eingeschlossene Gas wirkt dabei als pneumatische Feder, während das im Hals befindliche Gasvolumen als schwingende träge Masse fungiert. Basierend auf dieser Masse-Feder-Analogie ergibt sich die Eigenfrequenz des Resonators nach Gl. (2) [6]

$$f_0 = \frac{c}{2\pi} \sqrt{\frac{A}{V \cdot L_{\text{eff}}}}. \quad (2)$$

Die effektive Halslänge $L_{\text{eff}} = L + \Delta L$ enthält dabei eine akustische Endkorrektur ΔL , welche die mitschwingende Luftmasse außerhalb des Halses berücksichtigt. Der entscheidende Aspekt für die Sensoranwendung ist die direkte Proportionalität zwischen der Eigenfrequenz und der Schallgeschwindigkeit im eingeschlossenen Gas. Eine Änderung der CO_2 -Konzentration verändert c und verschiebt damit f_0 messbar [6]. Da die Messgröße die Resonanzfrequenz f_0 und nicht die Signalamplitude ist, weist das Verfahren eine inhärente Robustheit gegenüber amplitudenabhängigen Störeinflüssen wie Schwankungen der Mikrofonempfindlichkeit oder des Schalldruckpegels auf.

Dieses Prinzip wird mithilfe von MEMS-Mikrofonen (Infineon Technologies AG) umgesetzt, deren innere Geometrie einen solchen Resonator bildet. Ein MEMS-Mikrofon mit Dual-Backplate-Architektur besteht aus einer auslenkbaren Membran, die zwischen zwei perforierten Gegenelektroden (Backplates) eingespannt ist, wobei eine Backplate zur akustischen Anregung und die andere zur kapazitiven Signalauswertung dient. Der hinter der Membran eingeschlossene Raum bildet zusammen mit dem akustischen Einlass eine Kavität, der als Helmholtz-Resonator fungiert. Da sich das Messgas direkt in diesem Resonanzraum befindet, reagiert das System unmittelbar auf jede Änderung der lokalen Schallgeschwindigkeit, ohne dass ein separater externer Resonator erforderlich ist [7].

Es lassen sich zwei konstruktive Anordnungen unterscheiden. Die beiden Systeme sind in Abb. 1 dargestellt. Im Einzelmikrofonsystem übernimmt eine Backplate des Dual-Backplate-Mikrofons die akustische Anregung, während die andere Backplate das Signal kapazitiv ausliest. Das Rückvolumen des Chips bildet zusammen mit dem akustischen Einlass einen Helmholtz-Resonator, dessen Resonanzfrequenz direkt von der Schallgeschwindigkeit im eingeschlossenen Messgas abhängt. Im Doppelmikrofonsystem sind zwei Dual-Backplate-Mikrofone gegenüberliegend angeordnet, deren akustisch gekoppelte Kavitäten einen gemeinsamen Resonator bilden. Die in beiden Kammern erzeugten Druckwellen schwingen

gegenphasig, und das Ausgangsmikrofon erfasst die Resonanzfrequenz des gekoppelten Systems [8].

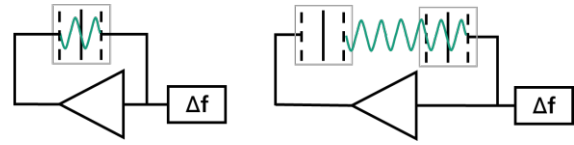


Abb. 1: Resonanzbasierte CO_2 -Messung mit Einzelmikrofon (links) und Doppelmikrofon (rechts). Die konzentrationsabhängige Frequenzverschiebung Δf dient als Messgröße.

Da CO_2 eine höhere molare Masse als N_2 besitzt, sinkt c mit zunehmendem CO_2 -Anteil, was zu einer messbaren Verschiebung von f_0 führt. Aufgrund der Wurzelabhängigkeit $c \propto 1/\sqrt{M_{\text{mix}}}$ ist dieser Effekt bei höheren CO_2 -Konzentrationen weniger ausgeprägt, was sich theoretisch in der abnehmenden Empfindlichkeit widerspiegelt.

Relevante Störgrößen sind insbesondere Temperaturschwankungen, da die Schallgeschwindigkeit gemäß Gl. (1) proportional zu $\sqrt{T}$ ist und sich eine Temperaturänderung daher direkt als scheinbare Konzentrationsänderung auswirkt. Ebenso beeinflusst die Luftfeuchte die Schallgeschwindigkeit, da Wasserdampf eine geringere molare Masse (18,02 g/mol) als Stickstoff (28,01 g/mol) als auch Kohlenstoffdioxid (44,01 g/mol) besitzt und somit c erhöht, was ebenfalls als Konzentrationsänderung interpretiert wird. Druckschwankungen beeinflussen die mechanische Vorspannung der MEMS-Membran und verschieben dadurch deren Resonanzfrequenz, was zu einem systematischen Mess-Offset führt. Bei den Einzelmikrofon- und Doppelmikrofonsystemen für die resonante Ultraschallmessung handelt es sich um rein akustische Messmethoden, die nicht selektiv gegenüber anderen Gasen mit anderer molarer Masse sind.

Messumgebung und Auswertung

Zur experimentellen Validierung wurden zwei Sensorplatinen entwickelt und aufgebaut, jeweils eine für das Einzelmikrofon- und das Doppelmikrofonsystem. Beide Platinen umfassen jeweils die vollständige Signalkette, von der akustischen Anregung bis zur digitalen Frequenzausgabe, und sind als Aufsteckmodule für ein PSoC5-Entwicklungsboard ausgeführt. Im Messaufbau werden die Sensorkonzepte seriell in einem gemeinsamen Strömungsweg betrieben, sodass das Einzel- und das Doppelmikrofonsystem demselben N_2/CO_2 -Gemisch ausgesetzt sind. Die Messungen wurden in einer definierten trockenen Stickstoffatmosphäre

durchgeführt. Dazu wurde dem N_2 -Trägergasstrom stufenweise CO_2 beigemischt, wobei die CO_2 -Konzentration in Schritten von 10.000 ppm (1 %) von 0 bis 50.000 ppm (5 %) erhöht wurde. Der Gesamtvolumenstrom wurde durch entsprechende Anpassung des N_2 -Anteils konstant gehalten, sodass ausschließlich die Gaszusammensetzung variierte. Die Konzentrationsrampe von 0 bis 50.000 ppm wurde dreimal nacheinander durchlaufen, wobei zwischen den Zyklen jeweils auf 0 ppm zurückgekehrt wurde, um die Reproduzierbarkeit der Messsignale zu bewerten und einen möglichen Offset-Drift zu erfassen. Die Laborumgebung wies dabei eine stabile Temperatur sowie eine vernachlässigbare Restfeuchte auf, sodass die im vorherigen Kapitel beschriebenen Störgrößen Temperatur und Luftfeuchte für diese Messreihe als konstant betrachtet werden können. Temperatur, relative Feuchte und Druck werden dennoch kontinuierlich mit einem BME280 (Bosch Sensortec) gemessen und aufgezeichnet, sie werden derzeit jedoch nicht zur aktiven Kompensation der Messwerte verwendet. Änderungen der CO_2 -Konzentration beeinflussen die Schallgeschwindigkeit und damit die Resonanzfrequenz. Aus der Frequenzverschiebung Δf der Einzel- bzw. Doppelmikrofon-Resonanzen wird die CO_2 -Konzentration abgeleitet. Der in Abb. 2 dargestellte Messaufbau umfasst jeweils eine Sensorplatine für die Einzel- und Doppelmikrofonmessung, die auf ein PSoC5-Entwicklungsboard aufgesteckt sind. Die Platine übernimmt Anregung, Rückkopplung sowie die analoge und digitale Signalverarbeitung. Resonanzfrequenzen, Temperatur, relative Feuchte und Druck werden anschließend an einen PC zur Auswertung übertragen.



Abb. 2: PSoC 5-Entwicklungsboard mit aufgesetzter Sensorplatine. Links befindet sich das Einzelmikrofonsystem, rechts das Doppelmikrofonsystem. Ein 3D-gedruckter Gehäusekopf mit Gasanschlüssen dient zur Integration der Sensoren in den Messplatz. Die Dual-Backplate-Mikrofone sind auf der Sensorplatine montiert.

Diskussion

CO_2 -Konzentrationsmessungen von 0 bis 50000 ppm zeigen in beiden Systemen eine klare und reproduzierbare Signaländerung, was die grundlegende Funktionalität beider Sensorkonzepte bestätigt. Die Messergebnisse sind in Abb. 3 dargestellt. Zur besseren Vergleichbarkeit wurden die Signale auf ihre jeweilige Resonanzfrequenz bei 0 ppm CO_2 normiert, sodass die dargestellten Werte die relative Frequenzverschiebung repräsentieren.

Das Doppelmikrofonsystem weist sowohl eine höhere Signalstabilität als auch geringere Driftphänomene auf. Kurzzeitige Druckstöße beim Schalten der Massendurchflussregler sind im Doppelmikrofonsignal als deutliche Transienten erkennbar.

Eine abnehmende Empfindlichkeit bei höheren CO_2 -Anteilen wäre gemäß Gl. (1) zu erwarten, ist jedoch im vorliegenden Messbereich von 0-50.000 ppm nicht erkennbar. Die Nachweisgrenze beträgt für das Einzelmikrofonsystem $3\sigma = 2.732$ ppm sowie für das Doppelmikrofonsystem $3\sigma = 1.040$ ppm. Über die Messdauer von 333 min ist ein Offset-Drift von 5.490 ppm feststellbar, der auf eine Temperaturänderung zurückzuführen ist. Insgesamt bestätigen die Ergebnisse die prinzipielle Eignung beider Ansätze für die CO_2 -Leckdetektion, wobei das Doppelmikrofonsystem hinsichtlich Signalqualität und dynamischem Verhalten Vorteile aufweist.

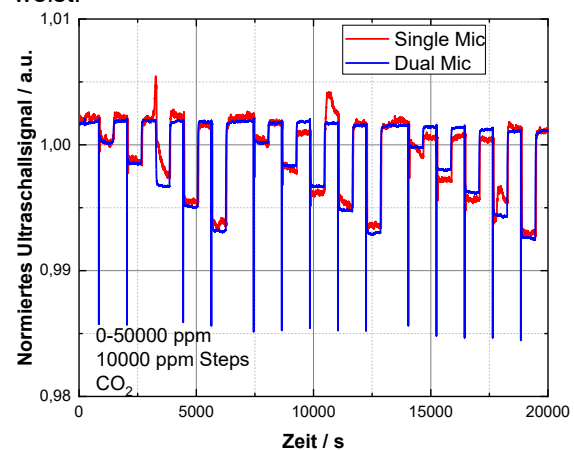


Abb. 3: Normierte Mikrofonensignale bei CO_2 -Konzentrationen von 0 bis 50.000 ppm in Schritten von 10.000 ppm in trockener Stickstoffatmosphäre. Rot: Einzelmikrofon ($3\sigma = 2.732$ ppm), Blau: Doppelmikrofon ($3\sigma = 1.040$ ppm). Das Doppelmikrofonsystem weist eine höhere Signalstabilität auf. Ein temperaturbedingter Offsetdrift von 5.490 ppm ist über die Messdauer von 333 min erkennbar.

Fazit & Ausblick

Die vorliegende Arbeit zeigt, dass als Helmholtz-Resonatoren aufgebaute MEMS-Mikrofone im Einzel- und Doppelmikrofonsystem prinzipiell zur akustischen CO₂-Leckdetektion in R744-Kälteanlagen geeignet sind. Die untersuchten Konzentrationsstufen von 10.000 ppm (1 %) sind mit beiden Aufbauten klar auflösbar, wobei das Doppelmikrofonsystem hinsichtlich Signalstabilität und geringeren Driftphänomenen Vorteile gegenüber dem Einzelmikrofonsystem zeigt.

Zukünftige Arbeiten adressieren zunächst eine Erweiterung des Messbereichs auf niedrigere CO₂-Konzentrationen mit feinerer Stufung, um die tatsächliche Auflösungsgrenze des Systems besser zu charakterisieren und die Nachweisgrenze von derzeit $3\sigma = 2.732$ ppm des Einzelmikrofon und $3\sigma = 1.040$ ppm des Doppelmikrofon zu prüfen. Darüber hinaus soll eine aktive Kompensation der erfassten Umgebungsgrößen Temperatur, relative Feuchte und Druck implementiert werden, um den Sensor auch unter variablen Umgebungsbedingungen zuverlässig betreiben zu können. Eine erweiterte Elektronik ist geplant, die auch die Auswertung von Signalamplitude und Dämpfung ermöglicht. Zusätzlich soll sie flexible Anregungssignale wie Sinus- und Pulsanregung unterstützen, wodurch ergänzend akustische Laufzeitmessungen (Time-of-Flight) zur Schallgeschwindigkeitsbestimmung realisiert werden können.

Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Energie (BMWE) für die Förderung des Projekts EiSBaEr im Rahmen des 8. Energieforschungsprogramms (Förderkennzeichen 03EN5048D).

Literaturnachweis

- [1] CEN, „EN 378-1 bis EN 378-4:2016 – Refrigerating systems and heat pumps – Safety and environmental requirements,” European Committee for Standardization, Brussels, Belgium, 2016.
- [2] Infineon Technologies AG, „Gas sensor containing an ultrasonic resonator and use for the detection of gases,” U.S. Patent US12326422B2, Jun. 2025.
- [3] A. D. Pierce, *Acoustics: An Introduction to Its Physical Principles and Applications*, Springer, New York, NY, USA, 2019.
- [4] National Institute of Standards and Technology, „Carbon Dioxide – Gas Phase Thermochemistry Data,” NIST Chemistry WebBook, SRD 69. [Online]. Verfügbar: <https://webbook.nist.gov/cgi/cbook.cgi?ID=124-38-9> [Zugriff am: Apr. 2025].
- [5] D. R. Lide (Hrsg.), *CRC Handbook of Chemistry and Physics*, 85. Aufl., CRC Press, Boca Raton, FL, USA, 2004.
- [6] L. E. Kinsler, A. R. Frey, A. B. Coppers, und J. V. Sanders, *Fundamentals of Acoustics*, 4. Aufl., Wiley, New York, NY, USA, 2000.
- [7] S. Anzinger, C. Bretthauer, H. S. Wasisto, und A. Dehé, „Dual-Backplate CMUTs With Wide Bandwidth and Low Driving Voltage for Airborne Applications,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, Bd. 70, Nr. 10, S. 1286–1294, Okt. 2023.
- [8] J. Rouxel, J.-G. Coutard, S. Gidon, O. Lartigue, S. Nicoletti, B. Parvitte, R. Vallon, V. Zéninari, und A. Glière, „Miniaturized Differential Helmholtz Resonators for Photoacoustic Trace Gas Detection,” *Sens. Actuators B: Chem.*, Bd. 236, S. 1104–1110, Nov. 2016.

Triboakustischer Pin: ein neuer Sensor für die *in-situ* Schmierfilmüberwachung in tribologischen Kontakten

Schirru Michele¹, Tatzgern Fabio¹, Varga Markus¹

¹AC2T research GmbH, Viktor-Kaplan-Straße 2/C 2700 Wiener Neustadt, Österreich

Michele.schirru@ac2t.at, Fabio.tatzgern@ac2t.at, Markus.varga@ac2t.at

Zusammenfassung

Diese Arbeit präsentiert einen neuartigen Ultraschallsensor zur *In-situ*-Messung der Schmierfilmdicke in tribologischen Kontakten, mit besonderem Fokus auf Anwendungen in Tribometern und Prüfständen. Der Ansatz basiert auf der Transformation eines konventionellen Pins in ein aktives Sensorelement durch Integration miniaturisierter piezoelektrischer Wandler. Dadurch wird eine Echtzeitüberwachung von Schmierstoffeigenschaften wie Filmdicke, Viskosität und Schmierregime mit hoher zeitlicher und räumlicher Auflösung ermöglicht. Validierungsexperimente zeigen eine hohe Übereinstimmung zwischen Ultraschallmessungen, optischen Referenzmethoden und theoretischen Modellen. Darüber hinaus erlaubt der Sensor die Beobachtung der Bildung und Degradation von Tribofilmen während des Betriebs. Der präsentierte triboakustische Pin lässt sich einfach in bestehende Messaufbauten integrieren und liefert wertvolle zusätzliche Einblicke in das Verhalten von Schmierstoffen unter realitätsnahen Betriebsbedingungen.

Keywords: Ultraschall, Sensorik, Schmierstoffprüfung, Akustik, Piezoelektrik.

Einleitung

Moderne mechanische Systeme wie Fahrzeugantriebe, Lager und industrielle Maschinen arbeiten unter zunehmend anspruchsvollen Bedingungen, darunter hohe Lasten, Temperaturen und Geschwindigkeiten. Unter diesen Bedingungen werden Leistung und Zuverlässigkeit der Komponenten maßgeblich durch das Verhalten der Schmierstoffe in stark eingegengten tribologischen Kontakten bestimmt. Häufig reduzieren sich die Schmierfilme auf Nanometerskalen, wo ihre Eigenschaften erheblich vom makroskopischen Verhalten abweichen.

Trotz der großen Bedeutung dieser Effekte stellt die direkte Messung von Schmierfilmeigenschaften unter realistischen Betriebsbedingungen eine erhebliche Herausforderung dar – sowohl in der Forschung als auch in industriellen Anwendungen.

Aktuelle experimentelle Ansätze, typischerweise basierend auf konventionellen Tribometern, ermöglichen zwar die Reproduktion von Kontaktbedingungen, liefern jedoch nur begrenzte Informationen über die *In-situ*-Entwicklung von Schmierfilmen. Ein Tribometer ist ein Laborgerät zur Untersuchung von Reibung, Verschleiß und Schmierung zwischen zwei Kontaktflächen. In

seiner einfachsten Form wird eine definierte Probengeometrie, beispielsweise eine Kugel oder ein Pin, unter kontrollierten Last-, Bewegungs- und Umgebungsbedingungen gegen eine Gegenfläche gedrückt. Durch Variation von Parametern wie Geschwindigkeit, Temperatur und Kontaktkraft lassen sich realistische Betriebsbedingungen technischer Komponenten simulieren.

In der Industrie werden Tribometer eingesetzt, um Schmierstoffe und Materialien zu bewerten und zu vergleichen, Bauteildesigns zu optimieren sowie Leistungsfähigkeit und Lebensdauer vor dem Einsatz im realen System sicherzustellen. Diese Ansätze basieren jedoch häufig auf indirekten Messungen oder Nachuntersuchungen und erfassen keine transienten Effekte oder Echtzeitänderungen der Schmierregime.

Zur Überwindung dieser Einschränkungen wurde ein neuartiger Ultraschallsensor entwickelt, der eine Echtzeit- und *In-situ*-Charakterisierung von Schmierfilmen direkt im tribologischen Kontakt ermöglicht. Anstatt vollständig neue Prüfgeräte zu entwickeln, folgt der Ansatz dem aktuellen Trend in der Tribologie, bestehende, etablierte Systeme um zusätzliche Sensorik zu erweitern. Dies entspricht der industriellen Praxis, in der standardisierte Tribometer und Prüfverfahren

über viele Jahre hinweg entwickelt und validiert wurden.

Durch die Integration ultraschallbasierter Sensorelemente in eine standardisierte Pin-Geometrie ermöglicht der vorgestellte Ansatz eine einfache Implementierung in bestehende Tribometer, ohne etablierte Prüfmethoden zu verändern. Dadurch können konventionelle Messungen durch direkte Echtzeitinformationen über das Verhalten von Schmierfilmen ergänzt werden. Die Entwicklung stellt somit ein praxisnahes und skalierbares Werkzeug dar, das das Verständnis von Schmiermechanismen verbessert und zur Optimierung von Schmierstoffen, Komponenten sowie zur Steigerung von Effizienz, Zuverlässigkeit und Lebensdauer mechanischer Systeme beiträgt.

Materialien: Der Triboakustische Pin

Das zentrale Element dieser Arbeit ist ein modifizierter tribologischer Pin, der als Triboakustischer Pin (TaP) bezeichnet wird und ultraschallbasierte Sensoreigenschaften in eine standardisierte Geometrie integriert.

Abbildung 1 zeigt schematisch die Anpassung eines konventionellen kugelförmigen Pins zur Aufnahme von zwei miniaturisierten piezoelektrischen Wandlern, die im Frequenzbereich von 5 bis 20 MHz arbeiten. Diese Wandler sind unter einem definierten Winkel im Pin-Körper eingebettet und ermöglichen die Aussendung und den Empfang von Ultraschallwellen durch die Kontaktzone.

Ein Wandler fungiert als Sender und erzeugt Ultraschallwellen, die sich durch den Pin in den geschmierten Kontakt ausbreiten. Der zweite Wandler dient als Empfänger und erfasst die an der Grenzfläche reflektierten Signale. Aufgrund der schrägen Anordnung werden sowohl longitudinale als auch transversale Wellenanteile detektiert, wodurch verschiedene Schmierstoffeigenschaften, einschließlich Filmdicke und scherabhängigem Verhalten, charakterisiert werden können.

Der Pin ist in einer speziellen Halterung montiert, die über federbelastete elektrische Kontakte verfügt. Diese gewährleisten eine zuverlässige Signalübertragung und ermöglichen gleichzeitig einen schnellen Austausch des Pins zwischen verschiedenen Experimenten. Die äußere Geometrie und die mechanischen Eigenschaften standardisierter Pins bleiben erhalten, sodass eine Integration

in bestehende Tribometer ohne Anpassung des Prüfaufbaus möglich ist.

Abbildung 2 zeigt die Integration des Pins in ein Tribometer. Der Triboakustische Pin kann in unterschiedlichen Tribometerkonfigurationen eingesetzt werden, darunter Systeme mit kontinuierlicher Gleitbewegung sowie oszillierende Systeme mit hin- und hergehender Bewegung. Diese Flexibilität gewährleistet die Kompatibilität mit etablierten Prüfverfahren und ermöglicht den Einsatz in einer Vielzahl tribologischer Untersuchungen.

Abbildung 3 zeigt die Messkette. Der Triboakustische Pin ist mit dem Pinhalter verbunden, der die elektrische Verbindung zu einer Sende-/Empfangeinheit herstellt. Für die in dieser Arbeit durchgeführten Experimente wurde ein Ultraschall-Puls-Empfangssystem (UPR) vom Typ Picoscope 5000 verwendet. Ein Vorverstärker verstärkt das Anregungssignal auf bis zu 20 Vpp. Eine Python-basierte grafische Benutzeroberfläche steuert die Messkette von einem Steuerrechner aus, insbesondere die Pulsparameter des Ultraschallbursts (Frequenz, Zyklenzahl) sowie die Erfassung des reflektierten Signals.

Die Einheit ermöglicht eine schnelle Datenerfassung mit bis zu 100 kS/s innerhalb eines hochauflösenden Signalpakets alle 15 Sekunden. Dadurch können sowohl zeitliche Trends der Filmdickenentwicklung als auch orts aufgelöste Messungen entlang der Oszillationsbewegung durchgeführt werden.

Datenanalyse, Validierung und Nachverarbeitung

Das Signal in Abbildung 4 ist durch eine Reihe von Reflexionen gekennzeichnet, die auf eine Modenkonzentration der einfallenden longitudinalen Welle in sowohl longitudinale als auch transversale Komponenten zurückzuführen sind [1]. Der Fokus dieser Arbeit liegt auf der hervorgehobenen ersten longitudinalen Reflexion, die mit der Dicke der Zwischenschicht zwischen dem TaP und dem Gegenkörper in Zusammenhang steht. Es ist jedoch ebenfalls möglich, durch eine entsprechende Analyse der transversalen Reflexionskomponente die Schersteifigkeit und damit die viskoelastischen Eigenschaften der Zwischenschicht zu bestimmen [2].

Die Dicke der Zwischenschicht wird wie folgt berechnet: Zunächst wird der Reflexionskoeffizient bestimmt. Dieser entspricht dem Verhältnis der Amplitude der

während des Versuchs gemessenen longitudinalen Reflexion zu einem Referenzwert. In der Regel handelt es sich dabei um eine Messung derselben Amplitude unter Referenzbedingungen (z. B. in Luft) [3]. Für den TaP ist das Referenzverfahren durch eine Patentanmeldung geschützt und umfasst mehrere Schritte, darunter die Standardmessung in Luft sowie definierte Belastungsschritte des Pins gegen den Gegenkörper zur Bestimmung der Nullschichtdickenantwort.

$$R = \frac{A_m}{A_r}$$

Gleichung 1

Wobei A_m die während des dynamischen Versuchs gemessene reflektierte Amplitude und A_r die gemäß Referenzverfahren bestimmte Referenzamplitude ist. Anschließend wird der Reflexionskoeffizient mit der Steifigkeit der Zwischenschicht korreliert [4]. Die Steifigkeit kann direkt aus dem Reflexionskoeffizienten berechnet werden:

$$K = \frac{Z_s}{2\omega} \left(\frac{1}{R} - 1 \right)$$

Gleichung 2

Wobei K die Schichtsteifigkeit, Z_s die akustische Impedanz des Triboakustischen Pins und ω die Kreisfrequenz der Ultraschallwelle ist.

Im Allgemeinen ist die absolute Steifigkeit der Schicht direkt mit ihren mechanischen Eigenschaften verknüpft [1]. Aufgrund der kugelförmigen Geometrie und der kleinen Kontaktfläche hat sich jedoch gezeigt, dass zur Maximierung der Sensitivität des Algorithmus gegenüber sehr dünnen Schichten die Bestimmung der Schichtdicke aus der Differenz zwischen der Steifigkeit im Nullschichtzustand und der während dynamischer Bedingungen gemessenen Steifigkeit geeigneter ist. Diese Steifigkeitsänderung wird anschließend mit der Gesamtabhebung des TaP vom Gegenkörper korreliert.

Der Algorithmus zur Bestimmung der Schichtdicke aus der Schichtsteifigkeit lautet somit:

$$h = ST \left(\frac{1}{K_0} - \frac{1}{K} \right) \cos(\theta) B$$

Gleichung 3

Wobei h die Schichtdicke, K_0 die Referenzsteifigkeit, K die während der dynamischen Messung bestimmte Steifigkeit, θ der Einfallswinkel der Ultraschallwelle und, und ST ein Koeffizient ist, der von der Größe der Kontaktfläche und der Frequenz des Wandlers abhängt. Dieser Koeffizient stellt eine fundamentale Größe dar, die im Zuge der Sensorentwicklung identifiziert wurde. Es ist zu beachten, dass ein solcher Faktor in der klassischen Ultraschall-Reflektometrie zur Schichtdickenmessung keine Rolle spielt. Es wurde jedoch festgestellt, dass die kugelförmige Grenzfläche einen Fokussierungseffekt hervorruft, der proportional zum Faktor ST ist. Diese sphärische Geometrie ermöglicht eine erhöhte Sensitivität im Vergleich zu anderen reflektometrischen Verfahren und erlaubt die Messung von Schichtdicken bis hinunter zu etwa 20 nm.

Der Algorithmus wurde durch Messungen an einem MTM-Tribometer in einem sogenannten Pin-on-Disc-Aufbau validiert, wie schematisch in Abbildung 5 dargestellt. In dieser Konfiguration ist der TaP fixiert, während sich der Gegenkörper dreht. Als Gegenkörper wurde eine Glasscheibe gewählt, sodass die Schmierfilmdicke gleichzeitig mittels interferometrischer Messung mit einer Genauigkeit von ± 1 nm bestimmt werden konnte. Der Versuch wurde mit vier kalibrierten PAO-Ölen im Viskositätsbereich von 10 cSt bis 1000 cSt durchgeführt. Jedes Öl wurde bei Drücken von bis zu 500 MPa und bei Scheibengeschwindigkeiten zwischen 50 mm/s und 1000 mm/s getestet. Ziel war es, einen breiten Bereich von Filmdicken mit unterschiedlichen Ölen abzudecken. Abbildung 6 zeigt die Ergebnisse der Validierung. Die experimentellen Ultraschall- und optischen Ergebnisse werden zudem mit der Hamrock-Dowson-Gleichung für die Filmdicke verglichen. Diese stellt ein etabliertes Verfahren in der Tribologie dar und beschreibt die Filmdicke im hydrodynamischen Schmierregime als Funktion von Drehzahl, Last, Temperatur und Mitnahmeviskosität des Öls. Die Validierung zeigt eine hohe Übereinstimmung zwischen optischen und ultraschallbasierten Ergebnissen: Über 75 % der Ultraschallmessungen liegen innerhalb von 10 % der optischen Messwerte, und über 90 % der Ergebnisse liegen zwischen der Hamrock-

Dowson-Vorhersage und der optischen Messung.

Ergebnisse in einem oszillierenden Prüfstandtribometer

Die Validierung des Algorithmus wurde zunächst in einem hydrodynamischen Kontakt durchgeführt (typisch beispielsweise für Gleit- oder Radiallager), bei dem eine vollständige Trennung zwischen TaP und Gegenkörper vorliegt. In realen Anwendungen treten jedoch häufig Grenz- oder Mischreibungszustände auf, insbesondere bei oszillierenden Komponenten wie Zahnrädern, Kolbenringen oder verschiedenen Lagerarten.

Zur Simulation solcher Bedingungen wird häufig ein oszillierendes Tribometer eingesetzt. In diesem bewegt sich der Pin zwischen zwei Endpunkten hin und her, während er gegen den Gegenkörper gedrückt wird (vgl. Abbildung 7). Unter diesen Bedingungen ist es schwieriger, eine vollständige hydrodynamische Trennung zu erreichen, sodass entweder Rauheitsspitzen (Asperities) in Kontakt treten oder sich ein Tribofilm bildet. Dieser Tribofilm ist eine viskoelastische Schicht, die durch den Verbrauch von Additiven im Schmierstoff infolge der Relativbewegung entsteht. Er kann gezielt ausgelegt werden, um die Trennung der Kontaktflächen zu verbessern, und seine Entwicklung ist entscheidend für die Bewertung der Schmierstoffleistung.

In diesem Abschnitt wird ein beispielhafter Datensatz von Messungen mit dem TaP vorgestellt, um die Leistungsfähigkeit des Sensors zu demonstrieren.

Abbildung 8 zeigt den Verlauf der Tribofilmdicken-Degradation über mehrere Wiederholungen für ein 0W-5-Schmieröl. Die Messungen des triboakustischen Pins werden mit *ex-situ* Elektronenmikroskopie-Messungen der Tribofilmdicke nach einem FIB-Schnitt verglichen. Der Vergleich zeigt eine gute Übereinstimmung sowie einen korrekten Trendverlauf über alle Wiederholungen hinweg.

Es ist zu beachten, dass jeder Datenpunkt in Abbildung 8 ein Mittelwert aus über 100.000 Einzelwellenformen ist, entsprechend etwa 50 Hübchen. Jeder Hub von 2 mm wird dabei mit etwa 2000 Messpunkten rekonstruiert. Abbildung 9 zeigt die hubaufgelöste Filmdicke für einen ausgewählten Punkt der Kurve aus Abbildung 8. Die Tribofilmdicke kann insbesondere an den Umkehrpunkten

bestimmt werden, da dort keine Relativbewegung zwischen Pin und Gegenkörper vorliegt und der flüssige Schmierstoff entweder aus dem Kontakt verdrängt wird oder keinen Beitrag zur Kontaktsteifigkeit leistet. Die Messung entspricht daher einem Mittelwert aus Tribofilm und festen Rauheitskontakten.

Sobald die Bewegung wieder einsetzt, nimmt die Filmdicke zu, was den Beitrag des flüssigen Schmierstoffs zur effektiven Kontakttrennung widerspiegelt. Da im Misch- bzw. Grenzreibungsregime nicht ausschließlich das Öl zur Trennung beiträgt, muss der Algorithmus aus Gleichung (3) angepasst werden, um mittlere Kontakteigenschaften zu berücksichtigen (z. B. den Kompressionsmodul der Stahlasperitäten statt der Öleigenschaften bei hohem Druck). Die Messung wird daher als effektiver oder äquivalenter Abstand zwischen TaP und Gegenkörper interpretiert.

Schlussfolgerung

Es wurde ein neuartiger ultraschallbasierter Triboakustischer Pin zur *In-situ*- und Echtzeitcharakterisierung von Schmierfilmen in tribologischen Kontakten vorgestellt. Durch die Integration miniaturisierter piezoelektrischer Wandler in eine standardisierte Pin-Geometrie ermöglicht der Sensor die direkte Messung der Filmdicke und weiterer schmierungsrelevanter Eigenschaften unter realistischen Betriebsbedingungen. Der Ansatz folgt dem aktuellen Trend, bestehende Tribometer durch zusätzliche Sensorik zu erweitern, anstatt neue Prüfsysteme zu entwickeln, und gewährleistet somit die Kompatibilität mit etablierten industriellen Verfahren.

Der Sensor wurde erfolgreich in verschiedenen Tribometerkonfigurationen implementiert und demonstriert eine hohe Vielseitigkeit sowie einfache Integration. Die Validierung zeigt eine sehr gute Übereinstimmung mit optischen Messmethoden und theoretischen Modellen und bestätigt die Zuverlässigkeit des Ansatzes. Darüber hinaus wurde gezeigt, dass der TaP aufgrund seiner Geometrie eine erhöhte Sensitivität gegenüber sehr dünnen Filmen (bis hinunter zu 20 nm) im Vergleich zu klassischen Reflektometrieverfahren aufweist. Zusätzlich ermöglicht die Fähigkeit zur Erfassung transienter Effekte wie Tribofilmbildung und -abbau wertvolle Einblicke, die mit konventionellen Methoden nicht zugänglich sind.

Insgesamt stellt der triboakustische Pin eine praxisnahe und skalierbare Lösung für die

fortgeschrittene Schmierstoffcharakterisierung dar und unterstützt die Entwicklung effizienterer, zuverlässigerer und langlebiger mechanischer Systeme.

Abbildungen

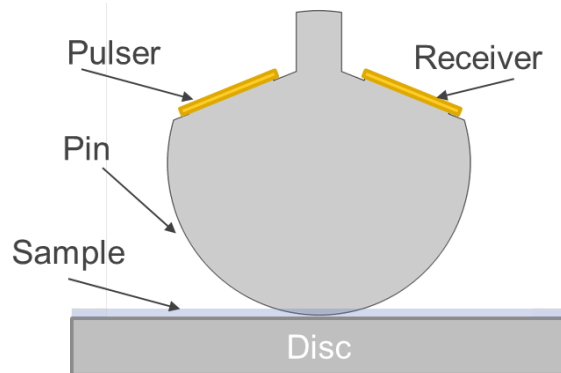


Figure 1: Schematische Darstellung des Triboakustischen Pins.

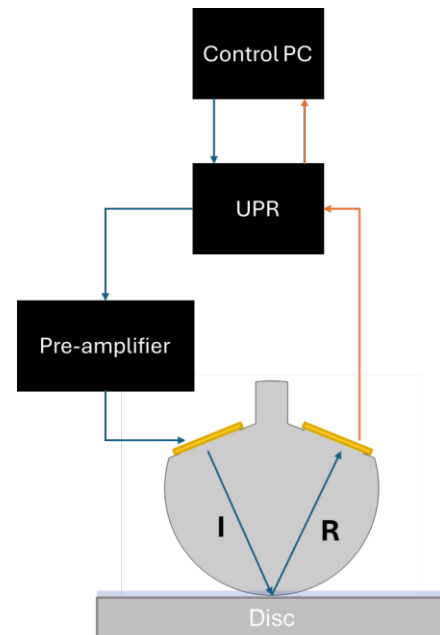


Figure 3: Messkette zur Datenerfassung.

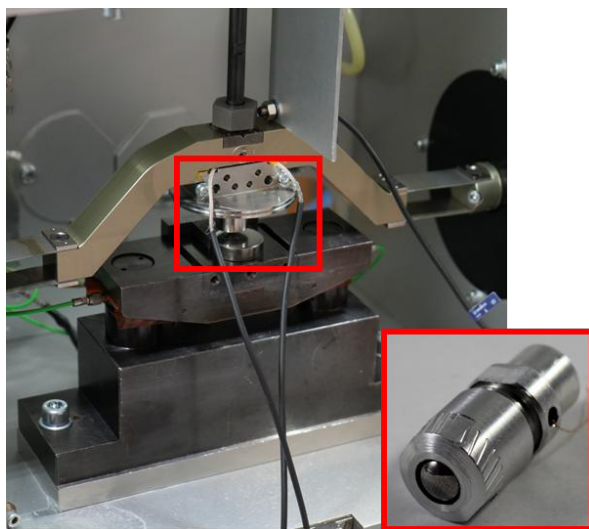


Figure 2: Der Triboakustische Pin eingebaut in einem Tribometer.

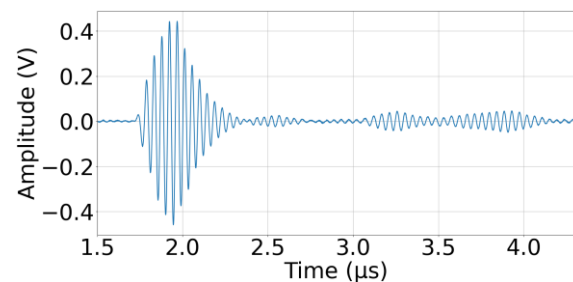


Figure 4: Beispiel Reflexion für den Triboakustischen Pin in Luft.

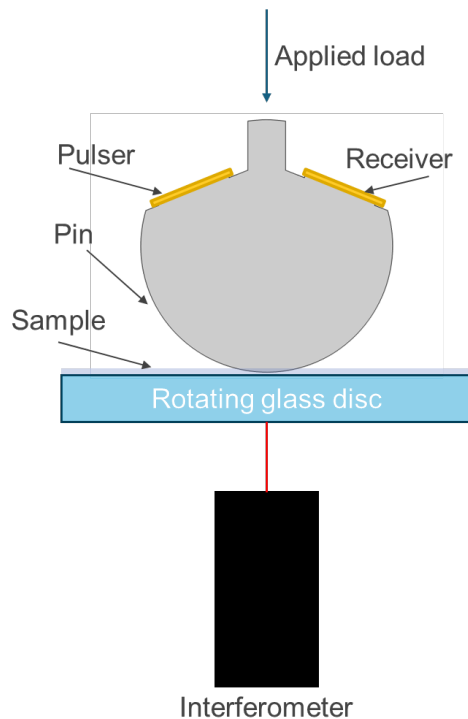


Figure 5: Schmatische Darstellung des aufbaus zur Validierung des Triboakustischen Pins.

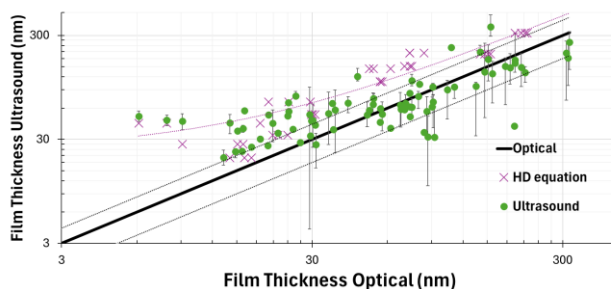


Figure 6: Validierung des Triboakustischen Pins mittels optischer Interferometrie.

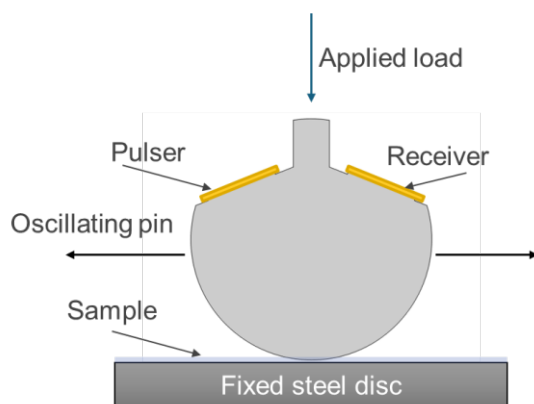


Figure 7: Schematische Darstellung des Triboakustischen Pins in einem oszillierenden System.

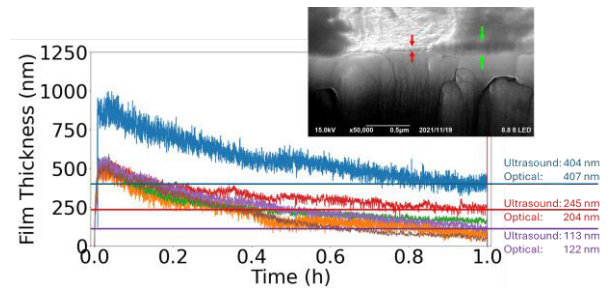


Figure 8: Tribofilm Bildung und Abbau gemessen mit dem Triboakustischen Pin in einem Oszillierenden Aufbau.

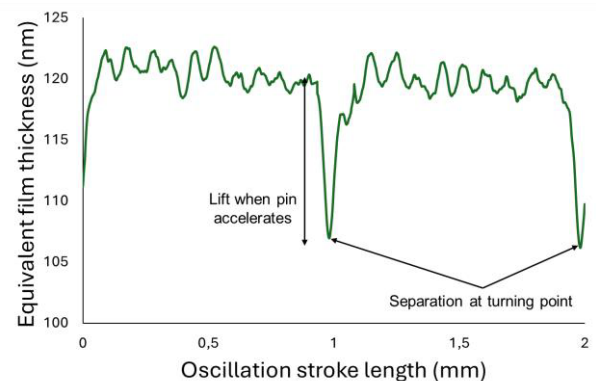


Figure 9: Orts- und hubaufgelöste Filmdicken messung, gemittelt über 40 Hübe.

Aknowledgment

Diese Arbeit wurde im Rahmen des COMET Zentrums InTribology (FFG-Nr. 906860), einem Projekt des „Kompetenzzentrums für Tribologie“ (AC2T research GmbH), durchgeführt. InTribology wird durch das COMET – Competence Centers for Excellent Technologies Programm von den Bundesministerien BMIMI und BMWETals auch den Bundesländern Niederösterreich und Vorarlberg auf Basis der finanziellen Unterstützung durch die beteiligten Projektpartner gefördert. COMET wird durch die Österreichische Forschungsförderungsgesellschaft (FFG) abgewickelt.

Literaturnachweis

- [1] Schirru, M., and M. Varga. "A review of ultrasonic reflectometry for the physical characterization of lubricated tribological contacts: history, methods, devices, and technological trends." *Tribology Letters* 70.4 (2022): 129..
- [2] Schirru, Michele, and Michael Adler. "An ultrasonic rheometer to measure gas absorption in ionic liquids: design, calibration and testing." *Sensors* 20.12 (2020): 3544.
- [3] Gray, William, and Robert Sean Dwyer-Joyce. "Ultrasonic resonance approach to measure rolling element bearing raceway films in-situ: A

practical approach." Tribology International
(2025): 111141.

- [4] Dwyer-Joyce, R. S., T. Reddyhoff, and J. Zhu.
"Ultrasonic measurement for film thickness and
solid contact in elastohydrodynamic lubrication."
(2011): 031501.

Compact Ion Mobility Spectrometer with Two Parallel Drift Tubes for Simultaneous Detection of Positive and Negative Ions

Tim Ostermeier¹, Timo Jähner¹, Alexander Nitschke¹, Moritz Hitzemann¹, Stefan Zimmermann¹

¹Leibniz University Hannover, Appelstr. 9a, 30167 Hannover, Germany
 ostermeier@geml.uni-hannover.de, +49 511 762 4222

Abstract

Rapid and reliable on-site detection of hazardous substances requires analytical instruments that combine high sensitivity, selectivity, and portability. In this work, a compact dual-polarity ion mobility spectrometer (IMS) is presented, allowing the simultaneous detection of positive and negative ions. The IMS is specifically designed with a focus on low weight, reduced power consumption, and compact size, making it suitable for mobile and drone-based applications. By integrating two parallel drift tubes and a single photon-based ionization source, the system shows efficient ion generation for both polarities. Experimental results demonstrate high resolving power and low limits of detection in the low parts per trillion range.

Keywords: ion mobility spectrometry, dual-polarity, parallel drift tubes, photon source

Many safety- and security-related applications require fast, sensitive, and compact measuring devices for on-site detection of hazardous substances. Furthermore, reliable detection of target substances in real environments requires appropriate selectivity. Here, ion mobility spectrometers (IMS) are well suited, as they reach detection limits in the low ppt_v range (parts per trillion) in short measurement time of just one second, while reaching high separation power in relation to instrumental size. The fundamental principle of IMS is the separation of ions based on their specific ion mobility in a neutral gas under the influence of an electric field. If the sample volume is limited or the application requires immediate response, simultaneous detection of positive and negative ions would be beneficial to gather full analytical information of the sample from one single measurement. This can be achieved by using a dual-polarity IMS with two parallel drift tubes. For later use on a drone, the dual-polarity IMS presented here is designed low-weight, low-power and compact targeting < 1400 g, < 14 W and < 15 x 20 x 25 cm³ with drift tubes made low-cost from standard printed circuit boards. Furthermore, the design requires just one

small high energy photon source efficiently generating positive and negative reactant and product ions to be injected in the two drift tubes.

The dual-polarity IMS reaches outstanding detection limits in the low-ppt_v range for the compounds tested, such as acetylacetone in both polarities (AcACE(+): 19.35 ppt_v, AcACE(-): 27.68 ppt_v), sevoflurane in negative polarity (SEVO(-): 28.67 ppt_v), and 2-hexanone in positive polarity (MBK(+): 118.39 ppt_v), all within one second of signal averaging. In addition, the device provides outstanding resolving power of $R_P = 80$. Furthermore, the IMS design allows for easy coupling with gas chromatography to allow for analyzing more complex samples or detecting target compounds in more complex environments.

Acknowledgment

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project number 540972827.

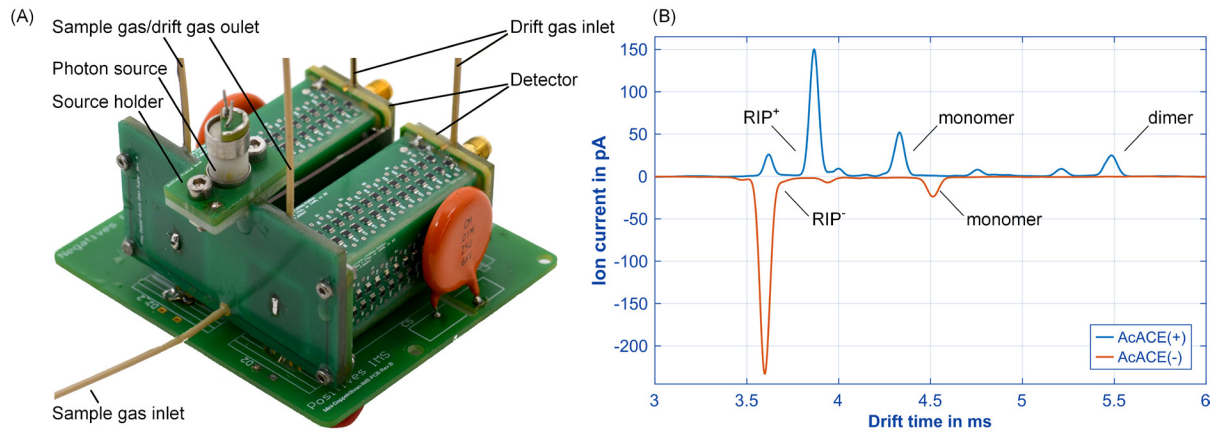


Figure 1: (A) Photo of the dual-polarity PCB-IMS with two drift tubes but just one single ionization source, here a small high energy photon source. (B) IMS spectra of acetylacetone in air in both polarities with the protonated monomer ($K_0 = 1.81 \text{ cm}^2/(\text{Vs})$) and proton-bound dimer ($K_0 = 1.43 \text{ cm}^2/(\text{Vs})$) peaks in positive polarity and the deprotonated monomer peak ($K_0 = 1.74 \text{ cm}^2/(\text{Vs})$) in negative polarity at a concentration of 517 ppt_v. The ion mobilities of the RIP^+ and RIP^- are $K_0 = 2.02 \text{ cm}^2/(\text{Vs})$ and $K_0 = 2.18 \text{ cm}^2/(\text{Vs})$.

Coated Blade Spray Ion Mobility Spectrometry for Medical Applications

*Sarah Deicke, Christian Thoben, Tim Ostermeier, Franziska A. Reddner, Stefan Zimmermann
Leibniz University Hannover, Institute of Electrical Engineering and Measurement Technology,
Department of Sensors and Measurement Technology, Appelstr. 9A, 30167 Hannover, Germany*

Abstract

Coated Blade Spray (CBS) is a user-friendly microextraction technique that uses polymer-coated blades to selectively enrich analytes, remove unwanted background matrix compounds, and serve as an electrospray emitter. In this study, CBS was coupled with a high-resolution ion mobility spectrometer, performing effectively in both positive and negative ion modes. Using this compact setup, a wide range of substances, including drugs, pesticides, and other pharmaceuticals, can be rapidly detected on-site with highest sensitivity in complex liquid samples.

Keywords: Coated Blade Spray, Ion Mobility Spectrometry, Microextraction, Electrospray Ionization, Medical Technology

Introduction

Coated Blade Spray (CBS) represents an easy to handle and robust microextraction technology introduced by Pawliszyn et al. that employs coated metal blades as both extraction device and emitter for electrospray ionization (ESI) [1]. Thus, the selective enrichment of analytes on a polymer coating, rinsing off complex background, followed by desorption by applying solvent enables direct ionization of analytes by ESI. This approach has now been combined with our previously described ESI-Ion Mobility Spectrometer (ESI-IMS) [2] giving a substantial advancement in the rapid and efficient analysis of complex samples and has already been successfully applied for the detection of drugs and pesticides in drinks and liquids [3].

Experimental section

The IMS used in this study has a drift length of only 75 mm and provides high resolving power of $R_P = 100$. To provide higher throughput, up to 10 blades can be sequentially positioned in front of the IMS by an autosampler. The autosampler also adds a defined volume of the desorption solvent to the coated area before turning on the ESI voltage. Here, Ringer's solution is selected as the sample matrix to substitute for blood, and the short-chain fatty acid isovaleric acid and the painkiller diclofenac are selected as analytes.

Results

The method demonstrates excellent capability for selective pre-concentration and subsequent detection by IMS. In addition to the positive ion mode, the CBS-IMS can also operate in negative ion mode.

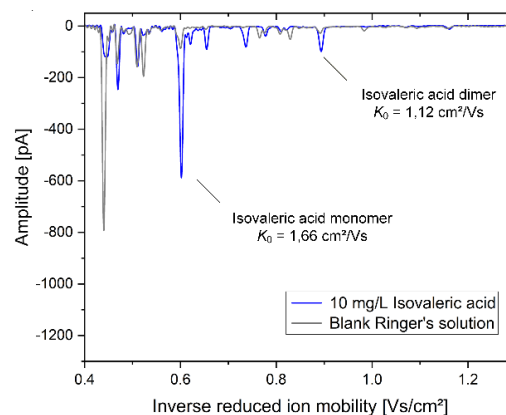


Abb. 1: Example of an ion mobility spectra from 10 mg/L isovaleric acid (blue) in 90:10 (v:v) Ringer's solution:methanol (gray).

Acknowledgments

We thank Restek Corporation for supplying the coated blades and PAS Technology Deutschland GmbH for providing the blade spray autosampler.

References

- [1] M. Tascon, G.A. Gómez-Ríos, N. Reyes-Garcés, J. Poole, E. Boyacı, J. Pawliszyn, *Anal. Chem.* 89, 8421–8428 (2017); doi: 10.1021/acs.analchem.7b01877.
- [2] C. Thoben, T. Werres, I. Henning, P.R. Simon, S. Zimmermann, T.C. Schmidt, T. Teutenberg, *Green Anal. Chem.* 1, 100011 (2022); doi: 10.1016/j.greeac.2022.100011.
- [3] C. Thoben, J.J. Stadler, P.R. Simon, C.-R. Raddatz, M. Sehlmeier, S. Zimmermann, *Anal. Chem.* 96, 3593–3599 (2024); doi: 10.1021/acs.analchem.3c05586

Entwicklung und Evaluierung kosteneffizienter Peak-Detektor-Schaltungen für das Online-Monitoring von Teilentladungen an Mittelspannungskabeln

Felix Grimm¹, Felix Kolodziej¹, Florian Wein¹, Christoph Andres¹, Ulrike Steinmann¹

¹Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg

Zusammenfassung

Ein flächendeckendes Online-Monitoring von Mittelspannungskabeln mittels Teilentladungsdiagnostik auf Basis induktiver Sensoren wie Hochfrequenz-Stromwandlern erfordert Sensorsysteme, die hochdynamisch und zugleich kosteneffizient sind. Bisherige Lösungen zur Peak-Detektion von Nanosekunden-Impulsen basieren jedoch auf teuren Präzisions-ICs, was die wirtschaftliche Skalierbarkeit im Verteilnetz stark einschränkt. Diese Arbeit präsentiert stattdessen modifizierte, diskrete Peak-Detektor-Topologien auf Basis von Standard-Operationsverstärkern. Die simulativ und messtechnisch validierten Schaltungen übertreffen teure Spezial-ICs in entscheidenden Parametern: Bei geringeren Materialkosten erzielen sie eine signifikant längere Haltedauer sowie eine deutlich niedrigere Detektionsschwelle für stark gedämpfte Signale. Damit schaffen sie die essenzielle Hardware-Basis, um transiente Impulse für kostengünstige Mikrocontroller zuverlässig erfassbar zu machen und großflächige Sensornetzwerke wirtschaftlich zu realisieren.

Keywords: analoge Peak-Detektion, Impulsamplitudenmessung, kosteneffektiv, Nanosekunden-Impuls, Teilentladung, Zustandsüberwachung

1. Einleitung und Motivation

Die fortschreitende Dezentralisierung der Stromerzeugung bedingt zunehmend bidirektionale Lastflüsse in elektrischen Verteilnetzen [1], [2]. Insbesondere die überalterten Mittelspannungskabel mit einer deutschlandweiten Gesamtlänge von rund 530.000 Kilometern [3] werden hierdurch unvorhergesehenen Belastungen ausgesetzt, was das Risiko für spontane Isolationsdurchschläge signifikant erhöht [4], [5]. Zur Gewährleistung der Versorgungssicherheit ist ein Paradigmenwechsel hin zu einer zustandsorientierten Instandhaltung unabdingbar. Die Erfassung von Teilentladungen (TE) hat sich dabei als zuverlässiger Frühindikator für beginnende Isolationsdefekte bewährt [6]. Einer flächendeckenden Skalierung dieses Verfahrens stehen derzeit jedoch die hohen Investitionskosten kommerzieller TE-Messsysteme entgegen, die einen wirtschaftlichen Einsatz an tausenden dezentralen Netzknoten verhindern. Diese Kosten resultieren maßgeblich aus der Natur der TE, bei denen es sich um sehr kurze Stromimpulse im Nanosekundenbereich handelt [6]. Um diese transienten Signale formgetreu zu erfassen, ist eine hohe Messbandbreite bis in den hohen Megahertz-

Bereich erforderlich. Zwar existieren entsprechend hochfrequente Analog-Digital-Wandler (ADCs), die diese Abstraten leisten können, jedoch führt deren Einsatz im kontinuierlichen Online-Betrieb zu immensen Datenmengen. Die Echtzeitverarbeitung, Zwischenspeicherung und Kommunikation dieser hochauflösenden Datenströme, übersteigt die Kapazitäten und Bandbreiten kostengünstiger Mikrocontroller.

Um eine wirtschaftliche Skalierung an tausenden Netzknoten zu ermöglichen, muss die anfallende Datenmenge bereits vor der Digitalisierung drastisch reduziert werden. Zur Umsetzung dieses Ziels können kosteneffiziente Peak-Detektor-Schaltungen (PDCs) als analoge Vorverarbeitungsstufe eingesetzt werden. Diese erfassen lediglich die Maximalamplitude der TE-Impulse und halten sie für die amplitudentreue Abtastung durch handelsübliche, langsamere Mikrocontroller-ADCs bereit. Ziel ist es, den Zielkonflikt aus hoher physikalischer Bandbreite, Messgenauigkeit und Kosteneffizienz aufzulösen, um so die Hardware-Basis für ein flächendeckendes Online-Monitoring zu schaffen.

2. Theoretischer Hintergrund und TE-Erfassung

Gemäß der internationalen Norm IEC 60270 sind TE lokale elektrische Entladungen, welche die Isolation zwischen Leitern nur teilweise überbrücken. Physikalisch handelt es sich dabei um elektrische Micro-Sparks in Isolationssystemen, die unter dem Einfluss hoher elektrischer Felder entstehen [7].

Solche TE-Aktivitäten sind häufig Symptome für herstellungsbedingte Mängel oder umweltbedingte Alterungsprozesse und können langfristig zur vollständigen Zerstörung der Isolation und damit zum Betriebsmittelausfall führen. Jede dieser TE resultiert physikalisch in einem Stromimpuls, der an den Anschlüssen des Prüfobjekts messbar ist [7].

In der messtechnischen Zustandsüberwachung werden diese Ereignisse als transiente Spannungsimpulse erfasst. Sie sind charakterisiert durch extrem kurze Impulsbreiten im Nanosekundenbereich sowie geringe Amplituden von meist unter 1 V. Diese Kombination aus nadelförmiger Signalform und minimaler Signalstärke macht die zuverlässige Detektion und exakte Amplitudenanalyse zu einer komplexen messtechnischen Herausforderung [8]. Die Auskopplung dieser Signale erfolgt dabei vorzugsweise induktiv mittels eines High-Frequency Current Transformers (HFCT), welcher den transienten TE-Stromimpuls in ein proportionales Spannungssignal übersetzt, das als Eingangsgröße für die PDC dient.

Die direkte Digitalisierung transientser TE-Impulse mittels High-Speed-ADCs erfordert kosten- und ressourcenintensive Abtastraten im Gigasample-Bereich und scheidet für dezentrale Low-Cost-Monitoringsysteme aus. Preiswerte Standard-Mikrocontroller sind aufgrund ihrer limitierten Abtastraten hingegen systembedingt oft nicht in der Lage, derart kurze Impulse im Nanosekundenbereich messtechnisch zu erfassen. Um diese hardwareseitigen Limitierungen der Abtastung zu umgehen, basieren etablierte Messmethoden nach IEC 60270 zumeist auf Quasi-Integrationsverfahren (Bandpassfilterung <1 MHz) [9]. Anstatt den schnellen Puls direkt aufzuzeichnen, wird hierbei durch Integration eine scheinbare Ladung ermittelt, welche anschließend mit einem Kalibrierungsfaktor multipliziert wird [7]. Da eine solche normgerechte Kalibrierung am installierten Kabelstrang jedoch in der Regel eine Anlagenabschaltung erfordert, findet dieses Verfahren vorwiegend in Offline-Messungen Anwendung. Zudem eliminiert die systembedingte Signalintegration die für eine

präzise Isolationsdiagnostik essenzielle Information über den tatsächlichen Spannungsspitzenwert des Ursprungsimpulses [8]. Um dennoch ein autarkes Online-Monitoring zu realisieren, verfolgt diese Arbeit daher einen grundlegend anderen Ansatz: Anstatt das Signal zu integrieren und zu kalibrieren, zielt die Schaltungsentwicklung auf die direkte Erfassung der reinen, hochfrequenten Spannungsimpulse ab [8].

Um diese extrem schnellen Pulse ohne die Generierung immenser Datenmengen zu erfassen, bietet sich die analoge Peak-Detektion (PD) als ressourcenschonende Lösung an. Das Grundprinzip solcher PDCs ist dabei sehr einfach: Eine elementare passive Grundschialtung aus Gleichrichterdiode und Speicherkondensator speichert das Spannungsmaximum theoretisch als Gleichspannung ab (vgl. Abbildung 1). Für TE-Pulse unter 1 V ist diese simple Topologie in der Praxis jedoch ungeeignet, da der Durchlassspannungsabfall der Diode sowie dynamische Rück- und stetige Leckströme größere Messfehler verursachen [8,10].

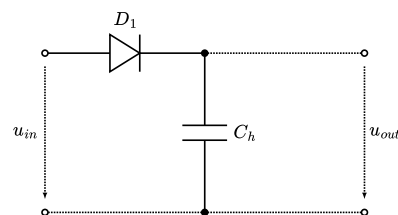


Abb. 1: passive Grundschialtung für PDCs

Aktive Adaptionen zur Kompensation dieser Fehler und deren spezifische Herausforderungen bei Nanosekunden-Impulsen werden von Fritsch et al. [8] detailliert betrachtet.

Gegenwärtige State-of-the-Art-Lösungen zur schnellen Spitzenwerterfassung weisen für ein kostensensibles Online-Monitoring erhebliche Nachteile auf:

- Vollständige Digitalisierung (Gigasample-ADCs & FPGAs): Scheidet für dezentrale Low-Cost-Sensoren aufgrund hoher Hardwarekosten, immensen Datenaufkommens und zu hoher Leistungsaufnahme aus [8].
- Integrierte CMOS-Schaltungen [11],[12],[13]: Bieten zwar höchste Präzision, erfordern jedoch enorme Entwicklungskosten und sind nicht als günstige Standardbauteile am Markt verfügbar [8].
- Spezialisierte SOTA-Bausteine (z. B. OPA615): Verarbeiten Nanosekunden-

Impulse zwar ohne Sättigungseffekte, treiben die Materialkosten pro Sensorknoten im Vergleich zu Standardbauteilen jedoch signifikant in die Höhe [8].

3. Entwicklung und Konfiguration der Peak-Detektor-Topologien

Um einen optimalen Kompromiss aus Kosten, Bandbreite und Messgenauigkeit zu erzielen, werden drei modifizierte, diskrete Topologien sowie eine kommerzielle Referenzschaltung zur TE-Spitzenwerterfassung evaluiert. Da analytische Berechnungen das nichtlineare Hochfrequenzverhalten bei Nanosekunden-Impulsen unzureichend abbilden [8], erfolgt die Dimensionierung iterativ mittels LTSpice-Simulationen. Diese dienen primär der exakten Bauteildimensionierung sowie der prinzipiellen Machbarkeitsprüfung der Schaltungen unter extremen dynamischen Anforderungen.

Modellierung und Simulation

Zur Nachbildung des Spektrums realer TE-Ereignisse in Mittelspannungsnetzen bilden drei synthetische Testimpulse die Basis der simulativen Evaluierung [14] (vgl. Abbildung 2):

- Standard-Impuls (160 mV, 60 ns): Typisches, gut messbares TE-Ereignis.
- Hochfrequenz-Impuls (180 mV, 10 ns): Prüft Bandbreiten- und Slew-Rate-Limitierungen bei extrem schnellen Transienten.
- Dispersions-Impuls (70 mV, 60 ns): Simuliert stark gedämpfte Signale nach längerer Kabellaufzeit.

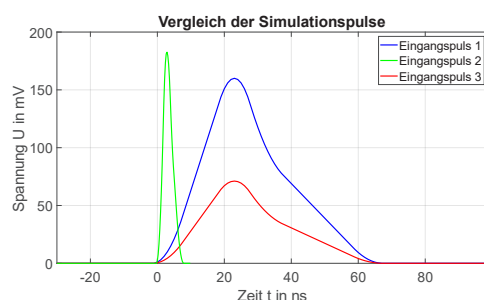


Abb. 2: Vergleich der drei synthetischen Eingangsimpulse für die simulative Untersuchung.

Ergänzende Simulationen unter Rausch- und Impulsfolgen-Bedingungen bestätigten diese Einzelanalysen qualitativ und werden aus Übersichtlichkeits- und Platzgründen nicht detailliert aufgeführt.

Bauteilselektion und spezifische Schaltungskonfiguration

Da in der Literatur empfohlene Schottky-Dioden (z. B. BAS40, BAT17) [8] den geforderten steilen Anstiegszeiten nicht genügten, wurde aufgrund besserer Dynamik und geringerer Vorwärtsspannungsverluste topologieübergreifend die Schottky-Barrierediode RB886CM implementiert. Als Speicherelement kam stets ein 22-pF-Kondensator zum Einsatz, der in Voruntersuchungen die besten Resultate lieferte. Da dieser das Systemverhalten maßgeblich beeinflusst, ist dessen detaillierte Optimierung in künftigen Untersuchungen angedacht (vgl. Ausblick). Die weitere Auslegung gestaltet sich wie folgt:

Schaltung 1: Weiterentwickelter Standard-Peak-Detektor

Diese Topologie ergänzt den klassischen aktiven Spitzenwertgleichrichter um eine zweite Diode im Feedback-Loop. Fällt die Eingangsspannung unter den Wert des Speicherkondensators, schließt diese Diode den Regelkreis. Dies verhindert ein Übersteuern des OPV in die negative Sättigung und minimiert die Einstellzeit. Ein Impedanzwandler entkoppelt abschließend den Speicherkondensator [8,15–17]. In der praktischen Ausführung wird als OPV der AD8047 eingesetzt (vgl. Abbildung 3).

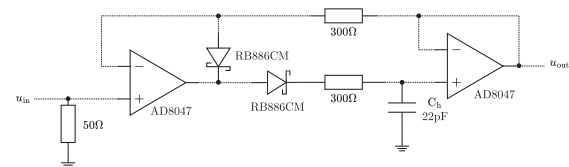


Abb. 3: Schaltplan des weiterentwickelten Standard-Peak-Detektors mit dem OPV AD8047.

Schaltung 2: Improved Active Diode Peak Detection

Zur Minimierung der Speicherkondensator-Entladung (Droop) durch Leckströme nutzt dieser Ansatz das „Ultra-Low Drift Peak Detector“-Prinzip [18]. Eine Bootstrap-Anordnung klemmt das Potenzial vor der Diode in der Haltephase auf das Ausgangsniveau. Die minimierte Sperrspannung reduziert Leckströme drastisch, was lange Haltezeiten erlaubt, jedoch potenziell die Bandbreite limitiert [19]. Die umgesetzte Schaltung nutzt den OPV LMH6642Q. Gegenüber der Literaturvorlage [18,19] wurde ein Feedback-Widerstand entfernt, sodass die Struktur Schaltung 1 ähnelt, jedoch mit serieller Diodenanordnung (vgl. Abbildung 4).

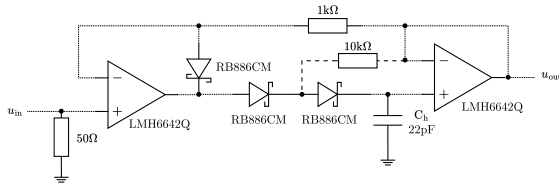


Abb. 4: Schaltplan der Improved Active Diode Peak Detection Topologie unter Verwendung des LMH6642Q.

Schaltung 3: Zweistufige Peak-Erfassung

Im Gegensatz zu den geschlossenen Regelkreisen der vorherigen Topologien trennt dieser Ansatz die Signalverstärkung von der eigentlichen Gleichrichtung. Die Schaltung gliedert sich funktional in einen linearen Vorverstärker, eine nachgeschaltete passive Gleichrichterstufe und einen abschließenden Impedanzwandler. Durch die lineare Vorverstärkung des schwachen TE-Signals auf ein höheres Spannungsniveau wird der relative Messfehler, der durch den nichtlinearen Durchlassspannungsabfall der passiven Diode entsteht, deutlich verringert. Der zentrale Systemvorteil dieser Entkopplung liegt in der hohen Stabilität: Die erste OPV-Stufe arbeitet als reiner Linearverstärker ohne limitierende nichtlineare Schaltelemente im Feedback-Loop. Für die Vorverstärkung sowie die Ausgangsverstärkung wird der OPV LMH6642Q verwendet und entsprechend den Datenblattspezifikationen beschaltet. Durch die Widerstände wird eine sechsfache Vorverstärkung und eine zweifache Ausgangsverstärkung erreicht. Die Schaltung könnte somit bei einer idealen Diode ohne Spannungsabfall eine Verstärkung von zwölf erreichen (vgl. Abbildung 5).

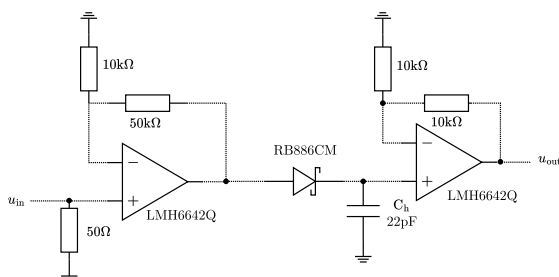


Abb. 5: Schaltplan der zweistufigen Peak-Erfassung mit linearer Vorverstärkung durch den OPV LMH6642Q.

Schaltung 4: Transkonduktanz-basierter Detektor

Als Benchmark für den Stand der Technik dient ein Sampling Operational Transconductance

Amplifier (SOTA). Hier lädt ein Transkonduktanzverstärker den Kondensator stromgesteuert (proportional zur Eingangsspannungsdifferenz). Diese Bausteine sind für Nanosekunden-Impulse optimiert und bieten extrem kurze Reaktionszeiten [8,10,14]. Als Referenz dient der SOTA OPA615, der gemäß Datenblatt die RB886CM-Diode treibt (vgl. Abbildung 6). Da seine hohen Kosten dem Ziel flächendeckender Sensorik widersprechen, dient er als Maßstab, um die Messgenauigkeit der kostengünstigen diskreten Alternativen (Schaltungen 1–3) vergleichend zu evaluieren.

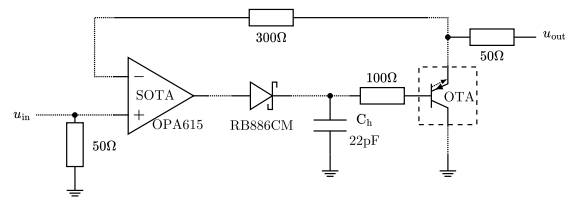


Abb. 6: Schaltplan des Transkonduktanz-basierten Detektors (Referenz) auf Basis des OPA615.

4. Messaufbau und -durchführung

Um die Leistungsfähigkeit und Sensitivität der verschiedenen Spitzenwerterfassungsschaltungen unter reproduzierbaren Bedingungen zu evaluieren, wurde der in Abbildung 7 dargestellte Messaufbau genutzt. Dieser setzt sich grundlegend aus folgenden vier Kernkomponenten zusammen:

(1) Funktions- und Pulsgenerator (RIGOL DG1022Z): Dieses Gerät dient als Signalquelle zur synthetischen Erzeugung der Eingangssignale. Um reale physikalische Bedingungen nachzubilden werden sehr kurze Nanosekunden-Impulse mit variierbaren, geringen Amplituden (< 1 V) generiert.

(2) Labornetzgerät (Aim-TTi CPX200D): Die duale Spannungsquelle stellt die symmetrische Gleichspannungsversorgung für die aktiven Bauelemente, insbesondere die Operationsverstärker, der jeweiligen Messschaltungen bereit.

(3) Evaluierungs-Platine: Hierbei handelt es sich um die eigentliche, zu untersuchende Schaltungstopologie. Die kompakten Platinen fungieren als austauschbare Module im Signalpfad, um einen direkten und unverfälschten Vergleich der verschiedenen Hardware-Designs zu ermöglichen.

(4) Digitales Oszilloskop (Rohde & Schwarz RTB2004): Das Oszilloskop dient der hochauflösenden parallelen Aufzeichnung und Auswertung der Signale. Es erfasst visuell und

messtechnisch simultan den vom Generator eingespeisten Ursprungsimpuls sowie die von der Testschaltung generierte und gehaltene Ausgangsspannung.

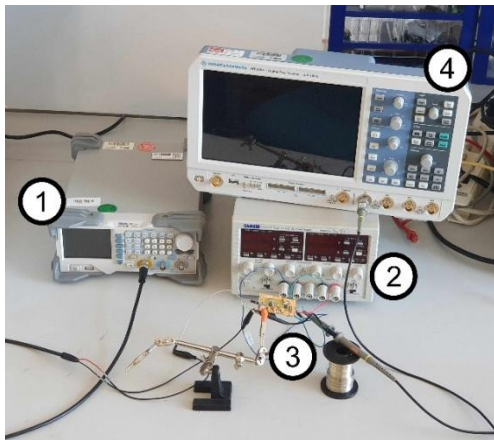


Abb. 7: Instrumentierter Messaufbau zur Evaluierung der PDCs. Ein Pulsgenerator (1) synthetisiert sehr kurze Testimpulse variabler Amplitude und Impulsdauer, welche in die gleichspannungsversorgte (2) Testplatine (3) eingespeist und anschließend mittels Oszilloskop (4) getriggert erfasst werden.

Während der praktischen Evaluierung speist der Pulsgenerator (1) die definierten transienten Stromimpulse direkt in die jeweilige, vom Netzgerät (2) versorgte Testplatine (3) ein. Das Oszilloskop (4) triggert auf die steigende Flanke des Eingangsimpulses und zeichnet das dynamische Verhalten der Schaltung auf. Durch eine Variation der eingespeisten Impulse, wird das Halteverhalten systematisch analysiert. Dabei wird über den Pulsgenerator (1) die Amplitude und Länge der Testimpulse entsprechend angepasst.

5. Messergebnisse und Validierung

Zur Evaluierung der entwickelten Low-Cost-Topologien (Schaltung 1–3) gegenüber der industriellen Referenz (Schaltung 4) wurden die Prototypen mit spezifischen Testimpulsen untersucht, die an reale TE-Ereignisse angelehnt sind [14]. Methodisch basieren diese auf den Signalen der vorangegangenen simulativen Untersuchung, wurden jedoch an die experimentellen Rahmenbedingungen (u. A. Limitationen der Signalgenerierung) adaptiert. Bewertet wurden anschließend die Signaltreue, Haltezeit, Detektionsschwellen und Kosteneffizienz.

Vor der quantitativen Auswertung ist das Ruhespannungsverhalten der Topologien zu berücksichtigen. Einige Schaltungen,

insbesondere Schaltung 2 und in geringerem Maße Schaltung 1, weisen im Ruhezustand einen statischen DC-Offset auf. Da dieser Offset konstant ist und durch einen DC-Filter oder im Rahmen der digitalen Nachverarbeitung durch den ADC softwareseitig trivial entfernt werden kann, wurden die Signale in der grafischen Evaluation um diesen Wert bereinigt und auf eine gemeinsame 0-V-Basislinie normiert. Ob die Ursache dieses Offsets in der Schaltungstopologie selbst, dem Messaufbau oder Effekten der Spannungsversorgung begründet liegt, wurde aufgrund der vernachlässigbaren systemischen Auswirkungen nicht abschließend analysiert.

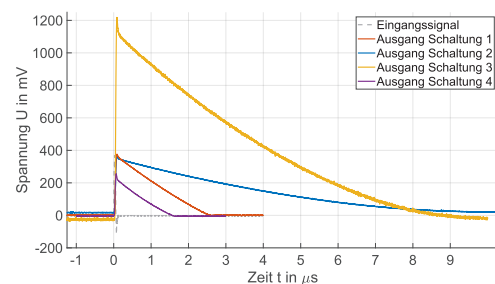


Abb. 8: Ausgangssignale der Platinen mit Eingangsimpuls 1 (250 mV, 50 ns) (Gesamt)

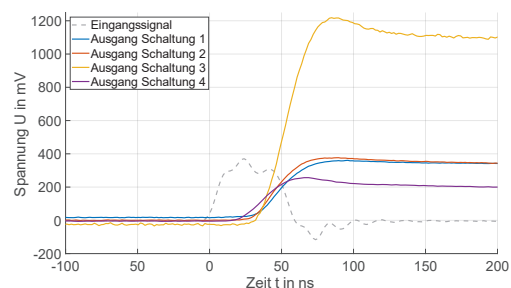


Abb. 9: Ausgangssignale der Platinen mit Eingangsimpuls 1 (250 mV, 50 ns) (Zoom)

Bei dem Eingangsimpuls 1 (vgl. Abbildung 9) folgen Schaltung 1 und 2 dem Signal hochpräzise. Die Referenzschaltung 4 schafft es nicht die Spannung des Eingangsimpulses nachzubilden, der Ausgangsspannungsspeak entspricht rund 80 % des Eingangsspannungsspeaks. Aufgrund der Topologie hebt Schaltung 3 das Signal durch die Verstärkung auf 1200 mV an. Dies begrenzt theoretisch den Eingangsimpuls in seiner Maximalspannung, da die Eingangsspannung des nachgeschalteten ADC begrenzt ist. Die Verstärkung wurde so gewählt, dass

realistische Impulse maximaler Amplitude detektiert werden können.

In der Haltephase überzeugen Schaltung 2 und 3 mit einer Signaldehnung von 8 μs , wobei Schaltung 2 eine minimale Droop-Rate von nur 43 mV/ μs aufweist. Im Gegensatz dazu entlädt sich Schaltung 4 mit 147 mV/ μs rapide und hält den Peak nur 1,5 μs . Schaltung 1 liegt mit 2,5 μs (144 mV/ μs) im Mittelfeld (vgl. Abbildung 8).

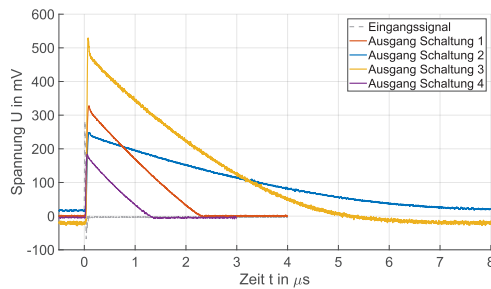


Abb. 10: Ausgangssignale der Platinen mit Eingangsimpuls 2 (250 mV, 16 ns) (Gesamt)

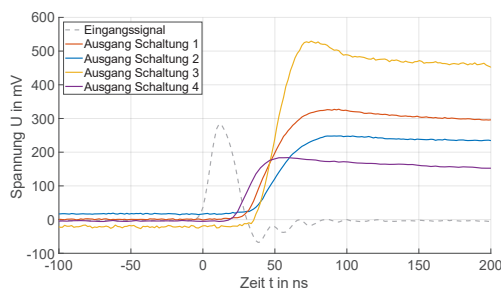


Abb. 11: Ausgangssignale der Platinen mit Eingangsimpuls 2 (250 mV, 16 ns) (Zoom)

Bei einer reduzierten Impulsbreite von 16 ns (Eingangspuls 2) unterliegen die Schaltungen 1, 2 und 4 noch keiner signifikanten Bandbreitenbegrenzung (vgl. Abbildung 11). Ihr transientes Antwortverhalten entspricht qualitativ dem des Standardimpulses 1. Die Haltedauern verkürzen sich lediglich marginal (um 6–13 %) auf 2,3 μs (Schaltung 1), 7,5 μs (Schaltung 2) und 1,3 μs (Schaltung 4).

Eine merkliche Slew-Rate-Limitierung zeigt ausschließlich die lineare Vorverstärkerstufe von Schaltung 3: Die erfasste Ausgangsamplitude sinkt im Vergleich zum Standardimpuls deutlich von 1200 mV auf 550 mV ab, analog reduziert sich die Haltedauer von 8,0 μs auf 5,0 μs . Der verringerte Spannungspegel ist für die Haltephase dennoch weiterhin ausreichend (vgl. Abbildung 10).

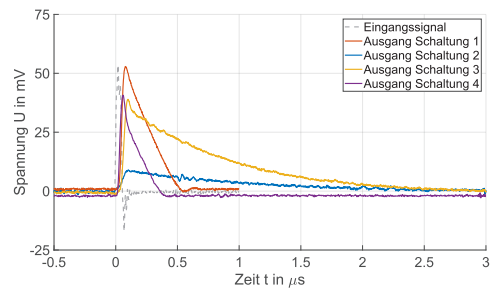


Abb. 12: Ausgangssignale der Platinen mit Eingangsimpuls 3 (50 mV, 50 ns) (Gesamt) (verschoben)

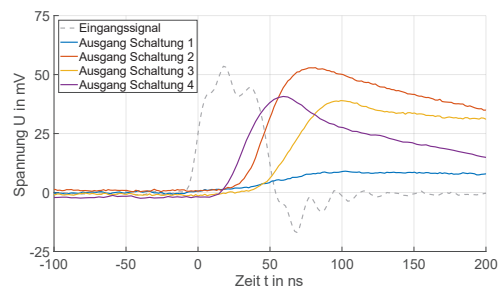


Abb. 13: Ausgangssignale der Platinen mit Eingangsimpuls 3 (50 mV, 50 ns) (Zoom) (verschoben)

Die Untersuchung mit einem stark gedämpften Signal mit geringer Amplitude (Eingangsimpuls 3) offenbart deutliche Limitationen bei allen Topologien. Schaltung 2 ist hier nicht in der Lage, das Signal adäquat zu erfassen und liefert einen Ausgangspeak von ca. 10 mV. Schaltung 1, 3 und die Referenzschaltung 4 können den Peak des Eingangssignals hingegen noch nachbilden. Aufgrund des Spannungsabfalls über der Diode erreicht Schaltung 3 trotz aktiver Verstärkung lediglich eine maximale Ausgangsspannung, die annähernd der Amplitude des Eingangssignals, also eine Verstärkung von rund eins entspricht (vgl. Abbildung 13).

Hinsichtlich der Haltephase zeigt sich ein anderes Bild: Schaltung 2 und 3 schneiden hier mit Haltedauern von 2,0 bis 2,5 μs am besten ab. Während Schaltung 2 den extrem kleinen Peak durch eine weiterhin exzellente Droop-Rate stabil hält, bietet Schaltung 3 in diesem Szenario den besten systemischen Kompromiss aus einer verwertbaren Peak-Spannung und solider Signalhaltezeit. Trotz einer guten initialen Signalerfassung erweisen sich die Schaltungen 1 und 4 in dieser Testreihe als am wenigsten geeignet. Aufgrund hoher Droop-Raten liegen ihre Haltezeiten unter 0,5 μs (vgl. Abbildung 12).

Da die vorangegangenen Versuchsreihen zeigten, dass die evaluierten Topologien bei

geringen Signalamplituden stärker limitiert sind als bei hohen Frequenzanteilen, wurde abschließend die untere Nachweisgrenze untersucht. Als Testsignal diente dabei ein Referenzimpuls mit einer Breite von 100 ns. Das maßgebliche Kriterium ist die minimal erforderliche Eingangsamplitude, bei der die Schaltung das Ausgangssignal für mindestens 1 μ s stabil hält. Diese Haltedauer wurde gewählt, da sie bei der exemplarisch anvisierten Abtastrate eines kostengünstigen Mikrocontroller-ADCs gemäß dem Abtasttheorem eine ausreichende Anzahl an Stützstellen generiert, um den Peak zuverlässig digital erfassen und auswerten zu können.

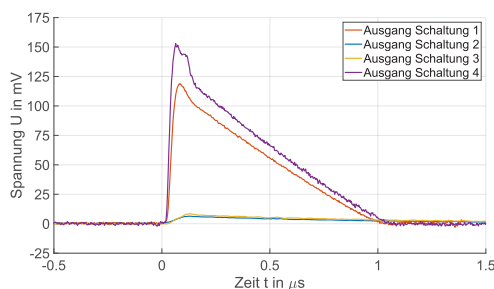


Abb. 14: Ausgangssignale der Platinen mit unterer Grenzspannung (Gesamt) (verschoben)

Die Ergebnisse (vgl. Abbildung 14) bestätigen die Tendenz der Messung der Signale mit kleinen Amplituden deutlich. Die niedrigsten Schwellen zur Detektion weisen Schaltung 3 (10 mV) und Schaltung 2 (25 mV) auf, die in dieser Untersuchung die geringsten Nachweisgrenzen erzielen. Schaltung 1 ordnet sich mit einer Grenzspannung von 90 mV im mittleren Bereich ein. Auffällig ist die Limitierung der Referenzschaltung 4: Mit einer erforderlichen Mindestamplitude des Eingangssignals von 150 mV erweist sie sich für die Erfassung stark gedämpfter Signale als weniger geeignet. Hinsichtlich dieses Parameters wird die Referenz von sämtlichen diskreten Low-Cost-Alternativen deutlich übertroffen.

Abschließender Vergleich

Die primäre Zielsetzung dieser Untersuchung ist die Identifikation eines optimalen Kompromisses aus messtechnischer Performance und ökonomischer Skalierbarkeit für den flächendeckenden Einsatz in Verteilnetzen. Tabelle 1 konsolidiert hierfür die wesentlichen technischen Leistungsparameter aus den vorangegangenen Versuchsreihen und stellt diese den Materialkosten gegenüber (Operationsverstärker und Dioden, bezogen auf prognostizierter Abnahmemengen).

Tabelle 1: Gegenüberstellung der messtechnischen Leistungsparameter und Materialkosten der untersuchten Schaltungstopologien.

Schaltung	Haltedauer (Standard)	Droop-Rate (Standard)	Min. Grenzspannung	Gesamtpreis OPV und Dioden
1	2,5 μ s	144 mV/ μ s	90 mV	9,00 €
2	8,0 μ s	43 mV/ μ s	25 mV	6,50 €
3	5,0 μ s	150 mV/ μ s	10 mV	5,50 €
4	1,5 μ s	147 mV/ μ s	150 mV	11,50 €

Die Gegenüberstellung der untersuchten Topologien zeigt deutliche wirtschaftliche und technische Unterschiede. Die kommerzielle Referenzschaltung 4 stellt mit 11,50 € die kostenintensivste Lösung dar. Aufgrund der höheren Schwelle für die Detektion (150 mV) und der kürzesten Haltedauer (1,5 μ s) erweist sie sich zudem als nachteilig für die Verarbeitung stark gedämpfter Signale durch langsame ADCs. Demgegenüber bieten die diskreten Topologien messbare Vorteile: Schaltung 3 positioniert sich mit 5,50 € als kosteneffizienteste und gleichzeitig Lösung mit der geringsten Schwelle zur Detektion (10 mV), weist jedoch eine stark vom Eingangssignal abhängige Verstärkung auf. Die ausgewogensten Eigenschaften zeigt Schaltung 1: Bei Materialkosten von 9,00 € liefert sie eine mittlere Haltedauer, moderate Schwelle für die Detektion, sowie eine gute Impulsabbildung.

6. Fazit und Ausblick

In dieser Arbeit wurden modifizierte, diskrete PD-Topologien konzipiert und evaluiert, die den Zielkonflikt aus Messgeschwindigkeit und Kosteneffizienz für ein autarkes TE-Monitoring lösen sollen. Die optimierten Standard-OPV-Schaltungen übertreffen teure Spezial-ICs bei kleinen Signalamplituden sowie hinsichtlich der Haltedauer und bilden eine wirtschaftliche Hardware-Basis für großflächige Sensornetzwerke. Zudem birgt die Anpassung des Haltekapazitors weiteres Optimierungspotenzial, um auch aufgrund zu geringer Haltedauer als unzureichend eingestufte Topologien, anwendungsspezifisch zu verbessern.

In der Praxis koppeln induktive HFCT-Sensoren die TE-Signale als Eingangsgröße für die PDCs aus. Für die Anwendbarkeit auf reale Setups ist jedoch kritisch zu reflektieren, dass die Nutzsignale stark durch die überlagerte 50-Hz-

Betriebsspannung sowie Rauschen und Störgrößen maskiert werden können. Die weitere Validierung dieser Erfassungskette erfolgt daher zweistufig: Auf Labortests mit HFCT und Pulskalibrator folgt die praxisnahe Evaluierung an Kabelprüfaufbauten unter echten Stör- und Dispersionsbedingungen. Parallel zur Hardware-Validierung und dem PCB-Design werden die ADC-Abtastung sowie Algorithmen zur Signalmustererkennung auf einem Mikrocontroller implementiert.

Literaturnachweis

- [1] BDEW German Association of Energy and Water Industries, Zukünftige Rolle des Verteilnetzbetreibers in der Energiewende.
- [2] Agora Verkehrswende, Agora Energiewende, Regulatory Assistance Project, Verteilnetzausbau für die Energiewende – Elektromobilität im Fokus.
- [3] B. Bundesnetzagentur, Monitoringbericht 2025, <https://www.bundesnetzagentur.de/DE/Fachthemen/ElektrizitaetundGas/Monitoringberichte/start.html>.
- [4] C. Mayoux, Degradation of insulating materials under electrical stress, *IEEE Transactions on Dielectrics and Electrical Insulation* 7, 590–601 (2000); doi: 10.1109/TDEI.2000.879355.
- [5] G. C. Montanari, L. Simoni, Aging phenomenology and modeling, *IEEE Transactions on Electrical Insulation* 28, 755–776 (1993); doi: 10.1109/14.237740.
- [6] S. Boggs, J. Densley, Fundamentals of partial discharge in the context of field cable testing, *IEEE Electrical Insulation Magazine* 16, 13–18 (2000); doi: 10.1109/57.871416.
- [7] G. C. Stone, A. Cavallini, G. Behrmann, C. Serafino, *Practical partial discharge measurement on electrical equipment*. IEEE Press; Wiley, Piscataway, NJ, Hoboken, New Jersey; 2023.
- [8] M. Fritsch, C. Andres, M. Wolter, Review on peak detect and hold circuits and their applicability in partial discharge detection, *at - Automatisierungstechnik* 72, 1041–1051 (2024); doi: 10.1515/auto-2024-0101.
- [9] IEC International Electrotechnical Commission Std., IEC 60270:2000: High-voltage Test Techniques - Partial Discharge Measurements (2000).
- [10] J. Shu, Y. Du, H. Zhao, K. Liu, J. Qiu, Modeling, Prediction, and Results Correction of PDSH Circuits for Nanosecond Pulse Peak Detection, *Applied Sciences* 13, 12597 (2023); doi: 10.3390/app132312597.
- [11] G. de Geronimo, P. O'Connor, A. Kandasamy, Analog CMOS peak detect and hold circuits. Part 1. Analysis of the classical configuration, *Nuclear Instruments & Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment* (2002); doi: 10.1016/S0168-9002(01)02059-9.
- [12] M. Kruiskamp, D. Leenaerts, A CMOS peak detect sample and hold circuit (1994); doi: 10.1109/23.281513.
- [13] H. Zheng, R. Ma, X. Wang, M. Liu, Z. Zhu, A CMOS Peak Detect and Hold Circuit With Auto-Adjust Charging Current for NS-Scale Pulse ToF Lidar Application, *IEEE Transactions on Circuits and Systems Part 1: Regular Papers* (2020); doi: 10.1109/TCSI.2020.3001168.
- [14] M. Fritsch, *Measurement of Partial Discharges on Power Cables*. Otto-von-Guericke University Magdeburg; 2024.
- [15] P. Horowitz, W. Hill, *The art of electronics*. Cambridge University Press, Cambridge, New York; 2024.
- [16] U. Tietze, C. Schenk, *Halbleiter-Schaltungstechnik*. Springer Berlin Heidelberg, Berlin, Heidelberg; 1985.
- [17] L. Vetrivel, B. M. Sivaram, Computer-controlled high-speed peak detector for use with pulsed lasers, *Review of Scientific Instruments* 66, 2394–2399 (1995); doi: 10.1063/1.1145636.
- [18] National Semiconductor (TI), Datasheet LF13741: Monolithic JFET Input Operational Amplifier, <https://www.alldatasheet.com/html-pdf/115354/NSC/LF13741/654/11/LF13741.html>. Accessed February 15, 2026.
- [19] R. Elliott, Peak Detection Circuits, <https://sound-au.com/appnotes/an014.htm>. Accessed February 15, 2026.

Nullfeld-Magnetfeldkamera

Ronja Rasser^{1,2}, Peter A. Koss¹, Svenja Knappe^{3,4} und Karsten Buse^{1,2}

¹*Fraunhofer Institut für physikalische Messtechnik IPM, Georges-Köhler-Allee 301, 79110 Freiburg im Breisgau, Deutschland*

²*Laboratory for Optical Systems, Institut für Microsystems Engineering IMTEK, University of Freiburg, Georges-Köhler-Allee 102, 79110 Freiburg im Breisgau, Deutschland*

³*Paul M. Rady Department of Mechanical Engineering, University of Colorado, Boulder, Colorado, 1111 Engineering Drive, 80309, USA*

⁴*FieldLine Inc., Boulder, 371 Centennial Parkway, Louisville, Colorado 80027, USA*

Zusammenfassung

Optisch gepumpte Magnetometer (OPMs) ermöglichen hochempfindliche Magnetfeldmessungen im Pikotesla-Bereich ohne kryogene Kühlung. In dieser Arbeit wird eine abbildende Magnetfeldkamera vorgestellt, die auf einem optisch gepumpten Nullfeldmagnetometer basiert und räumlich aufgelöste Magnetfeldmessungen erlaubt. Das System nutzt eine optische Lock-in-Kamera mit über 270.000 Pixeln zur bildgebenden Detektion des einem Lichtstrahl aufgeprägten Magnetfeldsignals einer MEMS-gefertigten Rubidiumdampfzelle. Ein speziell entwickeltes nichtmagnetisches Baukastensystem ermöglicht die Montage und Justage der Optik innerhalb der magnetischen Abschirmung. Erste Charakterisierungen zeigen Bandbreiten von bis zu 560 Hz und eine Rauschdichte von 270 pT/√Hz bei einer Mittelung über 20×20 Pixel. Diese Mittelung entspricht einem Dampfvolumen von 80 μm × 80 μm × 300 μm. Im Ausblick wird ein neuartiges Abschirmkonzept vorgestellt: Eine kompakte Spritzbeschichtungs-Abschirmung, die nur die Dampfzelle und die zu vermessende Probe umschließt. Dies ermöglicht zukünftig eine deutliche Reduzierung des Volumens von Systemen mit Nullfeldmagnetometern. Die Kombination aus orts aufgelöster Magnetfeldmessung und kompakter Bauweise eröffnet neue Anwendungsmöglichkeiten in der zerstörungsfreien Prüfung und Materialcharakterisierung.

Keywords: Magnetometer, Kamera, Magnetische Abschirmung, Magnetfeldbildgebung

Einleitung

Die hochempfindliche Messung magnetischer Felder ist in zahlreichen Anwendungsbereichen von Bedeutung – von der biomedizinischen Diagnostik über die zerstörungsfreie Prüfung bis zur Grundlagenforschung. Optisch gepumpte Magnetometer (OPMs) haben sich in den letzten Jahren als vielversprechende Alternative zu supraleitenden Quanteninterferenz-Detektoren (SQUIDs) etabliert, da sie vergleichbare Empfindlichkeiten erreichen können, ohne auf kryogene Kühlung angewiesen zu sein [1,2]. Neben Einzelpunktmessungen gewinnt die orts aufgelöste Abbildung von Magnetfeldern zunehmend an Bedeutung. Durch die Verwendung einer einzelnen großflächigen Dampfzelle in Kombination mit Mehrkanaldetektoren können räumliche Magnetfeldverteilungen simultan erfasst werden [3,4].

Das Funktionsprinzip eines OPMs basiert auf der Wechselwirkung zwischen polarisiertem Licht und einem atomaren Dampf, typischerweise Rubidium oder Cäsium. Ein

zirkular polarisierter Laserstrahl wird durch ein mit Alkalidampf gefülltes Volumen geleitet. Das optische Pumpen versetzt die Atome in einen spinpolarisierten Zustand: Durch die Absorption von Photonen wird der Drehimpuls des Photons auf die Alkaliatome übertragen und orientiert den atomaren Spin entlang der Strahlrichtung.

In Anwesenheit eines externen Magnetfeldes präzedieren die atomaren Spins um die Magnetfeldrichtung (Larmor-Präzession). Diese Präzession verändert die optischen Eigenschaften des Dampfes. Durch Beobachtung der Transmission kann das Magnetfeld gemessen werden. Alternativ kann die Drehung der Polarisation eines zweiten schwächeren linearpolarisierten Lasers zum Messen des Magnetfeldes genutzt werden.

Eine besonders empfindliche Betriebsart ist das Nullfeldmagnetometer. Hierbei wird das Magnetometer in einer Umgebung betrieben, in der das Hintergrundmagnetfeld auf wenige Nanotesla kompensiert ist. In diesem Regime werden störende Spin-Austausch-Relaxationsprozesse unterdrückt, was zu einer drastisch erhöhten Empfindlichkeit führt [5].

Die Notwendigkeit eines nahezu feldfreien Bereichs stellt jedoch eine zentrale Herausforderung dar: Das Erdmagnetfeld von etwa $50 \mu\text{T}$ muss um mehrere Größenordnungen abgeschirmt werden. Konventionelle Lösungen verwenden großvolumige, mehrlagige Abschirmkammern aus hochpermeablem Material (Mu-Metall), was die Systemgröße und -kosten erheblich beeinflusst.

Aufbau der Magnetfeldkamera

Die entwickelte Magnetfeldkamera erweitert das Prinzip des OPMs um die Fähigkeit zur orts aufgelösten Magnetfeldmessung. Anstelle eines einzelnen Detektors wird eine Lock-in-Kamera (Heliotis C4.0-S40U) mit 512×542 Pixeln eingesetzt. Da jeder Pixel das Magnetfeldsignal eines kleinen, ihm zugeordneten Bereichs des projizierten Dampfolumens detektiert, entsteht eine zweidimensionale Magnetfeldkarte. In die Dampfzelle wird zusätzlich Stickstoff gefüllt, um die Diffusion der Rubidiumatome zu verlangsamen und dadurch die Mischung der Signale benachbarter Pixel zu vermeiden.

Abbildung 1 zeigt den Aufbau des optischen Systems. Der Laserstrahl wird aus einer optischen Faser ausgekoppelt und auf einen Durchmesser von etwa 8 mm kollimiert. Nach Passage durch einen polarisierenden Strahlteiler und einer $\lambda/4$ -Wellenplatte wird das Licht zirkular polarisiert. Der Strahl durchläuft die Rubidiumdampfzelle, wird an einem rückseitigen dielektrischen Spiegel reflektiert und passiert die Zelle ein zweites Mal (Doppelpass-Konfiguration). Dies verdoppelt die effektive Wechselwirkungslänge. Nach erneuter Passage durch Wellenplatte und Strahlteiler wird der Strahl auf die Lock-in-Kamera abgebildet. Die Kamera befindet sich aufgrund ihres Magnetfelds außerhalb der Abschirmung. Die Koordinatenachsen sind so definiert, dass sich der Laserstrahl entlang der z-Achse ausbreitet. Das Magnetfeld wird in der xy-Ebene gemessen.

Da sich das optische System innerhalb der magnetischen Abschirmung befindet, müssen alle Komponenten nichtmagnetisch sein. Standardoptikhalterungen aus Stahl sind aufgrund des Eisengehalts ungeeignet. Auch Aluminiumlegierungen können problematisch sein, da große Metallvolumina durch Wirbelströme magnetisches Rauschen erzeugen können. Daher wurde ein modulares Baukastensystem aus gefrästen und 3D-gedruckten Kunststoffhalterungen entwickelt. Abbildung 2 zeigt ein Foto des optischen Systems, welches aus nichtmagnetischen Bauteilen konstruiert wurde.

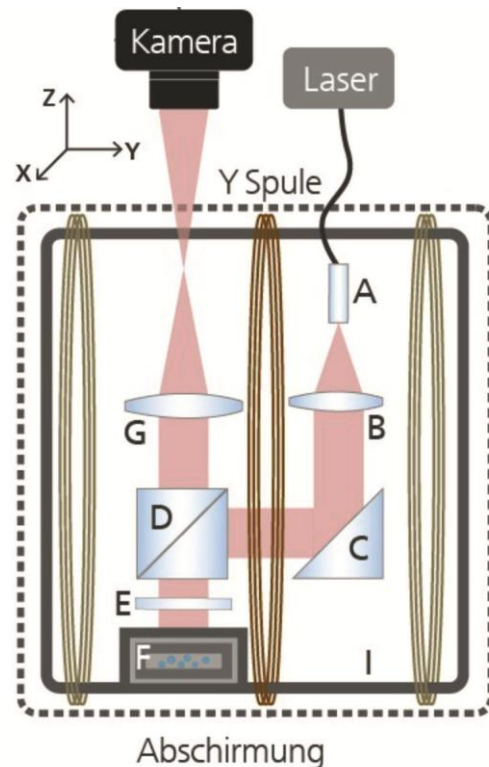


Abb. 1: Optisches System: Faserkoppler (A), Kollimationslinse (B), Spiegel (C), polarisierender Strahlteiler (D), $\lambda/4$ -Wellenplatte (E), Dampfzelle im Gehäuse (F), Abbildungslinse (G) und Halterungen für die Abschirmungsmontage (I).

Das optische System umfasst Halterungen für Linsen, Strahlteiler, Wellenplatten sowie einen Adapter für die optische Faser. Die Halterungen werden mit 6-mm-Stäben verbunden und sind kompatibel zum 30-mm-Cage-System, was die Verwendung von Standard-Ausrichtungswerkzeugen ermöglicht. Vor dem Einbau wurden alle Bauteile in einer geschirmten Kammer mit einem Fluxgate-Magnetometer (Auflösung: $0,1 \text{ nT}$) geprüft. Bauteile, die eine messbare Feldänderung im Nanoteslabereich erzeugten, wurden aussortiert.

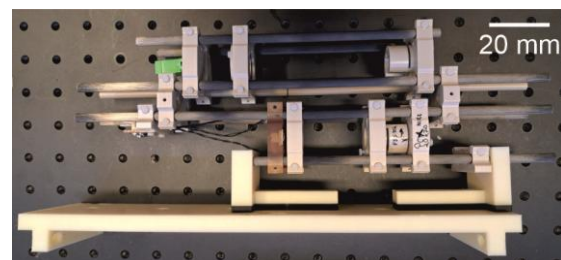


Abb. 2: Foto des nichtmagnetischen optischen Systems

Die Rubidiumdampfzelle wurde mittels MEMS-Technologie gefertigt. Sie besitzt eine nutzbare Fläche von 4 mm Durchmesser. Ein kleiner Abstand zu den zu untersuchenden Objekten und eine geringe Dicke der Zelle sind essenziell für eine gute räumliche Auflösung von Magnetfeldern. Daher wurde die Dicke der Dampfzelle auf 300 μm reduziert. Ein Kamerapixel misst die Magnetfeldreaktion eines Dampf Volumens von etwa $8\text{ }\mu\text{m} \times 8\text{ }\mu\text{m} \times 300\text{ }\mu\text{m}$. Die Zelle wird auf 120 – 160 °C Betriebstemperatur geheizt, um einen ausreichenden Rubidiumdampfdruck zu erzeugen. Die Heizung erfolgt mit einem 100-kHz-Wechselstrom, dessen hohe Frequenz außerhalb des Messbereichs liegt und somit wenige Störsignale verursacht.

Die magnetische Abschirmung besteht aus vier konzentrischen Mu-Metall-Zylindern mit beidseitigen Deckeln [6]. Der Schirmfaktor beträgt im Bereich 0–10 Hz etwa $3\text{--}4 \times 10^6$. Integrierte Spulen ermöglichen die Kompensation des Restfeldes sowie das Anlegen definierter Referenz- und Gradientenfelder. Für die Magnetfeldmessung wird ein Modulationsverfahren eingesetzt: Die zu messende Feldkomponente (Y-Komponente) wird mit einer definierten Frequenz moduliert und das resultierende Signal mittels Lock-in-Technik ausgewertet. Die Lock-in-Kamera führt diese Demodulation parallel für alle Pixel durch.

Erste Ergebnisse

Das System zeigt ein Tiefpassverhalten mit einer 3-dB-Grenzfrequenz von 390 Hz bei einer Modulationsfrequenz von 1330 Hz bzw. 560 Hz bei 2330 Hz Modulationsfrequenz. Die Rauschdichte verbessert sich durch Zusammenfassen mehrerer Pixel („Binning“) proportional zu $1/n$, wobei n die Anzahl der Pixel pro Bin ist. Dies deutet auf räumlich und zeitlich unkorreliertes Rauschen hin. Bei einer Binngröße von 20×20 Pixeln, bzw. Dampf volumen von $80\text{ }\mu\text{m} \times 80\text{ }\mu\text{m} \times 300\text{ }\mu\text{m}$, wird eine Rauschdichte von 270 pT/ $\sqrt{\text{Hz}}$ erreicht. Die Kamera kann räumliche Magnetfeldgradienten auflösen. Messungen eines oszillierenden Gradienten zeigen gute Übereinstimmung mit Simulationsergebnissen der verwendeten Gradientenspule.

Ausblick: Kompakte Spritzbeschichtungs-Abschirmung

Konventionelle magnetische Abschirmungen für OPMs umschließen das gesamte optische Messsystem. Dies führt zu großen Bauvolumina und schränkt die Anwendungsmöglichkeiten ein. Daher entwickeln wir derzeit ein neuartiges

Abschirmkonzept: Eine kompakte Abschirmung, die nur die Dampfzelle und die zu vermessende Probe umgibt (Abbildung 3).

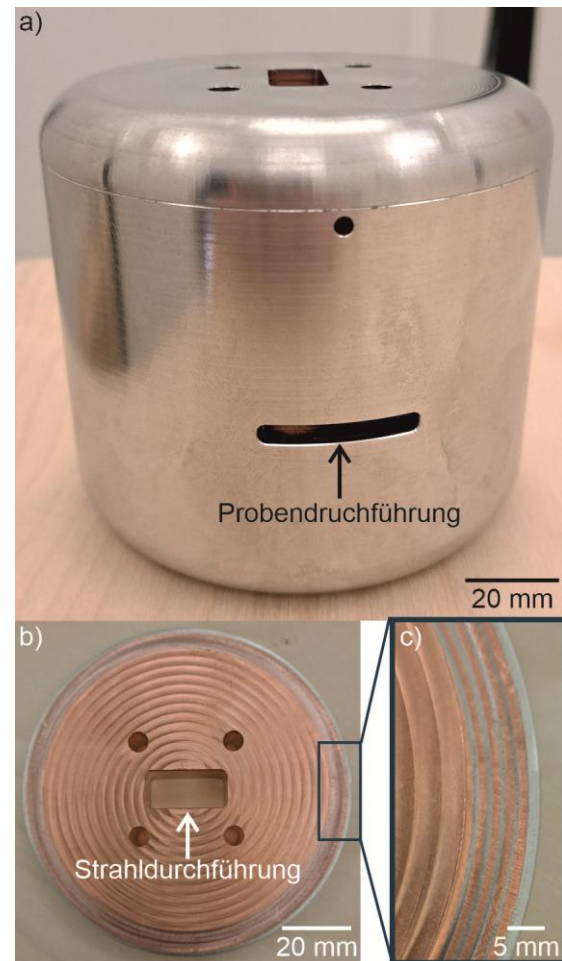


Abb. 3: Kompakte Abschirmung mit Spritzbeschichtung (a). Deckel der Abschirmung von innen (b) und Querschnitt durch Beschichtungen (c). Die Abschirmschichten werden auf einen Kupferträger aufgebracht und umschließen nur Dampfzelle und Probe.

Die Abschirmschichten werden mittels thermischer Spritzbeschichtung aufgebracht. Dieses Verfahren ermöglicht die Beschichtung komplexer Geometrien und bietet Flexibilität bei der Materialwahl und Schichtdicke. Der Schichtaufbau besteht aus vier alternierenden Lagen von 600 μm Reineisen und 1000 μm Reinkupfer. Zum Abschluss wurde eine Schicht 1.4404 Stahl der Dicke 1500 μm aufgetragen.

Die ferromagnetischen Schichten sorgen für die magnetische Abschirmung bei niedrigen Frequenzen durch Flussführung. Die leitfähigen Zwischenschichten tragen durch Wirbelstromeffekte zur Abschirmung bei höheren Frequenzen bei. Die Stahlschicht

dient in erster Linie als Schutzschicht für die Schirmung.

Erste Untersuchungen an Testproben zeigen vielversprechende Ergebnisse [7]: Spritzbeschichtete Eisenschichten auf Aluminiumsubstrat weisen einen signifikanten Abschirmeffekt auf. Die Abschirmwirkung verbessert sich mit zunehmender Schichtdicke und zeigt oberhalb von 100 Hz durch den Beitrag des leitfähigen Substrats eine deutliche Steigerung.

Zusammenfassung

Es wurde eine Magnetfeldkamera auf Basis eines optisch gepumpten Nullfeldmagnetometers entwickelt. Die Kombination aus Lock-in-Kamera-Detektion und nichtmagnetischem Optikkaukasten ermöglicht orts aufgelöste Magnetfeldmessungen mit hoher Pixelzahl. Im nächsten Entwicklungsschritt wird die konventionelle Abschirmung durch eine kompakte Spritzbeschichtungs-Abschirmung ersetzt. Diese umschließt nur die Dampfzelle und Probe, was das Systemvolumen deutlich reduziert und neue Anwendungsmöglichkeiten in der zerstörungsfreien Prüfung eröffnet.

Literaturnachweis

- [1] Budker, D., Romalis, M. Optical magnetometry. *Nature Phys* **3**, 227–234 (2007). DOI: 10.1038/nphys566
- [2] Kominis, I., Kornack, T., Allred, J. *et al.* A subfemtotesla multichannel atomic magnetometer. *Nature* **422**, 596–599 (2003). DOI: 10.1038/nature01484
- [3] Young Jin Kim, Igor Savukov, Shaun Newman; Magnetocardiography with a 16-channel fiber-coupled single-cell Rb optically pumped magnetometer. *Appl. Phys. Lett.* 8 April 2019; 114 (14): 143702. DOI: 10.1063/1.5094339
- [4] Rasser, R., Koss, P.A., Knappe, S. *et al.* Towards a multi-channel zero-field optically pumped magnetometer for imaging. *EPJ Quantum Technol.* **12**, 101 (2025). DOI: 10.1140/epjqt/s40507-025-00400-y
- [5] T. Tierney, N. Holmes, S. Mellor *et al.* Optically pumped magnetometers: From quantum origins to multi-channel magnetoencephalography, *Neuroimage* **199**, 598-608 (2019), DOI: 10.1016/j.neuroimage.2019.05.063.
- [6] Twinleaf MS-2 Magnetic Shield, <https://twinleaf.com/shield/MS-2/> Abgerufen: 10/04/2026
- [7] P. A. Koss *et al.*, Designing magnetic shielding for space applications using spray coating technology, International Astronautical Congress, IAC-24,C2,6,3,x82727 (2024)

Identification of Bacteria by Pyrolysis-Gas Chromatography-Ion Mobility Spectrometry

Daniel Röckrath¹, Moritz Hitzemann¹, Alejandra Vargas-Valderrama^{2,3}, Arne Höhne¹, Philip Timmermann¹, Stefan Zimmermann¹

¹Leibniz University Hannover, Institute of Electrical Engineering and Measurement Technology, Appelstr. 9A, 30167 Hannover, Germany

²Department of Chemistry, Royal Military Academy, Avenue de la Renaissance 30, 1000 Brussels Belgium

³Belgian Defence Laboratories (DLD), Quartier Major Housiau, Martelarenstraat 181, 1180 Vilvoorde

+49 511 762 4672, daniel.roeckrath@geml.uni-hannover.de

Abstract

This work presents a novel approach for rapid and reliable analysis of biological samples, crucial in medical, safety, and security sectors. The self-built device combines a pyrolyzer with a gas chromatograph and an ultra-fast ion mobility spectrometer. This portable system allows sensitive analysis of liquid and solid samples on-site, including viruses and bacteria. The sample undergoes pyrolysis under an inert atmosphere, breaking down larger non-volatile compounds into their volatile constituents. The gas chromatograph separates the constituents, while the ion mobility spectrometer provides an orthogonal dimension of separation and detection at the highest sensitivity. The method effectively differentiates between bacteria such as *Streptococcus Mitis* and *Streptococcus Suis*.

Keywords: Pyrolysis, GC, IMS, Bacteria, GC-IMS

The rapid and reliable analysis of biological samples is of utmost importance in various areas of interest, particularly in the medical, safety, and security sectors. We present a new approach using a self-built device that combines a low-cost pyrolyzer with a gas chromatograph and an ultra-fast polarity switching ion mobility spectrometer. This approach is field deployable and allows for rapid and sensitive analysis of liquid and solid samples, including viruses and bacterial pathogens. The sample is heated rapidly under an inert gas atmosphere in the absence of reactive gases, such as oxygen. This process, known as pyrolysis, transfers the solid or liquid sample into the gas phase while dissociating larger non-volatile compounds into more volatile constituents. The gas chromatograph then separates the complex sample matrix, and the eluting substances are detected using a dual-polarity ion mobility spectrometer, adding a second dimension of separation. The practical use case is the identification of bacteria, as shown with the two topographical plots for *Streptococcus Mitis* and *Streptococcus Suis* in **Fehler! Verweisquelle konnte nicht gefunden werden.** As is visible,

the two different bacteria of the same class can be clearly differentiated.

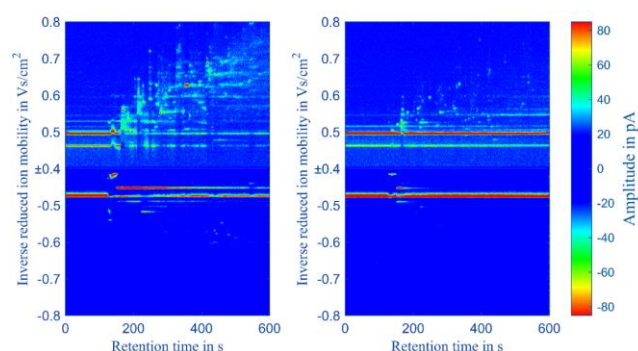


Figure 1. 2D-chromatogram of *Streptococcus Mitis* (left) and *Streptococcus Suis* (right) measured with the PY-GC-IMS at a concentration of 4×10^8 cfu/ μ l.

Acknowledgments

Funded by the European Union. Views and opinions expressed are those of the author(s) and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them. This project has received funding from the

European Defence Fund under grant
agreement No 101103176 (TeChBioT).

Label-Free Electrochemical Biosensor for the Detection of Glycophorin A (CD235a)

Sahar M. Hesari¹, Charlotte von Hecke², Juan Nicolas Bejarano Rios², Liudmila Deckert¹, Ulrike Steinmann¹

¹Chair Measurement Technology, OVGU Magdeburg, 39106 Magdeburg, Germany

²Core Facility Tissue Engineering, OVGU Magdeburg, 39106 Magdeburg, Germany
sahar.mohammadkhan@ovgu.de, charlotte.vonheckel@ovgu.de, juan.bejarano@ovgu.de
liudmila.deckert@ovgu.de, ulrike.steinmann@ovgu.de

Abstract

Electrochemical biosensors offer an alternative method for identification of specific proteins in a biological sample by producing an electrical signal that is dependent on molecular recognition. Herein, we present a label-free electrochemical biosensor that focuses on Glycophorin A (CD235a). The biosensor is based on gold microelectrodes which were functionalized with a self-assembled monolayer (SAM) of mercaptoundecanoic acid (MUA), then activated with EDC/NHS. Non-specific binding on the oriented immobilized anti-CD235a antibodies was blocked using bovine serum albumin (BSA). These sequential surface modifications yielded a very low characteristic current by cyclic voltammetry (CV): The peak current significantly decreased by 93.4% (anodic) and 96.9% (cathodic) between the bare electrode and BSA-blocked surface. Analysis of the impedance spectra revealed a charge transfer resistance (R_{ct}) of 6 k Ω for bare electrode. This increased significantly to 434 k Ω following the detection of CD235a, indicating successful target binding. This substantial increase in R_{ct} confirms the biosensor's high sensitivity. Furthermore, the biosensor demonstrated high specificity towards the CD235a surface antigen.

Keywords: biosensor, electrochemistry, glycophorin A, impedance spectroscopy

Introduction

Identifying the presence of specific proteins in a biological sample is an essential task for both basic and applied research. Although classical analytical methods such as flow cytometry, Enzyme-Linked Immunosorbent Assay (ELISA) and western blotting are accurate and can conduct multi-parameter analysis, they have several disadvantages, such as invasiveness and time-consuming sample preparation. They are also not useful for continuous monitoring, since they only provide discrete snapshots of the status [1, 2]. This is critical for dynamic bioprocess control, where continuous data is required.

Electrochemical biosensors offer a promising approach to overcoming these limitations, as they allow for label-free, sensitive, real-time detection [3]. They convert molecular recognition events into measurable electrical signals, providing quantitative information about the concentration of the target analyte. Such sensors capable of continuously capturing biological signals are important for systems that rely on feedback-driven adaptation. This includes biological-technology interfaces, where both processes interact in a dynamic and reciprocal manner. Within the

concept of co-evolution, innovative technologies and sensor-based approaches which are able to adapt to alterations in biological systems, lay the groundwork for self-learning systems.

A crucial step for electrochemical sensors is the proper immobilization of biorecognition elements (typically antibodies) to guarantee their effectiveness, as this retains their specificity for target analytes and ensures they are oriented in an optimal position for binding. Self-assembled monolayers (SAMs) of organothiols on gold electrodes, coupled using carbodiimide chemistry (EDC/NHS), are a well-established platform for developing electrochemical immunosensors. This method enables the stable covalent attachment of antibodies while maintaining bioactivity and oriented immobilization [4, 5].

Here, we describe the development and characterization of a label-free electrochemical biosensor targeting glycophorin A (CD235a), a major transmembrane sialoglycoprotein of erythrocytes which is widely used as a marker of commitment to an erythroid lineage [6, 7]. Glycophorin A is expressed at a high density on mature red blood cells (RBCs) (approximately one million copies per cell).

To the best of our knowledge, this study presents the first practical demonstration of a label-free electrochemical immunosensing platform for the detection of CD235a.

2. Materials and Methods

2.1 Materials and Reagents

Screen-printed gold electrodes (DRP-C223BT-U75, Metrohm, Germany) served as substrates for surface functionalization. The following chemicals were used: mercaptoundecanoic: mercaptoundecanoic acid (MUA, 95%), 2-(N-Morpholino) ethanesulfonic acid (MES, $\geq 99\%$), 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC, $\geq 99.0\%$), N-hydroxysuccinimide (NHS, 98.0%), phosphate-buffered saline (PBS, pH 7.4), potassium hexacyanoferrate (III) ($K_3[Fe(CN)_6]$), and potassium hexacyanoferrate (II) trihydrate ($K_4[Fe(CN)_6] \cdot 3H_2O$) all purchased from Sigma-Aldrich (Merck, Germany). A recombinant rabbit monoclonal anti-CD235a antibody (clone 27A8, Thermo Fisher Scientific, USA) was selected as the biorecognition element and Bovine serum albumin (BSA) was applied as a blocking agent, and potassium chloride (KCl) functioned as a supporting electrolyte. The electrodes were cleaned with absolute ethanol ($\geq 99.8\%$). The recombinant human Glycophorin A protein (Sino Biological Europe GmbH) served as the target biomarker (antigen) in this study. All aqueous solutions were prepared using deionized (DI) water, and all chemicals were used without further purification.

2.2 Electrode Preparation and Surface Functionalization

The gold microelectrodes were cleaned by rinsing them with ethanol, followed by a thorough wash with ultrapure water, then drying them under a stream of nitrogen. Surface functionalization was carried out using a sequential modification strategy (see Figure 1). First, a self-assembled monolayer (SAM) structure was formed by immersing the cleaned electrodes in a 1 mM solution of mercaptoundecanoic acid (MUA) in ethanol for 18 hours at room temperature in the dark to promote the formation of a well-organized monolayer. The thiol group of MUA creates stable Au-S bonds with the gold surface, and the terminal carboxyl groups serve as functional sites for further antibody coupling. The electrodes were then washed with ethanol, and ultrapure water, and dried under nitrogen.

The carboxyl groups were activated using a freshly prepared EDC/NHS solution (200 mM EDC and 50 mM NHS in MES at pH 6) by immersing the electrodes in it for 30 minutes at room temperature. This carbodiimide coupling chemistry generates reactive NHS esters from the carboxyl groups, which then react with the primary amines of the antibody's lysine residues to form stable amide bonds. After washing with PBS to remove any remaining reagents, 20 μ L of an anti-CD235a solution (100 μ g/mL in PBS) was dropped onto the electrode surface and left to incubate for 2 hours at room temperature. This approach enables the stable, covalent immobilization of antibodies on the electrode surface. The electrodes were then rinsed with PBS to remove any unbound antibody.

To block non-specific binding sites, the electrodes were incubated in 1% BSA in PBS for 30 minutes, afterwards they were rinsed three times with PBS. The BSA then acts as a passivating agent for any residual reactive sites, reducing non-specific protein adsorption in the next step of antigen detection. Subsequently, 20 μ L of recombinant human Glycophorin A protein (0.1 ng/mL in PBS) was pipetted onto the working electrode surface and incubated for 30 minutes at room temperature. After incubation, the electrodes were rinsed three times with PBS to remove any unbound antigen before electrochemical measurements.

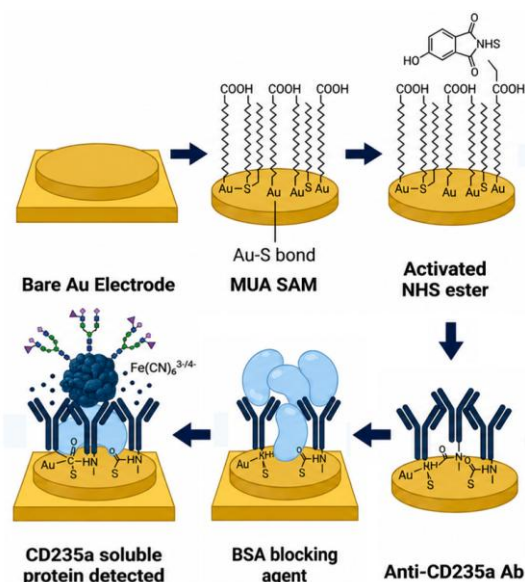


Figure 1: A schematic diagram of the Glycophorin A modification method and its detection.

2.3 Electrochemical Characterization

Electrochemical measurements were performed using a Gamry Interface 1010E potentiostat (Gamry Instruments, USA) in a three-electrode configuration consisting of a 1.6 mm diameter screen-printed gold working electrode (WE), a gold counter electrode (CE), and a Ag/AgCl pseudo-reference electrode. Measurements were performed at room temperature in a 5 mM $[\text{Fe}(\text{CN})_6]^{3-/4-}$ solution prepared in PBS (pH 7.4).

Cyclic voltammetry (CV) was conducted over a potential range of -0.2 to $+0.6$ V versus Ag/AgCl. EIS measurements were performed at a DC potential of $+0.23$ V (vs. Ag/AgCl) which is the formal potential of the redox couple, with an AC amplitude of 10 mV, over a frequency range of 100 kHz to 0.01 Hz. Echem Analyst software (Gamry Instruments) was used to analyze the impedance data and fit the R_{ct} values obtained from a Randles equivalent circuit model.

3. Results

3.1 Cyclic Voltammetry Analysis

Cyclic voltammetry measurements were performed using the $[\text{Fe}(\text{CN})_6]^{3-/4-}$ redox probe to evaluate the changes in electron transfer properties during the stepwise modification of the electrode surface (see Figure 2). The bare gold electrode produced well-defined, and quasi-reversible redox peaks with high current intensity (anodic peak current $I_{pa} = +46.0$ μA and cathodic peak current $I_{pc} = -41.8$ μA) as well as small peak-to-peak separation. Following EDC/NHS-based activation of the MUA SAM, the peak current decreased considerably ($I_{pa} = +8.4$ μA and $I_{pc} = -8.8$ μA), representing an 81.7% and 78.9% decrease in the anodic and cathodic currents, respectively compared to the bare electrode. This resulted in a marked increase in peak separation. Additionally, immobilizing the anti-CD235a antibody decreased the peak currents to $I_{pa} = +7.0$ μA and $I_{pc} = -8.7$ μA (reductions of 84.9% and 79.1%, respectively), compared to the bare electrode, followed by antibodies immobilization.

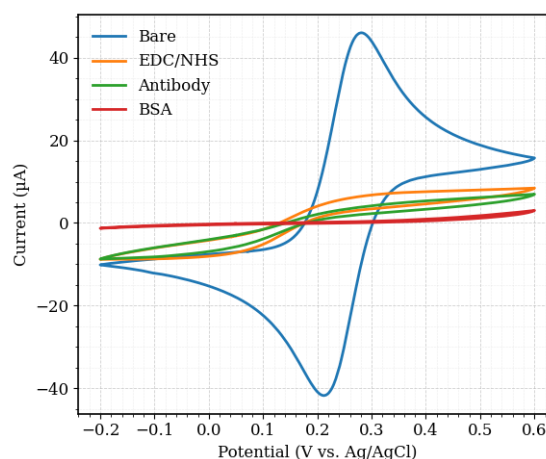


Figure 2: Cyclic voltammograms with modification of the electrode surface from bare gold through the formation of a self-assembled monolayer, the activation with EDC/NHS, the immobilization of antibodies and the blocking with BSA.

The final BSA blocking yielded the smallest measured currents ($I_{pa} = +3.0$ μA , $I_{pc} = -1.3$ μA), representing total reductions of 93.4% and 96.9% from the values for the bare electrode, respectively.

3.2 Electrochemical Impedance Spectroscopy

Electrochemical impedance spectroscopy (EIS) was utilized to analyze the changes in interfacial charge transfer resistance that occurred during the consecutive functionalization steps. This provided complementary quantitative data (see Figure 3). The Nyquist plots ($-Z''$ versus Z') showed an increase in semicircle diameter as the charge transfer resistance (R_{ct}) values increased. In Nyquist plots, the semicircle diameter is proportional to R_{ct} , reflecting resistance to electron transfer across the electrode–electrolyte interface. For the bare gold electrode, only a small semicircle with a low R_{ct} value of 6 k Ω appeared, due to the fast electron transfer occurring on the clean metal surface. Following MUA SAM formation (see Figure 3), the R_{ct} increased to 39 k Ω . Subsequent to EDC/NHS activation of the carboxyl-terminated SAM, R_{ct} increased further to 74 k Ω . The immobilization of the anti-CD235a antibody led to an increase in R_{ct} to 151 k Ω .

After BSA blocking, the R_{ct} increased slightly to 342 k Ω .

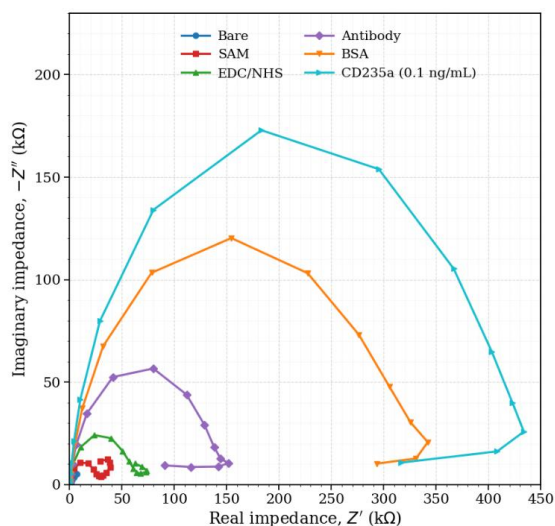


Figure 3: Nyquist plots with impedance changes during sequential electrode functionalization and CD235a detection.

3.3 Biosensor Performance for CD235a Detection

A continued increase in R_{ct} to 434 kΩ was observed after BSA-blocked electrodes were incubated with a 0.1 ng/mL solution of CD235a antigen in PBS (see Figure 4). This increase reflects the impact of antigen binding on interfacial resistance. An 27% increase in R_{ct} relative to the BSA-blocked state (corresponding to an absolute change of 92 kΩ) indicates the presence of specific antigen–antibody interactions. Although this increase is small compared to previous functionalization steps, it demonstrates molecular recognition between the immobilized anti-CD235a antibodies and their cognate antigen and confirming the sensitivity of the sensor toward low antigen concentrations.

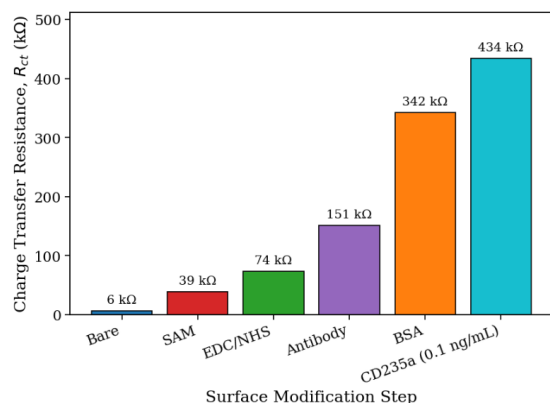


Figure 4: The charge transfer resistance (R_{ct}) values at each functionalization step.

4. Discussion

4.1 Surface Functionalization Strategy

The bare gold electrode exhibited quasi-reversible redox behavior with high peak currents and a small peak-to-peak separation (ΔE_p), indicating efficient electron transfer and a clean, electrochemically active gold surface. Such a surface provides an ideal foundation for the formation of a well-ordered self-assembled monolayer (SAM), which is essential for the stable immobilization of biomolecules. Accordingly, a carboxyl-terminated surface was established using a mercaptoundecanoic acid (MUA) self-assembled monolayer. Alkanethiol SAMs spontaneously assemble on gold through the formation of strong Au–S covalent bonds [8], resulting in highly stable monolayers. The 11-carbon alkyl chain promotes ordered molecular packing through van der Waals interactions, while the terminal carboxyl groups provide reactive sites for subsequent antibody immobilization.

The successful stepwise surface modification observed by cyclic voltammetry was further confirmed by electrochemical impedance spectroscopy, which provided complementary information on the interfacial electron-transfer properties during each functionalization step. The significant increase in R_{ct} following SAM formation indicates successful surface coverage by an insulating alkanethiol monolayer, which hinders electron transfer between the redox probe and the electrode surface [9,10]. This confirms the formation of a stable self-assembled monolayer suitable for subsequent biofunctionalization.

Activation of the terminal carboxyl groups using EDC/NHS chemistry enabled the covalent immobilization of antibodies through amide bond formation with their primary amine groups. The subsequent increase in R_{ct} after antibody immobilization confirms the successful attachment of the antibody layer. The relatively large size and insulating nature of the antibody molecules further impede interfacial charge transfer while preserving the integrity of the underlying SAM [11].

Finally, blocking with bovine serum albumin (BSA) occupied the remaining active sites, thereby minimizing non-specific adsorption [12]. The comparatively small increase in R_{ct} after BSA treatment indicates effective surface passivation without substantially altering the architecture of the biofunctional layer. Overall, the progressive increase in R_{ct} throughout each modification step confirms the successful stepwise fabrication of the biosensor and the formation of a stable biofunctional interface.

suitable for the specific recognition of the target antigen.

4.2 Biosensor Sensitivity and CD235a Detection

These results, along with the concurrent improvement in R_{ct} demonstrate the sensitivity of the biosensor and the successful incorporation of all functionalization steps from the bare electrode to the recognition of CD235a. Although the observed 27% increase in impedance upon antigen binding is modest relative to previous functionalization steps, it is indicative of specific molecular recognition and provides a reliable and distinguishable signal response. This increase is comparable to those reported for other electrochemical immunosensors for protein biomarkers and supports the effectiveness of the antibody immobilization strategy [3, 7].

Glycophorin A, the single most abundant protein in erythrocyte membranes, is present at approximately one million copies per mature red blood cell (RBC) [13] and accounts for ~2% of total membrane protein [14]. It is first expressed at the proerythroblast stage, and its levels increase during erythroid differentiation. With an effective concentration of 0.1 ng of CD235a, this biosensor shows potential for detecting CD235a expression.

The specificity of the biosensor is primarily attributed to the high affinity and selectivity of the recombinant anti-CD235a antibody (clone 27A8). Sensor performance was evaluated using a recombinant human CD235a (glycophorin A) protein corresponding to the extracellular domain (Met1–Glu91) as the target antigen [15]. Furthermore, the BSA blocking step minimized nonspecific adsorption by passivating the remaining active sites on the electrode surface, thereby reducing false-positive signals and improving the overall specificity of the biosensor.

Several factors contribute to the increase in impedance following antigen binding. First, the binding of CD235a, a heavily glycosylated membrane glycoprotein with an apparent molecular mass of approximately 40–45 kDa, increases the steric hindrance at the electrode interface, thereby restricting electron transfer [16]. Second, the extensive glycosylation of CD235a introduces a highly hydrated carbohydrate layer that further impedes charge transfer [17]. In addition, antigen binding may induce conformational changes in the antibody–antigen complex and modify the local dielectric environment at the electrode surface,

resulting in a further increase in the charge transfer resistance [18]. The clear distinction between the BSA-blocked and antigen-bound states demonstrates that the biosensor is capable of detecting CD235a at nanogram-per-milliliter concentrations, confirming the effectiveness of the antibody immobilization strategy.

4.3 Limitations of the biosensor

This study is a proof of concept showing the functionality of the biosensor under idealized conditions with a purified CD235a antigen in a buffer solution. However, prior to practical implementation, several key aspects must be addressed. First, a systematic dose-response characterization is necessary to determine the detection limit, linear dynamic range, and saturation concentration. Although preliminary results indicate sensitivity at the sub-nanogram per milliliter (ng/mL) level, a systematic dose-response analysis is required to accurately determine the detection limit and dynamic range. Second, the biosensor's performance should be evaluated in complex biological samples containing mixtures of different components such as serum proteins, growth factors and/or metabolites that could contribute to nonspecific interactions or interfere with antigen-antibody interactions. The BSA blocking step is expected to mitigate this interference; however, it must be validated under realistic conditions. Third, the long-term stability and reproducibility on different electrodes, as well as the possibility of fully regenerating the sensor, should be evaluated. Because the binding of antibody to antigen is irreversible under physiological conditions, the biosensor may function as a single-use device. Finally, validation in biologically relevant systems is essential. Such validation would demonstrate the ability of the biosensor to detect membrane-associated CD235a from live cells, as opposed to only soluble antigen.

4.4 Future directions: Real-time monitoring of CD235a

The reduced complexity associated of the label-free, real-time detection renders this biosensor beneficial for future monitoring systems, in which changes to the concentration of CD235a may be time dependent. Unlike ELISA, which requires periodic sampling, antibody labeling, and specialized instrumentation, electrochemical detection enables non-invasive, continuous, in situ observation. Making the developed biosensor an ideal tool for real-time monitoring of CD235a.

Continuous impedance measurements over a frequency range (100 kHz to 0.01 Hz) could provide information about CD235a levels. Additionally, the electrical readout allows for integration with automated control systems for closed-loop bioprocess control, which combined with the small (1.6 mm diameter) electrode footprint, enables easy integration into microfluidic or miniaturized bioreactor systems. Furthermore, automated strategies for electrode regeneration, reagent addition, sampling and testing under dynamic flow conditions could also be considered. Ultimately, developing multiplexed sensor arrays that monitor multiple biomarkers simultaneously may provide a comprehensive bioprocess characterization information and enable advanced control strategies.

5. Conclusion

We demonstrated the development of an electrochemical biosensor that enables the label-free electrochemical detection of glycophorin A (CD235a), a key erythroid differentiation marker. This biosensor exhibits high sensitivity and specificity. The sequential functionalization strategy involves mercaptoundecanoic acid (MUA) self-assembled monolayer formation, EDC/NHS activation, anti-CD235a antibody immobilization, and bovine serum albumin (BSA) blocking. This results in a stable biofunctional interface favorable for specific antigen recognition. Progressive surface modification via cyclic voltammetry resulted in a consistent decrease in peak currents, reaching up to 93.4% (anodic) and 96.9% (cathodic) from the bare electrode to the BSA-blocked surface. This indicates successful layer-by-layer assembly. The electrochemical impedance spectroscopy showed that charge transfer resistance increased significantly, from 6 k Ω for the bare electrode to 434 k Ω after CD235a detection, suggesting high sensitivity of the sensing platform. Clear differentiation from the BSA-blocked baseline upon the detection of 0.1 ng of the CD235a antigen further supports the specificity of the antigen-antibody interactions. Unlike conventional monitoring techniques, such as flow cytometry, this platform enables continuous, noninvasive, real-time monitoring without cell labeling, sampling, or disruption of the culture environment. Its small electrode footprint and electrical readout allow for integration into microfluidic devices with automated controls. This modular platform can easily be adapted to other biomarkers, making

it promising for new bioprocess monitoring technologies.

Funding

The author(s) declare that financial support was received for the research and publication of this article. This work was supported by Europäischer Fonds für regionale Entwicklung (European Regional Development Fund) as part of the TACTIC project (Towards co-evolution in human-technology interfaces, grant number: ZS/2023/12/182067).

6. References

- [1] N. G. de Isla, B. D. Riquelme, R. J. Rasia, J. R. Valverde, J. F. Stoltz, *Transfusion* 43, 1145–1152 (2003); doi: 10.1046/j.1537-2995.2003.00471.x
- [2] C. Berens, J. Oldenburg, B. Pötzsch, J. Müller, *Clin. Chem. Lab. Med.* 58, 2081–2087 (2020); doi: 10.1515/cclm-2020-0633
- [3] S. Damiani, C. Haslam, S. Sjøpstad, M. Peacock, T. Whitley, P. Davey, *Biosensors* 9, 97 (2019); doi: 10.3390/bios9030097
- [4] H. Shen, J. E. Mark, C. J. Seliskar, H. B. Mark, Jr., W. R. Heineman, *J. Solid State Electrochem.* 1, 148–154 (1997); doi: 10.1007/s100080050040
- [5] X. Hu, I. O'Connor, J. G. Wall, *Antibody Engineering*, pp. 90–104 (2011); doi: 10.1007/978-3-642-01144-3_7
- [6] C. G. Gahmberg, M. Ekblom, L. C. Andersson, *Proc. Natl. Acad. Sci. U.S.A.* 81, 6752–6756 (1984); doi: 10.1073/pnas.81.21.6752
- [7] U. Karsten, G. Butschak, R. Stahn, S. Goletz, *Int. Immunopharmacol.* 10, 1354–1360 (2010); doi: 10.1016/j.intimp.2010.08.003
- [8] C. Vericat, M. Vela, G. Benitez, P. Carro, R. Salvarezza, *Chem. Soc. Rev.* 39, 1805–1834 (2010); doi: 10.1039/b907301a
- [9] J. J. Calvente, G. López-Pérez, P. Ramírez, H. Fernández, M. A. Zón, W. H. Mulder, R. Andreu, *J. Am. Chem. Soc.* 127, 6476–6486 (2005); doi: 10.1021/ja050265j
- [10] C. M. A. Brett, *Molecules* 27, 1497 (2022); doi: 10.3390/molecules27051497
- [11] S. K. Vashist, C. K. Dixit, B. D. MacCraith, R. O'Kennedy, *Analyst* 136, 4431–4436 (2011); doi: 10.1039/c1an15325k
- [12] Y. L. Jeyachandran, J. A. Mielczarski, E. Mielczarski, B. Rai, *J. Colloid Interface Sci.* 341, 136–142 (2010); doi: 10.1016/j.jcis.2009.09.007
- [13] H. Hassoun, T. Hanada, M. Lutchman, K. E. Sahr, J. Palek, M. Hanspal, A. H. Chishti, *Blood* 91, 2146–2151 (1998); doi: 10.1182/blood.V91.6.2146
- [14] J. A. Chasis, N. Mohandas, *Blood* 80, 1869–1879 (1992); doi: 10.1182/blood.V80.8.1869.1869

- [15] Sino Biological Inc., Human GYPA/CD235a/Glycophorin A Protein (His Tag), Catalog No. 16018-H08H, Product Datasheet (accessed 2026).
- [16] F. Lisdat, D. Schäfer, *Anal. Bioanal. Chem.* 391, 1555–1567 (2008); doi: 10.1007/s00216-008-1970-7
- [17] J. Poole, *Blood Rev.* 14, 31–43 (2000); doi: 10.1054/blre.1999.0124
- [18] J. S. Daniels, N. Pourmand, *Electroanalysis* 19, 1239–1257 (2007); doi: 10.1002/elan.200603855

Modular, 3D-Printed Organ-on-Chip Platform with Integrated Optical Sensing for Real-Time Monitoring of the Tumor Microenvironment

Meghna Wilfred¹, Liudmila Deckert¹, Ulrike Steinmann¹

¹Chair Measurement Technology, Faculty of Electrical Engineering and Information Technology
Institute of Automation Technology (IFAT), Otto-von-Guericke-University Magdeburg,
Universitätsplatz 2, 39106 Magdeburg, Germany

Abstract

Conventional two-dimensional (2D) cancer culture systems fail to reproduce key aspects of the tumor microenvironment (TME), including oxygen gradients, biomechanical cues, and three-dimensional (3D) architecture. Although 3D spheroids and organoids provide improved physiological relevance, their static culture conditions limit dynamic perfusion and hinder real-time environmental monitoring. Existing microfluidic Organ-on-chip (OoC) platforms address these limitations but remain inaccessible to many laboratories due to high fabrication costs, reliance on cleanroom facilities, and lack of intermediate access to cultures. To overcome these challenges, we present a modular, reusable OoC platform fabricated using stereolithography (SLA) 3D printing. This approach enables rapid, cost-effective production of a perfusion system without specialised infrastructure. The system incorporates plug-and-play elements, including a screw top, polydimethylsiloxane (PDMS) membrane support, and central perfusion chamber, allowing easy assembly, sterilisation, and direct access to cultures during experiments. An integrated optical oxygen sensor enables continuous, non-invasive monitoring of oxygen partial pressure (pO_2), providing quantitative insight into metabolic activity and hypoxia formation under perfusion. This platform supports long-term perfusion culture of 3D cancer organoids, enhances environmental monitoring, and improves accessibility of advanced in vitro models. It represents a significant step toward scalable, instrumented systems for cancer research and drug screening.

Keywords: Organ-on-Chip, Dynamic Culture Systems, Stereolithography 3D Printing, Oxygen Sensing, Modular Biofabrication

Introduction

Cancer remains one of the leading causes of mortality worldwide, underscoring the urgent need for improved strategies in drug development and translational oncology [1, 2]. Traditional 2D cell culture models, while widely employed in early-stage compound screening, fail to recapitulate the complexity of tumor biology observed in vivo. Tumor cells in the human body are exposed to dynamic gradients of nutrients and oxygen, interact with a heterogeneous extracellular matrix (ECM), and form 3D structures that influence their growth, metabolism, and drug sensitivity [3].

Over the past decade, 3D cell culture models, including multicellular tumor spheroids, patient-derived organoids, and engineered tissue constructs, have gained considerable prominence due to their capacity to reproduce key aspects of tissue architecture and function [4]. These models offer improved physiological relevance compared to monolayer cultures,

enabling investigation of tumor growth, invasion, and drug response under conditions that better reflect the in vivo TME. Despite these advantages, most 3D models are maintained under static conditions, which limits nutrient and oxygen exchange and creates poorly controlled concentration gradients [5].

In response to these limitations, microfluidic OoC platforms have emerged as powerful tools to introduce controlled perfusion, dynamic shear, and precise transport of solutes [6]. By integrating peristaltic or pressure-driven flow, these systems create physiologically relevant environments, reduce diffusion-limited gradients, and extend the viability of 3D constructs over prolonged culture periods. Furthermore, the integration of real-time sensing technologies, most notably luminescence-based optical oxygen sensors, enables continuous monitoring of metabolic parameters critical to tumor biology, including hypoxia-induced signaling and drug resistance mechanisms [7].

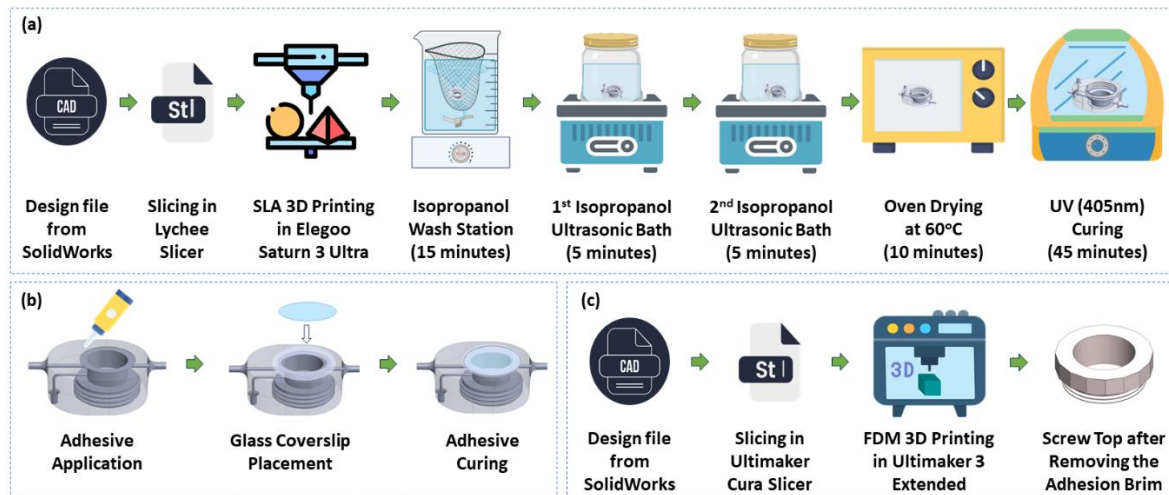


Fig. 1: Fabrication workflow for the 3D-printed perfusion chamber components. (a) SLA 3D printing of the chamber base, (b) Coverslip bonding process, (c) Fused Deposition Modeling (FDM) 3D-printing of the screw top. (Icons from <https://www.freepik.com/>).

However, conventional OoC platforms fabricated via soft photolithography are hindered by high costs, specialised cleanroom requirements, and irreversible sealing techniques that prevent intermediate access to cultures [8]. These barriers significantly restrict adoption in standard biological laboratories. There is, therefore, a pressing scientific and technical need for a simple, low-cost, reusable platform that combines the accessibility of conventional methods with the analytical power of dynamic perfusion and integrated biosensing. This paper presents the design, fabrication, and experimental validation of a novel modular OoC system that addresses these challenges. The platform leverages SLA 3D printing for rapid and cost-effective production of the core fluidic components, incorporates plug-and-play assembly for ease of use, and embeds an optical pO_2 sensor for real-time, non-invasive environmental monitoring. The system is validated using Hepatoblastoma G2 (HepG2) hepatocellular carcinoma 2D monolayers and 3D spheroids embedded in Basement Membrane Extract (BME), with continuous oxygen and viability tracking over up to 15 days.

Platform Architecture and Component Overview

The perfusion chamber consists of four primary components as shown in Figure 2: (1) the chamber body, (2) a thin PDMS gas-permeable membrane, (3) a polypropylene (PP) screw-top lid, and (4) a borosilicate glass coverslip affixed to the base.

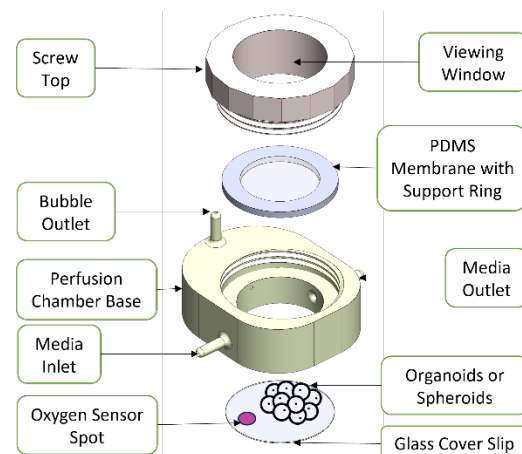


Fig. 2: Exploded view of the 3D-printed perfusion chamber assembly showcasing all the features and components.

Chamber Body Design and SLA Fabrication

The chamber body was designed in SolidWorks 2023 as a cylindrical well with an inner diameter (ID) of 16 mm and a height of 8 mm, yielding a culture volume of approximately 1.61 mL. Above the well, a flat annular surface was modelled to seat the PDMS membrane, with M27 × 2.0 metric threads incorporated into the surrounding side walls to accept the PP screw top. Inlet and outlet microchannels (1.5 mm ID) were positioned 2 mm from the chamber floor to facilitate efficient media exchange. An integrated bubble-outlet vent was incorporated to permit removal of trapped air via syringe aspiration without disrupting the fluidic circuit.

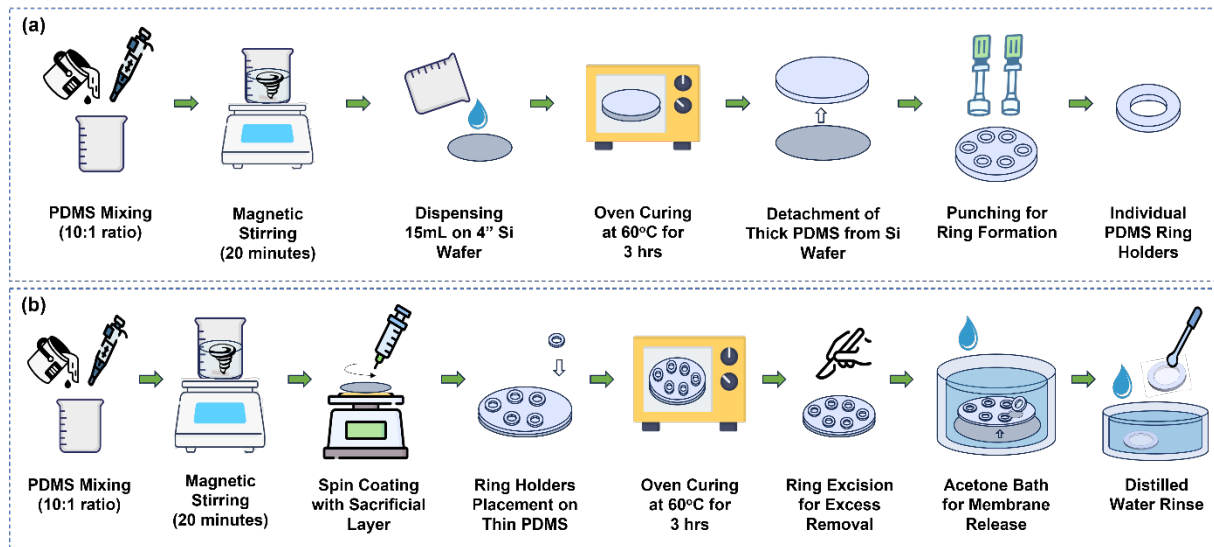


Fig. 3: Schematic representation of the PDMS membrane fabrication process. (a) PDMS ring holder and (b) composite thin membrane. (Icons from <https://www.freepik.com/>).

The chamber body was fabricated from Liqcreate Bio-Med Clear photopolymer resin using an SLA printer (Elegoo Saturn 3 Ultra, China). Design files were exported as STL and imported into Lychee Slicer 6.0.2. Models were oriented at approximately 45° to the build plate to minimise stair-stepping artefacts on critical surfaces. Following printing, parts were washed sequentially in three isopropanol (99%) baths, two of which included 5-minute ultrasonic cleaning cycles then dried at 60°C for 10 minutes, and finally ultraviolet (UV) cured for 45 minutes in an Elegoo Mercury X curing station. The process is summarized in Figure. 1.

PDMS Membrane Fabrication

PDMS membranes were prepared from Sylgard 184 silicone elastomer (Dow, USA) mixed at a 10:1 (base:curing agent) mass ratio as shown in Figure. 3. The mixture was stirred at 500 rpm for 20 minutes before use. A thin membrane (~22 µm) was produced by dispensing PDMS onto a 2-inch silicon wafer coated with AZ1512 photoresist as a sacrificial release layer, followed by spin-coating at 1800 rpm for 120 seconds (POLOS SPIN150i, SPS Netherlands). A pre-punched thick PDMS ring-holder (Outer diameter (OD): 22 mm, ID: 15 mm, ~1.5 mm thick) was placed concentrically over the uncured thin membrane to provide structural support during handling. The assembly was cured at 60°C for 3 hours. After curing and trimming the outer edges of the holder, the wafer was immersed in acetone to dissolve the sacrificial photoresist layer, releasing the composite membrane. Membrane thickness was verified by spectrophotometry (Filmetrics F20-UVX, USA).

Assembly and Sterilisation

The glass coverslip was bonded to the chamber base using SILPURAN® 4200 (Wacker Chemie, Germany), a biocompatible silicone adhesive, applied via a precision dispensing needle (Nordson EFD, USA, 1.5 mm, 14-GA) and cured at room temperature for 48 hours. Assembly was performed in a laminar flow hood; all reusable components were autoclaved prior to use, while the self-adhesive oxygen sensor spot (PreSens, Germany) was applied immediately before the experiment. The plug-and-play screw-top design allowed the chamber to be opened, accessed, and resealed between experimental time points without compromising the circuit's integrity.

Perfusion System Configuration

The perfusion circuit comprised a DURAN® GL 45 glass bottle as the media reservoir, Tygon® LMT-55 tubing (Saint-Gobain, France, 3 mm OD, 1.6 mm ID), a 3D-printed bubble trap, and a digital peristaltic pump (Ismatec Masterflex, Germany) and tubing (Ismatec Masterflex pump tubing, 0.25 mm ID, orange and blue, three stoppers). The media bottle was equipped with a three-port cap: two GL14 ports served as inlet and outlet (fitted with 3 mm OD tubing inserts), and the third port was connected to a 0.22 µm PTFE membrane filter (Sartorius, Germany) for sterile gas exchange. Figure. 4 depicts the assembly of the perfusion system. A standard flow rate of 50 µL/min was established and maintained continuously. A 3D-printed incubator door adaptor routed tubing without kinking, allowing the pump to remain outside the

incubator while the chamber, bubble trap, and media reservoir were maintained at 37°C, 5% CO₂, and 95% relative humidity (Thermo Scientific BBD 6220, USA).

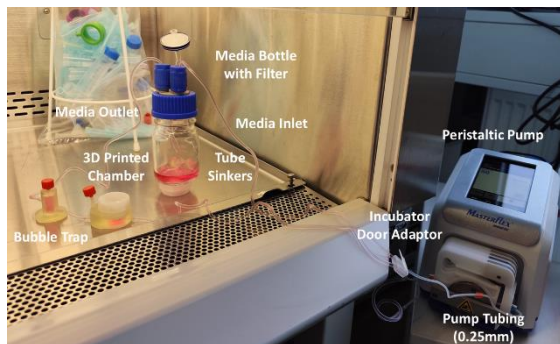


Fig. 4: Assembly of the perfusion system under the laminar flow hood.

Integrated Optical Oxygen Sensing

Dissolved oxygen was measured non-invasively using optical sensor spots operating on the principle of dynamic luminescence quenching (Stern-Volmer relationship):

$$I_0/I = \tau_0/\tau = 1 + K_{SV} [O_2] \quad (1)$$

where I_0 and τ_0 are the luminescence intensity and lifetime in the absence of the quencher, I and τ are the respective values in the presence of molecular oxygen, K_{SV} is the Stern-Volmer quenching constant, and $[O_2]$ is the local oxygen concentration [9].

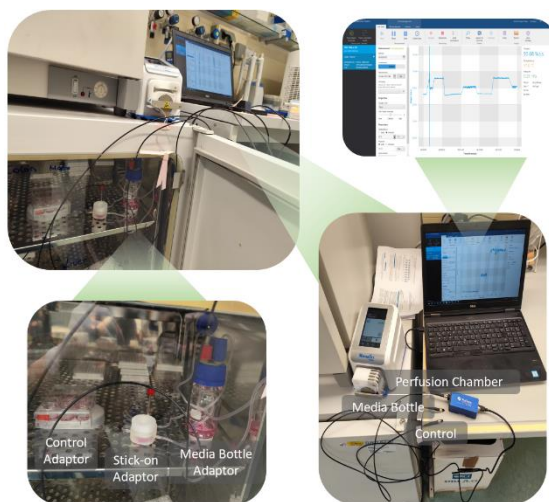


Fig. 5: Experimental oxygen measurement setup from different viewing angles.

Two sensor types were used: SP-Pst3-YAU (autoclavable glass-backed sensor) for the media reservoir, and SP-Pst3-SA (single-use self-adhesive polyester sensor) adhered to the culture chamber coverslip and to a control well

in a 12-well plate. Measurements were performed using an OXY-1-SMA single-channel fiber optic reader (PreSens, Germany) connected to PreSens Measurement Studio 2 software, which provided real-time graphical display, event annotation, and data export [10]. Custom 3D-printed adaptors maintained stable and precise alignment of fiber optic cables beneath each sensor spot, minimising measurement drift (Figure. 5).

Cell Culture and 3D Spheroid Formation

HepG2 human hepatocellular carcinoma cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) or Roswell Park Memorial Institute (RPMI) medium supplemented with 10% Fetal Calf Serum (FCS) and 1% penicillin/streptomycin at 37°C and 5% CO₂. Following routine 2D culture and at least three passages, cells were harvested via trypsinization and prepared for 3D spheroid formation using the hanging-drop method. Specifically, 10 µL drops containing 5,000-50,000 cells were resuspended in a mixture of 20% FCS, 40% DMEM, and 40% methylcellulose before being pipetted onto inverted Petri dish lids and incubated for three days. Once formed, individual spheroids were harvested, washed, and embedded in 10 µL of BME on 13 mm plastic coverslips (Sarstedt, Germany) [11]. After polymerization at 37°C for 30 minutes, the coverslips were transferred into 3D-printed perfusion chambers for downstream applications

Design Validation

The SLA-fabricated chamber body demonstrated excellent dimensional fidelity, with channel geometries and thread profiles matching design specifications. The M27 × 2.0 screw-top interface provided a repeatable, leak-free seal across multiple assembly/disassembly cycles, confirming the reusability of the platform. The SILPURAN® 4200-bonded coverslip showed no signs of delamination or media leakage throughout extended culture periods of up to 15 days under 50 µL/min perfusion, validating the adhesive selection for biocompatible sealing.

The spin-coated PDMS membranes achieved a mean thickness of 22 ± 2 µm as verified by spectrophotometry, sufficient to support passive gas exchange while maintaining mechanical integrity under the applied flow pressures. The plug-and-play design enabled assembly in under 10 minutes by a trained operator within a laminar flow hood, with no cleanroom access required, a considerable practical advantage over conventional photolithography-based OoC platforms [8].

HepG2 2D Monolayer Perfusion: Oxygen Dynamics and Flow Rate Effects

In the HepG2 2D monolayer experiment illustrated in Figure. 6, oxygen dynamics were observed over the course of two days in the perfusion chamber. During the initial perfusion phase, the system was maintained at a standard flow rate of 50 $\mu\text{L}/\text{min}$, where dissolved oxygen levels stabilised as a steady balance was reached between continuous oxygen delivery and cellular consumption. However, when the flow rate was momentarily increased to 600 $\mu\text{L}/\text{min}$, sharp upward spikes in pO_2 occurred due to the rapid replenishment of oxygenated media. While this intervention led to an increase in oxygen levels, the high-shear environment proved detrimental, causing significant cell detachment. This resulted in an unexpected rise in baseline oxygen levels by day two, reflecting a decrease in metabolic demand from the reduced cell population.

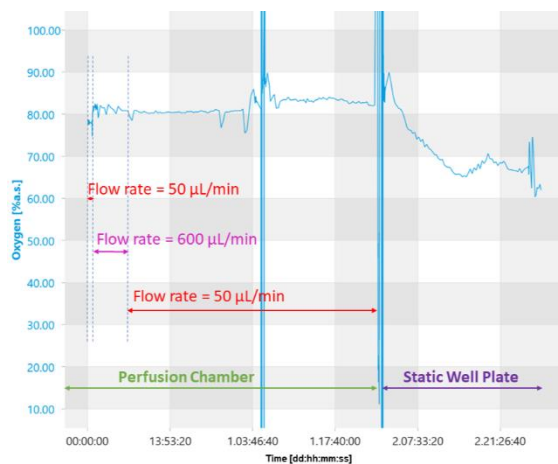


Fig. 6: Impact of flow rate on oxygenation and the variation in oxygen levels between the static control and perfusion chamber with HepG2 monolayer culture.

After two days, the input to the fiber optic oxygen meter was transitioned from the perfusion chamber to a static well plate control to observe the effects of stagnant conditions. Following the transition to the static environment, a loss of previous stability was observed, accompanied by a steady and monotonic decline in oxygen levels. By the final day of measurement, the pO_2 had reached near-depletion levels, as the dense, confluent HepG2 monolayer had exhausted the available oxygen without the benefit of active replenishment. This sharp contrast confirms that static culture conditions are insufficient to sustain the high metabolic demands of dense cell populations over time [5].

HepG2 3D Spheroid Culture: Long-Term Viability and Oxygenation

The HepG2 spheroid experiment represented the longest culture period in this study (15 days). As illustrated in Figure. 7, during the early phase (days 0–3), pO_2 in the perfusion chamber stabilised at approximately 85–86% air saturation, while the static control maintained $\sim 90\%$. In the intermediate phase (days 3–10), both conditions remained relatively stable. By the late phase (days 10–15), static well pO_2 declined sharply to $\sim 65\%$, indicative of oxygen depletion driven by the high metabolic activity of a densely packed spheroid culture with no active media replenishment. In contrast, the perfusion chamber maintained a stable $\sim 86\%$ pO_2 , confirming that continuous media exchange effectively prevented hypoxic depletion. The media reservoir consistently exhibited the highest oxygen levels (96–97%), confirming that observed variations arose from cellular consumption rather than environmental fluctuations.

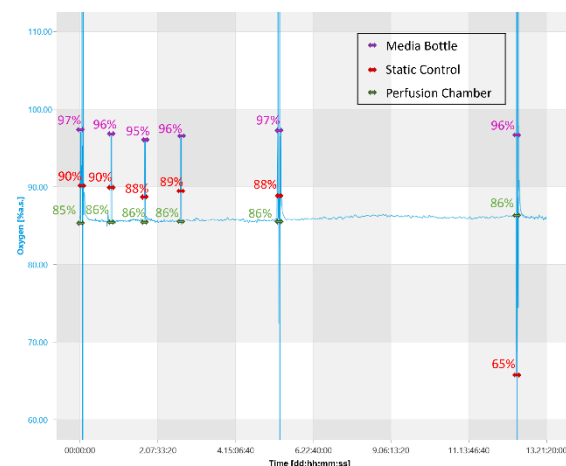


Fig. 7: Oxygenation differences among the static control, media bottle, and perfusion chamber in HepG2 spheroid culture over 15 days.

These findings highlight the platform's capability to distinguish between oxygen depletion regimes in real time, a capability not available in static systems and to sustain physiological oxygen tensions within 3D tumor models over extended culture periods. The approximately 4–6% difference in oxygen saturation between perfused and static conditions may appear modest but can be physiologically significant: even mild hypoxia can shift cellular metabolism, alter transcriptional programmes, and modulate drug sensitivity in tumor cell lines [12].

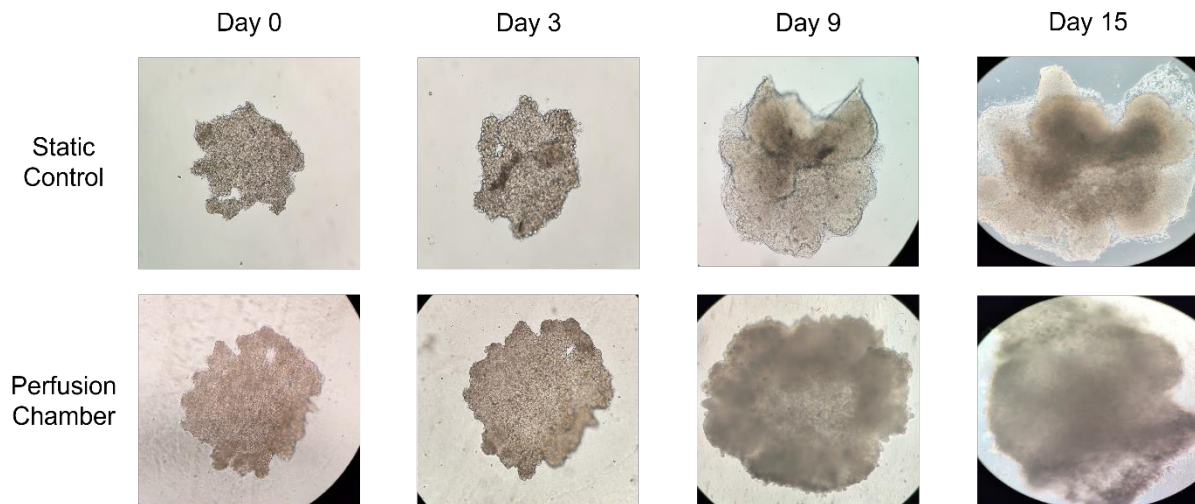


Fig. 8: Microscopic images of HepG2 spheroids in the static and perfusion conditions over the experimental period.

Morphological Assessment of HepG2 3D Spheroid Culture

Morphological assessment of HepG2 spheroids by brightfield microscopy (Zeiss Axiovert 25, Germany) revealed clear differences between static and perfusion culture conditions over the 15-day period (Figure. 8). In the static condition (top row), spheroids progressively increased in size and adopted irregular, lobulated morphologies with increasingly darkened and opaque cores. This pattern is consistent with the development of central hypoxia and the onset of necrotic core formation, a well-documented hallmark of avascular 3D tumor models under diffusion-limited, static conditions [5, 12]. The progressive spreading and irregular peripheral margins observed at later time points further reflect the unrestricted outgrowth characteristic of static 3D cultures lacking mechanical or fluidic constraints.

In contrast, spheroids maintained under continuous perfusion (bottom row) retained a more compact, rounded architecture with smoother boundaries throughout the culture period. The cores remained more translucent and morphologically uniform compared to their static counterparts, suggesting more homogeneous oxygenation across the spheroid volume, an observation directly supported by the stable pO_2 readings ($\sim 86\%$ air saturation) recorded in the perfusion chamber throughout the culture period (Figure. 7). The absence of progressive core darkening in perfused spheroids indicates that continuous media exchange was sufficient to prevent the oxygen depletion and associated cellular deterioration observed in the static condition.

Together, these morphological findings corroborate the oxygen data and reinforce the conclusion that the perfusion platform sustains a physiologically more controlled tumor microenvironment over extended culture periods, one that more faithfully reflects the constrained, flow-regulated conditions experienced by tumor nodules in vivo, compared to the unrestricted and increasingly hypoxic conditions that develop in static culture.

Platform Accessibility and Broader Applicability

A key outcome of this work is the demonstrated accessibility of the platform. By leveraging SLA 3D printing with a commercially available biocompatible resin, fabrication is achievable within hours and at a fraction of the cost of photolithographic alternatives. The modular, screw-top design not only enables intermediate culture access for direct sampling, microscopy, or sensor spot replacement, but also allows the chamber body to be reused after cleaning and re-sterilisation, reducing per-experiment consumable costs. Compatibility with standard borosilicate glass coverslips ensures seamless integration with conventional inverted fluorescence microscopes, enabling high-resolution live-cell imaging without platform modification.

The same architecture is directly transferable to other 3D tumor models embedded in BME, broadening the platform's applicability across oncology research. Future developments will focus on parallelisation of chambers for throughput enhancement, integration of additional electrochemical sensors (pH, glucose), and incorporation of co-culture chambers.

Conclusion

We have presented a modular, reusable OoC platform fabricated entirely by accessible 3D printing technologies, SLA for the primary chamber body and FDM for auxiliary components and validated for long-term 3D tumor culture perfusion culture with integrated optical oxygen sensing. The platform eliminates cleanroom requirements, reduces fabrication costs, and enables intermediate culture access while maintaining a controlled, physiologically relevant microenvironment. Experimental validation with HepG2 monolayers and spheroids over up to 15 days demonstrated robust oxygen tracking across perfused and static conditions, revealing key differences in metabolic dynamics that static systems cannot resolve. These results position the platform as a practical and scientifically rigorous tool for advanced in vitro oncology research, drug screening, and translational studies in standard biological laboratories.

Acknowledgements

The authors gratefully acknowledge the financial support provided by the European Union and the State of Saxony-Anhalt under the EFRE programme [IB ZS/2023/12/186743]. We would also like to thank Ahmed Sanin and Prof. Ulf Kahlert (Molecular and Experimental Surgery, University Medical Center Magdeburg) for their valuable technical assistance and expertise regarding cell culture.

References

1. A. Boire *et al.*, "Why do patients with cancer die?," *Nat. Rev. Cancer*, vol. 24, no. 8, pp. 578–589, Aug. 2024, doi: 10.1038/s41568-024-00708-4.
2. D. Hanahan and R. A. Weinberg, "Hallmarks of Cancer: The Next Generation," *Cell*, vol. 144, no. 5, pp. 646–674, Mar. 2011, doi: 10.1016/j.cell.2011.02.013.
3. F. Antonelli, "3D Cell Models in Radiobiology: Improving the Predictive Value of In Vitro Research," *Int. J. Mol. Sci.*, vol. 24, no. 13, Art. no. 13, Jan. 2023, doi: 10.3390/ijms241310620.
4. A. Borah and D. S. Kumar, "7 - Overcoming the barriers of two-dimensional cell culture systems with three-dimensional cell culture systems: techniques, drug discovery, and biomedical applications," in *Biomedical Product and Materials Evaluation*, P. V. Mohanan, Ed., Woodhead Publishing, 2022, pp. 179–229. doi: 10.1016/B978-0-12-823966-7.00003-7.
5. D. Egger, C. Tripisciano, V. Weber, M. Dominici, and C. Kasper, "Dynamic Cultivation of Mesenchymal Stem Cell Aggregates," *Bioengineering*, vol. 5, no. 2, Jun. 2018, doi: 10.3390/bioengineering5020048.
6. L. A. Low, C. Mummery, B. R. Berridge, C. P. Austin, and D. A. Tagle, "Organs-on-chips: into the next decade," *Nat. Rev. Drug Discov.*, vol. 20, no. 5, pp. 345–361, May 2021, doi: 10.1038/s41573-020-0079-3.
7. E. S. Ildiz, A. Gvozdenovic, W. J. Kovacs, and N. Aceto, "Travelling under pressure - hypoxia and shear stress in the metastatic journey," *Clin. Exp. Metastasis*, vol. 40, no. 5, pp. 375–394, 2023, doi: 10.1007/s10585-023-10224-8.
8. S. Halldorsson, E. Lucumi, R. Gómez-Sjöberg, and R. M. T. Fleming, "Advantages and challenges of microfluidic cell culture in polydimethylsiloxane devices," *Biosens. Bioelectron.*, vol. 63, pp. 218–231, Jan. 2015, doi: 10.1016/j.bios.2014.07.029.
9. Kerry. P. McNamara, X. Li, A. D. Stull, and Z. Rosenzweig, "Fiber-optic oxygen sensor based on the fluorescence quenching of tris (5-acrylamido, 1,10 phenanthroline) ruthenium chloride," *Anal. Chim. Acta*, vol. 361, no. 1, pp. 73–83, Mar. 1998, doi: 10.1016/S0003-2670(97)00703-4.
10. "Optical O2 Sensor Products." Accessed: Apr. 01, 2026. [Online]. Available: <https://www.presens.de/products/o2>
11. T. Hurrell, A. A. Ellero, Z. F. Masso, and A. D. Cromarty, "Characterization and reproducibility of HepG2 hanging drop spheroids toxicology *in vitro*," *Toxicol. In Vitro*, vol. 50, pp. 86–94, Aug. 2018, doi: 10.1016/j.tiv.2018.02.013.
12. A. Asthana and W. S. Kisaalita, "Microtissue size and hypoxia in HTS with 3D cultures," *Drug Discov. Today*, vol. 17, no. 15, pp. 810–817, Aug. 2012, doi: 10.1016/j.drudis.2012.03.004.

Ultradünne Lithiumtantalat-Chips mit Absorptionsschicht für hochempfindliche pyroelektrische Infrarotsensoren

Agnes Eydam¹, Gerald Gerlach¹

¹Technische Universität Dresden, Institut für Festkörperelektronik, 01062 Dresden, Deutschland

Zusammenfassung

In dieser Arbeit wird die Herstellung ultradünner Lithiumtantalat-Chips mit einer Dicke von ca. 1 μm beschrieben. Die Abdünnung erfolgt in mehreren Schritten mittels Ionenstrahlätzen, wobei die Inhomogenität des Ausgangswafers als auch des Ätzprozesses berücksichtigt werden müssen. Dabei wird ein Kompromiss zwischen Ausbeute und Bearbeitungszeit gefunden. Neben der Dickentoleranz wird auch der dielektrische Verlustfaktor des Ausgangsmaterials bestimmt, der möglichst gering sein soll. Mittels Schrägwinkelbedampfung wird eine Absorptionsschicht aus Aluminium auf den Chips abgeschieden. Eine Gradientenschicht erreicht dabei Absorptionsgrade von (80 ... 90) % für Wellenlängen von (2 ... 14) μm .

Keywords: Infrarotsensoren, Lithiumtantalat, Ionenstrahlätzen, GLAD-Verfahren, Absorptionsschicht.

Einleitung

Pyroelektrische Infrarotsensoren aus Lithiumtantalat (LiTaO_3) sind weit verbreitete Strahlungssensoren, die in vielen Bereichen wie der Gasanalytik, der Spektroskopie und der Pyrometrie eingesetzt werden. Bestimmte Anwendungsgebiete, z.B. der Nachweis sehr geringer Gaskonzentrationen in der Umwelt- und Prozessmesstechnik, erfordern immer empfindlichere Detektoren. Ziel dieser Arbeit ist deshalb die Entwicklung von pyroelektrischen Sensoren mit einem extrem hohen Signal/Rausch-Verhältnis. In [1] war gezeigt worden, dass die spezifische Detektivität von pyroelektrischen LiTaO_3 -Sensoren mit sinkender Elementdicke steigt. Deshalb ist der erste Lösungsansatz eine Technologie zur reproduzierbaren Fertigung von ultradünnen Sensorchips mit einer Dicke des empfindlichen Elements von etwa 1 μm . Der zweite Ansatz besteht in der Entwicklung einer an solche extrem dünnen Chips angepassten Absorptionsschicht aus Aluminium.

Stand der Technik

Da LiTaO_3 aufgrund seiner Zusammensetzung und seines Aufbaus chemisch nahezu inert ist, lassen sich entsprechende Wafer praktisch nicht sinnvoll durch chemische Ätzverfahren abdünnen. Deshalb liegt der Schwerpunkt der Arbeiten zur Schaffung von extrem dünnen Sensorelementen mit Dicken im Bereich einiger μm nicht auf solchen subtraktiven Verfahren, sondern auf additiven Abscheidungsverfahren von LiTaO_3 -Dünnschichten auf Substraten wie z.B. das

Sol-Gel-Verfahren, Sputtern und epitaktisches Aufwachsen [2]. Diese Schichten weisen jedoch aufgrund ihrer polykristallinen Struktur einen höheren dielektrischen Verlustfaktor und einen geringeren pyroelektrischen Koeffizienten auf als einkristallines LiTaO_3 . Ein weiteres Verfahren ist die Ion-Cut-Technologie, mit der eine Schichtdicke von 770 nm erreicht wurde [3]. Das Rauschen ist bei diesem Verfahren jedoch vergleichsweise hoch und die spezifische Detektivität entsprechend gering.

Der in dieser Arbeit verwendete Ansatz basiert hingegen auf der Abdünnung von einkristallinen LiTaO_3 -Wafers durch Ionenstrahlätzen (Ion Beam Etching - IBE), wodurch sich dünne, freitragende Sensorelemente mit einer Dicke im Bereich bis zu 1 μm fertigen lassen [4]. Unser Schwerpunkt lag dabei neben der möglichst großen Detektivität auch auf einer hohen Ausbeute und Reproduzierbarkeit, die für eine industrielle Fertigung unabdingbar sind.

Damit die Infrarotstrahlung möglichst vollständig und wellenlängenunabhängig vom LiTaO_3 -Chip absorbiert wird, muss zusätzlich eine geeignete Absorptionsschicht aufgebracht werden. In der Vergangenheit wurden durch Schrägwinkelbedampfung (Glancing Angle Deposition – GLAD) nanostrukturierte Nickel-Chrom (NiCr)-Schichten mit einem Absorptionsgrad von 90 % für den Wellenlängenbereich von (1 ... 15) μm entwickelt [5]. Neuere Entwicklungen nehmen carbonbasierte Absorptionsschichten in den

Fokus, die ähnliche Eigenschaften aufweisen [6].

In dieser Arbeit werden auf Basis des GLAD-Verfahrens dünne Absorptionsschichten aus Aluminium entwickelt, die im Vergleich zu den vorher genannten Lösungen eine deutlich reduzierte Wärmekapazität aufweisen, was deutliche Vorteile für ultradünne Chips und den nutzbaren Frequenzbereich bietet.

Charakterisierung des Ausgangsmaterials

Ausgangspunkt der technologischen Untersuchungen war die Charakterisierung des verwendeten Ausgangsmaterials. Ultradünne LiTaO₃-Wafer sind kommerziell mit Dicken von 20 µm oder 25 µm und Durchmessern von (2,5 ... 3) Zoll erhältlich. Für die Nutzung von Ätzverfahren zur Abdünnung der Sensorelemente ist die eingeschränkte Planparallelität der Ausgangswafer problematisch, da dies bei gleichmäßiger Ätzrate über dem Gesamtwufer zum Durchätzen von Chips führen kann. Deshalb wurde die Dickenverteilung der dünnen Wafer mit dem spektroskopischen Messgerät NanoCalc-2000-XR von Ocean Optics (im Folgenden NanoCalc genannt) bestimmt. Dazu wurden die Wafer in kleinere Teilwafer der Größe (22 x 22) mm² gesägt, aus denen sich 16 Chips der Größe (4 x 4) mm² ergeben. Tabelle 1 zeigt die mittlere Dicke und die Spannweite über die Teilwafer verschiedener Hersteller. Die Spannweite ist dabei als die Differenz zwischen maximaler und minimaler Dicke definiert.

Tab. 1: *Dickenverteilung über Teilwafer der Größe (22 x 22) mm² verschiedener Hersteller.*

Nr.	Mittelwert in µm	Spannweite in µm
1a	20,41	1,20
1b	20,06	2,30
1c	19,79	2,07
2	24,07	1,38
3	20,18	1,00
4	26,18	3,80

Wafer Nr. 3 hat mit 1 µm Spannweite die geringsten Dickenunterschiede, wohingegen Nr. 4 wesentlich inhomogener als die anderen Wafer ist. Nr. 1a bis 1c sind aus demselben Material, d.h. es kann Unterschiede bis zu 1 µm geben.

Neben der Dicke wurden wichtige Materialeigenschaften bestimmt. Der pyroelektrische Koeffizient p beschreibt die Änderung der spontanen Polarisation bei einer Temperaturänderung und bildet die Grundlage des pyroelektrischen Effekts. Er wird bestimmt, indem der Chip durch ein Peltier-Element harmonisch erwärmt und der Kurzschlussstrom gemessen wird. Die relative Permittivität bzw. Dielektrizitätszahl ϵ_r wird aus der Kapazität C des Chips berechnet. Die Messung erfolgte mit einem Impedanzanalysator Solartron SI1260 (Solartron Analytical) unter Anwendung des Vierleiterprinzips. Der dielektrische Verlustfaktor $\tan \delta$ wurde gemeinsam mit der Kapazität gemessen. Der pyroelektrische Chip stellt keine ideale Kapazität dar, sondern eine Parallelschaltung mit einem Wechselstromwiderstand. Dessen Größe wird durch $\tan \delta$ bestimmt und er ist eine wesentliche Rauschquelle in pyroelektrischen Sensoren. Die Messergebnisse sind in Tabelle 2 zusammengefasst.

Tab. 2: *Pyroelektrischer Koeffizient p (bei 20 °C), relative Permittivität ϵ_r und dielektrischer Verlustfaktor $\tan \delta$ (jeweils bei 1 kHz) der Teilwafer verschiedener Hersteller.*

Nr.	p in nC/(cm ² K)	ϵ_r	$\tan \delta$
1a	21,3	46	$1,47 \cdot 10^{-3}$
1b	21,3	47	$1,54 \cdot 10^{-3}$
1c	21,4	47	$1,45 \cdot 10^{-3}$
2	21,8	47	$1,90 \cdot 10^{-3}$
3	21,0	47	$2,04 \cdot 10^{-3}$
4	nicht bestimmbar	46	$1,76 \cdot 10^{-3}$

Der pyroelektrische Koeffizient beträgt für alle Wafer (21 ... 22) nC/(cm²K). Von Wafer 4 konnte er bisher nicht sicher bestimmt werden, weil aufgrund von Unebenheiten auf der Chipoberfläche der elektrische Kontakt während der Messung unzuverlässig war. Die relative Permittivität ist mit einem Wert von etwa 47 ebenfalls für alle Hersteller gleich. Unterschiede gibt es beim dielektrischen Verlustfaktor, wobei das Material von Hersteller 1 mit ca. 0,0015 gegenüber den anderen einen etwas geringeren Wert aufweist. Es wird deshalb bevorzugt verwendet.

Herstellungstechnologie für ultradünne Chips

Normalerweise bilden drei Teilwafer eine Chipserie, die gleichzeitig bearbeitet wird. Von jedem Wafer wird vor Bearbeitungsbeginn die Dickenverteilung gemessen. Der Wafer wird fotolithografisch strukturiert, um Schichten entweder durch thermische Bedampfung abzuscheiden oder durch Ionenstrahlätzen abzutragen. Jeder Chip benötigt eine Front- und Rückelektrode, jeweils mit einem verstärkten Bereich zur Kontaktierung des Chips. Der sich überlappende Bereich beider Elektroden bildet das empfindliche Element, das durch IBE abgedünnt wird. Dadurch entsteht eine freitragende Membran in Form einer Schüssel, die durch einen 20 µm dicken Trägerrahmen mechanisch stabilisiert wird. Ein solcher Chip ist in Abbildung 1 dargestellt.

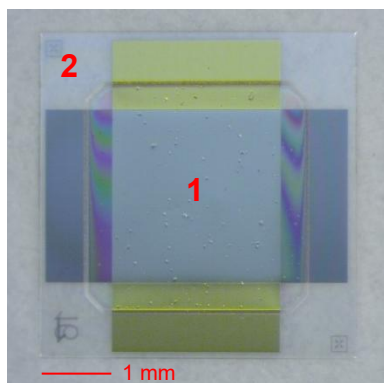


Abb. 1: Fotografie eines freitragenden LiTaO₃-Chips mit einer Dicke des empfindlichen Elements von 1,6 µm.

1 Schüssel, 2 Trägerrahmen.

Die Herausforderung besteht nun darin, bei Dickenunterschieden von 2 µm und mehr über einen Teilwafer und einer Zielenddicke von 1 µm eine relativ hohe Ausbeute zu erzielen. Zu der Inhomogenität des Wafers im Ausgangszustand kommt noch eine prozessbedingte Inhomogenität des IBE hinzu.

Die Wafer wurden in mehreren Teilschritten geätzt. Im ersten Schritt erfolgte immer eine Abdünnung auf (5 ± 1) µm, weil sich diese Dicke zuverlässig und reproduzierbar herstellen lässt. Anschließend wurden alle Chips wieder mit dem NanoCalc ausgemessen. Dann wurde in kleinen Schritten tiefer geätzt. Vor jedem Schritt wurden die Chips, die schon die gewünschte Enddicke erreicht hatten, manuell mit Fotoresist abgedeckt.

In der ersten Phase sollte zunächst eine Enddicke von 2 µm erreicht werden. Diese

wurde nach der Ätzung auf 5 µm in weiteren drei Schritten mit Ätztiefen von $(1 \dots 2)$ µm erzielt. Tabelle 3 zeigt beispielhaft die Dickenverteilung für einen Wafer, der einen sehr homogenen Ausgangszustand aufwies. Durch die Inhomogenität des Ätzprozesses vergrößerte sich die Spannweite jedoch wesentlich, d.h. die Ätzung hat einen großen Einfluss.

Tab. 3: Dickenverteilung von Wafer 1 der Chipserie I:

Ausg – Ausgangszustand,
1.Ätz – nach 1. Ätzung,
4.Ätz – nach 4. Ätzung.

Dicke in µm	Mittelwert	Max.	Min.	Spannweite
Ausg	18,7	19,2	18,3	0,8
1.Ätz	4,3	4,7	3,6	1,1
4.Ätz	2,0	2,6	0,9	1,7

In dieser Phase wurden keine Chips durchgeätzt, allerdings wurde dafür eine große Zahl von Ätzschritten gebraucht. Für eine industrielle Fertigung muss ein Kompromiss zwischen Ausbeute und Bearbeitungszeit gefunden werden.

In der zweiten Phase wurde deshalb die maximale Zahl der Ätzschritte auf vier begrenzt, um eine Zielenddicke von 1 µm zu erreichen. Die Ausbeute betrug mehr als 50 %. Tabelle 4 zeigt beispielhaft die Dickenverteilung einer Chipserie im Ausgangs- und Endzustand. Die im Ausgangszustand vorhandene Inhomogenität verringerte sich meistens nur wenig. In einigen Fällen vergrößerte sie sich sogar.

Tab. 4: Dickenverteilung der Chipserie II mit drei Wafern im Ausgangs- und Endzustand.

	Ausgangszustand		Endzustand	
Dicke in µm	Mittelwert	Spannweite	Mittelwert	Spannweite
W1	25,0	2,6	0,8	2,0
W2	26,2	1,6	1,1	1,9
W3	26,1	2,7	1,3	2,5

Die Dickenverteilung über die Wafer nach der ersten langen Ätzung verdeutlichte, dass die Ätzrate in der Mitte des Ionenstrahls höher ist

als am Rand. Die Bearbeitungsronen mit den Wafern wurden auf einen runden Träger geschraubt. Der Teil auf der Innenseite des Trägers war normalerweise tiefer geätzt. Für eine Sollätztiefe von ca. 15 µm lag die Spannweite der Ätztiefe dabei meistens bei (2 ... 3) µm. Die Inhomogenität der ersten Ätzung wurde in den folgenden Ätzschritten größtenteils beibehalten.

In der dritten Phase wurde der erste Ätzschritt deshalb in zwei Schritte geteilt, wobei der Wafer nach der halben Ätztiefe um 180° gedreht wurde, um die Inhomogenität auszugleichen. Ein späteres Drehen in den kurzen Ätzungen brachte aufgrund der geringen Ätzzeit keinen Ausgleich. Tabelle 5b zeigt, dass sich die Spannweite nach der zweiten Ätzung, d.h. nach der Drehung, verringert hat. Das ist hauptsächlich für Wafer 1 vorteilhaft, dessen Inhomogenität sich durch die erste Ätzung wesentlich verschlechtert hatte. Aufgrund der guten Homogenität genügte dann zwei weitere kurze Ätzschritte, um die Enddicke von 1 µm mit einer sehr guten Ausbeute zu erreichen. Für solch geringe Ätztiefen von (1 ... 2) µm ist die Spannweite der Ätzung meist kleiner als 0,5 µm. Generell verringert sich die Inhomogenität des IBE mit der Ätzdauer.

Tab. 5: *Dickenverteilung der Chipserie III mit drei Wafern: (a) Mittelwerte und (b) Spannweiten.*

*Ausg – Ausgangszustand,
1.Ätz – nach 1. Ätzung,
2.Ätz – nach 2. Ätzung,
End – Endzustand.*

(a)

Dicke in µm	Ausg	1.Ätz	2.Ätz	End
W1	19,1	11,4	4,8	1,2
W2	18,7	11,0	4,6	1,0
W3	20,0	12,0	5,5	1,1

(b)

Dicke in µm	Ausg	1.Ätz	2.Ätz	End
W1	1,0	1,7	1,1	0,9
W2	1,6	1,4	1,1	1,2
W3	2,2	2,1	1,9	0,7

Die Ätzrate lag dabei bei (27 ... 28) nm/min. Sie sollte möglichst stabil sein, da die Ätzdauer aus der gewünschten Ätztiefe und der Rate berechnet wird. Bevor die eigentliche Ätzung beginnt, läuft die Ionenstrahlquelle nach der Plasmazündung zunächst warm. Nach Einschalten der Strahl- und Beschleunigungsspannung wird der Shutter nicht sofort geöffnet, damit sich der Ionenstrahl zunächst stabilisieren kann und eventuelle Kontaminationen verschwinden. Die kurzen Ätzschritte führten zu reproduzierbareren Ätzraten, nachdem diese Zeiten verlängert wurden.

Als Ätzgas wurde Argon eingesetzt und der Einfallswinkel des Ionenstrahls betrug 5°.

Zu beachten ist die Schädigung der Kristalloberfläche durch das IBE. Ionen dringen in den Oberflächenbereich des verbleibenden einkristallinen Materials ein und schädigen diesen, was den dielektrischen Verlustfaktor und damit das Sensorrauschen erhöht. Durch einen zusätzlichen Nachätzprozess mit angepassten Ätzparametern wurde die Kristallstruktur geschont. Außerdem wurden entstandene Schäden durch Temperung der fertigen Sensorchips ausgeheilt. Messungen des dielektrischen Verlustfaktors an den fertigen Chips sind geplant.

Absorptionsschicht

Für eine hohe und spektral gleichmäßige Empfindlichkeit der Sensoren im Wellenlängenbereich von (2 ... 14) µm wird auf dem empfindlichen Element eine Absorptionsschicht aufgebracht. Es handelt sich um Aluminiumschichten, die unter Nutzung des GLAD-Verfahrens in einem thermischen Verdampfer unter Hochvakuum abgeschieden werden. Bei dem GLAD-Verfahren entsteht eine poröse Schicht mit einem kolumnaren Schichtaufbau, indem die Schicht unter sehr hohem Einfallswinkel bei gleichzeitiger Substratrotation abgeschieden wird [7]. Durch den sich ergebenden graduellen Übergang des Brechungsindex an der Grenzfläche zu Luft verringert sich die Reflexion und die Absorption erhöht sich.

Für besonders dünne Chips ist es wichtig, dass die Wärmekapazität der Absorptionsschicht wesentlich geringer ist als die der LiTaO₃-Schicht. Aluminium (Al) ist aufgrund seiner geringen volumenbezogenen spezifischen Wärmekapazität von 2,4 J/(cm³K) besser dafür geeignet als das bisher verwendete NiCr mit 3,8 J/(cm³K).

Zur Bestimmung der Absorption wurde das Aluminium auf ein Siliziumsubstrat mit einer Goldschicht aufgedampft, dessen diffuse Reflexion mit einem FTIR-Spektrometer Spectrum 2000 (Perkin Elmer) gemessen wurde. Gold ist im interessierenden Wellenlängenbereich von (2 ... 14) μm ein nahezu idealer Reflektor, weshalb der Absorptionsgrad direkt aus dem Reflexionsgrad ermittelt werden kann.

Zunächst wurden bei einem Abscheidewinkel von 83° Schichten mit verschiedenen Dicken hergestellt. Mit 83° wurden bisher bei NiCr gute Ergebnisse erzielt. Auf der einen Seite ist eine gewisse Schichtdicke notwendig, weil die Transmission aufgrund der Porosität der Schicht sonst zu groß ist. Andererseits steigt jedoch die Wärmekapazität mit zunehmender Schichtdicke. Deshalb wurden Dicken von (1 ... 2) μm gewählt. Die gemessene Absorption ist in Abb. 2 dargestellt. Wie zu erwarten war, erzielte die dickste Schicht mit einem Absorptionsgrad größer als 80 % bis zu einer Wellenlänge von ca. 14 μm den höchsten Wert. Außerdem wird die Kurve glatter, je dicker die Schicht ist.

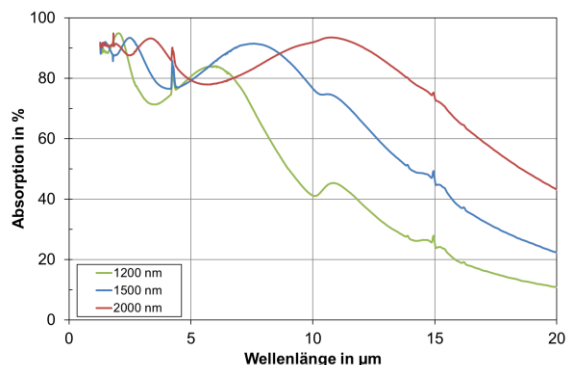


Abb. 2: Absorption von Al-Schichten mit einem Abscheidewinkel von 83° und verschiedenen Dicken.

Hier ist anzumerken, dass die angegebenen Dicken die angestrebten Sollwerte darstellen. Während der Schichtabscheidung wurde die Dicke mit einem Schwingquarz gemessen. Die Messung erfolgte jedoch bei einem Einfallswinkel von 0° , weil der Quarz nicht geneigt ist. Die Schichtdicke auf dem Substrat ist aber vom tatsächlichen Abscheidewinkel abhängig. Deshalb wurden die Dicken näherungsweise mit einem Profilometer Dektak 8000 (Veeco) gemessen. Diese Werte sind jedoch auch nur Schätzungen, weil die Messnadel die weiche Absorptionsschicht leicht eindrückt. Der sogenannte Toolingfaktor beschreibt das Verhältnis aus der

tatsächlichen Schichtdicke und der vom Quarz gemessenen.

Im nächsten Schritt wurden Schichten mit einer gleichen Dicke von 2 μm unter verschiedenen Winkeln abgeschieden. Wie in Abb. 3 zu erkennen ist, erhöht sich die Absorption bis zu einem Winkel von 83° zunächst, weil die Schicht poröser wird und dadurch die Reflexion an der Oberfläche abnimmt. Bei 85° sinkt die Absorption jedoch wieder ab, weil aufgrund steigender Transmission jetzt zu viel Strahlung hindurch gelassen wird.

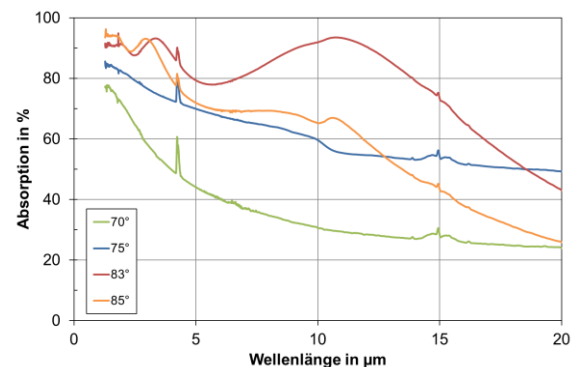


Abb. 3: Absorption von Al-Schichten mit einer Dicke von 2 μm und verschiedenen Abscheidewinkeln.

Um die Absorption weiter zu erhöhen und zu glätten, wurden im nächsten Schritt Gradientenschichten hergestellt, bei denen der Winkel während der Beschichtung variiert wird. Dadurch wird der graduelle Übergang des Brechungsindex an der Grenzfläche zu Luft noch verstärkt. Begonnen wird mit einer dünnen Schicht bei einem kleineren Winkel zwischen 60° und 70° . Darauf folgt dann die nächste dickere und porösere Schicht. In Abb. 4 ist der Absorptionsgrad von zwei Gradientenschichten dargestellt, die beide ähnliche Ergebnisse erzielen. Die erste Schicht besteht aus vier Lagen mit Winkeln von 60° bis 85° , die zweite aus nur 3 Lagen mit Winkeln zwischen 70° und 83° . Im Vergleich zu der bei 83° abgeschiedenen 2 μm dicken Monoschicht steigt die Absorption im Bereich von (5 ... 10) μm um ca. 5 % an.

Von den Gradientenschichten wurde zusätzlich die Transmission bestimmt, indem die Schicht auf ein lichtdurchlässiges Substrat aus Kaliumbromid (KBr) abgeschieden wurde. KBr hat einen Transmissionsgrad von ca. 90 % in einem großen Wellenlängenbereich von (2 ... > 20) μm . Die Messungen mit dem FTIR ergaben eine vernachlässigbare Transmission. Bei dünneren Schichten, die eine gewisse

Maß an Transmission aufweisen, muss das Gesamtsystem aus Absorber-Frontelektrode-LiTaO₃-Rückelektrode betrachtet werden. Mit Hilfe der Eigenabsorption des LiTaO₃-Chips kann die Schichtdicke dann verringert werden, ohne die Gesamtabsorption zu senken.

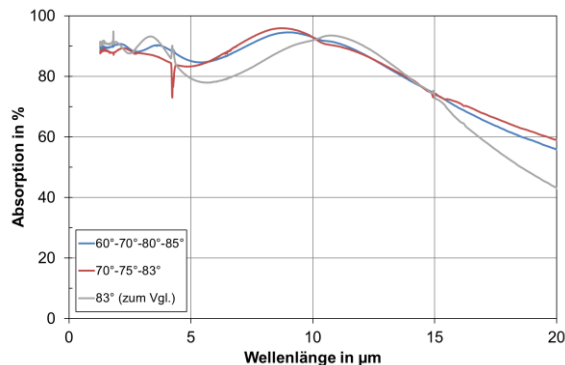


Abb. 4: Absorption von Al-Gradientenschichten mit einer Dicke von 2 µm. Die Monoschicht bei 83° mit derselben Dicke ist zum Vergleich dargestellt.

Schlussfolgerungen

Ultradünne LiTaO₃-Wafer mit Dicken von 20 µm oder 25 µm zeigen im Durchschnitt Dickenunterschiede von (1 ... 2) µm über eine Fläche von (22 x 22) mm². Durch Ionenstrahlätzen konnten 1 µm dünne Chips mit guter Ausbeute hergestellt werden, indem in mehreren Schritten geätzt wurde und Chips, die schon die gewünschte Enddicke erreicht hatten, manuell abgedeckt wurden. Die im Ausgangszustand vorhandene Inhomogenität verringerte sich dabei meistens nur wenig. Die zusätzliche Inhomogenität des Ätzprozesses konnte durch Drehen der Wafer nach halber Ätztiefe ausgeglichen werden.

Weiterhin wurden Materialeigenschaften der Ausgangswafer charakterisiert, wobei der dielektrische Verlustfaktor für das Sensorrauschen von großer Bedeutung ist. Der Einfluss des IBE auf den Verlustfaktor muss noch untersucht werden.

Die auf Basis des GLAD-Verfahrens entwickelte Absorptionsschicht aus Aluminium erreichte einen Absorptionsgrad von (80 ... 90) % für Wellenlängen von (2 ... 14) µm. Das beste Ergebnis erzielte eine 2 µm dicke Gradientenschicht. Im nächsten Schritt werden ultradünne Chips beschichtet, deren schwarze und spektrale Empfindlichkeit gemessen wird.

Danksagung

Die Forschungsarbeiten werden im Rahmen des SAB-Verbundprojekts „PyroGas“ (EFRE-Technologieförderung) mit Mitteln der Europäischen Union kofinanziert und durch Steuermittel auf Grundlage des vom Sächsischen Landtags beschlossenen Haushaltes mitfinanziert.

Literaturnachweis

- [1] M. Schossig, V. Norkus, G. Gerlach, Dielectric and pyroelectric properties of ultrathin, monocrystalline lithium tantalate, *Infrared Physics and Technology*, 63, 5-41 (2014); doi: 10.1016/j.infrared.2013.12.005
- [2] X. Xiao, S. Liang, J. Si, Q. Xu, H. Zhang, L. Ma, C. Yang, X. Zhang, Performance of LiTaO₃ crystals and thin Films and their application, *Crystals*, 13, 1233 (2023); doi: 10.3390/cryst13081233
- [3] J. Luo, W. Luo, K. Zhang, X. Sun, Y. Shuai, T. Wang, C. Wu, W. Zhang, High specific detectivity infrared detector using crystal ion slicing transferred LiTaO₃ single-crystal thin films, *Sensors and Actuators A: Physical*, 300, 111650 (2019); doi: 10.1016/j.sna.2019.111650
- [4] M. Schossig, Ultradünne, freitragende Lithiumtantalat-Elemente für hochdetektive Infrarotsensoren, *Dissertation, Technische Universität Dresden* (2012).
- [5] M. Schossig, V. Norkus und G. Gerlach, Infrared responsivity of pyroelectric detectors with nanostructured NiCr thin-film absorber, *IEEE Sensors Journal*, 10, 1564-1565 (2010); doi: 10.1109/JSEN.2010.2046162
- [6] L. Zhao, X. He, J. Li, X. Gao, J. Jia, Electrospayed carbon-based black coatings for pyroelectric detectors, *Sensors and Actuators A: Physical*, 196, 16-21 (2013); doi: 10.1016/j.sna.2013.03.030
- [7] M. Taschuk, M. Hawkeye und M. Brett, Glancing Angle Deposition. In: P. Martin (Hrsg.): *Handbook of Deposition Technologies for Films and Coatings*, 3. Aufl., Kapitel 13, William Andrew Publishing, S. 621-678 (2010).

Verbesserte Schätzung der Temperaturverteilung in gasdurchströmten Festbetten mittels elektrischer Kapazitätstomografie

Michael Weiss¹, Andreas Jess², Gerhard Fischerauer¹

¹*Lehrstuhl für Mess- und Regeltechnik und Zentrum für Energietechnik (ZET),
Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth*

²*Lehrstuhl für chemische Verfahrenstechnik und Zentrum für Energietechnik (ZET),
Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth*

Zusammenfassung

Diese Arbeit zeigt eine neue Möglichkeit auf, wie die auf Basis der Elektrischen Kapazitätstomografie (ECT) rekonstruierte Temperaturverteilung in einem zylindrischen Festbett durch Einbezug der physikalischen Eigenschaften des realen Systems besser geschätzt werden kann. Im Gegensatz zu bisherigen Beiträgen wird die radial ortsabhängige Porosität innerhalb der Schüttung bei der Erstellung der zur Rückrechnung verwendeten Empfindlichkeitsmatrix in das zweidimensionale Finite-Elemente-Modell einer ECT-Elektrodenanordnung einbezogen. Dabei wird die gemessene Temperaturabhängigkeit der relativen Permittivität des Schüttguts mittels einer Effektivmedientheorie auf lokale Bereiche konstanter Porosität übertragen, für die sich daraus jeweils ein eigener funktionaler Temperatur-Zusammenhang ergibt. Diese lokal gültigen Temperaturfunktionen fließen mit in das FE-Modell ein, welches in dieselben Porositätsbereiche unterteilt ist. Unter pixelweiser Zuordnung der Temperaturfunktionen wird die Empfindlichkeitsmatrix für zwei Referenztemperaturen bestimmt. Schließlich werden auf dieser Basis verschiedene Rückrechenalgorithmen durch Rekonstruktion simulativ generierter Temperaturverteilungen („Phantome“) miteinander verglichen. Die Ergebnisse zeigen, dass die Berücksichtigung der lokalen Porositätsunterschiede zu einer Verbesserung der Rückrechenqualität führt und die tatsächliche Temperaturverteilung besser geschätzt werden kann.

Keywords: Elektrische Kapazitätstomografie, Temperaturprofil, Festbett, Effektivmedientheorie, Inverses Problem.

1. Einführung

Die Untersuchung der Temperaturverhältnisse in gasdurchströmten Festbetten ist bereits seit vielen Jahrzehnten Gegenstand wissenschaftlicher Arbeiten [1, 2]. Insbesondere im Bereich der chemischen Verfahrenstechnik spielt die Kenntnis des Wärmetransports und der resultierenden axialen und radialen Temperaturprofile innerhalb zylindrischer Schüttungen eine entscheidende Rolle für den Betrieb von Festbettreaktoren bei stark exothermen, heterogen katalysierten Gasphasenreaktionen. Bisher wurden neben der Messung der Temperatur mithilfe von Thermoelementen verschiedene alternative, zum Teil nichtinvasive Messmethoden erforscht. Darunter befindet sich auch die Elektrische Kapazitätstomografie (ECT), für welche in einer vorangegangenen Arbeit gezeigt wurde, dass das Messverfahren grundsätzlich in der Lage ist, Temperaturgradienten innerhalb einer Schüttung zu erfassen [3]. Dabei wichen allerdings die mittels

der ECT rückgerechneten Temperaturen teils stark von Vergleichswerten ab, die mit Thermoelementen durchgeführt worden sind. Eine Ursache hierfür mag der Umstand sein, dass die Porosität innerhalb zylindrischer Festbetten radial ortsabhängig ist. Damit geht einher, dass auch die Verteilung der relativen Permittivität, welche durch das ECT-Verfahren berechnet wird, jene radiale Ortsabhängigkeit aufweist. Dieser Aspekt wurde jedoch in den bisherigen Untersuchungen vernachlässigt. Dieser Beitrag beschäftigt sich daher damit, wie die Porositätsverteilung eines zylindrischen Festbetts bei der Schätzung der Temperaturverteilung mittels ECT berücksichtigt werden kann, um die Ergebnisqualität zu verbessern.

2. Die Elektrische Kapazitätstomografie

An diesem Messverfahren wird bereits seit einigen Jahrzehnten geforscht, es findet bislang überwiegend zur Überwachung von Mehrphasengemischen in Rohrleitungen auch kommerzielle Anwendung [4, 5].

Der typische Messaufbau eines ECT-Messsystems umfasst eine Elektrodenanordnung, eine Umschaltel Elektronik sowie eine Recheneinheit. Die m Elektroden sind dabei beispielsweise ringförmig um ein Rohr angeordnet. Über die Umschaltel Elektronik wird nacheinander jede der M möglichen unabhängigen Kombinationen von Elektrodenpaaren durchgeschaltet und die Kapazität zwischen den jeweiligen Elektroden über die Beaufschlagung mit einer bekannten Wechselspannung gemessen. Aus den gemessenen Kapazitäten schätzt anschließend die Recheneinheit mittels eines geeigneten Rückrechenalgorithmus die Verteilung der relativen Permittivität $\varepsilon_r(x, y)$ innerhalb des Rohrquerschnitts („Detektionsbereich“).

Für die meisten Rückrechenalgorithmen wird die geometrieabhängige Empfindlichkeitsmatrix $\mathbf{S}$ der Elektrodenanordnung benötigt. Sie beschreibt die Kapazitätsänderungen ΔC infolge lokaler Veränderungen $\Delta \varepsilon_r$ der relativen Permittivität im Detektionsbereich [6]:

$$\frac{\Delta C}{M \times 1} = \mathbf{S} \cdot \frac{\Delta \varepsilon_r}{N \times 1} \quad (1)$$

$\mathbf{S}$ wird in der Praxis simulativ bestimmt, indem die Elektrodenanordnung in einem Finite-Elemente-(FE-)Programm modelliert, der Detektionsbereich in N Pixel diskretisiert und mit einer unteren Referenzpermittivität $\varepsilon_{r, \min}$ belegt wird. Anschließend werden die Pixel nacheinander einzeln mit der gleichen oberen Referenzpermittivität $\varepsilon_{r, \max}$ belegt und jeweils alle M Kapazitäten numerisch berechnet. Aus den bekannten Kapazitäts- und Permittivitätsänderungen werden anschließend M Empfindlichkeitskoeffizienten s für jedes der N Pixel berechnet – die Einträge der $M \times N$ -Matrix $\mathbf{S}$ [7]. Aufgabe der Rückrechenalgorithmen ist dann die Lösung des inversen Problems von Gl. (1). Hierbei werden die Kapazitäten, die Zeilensummen der Empfindlichkeitsmatrix und die Permittivitäten üblicherweise auf einen Wertebereich von 0 bis 1 normiert. Die normierten Größen werden im Nachfolgenden mit $\bar{\lambda}$, $\hat{\mathbf{S}}$ und $\bar{g}$ bezeichnet. In der Literatur ist bereits eine Vielzahl von Rückrechenalgorithmen vorgestellt worden. Für die hier angestellten Betrachtungen wurden zu Vergleichszwecken ein direkter sowie drei etablierte iterative Algorithmen ausgewählt, welche im Folgenden angegeben werden.

Als direktes Verfahren wird die *sensitivity coefficient back-projection* (SBP) nach Kotre [8] und Wang [9] verwendet. Hierbei berechnet sich die normierte Permittivität des k -ten Pixels gemäß

$$g_k \approx \frac{\sum_{j=1}^M \lambda_j \cdot \hat{\mathbf{S}}_{j,k}}{\sum_{j=1}^M \hat{\mathbf{S}}_{j,k}} \quad (2)$$

Die verwendeten iterativen Verfahren sind der *projizierte Landweber-Algorithmus* (PLW) [10]

$$\bar{g}_{i+1} = \bar{g}_i + \alpha \cdot \hat{\mathbf{S}}^T \cdot (\bar{\lambda}^T - \hat{\mathbf{S}} \cdot \bar{g}_i) \quad (3)$$

mit der Projektion für das k -te Pixel

$$P[g_{i+1}^k] = \begin{cases} 0 & \text{falls } g_{i+1}^k < 0 \\ g_{i+1}^k & \text{falls } 0 \leq g_{i+1}^k \leq 1 \\ 1 & \text{falls } g_{i+1}^k > 1 \end{cases} \quad (4)$$

die *iterative Tikhonov-Regularisierung* (ITK) [6]

$$\bar{g}_{i+1} = \bar{g}_i + (\hat{\mathbf{S}}^T \hat{\mathbf{S}} + \gamma \mathbf{I})^{-1} \cdot \hat{\mathbf{S}}^T (\bar{\lambda} - \hat{\mathbf{S}} \cdot \bar{g}_i) \quad (5)$$

sowie die *simultaneous iterative reconstruction technique* (SIRT) [6, 11]

$$\bar{g}_{i+1} = \bar{g}_i - \beta \hat{\mathbf{S}}^T \cdot \frac{\hat{\mathbf{S}} \cdot \bar{g}_i - \bar{\lambda}}{\text{diag}(\hat{\mathbf{S}} \cdot \hat{\mathbf{S}}^T)} \quad (6)$$

Dabei sind α , β und γ Regularisierungs- bzw. Wichtungsparemeter, $\mathbf{I}$ die Einheitsmatrix sowie i der Index des aktuellen Iterationsschritts. Tab. 1 enthält die hier für Berechnungen verwendeten Werte der Parameter.

Tab. 1: Verwendete Parameter für die iterativen Algorithmen

PLW	SIRT	ITK
$\alpha = 1$	$\beta = 1$	$\gamma = 0,1$
jeweils 200 Iterationsschritte		

Für eine tiefergehende Betrachtung des ECT-Verfahrens und der Rückrechenalgorithmen wird an dieser Stelle auf die einschlägige Fachliteratur verwiesen [6, 12].

3. Effektive relative Permittivität relevanter Schüttgüter u. radiales Porositätsprofil

Um mit der ECT auf die Temperaturverteilung innerhalb einer Festbettschüttung schließen zu können, muss die relative Permittivität des Schüttguts ausreichend stark von der Temperatur abhängen. In vorangegangenen Arbeiten wurden bereits verschiedene Nickelkatalysatoren mittels impedimetrischer Messungen auf ihre Temperaturabhängigkeit hin untersucht [3, 13]. Diese konnte zwar nachgewiesen werden, die Form der einzelnen Katalysatorpartikel (nadelförmig, zylindrische Pellets) erschwert jedoch die weiteren Betrachtungen. Daher haben wir auf dieselbe Art und Weise zwei Partikelfraktionen des kugelförmigen, kommerziellen Nickel-Katalysators

Meth135 der Firma Clariant untersucht. Abb. 1 zeigt die ermittelte Temperaturabhängigkeit der effektiven relativen Permittivität $\varepsilon_{r,eff}(\vartheta)$ im Vergleich zu den bisher untersuchten Katalysatoren. Bezogen auf den Wert bei Raumtemperatur hängen die beiden Partikelfraktionen des Meth135 am stärksten von der Temperatur ab, wobei die Kurve der Schüttung aus größeren Partikeln leicht unterhalb der Kurve der kleineren Partikel liegt.

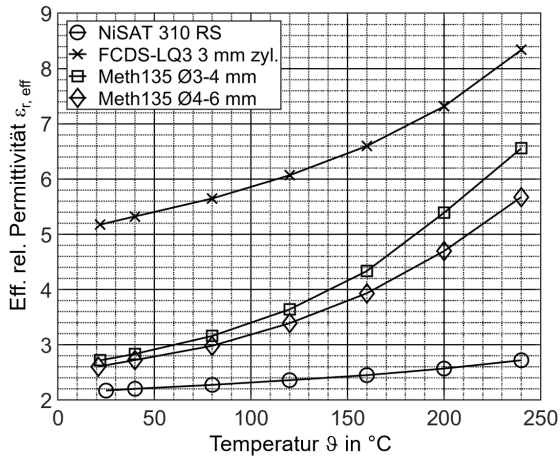


Abb. 1: Temperaturabhängigkeit der eff. rel. Permittivität von Nickel-Trägerkatalysatoren in vollständig oxidiertem Zustand (angelehnt an [3]).

Für alle nachfolgenden Betrachtungen wurde lediglich die Kurve der größeren Fraktion mit Partikeldurchmessern zwischen 4 mm und 6 mm verwendet, da diese den weitaus größeren Teil am gesamten Schüttgut ausmacht und sich daher besser für weiterführende Untersuchungen eignet.

Bei den Katalysatorpartikeln handelt es sich um kugelige, raue, leicht unrunde Partikel. In der Literatur findet man für diesen Fall bei Giese [14] die empirisch ermittelte Näherungsformel für das radiale Porositätsprofil einer in z-Richtung unendlich ausgedehnten, monodispersen Rohrschüttung:

$$\psi(r) = \psi_{\infty} \left(1 + 1,36 \cdot \exp \left[-5 \cdot \frac{R-r}{d_p} \right] \right). \quad (7)$$

Dabei ist ψ_{∞} die Schüttungsporosität der unendlich ausgedehnten Schüttung ohne Wand Einfluss, d_p der Partikeldurchmesser, R der Rohrrinnenradius sowie r die radiale Ortskoordinate mit Ursprung auf der Rohrachse. Aus einer Stichprobe von 100 Partikeln der betrachteten Fraktion hat sich mit einer näherungsweise Normalverteilung für den mittleren Partikeldurchmesser ein Wert von 5,00 mm ergeben, vgl. Abb. 2.

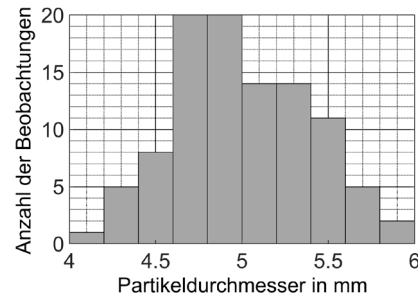


Abb. 2: Verteilung des Partikeldurchmessers der Fraktion Ø4 bis 6 mm von Meth135 in einer Stichprobe von 100 zufällig ausgewählten Partikeln.

Die effektive Bettporosität wurde gemäß der angegebenen Näherungsformel zu

$$\psi_{eff} \approx 1 - \frac{\bar{\rho}_b}{\bar{\rho}_p} = 1 - \frac{1,05 \text{ g/cm}^3}{1,81 \text{ g/cm}^3} \approx 0,42 \quad (8)$$

bestimmt, wobei $\bar{\rho}_p$ die aus der Masse und den aufsummierten Partikelvolumina der Stichprobe errechnete mittlere, scheinbare Partikeldichte und $\bar{\rho}_b$ die aus zehn Wiederholungsmessungen gemittelte Schüttdichte ist [15].

Der Rohrradius wurde im Folgenden zu $R = 20$ mm gewählt, da dies dem Innenwandradius der Temperatur-Messzelle sowie des in Abschnitt 5 gezeigten FE-Modells eines ECT-Messaufbaus entspricht. Die Porosität ψ_{∞} wurde schließlich als frei wählbarer Parameter so gewählt, dass die mit Gl. (7) über

$$\overline{\psi(r)} = \frac{2}{r_{max}^2 - r_{min}^2} \int_{r_{min}}^{r_{max}} r \cdot \psi(r) dr \quad (9)$$

berechnete mittlere Porosität dem Wert aus Gl. (8) entspricht. Dabei wurde $r_{max} = R = 20$ mm sowie $r_{min} = 1,5$ mm gewählt, was dem Radius des keramischen Führungsrohrs im Detektionsbereich des in Abschnitt 5 gezeigten FE-Modells entspricht. Dies führt auf einen Wert von $\psi_{\infty} \approx 0,371$.

Das daraus resultierende Porositätsprofil $\psi(r)$ ist Ausgangspunkt für die Berücksichtigung der lokalen Porositätsunterschiede im Festbett bei der simulativen Generierung der Empfindlichkeitsmatrix. In Abb. 3 ist das berechnete Porositätsprofil unter Vernachlässigung des Einflusses des Führungsrohrs in der Rohrmittte (vgl. Abb. 5) über dem Abstand r zur Rohrachse aufgetragen. An der Rohrwand erreicht die Porosität ihr Maximum, welches in dieser empirischen Näherungsformel unterhalb des idealen Werts von 1 liegt. In Wandnähe nimmt die Porosität zur Schüttungsmittte hin stark ab und strebt gegen den Wert von ψ_{∞} .

Die rote Line in Abb. 3 weist einen treppenförmigen Verlauf auf und stellt die abschnittsweise Mittelung der blauen Kurve innerhalb des Detektionsbereichs des FE-Modells von Abschnitt 5 dar. Dazu wurde der Bereich zwischen $r_{\min}$ und $r_{\max}$ in 20 gleich große Abschnitte („Ringe“) unterteilt, innerhalb derer die Porosität analog zu Gl. (9) gemittelt wurde.

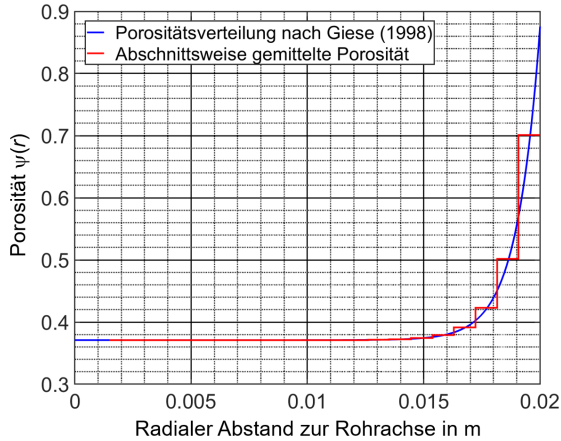


Abb. 3: Auf Basis von stichprobenartigen Partikelmessungen ermitteltes radiales Porositätsprofil für die Partikelfraktion Ø4 bis 6 mm von Meth135.

4. Effektivmedientheorie zur Berücksichtigung lokaler Porositätsunterschiede

Auf Basis der Porositäts-Abschnittswertelwerte lässt sich für jeden Abschnitt eine relative Permittivität abschätzen, welche den jeweiligen Luftanteil in der Schüttung berücksichtigt. Dafür wird eine Effektivmedientheorie benötigt, welche das Verhalten einer dicht gepackten Kugelschüttung möglichst genau beschreibt. Sihvola und Kong [16] haben eine Effektivmedienformel hergeleitet, welche mit einem empirischen Parameter nach Robinson und Friedman [17] auch zufällig und dicht gepackte, monodisperse Kugelschüttungen gut beschreibt. Demnach ergibt sich die effektive rel. Permittivität $\varepsilon_{r,\text{eff}}$ aus der rel. Permittivität der $\varepsilon_{r,p}$ der Partikel und der rel. Permittivität des umgebenden Mediums $\varepsilon_{r,m}$ mit der impliziten Formel

$$\varepsilon_{\text{eff}} = \varepsilon_m + \frac{f \cdot (\varepsilon_p - \varepsilon_m) \cdot [\varepsilon_m + a \cdot (\varepsilon_{\text{eff}} - \varepsilon_m)]}{\varepsilon_m + a \cdot (\varepsilon_{\text{eff}} - \varepsilon_m) + \frac{1}{3}(\varepsilon_p - \varepsilon_m)}, \quad (10)$$

$$1 - \frac{\frac{1}{3} \cdot f \cdot (\varepsilon_p - \varepsilon_m)}{\varepsilon_m + a \cdot (\varepsilon_{\text{eff}} - \varepsilon_m) + \frac{1}{3}(\varepsilon_p - \varepsilon_m)}$$

wobei $f = 1 - \psi$ der Feststoffanteil im effektiven Medium ist. Robinson und Friedman

schlagen für den Parameter a für dicht gepackte Kugeln einen Wert von $a = 0,2$ vor.

Gl. (10) wurde nun in dieser Arbeit verwendet, um von der gemessenen eff. Permittivität der Schüttung sowie der rel. Permittivität von Luft ($\varepsilon_{r,L} \approx 1$) zunächst auf die unbekannte rel. Permittivität der Partikel zu schließen. Die leichte Polydispersität der Schüttung (vgl. Abb. 2) wird dabei vernachlässigt.

Setzt man also für ε_{eff} die Werte von $\varepsilon_{r,\text{eff}}(\vartheta)$ aus Abb. 1, $\varepsilon_{r,L}$ für ε_m sowie die eff. Bettporosität ψ_{eff} aus Gl. (8) für ψ ein, erhält man durch numerische Lösung für alle betrachteten Temperaturen einen Schätzwert für die rel. Permittivität der Katalysatorpartikel, also den Zusammenhang $\varepsilon_{r,p}(\vartheta)$.

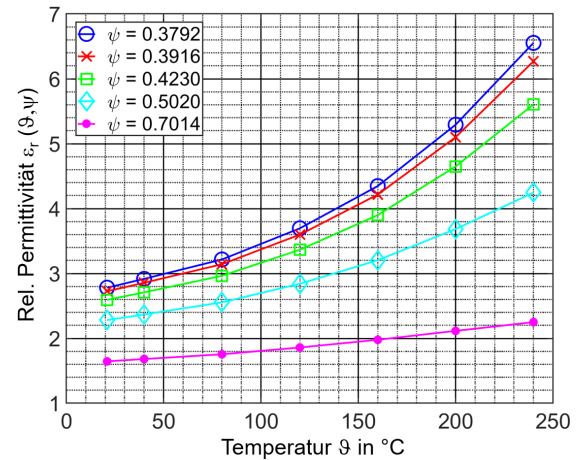


Abb. 4: Auf Basis der Effektivmedientheorie aus Gl. (10) mit $a = 0,2$ berechnetes Kennlinienfeld $\varepsilon_r(\vartheta, \bar{\psi}_i)$ für die Temperaturabhängigkeit der rel. Permittivität in den einzelnen Ringabschnitten aus Abb. 3.

Dieser Temperatur-Zusammenhang wird nun wiederum in Gl. (10) eingesetzt, wobei nun die rel. Permittivität des effektiven Mediums für beliebige Porositäten berechnet werden kann. Dies geschieht nun mit den abschnittsweise berechneten mittleren Porositäten aus Abb. 3, sodass für jeden der 20 über den Radius des Detektionsbereichs verteilten Abschnitte eine eigene Temperaturfunktion für die lokal vorherrschende rel. Permittivität berechnet wird. Es ergibt sich insgesamt ein Kennlinienfeld $\varepsilon_r(\vartheta, \bar{\psi}_i)$, welches sowohl von der Temperatur ϑ als auch von der lokalen mittleren Porosität $\bar{\psi}_i$ im Bereich i abhängt.

Da sich die Porosität nur in Wandnähe erkennbar ändert, kommt es auch nur dort zu großen Unterschieden der rel. Permittivität. Abb. 4 zeigt daher nur die Kurven von $\varepsilon_r(\vartheta, \bar{\psi}_i)$, welche zu den fünf Abschnitten

gehören, die der Wand am nächsten sind. Die Abbildung verdeutlicht, dass die rel. Permittivität an der Rohrwand infolge des hohen Luftanteils nur noch eine sehr schwache Temperaturabhängigkeit aufweist, wohingegen die letztere in der Schüttungsmitte noch höher ist, als $\varepsilon_{r,\text{eff}}(\vartheta)$ aus Abb. 1 zunächst hätte vermuten lassen.

5. Generierung einer porositätssensitiven Empfindlichkeitsmatrix

Um die lokalen Porositätsunterschiede auch in der ECT-Rückrechnung zu berücksichtigen, wird zunächst das zweidimensionale FE-Modell an einen bereits realisierten ECT-Aufbau angepasst. Das Modell unterscheidet sich von der bisherigen Arbeit [3] dadurch, dass der Detektionsbereich in dieselben 20 ringförmigen Abschnitte unterteilt wird, wie dies bereits in Abb. 3 bei der Berechnung der lokalen mittleren Porositäten erfolgt ist.

Abb. 5 zeigt das in der Software Ansys Mechanical APDL erstellte Flächenmodell des Detektionsbereichs vor der Vernetzung mit dem modellierten, luftgefüllten (rot) Keramikrohr (grün) im Zentrum sowie der Rohrwand (beige, äußerster Ring) und zwölf Elektroden (ockerfarben) an der Außenwand. Die 20 Ringe des Detektionsbereichs werden für bessere Lesbarkeit mit zwei verschiedenen Blautönen dargestellt. Nach Vernetzung enthält dieser Bereich insgesamt 2780 Pixel, welche ringweise ausgewählt und mit unterschiedlichen Permittivitäten belegt werden können.

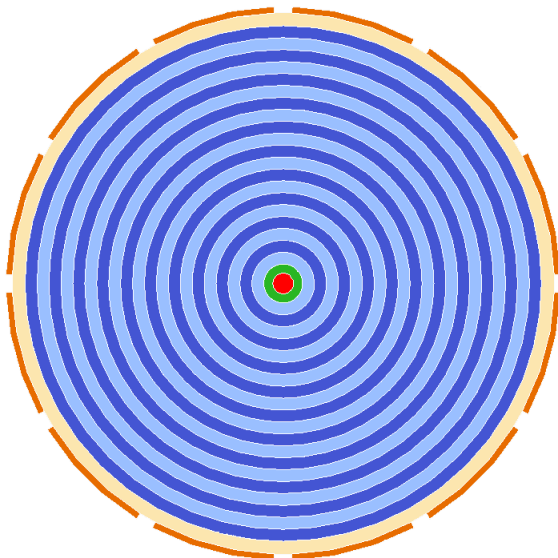


Abb. 5: Flächenmodell des Detektionsbereichs vor der Vernetzung. Zu den Farben siehe Erklärung im Text.

Zu Vergleichszwecken sind zwei verschiedene Empfindlichkeitsmatrizen, auf die in Abschnitt 2 geschilderte Weise generiert worden. Dabei galt in beiden Fällen, dass die untere Referenzpermittivität $\varepsilon_{r,\text{min}}$ dem Wert von Meth135 bei Raumtemperatur (21 °C) und die obere Referenzpermittivität $\varepsilon_{r,\text{max}}$ einer Katalysatortemperatur von 200 °C entsprechen soll. Im ersten Fall wurde jedoch für die Belegung der einzelnen Pixel des Detektionsbereichs jeweils eine konstante, ortsunabhängige rel. Permittivität vorgegeben, die den Werten von $\varepsilon_{r,\text{eff}}(\vartheta)$ für die Partikelfraktion Ø4 mm bis 6 mm in Abb. 1 entspricht. Dies entspricht dem üblichen Vorgehen beim ECT-Verfahren.

Im zweiten Fall wurde vor der Belegung eines Pixels geprüft, in welchem Ring i des Detektionsbereichs sich das Pixel befindet, um es dann mit der rel. Permittivität aus dem Kennlinienfeld $\varepsilon_r(\vartheta = 200\text{ °C}, \bar{\psi}_i)$ zu belegen. Alle anderen Pixel sind dabei entsprechend ihrer Position mit $\varepsilon_r(\vartheta = 21\text{ °C}, \bar{\psi}_i)$ belegt.

In Abb. 6 sind die niedrigen und hohen Referenzpermittivitäten über dem radialen Abstand zur Rohrachse aufgetragen. Es lässt sich erkennen, dass die Empfindlichkeit auf Permittivitätsänderungen an der Rohrwand über- und im Schüttungskern unterschätzt wird, wenn die radiale Ortsabhängigkeit der Schüttungsporosität bei der Erstellung der Empfindlichkeitsmatrix vernachlässigt wird.

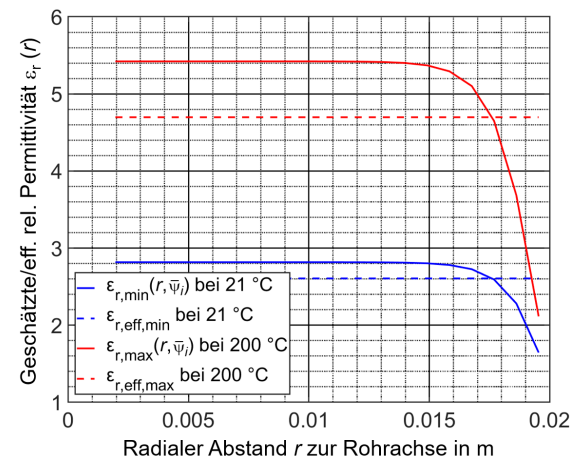


Abb. 6: Für die Generierung der Empfindlichkeitsmatrix verwendete Referenzpermittivitäten, aufgetragen über dem radialen Abstand zur Rohrachse.

6. Rückrechnung simulativ generierter Temperaturverteilungen

Um die Rückrechenqualität objektiv beurteilen zu können, muss die zu rekonstruierende Permittivitätsverteilung bzw. hier Temperaturverteilung bereits im Voraus bekannt sein. Dazu werden mithilfe des FE-Modells

Temperaturverteilungen simulativ generiert (sog. „Phantome“). Dabei werden die Ringe des Detektionsbereichs mit der Permittivitätsverteilung gemäß des Kennlinienfelds $\varepsilon_r(\vartheta, \bar{\psi}_i)$ belegt, welche dem gewünschten Temperaturprofil entspricht. Damit wird die Ortsabhängigkeit der Porosität berücksichtigt, denn diese ist auch in einer realen Messung vorhanden und hat einen unvermeidbaren Einfluss auf die gemessenen Kapazitäten. Um alle Ringe für beliebige Temperaturwerte mit den korrespondierenden rel. Permittivitäten belegen zu können, werden alle Temperaturkurven des Kennlinienfelds mit Polynomen 3. Ordnung interpoliert, so wie dies für die eff. rel. Permittivität bereits in einer vorangegangenen Arbeit getan wurde [3].

Abb. 7 zeigt die für die Beurteilung verwendeten Temperaturphantome und deren zugehörige Permittivitätsverteilungen. Die Phantome entsprechen jeweils einem linearen, einem quadratischen sowie einem kubischen Temperaturprofil. Die minimale Temperatur beträgt jeweils $\vartheta_{\min} = 100^\circ\text{C}$ und tritt im Falle des linearen Profils im innersten Ring des Detektionsbereichs, in den anderen beiden Fällen auf der Rohrachse ($r = 0$) auf. Die maximale Temperatur beträgt jeweils $\vartheta_{\max} = 200^\circ\text{C}$ und tritt jeweils im äußersten Ring des Detektionsbereichs auf.

Für die Auftragung der Temperaturen und Permittivitäten über dem Abstand r zur Rohrachse wurden jeweils die Abschnittsmittelpunkte der einzelnen Ringabschnitte verwendet.

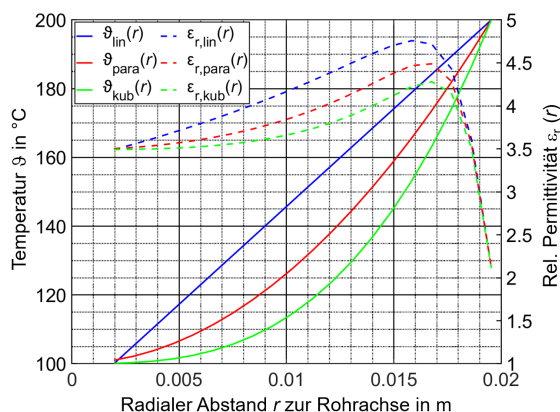


Abb. 7: Für die Beurteilung der Rückrechnungsqualität verwendete Temperaturphantome sowie zugehörige Permittivitätsverteilungen, welche sich durch Berücksichtigung der Ortsabhängigkeit der Porosität ergeben.

7. Ergebnisse der Temperaturschätzung

Die Temperaturumrechnung der aus der ECT erhaltenen normierten Pixel-Permittivitäten er-

folgt ebenfalls analog zu der vorangegangenen Arbeit [3]: die entsprechende $\varepsilon_{r,\text{eff}}(\vartheta)$ -Kurve aus Abb. 1 bzw. die Kurven des Kennlinienfelds $\varepsilon_r(\vartheta, \bar{\psi}_i)$ werden einzeln jeweils auf einen Wertebereich von 0 bis 1 normiert. Anschließend werden die Kennlinien invertiert und erneut durch Polynome 3. Grades interpoliert, in welche schließlich die berechneten normierten Pixel-Permittivitäten eingesetzt werden, um die Pixel-Temperaturen zu berechnen.

Abb. 8 vergleicht die geschätzten Temperaturverteilungen am Beispiel des PLW-Algorithmus. Dabei entsprechen die durchgezogenen Linien dem vorgegebenen linearen (blau), quadratischen (rot) und kubischen (grün) Temperaturphantom. Die fein gepunkteten Linien entsprechen der jeweils geschätzten Temperaturverteilung, wenn die Empfindlichkeitsmatrix verwendet wird, welche auf Basis der konstanten, ortsunabhängigen Referenzpermittivitäten erstellt wurde. Die gestrichelten Linien ergeben sich, wenn mit der Empfindlichkeitsmatrix gerechnet wird, welche die Ortsabhängigkeit der Porosität des Festbetts berücksichtigt.

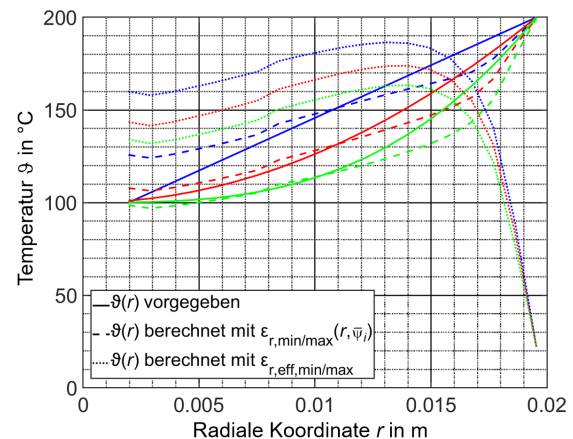


Abb. 8: Vorgegebene Temperaturphantome sowie mit und ohne Berücksichtigung der Ortsabhängigkeit der Porosität des Festbetts unter Verwendung des PLW-Algorithmus geschätzte Temperaturverteilungen (linear = blau, quadratisch = rot, kubisch = grün).

Es ist erkennbar, dass es bei Verwendung konstanter Referenzpermittivitäten nach Rückrechnung im Wandbereich zu einer deutlichen Unterschätzung der relativen Permittivität und damit der vorherrschenden Temperatur kommt. In der Schüttungsmitte wird die Temperatur hingegen überschätzt. Dieses Verhalten folgt aus unmittelbar aus der Betrachtung von Abb. 6 und Abb. 7.

Berücksichtigt man bei der Rückrechnung jedoch die porositätsbedingte Ortsabhängigkeit der Referenzpermittivitäten, wird das vorgegebene Temperaturphantom gut aus den Phantomkapazitäten berechnet. Lediglich beim linearen Temperaturprofil kommt es zu einer größeren Abweichung in der Mitte der Schüttung. Die anderen Rückrechenalgorithmen liefern überwiegend ähnliche Ergebnisse für die betrachteten Fälle.

Um die Qualität der Rückrechnung bewerten zu können, wird der in der Literatur übliche Bildfehler (*image error*, *IE*) [6]

$$IE = \frac{\|\hat{\vartheta} - \bar{\vartheta}\|}{\|\bar{\vartheta}\|} \quad (11)$$

mit den Vektoren der geschätzten Pixel-Temperaturen $\hat{\vartheta}$ sowie der vorgegebenen Pixel-Temperaturen $\bar{\vartheta}$ des jeweiligen Temperaturphantoms, verwendet. Außerdem wird der Korrelationskoeffizient

$$CC = \frac{\sum_{n=1}^N (\hat{\vartheta}_n - \bar{\vartheta})(\vartheta_n - \bar{\vartheta})}{\sqrt{\sum_{n=1}^N (\hat{\vartheta}_n - \bar{\vartheta})^2 \cdot \sum_{n=1}^N (\vartheta_n - \bar{\vartheta})^2}} \quad (12)$$

herangezogen [6]. Hierbei stellen $\bar{\vartheta}$ und $\bar{\vartheta}$ die Mittelwerte aller jeweiligen Pixel-Temperaturen dar, n ist der Pixel-Index.

Tabelle 2 stellt die berechneten Bildfehler und Korrelationskoeffizienten für die verwendeten Rückrechenalgorithmen sowohl für den Fall konstanter also auch für den Fall ortsabhängiger Referenzpermittivitäten gegenüber.

Tab. 2: Rückrechenqualität verschiedener Algorithmen bei Vernachlässigung bzw. Berücksichtigung der radialen Porositätsunterschiede im Festbett.

	$\varepsilon_{r,\text{eff,min/max}}$		$\varepsilon_{r,\text{min/max}}(r, \bar{\psi}_i)$	
Alg.	IE / %	CC / %	IE / %	CC / %
Lineares Temperaturprofil				
SBP	31,55	-69,67	8,67	96,28
PLW	41,20	-54,54	6,67	93,47
SIRT	50,80	-36,47	14,07	72,67
ITK	15,74	-50,53	12,87	71,03
Parabolisches Temperaturprofil				
SBP	35,43	-78,84	8,73	99,14
PLW	44,68	-63,22	7,00	94,94
SIRT	54,00	-43,82	16,54	73,96
ITK	17,09	-59,23	13,72	78,66
Kubisches Temperaturprofil				
SBP	38,77	-86,08	8,78	99,65
PLW	47,78	-72,13	8,46	94,95
SIRT	56,70	-51,85	18,75	73,09
ITK	18,21	-67,90	14,60	81,53

Der Vergleich der Werte verdeutlicht, dass insbesondere der direkte SBP-Algorithmus sowie der iterative PLW-Algorithmus zu den qualitativ besten Rückrechenergebnissen führen. Die in Abb. 8 ersichtlichen Abweichungen führen für den Fall konstanter Referenzpermittivitäten zu großen Bildfehlern und geringeren Korrelationskoeffizienten. Letztere weisen hier ein negatives Vorzeichen auf, was daher rührt, dass die rückgerechneten Pixel-Temperaturen in Wandnähe stark absinken, obwohl die wahre Verteilung einen Temperaturanstieg vorgibt.

Insgesamt lässt sich zusammenfassen, dass die vorgestellte Methode die Rückrechnung auf Permittivitäts- und Temperaturverteilungen in Rohrschüttungen aus kugeligen Partikeln, in denen die radiale Porositätsverteilung in Wandnähe eine starke Ortsabhängigkeit aufweist, erheblich verbessert.

8. Fazit

Es konnte anhand simulationsbasierter Untersuchungen gezeigt werden, dass die Berücksichtigung des Porositätsprofils eines Festbetts aus kugeligen Partikeln bei der Generierung der Empfindlichkeitsmatrix potenziell zu einer signifikanten Verbesserung der Schätzung der Temperaturverteilung mittels ECT beitragen kann. Insbesondere bei thermischen Laborversuchen mit gasdurchströmten Schüttungen hat die radiale Ortsabhängigkeit der Festbettporosität einen unvermeidbaren Einfluss auf die aufgenommenen Messdaten. Falls es gelingt, die physikalischen Eigenschaften des vorliegenden Messaufbaus detailgetreu mit der Finite-Elemente-Methode zu modellieren sowie Störeinflüsse beispielsweise durch Schirmung sowie Signalverarbeitung weitestgehend zu eliminieren, kann der hier vorgestellte Ansatz zu einer erheblichen Verbesserung der Rückrechenqualität beitragen.

Danksagung

Die Autoren danken dem Bayerischen Staatsministerium für Wissenschaft und Kunst für die finanzielle Unterstützung im Rahmen des Graduiertenkollegs „Energieautarke Gebäude“ der Technologieallianz Oberfranken (TAO).

Literaturnachweis

- [1] A.P. de Wasch, G.F. Froment, *Chem. Eng. Sci.* 27, 567-576 (1972); doi: 10.1016/0009-2509(72)87012-X
- [2] A. Zenner *et al.*, *Chem. Eng. Sci.* 201, 424-436 (2019); doi: 10.1016/j.ces.2019.02.028
- [3] M. Weiss, A. Fischerauer, A. Jess, G. Fischerauer, *Meas. Sci. Technol.* 35, 095407 (2024); doi: 10.1088/1361-6501/ad5224

- [4] S.M. Huang, A.B. Plaskowski, C.G. Xie, M.S. Beck, *J. Phys. E: Sci. Instrum.* 22, 173-177 (1989); doi: 10.1088/0022-3735/22/3/009
- [5] H. Wang, W. Yang, *Appl. Therm. Eng.* 176, 115311 (2020); doi: 10.1016/j.applthermaleng.2020.115311
- [6] W. Q. Yang, L. Peng, *Meas. Sci. Technol.* 14, R1-R13 (2003); doi: 10.1088/0957-0233/14/1/201
- [7] Q. Marashdeh, F.L. Teixeira, *IEEE T. Magn.* 40, 1204-1207 (2004); doi: 10.1109/TMAG.2004.825039
- [8] C.J. Kotre, *Physiol. Meas.* 15, A125-A134 (1994); doi: 10.1088/0967-3334/15/2A/017
- [9] M. Wang, *Meas. Sci. Technol.* 13, 101-117 (2002); doi: 10.1088/0957-0233/13/1/314
- [10] W.Q. Yang, S. Liu, *Electron. Lett.* 35, 313-317 (1999); doi: 10.1049/el:19990246
- [11] B. Su *et al.*, *Chem. Eng. J.* 77, 37-41 (2000); doi: 10.1016/S1385-8947(99)00134-5
- [12] W. Yang, *Meas. Sci. Technol.* 21, 042001 (2010); doi: 10.1088/0957-0233/21/4/042001
- [13] M. Weiss, A. Jess, G. Fischerauer, *Proc. SMSI 2023 Conf.: Sensor and Measurement Science Int. (Nuremberg, Germany)*, 259-260 (2023); doi: 10.5162/SMSI2023/E5.3
- [14] M. Giese, Dissertation, TU München (1998)
- [15] M. Stieß, *Mechanische Verfahrenstechnik – Partikeltechnologie* 1, 3. Aufl., Springer-Verlag, Berlin (2009); doi: 10.1007/978/3-540-32552-9
- [16] A.H. Sihvola, J.A. Kong, *IEEE T. Geosci. Remote* 26, 420-429 (1988); doi: 10.1109/36.3045
- [17] D.A. Robinson, S.P. Friedman, *Physica A* 358, 447-465 (2005); doi: 10.1016/j.physa.2005.03.054

Hochtemperatursensoren auf Basis von Keramik-Kompositwerkstoffen

*Bing Feng¹, Hans-Peter Martin¹, Paul Gierth¹, Lars Rebenklau¹, Alexander Michaelis¹,
¹Fraunhofer-Institut für Keramische Technologien und Systeme IKTS, Winterbergstr. 28, 01277
 Dresden, Deutschland*

Zusammenfassung

Keramische Temperatursensoren auf Basis von B_4C/TaB_2 -Kompositwerkstoffen wurden für den Einsatz bei Temperaturen bis 2200 °C entwickelt. Durch Variation des TaB_2 -Gehalts lassen sich die Seebeck-Koeffizienten des B_4C/TaB_2 -Sinterkörpers gezielt einstellen. Zwei Composite mit 4 Gew.-% bzw. 8 Gew.-% TaB_2 wurden als einzelne Thermoelementkomponenten extrudiert und anschließend durch Diffusionsfügen zu einem Thermoelement verbunden. Eine integrierte beheizbare Referenzstelle ermöglicht eine exakte Einstellung der Referenztemperatur bis 800 °C. Funktionstests unter Vakuumbedingungen zeigen ein kontinuierlich ansteigendes Messsignal bis 1925 °C mit einem maximalen Wert von 68 mV. Zyklustests zeigen zudem eine hohe Reproduzierbarkeit und Signalstabilität des Sensors. Die Ergebnisse demonstrieren das Potenzial keramischer Thermoelemente als zuverlässige Sensoren für Berührungsmessung bei Ultrahochtemperaturen. Die Einsatzmöglichkeiten keramischer Thermoelemente können durch Einsatz anderer geeigneter keramischer Werkstoffe noch deutlich hinsichtlich Temperaturbereich, Atmosphäre oder Messauswertung erweitert werden.

Keywords: Hochtemperatursensor, Thermoelement, Keramik, Borcarbid, Referenzstelle

1. Einleitung

Eine präzise und zuverlässige Temperaturmessung ist für die Prozesskontrolle und Produktqualität in zahlreichen industriellen Anwendungen von entscheidender Bedeutung. Thermoelemente gehören aufgrund ihrer Robustheit und ihrer einfachen Bauweise zu den am weitesten verbreiteten Temperatursensoren in der Industrie. Je nach verwendeter Materialkombination decken sie Einsatzbereiche bis 1800 °C ab, beispielsweise Cu–CuNi (max. 400 °C), NiCr–CuNi (max. 900 °C), NiCr–Ni (max. 1300 °C) und PtRh30 %–PtRh6 % (max. 1800 °C) [1-3]. Mit dem zunehmenden Einsatz moderner Hochtechnologien entstehen Prozesse, bei denen die Temperaturen deutlich über 1800 °C liegen, zum Beispiel bei der Kristallzucht, in der Werkstoffproduktion, in der Luft- und Raumfahrt oder in der Energieerzeugung. Eine zuverlässige Temperaturerfassung in diesem Bereich stellt mit den derzeit verfügbaren Messtechnologien nach wie vor eine große Herausforderung dar. Für Messungen über 1800 °C kommen gegenwärtig vorwiegend Thermoelemente auf Basis von Wolfram-Rhenium-Legierungen zum Einsatz. Unter inerten Gasen oder im Vakuum ermöglichen diese Sensoren Messungen bis

2300 °C [4]. Allerdings unterliegen W-Re-Thermoelemente bei hohen Temperaturen ausgeprägten Rekristallisationsprozessen, die zu einer Versprödung der Drähte sowie zu thermoelektrischer Drift führen. Die daraus resultierenden Materialinhomogenitäten reduzieren die Langzeitstabilität und Messgenauigkeit erheblich [5]. Im Temperaturbereich von 2000 – 2300 °C ist daher nur ein kurzzeitiger Betrieb möglich. Alternativ bieten Strahlungspyrometer die Möglichkeit, Temperaturen bis 3000 °C zu erfassen [6]. Ihre Genauigkeit hängt jedoch maßgeblich von der exakten Kenntnis des Emissionskoeffizienten sowie von einem freien optischen Zugang ab. Unter realen Prozessbedingungen können Rauch, Staub, Dampf oder Ablagerungen auf Sichtfenstern den optischen Zugang behindern, wodurch die Messgenauigkeit deutlich reduziert oder eine Messung vollständig verhindert werden kann. Diese Einschränkungen führen zu einem Bedarf an zusätzlichen Messsystemen für Ultrahochtemperaturen. Bereits in den 1990er Jahren wurden keramische Thermoelemente aus Borcarbid und Graphit entwickelt, die Temperaturmessungen bis etwa 2200 °C erlaubten [7]. B_4C zeichnet sich durch hohe thermische Stabilität und einen Seebeck-Koeffizienten von bis zu 300 $\mu V/K$ aus [8]. Frühere Untersuchungen zeigten außerdem,

dass sich der Seebeck-Koeffizient von Borcarbid durch die Zugabe von Metallboriden gezielt einstellen lässt [9].

In der vorliegenden Arbeit wird ein Thermoelement auf Basis von B_4C/TaB_2 -Kompositwerkstoffen entwickelt und experimentell hinsichtlich seiner thermoelektrischen Eigenschaften, der Referenzstelle, der Funktionalität sowie in Zyklustests bis $1940\text{ }^{\circ}\text{C}$ charakterisiert.

2. Experimente

Die Probekörper aus B_4C/TaB_2 -Kompositwerkstoffen zur Charakterisierung wurden mittels drucklosen Sinterns hergestellt. Der TaB_2 -Masseanteil variierte zwischen 2 und 40 Gew.-%. Die Ausgangspulver wurden zunächst in einem Trommelmischer homogenisiert und anschließend zu Grünlingen gepresst. Nach der Pyrolyse erfolgte das Sintern bei $2250\text{ }^{\circ}\text{C}$ unter Argonatmosphäre, wodurch dichte Verbundkörper entstanden.

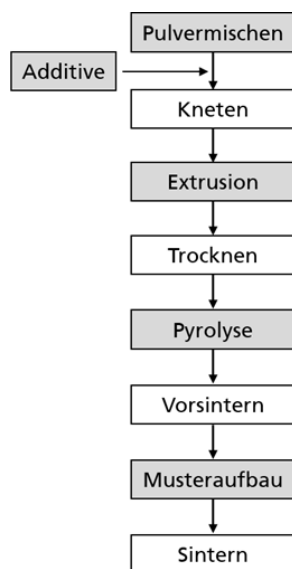


Abb. 1: Schematische Darstellung der Herstellungsschritte von B_4C-TaB_2 -Thermoelement-Komponenten

Die Fertigung des Thermoelements erfolgte nach den in Abbildung 1 dargestellten Arbeitsschritten. Für den Aufbau wurden zwei unterschiedliche Materialkomposite verwendet: ein Komposit mit 4 Gew.-% TaB_2 und ein Komposit mit 8 Gew.-% TaB_2 . Zur Vorbereitung für die Extrusion wurde die Pulvermischung zunächst mit Bindern/Additiven zu einer homogenen Masse geknetet. Die daraus extrudierten Rohre (Schenkel 1) und Stäbe (Schenkel 2) wurden anschließend ausgeheizt und vorgesintert (Abb. 2). Durch mechanische Bearbeitung

wurde ein Formkörper zur Verbindung der beiden Schenkel hergestellt, dessen stoffliche Zusammensetzung der des Rohres (Schenkel 1) entsprach. Nach dem Zusammenbau der Komponenten wurden die beiden Schenkel des Thermoelements bei $2250\text{ }^{\circ}\text{C}$ durch Diffusionsfügen an einem Ende miteinander verbunden [10].



Abb. 2: Vorgesinterte extrudierter Rohre und Stäbe aus B_4C/TaB_2 -Kompositwerkstoffen.

Für die Referenzstelle wurde ein zylindrischer Al_2O_3 -Formkörper konstruiert, wobei die beiden Schenkel darin positioniert und fest mit dem Formkörper verbunden wurden. Jeder Schenkel wurde über einen separaten elektrischen Draht angeschlossen. Die Verbindung der Schenkel mit dem Formkörper und mit den elektrischen Messleitungen erfolgte mithilfe eines Aktivlotes, wodurch eine stabile mechanische und elektrische Integration der Referenzstelle erreicht wird. Zur weiteren Funktionalisierung wurde auf der Oberfläche des Al_2O_3 -Zylinders mittels Dickschichttechnologie eine heizbare Platinschicht aufgebracht, die eine gezielte Temperierung der Referenzstelle im Bereich von 100 bis $800\text{ }^{\circ}\text{C}$ ermöglicht. In Abbildung 3 ist das fertige Thermoelement dargestellt.



Abb. 3: Thermoelement aus B_4C/TaB_2 -Kompositwerkstoffen.

Zur Untersuchung der Funktionalität des Thermoelements wurde dieses in eine Heizanlage eingebaut. Ein W-Heizer diente dabei als Wärmequelle. Die thermoelektrischen Spannungsmessungen sowie die Zyklentests wurden unter Vakuum (10^{-6} mbar) bis $1940\text{ }^{\circ}\text{C}$ durchgeführt.

3. Ergebnisse

Abbildung 4 zeigt die Temperaturabhängigkeit des Seebeck-Koeffizienten (S) der gesinterten Kompositwerkstoffe. Im untersuchten Temperaturbereich steigt S nur geringfügig mit der Temperatur, während der Gehalt an TaB_2 im Ausgangspulver seine Höhe maßgeblich bestimmt. Mit zunehmender TaB_2 -Zugabe fällt der Seebeck-Koeffizient deutlich ab und variiert über einen Bereich bis zu $275 \mu V/K$. Die metallähnliche Phase TaB_2 weist im Gegensatz zu B_4C negative Seebeck-Werte auf. Ein höherer Anteil von TaB_2 in den Verbundproben führt zu einer teilweisen Kompensation der positiven Beiträge von B_4C , wodurch der Gesamt-Seebeck-Koeffizient reduziert wird. Damit lässt sich der Seebeck-Koeffizient der Kompositwerkstoffe gezielt über die Zusammensetzung steuern und eröffnet Möglichkeiten bei der Auswahl von Thermoelement-Schenkeln mit unterschiedlichen Seebeck-Koeffizienten.

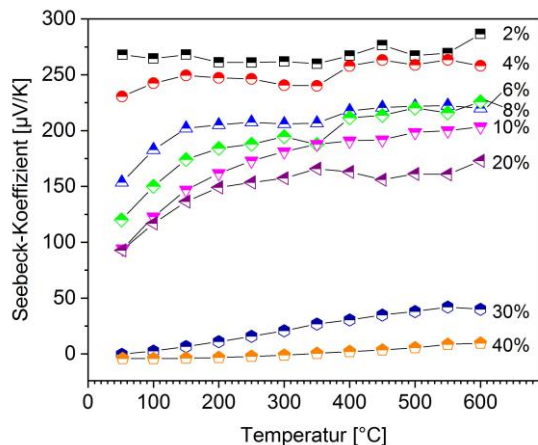


Abb. 4: Seebeck-Koeffizient von B_4C - TaB_2 -Kompositwerkstoffen in Abhängigkeit von Temperatur und TaB_2 -Gehalt der Ausgangspulvermischung.

Die Heizfunktion der mittels Siebdruck aufgetragenen Platinschicht an der Referenzstelle wurde experimentell bis zu einer Temperatur von $800^\circ C$ untersucht (Abb. 5). Innerhalb dieses Temperaturbereichs weist das Referenzstellenelement sowohl thermische als auch mechanische Stabilität bei gleichzeitig ausreichender thermischer Leitfähigkeit auf. Durch die aktive Regelung der Referenztemperatur kann der thermische Fluss entlang des keramischen Thermoelements gezielt reduziert werden. Darüber hinaus ermöglicht die Variation der Referenztemperatur die Definition mehrerer Bezugspunkte, wodurch die Signalstabilität und die Messgenauigkeit verbessert werden können.

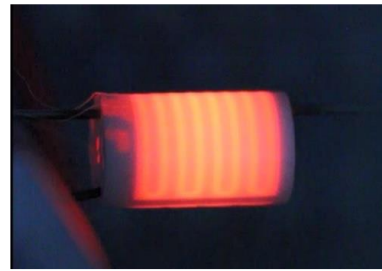


Abb. 5: Test des Referenzstellenelements bei $800^\circ C$.

Der Funktionalitätstest des Thermoelements wurde bis zu einer Temperatur von $1925^\circ C$ durchgeführt. Abbildung 6 zeigt die Abhängigkeit der thermoelektrischen Spannung des Thermoelements von der Temperaturdifferenz ΔT . Da das untersuchte Thermoelement nicht kalibriert ist, wird in der Darstellung ausschließlich der mit der Temperatur korrelierende Spannungsverlauf wiedergegeben. Die Funktionsfähigkeit des aufgebauten Temperatursensors konnte dabei nachgewiesen werden. Es zeigt sich ein mit zunehmender Temperaturdifferenz ansteigender Verlauf der Thermospannung. Bei einer Temperaturdifferenz von etwa $1900 K$ wurde eine thermoelektrische Spannung von ca. $68 mV$ erreicht. Im Vergleich dazu weist ein W-Re-Thermoelement bei einer vergleichbaren Temperaturdifferenz eine Thermospannung von etwa $32 mV$ auf [11]. Der B_4C/TaB_2 -Sensor zeigt somit ein annähernd doppelt so großes thermoelektrisches Signal, was auf eine im Vergleich erhöhte thermoelektrische Empfindlichkeit schließen lässt.

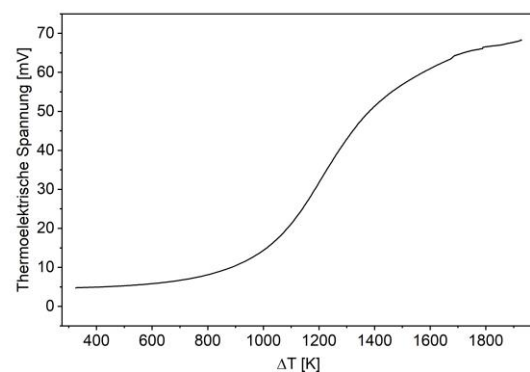


Abb. 6: Thermoelektrische Spannung des Thermoelements in Abhängigkeit von der Temperaturdifferenz ΔT für die Materialkombination $B_4C + 4 \text{ Gew.-% } TaB_2$ und $B_4C + 8 \text{ Gew.-% } TaB_2$.

Der Zyklentest des Thermoelements wurde in vier Messreihen mit jeweils drei Zyklen unter definierten Temperaturintervallen durchgeführt. Die thermoelektrische Spannung wurde in

Abhängigkeit von der Zeit sowie von der Temperaturdifferenz ΔT ausgewertet (Abb. 7). Die Ergebnisse zeigen eine hohe Übereinstimmung der Spannungsverläufe innerhalb der einzelnen Zyklen. Insbesondere stimmen die jeweiligen Maximal- und Minimalwerte der einzelnen Zyklen nahezu überein. Dies weist auf eine gute Reproduzierbarkeit der Messergebnisse sowie auf ein stabiles Verhalten des Thermoelements unter zyklischer thermischer Belastung hin. Die Auswertung der thermoelektrischen Spannung in Abhängigkeit von ΔT zeigt eine geringfügige Abweichung zwischen den einzelnen Zyklen. Diese Abweichungen lassen sich auf einen inhomogenen Temperaturverlauf innerhalb der Testanlage sowie auf leichte Änderung der Referenztemperatur an der Kaltseite zurückführen.

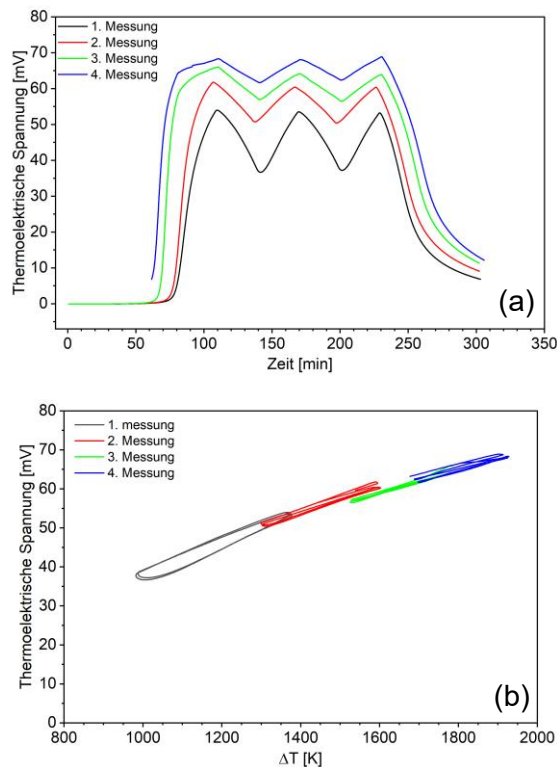


Abb. 7: Vergleich der thermoelektrischen Spannung des Thermoelements im Zykientest in Abhängigkeit von (a) der Zeit und (b) der Temperaturdifferenz ΔT .

4. Ausblick

Funktionalitäts- und Zyklostests zeigen, dass das Thermoelement aus B_4C-TaB_2 -Kompositwerkstoffen reproduzierbar und zuverlässig arbeitet. Für eine vollständige Charakterisierung des Temperatursensors sind weitere Untersuchungen notwendig. Besonders wichtig sind Langzeitstabilitätstests,

die mögliche Drift der Thermospannung bei hohen Betriebstemperaturen ($\geq 1800^\circ C$) über längere Zeiträume aufzeigen können. Darüber hinaus sollte das Thermoelement präzise kalibriert werden, um eine normungsfähige Kennlinie zwischen thermoelektrischer Spannung und Temperaturdifferenz ΔT zu erhalten.

Literaturnachweis

- [1] L., Yang, Y. Zhao, C. Feng, H. Zhou, *Physics Procedia* 22, 95-100 (2011); doi: 10.1016/J.PHPRO.2011.11.015
- [2] P.R.N. Childs, J.R. Greenwood, C.A. Long, *Review of Scientific Instruments* 71, 2959–2978 (2000); doi: 10.1063/1.1305516
- [3] The European Standard EN 60584-1 (1995), Thermocouples - Part 1: Reference Tables. Beuth Verlag: Berlin, Germany,
- [4] J.V. Pearce, C.J. Elliott, G. Machin, O. Ongrai, *AIP Conference Proceedings* 1552, 595-600 (2013); doi: 10.1063/1.4821396
- [5] O. Ongrai, J.V. Pearce, G. Machin, S. Sweeney, S. *Measurement Science and Technology* 22, 105103 (2011); doi: 10.1088/0957-0233/22/10/105103
- [6] J. Hollandt, J. Hartmann, O. Struß, R. Gärtner, *Experimental Methods in the Physical Sciences* 43, 1-56 (2010); doi: 10.1016/S1079-4042(09)04301-X
- [7] K. Hunold, *Advanced Materials & Processes* 9, 94-5 (1986)
- [8] B. Feng, H.-P. Martin, F.-D. Börner, W. Lippmann, M. Schreier, K. Vogel, A. Lenk, I. Veremchuk, M. Dannowski, C. Richter, *Advanced Engineering Materials* 16, 1252-1263 (2014), doi: 10.1002/adem.201400183
- [9] B. Feng, H.-P. Martin, A. Michaelis, *Journal of Electronic Materials* 42, 2314-2319 (2013); doi: 10.1007/s11664-012-2413-y
- [10] H.-P. Martin, B. Feng, L. Rebenklau, P. Gierth, Patent DE102018200548 B3, Keramisches Thermoelement sowie Verfahren zu seiner Herstellung
- [11] R. R. Asamoto, P. E. Novak, *Review of Scientific Instruments* 38, 1047-1052 (1967); doi: 10.1063/1.1720964

Representing Measurement Uncertainty by Gaussian Mixtures

Luisa Hoffmann¹ ; Mohamed Yamen Alayoubi¹; Ali Darijani^{2,3} ; Jürgen Beyerer^{2,3} ;
Michael Heizmann¹ 

¹ Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT)

² Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB)

³ Karlsruhe Institute of Technology, Vision and Fusion Laboratory (IES)

Abstract

This paper presents a probabilistic framework for uncertainty quantification in nonlinear measurement systems based on Gaussian mixture models. Instead of relying on local linearization as often used according to the Guide to the Expression of Uncertainty in Measurement (GUM), the proposed approach represents the measurement model as a joint probability density function of input and output quantities. This joint distribution is approximated using a Gaussian mixture, the parameters of which are estimated from data via the Expectation–Maximization (EM) algorithm. Posterior distributions of the measurand are obtained by conditioning the Gaussian mixture on observed input values, enabling the characterization of non-Gaussian, asymmetric, and multimodal uncertainty structures. This formulation allows for a flexible and computationally efficient alternative to classical uncertainty propagation methods, particularly in the presence of nonlinear relationships. The approach is demonstrated using fatigue strength estimation of materials, where uncertainties in hardness, residual stress, and surface roughness are propagated through a nonlinear model. Monte Carlo simulation is used as a reference, and the resulting distributions are approximated by Gaussian mixtures. The results show that the proposed method provides a consistent and compact representation of uncertainty while preserving relevant distributional features.

Keywords: Gaussian mixture, Measurement uncertainty, Circular Factory

1 Introduction

1.1 Motivation

In the Circular Factory¹, used products are reintegrated into production and reconditioned for subsequent product generations. Products in such a system are subject to numerous sources of uncertainty, including unknown prior usage conditions, measurement inaccuracies, incomplete material knowledge, and imperfect models. These uncertainties must be rigorously accounted for in decision-making processes. Probability density functions (PDFs) provide a principled mathematical framework for representing uncertainty [1]. In contrast to single-value uncertainty measures, PDFs capture the full stochastic behavior of a quantity of interest. However, their practical application requires representations that are both expressive and computationally efficient. Gaussian mixtures (GMs) constitute such a representation. As a parametric family characterized by weights, means, and variances, they enable compact descriptions of complex datasets. GMs can approximate arbitrary distributions, including multimodal and asymmetric ones, while maintaining computational tractability [2]. Moreover, many key oper-

ations such as marginalization, conditioning, and Bayesian fusion can be realized via closed-form or efficiently computable solutions in the Gaussian mixture domain. This makes them particularly suitable for nonlinear estimation and sensor fusion tasks [1]. A fundamental component of uncertainty quantification is the measurement model, which describes the relationship between an observed value and the underlying true quantity. Such models may be defined analytically (e.g., density derived from mass and volume), specified in datasheets, or obtained empirically from measurement data.

1.2 Related Work

The current state of the art in measurement uncertainty evaluation is the *Guide to the Expression of Uncertainty in Measurement* (GUM) [3], published by the International Organization for Standardization. The classical GUM framework relies on linearization of the measurement model and assumes approximately Gaussian output distributions. These assumptions are violated for strongly nonlinear models or asymmetric distributions.

To overcome these limitations, GUM Supplement 1 introduces a Monte Carlo method for propagating full probability distributions numerically. This approach avoids linearization and Gaussian assumptions, and instead approximates the output distribution through repeated sampling. Harris and Cox

¹The vision of the Collaborative Research Center 1574 "Circular Factory for the Perpetual Product" is to enable integrated linear and circular production on an industrial scale. (www.sfb1574.kit.edu)

[4] describe this method as a numerical implementation of the "propagation of distributions" and discuss practical aspects such as convergence and computational cost. Comparative studies, for example [5] [6], show that GUM and Monte Carlo methods yield consistent results for linear or near-linear models, while Monte Carlo methods become essential for significantly nonlinear systems.

However, while the Monte Carlo approach provides accurate distributional propagation, it typically yields only numerical representations of the resulting distributions. This can limit interpretability and computational efficiency in downstream processing, motivating the need for compact parametric representations of uncertainty.

In this paper, we propose a method to determine the posterior uncertainty of a true quantity given a single observed measurement and an arbitrary measurement model. The resulting uncertainty is represented as a probability density function in the form of a Gaussian mixture. This approach generalizes classical GUM-based uncertainty propagation by accommodating non-Gaussian, nonlinear, and multimodal structures while preserving computational efficiency.

As a practical example, we consider measurement data related to the fatigue strength of materials. This example demonstrates how complex measurement uncertainties can be systematically represented and propagated within the proposed Gaussian mixture framework. We compare the method to the GUM uncertainty propagation method.

2 Method

2.1 Estimating the model

The objective of the proposed method is to represent the entire measurement model as a joint PDF of the output quantity and all relevant input quantities using a multivariate GM. Let $\mathbf{x} = [x_1, x_2, \dots, x_i]^T$ denote the unknown output quantity of interest, the measurand. The input variables, which are measured or observed, are described by $\mathbf{y} = [y_1, y_2, \dots, y_j]^T$. Instead of describing the measurement model as a deterministic function, the stochastic relationship between input and output quantities is modeled through the joint distribution. This joint density is approximated by a Gaussian mixture model using the random variables $\mathbf{y}$, $\mathbf{x}$. This formulation allows modeling complex and nonlinear dependencies between the variables to be represented within a compact parametric framework:

$$p(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{x}, \mathbf{y} \mid \mu_k, \Sigma_k) \quad (1)$$

where the mixture weights satisfy

$$\pi_k \geq 0, \quad \sum_{k=1}^K \pi_k = 1. \quad (2)$$

The mean vector and covariance matrix of the k -th component are given by

$$\mu_k = \begin{bmatrix} \mu_{k,x} \\ \mu_{k,y} \end{bmatrix}, \quad \Sigma_k = \begin{bmatrix} \Sigma_{k,xx} & \Sigma_{k,xy} \\ \Sigma_{k,yx} & \Sigma_{k,yy} \end{bmatrix}. \quad (3)$$

This representation allows us to model arbitrary nonlinear and multimodal dependencies between variables.

To estimate the parameters of the Gaussian mixture, the Expectation-Maximization (EM) algorithm [7], [8] is employed. The EM algorithm is a standard and efficient approach for maximum likelihood estimation of mixture models and provides closed-form updates for the mixture parameters. The algorithm iteratively alternates between two steps. In the expectation step (E-step), the posterior probability that a data sample belongs to each mixture component is computed. In the maximization step (M-step), the mixture parameters, including component weights, means, and covariance matrices, are updated based on these findings. Thus, the parameters of the mixture model are learned directly from the observed data samples by maximizing the likelihood of the dataset.

To prevent overfitting, the number of Gaussian components is not chosen arbitrarily but determined using model selection criteria. In particular, the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC) [8] are evaluated for models with different numbers of components. Both scores penalize the number of components. The model with the minimum information criterion is selected, providing a balance between model accuracy and complexity.

The AIC is defined as

$$\text{AIC} = 2p - 2\ln(\mathcal{L}), \quad (4)$$

where p denotes the number of free parameters of the model and $\mathcal{L}$ the maximized likelihood of the data.

The BIC is given by

$$\text{BIC} = p\ln(N) - 2\ln(\mathcal{L}), \quad (5)$$

where N is the number of data samples. The preferred model is the one that minimizes the respective information criterion. Figure 2 shows an example of a nonlinear measurement curve and its joint GM fit. The number of components was set to five, based on the minimum AIC score shown in Figure 1.

2.2 Calculate the posterior distribution

Once the joint probability density $p(\mathbf{x}, \mathbf{y})$ has been estimated, inference can be performed by conditioning on observed input values. Conditioning corresponds to observing specific measurement values and evaluating the resulting posterior distribution of the output quantity. Conceptually, this can be

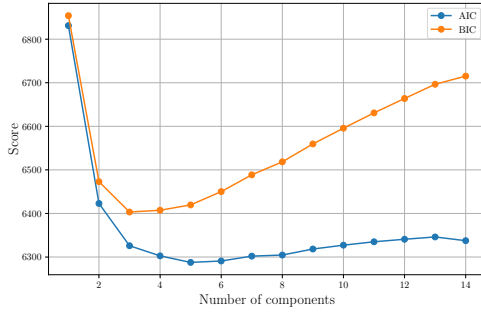


Figure 1: AIC and BIC scores of the GM for different numbers of components

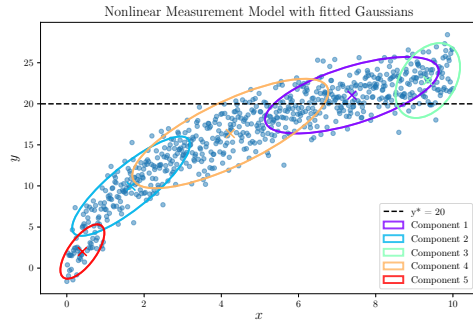


Figure 2: Measurement model with an observed value $y^* = 20$

interpreted as taking a slice through the joint Gaussian mixture along the observed dimensions. Assume that the input variables are observed as $\mathbf{y}^* = [y_1^*, y_2^*, \dots, y_i^*]^T$. The posterior distribution of the output quantity x is then obtained using Bayes' rule

$$p(\mathbf{x} | \mathbf{y}^*) = \frac{p(\mathbf{x}, \mathbf{y}^*)}{p(\mathbf{y}^*)}. \quad (6)$$

The denominator represents the marginal density of the joint distribution evaluated at the observed input values. It is obtained by integrating the joint density over the output variable

$$p(\mathbf{y}) = \int p(\mathbf{x}, \mathbf{y}) d\mathbf{x}. \quad (7)$$

Due to the properties of Gaussian mixtures, this conditioning step can be performed analytically for each mixture component. As a result, the posterior distribution $p(\mathbf{x} | \mathbf{y}^*)$ is again represented as a Gaussian mixture of lower dimension. This enables efficient inference of the output uncertainty for arbitrary observations of the input variables.

The conditional distribution of the output variable for component k is therefore given by (see [2] [1]):

$$\mu_{k,x|y} = \mu_{k,x} + \Sigma_{k,xy} \Sigma_{k,yy}^{-1} (y^* - \mu_{k,y}), \quad (8)$$

$$\Sigma_{k,x|y} = \Sigma_{k,xx} - \Sigma_{k,xy} \Sigma_{k,yy}^{-1} \Sigma_{k,yx}. \quad (9)$$

The mixture weights are updated according to the likelihood of the observation under each component

$$L_k = \mathcal{N}(y^* | \mu_{k,y}, \Sigma_{k,yy}), \quad (10)$$

leading to the normalized conditional weights

$$\tilde{\pi}_k = \frac{\pi_k L_k}{\sum_{j=1}^K \pi_j L_j}. \quad (11)$$

The posterior distribution $p(\mathbf{x} | \mathbf{y}^*)$ is therefore again a Gaussian mixture with updated parameters $\{\tilde{\pi}_k, \mu_{k,x|y}, \Sigma_{k,x|y}\}_{k=1}^K$.

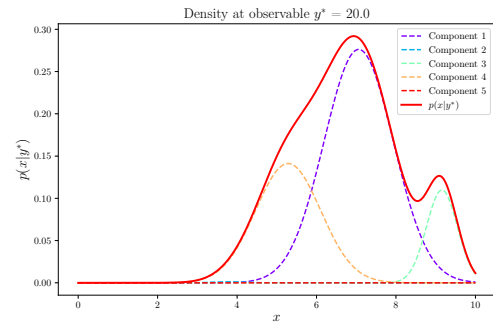


Figure 3: Conditional posterior $p(x | y^*)$ GM for an observed value $y^* = 20$

Figure 3 shows the posterior PDF after an observation for the given exemplary measurement model.

3 Application

3.1 Generation of fatigue strength samples

Determining the local fatigue strength $R_w(x)$ of materials in used components is an important quality indicator for assessing the product's remaining service life and plays an important role in the context of the circular factory in determining whether a material can be reused. Fatigue strength is a quantity that depends on the depth r into the material. However, fatigue strength cannot be measured directly, instead, other parameters must be measured to infer this quantity. Those measurable parameters are surface roughness R_z , hardness HV , and residual stress σ_{ES} .

In the experiment, each of these quantities was measured repeatedly, yielding a mean value μ and a standard deviation σ that describe the measurement uncertainty. Table 1 summarizes the obtained statistics. In this experiment, all quantities are considered at a depth of $r = 0$, corresponding to the material surface. The fatigue strength was calculated using nonlinear analytical relations that describe the influence of the considered surface and material parameters. The formulas to calculate fatigue strength stem from Winderlich [9]:

Table 1: Measured input variables (depth $r=0$).

Variable	Mean	Standard deviation
R_z	$6.66 \mu\text{m}$	$0.89 \mu\text{m}$
HV	677.5 HV	11.4 HV
σ_{ES}	348.4 MPa	4.15 MPa

The static tensile strength is proportional to the hardness:

$$R_m(r) = 3.3 \cdot HV(r) \quad (12)$$

Subsequently, the surface influence factor is:

$$\omega = 1 - 0.22 \cdot \log(R_z) \left(\frac{\log(R_m(r=0))}{20} - 1 \right). \quad (13)$$

The following quantity named Local stress-free alternating load resistance is computed:

$$R_w^0(r) = \frac{1.98HV(r) - 0.0011 \cdot HV(r)^2}{1 + 20.7/HV(r)} \quad (14)$$

Finally, the fatigue strength is:

$$R_w(r) = \omega \cdot R_w^0 \cdot \log(R_z) \cdot \left(1 - \frac{\sigma_{ES}(r)}{R_m(r)} \right). \quad (15)$$

Because the model equations are nonlinear, analytical uncertainty propagation is not straightforward. Therefore, a Monte Carlo simulation was employed to propagate the input uncertainties through the fatigue strength model. Each input quantity was modeled as a normal distribution $\mathcal{N}(\mu, \sigma)$ defined by its mean and standard deviation, from which samples were drawn for every simulation run. The sampled inputs were then inserted into the fatigue strength equations, resulting in a large set of simulated fatigue strength realizations. This procedure yields a numerical approximation of the joint distribution of the input variables and the resulting fatigue strength. In this case, the output quantity depends on three input variables. Therefore the measurement model is four-dimensional.

3.2 Estimation of the probabilistic measurement model

The generated dataset of input-output samples was used to estimate a probabilistic measurement model of the fatigue strength. For this purpose, the joint distribution of the fatigue strength and the input variables was approximated by a four-dimensional GM. The parameters of the mixture model were estimated using the EM algorithm. To determine the optimal number of components, the AIC and BIC were calculated for different numbers of Gaussian components. The trend of the scores over the number of components is shown in Figure 5. The optimal number of components according to BIC is 30. For this case, the joint PDF is demonstrated in Figure 4. To visualize the four-dimensional distribution

effectively, the marginal distribution was plotted for each input variable.

The resulting Gaussian mixture provides a compact parametric representation of the joint probability density of the fatigue strength and the influencing input variables. In contrast to deterministic models, this probabilistic representation captures nonlinear relationships as well as possible multimodal characteristics of the distribution.

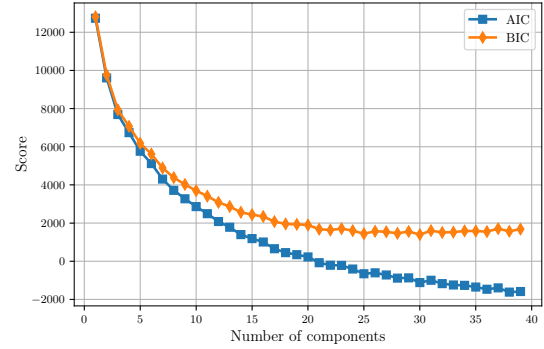


Figure 5: AIC and BIC score for different numbers of components of the 4-dimensional GM

3.3 Posterior inference of fatigue strength

Once the probabilistic measurement model has been estimated, it can be used to infer the probability distribution of the fatigue strength for observed input parameters. Observations of surface roughness, hardness, and residual stress correspond to conditioning the joint Gaussian mixture model on the measured input variables.

By conditioning the joint model on the observed values of the input variables, the posterior distribution of the fatigue strength is obtained. This posterior distribution represents the uncertainty of the fatigue strength given the measured material parameters. Figure 6 shows an example of the posterior probability density of the fatigue strength for a specific set of observed input variables, meaning $HV^* = 670 \text{ MPa}$, $\sigma_{ES}^* = 345 \text{ MPa}$ and $R_z^* = 7 \mu\text{m}$. The resulting distribution reflects both the uncertainty of the measurements and the nonlinear propagation of uncertainty through the fatigue strength model. The results demonstrate that the proposed probabilistic framework allows the fatigue strength distribution to be efficiently inferred for arbitrary combinations of observed material parameters. This provides a flexible measurement model that can directly quantify the uncertainty of the fatigue strength based on experimentally observed input quantities. [9].

4 Evaluation

In this example, a direct comparison between the proposed method and the GUM framework is not

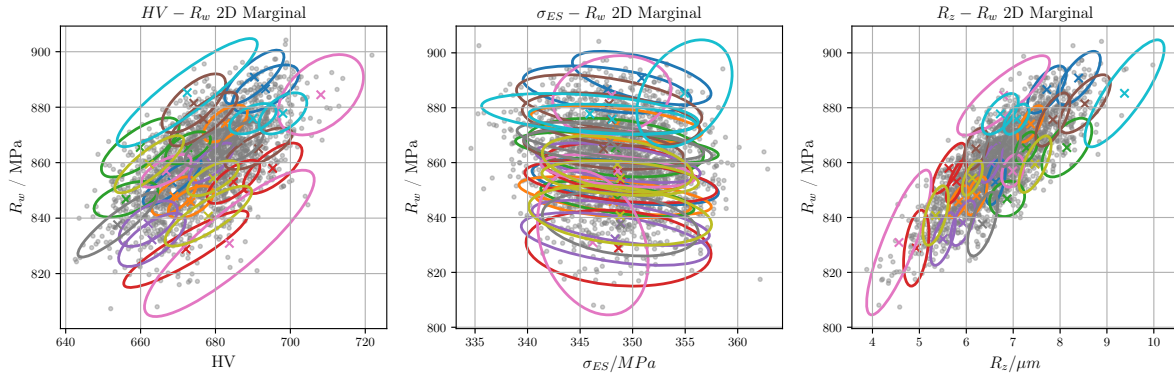


Figure 4: Joint PDF of the Fatigue strength measurement model with 30 GM components

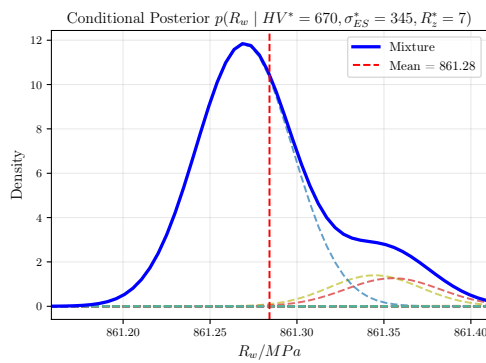


Figure 6: Conditional Posterior with 30 GM components

straightforward, as the two approaches address fundamentally different problem formulations. The classical GUM approach focuses on forward uncertainty propagation, where uncertainties in input quantities are propagated through a measurement model to obtain the uncertainty of an output quantity. In contrast, the method proposed in this work addresses an inverse problem. Starting from an observed measurement, it aims to determine the posterior uncertainty of the underlying true quantity. To this end, both the measurement values and their associated uncertainties are incorporated directly into the probabilistic model, resulting in a multi-dimensional Gaussian mixture representation. From this model, conditional posterior distributions can be derived for given observations.

Applying the GUM framework in this context would require predefined input quantities along with their associated uncertainties. In this example, the standard deviations reported in Table 1 can be interpreted as measurement uncertainties and used within a GUM-based propagation. However, such an approach does not account for the original measurement in a probabilistic inference sense and therefore does not yield a posterior distribution of the true quantity.

Consequently, while both approaches operate on similar data, they differ fundamentally in their objective: the proposed method performs probabilistic inference of the true quantity given an observation, whereas the GUM framework propagates uncertainty from inputs to outputs.

Figure 7 shows the PDF of the observed values including uncertainty. The Monte Carlo-based kernel density estimate (KDE) is approximately Gaussian in this case. Compared to the posterior distribution obtained with the proposed method (Figure 6), differences in the estimated uncertainty become apparent. In particular, the standard deviation obtained using the GUM-based approach is 14.8, while the Monte Carlo method yields a slightly larger value of 15.3, while the standard deviation of the Conditional Posterior in Figure 6 is 0.04.

These differences can be attributed to the fundamentally different treatment of uncertainty. While the GUM and Monte Carlo approaches describe the uncertainty of the observed quantity via forward propagation, the proposed method estimates the posterior distribution of the underlying true quantity. As a result, the proposed approach can lead to a different uncertainty characterization, particularly in the presence of nonlinearities or non-Gaussian effects.

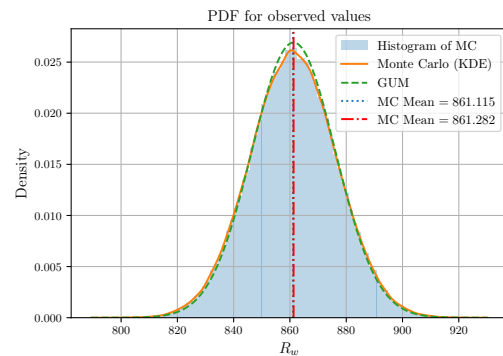


Figure 7: PDF for the GUM and Monte-Carlo case

5 Conclusion

This study presents a robust probabilistic framework for quantifying measurement uncertainty in complex, nonlinear systems, specifically tailored for the challenges of the Circular Factory. The proposed method overcomes significant limitations of conventional approaches by representing the entire measurement model, including the measurand and influencing input quantities, as a joint probability density function approximated by a GM.

In contrast to the standard GUM linearization approach, which relies on local approximations and may become inaccurate for strongly nonlinear or non-Gaussian systems, the proposed method enables a more flexible representation of uncertainty. By leveraging the EM algorithm, the mixture parameters can be estimated directly from data, while model selection criteria provide guidance on the appropriate model complexity. Furthermore, the analytical conditioning of Gaussian mixtures allows the efficient computation of posterior distributions for given observations, offering a computationally attractive alternative to repeated Monte Carlo simulations in scenarios requiring multiple evaluations. The practical application to fatigue strength estimation demonstrates the framework's versatility. By integrating surface roughness, hardness, and residual stress into a unified probabilistic model, we successfully propagated uncertainties through highly nonlinear analytical relations. The results confirm that the method not only yields accurate point estimates but also provides a comprehensive description of the uncertainty distribution, including skewness and multimodality, which are critical for decision-making in reconditioned components. Ultimately, this approach establishes a foundational tool for data-driven decision-making in circular production systems, ensuring that uncertainties are rigorously accounted for when determining the remaining service life of reused components.

Acknowledgement

This research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the collaborative research center (CRC) 1574 "Circular Factory for the Perpetual Product" with the project ID 471687386.

References

- [1] Ali Darijani et al. *Comprehensive Description of Uncertainty in Measurement for Representation and Propagation with Scalable Precision*. 2026. arXiv: 2603.20365 [stat.ML]. URL: <https://arxiv.org/abs/2603.20365>.
- [2] Christopher M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 0387310738.
- [3] Joint Committee for Guides in Metrology (JCGM/WG 1). *Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. 1st edition, corrected version 2010. JCGM 100:2008 (GUM 1995 with minor corrections). Sèvres, France: Bureau International des Poids et Mesures (BIPM), 2008. URL: https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf.
- [4] P M Harris and M G Cox. "On a Monte Carlo method for measurement uncertainty evaluation and its implementation". In: *Metrologia* 51.4 (2014), S176. DOI: 10.1088/0026-1394/51/4/S176. URL: <https://doi.org/10.1088/0026-1394/51/4/S176>.
- [5] Radoslav Marinov. "Comparison of GUM and Monte Carlo Methods for Evaluation the Measurements Uncertainty of DC Voltage, DC Magnetic Flux Density and Dc Resistance". In: *2025 XXXV International Scientific Symposium Metrology and Metrology Assurance (MMA) (2025)*, pp. 1–4. URL: <https://api.semanticscholar.org/CorpusID:284358630>.
- [6] Antonio Piratelli-Filho and Benedito Giacomo. "Uncertainty evaluation in small angle calibration using ISO GUM approach and Monte Carlo method". In: (Jan. 2003).
- [7] Jürgen Beyerer, Raphael Hagmanns, and Daniel Stadler. *Pattern Recognition*. Berlin, Boston: De Gruyter Oldenbourg, 2024. ISBN: 9783111339207. DOI: 10.1515/9783111339207. URL: <https://doi.org/10.1515/9783111339207>.
- [8] Kevin P. Murphy. *Probabilistic Machine Learning: An introduction*. MIT Press, 2022. URL: <http://probml.github.io/book1>.
- [9] B. Winderlich. "Das Konzept der lokalen Dauerfestigkeit und seine Anwendung auf martensitische Randschichten, insbesondere Laserhärtungsschichten". In: *Materialwissenschaft und Werkstofftechnik* 21.10 (1990), pp. 378–389. DOI: <https://doi.org/10.1002/mawe.19900211006>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mawe.19900211006>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/mawe.19900211006>.

Towards Hardware Implementation of Homogeneous Multi-Sensor Self-X System through Simulation-Based Virtual Sensor Modeling

Elena Gerken¹ and Andreas König¹

¹ *Cognitive Integrated Sensor Systems, University of Kaiserslautern-Landau, Deutschland,
e.gerken@rptu.de*

Summary:

Modern sensor systems are expected to provide reliable and consistent measurement quality even under non-ideal mechanical, environmental, or operational conditions. In many applications, achieving sustained high measurement quality with a single sensor depends on a high-quality sensing element, precise mechanical integration, and application-specific calibration or compensation, which increases system cost and design complexity.

This work investigates a Self-X sensor system in which several low-cost and imperfect homogeneous sensors are used instead of relying on a single high-quality and precisely mounted sensor. The main objective is to improve and stabilize the estimation of the target measurand through algorithmic fusion of redundant measurement channels. While previous work mainly focused on testing this hypothesis using simulation-based models, the present study represents a first step toward hardware realization. For this purpose, measurements from one real tunnel magnetoresistance (TMR) sensor were acquired at three different positions around a rotating magnet and used to emulate a homogeneous multi-sensor configuration. After alignment and resampling, the resulting channels were combined using dimensionality reduction to generate a virtual sensor representation.

Compared with the best individual sensor, the virtual sensor reduced mean absolute error (MAE) from 2.214° to 0.953° and root mean square error (RMSE) from 2.854° to 1.288° . This corresponds to improvements of approximately 57% and 55%, respectively. The results indicate that multi-sensor fusion can compensate for position-dependent distortions and provide a useful basis for further hardware development of the Self-X sensor concept.

Future work will focus on moving towards a simultaneous multi-channel hardware configuration, including configurations with more than three sensors and extended measurement campaigns under various operating conditions.

Keywords: Self-X System, TMR Sensor Array, Virtual Sensor, Dimensionality Reduction, Multi-Sensor Integration, Real-Measurement-Based Emulation.

Introduction

In modern Industry 4.0 environments, sensor systems are expected to operate autonomously, reliably, and with minimal human intervention [1, 2]. A homogeneous multi-sensor architecture can support these requirements by allowing several identical sensors to measure the same physical variable simultaneously. In contrast to a solution based on a single high-quality and carefully mounted sensor, such architecture can exploit redundancy among several low-cost and imperfect sensing units. However, redundancy alone does not ensure reliable operation, since individual channels may be affected by drift, misalignment, installation tolerances, amplitude scaling, phase shifts, and other sensor-specific distortions [3–5].

As a result, the potential benefits of multiple sensors cannot be fully exploited by conventional aggregation methods alone.

This limitation creates the need for more intelligent methods capable of evaluating sensor behavior, identifying inconsistencies, and adapting the measurement process accordingly. In this context, the Self-X approach provides a suitable framework for next-generation intelligent sensing solutions. Self-X properties, such as self-monitoring, self-calibration, and self-healing, allow sensor systems to detect internal deviations, compensate for partial failures, and maintain functionality under changing operating conditions. Within such architecture, a virtual sensor can provide a global estimate by combining information from multiple physical channels while accounting for their current condition and mutual consistency.

The present study is a logical continuation of previous work, where the Self-X-based fusion concept for homogeneous sensor systems was investigated mainly using simulation-based models [6]. In contrast, this work introduces real

measurement data as a first hardware-oriented step toward the realization of such a system. TMR sensors are used as a representative use case for angle estimation. The measured TMR signal is sensitive to the sensor placement relative to the magnetic source [7]. Even small deviations between the sensor and the magnet can introduce amplitude and phase distortions, thereby reducing angle estimation accuracy. Therefore, this use case is suitable for studying the compensation of position-dependent channel imperfections. However, the general Self-X concept is not restricted to this sensor type. Instead of implementing a complete multi-sensor hardware prototype with several concurrently operating sensors at this stage, a simplified proof-of-principle setup is used. A single real TMR sensor is measured at several different spatial positions around a rotating magnet, and the resulting signals are treated as individual channels of an emulated homogeneous multi-sensor configuration. The choice of three positions serves the current experimental design and does not represent a fundamental limitation of the concept. The individual channels are then combined into a virtual sensor using dimensionality-reduction methods, with Principal Component Analysis (PCA) used as an illustrative example. The combined latent representation is subsequently used to estimate the angle and is quantitatively compared with the individual sensor channels.

The main contributions of this work are threefold. First, a real-measurement-based proof-of-principle emulation framework is introduced as a first hardware-oriented step toward a homogeneous multi-sensor Self-X system. Second, dimensionality-reduction-based virtual sensor modeling is applied to real TMR signals obtained from multiple spatial positions. Third, the achieved improvement is quantified using standard regression error metrics applied to the angle estimation task and discussed with respect to the further development of a complete multi-sensor hardware implementation.

Concept of the Homogeneous Multi-Sensor Self-X System

The proposed concept is represented by a multi-layer architecture, as shown in Figure 1. This architecture represents the intended complete structure of a homogeneous multi-sensor Self-X system, which is planned to be realized in a full hardware implementation with several concurrently operating sensor channels. In this complete realization, sensor readings are gradually transformed from raw data into synchronized multi-sensor measurements, then into a fused virtual sensor representation, and finally into estimated values of the target parameter. Self-X

capabilities are then used to evaluate system behaviour and introduce correction or adaptation when deviations are detected.

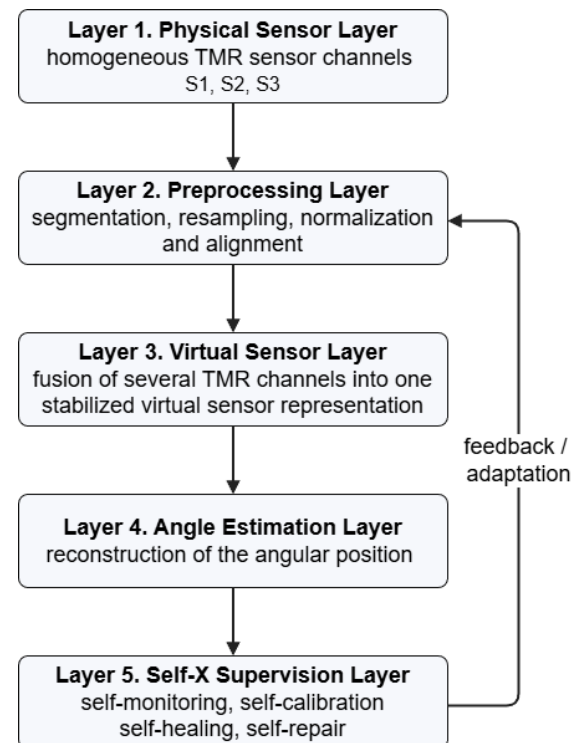


Fig. 1. General layer-based representation of the proposed homogeneous multi-sensor Self-X architecture.

The physical sensor layer represents the data collection stage of the architecture. In a complete hardware realization, this layer consists of homogeneous sensor channels that provide individual measurements of the same quantity or state of the physical system. Their purpose is to provide measured data for further processing.

The preprocessing layer converts the raw channel outputs into a form suitable for data fusion. These operations are required to ensure that the data compared across channels represent equivalent physical states or comparable measurement conditions. Depending on the measured process, these operations may include segmentation, resampling, normalization, filtering, synchronization, and alignment with a reference. The virtual sensor layer acts as the central transformation stage of the architecture. Its purpose is to generate a representative signal from the prepared multi-channel input. Instead of selecting one physical channel, this layer combines the available measurements into a system-level representation. In the present study, this layer is the main object of investigation, since it determines how several related sensor channels can be transformed into a more stable signal basis.

The estimation layer receives the virtual sensor output and converts it into the required target quantity q .

In the TMR-based use case considered in this work, this quantity corresponds to the estimated angular position θ .

The Self-X supervision layer forms the decision-making part of the architecture. It evaluates the behaviour of the preceding layers and initiates feedback, correction, or adaptation when deviations are detected. In a later hardware realization, this layer may adjust preprocessing parameters, update fusion settings, identify unstable channels, or trigger recalibration procedures. Current work focuses on a first hardware-oriented proof-of-principle step. At this stage, data from one real TMR sensor are acquired, preprocessed, and used to emulate a homogeneous multi-sensor configuration. The main focus is therefore placed on the signal-processing part of the architecture, namely virtual sensor formation from real measured data.

Experimental Setup with Multi-Sensor Emulation

The data used in this study were obtained using a previously developed setup for measuring TMR sensor outputs [4]. The experimental setup is shown in Figure 2 and consists of a rotating permanent magnet driven by an electric motor and a single TMR sensor placed near the magnet.

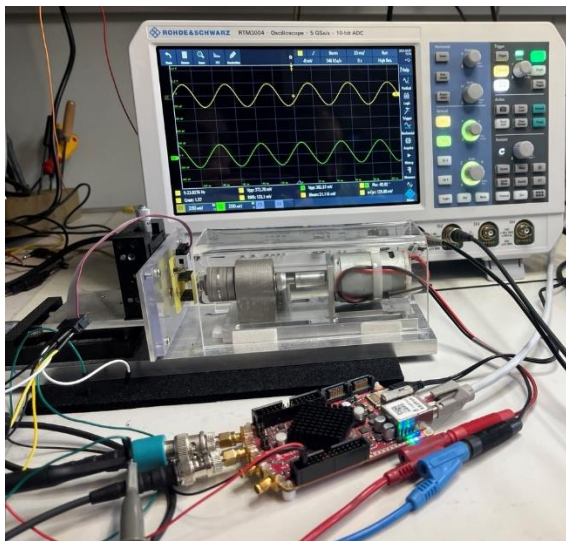


Fig. 2. Experimental setup.

During operation, the sensor records the variation of the magnetic field caused by the rotating magnet. This produces a periodic signal whose structure depends on the relative geometry between the sensor and the magnetic source. For the experiments, the motor operated at approximately 1200 rpm. Since the rotational speed was not perfectly constant, the number of recorded samples per revolution varied from 4140 to 4236, with a mean value of about 4185 samples. To ensure comparability between revolutions and

sensor positions, each revolution was segmented individually and resampled to a fixed length of 5000 samples. This preprocessing step was used only to obtain a consistent temporal representation for subsequent comparison and fusion.

To emulate a homogeneous three-channel sensor configuration, a single TMR sensor was sequentially placed at three different positions relative to the rotating magnet. Additional placement-related imperfections were introduced by tilting the sensor board during positioning, as shown in Figure 3.

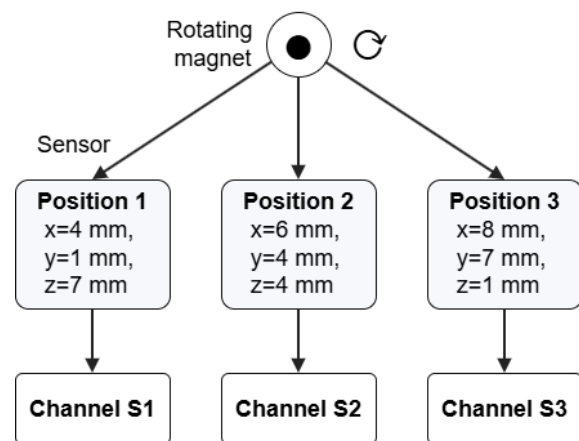


Fig. 3. Real-measurement-based emulation.

At each position, a separate measurement series was recorded and treated as one channel of an emulated homogeneous multi-sensor configuration.

The number of three channels was chosen for experimental feasibility only and is not a fundamental limitation of the proposed approach. The same fusion concept can be extended to a larger number of homogeneous sensor channels.

The present emulation has several limitations. The three channels were acquired sequentially rather than simultaneously; therefore, synchronization effects and possible sensor-to-sensor interactions were not represented. In addition, only one physical sensing element was used, so manufacturing variability between different sensors was not captured. The setup was also limited to one rotational regime and one magnetic configuration. Nevertheless, for a first hardware-oriented proof-of-principle step, the proposed emulation provides a feasible basis for assessing the benefit of virtual sensor representation using real measured data.

Virtual Sensor Modelling Methodology

Figure 4 presents the processing sequence used to derive the virtual sensor. The pipeline starts with synchronized sine-like and cosine-like outputs from the three TMR sensor positions.

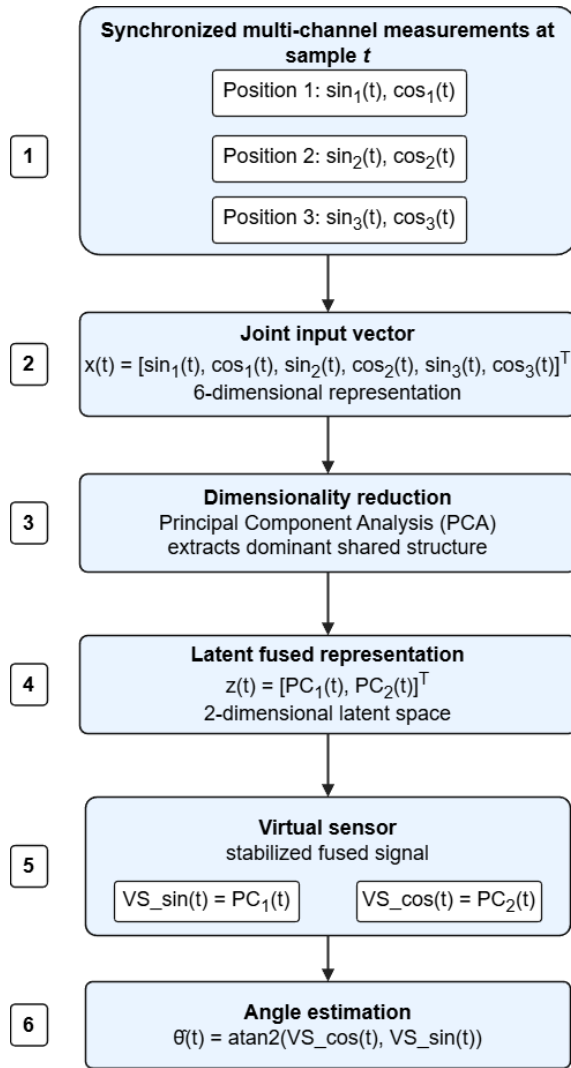


Fig. 4. Dimensionality-reduction pipeline for virtual sensor modelling.

After preprocessing, the outputs from the three sensor positions are organized into a common multi-channel representation. At each sample point, the input vector contains the sine-like and cosine-like components from all available emulated channels. Thus, the same rotational state is described by several spatially different observations rather than by one sensor position alone. The individual sensor-position signals are then processed jointly in the dimensionality-reduction stage. PCA is used in this study as a representative dimensionality-reduction method. It transforms the synchronized multi-channel data into a lower-dimensional latent representation that emphasizes the common rotational pattern shared across the channels and reduces the influence of position-specific signal distortions.

The obtained latent representation is interpreted as the virtual sensor output. Instead of selecting one physical channel or manually correcting amplitude deviations, local distortions, or residual

phase differences in each channel separately, the fusion stage extracts a system-level representation from the combined behaviour of multiple channels. In this sense, the virtual sensor provides a stabilized signal basis derived from redundant measurements.

Based on the virtual sensor representation, angle estimation is performed and compared with the estimates derived from the individual measurement channels.

Experimental Results

Figure 5 shows the comparison between the three individual TMR sensor positions and the fused virtual sensor representation obtained by dimensionality reduction. The individual channels exhibit visible differences in amplitude, phase, and signal shape, which are caused by the different spatial positions of the sensor relative to the rotating magnet. These placement-dependent differences also affect the reconstructed angle and lead to channel-specific error patterns.

The virtual sensor does not simply reproduce one of the individual channels. Instead, it represents a fused latent signal derived from the common structure of the three measured channels. Although the fused representation differs from the individual sensor outputs in its signal form and scale, it provides a more consistent basis for angle reconstruction.

The quantitative comparison is presented in Figure 6. The evaluation was performed over 100 consecutive rotations using standard regression error metrics computed between the estimated and reference angles. Sensor position 1 achieved MAE of 6.652° and RMSE of 8.689° , Sensor position 2 achieved MAE of 2.327° and RMSE of 2.933° , and Sensor position 3 achieved MAE of 2.214° and RMSE of 2.854° . In comparison, the virtual sensor achieved MAE of 0.953° and RMSE of 1.288° .

Compared with the best-performing individual channel, the virtual sensor reduced MAE by approximately 57% and RMSE by approximately 55%. These results show that dimensionality-reduction-based fusion provides a more accurate signal basis for angle estimation than any individual TMR sensor position in the considered proof-of-principle setup. The lower RMSE also indicates that the virtual sensor representation is less affected by larger deviations than the individual channels. Therefore, the results support the use of redundant imperfect measurement channels for improving the quality of the estimated target quantity within the proposed Self-X sensor concept. At the same time, the results should be interpreted in the context of the current real-measurement-based emulation, where the

channels were acquired sequentially rather than by a complete multi-sensor hardware prototype.

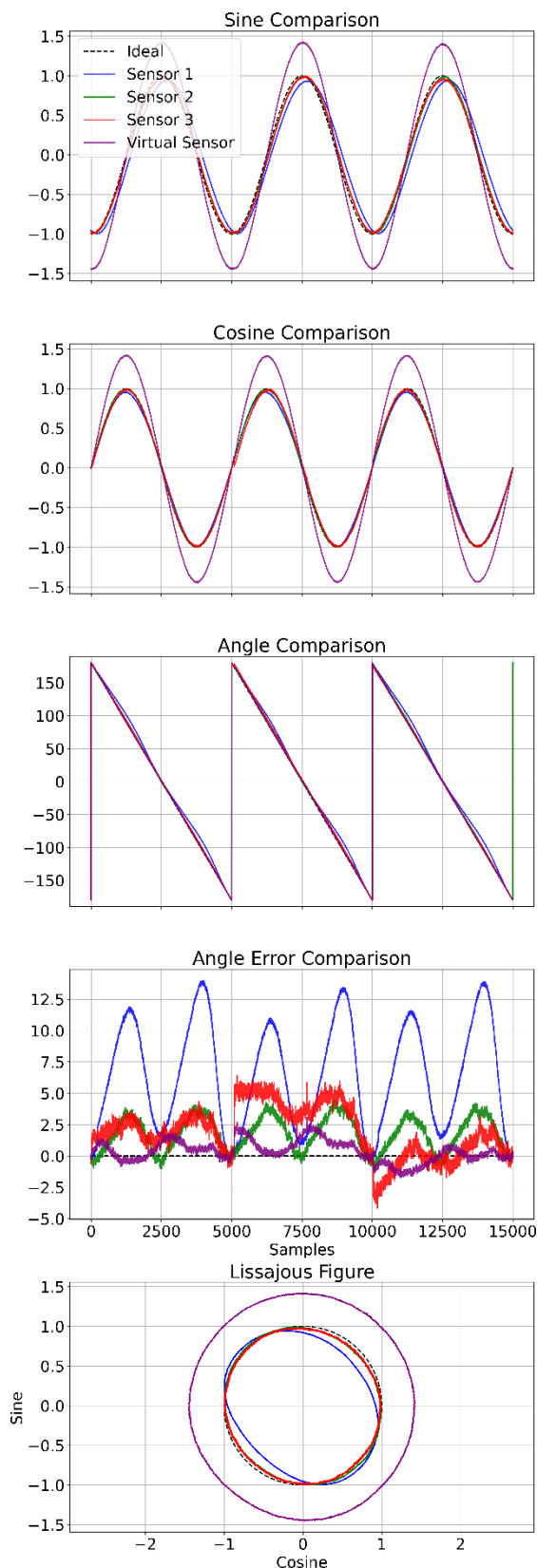


Fig. 5. Measured signals from three individual TMR sensor positions compared with the fused virtual sensor obtained via dimensionality reduction.

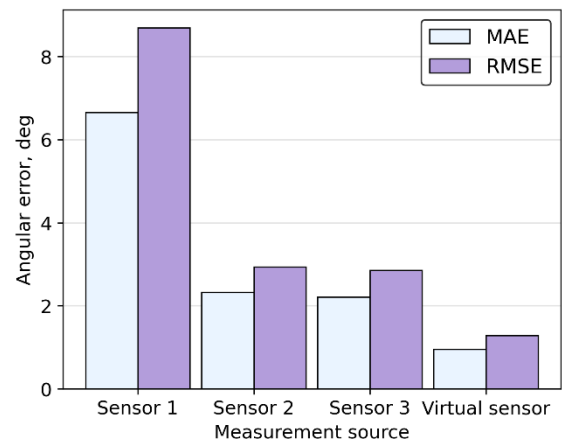


Fig. 6. Error comparison for three individual TMR sensor positions and the virtual sensor.

Conclusion

This study brings the previously developed Self-X sensor system [6] closer to practical hardware realization by introducing a real-measurement-based emulation of a homogeneous multi-sensor configuration. Instead of relying on one ideal sensor, the approach combines redundant information from imperfect homogeneous sensing channels and transforms it into a virtual sensor representation.

For the considered TMR-based angle estimation task, the virtual sensor provided a more accurate reconstruction than any individual measured channel. Compared with the best single channel, MAE decreased from 2.214° to 0.953° , and RMSE from 2.854° to 1.288° over 100 consecutive rotations, corresponding to improvements of about 57% and 55%, respectively. These results support the use of multi-channel fusion as an intermediate step toward a hardware-based Self-X sensor system.

The concept is not limited to TMR sensors or to three sensor channels. Future work should move from the present sequential emulation to a simultaneous multi-channel hardware setup with more than three physical sensor channels, different spatial arrangements, and longer measurement campaigns. Further experiments should also include varying operating conditions and more severe sensor faults, such as scale and offset errors, nonlinear distortions, clipping, drift, aging effects, and gradual loss of sensor quality.

References

- [1] Khan M.I., Yasmeen T., Khan M., Hadi N.U., Asif M., Farooq M., Al-Ghamdi S.G. Integrating industry 4.0 for enhanced sustainability: Pathways and prospects. *Sustainable Production and Consumption*. 2025; 54:149-189. ISSN 2352-5509. <https://doi.org/10.1016/j.spc.2024.12.012>.
- [2] Gosavi, A., Gosavi, A. A simulation-based digital twin for data-driven maintenance scheduling of risk-prone production lines via actor critics. *Flex Serv*

- Manuf J (2024). <https://doi.org/10.1007/s10696-024-09579-1>.
- [3] Nemec D, Andel J, Simak V, Hrbcek J. Homogeneous Sensor Fusion Optimization for Low-Cost Inertial Sensors. *Sensors*. 2023; 23(14):6431. <https://doi.org/10.3390/s23146431>.
 - [4] Kajmakovic A, Diwold K, Römer K, Pestana J, Kajtazovic N. Degradation Detection in a Redundant Sensor Architecture. *Sensors*. 2022; 22(12):4649. <https://doi.org/10.3390/s22124649>
 - [5] Giordano, M.R., Malings, C., Pandis, S.N., Presto, A.A., McNeill, V.F., Westervelt, D.M., Beekmann, M., & Subramanian, R. (2021). From low-cost sensors to high-quality data: A summary of challenges and best practices for effectively calibrating low-cost particulate matter mass sensors. *Journal of Aerosol Science*, 158, 105833.
 - [6] Gerken E, König A. Enhancing Reliability in Redundant Homogeneous Sensor Arrays with Self-X and Multidimensional Mapping. *Sensors*. 2025; 25(13):3841. <https://doi.org/10.3390/s25133841>.
 - [7] TDK Corporation. TMR angle sensor Guide. Technical Resources: Advice on optimal placement according to the magnetic field. 2025. <https://product.tdk.com/en/products/sensor/angle/tmr-angle/technote/menu.html>.

Kalibrierungsfreier Stromsensor für Hochspannungsfreileitungen

P. Arnold¹, C. Brockmann¹, H. Pötter¹, Prof. Dr.-Ing. U. Ganesh¹

¹Fraunhofer Institut für Zuverlässigkeit und Mikrointegration, Berlin (Deutschland)

Zusammenfassung

Bestehende nicht-invasive und indirekte Messsysteme zur Erfassung der Stromstärke von Hochspannungsfreileitungen, wie Hallsensoren oder Spulen mit magnetisch bündelndem Kern müssen für jeden Leiterseildurchmesser kalibriert werden. Das in dieser Arbeit beschriebene Messprinzip erübrigt diese Kalibrierung durch den Einsatz einer leiterplattenintegrierten Rogowskispule. Damit kann der Stromsensor unter bestimmten Randbedingungen auf beliebigen Leiterseildurchmessern eingesetzt werden. Die Messsignale werden durch ein analoges Frontend vorverarbeitet und im Anschluss digitalisiert, um die Messwerte weiterverarbeiten zu können. Für den Einsatz des Messverfahrens in energieautarken drahtlosen Sensorsystemen ist die gesamte Schaltung hinsichtlich des Energiebedarfs optimiert. Das angestrebte Anwendungsszenario für den Stromsensor umfasst Hochspannungsfreileitungen, welche mit einer Netzfrequenz von 50Hz arbeiten. Außerdem wird die Spannungsebene bis 420kV mit einer Leiterseilstromstärke bis 1500A betrachtet.

Keywords: Stromsensorik, kalibrierungsfrei, nicht-invasiv, indirekt, Hochspannungstechnik

Einleitung

Die Grundlage für die Messung der Stromstärke durch einen Leiter wurde 1912 in einem Artikel von W. Rogowski und W. Steinhaus gelegt [1]. Hier wurden mit einer handgewickelten Spule ohne permeablen Kern erste Versuche gemacht, die magnetische Flussdichte einer stromdurchflossenen Spule zu messen und auf die erregende Stromstärke Rückschlüsse zu ziehen.

Um die Präzision und Genauigkeit zu steigern wurde in [2] die Rogowskispule in eine Leiterplatte integriert. Ziel hier war neben der hohen Genauigkeit auch der Einsatz in der Hochspannungstechnik als Messgerät und Schutzeinrichtung.

Eine weitere Arbeit zu leiterplattenintegrierten Rogowskispulen beschäftigte sich mit der Behandlung von kapazitiv einkoppelnden elektrischen Feldern [3]. Hier wurden zwei Luftspulen mit entgegengesetztem Wicklungssinn ineinander gewickelt, um die kapazitiven Einkopplungen durch Addition der Ausgangssignale der Luftspulen auszulöschen. In [4] wird eine Möglichkeit beschrieben, die Rogowskispule zu öffnen und zu schließen, um sie an einem Leiter zu befestigen, ohne diesen auftrennen zu müssen.

Neben diesen Arbeiten existieren viele weitere, welche sich in den Grundzügen ähneln. Allen gemein ist, dass sie nicht alle erarbeiteten

Konzepte vereinen, sondern sich meist auf einen Aspekt der Optimierung konzentrieren.

In der vorliegenden Arbeit werden die Konzepte der genannten Referenzen wie die Leiterplattenintegration und die Immunität gegenüber kapazitiv einkoppelnder elektrischer Felder übernommen. Die Erweiterung gegenüber den Arbeiten liegt in der dauerhaften Öffnung der Spule. In der Vergangenheit wurden Rogowskispulen immer wieder geschlossen, um die Vorteile des Messprinzips beizubehalten. In der vorliegenden Arbeit wird diese Öffnung untersucht und mittels definierten Randbedingungen validiert, um einen kalibrierungsfreien Stromsensor zu erhalten.

Grundlagen

Analog zu anderen nicht-invasiven und indirekten Messverfahren, welche auf dem Prinzip der Induktion beruhen, stützt sich die Erfassung der Stromstärke durch einen Leiter mittels Rogowskispule auch auf das vom Stromfluss durch den Leiter hervorgerufene Magnetfeld. Dieses Feld koppelt in die Spule ein und induziert eine Spannung in den Wicklungen. Diese Spannung ist proportional zur Stromstärke, welche durch den Leiter fließt. Abb. 1 verdeutlicht diesen Zusammenhang.

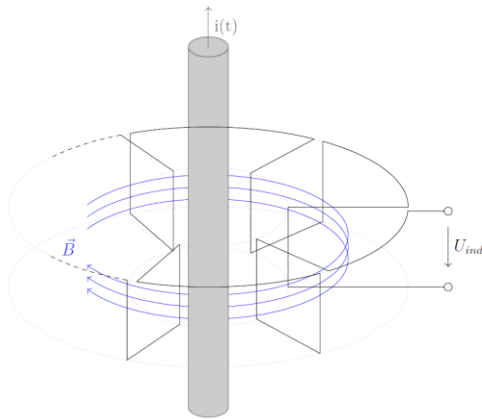


Abb. 1: Anordnung der Wicklungen der Rogowskispule um einen stromdurchflossenen Leiter. Der Stromfluss erzeugt das Magnetfeld und die resultierende magnetische Flussdichte, welche eine Spannung über der Spule induziert.

Die Wicklungen der Rogowskispule müssen so angeordnet sein, dass die durch die Windungen aufgespannten Flächen vom Magnetfeld (respektive der magnetischen Flussdichte) senkrecht durchdrungen werden. Die Flächennormale muss demnach tangential zum Verlauf des Magnetfeldes liegen. Nach dem Durchflutungsgesetz nach Ampère ergibt sich für ein geschlossenes Linienintegral um einen stromführenden Leiter folgender Zusammenhang Gl. (1).

$$\oint_{\partial A} \vec{H} \cdot d\vec{s} = I_{\text{eing.}}$$

Das Ergebnis dieser Integration entspricht der Stromdichte (respektive der eingeschlossenen Stromstärke) durch jene Fläche, über deren Rand integriert wird. Somit ergibt sich eine Unabhängigkeit der Verteilung der Stromdichte über der Fläche. Solange der Weg, über den integriert wird, geschlossen ist, kann die Verteilung der Stromdichte willkürlich sein. Diese Verteilung kann aus mehreren stromführenden Leitern bestehen, wobei Form und Dimension der Leiter beliebig sein können. Hierin liegt der wesentliche Vorteil des kalibrierungsfreien Stromsensors.

Bei der Rogowskispule handelt es sich um eine Luftspule, denn sie besitzt keinen permeablen Kern. Dies hat den Vorteil, dass es zu keiner Sättigung des Kerns kommen kann. Somit gibt es auch keine Hystereseeffekte und keine Remanenzflussdichte. Die Spule kann dadurch auch als kurzschlussfest bezeichnet werden. Es können beliebig große Überströme fließen, welche deutlich außerhalb des Messbereichs liegen, da beliebig hohe Magnetfelder durch die Luftspule fließen können. Somit ist die

messbare Stromstärke nicht begrenzt, insofern es die auswertende Elektronik verarbeiten kann. Der fehlende magnetisch bündelnde Kern bedingt aber auch einen großen Nachteil von Rogowskispulen. Da die Permeabilitätszahl von Luft nur geringfügig größer als die magnetische Feldkonstante μ_0 ist, hat die magnetische Kennlinie eine Steigung von eins. Demnach wird das Magnetfeld in der Spule nicht gebündelt und die induzierte Spannung in der Rogowskispule fällt deutlich geringer aus als bei Spulen mit permeablem Kern. Dies kann durch ein geeignetes analoges Frontend und eine hochauflösende Digitalwandlung kompensiert werden.

Aufgrund der geringen Empfindlichkeit der Luftspule ist eine Hohe Wicklungsanzahl anzustreben, um die induzierte Spannung so groß wie möglich zu gestalten. Für eine Wicklung einer Spule gilt nach dem Induktionsgesetz nach Faraday Gl. (2).

$$\oint_{\partial A} \vec{E} \cdot d\vec{s} = -\frac{d}{dt} \iint_A \vec{B} \cdot d\vec{A} = -\frac{d\phi}{dt}$$

Die Grundlage für die Induktion ist der sich über die Zeit verändernde Strom di/dt . Dieser ruft ein Magnetfeld H hervor, woraus unter Berücksichtigung von μ_r die magnetische Flussdichte B entsteht. Wenn diese die Wicklungen der Rogowskispule durchdringt, entspricht die Ausgangsspannung der Luftspule der Summe der Induktionsspannungen der einzelnen Wicklungen der Spule. Die Induktionsspannung kann mit folgender Formel berechnet werden Gl. (3).

$$u_{\text{ind}}(t) = -M \frac{di(t)}{dt}$$

Dabei entspricht M der Selbstinduktion respektive der magnetischen Kopplung. Die Kopplung ist ein Maß für die Empfindlichkeit der Rogowskispule.

Eine steigende Wicklungszahl erhöht gleichermaßen die Trägheit der Luftspule und führt dazu, dass sich schnell ändernde Ströme schlecht oder gar nicht erfasst werden können. Im hier angestrebten Anwendungsfall einer Hochspannungsfreileitung mit niedriger und konstanter Netzfrequenz von 50Hz kommt dieser Effekt nicht zum Tragen.

Eine grundlegende Anforderung an die Wicklungen der Rogowskispule ist deren möglichst identische Geometrie. Nur damit ist die homogene Integration über das Magnetfeld gegeben. Dies lässt sich nur unter sehr hohem Aufwand mit drahtgewickelten Spulen realisieren, selbst wenn ein nicht permeabler Kern verwendet wird. Den höchstmöglichen Grad an identischen Wicklungen kann erreicht

werden, indem die Rogowskispule in eine Leiterplatte integriert wird. Aktuelle Leiterplattenfertigungstechnologien unterliegen sehr geringen Toleranzen und bieten die Möglichkeit mit vergleichsweise geringem Bauvolumen eine hohe Anzahl an Wicklungen zu realisieren.

Bis hierhin wurde eine um den stromführenden Leiter geschlossenen Spule betrachtet. Dieser Ansatz ist im Einsatz des Stromsensors an bestehenden Freileitungen unzureichend, da die Leiterseile nicht geöffnet werden können, um den Sensor zu montieren. Selbst bei Neubauten ist eine geschlossene Spule nachteilig, da sie über die gesamte Seillänge verschoben werden müsste. Demnach ist eine Öffnung der Rogowskispule notwendig, um den praktischen Einsatz zu ermöglichen. Die Öffnung zieht nach sich, dass das Linienintegral nicht mehr geschlossen ist und nicht die gesamte Stromdichte durch die Querschnittsfläche der Integration erfasst wird. Um dies zu kompensieren, muss der stromführende Leiter im Zentrum der Luftspule liegen. Außerdem muss die Spule kreisförmig ausgeführt sein, damit die Wicklungen homogen vom Magnetfeld durchdrungen werden. Für die Validierung des Messverfahrens ist eine einmalige Parameteridentifikation des Stromsensors notwendig. Im Nachgang gefertigte identische Rogowskispulen benötigen keine Kalibrierung, da die Geometrien identisch sind.

Auslegung

Im Fall einer einfach gewickelten Rogowskispule können elektrische Felder kapazitiv in die Luftspule einkoppeln. Diese Einkopplung kann eliminiert werden, indem zwei Spulen mit entgegengesetztem Wicklungssinn ineinander integriert werden [1]. Da die Spulen unterschiedlichen Windungssinn aufweisen, erhalten die Induktionsspannungen unterschiedliche Vorzeichen, während die kapazitiven Einkopplungen das gleiche Vorzeichen besitzen. Werden die Induktionsspannungen, welche jeweils durch die Einkopplung beaufschlagt sind voneinander abgezogen, löschen sich die Einkopplungen aus und die Induktionsspannungen addieren sich. Abb. 2 zeigt den Aufbau der Wicklungen zweier geschlossener Rogowskispulen mit entgegengesetztem Windungssinn.

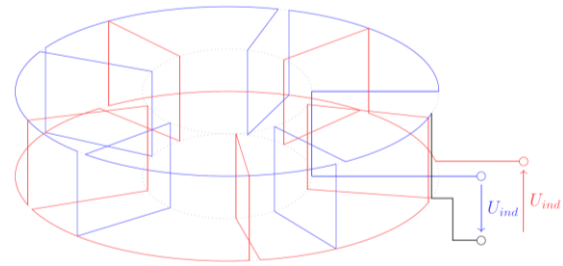


Abb. 2: Zwei ineinander gewickelte geschlossene Rogowskispulen mit entgegengesetztem Wicklungssinn zur Eliminierung von kapazitiven Einkopplungen.

Die beiden Spulen sind an ihren Enden mit der Schaltungsmasse verbunden. An den beiden Anfängen kann die Induktionsspannung gegenüber Masse abgegriffen werden. Da der Gesamtaufbau geöffnet werden muss, ergibt sich der in Abb. 3 dargestellte Aufbau.

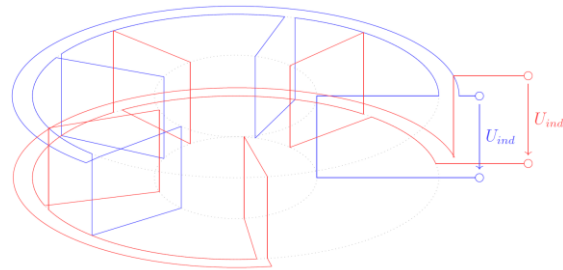


Abb. 3: Zwei ineinander gewickelte Rogowskispulen mit entgegengesetztem Wicklungssinn, Öffnung und Rückführung.

Hier dargestellt sind beide Wicklungen mit Rückführung. Die Rückführung der Spulenenden ist notwendig, um die Induktionsspannung in den Spulen abgreifen zu können. Dabei muss es nur eine Rückführung geben, da beide Spulenenden mit einem Bezugspotential, der Schaltungsmasse, verbunden werden.

Dimensionierung

Der Stromsensor wurde für eine bestehendes energieautarkes und drahtloses Sensorsystem für das Monitoring von Hochspannungsfreileitungen entwickelt. Alle folgenden Einschränkungen geometrischer Natur rühren aus den Restriktionen, welche durch das Gehäuse dieses Sensorsystems gegeben sind. Zur Validierung des Messprinzips wurde eine geöffnete Rogowskispule nach beschriebenem Aufbau hergestellt, welche den Bauraum im Gehäuse maximal nutzt. Dabei wurde eine zweilagige Leiterplatte verwendet. Die Wicklungen bestehen aus je einer Leiterbahn auf der Ober-

und Unterseite und zwei Durchkontaktierungen, welche die Leiterbahn zu einer Wicklung verbinden. Dabei endet eine Leiterbahn kurz vor der Durchkontaktierung und verläuft zur nächsten Windung. Damit der Stromsensor auf einem Leiterseil montiert werden kann, weist die Rogowskispule eine relativ große Öffnung auf. In Abb. 4 ist die gefertigte Leiterplatte dargestellt.



Abb. 4: Leiterplattenintegrierte Rogowskispule mit Öffnungswinkel ca. 117°, 2x174 Wicklungen und $M \approx 68 \text{ nH}$.

Diese enthält zwei Rogowskispulen mit je 174 Wicklungen. Die Gesamtwicklungszahl ergibt sich zu 348. Der Öffnungswinkel beträgt ca. 117°. Die Leiterplattendicke beträgt 2,4mm. Die magnetische Kopplung dieser Anordnung kann mit folgender Gleichung bestimmt werden Gl. (4) (vgl. [1]).

$$M = \frac{\mu_0}{2\pi} \cdot N \cdot h \cdot \ln\left(\frac{a}{d}\right)$$

Hierbei beschreibt N die Wicklungsanzahl, h die Höhe einer Wicklung beziehungsweise der Dicke der Leiterplatte, a den Außendurchmesser der Rogowskispulen und d den Innendurchmesser. Aus diesen Parametern ergibt sich für die magnetische Kopplung dieses Aufbaus $M \approx 68 \text{ nH}$. Zur Berechnung der zu erwartenden Induktionsspannung in den Rogowskispulen wird eine maximale Stromsteilheit beziehungsweise -änderung von $0,5 \text{ A}/\mu\text{s}$ angenommen. Daraus resultiert unter Anwendung von Gl. (3) die maximale Ausgangsspannung der Luftspulen nach Addition der Einzelspannungen von $U_a \approx 34 \text{ mV}$.

Schaltung

Die Ausgangssignale der Rogowskispulen umfassen den Wertebereich von wenigen Mikrovolt bis zu geringen zweistelligen Millivolt. Im hier angestrebten Einsatzszenario handelt es sich um zwei um 180° phasenverschobene sinusförmige Wechsignale mit einer Grundfrequenz von 50 Hz beziehungsweise der momentanen Netzfrequenz, welche nicht exakt bei 50 Hz liegen muss. Weiterhin sind

die Ausgangssignale der Luftspulen nicht belastbar, die Spulen weisen ein sehr hochimpedantes Verhalten auf.

Die Ausgangssignale müssen zunächst addiert, danach verstärkt, optional integriert und zum Schluss abgetastet werden. Abb. 5 zeigt das zugehörige Blockschaltbild, welches die Anordnung der Funktionsblöcke und den Signalpfad der Schaltung darstellt.



Abb. 5: Blockschaltbild der analogen Vorverarbeitung der Ausgangssignale der Rogowskispulen mit Addition, Verstärkung, Integration (optional) und Abtastung.

Da die Rogowskispulen Wechsignale erzeugen und der Wertebereich der digitalen Abtastung voll ausgeschöpft werden soll, muss das Potential, um das die Signale schwingen, angehoben werden. Hierfür werden in der Messschaltung zwei Massepotentiale verwendet. Einerseits eine allgemeine Schaltungsmasse, welchen den Bezug für die Versorgungsspannung darstellt. Andererseits eine virtuelle Masse, welche sich auf dem halben Potential zwischen Versorgungsspannung und Masse befindet. Die virtuelle Masse wird über einen hochohmigen Spannungsteiler generiert. Außerdem wird das Potential der virtuellen Masse vom Spannungsteiler mit einem Operationsverstärker (OPV) entkoppelt, damit durch folgende Komponenten kein belasteter Spannungsteiler entsteht. Diese Belastung würde die virtuelle Masse im Potential verändern. Mit Einführung der virtuellen Masse entstehen ihr gegenüber eine positive und eine negative Versorgungsspannung. Die Positive entspricht der eigentlichen Versorgungsspannung und die Negative der Schaltungsmasse. Die Funktionsblöcke, welche Wechsignale verarbeiten sollen, ohne die negativen Halbwellen abzuschneiden, werden mit diesen neuen Versorgungsspannungen und der virtuellen Masse betrieben. Außerdem wird ein Anschluss der Rogowskispulen mit der virtuellen Masse verbunden. Damit werden die Ausgangssignale der Luftspulen angehoben und schwingen genau um die Mitte der Arbeitsbereiche der verarbeitenden Komponenten. Dies gilt auch für die digitale Abtastung, welche mit der Schaltungsmasse als Versorgung und der normalen Versorgungsspannung betrieben wird.

Der erste Funktionsblock stellt die Addition dar. Hier werden die Signale der Rogowskispulen addiert. Die Addition wird als Subtraktion ausgeführt. Somit addieren sich die zueinander inversen Signale der Rogowskispulen. Dieser Schaltungsteil wird mit OPVs realisiert. Ein einfacher OPV genügt den Anforderungen nicht, da seine Eingänge nicht hochohmig genug sind. Am besten geeignet ist ein Instrumentationsverstärker, welcher schlussendlich aus drei OPVs besteht. Dieser stellt im Bereich der analogen Addition eine der höchsten Eingangsimpedanzen dar. Gleichzeitig bietet er eine hohe Gleichtaktunterdrückung. Diese dient zusätzlich der Subtraktion der kapazitiven Einkopplungen, welche vornehmlich durch die Verschaltung beider Rogowskispulen ausgelöscht werden. Weiterhin sind viele integrierte Schaltkreise (IC) dieser Art verfügbar, die nur geringen Bauraum benötigen und eine einstellbare Verstärkung aufweisen. Die Verstärkung ist notwendig, um die Amplitude des Ausgangssignals des Instrumentationsverstärkers anzuheben, damit die folgenden Komponenten besser in ihrem Arbeitsbereich betrieben werden können. In den meisten Fällen genügt ein einzelner Widerstand als externe Komponente, um den Verstärkungsfaktor einzustellen. Mit der integrierten Verstärkung kann auf einen nachfolgenden Verstärker verzichtet werden, was Energiebedarf, Rauschanfälligkeit und Kosten reduziert. Die Verstärkung darf nicht zu hoch gewählt werden, um das Eingangsrauschen des Verstärkers nicht unnötig mit zu verstärken. Es ist ein Trade-Off zwischen Wertebereichsausnutzung bei der digitalen Abtastung und der Verstärkung des Rauschens. In der hier angestrebten Anwendung wird das Signal nur soweit angehoben, dass die kleinste messbare Signaländerung den Anforderungen entspricht. Das Rauschen in den Ausgangssignalen der Luftspulen kann durch Filterkomponenten reduziert respektive gedämpft werden. Im einfachsten Fall erfolgt dies mit passiven Tiefpass- oder Bandpassfiltern. Die Reduktion des Messrauschens an dieser Stelle im Signalpfad ist nicht Bestandteil dieser Arbeit. Im Ergebnis der Addition und Verstärkung liegt die Summe beider Ausgangssignale der Rogowskispulen vor. Das Signal ist nahezu gleichtaktfrei und enthält keinen Gleichspannungsoffset. Der verbleibende Offset ist bedingt durch die verwendeten OPVs. Dieser Offset lässt sich durch eine Entkopplung mit einem Kondensator in Reihe im Signalpfad beheben.

Nach der Addition und Verstärkung folgt eine optionale Integration. Diese ist nur im angestrebten Anwendungsszenario optional, da die zu erfassenden Signale eine nahezu konstante Frequenz aufweisen. Ist die Frequenz variabel oder sind die Signale mit weiteren Frequenzanteilen beaufschlagt, ist die Integration notwendig, damit kein Informationsverlust in der analogen Vorverarbeitung geschieht. Im hier betrachteten Fall diene die Integration lediglich einer weiteren Verstärkung und der Korrektur der Phasendrehung. Die Phasendrehung passiert in den Rogowskispulen. Sie verschieben die Phasenlage zwischen Messsignal und erregender Stromstärke im ideellen Fall um 90° . Da dieser Phasenversatz hier konstant und empirisch bestimmbar ist, wird auf den Funktionsblock Integrierer verzichtet. Dies reduziert Bauraum, Rauschanfälligkeit und Kosten. Der nicht kompensierte Phasenversatz kann nachträglich in Software korrigiert werden, insofern die Phasenlage und nicht nur die Amplitude respektive der Effektivwert von Interesse sind.

Nach der analogen Vorverarbeitung ist das Messsignal offsetfrei und verstärkt. Um die Messwerte weiterverarbeiten zu können, wird das Ausgangssignal des Analogpfades mit einem Analog-Digital-Umsetzer (ADC) abgetastet. Der hier eingesetzte ADC arbeitet mit sukzessiver Approximation und weist eine Auflösung von 16 Bit auf. Die im angestrebten Szenario gewünschte Auflösung liegt bei einem Ampere. Um etwaiges Messrauschen zu kompensieren darf das Bit mit dem niedrigsten Stellenwert (LSB) maximal einem $1/4$ Ampere entsprechen. Die Amplitude des addierten Ausgangssignals der Spulen liegt bei 1A bei ca. $21\mu\text{V}$. Da der 16-Bit ADC hier mit einer Referenzspannung von 3,3V arbeitet, liegt seine minimale Auflösung bei ca. $50\mu\text{V}$. Somit müssen die $5\mu\text{V}$, welche $1/4$ Ampere entsprechen mit dem Faktor 10 verstärkt werden. Damit ist sichergestellt, dass mindestens zwei LSB verwendet werden, um das kleinste Signal zu erfassen. Durch Änderung der Verstärkung kann eine höhere Auflösung generiert werden oder aber der Messbereich der messbaren Stromstärke kann ohne jegliche Änderung gesteigert werden, wenn es das Anwendungsszenario erfordert. Wird die Verstärkung erhöht, ist zu bedenken, dass damit auch das Rauschen der Ausgangssignale der Rogowskispulen verstärkt wird. Für die Versuche wurde dennoch eine Verstärkung von ca. 100 eingestellt. Dies dient der vollen Ausnutzung des Wertebereichs des ADCs. Außerdem ist

der Rauscheintrag in den Versuchsaufbau minimal, da die Untersuchungen unter Laborbedingungen stattfanden. Der verwendete ADC kann mit bis zu 100kHz abtasten. Nach dem Nyquist-Shannon-Abtasttheorem muss ein Signal mit mehr als der doppelten Frequenz bezüglich der Grundfrequenz des Signals abgetastet werden. Unter der Annahme von 150Hz Abtastfrequenz würde eine Vollwelle der Stromstärke mit drei Messpunkten quantisiert. Da dies sehr gering und schnell zu Fehlabtastungen führen kann, sollte eine höhere Abtastfrequenz gewählt werden. Zu beachten ist hierbei, dass je schneller die Abtastung erfolgt, desto geringer ist die Zeit, in der sich die zu messende Kapazität im ADC aufladen kann. Ein weiterer Aspekt ist der gesamte Energiebedarf von ADC und Ansteuerung, welche bei höheren Taktfrequenzen steigt. Für diese Arbeit wird eine Abtastrate von 5kHz verwendet. Damit wird eine 50Hz Vollwelle mit 100 Messpunkten quantisiert und die Extrema können zuverlässig detektiert werden.

Der ADC wird von einem Mikrocontroller gesteuert. Die Schnittstelle ist ein Serial Peripheral Interface (SPI). Da der ADC keine Einstellmöglichkeiten vorsieht, entfällt das Master-Output-Slave-Input Signal. Das Taktsignal der SPI dient gleichzeitig als Taktgeber für die Abtastfrequenz, mit der der ADC arbeiten soll. Die in den Versuchen verwendete Firmware und die eingesetzten Algorithmen werden in dieser Arbeit nicht näher betrachtet.

Energieversorgung

Die Energieversorgung nutzt eine Versorgungsspannung von 3,3V. Diese werden im Anwendungsszenario vom Sensorsystem durch einen Linearregler zur Verfügung gestellt. Die Versorgungsspannung wird mit einem Ferrit im Stromsensor gefiltert, um hochfrequente Einkopplungen zu unterdrücken. 3,3V werden in der hier verwendeten Bestückungsvariante des Stromsensors benötigt, da die OPVs mindestens $\pm 1,5V$ Versorgungsspannung benötigen, wenn sie im Dual-Supply-Betrieb verwendet werden. Durch Einsatz anderer verfügbarer OPVs kann die Versorgungsspannung auf 2V gesenkt werden. Da zum Zeitpunkt dieser Arbeit keine Komponenten mit ähnlich geringem Drift, geringem Leckstrom und geringem Offset für diese Spannungsebene gefunden werden konnten, wurden die bestmöglichen Bauteile gewählt und hiernach die Versorgung bestimmt.

Die Stromaufnahme für den Stromsensor liegt bei ca. 5mA. Dies liegt einerseits am hohen Verstärkungsfaktor des Instrumentationsverstärkers. Andererseits an der Erzeugung der virtuellen Masse. Der hier verwendete OPV gibt konstant 1,65V aus und muss einen Teil der Schaltung versorgen. Der ADC benötigt maximal 1,3mA.

Der Energiebedarf kann gesenkt werden, in dem energiesparsamere Bauteile verwendet werden. Weiterhin kann der Verstärkungsfaktor reduziert werden. Durch den Verzicht auf die negativen Halbwellen entfällt die Notwendigkeit der virtuellen Masse, was die Stromaufnahme stark senkt. Eine weitere Maßnahme ist die Reduktion der Auflösung und Abtastrate des ADCs. Dies reduziert den Energiebedarf, verringert aber die Genauigkeit.

Versuch

Der entwickelte Stromsensor wurde in einem Hochstromteststand untersucht. Dieser Teststand ist mit einem Stell- und einem Prüftransformator ausgestattet. Sekundärseitig kann die Wicklung am Prüftransformator frei konfiguriert werden. In der verwendeten Konfiguration bestand die Sekundärseite aus einer kurzgeschlossenen Wicklung und es konnten Ströme von 0A bis 750A eingestellt werden. Zur Messung der eingestellten Stromstärke diente eine kommerziell erhältliche Rogowskispule. Um das Messprinzip zu validieren, wurden drei Rundstangen aus Aluminium mit den Durchmessern 20mm, 30mm und 40mm als Leiterseilnachbildung eingesetzt. **Abb. 6** zeigt einen Ausschnitt des Versuchsaufbaus.

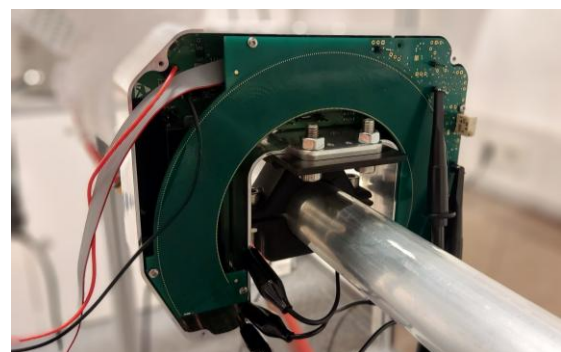


Abb. 6: Versuchsaufbau mit im Sensorgehäuse montierter Rogowskispule auf Leiterseilnachbildung mit 30mm Durchmesser.

Damit das Leiterseil im Zentrum der Rogowskispule verläuft wurden im Rapid-Prototyping-Verfahren drei Leiterseilklemmen gefertigt, welche die genaue Positionierung ermöglichen. Damit ist auch die Vergleichbarkeit mit den anderen

Leiterseilnachbildungen gegeben. Während der Versuche wurde das Sensorsystem mit einem Labornetzteil versorgt. Damit konnte auch der Energiebedarf bestimmt werden. Zusätzlich zur digitalen Auswertung der vom ADC ermittelten Werte wurde ein Oszilloskop verwendet, welches die Ausgangssignale der Rogowskispulen und des Instrumentationsverstärkers aufzeichnete. Da keine Referenz im Sinne einer vollständig geschlossenen Rogowskispule zur Verfügung stand wurden aus den Messwerten der drei Versuche die Mittelwerte gebildet, um die Abweichung der einzelnen Messungen zu den Mittelwerten aufzuzeigen. In **Tab. 1** sind die Mittelwerte dargestellt.

Tab. 1: Mittelwerte der drei Versuche. Die erste Spalte entspricht der eingestellten Stromstärke, die zweite Spalte der Ausgangsspannung des Instrumentationsverstärkers und die dritte Spalte dem digitalisierten und berechneten Messwert des ADCs.

Strom [A]	U_Inst. [V]	ADC [V]
5	0,03	0,05
10	0,05	0,07
25	0,10	0,12
50	0,19	0,21
100	0,38	0,39
200	0,75	0,76
300	1,12	1,13
400	1,49	1,49
500	1,87	1,87
600	2,24	2,24
750/738	2,54	2,52

Hierbei entspricht die erste Spalte der eingestellten Stromstärke, die zweite Spalte der gemessenen Spannung am Ausgang des Instrumentationsverstärkers und die dritte dem digital gewandelten und berechneten Messwert des ADCs. Die Abweichung zwischen der Spannung des Verstärkers und dem Messwert des ADCs sind einerseits durch Fehlabtastungen des ADCs bedingt. Andererseits durch Schwankungen im Messsignal des Oszilloskops. Die Werte wurden im Oszilloskop mit Hilfe von stochastischer Messwertverteilung ermittelt. Im Fall der Stromstärke von 750A ist anzumerken, dass die Leiterseilnachbildung mit Durchmesser 20mm diesen Strom nicht dauerhaft tragen konnte und der Wert nach Erwärmung des Leiters sank. Die Impedanz war zu groß und es konnten nur 738A erreicht werden.

Um die Kalibrierungsfreiheit des Messverfahrens nachzuweisen, müssen die Messwerte der Einzelmessungen mit den Mittelwerten verglichen werden. **Tab. 2** zeigt die entsprechende Übersicht.

Tab. 2: Vergleich der Messwerte der einzelnen Versuche mit den Mittelwerten über alle Versuche. Die erste Spalte entspricht der eingestellten Stromstärke, die zweite Spalte der digitalisierten und berechneten Spannung des ADCs und die Spalten drei bis fünf den Abweichungen der Versuche mit unterschiedlichen Leiterseildurchmessern gegenüber dem Mittelwert.

I [A]	ADC [V]	20mm [%]	30mm [%]	40mm [%]
5	0,05	0,6	1,2	1,8
10	0,07	1,3	0,5	1,9
25	0,12	0,5	1,1	1,9
50	0,21	0,7	0,9	1,1
100	0,39	0,7	0,7	0,9
200	0,76	0,2	1,0	0,0
300	1,13	-0,6	1,6	-0,2
400	1,49	0,4	0,6	-2,4
500	1,87	-0,1	0,5	-2,3
600	2,24	-0,2	0,2	-0,4
750/738	2,52	-1,1	-0,3	-1,1

Die Spalten drei bis fünf stellen die Abweichungen der einzelnen Versuche gegenüber dem Mittelwert aller Versuche dar. Dabei beträgt die Abweichung maximal 2,4% und im Mittel unter 1%. Auch im direkten Vergleich zwischen den Versuchen treten nur kleine Abweichungen von maximal 3% auf. Diese geringen Messabweichungen lassen den Schluss zu, dass der Leiterseilquerschnitt keinen Einfluss auf die Messwerte des Stromsensors hat. Die Abweichungen sind bedingt durch Fehlabtastungen, Messrauschen und durch die Ungenauigkeit beim Einstellen der Stromstärke. Die Stromeinstellung geschieht manuell und unterliegt einem Fehler von ca. +/-1A bis 300A und +/-4A zwischen 300A und 750A. Dementsprechend ist die Wiederholgenauigkeit zwischen den Versuchen gering.

Fazit

In dieser Arbeit konnte gezeigt werden, dass das Messprinzip einer leiterplattenintegrierten Rogowskispule unter Einhaltung der aufgezeigten Randbedingungen keine Kalibrierung bezüglich des

Leiterseilquerschnittes erfordert. Lediglich initial müssen die Messparameter identifiziert werden, um aus der gemessenen Induktionsspannung auf die erregende Stromstärke schließen zu können. Die Randbedingungen werden einerseits durch die Leiterplattenintegration einer kreisförmigen Luftspule garantiert. Andererseits sind sie im Sinne der Leiterseilführung im Zentrum der Rogowskispule konstruktiv sehr genau einstellbar und wiederholgenau.

Nicht Bestandteil der hier durchgeführten Untersuchungen ist der Einfluss von magnetischen Feldern, welche durch Leiter außerhalb des Stromsensors erzeugt werden und in die Rogowskispule einkoppeln. Im Kontext von Hochspannungsfreileitungen wird dies besonders interessant bei der Betrachtung von Bündelleitern.

Der hier entwickelte Stromsensor konnte nicht gegenüber einer Referenz im Sinne einer geschlossenen Rogowskispule verglichen werden. Die Untersuchung der geschlossenen und einer noch weiter geöffneten Spule, zum Beispiel mit einem Öffnungswinkel von 270° , bieten Potential, den Stromsensor kleiner und somit kostengünstiger herzustellen. Zudem kann ein Korrekturmodell geschaffen werden, um vom Referenzwert der geschlossenen Spule auf die Messwerte aller Varianten mit Öffnungen schließen zu können. Ob dieses Verfahren auch kalibrierungsfrei sein kann, müssen zukünftige Arbeiten zeigen.

Eine weitere mögliche Entwicklung besteht in der Senkung des Energiebedarfs bei bestehenbleibender Auflösung und Genauigkeit. Dies kann durch reduzierten Schaltungsaufwand und optimierte Bauteilwahl geschehen. In diesem Zuge kann auch die Versorgungsspannung gesenkt werden.

Literaturnachweis

- [1] W. Rogowski und W. Steinhaus. "Die Messung der magnetischen Spannung. (Messung des Linienintegrals der magnetischen Feldstärke.)", 1912. Archiv für Elektrotechnik, Band I, Heft 4, S.141ff.
- [2] Ljubomir A. Kojovic. „PCB ROGOWSKI COILS – HIGH PRECISION LOW POWER SENSORS“, 2004. Cooper Power Systems, USA.
- [3] Christoph Neeb Jan Niklas Fritz and Rik W. De Doncker. "A PCB Integrated Differential Rogowski Coil for Non-Intrusive Current Measurement Featuring HighBandwidth and dv/dt Immunity", 2015. Institute for Power Electronics and Electrical Drives, RWTH Aachen University Jaegerstr. 17/19, 52066 Aachen, Germany.

- [4] 田慧, 张亚, 堵文灿. „A kind of open-close ring PCB type pulse high current sensor“, 2021. Nanjing University of Science and Technology.

Automatic Sensor Data Fusion for the Detection of Bone Layer Transitions and Channel Length Measurement in Axial Oscillation Bone Drilling

Adriano Boaron¹, Dirk Stöbener^{1,2}, Andreas Fischer^{1,2}

¹ University of Bremen, Bremen Institute for Measurement, Automation and Quality Science, Germany

² MAPEX Center for Materials and Processes, University of Bremen, Bremen, Germany

Abstract

In surgery, bone drilling techniques depend largely on the surgeon's tactile perception, which can be subjective and unreliable. Indeed, accurate detection of bone layer transitions during drilling is critical in orthopedic surgery to avoid over-penetration, as well as thermal and mechanical damage to surrounding tissue. This study introduces automatic sensor data fusion approaches to improve the detection of transitions between cortical and cancellous bone during axial oscillation drilling. Here, multiple signals from a multisensory drive train are fused to enable automatic identification of layer boundaries and overall drill channel length. Experimental validation on pig rib bones shows that time-domain analysis achieves high accuracy, with a standard deviation of 0.22 mm and a mean systematic deviation of 0.26 mm relative to caliper references. The time-frequency approach exhibits larger deviations, with a mean systematic deviation of 1.05 mm and variability up to 1.52 mm. While drill breakthrough is consistently detected, internal layer transition detection reaches 60 % success. Overall, the results demonstrate reliable drill length estimation and highlight the need for improved algorithms for transition detection, supporting the potential of sensor fusion for precise, feedback-driven surgical drilling.

Keywords: Drilling, process monitoring, sensor-fusion, inhomogeneous materials, measurement uncertainty.

1. Introduction

Orthopedic procedures require accurate detection of bone layer transitions during drilling to prevent over-penetration, thermal and structural damage. Conventional bone drilling depends on the surgeon's tactile feedback, which is often subjective and unreliable in complex or minimally invasive interventions. Therefore, Sorg et al. [1] tested a method for estimating the drill channel length when drilling porcine bones by employing a multisensory drive system and an evaluation of the signal amplitudes in the time domain. The method employs a reflective photoelectric sensor, an optical distance sensor as well as a motor current sensor and was experimentally evaluated through drilling tests on pig rib bones. Although a promising measurement uncertainty of 0.3 mm ($k_p = 2$) was reported, validation of the determined bone layer transition using a well-referenced object is still pending.

Guan et al. [2] developed a method to identify the beginning and completion of drilling for spinal pedicle screws aiming at identifying different bone layers. By recording force and acoustic emission signals and analyzing them in the frequency range between 10 kHz and

15 kHz, the position of the transitional region between cortical and cancellous bone layers could be successfully determined.

Furthermore, a method for automatic drill breach identification during spinal pedicle drilling using vibroacoustic sensors was created by analyzing the signal spectrograms and applying a neural network [3]. Various vibroacoustic sensors (including a contact microphone, a free-field microphone, a tri-axial accelerometer, and a uni-axial accelerometer) along with an optical tracking system, were integrated into the experimental setup. By incorporating sensor data fusion and deep learning, the proposed method demonstrates potential to enhance the autonomy of robotic surgery.

Experimental results demonstrated that evaluating the measured force values enables the detection of different bone layers and the drill breakthrough as developed in 2012 by Aziz et al. [4]. The work introduced a method for real-time breakthrough detection during bone drilling using a CRS Catalyst-5 robot. It applies a breakthrough algorithm that analyzes signals from a six-degree-of-freedom force-torque sensor during the drilling process.

Zdero et al. [5] describe experimental methods for measuring force and torque during surgical drilling into whole bone, focusing on instrumentation, data acquisition, and parameter control (e.g., feed rate, spindle speed, drill geometry, and bone type). The study highlights how drilling conditions influence mechanical response and potential thermal damage affecting fixation quality. While force–torque signals may reflect structural variations, the work does not explicitly analyze bone layer transitions or drill channel length as independent variables.

Feldmann et al. [6] experimentally evaluate artificial cortical bone substitute materials for surgical tool development, training, and drill wear analysis. Through controlled drilling tests, thrust force and torque responses of various polymers and composites are compared with bovine cortical bone. Results indicate no material fully replicates bone behavior, revealing trade-offs in performance and durability. The study focuses on material response and does not explicitly analyze bone layer transitions or drill channel length as independent variables.

Torun and Öztürk [7] propose a novel breakthrough detection method for robotic orthopedic drilling based on closed-loop control signals and machine learning. Using features derived from control and force data, the approach achieves high detection accuracy ($\approx 97\% - 98\%$) and enables timely drill stopping within a small breakthrough margin, improving surgical safety. The study focuses on real-time breakthrough detection and drill state estimation, rather than explicitly analyzing bone layer transitions or drill channel length.

This study presents an automatic sensor data fusion method that integrates signals from a multisensory drive train to identify transitions between cortical and cancellous bone layers and the drill channel length during axial oscillation drilling and finally evaluating the respective measurement uncertainties.

Section 2 describes the measurement principle of the multisensory drilling system and the main components of the experimental setup. Section 3 outlines two realizations of a sensor data fusion. In Section 4, the experimental setup is presented and Section 5 shows and discusses the experimental results concerning the drill channel length measurements as well as the layer transition recognition into bone material. Finally, Section 6 concludes the article with closing remarks.

2. Principle of the solution approach

The measurement principle relies on a drilling system which comprises a sensory drive

train (originally introduced in [1]) equipped with various sensors (green), along with a sensory workpiece holder that includes an acoustic emission (AE) sensor, Fig. 1.

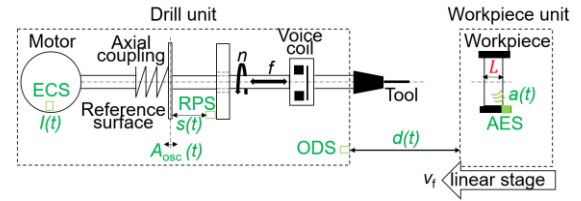


Fig. 1: Measurement principle of the multisensory drilling system, adapted from [1].

On the left-hand side of Fig. 1, the drill unit is schematically shown which remains fixed in a constant position during the drilling process. The electric motor drives the drill by rotating the connected shaft at a specified and controlled rotational speed n .

The motor current $I(t)$ is measured with an electric current sensor (ECS) from which the torque M can be determined via a calibrated constant k :

$$M = k I(t). \quad (1)$$

The rotational speed is measured by using a reflective photoelectric sensor (RPS), which detects a notch on a disk mounted to an axial coupling that rotates with the shaft.

The axial coupling, which links the motor-side input shaft to the tool-side output shaft, allows for small axial displacements of the output shaft relative to the input shaft. As a result, the drill rotates with an additional superimposed axial oscillation f , induced by an electrical voice coil mounted on the tool-side output shaft.

During the drilling process, the feed force F acting on the tool increases the measured distance $s(t)$ between the RPS and the disk, as the axial coupling behaves like a spring that compresses under load and returns to its original state when the load is removed. Consequently, the RPS provides an indirect measurement of the drilling feed force F . The mean feed force F is equal to the average spring force, which is derived from the low-pass-filtered displacement $s(t)$ and the spring constant D ,

$$F = D \cdot (\overline{s(t)} - s_0), \quad (2)$$

where s_0 is the initial distance between the RPS and the disk.

On the right-hand side of Fig. 1, the workpiece unit is schematically represented, which moves towards the drill by means of a linear stage with the set and controlled constant drilling infeed velocity v_f .

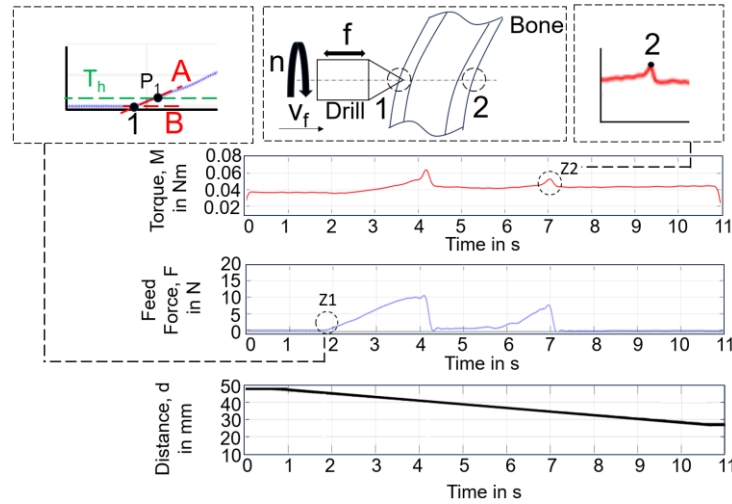


Fig. 2: Automatic sensor data fusion approach based on time domain features. Top left: Initial drill–bone contact. Top middle: Schematic diagram for the determined drill time instant points. Top right: Drill breakthrough determined by the maximum torque signal value. Bottom: Exemplary behavior of M , F and $d(t)$ during drilling.

The distance $d(t)$ between the drill unit and the workpiece unit is measured using an optical distance sensor (ODS) based on laser triangulation, which is used to indirectly measure v_f .

As the workpiece is subjected to F during drilling in addition to contact and friction occurring between the tool and the material, its lattice structure becomes locally distorted and rearranged. This deformation releases energy and produces elastic stress waves within the material, commonly referred to as acoustic emission (AE) waves with an amplitude $a(t)$ that can be captured by an AE sensor. To obtain these waves to derive additional information about the workpiece structural layers and tool breakthrough, an AE sensor (AES) is mounted on the mechanical vise to record the AE signals generated during the drilling process.

3. Sensor Data Fusion

The drill channel length L is automatically obtained through a developed signal processing analysis, which makes use of different sensors' outputs over time and frequency domain features. The time gap between two time instants - t_1 for the detected initial workpiece contact of the drill and t_2 for the drill breakthrough - defines the drilling duration. By using v_f as the average cutting tool infeed velocity, the drill channel length L is determined by calculating

$$L = v_f \cdot (t_2 - t_1) - \Delta L_{\text{spring}} \quad (3)$$

with ΔL_{spring} as the spring compression of the axial coupling at the time t_2 (indirectly determined via the RPS signal).

Two different approaches to determine the time instants t_1 and t_2 have been investigated, and the first one makes use of solely time domain features, whereas in the second approach, t_1 is defined through a time domain feature and t_2 is instead found based on a frequency domain feature. Both approaches are examined with regard to the quality of the channel length result in terms of the achieved systematic and random errors.

3.1 Sensor data fusion based on time domain features

Fig. 2 schematically shows the sensor data fusion algorithm based on time domain features.

The characteristic initial, almost linear rise observed in the feed force F corresponds to the period when the drill contacts the workpiece and the drill is increasingly pressed into the workpiece surface layer. The beginning of the force rise marks the time instant t_1 of the bone entrance. This feature is determined by using the algorithm shown in Fig. 2, in which a threshold T_h above the noise level (green dashed line) is applied, and the first point exceeding this threshold is identified as P_1 . Regression lines (red dashed lines A and B) are then approximated to the signal on both sides of P_1 using an evaluation window of 0.2 s centred

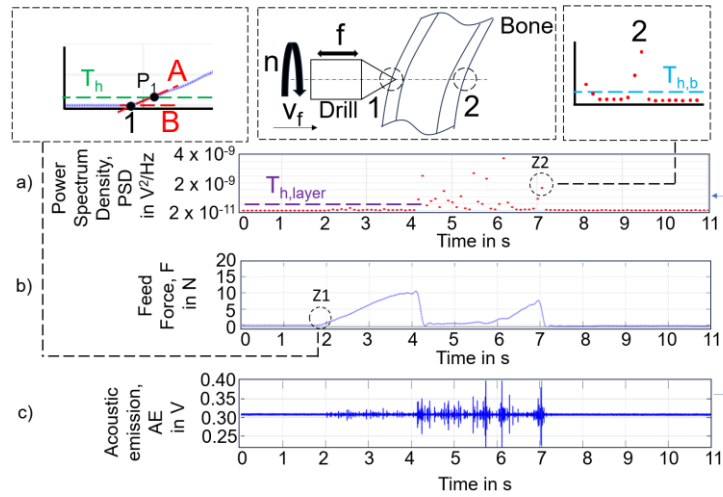


Fig. 3: Automatic sensor data fusion approach based on time and frequency domain features.
Top left: Initial drill–bone contact. Top middle: Schematic diagram for the determined drill time instant points. Top right: Drill breakthrough determined by Power Spectrum Density (PSD) function. Bottom: Measured behavior of PSD, F and AE during drilling.

at P_1 . The intersection of these regression lines defines t_1 , which marks the drill's initial contact with the workpiece.

The moment t_2 , corresponding to the breakthrough, is simply established from the peak value shown by the torque signal (Fig. 2, zoomed view Z2 of point 2).

3.2 Sensor data fusion based on time-frequency domain features

The sensor data fusion algorithm for the evaluation of L based on time-frequency domain features is schematically shown in Fig. 3. As the time instant t_1 cannot be determined from frequency domain features, it is derived in the time domain like in Section 3.1, as shown in Fig. 3b.

Only the time instant t_2 associated with the drill tool exit is obtained through the processing of the AE signal by evaluating its power spectrum density (PSD), see Fig. 3c. The automatic computation of the PSD is implemented in succeeding time intervals of 0.1 s for the complete digitized AE signal. It extracts the maximum value of the PSD, Fig. 3a. Based on the plot of the maximum PSD values over time, the drill breakthrough position t_2 is defined as the last maximum PSD value that oversteps the static limit $T_{h,b}$, as schematically shown in the upper right-hand side of Fig. 3. The definition of $T_{h,b}$ is

$$T_{h,b} = \mu_{\text{PSD_peak_end}} + C \cdot \sigma_{\text{PSD_peak_end}}, \quad (4)$$

where $T_{h,b}$ is the threshold for detecting the breakthrough instant t_2 , $\mu_{\text{PSD_peak_end}}$ refers to the peak mean value at time window Δt_{end} at the

last 1 s time duration of the PSD plot, C is a constant value and $\sigma_{\text{PSD_peak_end}}$ corresponds to the PSD peak standard deviation at time window Δt_{end} .

Although the AE signal is not suitable for detecting the drill entry into the bone, it can be used to identify the cortical-cancellous drill transition inside the bone.

This recognition works similar to the breakthrough detection. By identifying the first maximum PSD value that oversteps the static limit $T_{h,\text{layer}}$, as schematically shown by the pink dashed line of Fig. 3a, the time instant of the transition is defined. The necessary static limit is calculated based on statistical parameters,

$$T_{h,\text{layer}} = \mu_{\text{PSD_peak_initial}} + C \cdot \sigma_{\text{PSD_peak_initial}}. \quad (5)$$

Here $T_{h,\text{layer}}$ is the threshold for detecting the layer transition time instant $t_{\text{layer 1,end}}$ regarding the end of cortical layer, $\mu_{\text{PSD_peak_initial}}$ refers to the peak mean value at time window $\Delta t_{\text{initial}}$ at the first 1 s time duration of the PSD plot, C is a constant value and $\sigma_{\text{PSD_peak_initial}}$ corresponds to the PSD peak standard deviation at time window $\Delta t_{\text{initial}}$.

4. Experimental setup

The developed experimental setup of the multisensory drilling system is shown in Fig. 4. An AC motor from Nanotec (DB41M024030-A) with 24 V, 50 W, 0.5 Nm torque, and 3000 min^{-1} speed has been used for providing the cutting movement to the drill tool. The motor current is measured via the Hall-effect ECS from Allegro (ACS712), producing an AC voltage. Torque $M(t)$ is obtained by removing offset, computing

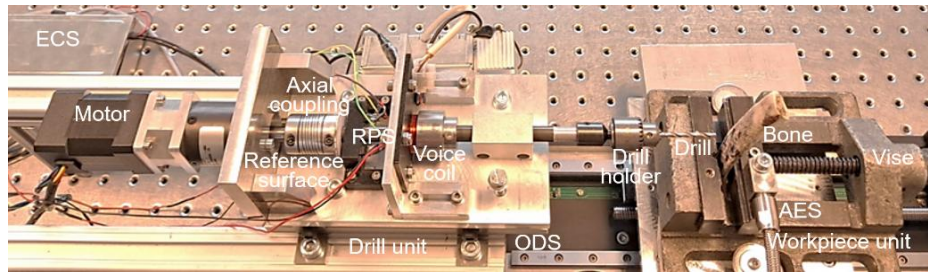


Fig. 4: Experimental setup and procedure.

RMS current, applying a Gaussian filter, and converting the signal using motor's data sheet.

The axial movement of the shaft is generated by a voice coil, driven by a sinusoidal signal with 80 Hz via an amplifier. The RPS to determine the shaft movements by distance measurements to the reference disk is from Vishay Semiconductors (CNY70, 950 nm). Feed force F is derived by applying a 5 Hz low-pass IIR filter to the raw data of the RPS, removing noise and isolating slow mechanical deflections.

The AES from Kistler Instruments GmbH (8152C0050500) measures acoustic emission during drilling and is clamped to the vise holding the pig rib bone. A constant 15 mm distance ensures consistent amplitude. Signals are sent to an AE coupler (5125C3) for amplification and low-pass filtering, by using an analog cut-off frequency of 100 kHz. The coupler outputs a voltage signal.

A linear stage from Physik Instrumente (M-414.2DG) moves the bone material toward the drilling unit at speeds up to 3 mm/s. An ODS from Micro-Epsilon (optoNCDT 1401) uses laser triangulation, offering a 50 mm range, 25 μ m resolution, and 90 μ m linearity.

The analog signals from ODS, ECS, RPS, AES, and voice coil oscillation are digitized using two devices from National Instruments. The NI-9234 samples ECS signals (3.4 kHz) at 51.2 kS/s, while the NI-9222 records ODS (1 kHz), RPS, voice coil oscillation and AES at up to 200 kS/s to prevent aliasing. Data acquisition is implemented in LabVIEW, with processing in MATLAB.

The machining procedure involves constant plunge drilling with axial oscillation. A standard twist drill (3.0 mm diameter, 135° tip angle) is used, and the process parameters are listed in Table 1.

The tool starts at a safe distance, positioned away from the workpiece. Drilling proceeds until a set position beyond the workpiece ensures full penetration. The tool then reverses its feed without rotation or oscillation, preventing further cutting.

Tab. 1: Drilling process parameters.

Drill type	Twist drill
Drill diameter	3 mm
Rotation speed n	600 min ⁻¹
Infeed cutting velocity v_f	2 mm/s
Axial oscillation frequency f	80 Hz

The workpiece thickness is measured with a caliper gauge (0.05 mm resolution) for reference and a total of five experiments are carried out.

5. Experimental Results

Fig. 5 shows the results from five L measurements by using both sensor fusion approaches, when drilling through a pig rib bone. For comparison, also the reference values from caliper gauge measurements of the bone thickness are shown, which are equal to the drill channel length measurements.

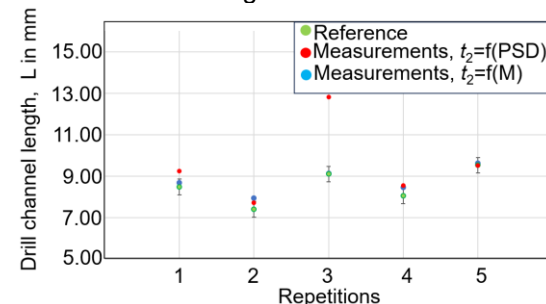


Fig. 5: Repeated drill channel length L measurements during drilling of bone material for both evaluation approaches (time domain and time-frequency domain features) including also caliper reference measurements.

The reference values for each borehole were obtained by averaging four subsequent thickness readings. The caliper gauge values serve as the baseline for evaluating measurement accuracy of the two presented approaches and their error bars show the standard deviation of the mean of the four thickness measurements. The thickness values

scatter over the five different boreholes due to the varying geometry of the bone.

When using the sensor fusion approach based on *time domain* features, a standard deviation of 0.22 mm for the differences between sensor and reference values results. In addition, a mean systematic deviation of 0.26 mm compared to the reference values is observed. These values indicate, that the time domain-based approach achieves good accuracy and precision.

For the sensor fusion approach based on *time-frequency* features the evaluated channel lengths also show that, with exception of repetition 3, the measurements follow the general trend of the caliper-based reference values, thus indicating good agreement across all repetitions. Both overestimation and underestimation are observed, as reflected in deviations ranging from -0.02 mm to 3.72 mm.

The mean systematic deviation between measurement and reference is 1.05 mm. This indicates a positive bias that needs to be corrected for the estimation of t_2 compared to the time domain approach, since t_1 is determined the same way in both approaches. The observed error deviations are generally moderate, (with exception of repetition 3) with most values lying within a standard deviation interval of the differences between measurement and reference values of 1.52 mm, highlighting a reasonable error that needs correction.

In summary, the L measurements found with the time-frequency domain sensor fusion approach also demonstrate overall good consistency with the caliper gauge reference values, characterized by a systematic overestimation of 1.05 mm and an spread measurement variability.

Regarding the detection of the bone-internal transition between the cortical and the cancellous layer, an accurate detection of the transition time instant could recently not be achieved. The detection was only successful for 60 % of the investigated boreholes, indicating that the evaluation algorithm has to be improved for the detection of internal bone layer transitions.

6. Conclusion and Outlook

This study demonstrates the feasibility and effectiveness of an automatic sensor data fusion approach for detecting bone layer transitions and measuring drill channel length during axial oscillation drilling. By integrating multisensory signals and alternative time-domain- and time-frequency-domain-based analyses, the proposed method provides a

reliable framework for real-time process monitoring.

When considering the time-domain analysis, the experimental results confirm that drill channel length can be estimated with sufficient quality for the application in surgical treatments, as indicated by low random measurement errors and acceptable systematic deviations relative to caliper-based references.

While drill breakthrough was consistently and accurately identified with the time-based analysis, corrections need to be implemented in the time-frequency domain approach, in which the detection of intermediate layer transitions showed moderate reliability, highlighting an area for further refinement. Nonetheless, the observed correlation between signal features and physical drilling events confirms the potential of the approach.

Overall, the proposed system enhances the objectivity and robustness of bone drilling procedures, reducing reliance on subjective tactile feedback. The proposed method establishes the foundation for intelligent, feedback-driven drilling systems capable of adapting to bone structure evaluation aiming at enhancing surgical precision and safety. Future research should focus on optimizing spectral feature extraction, reducing systematic errors, and validating the method under varying drilling conditions and bone properties to further strengthen its applicability.

Acknowledgement

The authors gratefully acknowledge the financial support of the current research project by the Volkswagen-Stiftung within the scientific project SMILE (Laboratory for sensitive machining of biological materials) – Funding ID: Momentum.

References

- [1] M. Sorg, J. Osmers, A. Fischer, Methodical Approach for Determining the Length of Drill Channels in Osteosynthesis. *Sensors* 19(16):3532, (2019); <https://doi.org/10.3390/s19163532>
- [2] F. Guan, Y. Sun, X. Qi, Y. Hu, G. Yu, J. Zhang, State Recognition of Bone Drilling based on Acoustic Emission in Pedicle Screw Operation, *Sensors* 18(5):1484 (2018); <https://doi.org/10.3390/s18051484>
- [3] A. Massalimova, M. Timmermans, N. Cavalcanti, D. Suter, M. Seibold, F. Carrillo, J. C. Laux, R. Sutter, M. Farshad, K. Denis, P. Frnstahl, Automatic Breach Detection during Spine Pedicle Drilling based on Vibroacoustic Sensing, *Artificial Intelligence in Medicine* 144:102641 (2023); <https://doi.org/10.1016/j.artmed.2023.102641>
- [4] M.H. Aziz, M.A. Ayub, R. Jaafar, Real-Time Algorithm for Detection of Breakthrough Bone Drilling. *Procedia Engineering - International Symposium on Robotics and Intelligent Sensors* 41,

- 352 – 359 (2012);
<https://doi.org/10.1016/j.proeng.2012.07.184>
- [5] R. Zdero, T. MacAvelia, F.J. Sharifi, Force and Torque Measurements of Surgical Drilling Into Whole Bone. *Experimental Methods in Orthopaedic Biomechanics* 6, 85 – 100 (2017);
<http://dx.doi.org/10.1016/B978-0-12-803802-4.00006-8>
- [6] A. Feldmann, M. Schweizer, S. Stucki, L. Nolte, Experimental Evaluation of Cortical Bone Substitute Materials for Tool Development, Surgical Training and Drill Bit Wear Investigations. *Medical Engineering and Physics* 66, 107–112 (2019);
<https://doi.org/10.1016/j.medengphy.2019.02.010>
- [7] Y. Torun, A. Öztürk, A New Breakthrough Detection Method for Bone Drilling in Robotic Orthopedic Surgery with Closed-Loop Control Approach. *Annals of Biomedical Engineering* 48(4), 1218–1229 (2020); <https://doi.org/10.1007/s10439-019-02444-5>

Quantitative Analyse von Störungen der Periodizität in strukturierten Materialien mittels breitbandiger geführter Wellen

Mareen Wippermann¹, Jonas Hölscher¹, Leander Claes¹

¹ Universität Paderborn, Warburger Straße 100, 33098 Paderborn, Deutschland

Zusammenfassung

In dieser Arbeit wird eine Methode zur Bestimmung der Struktureigenschaften periodischer Materialien mithilfe breitbandiger Ultraschallwellen vorgestellt. Materialien mit periodischen elastischen Eigenschaften, wie gewebeverstärkte Kunststoffe, deren Gewebestruktur maßgeblich die mechanischen Eigenschaften beeinflusst, sind aufgrund ihrer hohen Steifigkeit bei geringer Masse in industriellen Anwendungen von großer Bedeutung. Da die Gewebestruktur durch Umformprozesse beeinflusst werden kann, wird in dieser Arbeit eine quantitative Analyse von Störungen der Periodizität in strukturierten Materialien durchgeführt. Es werden Modellproben aus homogenen Polyamid 6 Platten hergestellt, in die nachträglich definierte Strukturen eingebracht werden. Untersucht werden sowohl Modellproben mit ungestörter Periodizität als Referenzsystem als auch Modellproben mit lokal gestörter Periodizität, um die Auswirkungen von Defekten auf das dispersive Verhalten systematisch zu analysieren. Zur Auswertung werden wiederholt auftretende Moden im Wellenzahlbereich mittels Autokorrelation untersucht. Dabei zeigt sich erwartungsgemäß, dass periodische Effekte im Wellenzahlbereich durch eine Störung der Periodizität in der Struktur der Probe geschwächt werden. Die Höhe des ersten Nebenmaximums quantifiziert die Stärke der periodischen Selbstähnlichkeit, während die Halbwertsbreite Variationen in der Korrelationslänge und damit in der räumlichen Struktur charakterisiert. Die Ergebnisse zeigen, dass die Methode geeignet ist, die Ausprägung von Störungen in Bauteilen mit periodischen Strukturen zu bewerten. Zudem ermöglicht sie eine quantitative Auswertung, sodass Unterschiede zwischen ungestörten und lokal gestörten Proben messbar werden.

Keywords: Faserverbundkunststoffe, Periodizität, geführte Wellen, Störungen von Periodizität, strukturierte Materialien

Motivation

Faserverbundkunststoffe werden in einer Vielzahl industrieller Anwendungen eingesetzt, da sie eine hohe mechanische Steifigkeit bei geringer Masse aufweisen. Zudem können die mechanischen Eigenschaften durch die gezielte Auswahl von Parametern wie dem Material der Faser und der Kunststoffmatrix, der Art des Gewebes, und der Anzahl der Lagen beeinflusst werden. Die flexible Auslegung der faserverstärkten Kunststoffe ermöglicht den Einsatz für spezifische industrielle Anwendungen, insbesondere in der Luft- und Raumfahrttechnik und im Automobil. Ein weiterer Einflussfaktor auf die resultierenden Eigenschaften eines faserverstärkten Kunststoffes ist die Geometrie des Fasergewebes. Die Art des Gewebes ist dabei von entscheidender Bedeutung [1]. Leinwandbindung und Köperbindung sind jeweils orthogonale Bindungsarten mit quadratischen Einheitszellen und unterscheiden sich in der Anzahl der Fasern, die in die jeweilige Richtung gekreuzt werden. Bei der Leinwandbindung

wird jeweils ein Faserbündel gekreuzt, bei der 2/2-Köperbindung hingegen in beide Richtungen zwei Faserbündel. Eine Einheitszelle beschreibt dabei die kleinst mögliche Wiederholeinheit einer in diesem Fall zweidimensionalen periodischen Struktur.

Abgesehen von vorbestimmten produktionsbedingten Unterschieden in den Geometrien der Einheitszelle besteht die Möglichkeit, dass die Geometrie auch durch Fehler im Herstellungsprozess oder durch eine Weiterverarbeitung beeinflusst werden kann. Die vorliegende Untersuchung ist motiviert durch die Umformung eines faserverstärkten Kunststoffes, aus der eine lokale Stauchung bzw. Dehnung der periodischen Struktur des Materials resultiert. Dies kann signifikante Auswirkungen auf die mechanischen Eigenschaften des Materials haben [1]. Darüber hinaus sind Fehlerfälle im Herstellungsprozess denkbar, bei denen das Gewebe beispielsweise verzogen ist. Aufbauend auf den genannten prozess- und anwendungsnahen Beispielen sollen lokale Störun-

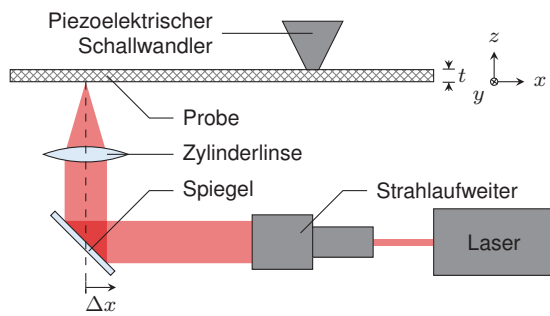


Abb. 1: Experimenteller Aufbau zur Anregung und Detektion breitbandiger Ultraschallwellen in plattenförmigen Proben [2].

gen der periodischen Struktur in Materialien untersucht werden. Für die Untersuchung und quantitative Analyse einer vorliegenden Störung der Periodizität werden Modellproben hergestellt und untersucht. Dabei handelt es sich um homogene Proben, in die periodische Strukturen mit definierten Störungen eingebracht werden. Untersuchungen auf Basis breitbandiger akustischer Wellen, ermöglichen eine Bewertung der Periodizität in Bezug auf eine Referenz. Das langfristige Ziel ist es, den Einfluss von Störungen in der Gewebestruktur eines faserverstärkten Kunststoffes zu quantifizieren, um daraus Rückschlüsse auf die mechanischen Eigenschaften der Materialien ziehen zu können.

Experimenteller Aufbau

Der in Abbildung 1 dargestellte Experimentalaufbau wird zur Untersuchung plattenförmiger Proben als akustischer Wellenleiter mittels breitbandiger Ultraschallwellen verwendet [2]. Die gepulste infrarote Laserstrahlung (FLARE NX 1030, Coherent) mit einer Pulslänge von $1,5 \text{ ns}$ und einer Pulsenergie von etwa $500 \mu\text{J}$ wird zunächst aufgeweitet, anschließend mittels eines Spiegels umgelenkt und schließlich durch eine Zylinderlinse auf die Probe fokussiert. Die pulsformige Erwärmung der Probe durch die Laserstrahlung und die darauffolgende Ausdehnung im Material resultieren in breitbandigen Ultraschallwellen, welche durch die Probe propagieren. Die angeregten Ultraschallwellen werden durch einen piezoelektrischen Schallwandler detektiert. Im Schallwandler wird eine streifenförmige Piezokeramik als aktives Element eingesetzt (PIC255, PI Ceramic) [3]. Eine örtliche Variation der Anregung entlang der x -Achse ermöglicht die Erfassung eines zweidimensionalen Datensatzes, welcher sowohl eine zeitliche Auflösung t als auch eine räumliche Auflösung x aufweist. Abbildung 2 zeigt die Messergebnisse im Zeit- und Ortsbereich für eine Probe mit drei Lagen Leinwandbindung aus

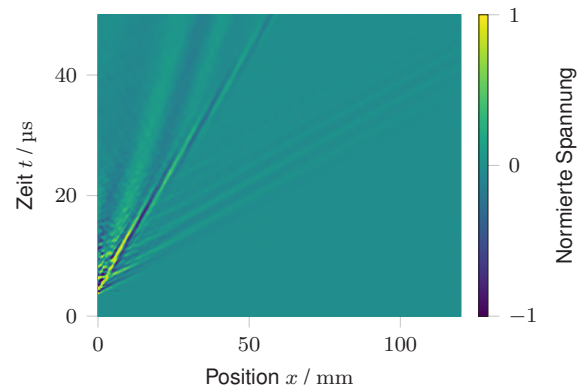


Abb. 2: Messdaten einer Probe mit drei Lagen Leinwandbindung aus Glasfasern, eingebettet in Polyamid 6, aufgenommen mit dem experimentellen Aufbau aus Abbildung 1.

Glasfasern, eingebettet in eine Matrix aus Polyamid 6. Für diese Messung werden 750 Positionen mit einer Schrittweite von $100 \mu\text{m}$ vermessen. Zur Datenverarbeitung wird im Ortsbereich der Offset entfernt und durch Zero-Padding die effektive Auflösung erhöht. Zur Minimierung des Leakage-Effekts wird ein Hann-Fenster angewendet. Mittels der zweidimensionalen Fourier-Transformation kann dieser verarbeitete Datensatz in den Frequenz-Wellenzahlbereich transformiert werden [4]. In Abbildung 3 ist der Betrag der Fourier-Transformierten dargestellt, in dem ausbreitende Moden als Grate sichtbar werden. Dabei werden zwei Quadranten der Fourier-Transformierten überlagert, um Moden mit sowohl positiver als auch negativer Gruppengeschwindigkeit darzustellen. Abschließend wird eine potenzierte Arcus-Tangens-Funktion angewendet, die einer Bildkompression entspricht und die Bilddaten normiert [5]. Mithilfe des vorgestellten Experimentalaufbaus ist es möglich, das dispersive Verhalten plattenförmiger Proben zu analysieren.

Bestimmung der Strukturgröße

Abbildung 3 zeigt Messdaten eines faserverstärkten Kunststoffes (drei Lagen Leinwandbindung aus Glasfasern, Polyamid 6 Matrix). Die Grate in der Darstellung repräsentieren die ausbreitungsfähigen Moden. In diesem Fall sind die A0- und S0-Mode sichtbar. Im Vergleich zu Messungen an einer homogenen Probe sind äquidistanten Wiederholungen im Wellenzahlbereich, insbesondere der A0-Mode, auffällig. Die periodischen Strukturen der vorliegenden Faserverbundkunststoffe lassen sich mathematisch mithilfe periodisch modulierter Materialparameter im Ortsbereich beschreiben. Dieser Ansatz ist analog zum Bloch-

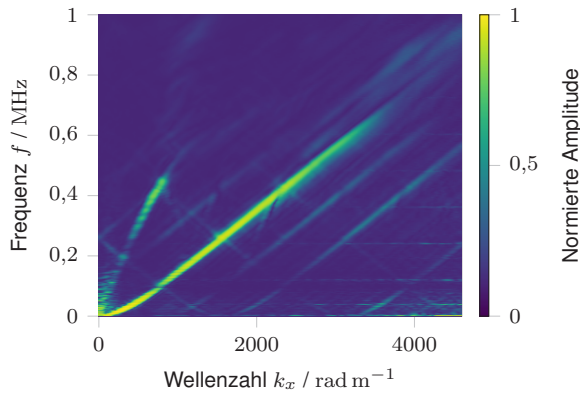


Abb. 3: Der Betrag der zweidimensionalen Fourier-Transformation der Messdaten (siehe Abbildung 2) visualisiert die dispersiven Eigenschaften der vorliegenden Probe. Aus der periodischen Struktur der Probe resultieren Wiederholungen der Moden im Wellenzahlbereich.

Theorem in der Festkörperphysik, das das Verhalten von Elektronen in periodischen Kristallgittern beschreibt [6] und einen Spezialfall des Floquet-Theorems [7] darstellt. Durch die Übertragung dieses Theorems können die vorliegenden periodischen Strukturen (Faserverbundkunststoffe) als phononische Kristalle modelliert werden [8]. Da ein eingeschwungener Zustand betrachtet wird, ergibt sich aus der Periodizität der strukturierten Proben mithilfe des Bloch-Theorems eine Periodizität im Wellenzahlbereich.

Charakterisiert werden die Wiederholungen über die Distanz Δk_x im Wellenzahlbereich (vgl. Abbildung 3). Aus dem Abstand Δk_x kann auf die Größe a der Einheitszelle im Ortsbereich, das heißt auf die Strukturgröße des Gewebes, rückgeschlossen werden. Der reziproke Zusammenhang aus Gleichung 1 ermöglicht, ohne vollständige Rücktransformation in den Ortsbereich, die Bestimmung der Länge a der Einheitszelle in x -Richtung [9].

$$a = \frac{2\pi}{\Delta k_x} \quad (1)$$

Für eine robuste quantitative Bestimmung der Strukturgröße a der Einheitszelle, wird die Autokorrelationsfunktion der Messdaten (Abbildung 3) in Richtung der Wellenzahl k_x berechnet (siehe Abbildung 4), sodass hohe Intensitäten (hier vertikale Linien) der normierten Autokorrelationsfunktion $\Phi(\Delta k_x)$ Wiederholungen repräsentieren. In einem weiteren Verarbeitungsschritt wird die Autokorrelationsfunktion $\Phi(\Delta k_x)$ über die Frequenz gemittelt, sodass in $\bar{\Phi}(\Delta k_x)$ jedes Maximum eine Wiederholung im Wellenzahlbereich repräsentiert.

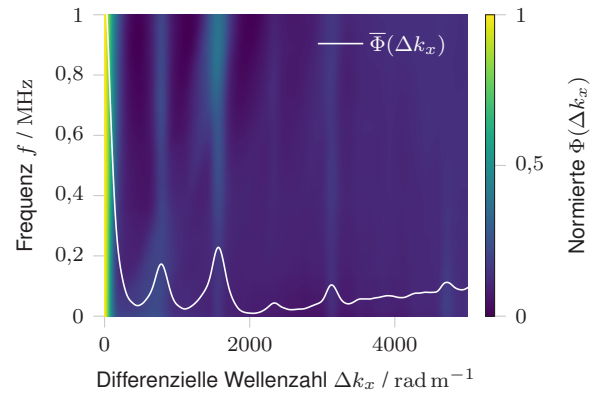


Abb. 4: Autokorrelationsfunktion der Messdaten aus Abbildung 3 berechnet in Richtung der Wellenzahl k_x . Über die Frequenz gemittelte Autokorrelationsfunktion $\bar{\Phi}(\Delta k_x)$ in weiß.

In Abbildung 4 ist die gemittelte Autokorrelationsfunktion $\bar{\Phi}(\Delta k_x)$ der zuvor erwähnten Probe mit Leinwandbindung (siehe Abbildung 3) gezeigt. Diese Methode zur Bestimmung der Größe einer Einheitszelle, also der Strukturgröße in eine Raumrichtung (Messrichtung, hier x -Richtung) wurde an unterschiedlichen faserverstärkten Kunststoffen evaluiert [10]. Die Proben unterscheiden sich in Bindungsart, Faser- und Matrixmaterial oder Anzahl der Lagen. Außerdem kann das Verfahren dahingehend erweitert werden, dass die Propagationsrichtung der geführten Wellen in der Platte variiert und eine winkelabhängig Darstellung ermöglicht wird, welche neben der Größe auch die Geometrie der Einheitszelle sichtbar macht [11].

Modellproben

In der vorliegenden Arbeit werden Modellproben untersucht. Sie dienen als Grundlage, um in Zukunft Fehlstellen in der Struktur von faserverstärkten Kunststoffen zu analysieren. Durch die gezielte Herstellung definierter Störungen wird es ermöglicht, deren Einfluss auf das disperse Verhalten systematisch zu charakterisieren und entsprechende Erwartungshaltungen abzuleiten. Die Untersuchung unterschiedlicher Störungsarten zielt darauf ab, ein grundlegendes Verständnis darüber zu entwickeln, wie sich spezifische Defekte im akustischen Verhalten periodischer Materialien manifestieren.

Die Untersuchung realer Bauteile aus faserverstärkten Kunststoffen ist derzeit mit zusätzlichen Herausforderungen verbunden. Zum einen sind viele strukturelle Störungen von außen nicht eindeutig sichtbar, wodurch eine gezielte Lokalisierung und Bewertung erschwert wird. Zum ande-

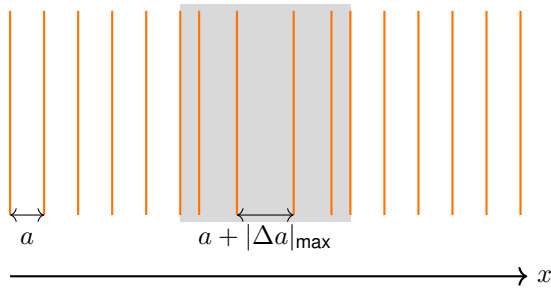


Abb. 5: Eindimensionale periodische Struktur mit Störung im grauen Bereich, welche durch Laserablation in homogene Polyamid 6 Platten realisiert wird.

ren erfordern insbesondere umgeformte Proben eine Anpassung des Messaufbaus, um eine Fokussierung auf die Bauteiloberfläche sicherzustellen. Diese Einschränkungen motivieren den Einsatz von Modellproben mit definierten Eigenschaften, um zunächst ein grundlegendes Prozess- und Systemverständnis zu erarbeiten.

Homogene Kunststoffplatten (Polyamid 6) mit einer Dicke von 1 mm dienen als Grundlage für die Untersuchung von strukturierten Modellproben mit gestörter Periodizität. Die Strukturierung erfolgt mittels Laserablation, bei dem das Material durch linienförmige Vertiefungen modifiziert wird. Die eingebrachten Strukturen haben eine Tiefe von etwa zwei Dritteln der Plattendicke. In die Modellproben werden eindimensionale periodische Strukturen sowohl in ungestörter Form zur Definition eines Referenzsystems als auch mit gezielt eingebrachten Störungen der Periodizität realisiert. Die ungestörte Struktur besteht aus äquidistant angeordneten Linien mit einem Abstand von a . Die gestörte periodische Struktur ist schematisch in Abbildung 5 dargestellt, wobei etwa ein Drittel der Gesamtstruktur von der Modifikation betroffen ist. Durch eine entsprechende Skalierung der Struktur wird im Randbereich der gestörten Zone eine Stauchung und im zentralen Bereich eine Dehnung erzeugt. Ziel dieser Gestaltung ist eine Annäherung an einen realitätsnahen Anwendungsfall: In faserverstärkten Kunststoffen können infolge von Umformprozessen lokale Störungen der Struktur auftreten, bei denen Randbereiche verdichtet (gestaucht) und Bereiche in Umformtälern gedehnt werden.

Auswertung

Die periodisch strukturierten Modellproben werden mit dem oben beschriebenen Experimental-aufbau vermessen. Die Anregung erfolgt auf der Seite der eingepprägten Struktur, während die Detektion auf der gegenüberliegenden Plattenseite durchgeführt wird. Die Struktur erstreckt sich über den gesamten Messweg von 75 mm. Im Falle von

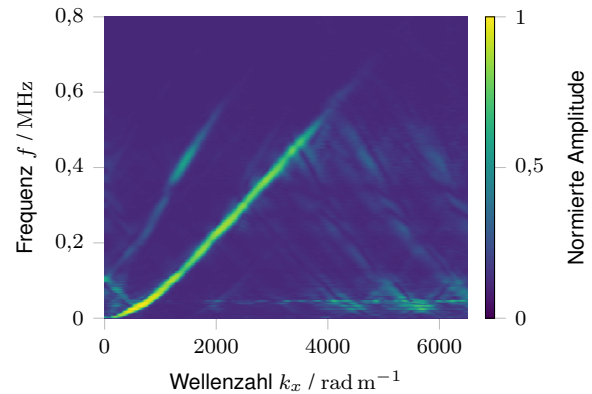


Abb. 6: Messdaten für die Dispersionseigenschaften von akustischen Wellen in einer Platte aus Polyamid 6 mit eingepprägter periodischer Struktur und lokaler Störung gemäß Abbildung 5.

Störungen in der Periodizität ist die Modifikation – wie in Abbildung 5 dargestellt – im mittleren Drittel des Messbereichs lokalisiert.

Im Folgenden wird die Modellprobe mit ungestörter Struktur (Strukturgröße von $a = 5$ mm) mit einer Probe mit einer maximalen Abweichung in der Struktur von $|\Delta a|_{\max} = 1,03$ mm verglichen. Abbildung 6 zeigt die Messdaten der Modellprobe mit gestörter Periodizität. Im Gegensatz zu den Messdaten eines faserverstärkten Kunststoffs (Abbildung 3), bei der von einer ungestörten Periodizität ausgegangen wird, ist erkennbar, dass die zu untersuchenden Wiederholungen im Wellenzahlbereich unschärfer ausfallen. Die über die Frequenz gemittelte Autokorrelationsfunktion beider Modellproben ist in Abbildung 7 dargestellt. Für die Modellprobe mit ungestörter Periodizität, d. h. mit äquidistanten Linien, sind die Maxima der Autokorrelationsfunktion deutlicher ausgeprägt. Zudem bleiben die Maxima vom Betrag auch für größere differentielle Wellenzahlen Δk_x nahezu konstant. Diese Beobachtungen bestätigen mit dem dispersiven Verhalten die ausgeprägte periodische Struktur dieser Probe. Im Gegensatz dazu zeigt die Autokorrelationsfunktion der Messdaten der Modellprobe mit gestörter Periodizität ein verändertes Verhalten. Zwar ist das erste Nebenmaximum nach dem Hauptmaximum noch deutlich erkennbar, jedoch sind die darauffolgenden Maxima schwächer ausgeprägt oder nicht mehr erkennbar.

Sowohl das erste Nebenmaximum der ungestörten Modellprobe als auch das entsprechende Maximum der gestörten Modellprobe treten bei $\Delta k_x = 1255 \text{ rad m}^{-1}$ auf (grün markiert in Abbildung 7). Unter Verwendung von Gleichung 1 ergibt sich daraus eine charakteristische Struktur-

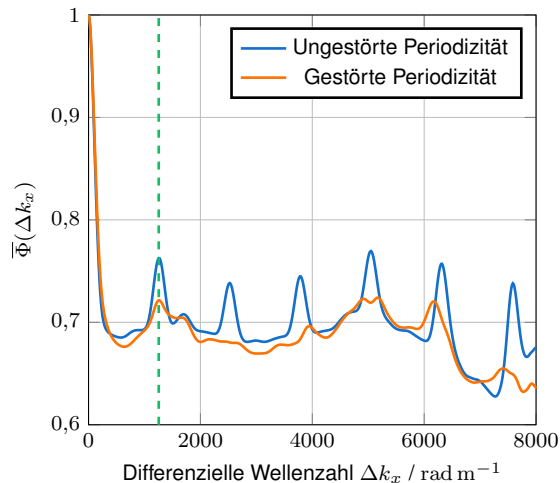


Abb. 7: Die über die Frequenz gemittelte Autokorrelationsfunktion $\bar{\Phi}(\Delta k_x)$ in Richtung der Wellenzahl für die Modellproben mit ungestörter und gestörter Periodizität.

größe von $a = 5,00 \text{ mm}$, was der vorgegebenen Periodizität und somit der Erwartung entspricht. Eine quantitative Betrachtung des ersten Nebenmaximums der gemittelten Autokorrelationsfunktion in Abbildung 8 liefert bei ungestörter Periodizität einen Maximalwert von 0,76, während die Modellprobe mit Störung der Periodizität mit 0,72 einen geringeren Maximalwert erreicht. Da beide Autokorrelationsfunktionen normiert sind, kann die Höhe des ersten Nebenmaximums als Maß für die Ausprägung der Periodizität herangezogen werden.

Neben der Ausprägung kann auch die Halbwertsbreite des Nebenmaximums als Kenngröße zur Beurteilung von Störungen in periodisch strukturierten Materialien dienen. Visuell ist erkennbar, dass das Nebenmaximum der Modellprobe mit ungestörter Periodizität deutlich schmäler ausgeprägt ist als das der gestörten Modellprobe. Die Halbwertsbreite ergibt sich aus dem Wellenzahlintervall, für das die Autokorrelationsfunktion größer ist als der halbe Abstand zwischen ihrem Maximum und dem globalen Minimum im betrachteten Bereich. Quantitativ ergibt sich für die Modellprobe mit ungestörter Periodizität eine Halbwertsbreite von 265 rad m^{-1} , während die Modellprobe mit gestörter Periodizität mit 697 rad m^{-1} eine deutlich größere Halbwertsbreite aufweist.

Eine Erweiterung der Studie um eine Modellprobe mit geringerer maximaler Abweichung der Strukturgröße ($|\Delta a|_{\max} = 0,55 \text{ mm}$) zeigt, dass für beobachtete Größen (Ausprägung des ersten Nebenmaximums und Halbwertsbreite) kontinuierliche Zusammenhänge zur Ausprägung der Störung zu erwarten sind (Tabelle 1). Der Wert des

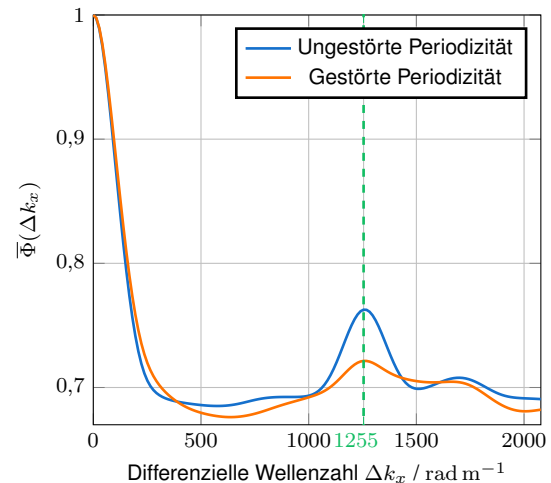


Abb. 8: Betrachtung des ersten Nebenmaximums der Autokorrelationsfunktion $\bar{\Phi}(\Delta k_x)$ für die zu vergleichenden Modellproben.

Tab. 1: Quantitative Gegenüberstellung der ersten Nebenmaxima der gemittelten Autokorrelationsfunktion $\bar{\Phi}(\Delta k_x)$ für drei Modellproben mit variierender maximaler Abweichung $|\Delta a|_{\max}$ der Strukturgröße a .

$ \Delta a _{\max}$ / mm	$\bar{\Phi}_{\max}(\Delta k_x \neq 0)$	Halbwertsbreite / rad m^{-1}
0	0,76	265
0,55	0,73	321
1,03	0,72	697

ersten Nebenmaximums $\bar{\Phi}_{\max}(\Delta k_x \neq 0)$ nimmt mit abnehmender Störung zu und liegt hier mit einem Wert von 0,73 zwischen der ungestörten und der stärker gestörten Modellprobe. Ein analoges Verhalten zeigt die Halbwertsbreite, die bei geringerer Störung deutlich kleiner ist und in diesem Fall ungefähr halb so groß ausfällt wie bei der stärker gestörten Modellprobe.

Die Ergebnisse zeigen, dass sowohl eine geringere relative Höhe als auch die vergrößerte Halbwertsbreite des ersten Nebenmaximums in der Autokorrelationsfunktion bezüglich der Wellenzahl auf eine Störung in der Periodizität der Probe hinweisen. Während die Abnahme der Maximalwerte eine geringere Selbstähnlichkeit der periodischen Struktur widerspiegelt, deutet die Verbreiterung des Peaks auf eine Variation in der Korrelationslänge und damit auf eine stärkere strukturelle Inhomogenität hin. Insgesamt bestätigen beide Kenngrößen konsistent die Beeinträchtigung der Fernordnung infolge der eingebrachten Störung.

Zusammenfassung und Ausblick

In dieser Arbeit wird eine Methode vorgestellt, um Abweichungen in der Strukturgröße periodisch strukturierten Materialien mithilfe breitbandiger Ultraschallwellen zu bestimmen. Die Motivation dieser Arbeit besteht in der Untersuchung periodisch strukturierter Materialien wie faserverstärkter Kunststoffe, deren mechanische Eigenschaften maßgeblich durch Geometrie und Größe der Gewebestrukturen beeinflusst werden, die insbesondere nach Umformprozessen nicht zwingend von außen erkennbar sind. Als Vorarbeit werden in dieser Studie die Grundlagen gelegt, um Störungen in der Periodizität und deren Einfluss auf das dispersive Verhalten solcher Materialien zu charakterisieren. Hierfür werden Modellproben mit ungestörter Periodizität mit Modellproben verglichen, bei denen die periodische Struktur definiert gestört wird. Die Auswertung erfolgt anhand von Autokorrelationsfunktionen im Wellenzahlbereich von Messdaten der akustischen Dispersionseigenschaften. Dabei zeigt sich, dass das erste Nebenmaximum der Autokorrelationsfunktion, welches durch wiederholende Moden hervorgerufen wird, empfindlich auf Störungen der Periodizität im Material reagiert und somit als einfaches, quantifizierbares Maß zur Bewertung solcher Störungen dienen kann.

Für zukünftige Arbeiten sollte diese Untersuchung erweitert werden, um eine größere Vielfalt an Störungsarten systematisch zu charakterisieren, bevor das gewonnene Wissen auf anwendungsnahe Proben (Faserverbundkunststoffe) übertragen wird, bei denen Störungen von außen nicht sichtbar sind. Langfristig ist eine Optimierung des Verfahrens anzustreben, um eine Anwendung in industriellen oder praxisrelevanten Szenarien zu ermöglichen.

Literaturnachweis

- [1] K. M. F. Hasan, P. G. Horváth und T. Alpár. „Potential fabric-reinforced composites: a comprehensive review“. In: *Journal of Materials Science* 56.26 (2021), S. 14381–14415. ISSN: 1573-4803. DOI: 10.1007/s10853-021-06177-6.
- [2] S. Johannesmann, J. Düchting, M. Weberßen, L. Claes und B. Henning. „An acoustic waveguide-based approach to the complete characterisation of linear elastic, orthotropic material behaviour“. In: *tm - Technisches Messen* 2018.85 (2018), S. 478–486. DOI: 10.1515/teme-2017-0132.
- [3] L. Claes, T. Meyer, F. Bause, J. Rautenberg und B. Henning. „Determination of the material properties of polymers using laser-generated broadband ultrasound“. In: *Journal of Sensors and Sensor Systems* 5.1 (2016), S. 187–196. ISSN: 2194-878X. DOI: 10.5194/jsss-5-187-2016.
- [4] D. Alleyne und P. Cawley. „A 2-dimensional Fourier transform method for the quantitative measurement of Lamb modes“. In: *IEE Symposium on Ultrasonics*. IEEE, 1990. DOI: 10.1109/ultsym.1990.171541.
- [5] L. Claes, S. Johannesmann, H. Zeipert und B. Henning. „Broadband acoustic waves in plate-like structures for acoustic material characterisation“. In: *2023 International Congress on Ultrasonics, Beijing, China*. Bd. 2822. IOP Publishing, Sep. 2024, S. 012171. DOI: 10.1088/1742-6596/2822/1/012171.
- [6] F. Bloch. „Über die Quantenmechanik der Elektronen in Kristallgittern“. In: *Zeitschrift für Physik* 52.7-8 (1929), S. 555–600.
- [7] G. Floquet. „Sur les équations différentielles linéaires à coefficients périodiques“. In: *Annales scientifiques de l'École normale supérieure*. Bd. 12. 1883, S. 47–88.
- [8] A. Khelif und A. Adibi. *Phononic Crystals: Fundamentals and Applications*. Jan. 2015, S. 1–245. ISBN: 978-1-4614-9392-1. DOI: 10.1007/978-1-4614-9393-8.
- [9] J. D. Jackson. *Classical electrodynamics*. John Wiley & Sons, 2021.
- [10] M. Wippermann, L. Claes, P. Brandes, E. Moritzer und B. Henning. „Determination of the unit cell geometry in fibre-reinforced polymer sheets using guided acoustic waves“. In: 2025. DOI: 10.71568/DASDAGA2025.116.
- [11] M. Wippermann und L. Claes. „Phononic crystallography of fibre-reinforced polymers using broadband acoustic waves“. In: *2025 International Congress on Ultrasonics, Paderborn, Germany*. AMA Service GmbH, 2025, S. 24–27. DOI: 10.5162/ultrasonic2025/a1-b1.

Improvements of the Accuracy of Generic Extrinsic Camera Calibration with ICP

Carsten Schmerbeck¹ ; Linus Dau¹; Michael Heizmann¹ 

¹Karlsruhe Institute of Technology, Institute of Industrial Information Technology (IIIT), Karlsruhe

Abstract

For the extrinsic calibration of a camera system, a precise estimation of a set of calibration points is necessary [1]. However, there are influences that reduce the accuracy of point localization, such as lighting, exposure time, and thereby resulting in the over- and underexposure of the recorded images. This paper demonstrates the influence of varying exposure times on the localization of ArUco calibration markers [2] for generic ICP-based extrinsic calibration [1, 3, 4]. We evaluated the perceived location of the marker corners and thus their perceived size for various exposure times due to blur and overexposure. Furthermore, we propose a method of enabling subpixel accuracy by combining adjacent rays with a weighted mean. For all settings, we quantified the effect on the extrinsic camera calibration. We found, that the markers appear smaller with increasing exposure times, which can be explained by blur of the marker borders.

Keywords: Camera calibration, ICP

1 Introduction

1.1 Classic intrinsic camera calibration

The intrinsic camera calibration is the process to determine the internal camera parameters. A precise determination of these parameters is paramount for computer vision and optical metrology applications. The intrinsic parameters are specific for a camera and need to be calibrated for each camera individually. These parameters include the focal length (f_x, f_y) and the optical centers (c_x, c_y), which can be used to create the camera matrix $\mathbf{K}$. Furthermore, there are distortion coefficients that can improve the quality of the calibration [2].

$$\mathbf{K} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (1)$$

1.2 Generic intrinsic camera calibration

For some applications, these few parameters are not enough to describe the camera sufficiently accurate. This finding is backed by the experiments of Schöps et al. [5]. This especially applies to non-central cameras such as plenoptic cameras [3].

Generic intrinsic calibration is a model-free black-box representation of a camera, consisting of a ray or raxel representation of each camera pixel individually, as introduced by Grossberg and Nayar [6].

1.3 Exposure in camera calibration

Illumination and exposure significantly influence the performance of camera calibration systems. This is backed by various publications, e.g.

- Cledat et al. [7] showed the change of perceived size of ArUco markers with varying exposure.

- Experiments of Mousavirad et al. [8] showed this relation for a distance estimation system.
- Agnus et al. [9] proposed an approach to estimate the effect of overexposure in rectangular ArUco patterns.
- Mondéjar-Guerra et al. [10] simulated the effect of overexposure on ArUco markers using a dilation operation.
- Berral-Soler et al. [11] proposed a deep learning-based framework for ArUco detection in challenging lighting conditions.

2 Hardware setup

2.1 Cameras

The cameras used in this experiment are made of Raspberry Pi 3 [12] and Raspberry Pi HQ Camera modules with C/CS-mount. These parts are built into an aluminum tube, which grants the cameras the ability to be stacked or easily mounted into the calibration setups. The system is explained in more detail in [1, 13].



(a) Intrinsic setup.

(b) Extrinsic setup.

Figure 1: Hardware setups for calibration.

2.2 Intrinsic calibration setup

For the intrinsic calibration, a single camera is placed in front of a monitor, which serves as an active reference target. The camera is mounted on a linear axis to be able to be moved automatically towards the monitor. This setup is put in an opaque enclosure whose walls are painted black to absorb any stray light from the monitor. The system is explained in greater detail in [1, 13].

2.3 Extrinsic calibration setup

For the extrinsic calibration, the intrinsically calibrated cameras are mounted on a metal frame in varying positions directed at a reference plane below them. On this plane, ArUco boards are placed as reference targets. The locations of the ArUco boards are known and are used to estimate the camera's position and orientation [14].

This is achieved by placing the reference targets on a flat reference plane, a precision-milled piece of cast aluminum¹ measuring $1000 \times 1000 \times 25$ mm. The material has a specified surface flatness of $130 \mu\text{m m}^{-1}$, as well as extremely low internal stresses [15]. The plate is set up on six equally spaced, adjustable feet that were adjusted until no more deviations could be detected using a 1000 mm precision steel straightedge² with straightness according to DIN874 GG1 ($\approx 20.67 \mu\text{m}$) [16, p. 181] and a $20 \mu\text{m}$ feeler gauge³. This procedure of measurement is in accordance to DIN876:2 [17, Sec. 4.3]. A placement of the plate on its three Bessel points [17, Sec. 3.1][18, Sec. 3.1.5, 5.4.2] did not produce a sufficiently flat surface.

The positioning of the targets is achieved through aligning them with a plus-shaped 3 mm sheet of aluminum, that was milled around its entire circumference to the required contour within a specified tolerance of $\pm 50 \mu\text{m}$ ⁴. The targets themselves are 4 mm aluminum composite plates, on which the ArUco markers were printed⁵. Beforehand, the markers were positioned in a milling machine by the corners of the prints using a centering microscope⁶, and their circumference was milled to match the desired contour within a specified tolerance of $\pm 50 \mu\text{m}$ ⁷.

3 Generic calibration procedure

3.1 Generic intrinsic calibration

The intrinsic parameters were determined in two steps identical to the description by Zörn et al. [1]: (i) The monitor displays a sequence of sinusoidal patterns, whereby the phase of the signal contains

¹G.AL C250 by Gleich Aluminium GmbH

²0465103 by Helios Preisser / Mahr GmbH

³W2868B by Jiechao Tools Co., Ltd.

⁴Milled with CAM 1020 Premium by vhf camfacture AG

⁵Unidirectional UV direct printing on 4 mm aluminum composite board by Hannes Industriegravuren GmbH

⁶MZM30 by Erdmann & Grün KG Wetzlar

⁷Milled with FP2 by Friedrich Deckel AG, measurements with LS388C by Dr. Johannes Heidenhain GmbH

coding on the pixel coordinate. For each pattern displayed, an image is taken from the camera at 10 equidistant known positions in front of the monitor. The positioning was automated using a motorized linear axis, moving the camera towards the monitor [1]. Through probabilistic phase unwrapping the registration between the camera pixels and monitor is determined. For this, a modified version of Uhlig's implementation [3, Ch. 4-5] is used.

(ii) An alternating minimization process is used to estimate both the set of rays intersecting the registered points on the monitor and to estimate the set of monitor poses ($\mathbf{R}_{\text{intr.}}, \mathbf{t}_{\text{intr.}}$) for the intrinsic calibration [3, Ch. 5].

3.2 Generic extrinsic camera calibration with ICP

The goal of the extrinsic calibration is to determine the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$, that describe the camera location and orientation in the extrinsic calibration setup, hereafter called scene. Contrary to the intrinsic calibration, the extrinsic calibration must be performed every time the camera is moved.

The ArUco marker corners on the boards are the known calibration points $\mathcal{P}$ in the 3D scene. The surrounding cameras take corresponding pictures of the scene. The calibration points $\mathcal{P}$ are detected in the image with the OpenCV `detectMarkers()` function [2]. This results in the 2D corner coordinates $\mathbf{h}$ of each marker in the given picture that correspond to the known 3D calibration points $\mathcal{P}$. With this the `SolvePnP()` function from OpenCV [2] is used to get an initial guess of the rotation matrix $\mathbf{R}_0$ and the translation vector $\mathbf{t}_0$. They are used as starting values for the ICP (iterative closest point) algorithm.

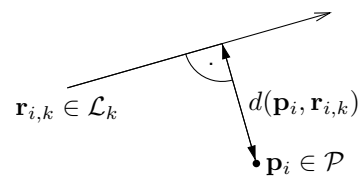


Figure 2: ICP method illustration for step k .

The ICP algorithm transforms the rays $\mathcal{L}$ to align with the known points $\mathcal{P}$. This process is described in [1] and uses a modified version of ICP originally introduced by Besl and McKay [19]. Furthermore it was adapted to align rays to points [20, Sec. 10.2.1]. This procedure is repeated for each camera individually. The procedure is explained in more detail in [1].

3.3 RMSE residuum in the ICP process

The extrinsic calibration accuracy is described by the RMSE of the minimum distance between a calibration target and the according calibrated camera ray for each ICP step:

$$\text{RMSE}_k(d(\mathbf{p}, \mathbf{r})) = \sqrt{\frac{1}{N_{\text{obs}}} \cdot \sum_i^{N_{\text{obs}}} [d(\mathbf{p}_i, \mathbf{r}_{i,k})^2]} \quad (2)$$

Here, N_{obs} is the number of observed points. $\mathbf{p}_i$ is the location of the i -th calibration point and $\mathbf{r}_{i,k}$ the according ray [1, 4, 14].

4 Improvements

4.1 Overexposure

Classical calibration approaches have mitigated this problem by using different patterns for calibration, namely checkerboard or ChArUco boards. The disadvantage of using such patterns is that they require a larger observed area without any occlusions [2, 9]. Our approach only requires individual ArUco markers to be fully visible. This was tested with camera poses that only see subsets of the markers and recordings with parts of the markers being obstructed [1, 4].

Furthermore we experienced in our setup, that glare due to spotlights and unsuitable camera positions negatively affected the detection quality. This is a common problem for various machine vision tasks, but especially for ArUco detection, as backed by the findings of Berral-Soler et al. [11]. This issue was addressed in the experiment setup by using diffuse lighting over a large surface area [13, Sec. 2.4.2]. We experienced, that for various exposure levels, the detected location of the marker corners slightly shifted towards the center of the markers with increasing exposure time.

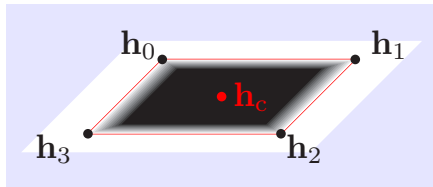


Figure 3: Schematic illustration of the blurred ArUco marker edge. The marker's internal structures are omitted. The red border is the real border, the shading illustrates the perceived brightness due to blur and defocus.

This can be explained by the blur recorded along the marker's boundaries (see Figure 3). To counteract this effect, we used the mean of the four corner coordinates as reference points, as we assume that the effect of blur is nearly equal for all four corners and the perceived shift of the four corners would cancel out.

$$\mathbf{h}_c = \frac{1}{4} \sum_{j=0}^3 \mathbf{h}_j \quad (3)$$

As opposed to the implementation of Agnus et al. [9] this proposed method does not require adjacent markers in a certain geometric position.

4.2 Subpixel mixture of rays

To utilize the **subpixel accurate corner detection** [2] to its full extent, we propose a method of mixing the four adjacent rays $\mathbf{r}_j = \mathbf{r}_0 + \lambda \mathbf{r}_a$ with $\mathbf{r}_0$ being a point on the ray and $\mathbf{r}_a$ its direction using bilinear interpolation.

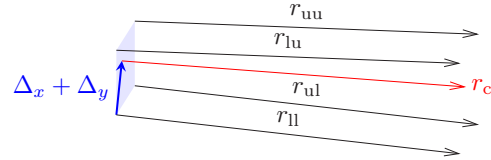


Figure 4: Illustration of the mixture of four adjacent rays.

The next lower or upper adjacent points are denoted with l and u respectively and the deltas denote the relative distance to the lower adjacent point $\mathbf{h}_{ll}$ [21, Ch. 2.3]. The combination resembles a weighted mean of the four rays. This procedure was performed both for the ray's origin $\mathbf{r}_0$ as well as its direction vector $\mathbf{r}_a$.

$$\begin{aligned} \mathbf{r}_{0,c} = & \mathbf{r}_{0,ll}(1 - \Delta x)(1 - \Delta y) \\ & \mathbf{r}_{0,ul}(\Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{0,lu}(1 - \Delta x)(\Delta y) + \\ & \mathbf{r}_{0,uu}(\Delta x)(\Delta y) \end{aligned} \quad (4)$$

$$\begin{aligned} \mathbf{r}_{a,c} = & \mathbf{r}_{a,ll}(1 - \Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{a,ul}(\Delta x)(1 - \Delta y) + \\ & \mathbf{r}_{a,lu}(1 - \Delta x)(\Delta y) + \\ & \mathbf{r}_{a,uu}(\Delta x)(\Delta y) \end{aligned} \quad (5)$$

Due to the rays being normalized by Uhlig's implementation [3, Sec. 5.2], the support vectors $\mathbf{r}_0$ are located near the optical center of the camera (where the rays are most dense). Therefore a weighted mean will find a suitable combined support vector $\mathbf{r}_{0,c}$. This method relies on the assumption, that the optical system of the camera is similar for adjacent pixels. Discontinuities between these pixels cannot be detected or treated. However, for regular camera geometries, this assumption is reasonable [3, Sec. 4.2].

5 Results

5.1 Overexposure

In Figure 5, we can observe the perceived position shift of the marker's corners viewed from an exemplary camera, based on its orientation in the scene. Each color represents one corner of a ArUco marker (top left, top right, bottom right, bottom left). We set the perceived position of the marker at the shortest exposure time as reference for a relative measurement.

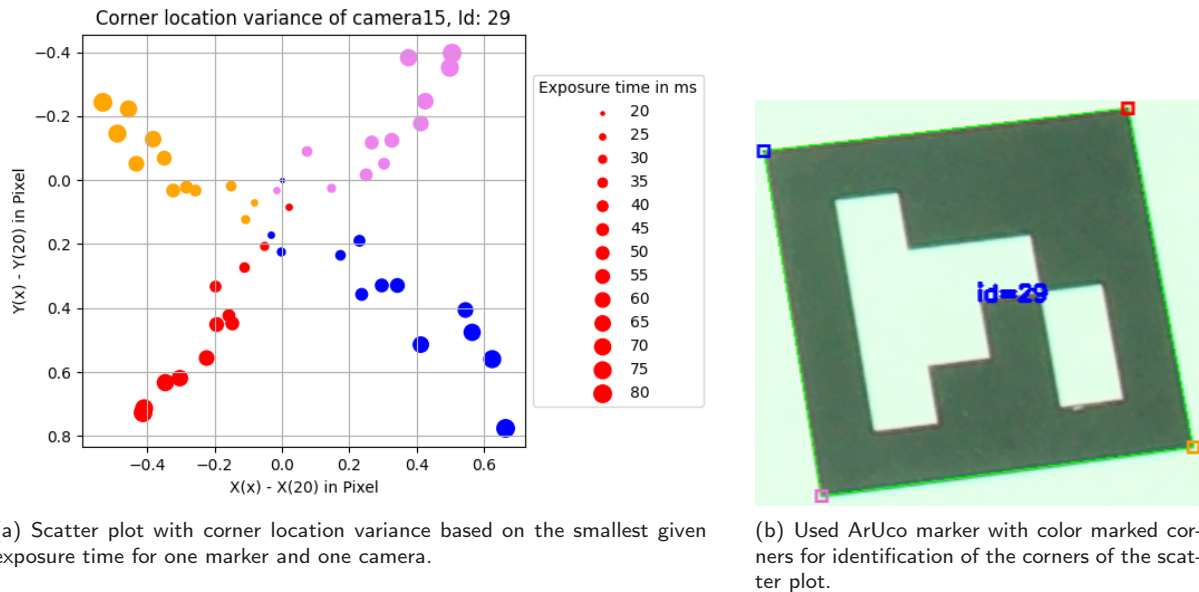


Figure 5: Corner location of an ArUco marker for different exposure times.

We can observe that each marker moves further away from its origin as exposure time increases. This hints to the fact that the perceived marker corners are detected more towards the center of the marker, due to the bright surroundings, introducing light to the nearby pixels. The observed effect of shift towards the marker center is the same as described by Agnus et al. [9] and Cledat et al. [7, Fig. 1].

The extrinsic calibration, using the corners of the marker as calibration points, results in a higher RMSE for higher exposure times. Especially if only one ArUco board is observed, the effect is the most noticeable.

5.2 Subpixel mixture of rays

To examine the influences of both measures (overexposure compensation by averaging of corners and usage of mixed rays) we present Figure 6. Here, the residual RMSE of the ICP process is plotted on the ordinate axis over various exposure times t_{exp} (abscissa).

The experiment was repeated for multiple individual cameras (columns), each individually intrinsically calibrated, and with the usage of the marker's corners directly (top row), as well as with the proposed weighted mean (bottom row). The graphs include both the values for utilizing full pixels and rays (blue) and the mixture of adjacent pixels (orange).

For the usage of the corners we can observe the rising RMSE associated with rising exposure times and thus most probably the effects of overexposure as described above. By utilizing the marker's centers, we can no longer observe this relation, as conjectured above.

Interestingly, the subpixel method does not result in any significant increase in calibration quality.

Furthermore, we can observe that the span of RMSE values is significantly smaller for the centers case as opposed to the usage of the marker corners. This is probably related to the process of averaging multiple observations of markers, thus reducing the measurement error.

Unfortunately the physical setup of the cameras does not allow for an accurate measurement of the optical centers of the cameras and thus the results cannot be verified exactly. However we measured the camera positions in the physical setup and considered them to be plausible.

6 Conclusion

In summary, we quantitatively evaluated the influence of overexposure on the process of generic extrinsic camera calibration.

Based on these findings, we improved the ICP method's results by utilizing the marker's centers and proposed a novel method for subpixel accurate ICP-based generic extrinsic calibration.

Our results indicate a strong dependency of the perceived marker corner location with varying exposure times. However, we found that the proposed approach of using the marker's center diminishes this relation and yields better RMSE results. Thus more robustness and accuracy for different lighting conditions can be expected for future experiments.

The usage of subpixels and mixed adjacent rays did not result in a significant decrease of the RMSE values. This might be an effect of image distortions as they are not corrected by the generic calibration. However for small distances in the image, the effect

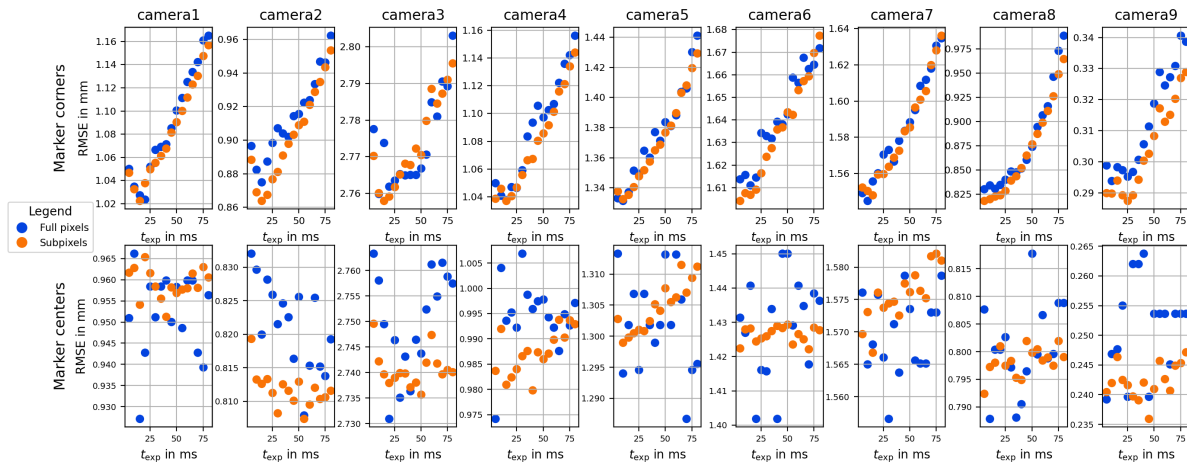


Figure 6: Resulting extrinsic RMSE for each camera for all combinations of calibration methods. Note the different scaling of the ordinate values.

is expected to be minor.

However there are still issues that future research should address: The extrinsic calibration result cannot be accurately verified with the current physical setup. Additionally, the current measure for accuracy, the RMSE residual distance between rays and calibration points might not be ideal as it allows for false local minima. Furthermore there might be other physical or optical effects influencing the accuracy of this process, such as image noise, chromatic aberration or systematic errors during the intrinsic calibration.

Moreover, the approach of overexposure-compensated calibration could be realized with different marker geometries, as proposed by Cledat et al. [7]. Especially the combination with generic camera calibration could result in more precise camera calibration.

Acknowledgments

- The authors would like to thank the Ministry of Science, Research and Arts of the Federal State of Baden-Württemberg for the financial support of the projects within the Innovation-Campus Future Mobility.
- The authors would like to thank the mechanical workshops at the IIIT and IRS for the manufacturing and technical support.

References

- [1] M. Zürn et al. “Robotic Wiring Harness Bin Picking Solution using a Deep-Learning-based Spline Prediction and a Multi-Stereo Camera Setup”. In: *SCAP2024* (2024).
- [2] G. Bradski. “The OpenCV Library”. In: *Dr. Dobb’s Journal of Software Tools* (2000).
- [3] David Uhlig. “Light Field Imaging for Deflectometry”. PhD thesis. Karlsruher Institut für Technologie (KIT), 2023. DOI: 10.5445/IR/1000156025.
- [4] Carsten Schmerbeck and Michael Heizmann. *Anwendung generischer Kamerakalibrierung für optische Messsysteme in der Fertigungstechnik*. German. 126. Jahrestagung der Deutschen Gesellschaft für angewandte Optik (DGaO 2025), Stuttgart, 11.–14. June 2025. 2025.
- [5] Thomas Schöps et al. *Why Having 10,000 Parameters in Your Camera Model is Better Than Twelve*. 2019. DOI: 10.48550/ARXIV.1912.02908.
- [6] M. D. Grossberg and S. K. Nayar. “A general imaging model and a method for finding its parameters”. In: *Proceedings 8. ICCV 2001*. Vol. 2. ICCV-01. IEEE Comput. Soc, 2001, pp. 108–115. DOI: 10.1109/iccv.2001.937611.
- [7] E. Cledat, M. Rufener, and D.A. Cucci. “Compensating over- and underexposure in optical target pose determination”. In: *Pattern Recognition* 116 (2021), p. 107930. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2021.107930. URL: <https://www.sciencedirect.com/science/article/pii/S0031320321001175>.
- [8] Seyed Jalaeddin Mousavirad et al. “Addressing Contextual Factors in ArUco Marker-Based Distance Estimation: A Machine Learning Approach”. In: *2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, May 2025, pp. 1–

6. DOI: 10.1109/i2mtc62753.2025.11079149.
- [9] Vincent Agnus, Stéphane Nicolau, and Luc Soler. "Illumination independent marker tracking using cross-ratio invariance". In: *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*. VRST '14. ACM, Nov. 2014, pp. 69–72. DOI: 10.1145/2671015.2671119.
- [10] Víctor Manuel Mondéjar-Guerra et al. "Classification of Fiducial Markers in Challenging Conditions with SVM". In: *Pattern Recognition and Image Analysis*. Ed. by Luís A. Alexandre, José Salvador Sánchez, and João M. F. Rodrigues. Cham: Springer International Publishing, 2017, pp. 344–352. ISBN: 978-3-319-58838-4.
- [11] Rafael Berral-Soler et al. "DeepArUco++: Improved detection of square fiducial markers in challenging lighting conditions". In: *Image and Vision Computing* 152 (Dec. 2024), p. 105313. ISSN: 0262-8856. DOI: 10.1016/j.imavis.2024.105313.
- [12] Raspberry Pi. *RPI Documentation - Camera*. 2024. URL: <https://www.raspberrypi.com/documentation/accessories/camera.html> (visited on 03/27/2024).
- [13] Carsten Schmerbeck and Michael Heizmann. "VAMOS-Cam: Versatile Affordable Modular Stackable Camera System". 2025. DOI: 10.17632/75x3v4jyr2.1. in preparation.
- [14] Carsten Schmerbeck, Hesheng Wang, and Michael Heizmann. "Optimal aperture settings for generic camera calibration". In: *tm - Technisches Messen* 92.s1 (2025), pp. 93–98. DOI: 10.1515/teme-2025-0037.
- [15] G.AL® C250 – Aluminium Präzisionsplatte. GLEICH Aluminium GmbH, 2023. URL: https://www.gleich-aluminium-shop.de/module/mysydeshop/webforms/download_file.php?file=23 (visited on 10/16/2025).
- [16] *Messtechnik-Katalog*. Helios Preisser, 2023. URL: <https://metrology.mahr.com/fileadmin/catalogs/OnlineCatalogs/Helios-Preisser/index.php?catalog=HP-Metrology-DE&lang=de> (visited on 10/16/2025).
- [17] DIN Deutsches Institut für Normung e.V. *DIN 876-2:2014-12 - Prüfplatten; Prüfplatten aus Gußeisen; Anforderungen, Prüfung*. German Institute for Standardization (DIN). Berlin, 1984.
- [18] H.-J. Warnecke and W. Dutschke. *Fertigungsmeßtechnik. Handbuch für Industrie und Wissenschaft*. Ed. by H.-P. Grode. Vol. 1. Springer, 1984, p. 814.
- [19] P. J. Besl and Neil D. McKay. "A method for registration of 3-D shapes". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.2 (1992), pp. 239–256. DOI: 10.1109/34.121791.
- [20] P. J. Schneider and D. H. Eberly. *Geometric Tools for Computer Graphics*. Elsevier Science, 2003.
- [21] Fernando Puente León. *Messtechnik. Grundlagen, Methoden und Anwendungen*. 11th ed. Springer Vieweg, 2019, p. 481.

Vision Guided Laser Tracking System for Dynamic Vibrometry using a Scanning Laser Doppler Vibrometer and ArUco Marker

Aparajita Talukder, Christian Rembe

TU Clausthal, Institut für Elektrische Informationstechnik, Leibnizstraße 28, 38678 Clausthal-Zellerfeld, aparajita.talukder@tu-clausthal.de

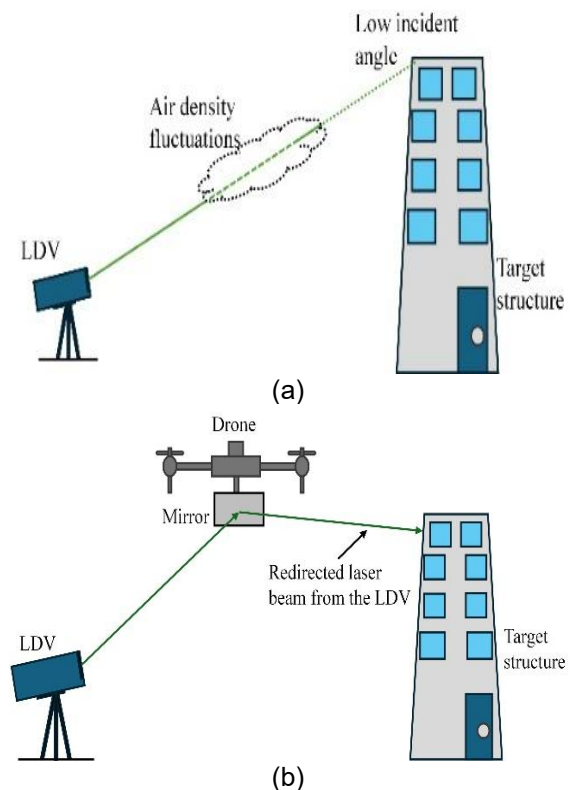
Abstract

In this paper, a vision-guided laser tracking system using a camera-based proportional feedback loop and a scanning laser Doppler Vibrometer (SLDV) is presented. An ArUco fiducial marker is used as a landmark for the system. The ArUco marker is mounted on a linear table that provides translational motion along the x -axis. The system detects the centroid of the ArUco marker at each image frame through image processing. Based on the updated centroid position, steering commands are sent to the galvo mirrors of the SLDV so that the laser beam can follow the moving target in real time. To evaluate the performance of the laser tracking system, experiments are conducted utilizing a linear table at a speed of 24 mm/s, 36 mm/s and 60 mm/s. The maximum tracking error, defined as the difference between the laser position and the marker centroid, is found to be 8 pixels (9.6 mm) at 24 mm/s, 10 pixels (12 mm) at 36 mm/s and 13 pixels (15.6 mm) at 60 mm/s speed. In addition, the root mean square (RMS) tracking error over the full measurement duration is calculated and discussed.

Keywords: Scanning Laser Doppler Vibrometer (SLDV), Image processing, ArUco marker detection, Translational Motion, Laser beam tracking.

1. Introduction

A laser Doppler vibrometer (LDV) is a non-contact measurement instrument widely used to measure surface vibrations based on the principle of optical interferometry [1, 2]. In recent years, LDV has gained increasing importance in the field of Structural Health Monitoring (SHM) [3]. However, measurements become challenging when the structure of interest is located far away from the LDV or the structure has complex geometric curves. In such cases, a ground-based LDV often cannot measure the vibration of the structure reliably [4]. This limitation mainly results from low angles of incidence and air-density fluctuations, as shown in Figure 1(a). One possible solution is to integrate a drone-mounted mirror to redirect the laser beam. This approach brings the measurement path closer to the structure and improves the angle of incidence [4, 5], as shown in Figure 1(b). To realize such a method, the measurement laser beam of the ground LDV must be able to track the drone-mounted mirror in real time during flight [5], as shown in Figure 1(c).



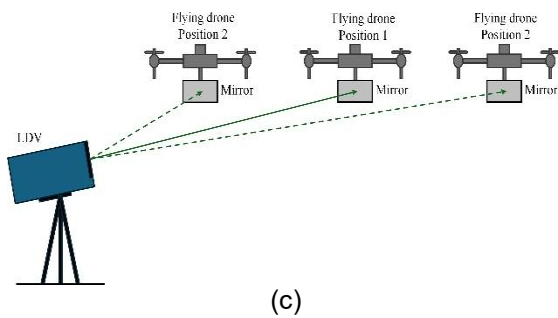


Fig. 1: (a) LDV arrangement for SHM without Flyable Mirror, (b) Introduction of Flyable Mirror to LDV arrangement, (c) Application of laser tracking in Flyable Mirror.

Several LDV configurations are available for vibration research. The choice of configuration strongly influences the design of an experiment. A conventional single-point LDV employs a stationary laser beam and measures vibration at a fixed point along the laser beam direction. A scanning LDV (SLDV) measures structural vibration at discrete points by steering the laser beam sequentially to predefined locations using two orthogonal scanning mirrors. Continuous beam motion along a prescribed path during measurement defines the continuously scanning LDV (CSLDV) mode. Active control of the measurement beam to follow the same material point on a moving object allows SLDV or CSLDV systems to operate in tracking mode [7]. In this paper, the feasibility of SLDV-based tracking is investigated through experiments on the translational motion of a linear table along the x -axis.

Conventional methods of laser tracking utilize two galvanoscanners to steer the laser beam along the x - and y -axes. Early implementations have relied on encoder-assisted synchronization and prior knowledge of the target trajectory to drive the scanning mirrors. Tracking accuracy has been evaluated in terms of the positional error between the laser beam and the target measurement point. Based on this approach, CSLDV tracking methods have been applied to rotating disks [8] and to a fan during coast-down [9] by measuring the rotational position with an encoder and synchronizing the scanning process with the rotation. In both studies, the galvanometric mirrors were driven by sinusoidal signals in the x - and y -directions generated by a digital-to-analog converter (DAC) board [8, 9]. Encoder-synchronized tracking using predefined mirror-drive waveforms was demonstrated in [10].

In [11], an encoder-based tracking system with rotor-angle feedback has been demonstrated, achieving an accuracy of ± 0.5 mm when tracking a selected point on a rotating bladed

disc. For translational motion tracking, an encoder-based laser-tracking system has been applied to measure vibrations of a pulley-driven timing belt running at approximately 4.8–5.0 m/s [12]. The target position measured by the encoder has been used to generate the set-points for two linear proportional–integral–derivative (PID) controllers that drive the scanning mirrors [11, 12].

Laser tracking without prior knowledge of the moving target has been implemented using image processing based on camera feedback. An image-based closed-loop laser tracking system has been introduced in [13], in which a charge-coupled device (CCD) camera has been employed to detect the centroid of the retroreflective tape attached to a windscreen wiper. A predictor-corrector approach has been employed to compensate for the dead time introduced by the image acquisition and processing. The study has reported a tracking accuracy of approximately ± 1.2 mm. In [14], a color-based landmark has been used to establish tracking using an SLDV. A proportional-integral (PI) control strategy has been applied to drive the mirrors. During experiments on a rotating wheel at a rotational frequency of 1\ Hz, the reported peak-to-peak error ranges from 1.8 mm to 5.2 mm. However, image-based approaches are sensitive to optical conditions. In [13], unintended bright or reflective features within the field of view of the camera can disturb the detection of the retroreflective marker. In [14], the high-intensity laser beam can alter the apparent color of the landmark when the measurement point is too close, which can compromise tracking accuracy and robustness.

This work proposes a real-time vision-guided laser tracking system that does not require encoder feedback or predefined trajectories of the moving target [12–14]. The approach is therefore suitable for objects with unpredictable motion and does not require additional sensors, since visual feedback alone is sufficient for tracking. An ArUco marker from the 4x4_50 dictionary with ID 0 is used as a landmark because its unique encoded identifier enables reliable and unambiguous detection. Consequently, the tracking process becomes insensitive to variations in laser beam intensity or spot location, as well as to the presence of other bright or similar-looking objects within the camera field of view.

This paper is organised as follows. Section 2 describes the main components used in this experiment. Section 3 presents the structure of the system. Section 4 reports the tracking results. Finally, section 5 draws a conclusion and outlines the future work of the experiment.

2. Components of the Laser Tracking System

The main component of the proposed laser tracking system is the Optomet SLDV, shown in Figure 2. The SLDV provides the measurement laser beam and enables beam steering through its internal scanning mechanism.

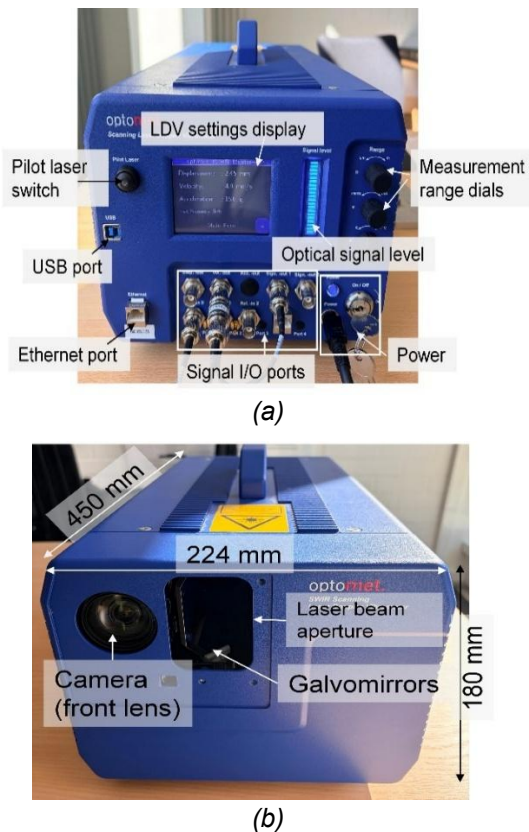


Fig. 2: Optomet SLDV (a) rear view (b) front view

The SLDV consists of two independent communication subsystems. First, an Ethernet-based subsystem provides access to the internal signal processing electronics, allowing access to the configuration of the measurement settings and transmission of the acquired measurement data. Second, a USB-based subsystem is used to interface with the peripheral components of the setup, specifically the camera, the mirror control systems, and the data acquisition unit. In this study, the USB-based subsystem has been used [15].

Figure 2(a) shows the rear view of the SLDV. It contains a touch screen display for adjusting configurations, an LED bar indicator for monitoring the optical signal level during laser focusing, and dials for selecting the displacement, velocity, and acceleration measurement ranges. The main electrical interfaces, including USB and Ethernet ports for access to the subsystems, analog signal input/output ports, a power socket, and the

on/off switch of the pilot laser, are also located in the rear panel of the SLDV.

The front view of the SLDV in Figure 2(b) shows the front lens of the integrated camera and the laser beam aperture, behind which the galvo mirrors for beam steering are located. In this paper, the integrated camera is used to provide visual feedback. It is a Complementary Metal-Oxide-Semiconductor (CMOS)-based zoom camera with full HD resolution. The camera resolution can be configured and for this experiment HD resolution has been maintained. A linear table, shown in Fig. 3, has been used to generate motion. It provides translational motion at a controllable speed and is therefore suitable for evaluating the tracking performance of the proposed system. The linear table actuator has a travel length of 400 mm and is driven by a NEMA 23 stepper motor with a screw pitch of 5 mm.

A unique landmark is required for the visual reference of the vision-guided system. A fiducial ArUco marker has been used, as shown in Figure 3. Further details of the landmark are provided in Section 3.

3. Structure of the system

The integrated camera of the SLDV provides image feedback at a frame rate of 60 fps. It continuously observes the landmark attached to the moving target and provides its position for tracking. The laser beam is steered by two galvanometer mirrors in the horizontal (x -axis) and vertical (y -axis) direction. As a visual landmark, a 4×4 , 50 mm^2 , ID-0 ArUco marker is mounted on the target. It enables reliable and unique detection of the target position. To generate controlled motion of the target, the landmark is mounted on the linear table actuator, which serves as the moving platform during the experiments. The overall configuration of the proposed system is illustrated in Figure 3.

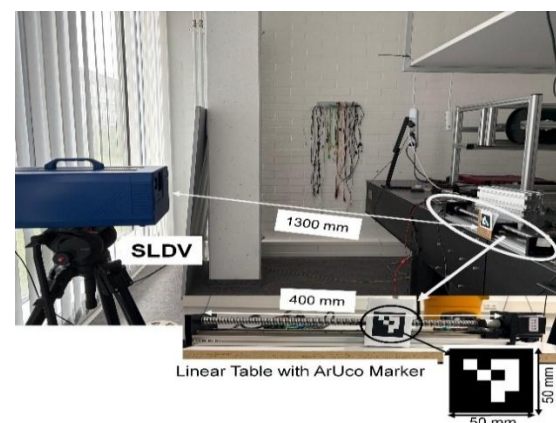


Fig. 3: Tracking system setup with the linear table and the ArUco marker.

The overall tracking integrates the following two subsystems:

3.1: image processing using computer vision

During the tracking experiments, the camera is operated in manual exposure mode to detect both the marker and the laser-beam position. For image processing, each captured image frame is first converted from blue-green-red (BGR) format to grayscale in OpenCV. A predefined ArUco dictionary provided by OpenCV is used to detect and identify the marker in the camera frame from the grayscale image. The marker centroid is used as the reference position. After detection, the four corner coordinates of the marker are extracted, and the center is computed as the mean of the corner positions. The overall image-processing procedure for ArUco marker detection is shown in Figure 4.

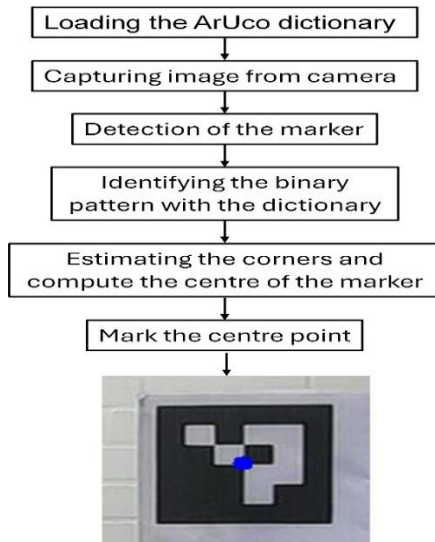


Fig. 4: Marker detection scheme

In addition, detection of the laser beam is required to verify the tracking result by comparing the laser-beam position with the marker centroid in real time. In this study, the laser-beam position is determined in each image frame by identifying the brightest pixel, while the camera exposure has been maintained at a low level to improve spot visibility. The camera is controlled via a serial interface using the VISCA protocol [16].

3.2. Mirror control

The galvo mirrors of the SLDV are precisely controlled to enable real-time tracking of the marker centroid during motion. The SLDV uses a Measurement Computing (MCC) Data Acquisition (DAQ) card to provide control output voltages to both the x -axis and y -axis mirrors within a range of ± 10 V.

Let the detected marker center in the camera image at frame k ($k = 1, 2, 3, \dots$) be,

$$\vec{P}_m(k) = \begin{bmatrix} x_m(k) \\ y_m(k) \end{bmatrix},$$

where, $x_m(k)$ and $y_m(k)$ are the pixel coordinates of the ArUco marker centroid.

If (x_c, y_c) denotes the coordinate of the image center, then,

$$\vec{P}_c = \begin{bmatrix} x_c \\ y_c \end{bmatrix} = \begin{bmatrix} W/2 \\ H/2 \end{bmatrix},$$

where, $W = 1280$ pixels and $H = 720$ pixels are the image width and height, respectively.

The ArUco marker is attached to the linear table as shown in Figure 3, so that it follows the translational motion of the table directly.

During the experiment, the displacement between the center of the detected marker and the center of the image frame is taken as the tracking error for the vision-guided control loop. The tracking error in the image plane is defined as,

$$e(k) = \vec{P}_m(k) - \vec{P}_c = \begin{bmatrix} x_m(k) - x_c \\ y_m(k) - y_c \end{bmatrix}.$$

This error is calculated in pixels. To make it independent of image resolution, the error is normalized as

$$e_x'(k) = \frac{x_m(k) - x_c}{W/2}$$

$$\text{and, } e_y'(k) = \frac{y_m(k) - y_c}{H/2},$$

Thus, e_x' and e_y' lie in the range $[-1, 1]$.

The control voltages applied to the two galvanometer mirrors are generated using proportional control:

$$V_x(k) = K_x e_x'(k)$$

$$\text{and, } V_y(k) = K_y e_y'(k).$$

Here, K_x and K_y are proportional gains where $K_x = K_y = 7$. Therefore, a proportional control strategy is implemented. Figure 5 illustrates the block diagram of the control scheme. The computed mirror-control voltages are clipped to the allowable DAQ range and subsequently converted into digital DAC commands for analog output to the scanner mirrors.

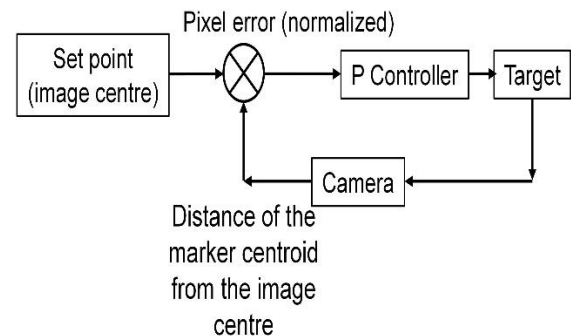


Fig. 5: The proportional control scheme.

4. Tracking result

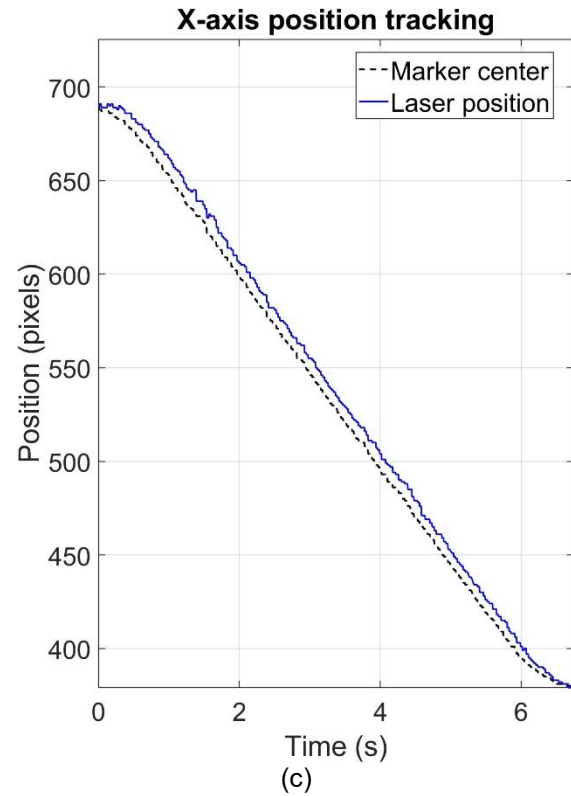
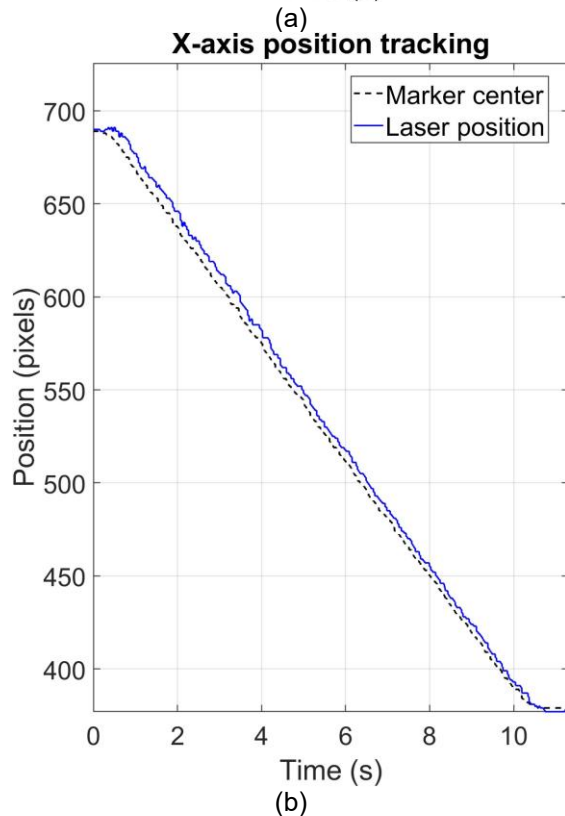
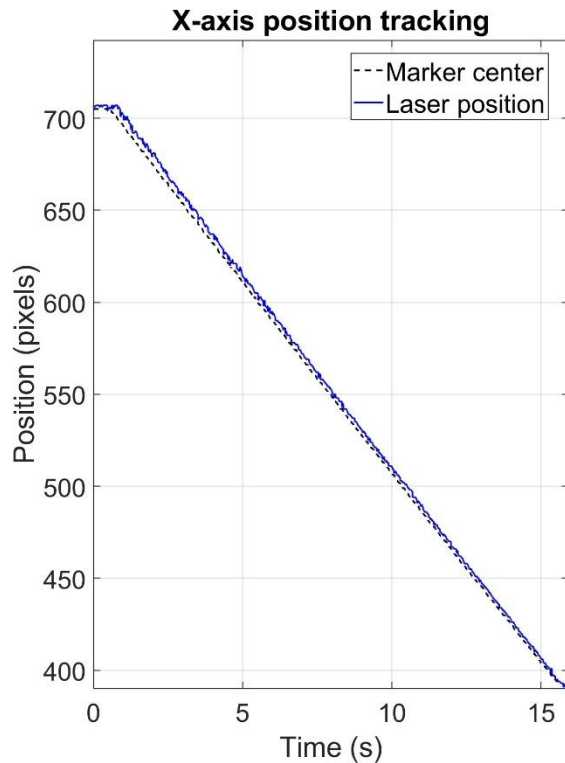


Fig. 6: Tracking performance at different translational speeds of motion - (a) 24 mm/s, (b) 36 mm/s, (c) 60 mm/s

Figure 6 illustrates the tracking performance during translational motion. The dashed and the solid lines represent the position of the center of the marker and the measured laser position, respectively, in pixels. The vertical axis in Figure 6 represents the position of the marker and the laser. The horizontal axis represents time in seconds in Figure 6. A small deviation between the target and actual laser position is visible due to the combined effect of image processing delay and mirror actuation dynamics. This deviation is defined as the tracking error. The root mean square (RMS) error is computed from the difference between the laser position and the marker position at each recorded sample during tracking. The error samples are squared to prevent positive and negative deviations from canceling each other and to assign greater weight to larger deviations. The mean of the squares is then calculated, and the square root of this mean is taken to obtain the RMS error. In the present setup, 1 pixel in the image frame corresponds to 1.2 mm in real-world distance. The marker and laser positions are recorded in integer pixels. Therefore, the tracking error is evaluated at pixel resolution.

The tracking performance is first evaluated at a linear table speed of 24 mm/s. Under this condition, the maximum tracking error is 8

pixels (9.6 mm), while the RMS error is 3.56 pixels (4.27 mm), as shown in Fig. 6(a). The experiment is then repeated at 36 mm/s, where the maximum tracking error increases to 10 pixels (12 mm) and the RMS error to 5.71 pixels (6.85 mm), as presented in Fig. 6(b). Finally, at 60 mm/s, the maximum error further increases to 13 pixels (15.6 mm), with an RMS error of 7.73 pixels (9.28 mm), as shown in Fig. 6(c). These results are obtained from the X-axis tracking data.

The summary of the tracking performance is given in Table 1.

Tab. 1: Summary of the tracking performance at different translational speed of motion: 24 mm/s, 36 mm/s and 60 mm/s

	Translational Motion		
	24 mm/s	36 mm/s	60 mm/s
Maximum tracking error, in mm	9.6	12	15.6
RMS error, in mm	4.27	6.85	9.28

5. Conclusion and Future Work

This paper has presented a vision-based laser tracking system using an SLDV, a proportional control feedback loop based on camera images, and an ArUco fiducial marker as a landmark and motorized a linear table to generate translational motion. The proportional control strategy has been successfully implemented to drive the galvo mirrors of the SLDV for laser beam steering. The result of the demonstration shows that the laser beam can follow the moving target with a small positional error. However, this approach relies on the clear and complete visibility of the landmark (ArUco marker). As a result, any short occlusion of the marker can interrupt the tracking process, which is not suitable for applications that require uninterrupted measurement with laser tracking under dynamic conditions. Future work of this study will therefore focus on the development of an advanced control strategy integrated with predictive methods to improve the accuracy of tracking at higher target speeds with enhanced robustness against temporary landmark occlusion.

References

- [1] S. J. Rothberg, M. S. Allen, P. Castellini, D. Di Maio, J. J. Dirckx, D. J. Ewins, et al., *Optics and Lasers in Engineering* 99, 11–22 (2017); doi: 10.1016/j.optlaseng.2016.10.023
- [2] P. Castellini, M. Martarelli, E. P. Tomasini, *Mechanical Systems and Signal Processing* 20, 1265–1285 (2006); doi: 10.1016/j.ymssp.2005.11.015
- [3] C. Rembe, B. J. Halkon, M. A. Ismail, *Advanced Devices & Instrumentation* 6, 0103 (2025); doi: 10.34133/adi.0103

- [4] M. Schewe, M. A. A. Ismail, C. Rembe, *Insight - Non-Destructive Testing and Condition Monitoring* 63, 280–282 (2021); doi: 10.1784/insi.2021.63.5.280
- [5] M. Schewe, M. A. Ismail, R. Zimmermann, U. Durak, C. Rembe, in: *Journal of Physics: Conference Series* 2698, 012007 (2024); doi: 10.1088/1742-6596/2698/1/012007
- [6] S. Hassani, U. Dackermann, *Sensors* 23, 2204 (2023); doi: 10.3390/s23042204
- [7] P. Chiariotti, C. Rembe, P. Castellini, M. Allen, in: R. Allemang, P. Avitabile (eds.), *Handbook of Experimental Structural Dynamics*, 1–45 (2021); doi: 10.1007/978-1-4939-6503-8_4-2
- [8] D. Di Maio, D. J. Ewins, *Mechanical Systems and Signal Processing* 24, 3013–3036 (2010); doi: 10.1016/j.ymssp.2010.06.012
- [9] M. Martarelli, P. Castellini, C. Santolini, E. P. Tomasini, *Measurement Science and Technology* 22, 115106 (2011); doi: 10.1088/0957-0233/22/11/115106
- [10] B. J. Halkon, S. J. Rothberg, in: *Proceedings of SPIE* 5503, 260–271 (2004); doi: 10.1117/12.579759
- [11] P. Castellini, N. Paone, *Review of Scientific Instruments* 71, 4639–4647 (2000); doi: 10.1063/1.1319862
- [12] P. Castellini, E. Cupido, N. Paone, E. P. Tomasini, in: *Proceedings of SPIE* 4072, 194–200 (2000); doi: 10.1117/12.386758
- [13] P. Castellini, E. P. Tomasini, *Review of Scientific Instruments* 75, 222–232 (2004); doi: 10.1063/1.1630859
- [14] D. Kim, H. Khalil, J. Nam, K. Park, *International Journal of Precision Engineering and Manufacturing* 16, 1717–1721 (2015); doi: 10.1007/s12541-015-0225-8
- [15] Optomet GmbH, *Scanning Laser Doppler Vibrometer*, <https://www.optomet.com/de/produkte/classic-serie/scanning-vibrometer/>
- [16] Sony, FCB-EV7520A Technical Manual, URL: <https://www.activesilicon.com/wp-content/uploads/FCB-EV7520A-Technical-Manual.pdf> (Accessed on 16.03.2026).

Robuste Befestigung von Ultraschallwandlern: Prüfstrategien für kryogene Einsatzbedingungen

Daniel Pollmann¹, Alexander Backer¹, Klaus Stefan Drese¹

¹Institut für Sensor und Aktortechnik (ISAT), Hochschule für angewandte Wissenschaften Coburg
Am Hofbräuhaus 1b, 96450 Coburg, Deutschland

Daniel.Pollmann@hs-coburg.de, Alexander.Backer@hs-coburg.de, Klaus.Drese@hs-coburg.de

Zusammenfassung

Für ultraschallbasierte Sensoren in kryogenen Anwendungen ist eine stabile akustische Ankopplung der piezoelektrischen Wandler an metallische Strukturen entscheidend. Temperaturänderungen zwischen Raumtemperatur und kryogenen Bedingungen führen jedoch zu erheblichen thermomechanischen Spannungen, die das Schwingungs- und Übertragungsverhalten der Wandler beeinflussen können, ohne dass es zu einem vollständigen Versagen der Verbindung kommt. In diesem Beitrag werden zwei komplementäre Messmethoden zur Bewertung der Temperaturänderungsbeständigkeit von Ultraschallankopplungen vorgestellt: die elektrische Impedanzmessung und die temperaturabhängige Ultraschallamplitudenmessung. Der Schwerpunkt liegt auf der Ultraschallamplitudenmessung als funktionsnaher Prüfmethode. Es wird gezeigt, dass Änderungen der Ankopplung anhand charakteristischer Amplitudenverläufe während Abkühl- und Aufwärmzyklen detektiert werden können. Unterschiede in den Materialeigenschaften der Koppelschichten sowie fertigungstechnische Einflüsse spiegeln sich deutlich in frequenz- und temperaturabhängigen Amplitudenverläufen wider. Die Ergebnisse verdeutlichen, dass Vergleiche bei einzelnen Frequenzen nicht ausreichend sind und eine breitbandige Betrachtung erforderlich ist, um belastbare Aussagen über die Stabilität der Ankopplung zu treffen.

Keywords: Kryosensorik, Ultraschallankopplung, elektrische Impedanz, Ultraschallamplitude, Prüfmethoden

Einleitung

Ultraschallbasierte Messverfahren gelten als vielversprechender Ansatz für die Sensorik in kryogenen Anwendungen, etwa zur Durchflussmessung verflüssigter Gase. Ihr Einsatz unter diesen Randbedingungen ist jedoch weiterhin Gegenstand aktueller Forschung, da die tiefen Temperaturen hohe Anforderungen an Wandler, Ankopplung und Signalstabilität stellen [1,2]. Ein wesentlicher Bestandteil solcher Sensoren ist die akustische Ankopplung piezoelektrischer Ultraschallwandler an metallische Strukturen, beispielsweise Edelstahlrohre. Die Zuverlässigkeit der Messung hängt dabei maßgeblich von der Stabilität dieser Ankopplung ab.

Bei Temperaturänderungen von Raumtemperatur bis hin zu kryogenen Bedingungen treten aufgrund unterschiedlicher thermischer Ausdehnungskoeffizienten der beteiligten Materialien hohe mechanische Spannungen auf. [3] Diese können zu Veränderungen der Koppelschicht führen, die sich unmittelbar auf das

Schwingungsverhalten des Wandlers und damit auf die akustische Signalübertragung auswirken. Für die Bewertung der Eignung von Ankoppelverfahren sind daher Messmethoden erforderlich, die solche Änderungen sensitiv und möglichst funktionsnah erfassen.

Ziel dieses Beitrags ist die Vorstellung und Einordnung zweier Messmethoden zur Bewertung der Temperaturänderungsbeständigkeit von Ultraschallankopplungen. Während die elektrische Impedanzmessung einen Überblick über Änderungen des gekoppelten Schwingungsverhaltens liefert, steht die Ultraschallamplitudenmessung während thermischer Zyklen im Fokus als zentrale Prüfmethode mit direktem Bezug zur sensorischen Funktion.

Versuchsaufbau

Für die Entwicklung der Prüfmethoden wurden Prüflinge mit geklebten oder gelöteten Piezokeramiken, wie in Abbildung 1 dargestellt, verwendet.

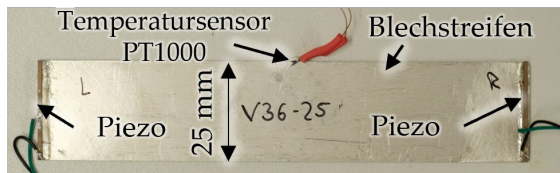


Abb. 1: Aufbau einer 125 mm langen Laufstrecke für Ultraschall-amplitudenmessungen mit 2 Piezos (Sender und Empfänger) und einem Temperatursensor Pt1000 auf einem Edelstahlblech vom Typ 1.4571 mit einer Dicke von 2 mm.

In Abbildung 2 ist die Vorrichtung für die Durchführung der Abkühlung mittels LN_2 dargestellt. Mit der Kranvorrichtung wurden die Prüflinge in der Gasphase über flüssigem Stickstoff (LN_2) auf bis zu -196°C abgekühlt. Vor, während und nach dem Abkühlvorgang wurden die Messungen für die Prüfmethode durchgeführt.

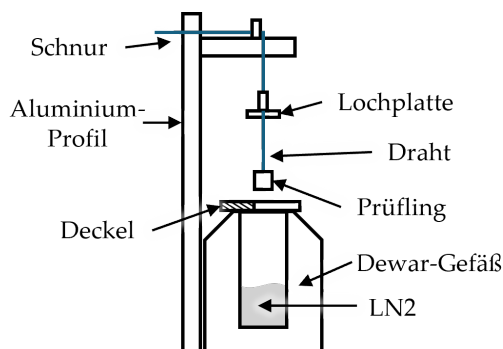


Abb. 2: Versuchsaufbau für Abkühl- und Eintauchversuche mit LN_2

Elektrische Impedanzmessung zur Bewertung der Ankopplung

Die elektrische Impedanz eines piezoelektrischen Ultraschallwandlers ist eng mit dessen mechanischem Schwingungsverhalten gekoppelt [4]. Änderungen der Randbedingungen, insbesondere der akustischen Ankopplung an ein Substrat, führen zu charakteristischen Veränderungen im Impedanzverlauf über der Frequenz.

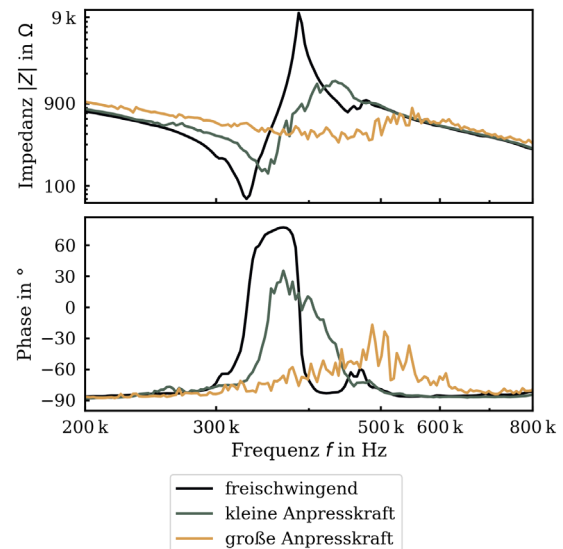


Abb. 3: Bode-Diagramm der Impedanzmessung einer Piezokeramik in freischwingendem und angepresstem Zustand.

Abbildung 3 zeigt exemplarisch den Einfluss unterschiedlicher Ankoppelzustände auf den Impedanzverlauf eines Ultraschallwandlers. Eine feste und homogene Ankopplung führt zu einer deutlichen Verschiebung und Dämpfung der Resonanzstrukturen im Vergleich zu einem nur losen oder unvollständig angekoppelten Wandler. Diese Veränderungen betreffen nicht nur einzelne Resonanzfrequenzen, sondern den gesamten betrachteten Frequenzbereich. Damit eignet sich der Impedanzverlauf als empfindlicher Indikator für Änderungen der Ankopplung, insbesondere auf Ebene einzelner Ultraschallwandler.

Einen Überblick über die Auswirkungen thermischer Zyklen auf unterschiedliche Ankoppelverfahren liefert Abbildung 4. Dargestellt sind Impedanzverläufe vor und nach dem Abkühlen auf kryogene Temperaturen für verschiedene herstellungsbedingte Ausprägungen der Ankopplung.

Die Abbildung verdeutlicht, dass sich Änderungen der Ankopplung nicht in jedem Fall als vollständiges Versagen äußern. Vielmehr treten häufig subtile, aber reproduzierbare Veränderungen der Resonanzlagen, der Resonanzstärken oder der Dämpfung auf. Diese Effekte unterscheiden sich deutlich in Abhängigkeit vom verwendeten Ankoppelverfahren und den damit verbundenen Materialeigenschaften der Koppelschicht. Die Impedanzmessung erlaubt somit eine qualitative Bewertung der Stabilität der Ankopplung über einen größeren Frequenzbereich und eignet sich insbesondere

zur Analyse einzelner Wandler sowie zum Vergleich verschiedener Ankoppelkonzepte.

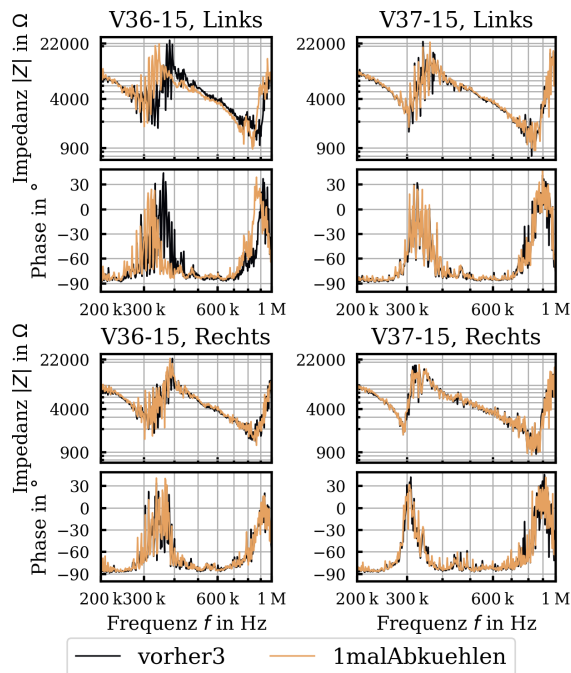


Abb. 4: V36-löten, V37-kleben Bode-Diagramm der elektrischen Impedanzverläufe der Piezokeramiken auf den Laufstrecken V36-15 und V37-15. Die Verläufe wurden vor und nach dem ersten Abkühlen gemessen.

Ultraschallamplitudenmessung als funktionale Prüfmethode

Zur Bewertung der akustischen Übertragung wird die Ultraschallamplitude als zentrales Maß herangezogen. Als Amplitudenkennwert wird $U(t_{A0})$ verwendet, definiert als der Wert der Einhüllenden des empfangenen Ultraschallsignals zum Zeitpunkt t_{A0} . Der Zeitpunkt t_{A0} entspricht der erwarteten Ankunftszeit des Nutzsignals und wird aus der bekannten Laufstrecke und der Schallausbreitung bestimmt.

Die Bestimmung von $U(t_{A0})$ erfolgt, indem das empfangene zeitabhängige Signal zunächst bandpassgefiltert und anschließend mittels Hüllkurvenbildung ausgewertet wird. Abbildung 5 zeigt die so bestimmten Auswertepunkte einiger Beispielsignale bei unterschiedlichen Frequenzen und Temperaturen. Der so gewonnene Amplitudenwert stellt ein robustes Maß für die akustische Kopplung zwischen Sender und Empfänger dar und ist direkt mit der Effizienz der Schallübertragung verknüpft.

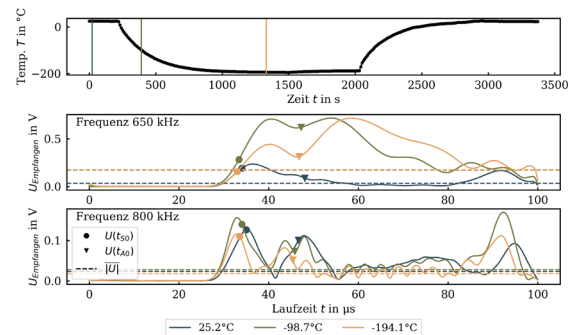


Abb. 5: Hüllkurven der Empfangssignale bei verschiedenen Temperaturen und Frequenzen, sowie Darstellung der Maße für die Amplitude

Ein Beispiel für einen auffälligen Amplitudenverlauf während eines thermischen Zyklus ist in Abbildung 6 dargestellt. Beim Abkühlen und anschließenden Aufwärmen zeigt sich ein deutlicher Sprung im Amplitudenverlauf. Zudem unterscheiden sich die Amplitudenwerte vor und nach dem thermischen Zyklus signifikant. Beide Beobachtungen sprechen für eine bleibende Änderung der Ankopplung.

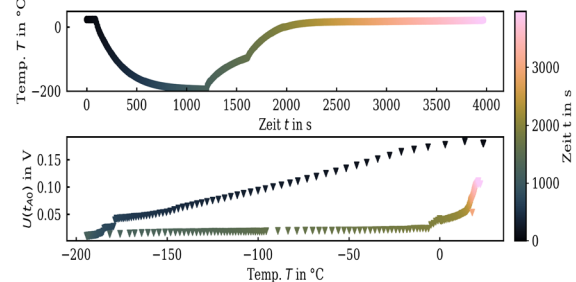


Abb. 6: Amplitudenverlauf in Abhängigkeit der Temperatur von V36-15-löten 1. mal Abkühlen bei 380kHz

Wesentlich ist, dass solche Änderungen auch dann erkannt werden können, wenn die Amplitude vor und nach dem Zyklus auf ähnlichem Niveau liegt. Ein sprunghafter Verlauf während des Abkühlens oder Aufwärmens weist auf eine Änderung der Schwingungs- und Kopplungseigenschaften hin, die allein durch einen Vorher-Nachher-Vergleich nicht zuverlässig erfasst werden. Die Betrachtung des vollständigen Temperaturverlaufs liefert daher einen entscheidenden Erkenntnisgewinn.

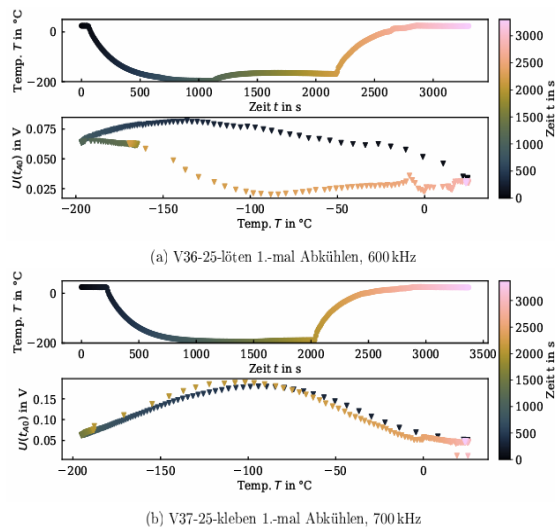


Abb. 7: Vergleich der Amplitudenverläufe von V36-25-löten und V37-25-kleben beim Abkühlen und Aufwärmen.

Abbildung 7 zeigt exemplarisch die Amplitudenverläufe für unterschiedliche Ankoppelverfahren während eines thermischen Zyklus. Dabei wird deutlich, dass sich geklebte und gelötete Ankopplungen in ihrem temperaturabhängigen Verhalten unterscheiden. Diese Unterschiede äußern sich in der Form der Amplitudenverläufe des Abkühlens und Aufwärmens, die bei den gelöteten Verbindungen voneinander abweichen.

Die beobachteten Effekte lassen sich auf die unterschiedlichen Materialeigenschaften der Koppelschichten zurückführen, etwa auf deren Elastizität, Dämpfung und Fähigkeit, thermisch induzierte Relativdehnungen auszugleichen. Die Ultraschallamplitudenmessung ermöglicht es somit, materialbedingte Unterschiede der Ankopplung direkt sichtbar zu machen.

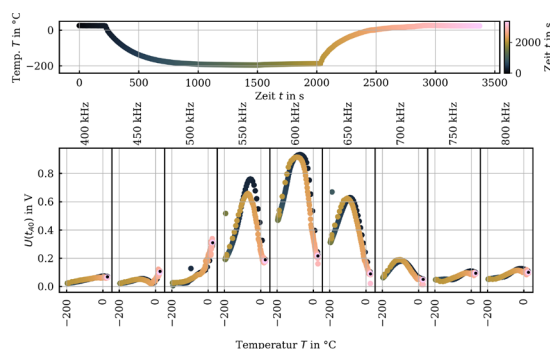


Abb. 8: Gesamtübersicht der Amplitudenverläufe in Abhängigkeit der Temperatur bei den gemessenen Frequenzen von V37-25-kleben.

Der Gesamtzusammenhang zwischen Frequenz, Temperatur und Ankoppelverfahren ist in Abbildung 8 dargestellt. Die Abbildung zeigt, dass die Amplitudenverläufe stark frequenzabhängig sind. Ursache hierfür sind Änderungen der Schwingeigenschaften des Gesamtsystems aus Wandler, Koppelschicht und Substrat, die sowohl durch die Materialeigenschaften als auch durch Fertigungstoleranzen beeinflusst werden. Diese ausgeprägte Frequenzabhängigkeit hat eine wichtige methodische Konsequenz: Vergleiche von Ultraschallamplituden bei nur einer einzelnen Frequenz sind nicht aussagekräftig. Eine belastbare Bewertung der Temperaturänderungsbeständigkeit von Ankoppelverfahren erfordert vielmehr die Betrachtung eines breiten Frequenzbereichs. Erst dadurch lassen sich systematische Unterschiede zwischen verschiedenen Ankoppelkonzepten identifizieren und zufällige Effekte durch lokale Resonanzverschiebungen erkennen.

Diskussion

Die Ergebnisse zeigen, dass die Ultraschallamplitudenmessung eine besonders geeignete Prüfmethode zur Bewertung der Temperaturänderungsbeständigkeit von Ultraschallankopplungen darstellt. Sie erlaubt die Detektion von Änderungen der Ankopplung sowohl durch Sprünge im Amplitudenverlauf während thermischer Zyklen als auch durch Unterschiede zwischen Zuständen vor und nach der Beanspruchung. Dabei ist hervorzuheben, dass selbst dann Änderungen erkannt werden können, wenn die Amplitudenwerte vor und nach dem Zyklus ähnlich sind.

Die elektrische Impedanzmessung ergänzt diesen Ansatz sinnvoll, indem sie einen Überblick über Änderungen des gekoppelten Schwingungsverhaltens über einen großen Frequenzbereich liefert. Sie eignet sich insbesondere zur Analyse einzelner Ultraschallwandler und zur qualitativen Bewertung unterschiedlicher Ankoppelzustände. Im Vergleich zur Amplitudenmessung ist sie jedoch weniger direkt an die spätere Sensorfunktion gekoppelt. Insgesamt zeigt sich, dass erst die Kombination beider Messmethoden eine umfassende Bewertung der Ankoppelstabilität ermöglicht. Während die Impedanzmessung strukturelle Änderungen des Wandlerverhaltens sichtbar macht, liefert die Amplitudenmessung den funktionsnahen Nachweis ihrer Relevanz für die akustische Signalübertragung.

Fazit

In diesem Beitrag wurden Messmethoden zur Bewertung der Temperaturänderungsbeständigkeit von piezoelektrischen Ultraschallankopplungen unter kryogenen Bedingungen vorgestellt. Der Schwerpunkt lag auf der Ultraschallamplitudenmessung als funktionsnaher Prüfmethode. Es konnte gezeigt werden, dass Änderungen der Ankopplung anhand charakteristischer Amplitudenverläufe während thermischer Zyklen zuverlässig detektiert werden können. Unterschiede in den Materialeigenschaften der Koppelschichten sowie fertigungstechnische Einflüsse spiegeln sich deutlich in der Frequenz- und Temperaturabhängigkeit der Amplituden wider.

Die Ergebnisse unterstreichen, dass eine breitbandige Betrachtung über Frequenz und Temperatur notwendig ist, um belastbare Aussagen über die Stabilität von Ankoppelverfahren zu treffen. Die vorgestellten Methoden bieten damit ein geeignetes Instrumentarium für die Entwicklung und Qualifikation robuster Ultraschallsensoren für kryogene Anwendungen.

Literaturnachweis

- [1] ISAT - ISAT Institut für Sensor- und Aktortechnik. HydrAmess: Akustische Sensorik zur Durchflussmessung von flüssigem Wasserstoff und anderen Energieträgern. url: <https://www.isat-coburg.de/projekt/hydrames-akustische-sensorik-zur-durchflussmessung-von-fluessigem-wasserstoff-und-anderen-fluessigenergetraegern/> (besucht am 24.08.2025).
- [2] Daniel A. Kiefer. Elastodynamic quasi-guided waves for transit-time ultrasonic flow metering. ISBN: 9783961475506. Diss. Erlangen-Nürnberg: Friedrich Alexander-Universität (FAU), 2022. url: <https://open.fau.de/handle/openfau/19767> (besucht am 25.08.2025).
- [3] Jan Dolkemeyer und Peter Loosen. Aufbau- und Löttechniken für die Montage von Festkörperlasern. Number: RWTH-CONV-124363. Diss. Publikationsserver der RWTHAachenUniversity, 2012. url: <https://publications.rwth-aachen.de/record/62868> (besucht am 08.10.2024)
- [4] Daniel Schmidt. Modenselektive Übertragung von Lambwellen in Faserverbundstrukturen. ISSN: 1434-8454 Publisher: Deutsches Zentrum für Luft- und Raumfahrt e.V. Berichtreihe. Technischen Universität Carolo-Wilhelmina zu Braun schweig, 2014. url: <https://elib.dlr.de/90214/> (besucht am 15.07.2025)

Integrated photonic and electronic building blocks for NV-diamond magnetometry

Lena Schäd¹, Steffen Epple¹, Tobias Hoose², Mathias Kaschel¹, Kevin Edelmann¹, Gregor Ronniger¹, Michael Thiel², Robert Rölver, Michal Kern³, Jens Anders^{1,3}

¹ *Institut für Mikroelektronik Stuttgart, 70569 Stuttgart, Germany*

² *Nanoscribe GmbH & Co. KG, 76344 Eggenstein-Leopoldshafen, Germany*

³ *Institute of Smart Sensors, University Stuttgart, 70569 Stuttgart, Germany*

Abstract

We present key building blocks for an integrated approach to NV-diamond magnetometry. The work combines application-specific integrated circuits (ASICs) for control and data processing with integrated optical waveguides within a compact architecture. The objective is to investigate miniaturized and scalable implementation strategies for diamond-based NV magnetic field sensors while maintaining suitable optical excitation conditions and compatibility with optical readout. Within a photonic integrated circuit (PIC), both the excitation path and the signal and reference paths were realized on a single chip. Two different photonic chip architecture concepts were implemented and characterized with respect to optical throughput and assembly complexity. Experimental results show that integrated Si₃N₄ waveguides, in combination with additively manufactured micro-optics, can provide efficient diamond excitation with power densities exceeding 1 kW/cm². In addition, preliminary feasibility experiments indicate that the presented approach is compatible with optically detected magnetic resonance (ODMR) measurements. The presented results provide a basis for future integrated NV-based sensing systems.

Keywords: magnetic field sensor, optical measurement methods, NV-diamond magnetometry, integrated photonics

Introduction

Applications in materials research and medical diagnostics demand high-performance magnetic field sensors. There is a particular need for sensors that can operate at room temperature, exhibit high magnetic field sensitivity in the picotesla range, and allow for significant miniaturization to enable integration into compact or portable systems. These requirements are becoming increasingly important, particularly in the context of human-machine interfaces in mobile diagnostic systems. Quantum sensors based on nitrogen-vacancy (NV) defects in diamond represent a promising technology for meeting these requirements [1]. NV color centers are point defects in the diamond lattice, consisting of a nitrogen atom in the immediate vicinity of a lattice vacancy. The ground state of this defect possesses an electron spin that can be optically detected and manipulated using microwave fields. This principle is utilized in the context of optically detected magnetic resonance (ODMR) [2]. A schematic image of the energy transitions is shown in Fig. 1 (a). If the spin state $m_s = 0$ is excited to the states $m_s = \pm 1$ by a resonant microwave field (2.87 GHz), the observed fluorescence intensity, indicated by the red arrow, is reduced due

to an increased transition to a non-radiative singlet state. The resulting dip in the fluorescence signal as a function of microwave frequency is shown in Fig. 1 (b). In the presence of an external magnetic field, the states $m_s = \pm 1$ are energetically split by the Zeeman effect, causing the characteristic fluorescence dips in the ODMR spectrum to split as well. The positions of these resonances depend directly on the strength of the external magnetic field, enabling precise determination of changes in the magnetic field. Due to the long coherence times of the spins, the high radiative quantum efficiency, and the optical accessibility of the NV centers, physical quantities, particularly magnetic fields, can be measured with very high sensitivity at room temperature [1]. NV-based quantum sensors thus offer a decisive advantage over other quantum sensors, such as optically pumped magnetometers (OPMs) and superconducting quantum interference devices (SQUID), which often have more restrictive operating conditions, such as shielding or cryogenic requirement [1]. For many applications, however, extensive miniaturization of the sensor system is required in addition to sensitivity. Whilst recent work has already integrated individual sub-components, a fully integrated system solution has so far been lacking [3], [4].

In particular, the integration of the various functional components like laser sources, optical filters, waveguides, microwave generators and resonators, as well as photodetectors, together with the necessary control and readout electronics presents a key challenge. At the same time, low-loss optical signal transmission and stable interfaces between the individual components must be ensured. Many current publications on miniaturized, portable NV-magnetometers continue to be based on free space optics. Whilst this allows sensor heads to be realized on a centimeter scale [3], [4], scalable and reproducible manufacturing of such systems is only possible to a limited extent. In this context, approaches toward photonic and electronic integration are of particular interest, as they may enable more compact and scalable implementations. However, achieving low-loss optical routing, stable interfaces, and efficient excitation within such integrated systems remains a key challenge.

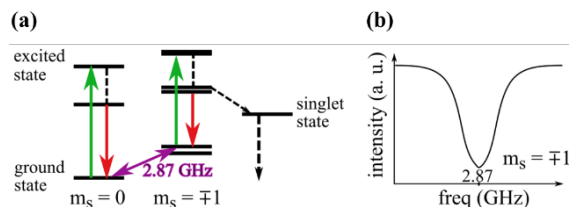


Fig. 1: (a) Simplified energy levels of the electron spin in the NV center. (b) Optical pumped and zero biased ODMR spectrum of the NV center according to microwave frequency. The graphic is based on [2].

We investigate a miniaturized approach based on integrated photonics and integrated electronics for NV-diamond magnetometry. Due to its broad transmission window of 400 nm to 4000 nm, Si_3N_4 is a suitable integrated photonic platform for visible wavelengths [5]. For in-plane fiber-to-chip coupling which is useful for scalable packaging, edge couplers with inverse tapers are widely used to match the large mode field diameter (MFD) of the fiber connected to a light source [5], [6]. Most developments focus on the short-wave infrared (SWIR) wavelength range. Since integrated waveguides become smaller with shorter wavelengths, the coupling requirements for 532 nm are more challenging. For instance, in [7], coupling losses of 7.5 – 11.3 dB/facet for TE were achieved in the wavelength range from 430 nm to 648 nm. Another group achieved around 4 dB per facet for 445 nm to 640 nm using a bi-layer coupling concept [8]. This illustrates that the fiber-to-chip interface is a critical component of visible-wavelength PICs. Therefore, this paper focuses in particular on the realization and

characterization of photonic integration concepts and their suitability for future integrated NV-diamond sensing systems.

Integrated system overview

The overall approach and the corresponding signal chains are shown Fig. 2. A central element is the optical transmission path, which transfers the excitation light (532 nm) from the light source through the diamond (D) to the signal photodiode (PD1) or reference photodiode (PD2). For this purpose, a photonic chip with integrated waveguides was manufactured at IMS CHIPS. Splitting the light into two optical paths enables balanced photodetection. To support future miniaturized implementations, three application-specific integrated circuits (ASICs) are being developed to drive, control, and read out the light source, microwave (MW), and photodiodes. A major challenge lies in integrating the various components — laser source, filters, waveguides, MW generators, resonator, and photodetectors — with the corresponding ASICs and peripheral electronics into a robust system. For this reason, a holistic view of the relevant subsystems is beneficial during implementation.

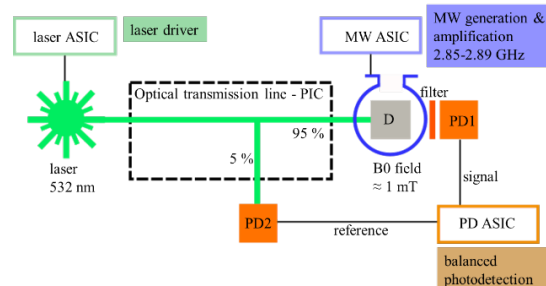


Fig. 2: System overview and signal chains of the integrated system concept.

The design of the optical signal paths influences the layout of the electronics and peripherals. A critical parameter for sensor performance is the diamond excitation power density. To balance thermal heating effects with sufficient excitation power below the saturation intensity, the power density is targeted between 1-10 kW/cm² [9], [10]. Based on this, an excitation density at the diamond of greater than 1 kW/cm² is targeted in this work. This gives rise to several requirements regarding the components to be used and their interfaces. Initially, the miniaturized passive components and subassemblies are operated using external laboratory equipment (a 532 nm fiber-coupled laser source and a MW generator) for optical and electrical supply during design and testing. In this work, the photonic chip is first developed for the optical transmission path and then combined with diamonds, filters, and photodiodes in a demonstrator

setup. The optical and electrical aspects are described separately below, with particular emphasis on the optical integration.

Optical Integration - Design

This approach uses integrated Si_3N_4 waveguides to enable the scalable splitting of signals and references on a chip, facilitating a compact optical architecture with dimensions on the order of a few millimeters. In addition, a hybrid approach involving on-chip printed microstructures for a robust, long-term assembly technology solution is being considered. This will be addressed in more detail later in the section. As previously mentioned, developing an efficient and robust solution for fiber-to-chip and chip-to-diamond coupling is the main challenge here. As shown in Fig. 3, two chip arrangements are being investigated: (1) a vertical arrangement with grating couplers (GC1 and GC2) and (2) a horizontal arrangement with edge outcoupling. The main part of the transmitted light should excite the diamond (D). For balanced photodetection, 5% of the transmitted light should be split to reach the reference diode (PD2) [11]. In arrangement (1), splitting occurs based on the grating coupler design, whereby only a small portion of the light passes GC1 without being decoupled. The partially etched grating coupler has a width of $1.4\text{ }\mu\text{m}$, a period of 250 nm , and a duty cycle of 0.7 . It consists of 150 periods. In (2), a directional coupler (DC) splits the light. It has a gap width of 400 nm and a coupling length of $18\text{ }\mu\text{m}$. Both concepts adopt an edge coupling approach on the input side. Because the fiber-chip interface is the most critical component, this work examines its development in greater detail.

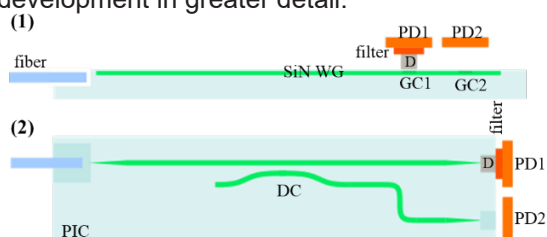


Fig. 3: Integrated optical concepts. (1) Side view of the vertical arrangement with grating couplers (GC1 and GC2); (2) top view of the horizontal arrangement with directional coupler (DC).

The proposed edge coupler is illustrated in Fig. 4 (a) and consists of two tapers (taper 1 and taper 2) ending in a single-mode waveguide (SM WG) that is 180 nm high and 280 nm wide. The performance of the edge coupler depends largely on the taper tip geometry and taper length. Taper 1 is designed to optimize mode overlap between the fiber and the waveguide

tip. To achieve the required mode expansion near the fiber, an additional shallow etch was applied to the entire taper 1 and the outer regions of taper 2, ensuring an adiabatic transition to the SM WG height level. A polarization-maintaining fiber (PMF, PM460-HP) was selected as the coupling fiber. Fig. 4 (b) shows the mode overlap as a function of the taper tip geometry. Dark red squares indicate the highest coupling efficiencies between the fiber mode and the waveguide tip fundamental transverse-electric (TE) mode. Considering manufacturing tolerances, a taper tip geometry with a height of 75 nm and a width of 90 nm was chosen, resulting in around 87% mode overlap.

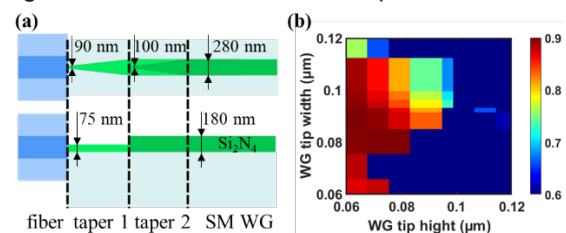


Fig. 4: (a) Schematic of fiber to Si_3N_4 edge coupler. (b) Mode field overlap between fiber and Si_3N_4 taper tip related to tip height and width.

Lumerical Eigenmode Expansion (EME) was used to simulate beam propagation and determine the taper lengths. Fig. 5 (a) displays a side view of the field distribution from the fiber to the SM WG. The transmitted power of the fundamental TE mode, as a function of the taper 1 length, is shown in Fig. 5 (b). Transmission only increases minimally beyond a length of $200\text{ }\mu\text{m}$. Therefore, the length of taper 1 was set to $200\text{ }\mu\text{m}$. For the selected geometry, Fig. 5 (c) displays the field profiles at the marked positions. It is evident that the mode over the tapered edge coupler experiences increased confinement within the waveguide. Position 1 belongs to the tip of taper 1 and its MFD has dimensions of $1.53\text{ }\mu\text{m}$ and $1.82\text{ }\mu\text{m}$ in the horizontal and vertical directions, respectively. Using the same procedure, the length of the second taper was set to $100\text{ }\mu\text{m}$, with a taper tip 100 nm wide and 280 nm high. The input coupler described above was simply mirrored for the output edge couplers without the shallow etch step.

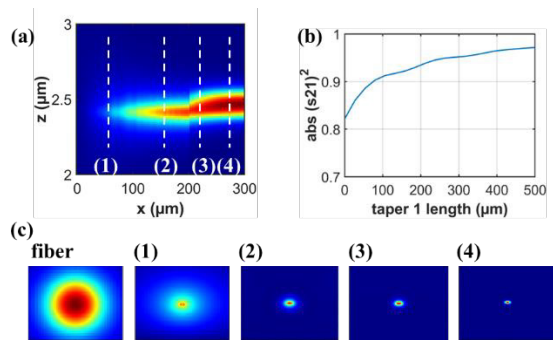


Fig. 5: (a) Sideview of the field distribution along the two-level taper from the fiber to the SM WG (E_{abs}) (b) Transmission as $\text{abs}(s_{21})^2$ in relation to taper 1 length. (c) WG mode fields along edge coupler.

To ensure a secure, long-term connection between the PIC and the laser fiber for a scalable setup, 3D-printed microstructures are employed. Microfiber clamps are used to mechanically position the fiber on the PIC. As a further solution to make the fiber-chip coupling more robust, the approach of using microlenses between the fiber facet and the chip facet, in combination with the micro clamps, is being investigated. The focusing microlens is designed and optimized to match the fiber mode field to the WG tip mode field (Fig. 5 (c) position1) and to provide increased optical assembly tolerance in the demonstrator setup, which is intended to enable scalable production. The simulated maximum achievable coupling efficiency from the fiber to the WG tip of the designed lens is 88 %. This represents a slight improvement over the simulation without the lens; however, it is expected that this will yield a benefit in terms of stable assembly technology.

Optical Integration - Fabrication and Measurement

To measure and characterize the photonic structures, the optical transmission lines were arranged on a 6-inch wafer in a chip-wise manner. The specified chip size is 12.5 mm x 3 mm. An exemplary chip layout is shown in Fig. 6 (a). The devices were fabricated and diced at IMS CHIPS. An oxide etch process in combination with a Si deep etch process was used to create smooth, vertical edge surfaces for in- and out-coupling. Samples of concepts (1) and (2) were fabricated on a silicon wafer coated with a 2.38 μm thick buried oxide (BOX) layer and a 180 nm thick Si_3N_4 film. First, shallow etching of the initial taper 1 and the grating coupler was performed using an e-beam lithographic process. Then, in the second step, full-depth waveguide etching was performed using the same procedure. A 2 μm SiO_2 layer was then

deposited as top oxide (TOX) which acts as cladding. An etch trench through the SiO_2 layers into the Si substrate was finally performed to realize the input and output ports of the devices. In Fig. 6 (b), the deep-etched in-coupling facet and input taper is visible. The scanning electron microscope (SEM) image of section A-A presents clearly a vertically edged oxide facet.

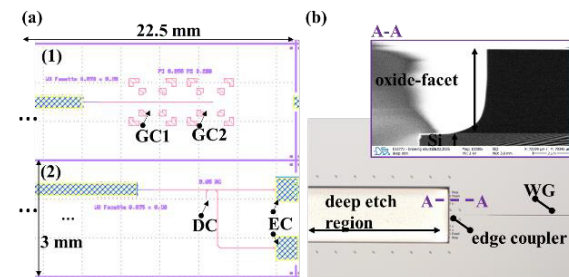


Fig. 6: (a) Layout image of the designed photonic concepts (1) and (2). The chip dimensions and the splitting or coupling structures are marked. (b) Microscope images of fabricated deep-etched in-coupling facet and edge coupler. Section A-A shows a SEM image of the etched oxide facet.

The chips were characterized using the setup shown in Fig. 7. Light from a laser source (RLTMGL-532UPM-250-5, Roithner) was coupled into the input facet via a PMF. Fiber alignment was performed using a microscope and a 6-axis stage. Light emitted from the photonic chip's output coupler was collected with a 400 μm core-diameter multimode fiber (MM), which was connected to an optical power meter (PM100USB, S154C, Thorlabs). When using a grating coupler as the outcoupling port, the MM fiber was placed vertically above the coupler. The transmitted light of the PMF-PIC-MM transmission lines for the signal and reference paths were measured, respectively. The measured power ratio is relative to the fiber-to-fiber (PMF-MM) coupled power. Due to system-level considerations, this publication considers only the overall transmission. The total throughput is calculated by adding the power emitted by the reference and signal paths. The results of concepts (1) and (2) are compared in Tab. 1. Clearly, Concept 2's total throughput is much higher than that of the grating coupler. It was found that a large proportion of the light is guided in the direction of the beam without being decoupled by the grating couplers. Clearly, Concept 2's total throughput is much higher than that of the grating coupler. It was found that a large proportion of the light is guided in the direction of the beam without being decoupled by the grating couplers.

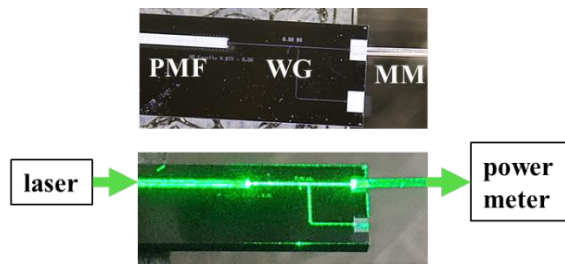


Fig. 7: Schematic measurement setup of concept (1) without and with excitation light

Tab. 1: Characterization results.

Concept	Reference path (%)	Signal path (%)	Total throughput (%)
(1)	0.3	10	10.3
(2)	2	27	29

Since it is not possible to accurately measure the power density with the employed setup, it is estimated from the radiating area and the out-coupled power. Assuming a distance of 2-3 μm from the decoupler to the diamond and 10 mW laser excitation, the radiating area for the grating coupler approach is approximately 20-30 μm^2 , leading to a power density in the range of 3-5 kW/cm². For the edge coupler, the estimated radiating area is 3-5 μm^2 , and the power density is 50-100 kW/cm². Despite the edge coupler being more efficient, both concepts reach the target value of greater than 1 kW/cm². Due to the higher efficiency, only the edge coupler architecture will be considered for the system demonstrator.

As part of the hybrid mechano-optical interface solution, the microstructures were directly printed onto the Si₃N₄ chips with a Quantum X align system, making use of Nanoscribe's Aligned 2-Photon Lithography (A2PL®) micro-fabrication technology. Initial printing trials and tests were conducted to print the structures into the deep-etched trenches and couple light into the waveguide. Fig. 8 shows the printed structures placed in the deep-etched groove at the coupling facet. The marks in the Si₃N₄ layer are used to align the structures relative to the Si₃N₄ waveguide during printing. The clamp, into which the PMF can be easily clipped, manages the x-, y-, and z-positioning of the fiber. As can be seen, these structures enable fiber positioning and light coupling into the waveguide without hexapod positioning. Additionally, clamp-based fixation enables robust adhesion of the fiber to the photonic chip.

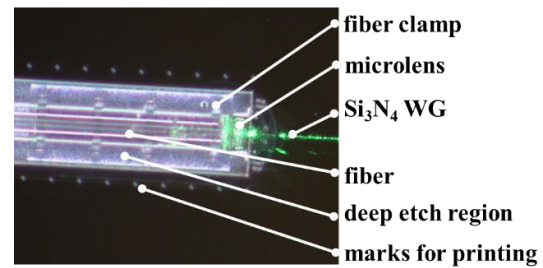


Fig. 8: Fiber-to-chip interface with a hybrid mechano-optical solution. 3D-printed microstructures at the facet of the PIC. A fiber clamp is used for mechanical positioning, and an in-coupling microlens increases assembly tolerances.

Electrical Integration

As already described, the electrical signal transmission involves driving the laser source, reading out and driving the photodiodes, and supplying the microwave power. To achieve the desired integrated system design from Fig. 2, the three ASICs must be integrated with the peripheral electronics.

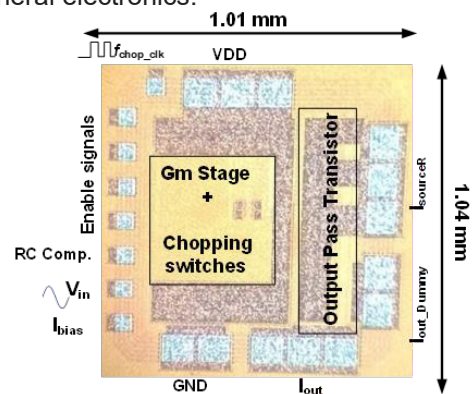


Fig. 9: Floorplan of the laser driver ASIC.

To enable system integrated NV-diamond magnetometry, it is necessary to precisely set the DC operating of deployed laser diodes. The noise performance needs to be close to the electrical shot-noise limit of the photodiodes. For realizing this, the laser driver ASIC shown in Fig. 9 is implemented. It is manufactured in a 130 nm SOI CMOS technology with a chip area of about 1 mm² and a supply voltage of 2.5 V. Driving currents in the range of 50 mA – 300 mA are possible, while the output noise is 210 pA/ $\sqrt{\text{Hz}}$ in the low-noise mode. Driving the system in DC mode is possible, as well sinusoidal and pulse modulation. The modulation frequencies of the laser current around the DC operating point range from 1 kHz to 40 MHz. The ASIC is integrated on a PCB with all peripherals for power supply, control and modulation within a size of about 3 x 3 cm².

Optical readout of the magnetic field sensor is done by commercial photodiodes (LAPD45HP1) connected to an ASIC-integrated Current-to-digital ADC with about 12-bit ENOB, 75 dB SNR, corresponding to an input-referred noise of about $32 \text{ pA}/\sqrt{\text{Hz}}$. The ASIC also implements an autobalancing circuit to reduce common-mode noise from the optical system. The ASIC has a power consumption of 50 mW at a size of 1.2 mm^2 in a 130 nm SOI technology.

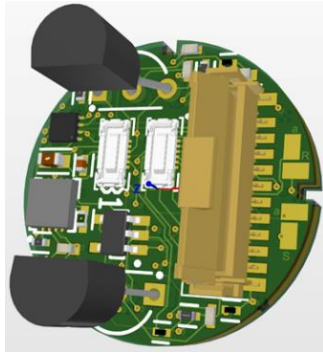


Fig. 10: PCB for balanced photodetection.

Photodiodes (PD1 and PD2) and the ASIC are combined on a round PCB shown in Fig. 10 with a diameter of 2 cm and all peripherals and connectors for external control and further signal processing e.g. lock-in detection.

The concept for stimulating the NV-diamond with microwaves in the GHz range is based on an ASIC-integrated PLL, combined with a mixer and a power amplifier, to synthesize the base-band signal. It shall enable system-integrated hyperfine stimulation and frequency modulation for lock-in detection. Finally, the GHz signal is delivered to the NV diamond by a customized patch resonator.

As an intermediate step towards system integration, the microwave is supplied by the commercial Windfreak Synth USB3 RF signal generator and a Mini Circuits ZX60-63GLN+ amplifier.

ODMR-Demonstrator

To assess the compatibility of the presented photonic and electronic building blocks with ODMR-based sensing, a preliminary demonstrator was assembled, as shown schematically in Fig. 11. The setup combines the developed PIC with a diamond-filter-photodiode arrangement as well as MW and photodiode readout electronics. At this stage, the arrangement serves as a feasibility platform for evaluating optical excitation, fluorescence collection, and microwave-driven spin manipulation under conditions relevant for future integrated implementations. According to the schematic configuration, the demonstrator was set up

except for the absence of the Halbach ring for the B0 field.

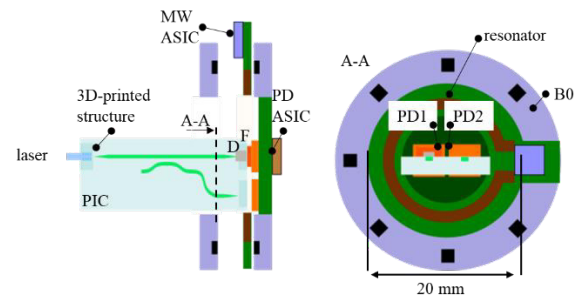


Fig. 11: Schematic ODMR demonstrator setup. On the left side, top view of the demonstrator setup. On the right side, cut A-A view.

The PIC was assembled with a diamond-filter-photodiode stack, the photodiode readout, and the MW electronics. Due to the horizontal edge coupling approach, fixing the diamond-filter-photodiode stack is more difficult, as the PIC cannot serve directly as a mechanical sub-mount. As a filter, a 550 nm long-pass filter (15-446, Edmund Optics) was cut to the target geometry of $1.4 \text{ mm} \times 1 \text{ mm}$ and glued to the signal photodiode using an optical adhesive. The diamond (DNL-B14) with a cube edge length of $500 \mu\text{m}$ was glued to the filter using an optical adhesive. For the input connection, the PM laser fiber was fixed using the on-chip 3D-printed micro-fiber clamps. The resonator and photodiode boards were connected via 3D-printed structures to position the diamond stack at the center of the resonator and at the outcoupling edge coupler of the PIC. Fig. 12 presents the assembled demonstrator. As shown on the right side, optical excitation of the diamond and the generation of a measurable fluorescence signal were achieved.

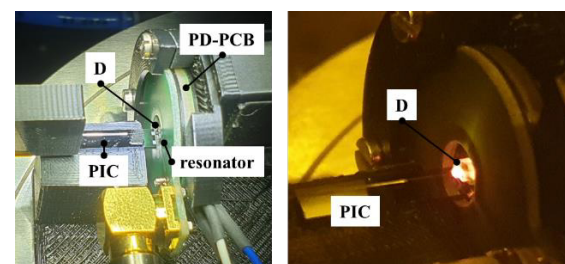


Fig 12: ODMR demonstrator setup. On the left side assembly of the sensor head. On the right side long-pass filtered image of the sensor to suppress the source light.

For an initial ODMR feasibility measurement, the diamond was adhered to a multimode fiber and positioned at the center of the MW resonator. In this preliminary configuration, the signal

photodiode and filter were placed outside the sensor head. A total of ten consecutive frequency sweep measurements were recorded in the range from 2.850 GHz to 2.890 GHz using a microwave power of 6 dBm. Fig. 13 shows the ODMR dip in the individual measurements (top) and in the average signal (bottom). Despite the absence of an integrated B₀ field, a splitting of the dip is observable, indicating the presence of an ambient magnetic field. These results confirm the general compatibility of the employed optical excitation concept with ODMR measurements, while a fully integrated quantitative evaluation remains the subject of future work.

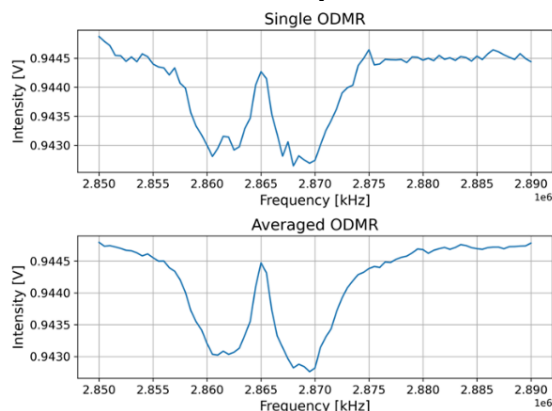


Fig 13: ODMR demonstrator measurement result. Top: result of a single ODMR measurement; bottom: average of 10 consecutive single ODMR measurements

For future ODMR measurements with the PIC-diamond-filter-photodiode stack, it is essential to protect the signal photodiode from unfiltered stray light. In particular, light incident from the side and bypassing the filtering layer is problematic. Future work will therefore include an optimized shielding and integration concept for the photodiode in order to enable quantitative measurements in a more highly integrated configuration.

Conclusion

In conclusion, this work presented key photonic and electronic building blocks for a miniaturized approach to NV-diamond magnetometry. The photonic integrated circuit was described in detail, and two coupling concepts were investigated and compared. Due to its superior transmission efficiency, the edge coupling approach was selected for further development despite its increased assembly complexity. The realized photonic transmission line achieves a power density exceeding 1 kW/cm², which is sufficient for optical excitation of diamond-based sensing elements.

The presented approach offers a path toward compact and scalable sensor implementations with dimensions on the order of only a few millimeters. In particular, the achieved excitation power density and the compatibility with compact optical integration are promising with regard to future miniaturized NV-based sensing systems. Preliminary ODMR feasibility measurements further indicate that the developed optical excitation concept is suitable for ODMR-based operation.

Future work will focus on quantitative evaluation of lens-to-chip coupling strategies in order to further improve optical efficiency. In addition, optimized integration of photodiodes and dedicated ASICs will be investigated to enable quantitative system-level characterization in more highly integrated demonstrator configurations.

Acknowledgement

This work was supported by the Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) within the project “QHMI2” (FKZ 03ZU2110FF).

References

- [1] C. Zhang *et al.*, “Optimizing NV magnetometry for Magnetoneurography and Magnetomyography applications,” *Front. Neurosci.*, vol. 16, 2023, doi: 10.3389/fnins.2022.1034391.
- [2] R. Katsumi, K. Takada, F. Jelezko, and T. Yatsui, “Recent progress in hybrid diamond photonics for quantum information processing and sensing,” *Commun. Eng.*, vol. 4, no. 1, pp. 1–19, 2025, doi: 10.1038/s44172-025-00398-2.
- [3] F. M. Stürner *et al.*, “Compact integrated magnetometer based on nitrogen-vacancy centres in diamond,” *Diam. Relat. Mater.*, vol. 93, pp. 59–65, 2019.
- [4] K. Huang *et al.*, “A Portable and Highly Integrated Solid-State Quantum Magnetometer Module Based on the Diamond NV Color Centers,” *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–9, 2024, doi: 10.1109/TIM.2024.3470972.
- [5] K. A. Buzaverov *et al.*, “Silicon Nitride Integrated Photonics from Visible to Mid-Infrared Spectra,” *Laser Photonics Rev.*, vol. 18, no. 12, 2024, doi: 10.1002/lpor.202400508.
- [6] B. Bhandari, C. S. Im, K. P. Lee, S. M. Kim, M. C. Oh, and S. S. Lee, “Compact and broadband edge coupler based on multi-stage silicon nitride tapers,” *IEEE Photonics J.*, vol. 12, no. 6, 2020, doi: 10.1109/JPHOT.2020.3036498.
- [7] W. D. Sacher *et al.*, “Visible-light silicon nitride waveguide devices and implantable neurophotonic probes on thinned 200 mm silicon wafers,” *Opt. Express*, vol. 27, no. 26, p. 37400, 2019, doi: 10.1364/oe.27.037400.
- [8] Y. Lin *et al.*, “Low-loss broadband bi-layer edge couplers for visible light,” *Opt. Express*, vol. 29,

- no. 21, p. 34565, 2021, doi: 10.1364/oe.435669.
- [9] J. M. Schloss, J. F. Barry, M. J. Turner, and R. L. Walsworth, "Simultaneous Broadband Vector Magnetometry Using Solid-State Spins," *Phys. Rev. Appl.*, vol. 10, no. 3, p. 1, 2018, doi: 10.1103/PhysRevApplied.10.034044.
- [10] T. Wolf *et al.*, "Subpicotesla diamond magnetometry," *Phys. Rev. X*, vol. 5, no. 4, pp. 1–10, 2015, doi: 10.1103/PhysRevX.5.041001.
- [11] B. Baradaran, A. J. Black, S. R. Heilbronner, and T. Akkin, "Visible light polarization-sensitive optical coherence tomography with balanced detection," *J. Biomed. Opt.*, vol. 30, no. 03, pp. 1–11, 2025, doi: 10.1117/1.jbo.30.3.036002.

Pyroelektrische DLaTGS-Detektoren für FTIR-Spektrometer – Moderne Technologie oder alter Hut?

Martin Ebermann, Doris Mutschall, Eric Müller
InfraTec GmbH Infrarotsensorik und Messtechnik, Dresden

Zusammenfassung

Im vorliegenden Beitrag soll eine Übersicht zu IR-Detektoren gegeben werden, die auf dem pyroelektrischen Material DLaTGS basieren. Nach einer kurzen Einführung in die Funktionsweise werden die Anforderungen diskutiert, die sich aus ihrer Anwendung in FTIR-Spektrometern ergeben. Dabei wird ein Vergleich des Detektormaterials DLaTGS mit dem anderweitig etablierten Material Lithiumtantalat gezogen. Im Weiteren wird auf ausgewählte Designaspekte und Lösungsansätze eingegangen, die DLaTGS-Detektoren von InfraTec gegenüber dem bisherigen Stand der Technik auszeichnen.

Keywords: Infrarotsensoren, Fourier-Transform-Spektrometer, modifiziertes Triglycinsulfat, Lithiumtantalat.

Einleitung

Die Infrarot-Spektroskopie ist eine Standardmethode in der chemischen Analytik. Fourier-Transform-Infrarot-(FTIR)-Spektrometer gehören zur Grundausstattung eines jeden Analytiklabors. An den IR-Detektor für ein solches Gerät werden spezielle und widerstreitende Anforderungen gestellt: Die Analytik verlangt neben großer spektraler Breite eine hohe Empfindlichkeit, Auflösung und niedrige Nachweisgrenzen, wofür das höchstmögliche Signal-Rausch-Verhältnis benötigt wird. Das Funktionsprinzip des FTIR-Spektrometers bringt zusätzliche Forderungen mit sich: Die Belastbarkeit mit sehr hohen Strahlungsleistungen sowie einen großen Dynamikbereich bei gleichzeitig sehr guter Linearität. Seit mehreren Jahrzehnten gibt es hierfür eine Standardlösung: den pyroelektrischen Detektor auf Basis von einkristallinem Triglycinsulfat bzw. dessen Modifikationen durch Deuterierung und Dotierung (DLaTGS).

Aufbau und Funktionsweise pyroelektrischer Detektoren

Pyroelektrische Detektoren haben weite Verbreitung als Strahlungsempfänger in infrarot-optischen Messsystemen und Spektrometern. Sie gehören zur Klasse der thermischen Sensoren und besitzen damit eine breitbandige spektrale Empfindlichkeit. Sie sind außerdem robust, preiswert und können ohne Kühlung betrieben werden. Eine vollständige Darstellung der Funktionsweise pyroelektrischer Detektoren - über die im Folgenden angebotene stark verkürzte Zusammenfassung hinaus - ist in der Literatur nachzulesen [1].

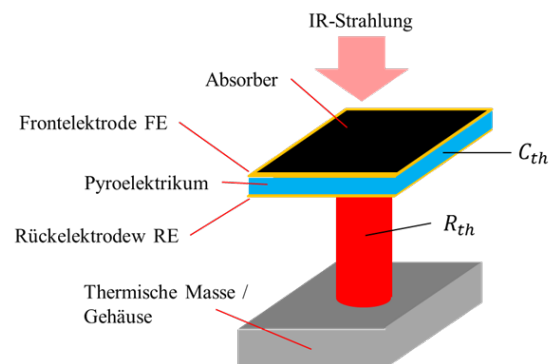


Abb. 1: Schema des Sensorelement und seiner thermischen Ankopplung in einem pyroelektrischen Detektor.

Das zweistufige Wandlungsprinzip lässt sich kurz wie folgt beschreiben: Die einfallende Strahlung wird auf einem flächigen und dünnen Sensorelement absorbiert und verursacht eine Temperaturerhöhung. Das Element besteht aus einem pyroelektrischen Material, welches auf Vorder- und Rückseite mit Elektroden beschichtet ist (Front- und Rückelektrode, FE und RE), so dass es einen elektrischen Plattenkondensator bildet (Abb. 1). Die Temperaturänderung führt aufgrund des pyroelektrischen Effekts zu einer Änderung der spontanen Polarisierung im Material. Dadurch werden die Elektroden aufgeladen bzw. wird ein Stromfluss erzeugt, welcher zur Weiterverarbeitung in ein Spannungssignal (Sensorsignal) gewandelt wird. Eine Besonderheit dieses Wandlungsprinzips besteht darin, dass nur Änderungen der Temperatur zu einem Signal führen.

Pyroelektrische Detektoren sind somit nur für modulierte Strahlung empfindlich nicht jedoch für Gleichlicht (Continuous Wave - CW) – im Unterscheid zu anderen thermische IR-Sensoren wie z.B. Thermopiles oder Bolometern. Das kann in vielen Anwendungen als Vorteil ausgenutzt werden.

Zur Maximierung der Strahlungsabsorption ist das Element oftmals mit einer zusätzlichen dünnen absorbierenden Schicht (Absorber) versehen. Die Empfindlichkeit des Detektors wird weiterhin dadurch maximiert, dass das Element besonders dünn ausgeführt wird, wodurch es eine geringe Wärmekapazität besitzt. Das Element muss zudem im Detektoraufbau thermisch isoliert montiert bzw. elektrisch kontaktiert werden. Die Art der Montage bzw. das umgebende Gas (typischerweise Luft oder Stickstoff) leiten die absorbierte Wärme in das Gehäuse bzw. die Umgebung ab (Abb. 1). Dies lässt sich vereinfacht durch einen thermischen Übergangswiderstand R_{th} beschreiben. Zusammen mit der Wärmekapazität des Sensorelements C_{th} entsteht ein thermisches Tiefpassverhalten mit der Zeitkonstante

$$\tau_{th} = R_{th} \cdot C_{th}$$

bzw. der thermischen Eckfrequenz

$$f_{th} = 1/(2\pi \cdot \tau_{th}).$$

Im Falle von CW-Bestrahlung bzw. für niedrige Modulationsfrequenzen unterhalb der thermischen Eckfrequenz erfährt das Element die maximale Temperaturänderung, welche ausschließlich durch R_{th} bestimmt wird. Die Wärmekapazität des Sensorelements C_{th} ist maßgeblich dafür, wie schnell sich seine Temperatur ändert. Nach dem bereits erwähnten differenzierenden Verhalten ist der pyroelektrische Strom proportional zur Geschwindigkeit der Temperaturänderung. Demnach besitzt der Sensor eine hohe Empfindlichkeit, wenn die Modulationsfrequenz der Strahlung oberhalb der thermischen Eckfrequenz liegt. Für niedrigere Frequenzen fällt die Empfindlichkeit ab und verschwindet bei unmodulierter Strahlung (CW).

Diese Betrachtung gilt zunächst nur für die Generierung des pyroelektrischen Stroms, die sogenannte Stromempfindlichkeit R_i (Abb. 2), sie lässt die nachfolgende elektrische Signalwandlung außer Acht. Letztere kann unterschiedlich ausgeführt sein und prägt dem Sensor ein zusätzliches Frequenzverhalten auf. Viele Anwendungen von pyroelektrischen Detektoren, z.B. in der NDIR-Gasanalyse, nutzen thermische Strahler die entweder gepulst angesteuert oder gehoppert werden.

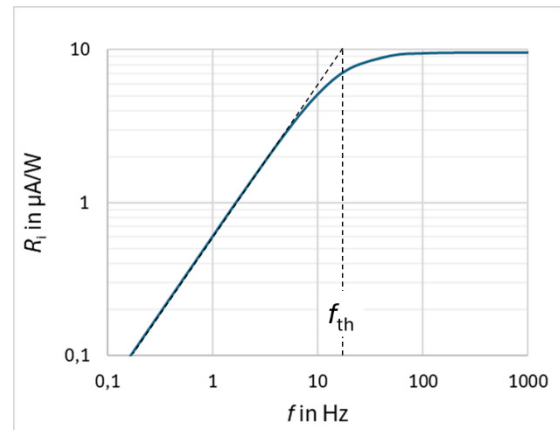


Abb. 2: Typische Frequenzabhängigkeit der Stromempfindlichkeit R_i , beispielhaft für einen DLaTGS-Detektor mit einer thermischen Eckfrequenz im Bereich von 10...20 Hz.

Die Modulationsfrequenzen liegen oft im Bereich von 1...10 Hz. In diesen Fällen ist eine besonders gute thermische Isolation des Sensorelements anzustreben, die thermischen Zeitkonstanten liegen typischerweise im Bereich von 100...200 ms. Dies wird durch eine Montage und Kontaktierung an nur wenigen kleinen Punkten mittels Klebstoffs erreicht, wobei das Element im Übrigen von einer dicken Gasschicht umgeben ist. Die Frontelektrode wird meist durch einen Bonddraht elektrisch kontaktiert. Einfache Ausführungsformen nutzen ein Loch in einem Substrat als Unterlage (z.B. Leiterplatte) andere Ausführungen nutzen mikrostrukturierte Chiphalter [2], oder Bonddrahtbrücken zur gleichzeitigen mechanischen und thermischen Entkopplung sowie zur elektrischen Kontaktierung [3].

Anwendung in FTIR-Spektrometern

Ein wichtiges Anwendungsgebiet pyroelektrischer Detektoren sind FTIR-Spektrometer. Die Randbedingungen unterscheiden sich hier in mehreren Aspekten sehr deutlich von den oben genannten Anwendungsfällen. Kernstück eines FTIR ist ein Zweistrahlinterferometer (z.B. Michelson-Interferometer), welches die Strahlung von einer breitbandigen thermischen IR-Quelle durch Interferenz der beiden Teilstrahlen im Interferometer in der Intensität moduliert, so dass die Wellenlängeninformation in den Frequenzbereich transformiert wird. Durch den Detektor wird schließlich ein Interferogramm aufgezeichnet, aus dem durch Fouriertransformation das Spektrum gewonnen wird [4]. Für ein Gerät typischer Bauart für den mittleren und langwelligen IR-Bereich (2 – 30 µm) liegt der Signalfrequenzbereich des Interferogramms im Bereich 100 Hz - 2 kHz.

Die Geräte sind normalerweise so ausgelegt, dass ein maximaler optischer Durchsatz

erreicht wird. Dies beinhaltet eine mehrstufige Abbildung bzw. Strahlformung, vom IR-Emitter durch das Interferometer und die zu messende Probe und schließlich auf den Detektor, im Idealfall exakt auf die Größe der Empfängerfläche angepasst.

Die Gesamtstrahlungsleistung der Quelle beträgt in der Regel $\gg 100$ mW. Davon wird letztlich zwar nur ein sehr kleiner Anteil in eine auswertbare Intensitätsmodulation im Interferogramm übersetzt, diese besitzt jedoch einen sehr großen Dynamikbereich (Abb. 3). Ein großer Anteil (einige zehn Milliwatt) gelangt als Gleichlicht (CW) auf das Sensorelement. Dieser Aspekt wird oft vernachlässigt, spielt jedoch eine sehr wichtige Rolle.

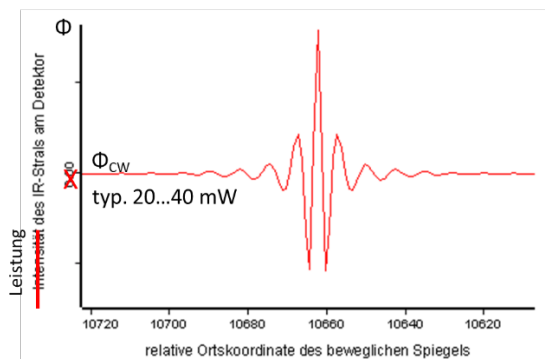


Abb. 3: Typischer Verlauf des Interferogramms in einem FTIR-Spektrometer, modif. Darstellung nach <https://de.wikipedia.org/wiki/FTIR-Spektrometer> (letzter Aufruf am 29.05.2026).

Pyroelektrische Detektoren sind unter anderem deshalb die bevorzugten Empfänger für FTIR-Spektrometer, weil sie nur auf die modulierten Strahlungsanteile ansprechen. Der CW-Anteil erzeugt dagegen kein Signal. Gleichwohl muss er im thermischen Verhalten der Detektoren berücksichtigt werden. Die Wärme muss vom Element sicher und definiert abgeführt werden. In der oben beschriebenen klassischen Bauform pyroelektrischer Detektoren mit großer Zeitkonstante würde die Erwärmung einige 10 K betragen. Bei einem DLaTGS-Detektor hätte dies entweder den vollständigen Verlust der Empfindlichkeit durch Entpolung, im schlimmsten Fall eine Zerstörung des Elements, mindestens aber ein stark nichtlineares Sensorverhalten zur Folge [6]. Die Gründe werden im folgenden Abschnitt näher erläutert.

Detektormaterial DLaTGS

Für einen fundierten Vergleich verschiedener pyroelektrischer Materialien, deren Materialparameter und Leistungskenngrößen (Figure-of-Merits) sei nochmals auf die Literatur verwiesen [1]. Im Folgenden wird nur auf spezielle Aspekte eingegangen.

Die Spannungsempfindlichkeit des Detektors ist im Wesentlichen proportional zum pyroelektrischen Koeffizienten und umgekehrt proportional zur Permittivität und damit sowohl material- als auch temperaturabhängig

$$R_v \sim p(T)/\varepsilon_r(T).$$

Die relative Lage der Curie-Temperatur T_C zum Betriebstemperaturbereich des Detektors bestimmt die Höhe des Temperaturkoeffizienten der Materialparameter. Bei der Curietemperatur findet ein Phasenübergang von der ferroelektrischen in die paraelektrische Phase statt, in der die pyroelektrischen Eigenschaften verschwinden. Weiter entfernt besitzen p und ε_r geringe bzw. leicht positive Temperaturkoeffizienten. Steigt die Temperatur in die Nähe von T_C dann steigen p und ε_r überproportional stark und erreichen in unmittelbarer Umgebung von T_C ein Maximum (Abb. 4).

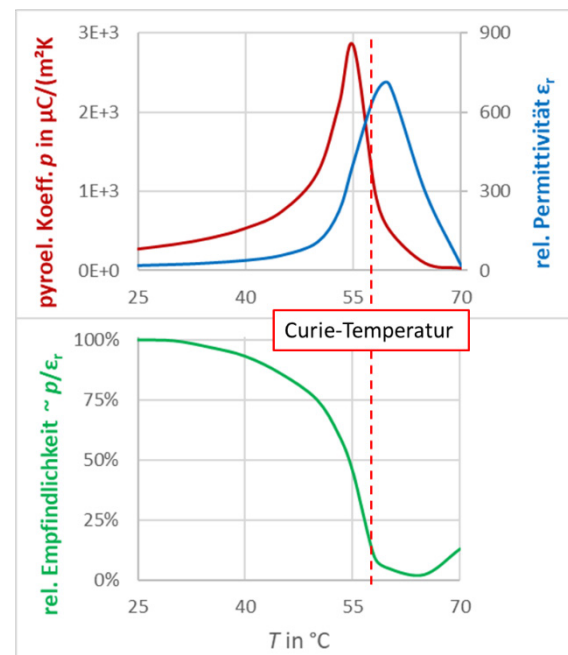


Abb. 4: Typische Temperaturabhängigkeiten eines pyroelektrischen DLaTGS-Detektors: pyroelektrischer Koeffizient p und Permittivität ε_r von einkristallinem DLaTGS (oben), sowie relative Empfindlichkeit des Sensors (unten).

Die Curietemperatur pyroelektrischer Materialien unterscheidet sich mitunter sehr stark. Bei Lithiumtantalat (LiTaO_3) - als eines der am häufigsten verwendeten Pyroelektrika - liegt sie bei > 600 °C. Damit ist die Temperaturabhängigkeit im Bereich normaler Betriebsbedingungen sehr gering und kann in der Regel vernachlässigt werden. DLaTGS besitzt dagegen eine sehr niedrige Curietemperatur im Bereich von 55...62 °C. Wenn sich die Temperatur eines DLaTGS-Sensorelements

nur um einige Kelvin ändert - infolge der Bestrahlung oder durch die Umgebung - ändert sich die Empfindlichkeit des Detektors merklich, was letztlich zu Messabweichungen oder Nicht-linearitäten führt.

Für FTIR-Detektoren ist DLaTGS besonders bevorzugt, weil es von allen bekannten Pyroelektrika das beste SNR bietet, insbesondere im Frequenzbereich > 100 Hz. Seine große Temperaturempfindlichkeit und seine niedrige Curietemperatur werden dabei zugunsten des hohen SNR in Kauf genommen.

Erschwerend kommt hinzu, dass DLaTGS wasserlöslich ist und die Detektoren somit sehr feuchteempfindlich sind. In Spezialfällen wird auch LiTaO_3 als Detektormaterial für FTIR-Spektrometer verwendet, insbesondere dann, wenn es auf höchste photometrisch Genauigkeit ankommt, was eine hohe Linearität des Detektors erfordert, oder wenn eine hohe Robustheit gegen Feuchte und Temperatur verlangt wird.

Tab. 1: Vergleich von LiTaO_3 - und DLaTGS-Detektoren anhand FTIR-spezifischer Anforderungen.

Aus Analytik	Anforderung an Detektor	LiTaO_3	DLaTGS
Empfindlichkeit, Auflösung und Nachweisgrenzen	Höchstmögliches Signal-Rausch-Verhältnis		!
Universell, flexibel	Breiter Wellenlängenbereich		
Aus Messverfahren			
Große Spitzenwerte, Information ist noch in kleinsten Intensitätsmodulationen enthalten	Großer Dynamikbereich, Signalfrequenzbereich 100 Hz ... 10 kHz		
Hohe CW-Bestrahlung ohne Informationsgehalt	Hohe Zerstörschwelle, gute Wärmeabfuhr, CW-Unterdrückung		
Weitere Anforderungen			
Kostengünstig			
Robust, langlebig, stabil			

DLaTGS-Detektoren sind seit Jahrzehnten am Markt etabliert und im Wesentlichen bewährt. Weiterentwicklungen und Optimierungen fanden jedoch kaum noch statt, der technologische Stand war Jahrzehnte alt und Wettbewerb kaum vorhanden. Mit der Vorstellung einer eigenen, von Grund auf neu entwickelten Produktreihe von DLaTGS-Detektoren bietet die Firma InfraTec eine modernisierte Technologie bei gleichzeitig hoher und reproduzierbarer Leistung und Qualität an. Es werden hierbei insbesondere die Aspekte adressiert, bei denen DLaTGS-Detektoren bisher sehr deutlich im Nachteil waren (siehe Tab.1). InfraTec bietet zudem weiterhin auch FTIR-Detektoren auf Basis von LiTaO_3 an.

Kristallzüchtung und Herstellung der Sensorelemente

Die Herstellung des Grundmaterials beginnt mit der Synthese von Triglycinsulfat (TGS) aus hochreinen Ausgangsstoffen. Anschließend wird das Material mehrfach in schwerem Wasser (D_2O) umkristallisiert und somit deuteriertes Material (DTGS) gewonnen. Damit wird eine übersättigte Züchtungslösung in D_2O hergestellt, in die ein Keimkristall eingesetzt wird. Durch sehr langsame und präzise geregelte Absenkung der Temperatur lagert sich das Material aus der Lösung am Keim an und der Kristall beginnt zu wachsen. Durch Zugabe von L- α -Alanin in die Lösung wird eine zusätzliche Dotierung erreicht, die die Materialeigenschaften entscheidend verbessert [5][6]. Auf diese Weise werden DLaTGS Kristalle bei InfraTec herstellt (Abb. 5).



Abb. 5: Foto eines DLaTGS-Kristalls nach der Entnahme aus der Züchtungslösung.

Der Kristall wird nun auf einer Drahtsäge in Scheiben (Wafer) geschnitten. Anschließend werden die Wafer durch Läpp- und Polierprozesse auf die Zieldicke im Bereich von $< 20 \mu\text{m}$ abgedünnt. Sie werden vorder- und rückseitig über Schattenmasken mit Elektroden beschichtet und zu Sensorelementen vereinzelt. All diese Prozesse müssen wasserfrei, mit sehr geringer Temperaturbelastung und innerhalb sehr enger Prozessfenster ausgeführt werden, damit eine gleichbleibende Materialqualität und Ausbeute erreicht wird.

Detektoraufbau und Montage der Sensorelemente

Der Grundaufbau von Detektoren bzw. die Montage der Sensorelemente für geringe Strahlungsleistung und niedrige Modulationsfrequenzen (große thermische Zeitkonstante) wurde bereits oben beschrieben. Für hohe CW-Strahlungsleistungen in FTIR-Spektrometern müssen die Elemente anders montiert werden, so dass die absorbierte Strahlungswärme schneller und gleichmäßiger abgeführt und deren Temperatur begrenzt wird. Front- und Rückelektroden des Elements sind üblicher-

weise so strukturiert, dass sich eine auf die Optik der Geräte angepasste empfindliche Fläche ergibt. Typische Größen sind im Bereich $\varnothing 1,3 \dots 2,0$ mm.

In [8] ist ausführlich dargestellt und analysiert, wie das nach dem bisherigen Stand der Technik bei DLaTGS-Detektoren realisiert wird und welche Vor- und Nachteile sich daraus ergeben. In einer typischen Ausführungsform ist eine dünne Luftschicht unter dem Element ausgebildet, über den die Wärme in ein darunter liegendes Trägersubstrat abgeführt wird. Das Element ist hierbei mit einem einzigen mittigen und punktförmigen Indium-Bondkontakt mechanisch fixiert und elektrisch kontaktiert [9]. In einer anderen Ausführungsform wird das Element umlaufend an seinen Außenkanten verklebt. Die Wärmeabfuhr erfolgt hier im Wesentlichen durch die laterale Wärmeleitung innerhalb des Elements, von der bestrahlten Fläche im Zentrum zu den Kanten. Die elektrische Kontaktierung der Rückelektrode erfolgt durch einen angeklebten Bonddraht innerhalb der empfindlichen Fläche [8]. In beiden Varianten wird der elektrische Kontakt zur Frontelektrode durch einen leitfähig geklebten Bonddraht hergestellt, wobei die Kontaktstelle inner- oder außerhalb der empfindlichen Fläche liegen kann.

Alle genannten Montagevarianten haben den Nachteil, dass sich durch die Platzierung der Montage- bzw. Kontaktstellen innerhalb der empfindlichen Sensorfläche eine inhomogene Temperatur- bzw. Empfindlichkeitsverteilung ergibt. Bei der zentralen Chipmontage ist es zudem sehr schwierig den Abstand bzw. die Luftschicht präzise und reproduzierbar im Montageprozess herzustellen. Zusätzlich besteht die Gefahr, dass das Element verkippt montiert wird, im Extremfall kann es sogar zur Berührung der Ecken oder Kanten mit dem darunter liegenden Substrat kommen. Problematisch ist auch hier wieder eine ungleichmäßige Temperatur- und Empfindlichkeitsverteilung. Schlechte Reproduzierbarkeit und große Fertigungsstreuungen der Detektoreigenschaften sind die Folge. Schon ein kleiner Versatz des Strahlfokus auf dem Sensorelement, etwa durch eine Rückwirkung der Probe auf den Strahlengang im Gerät, kann zu Messfehlern führen.

Bei genauer Analyse wird klar: Die funktionsbestimmenden Aspekte des Aufbaus – die mechanische Fixierung des Elements, die thermische Isolierung und die elektrische Kontaktierung – sind derart miteinander verkoppelt, dass sie sich nicht unabhängig voneinander optimieren lassen. Der hier vorgestellte Lösungsansatz besteht im Kern darin, dass die genannten Funktionen voneinander entkoppelt werden, und somit unabhängig voneinander

optimierbar sind. Abb. 6 zeigt das Grundprinzip in einer stark vereinfachten Weise: Schlüsselement ist ein mit Si-MEMS-Technologie hergestellter mikrostrukturierter Chiphalter. Er besitzt ein durch Ätzprozesse definiertes abgestuftes Höhenprofil: In den Randbereichen befinden sich Podeste, auf denen das Element flächig aufliegt, sowie tief geätzte Gebiete mit metallischen Kontaktflächen (Pads) für Fixierung und Kontaktierung des Elements. Auf dem Element sind entsprechende Pads an den Elektroden nach außen geführt. Im zentralen Bereich der empfindlichen Fläche befindet sich eine flach geätzte Kavität. Durch sie wird unter dem Element eine sehr homogene einige Mikrometer dicke Luftschicht ausgebildet.

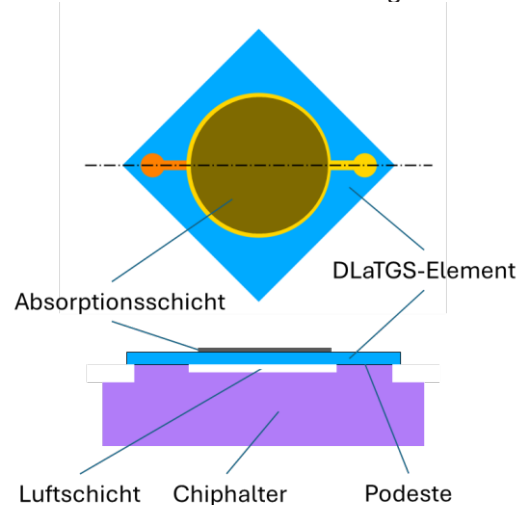


Abb. 6: Montage des Sensorelements mit Si-MEMS-Chiphalter, Klebe, Kontakt- und Bondstellen zur Vereinfachung weggelassen

Der Entwurf wurde durch FEM-Simulationen (COMSOL®) unterstützt. Hierbei wurde nicht nur die Temperaturentwicklung im Zeit- und Frequenzbereich analysiert und optimiert, sondern auch thermo-mechanische Spannungen und Verwölbungen betrachtet. Je nach Anwendungsfall erfolgt die Optimierung der Anordnung und der Luftschichtdicke für ein bestimmtes Leistungsniveau der Bestrahlung. Im Endergebnis ergeben sich Zeitkonstanten τ_{th} , die im Bereich von 10 ... 50 ms liegen.

Das Simulationsergebnis in Abb. 7 zeigt, was ein Betrieb über der im Entwurf angenommenen maximalen Strahlungsleistung zur Folge haben kann: Bei 50 mW beginnt sich das Element bis zu 30 μm nach oben zu verwölben und es entwickelt sich eine Temperaturüberhöhung von bis zu 40 K. Das würde mit Sicherheit eine Entpolung des Elements verursachen. Bei moderater Strahlungsbelastung von 10 mW verhält sich das System gutmütig und linear.

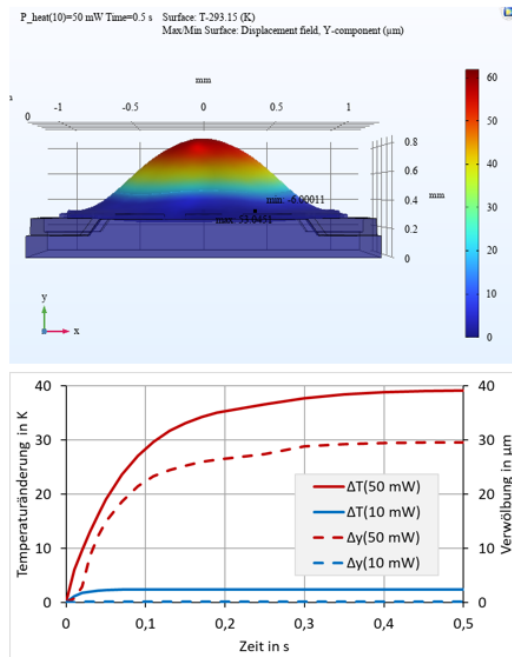


Abb. 7: Thermo-mechanische FEM-Analyse der Elementmontage nach Abb.6 mit sprungförmiger Anregung mit 10 mW und mit 50 mW konzentriert auf die empfindliche Fläche mit Ø1,3 mm: räumliche (überhöhte) Darstellung der Verwölbung (oben, Farbskala in µm, 50 mW), zeitlicher Verlauf der mittleren Elementtemperatur und Verwölbung (unten).

Klebeverbindungen und rückseitige Kontaktierung des Elements sind durch eine Kombination von festen und flexiblen Klebstoffen ausgeführt. Als Absorber wird eine bewährte Metallschwarzschrift genutzt, welche eine sehr hohe und spektral flache Absorption besitzt. Im Gehäuse ist eine rauscharme Vorverstärkerschaltung mit diskreten Bauelementen hybrid integriert und zusätzlich vergossen. Der gesamte Aufbau wird in weitgehend automatisierten Prozessen mit Chip- und Drahtbondautomaten realisiert.

Abb.8 zeigt den offenen Innenaufbau des Detektors mit strukturiert abgeschiedenem Absorber, sowie den verschlossenen Detektor mit aufgeschweißter Fensterkappe.



Abb. 8: Gesamtaufbau ohne (links) und mit KBr-Fensterkappe im TO39-Gehäuse (rechts).

Tab. 2: Vergleich der wichtigsten Kenngrößen von pyroelektrischen Detektoren mit LiTaO_3 (Typ LIE-332f) und DLaTGS (Typ DSV-7311), (Elementgröße Ø 1,3 mm).

Parameter	f	LIE-332f	DSV-7311	Einheit
R	100 Hz	140	650	V/W
D^*	100 Hz	3,0E8	5,2E8	Jones
R	1 kHz	13	65	V/W
D^*	1 kHz	1,3E8	3,3E8	Jones
τ_{th}		20	12	ms

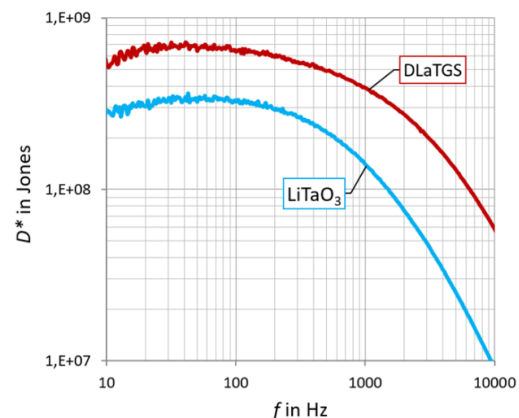


Abb. 9: Vergleich der frequenzabhängigen spezifischen Detektivität D^* .

Charakterisierung der Detektoren

Die Detektoren werden vor Auslieferung zu 100% geprüft bzw. charakterisiert. Typische Kenndaten der beiden Detektortypen LIE-332f (LiTaO_3) und DSV-7311 (DLaTGS) sind in Tab. 2 angegeben, Abb. 9 vergleicht die frequenzabhängige spezifische Detektivität D^* als Maß für das SNR. Im relevanten Frequenzbereich von 100 Hz - 2 kHz bietet DLaTGS um ca. Faktor 2 - 3 bessere Leistungswerte.

Die Messung der Empfindlichkeit (Responsivity R) erfolgt mit einem Blackbody-Strahler bei 500 K. Für die Messung des Rauschdichtespektrums und die anschließende Berechnung des D^* wird wie üblich auf eine Rauschbandbreite von 1 Hz normiert. Alle Messungen wurden nahe bei RT durchgeführt (25 ... 30°C).

Messung von Stabilität und Linearität

Die besonderen Einsatzbedingungen in FTIR-Spektrometern, wie beispielsweise starke und schnell wechselnde Bestrahlungsleistungen, müssen für eine aussagekräftige Charakterisierung der Detektoren möglichst gut nachgestellt werden. Dazu wurde ein spezieller Messstand entwickelt, der durch den optischen Grundaufbau von FTIR-Spektrometern bzw. ähnlichen Aufbauten in der Literatur inspiriert wurde [6], [10]. Mit ihm ist es möglich die Detektoren über einen weit einstellbaren Bereich mit bis zu

100 mW CW-Strahlungsleistung zu beaufschlagen, diese Leistung präzise auf einen definierten Messfokus variabler Größe ($\varnothing$ 0.5 ... 3 mm) – angepasst auf verschiedene Elementgrößen – abzubilden und dabei die Gehäusetemperatur des Detektors über einen weiten Bereich (10 ... 80°C) zu variieren. Das System ermöglicht die Untersuchung von Detektoren hinsichtlich ihrer Linearität und Stabilität sowie die Bewertung ihres Temperaturverhaltens und ihrer Langzeitstabilität unter Dauerbelastung.

Abb. 10 zeigt den schematischen Messaufbau: Als Quelle dient ein keramischer CW-Strahler identischer Bauart wie in FTIR-Spektrometern verwendet. Über Off-Axis-Parabolspiegel wird eine zweistufige Abbildung von der Quelle zum Detektor realisiert. Eine austauschbare präzise Lochblende (Pinhole) steuert dabei die Größe des Fokus in der Detektorebene. Über eine motorisierte Irisblende kann die Strahlöffnung und damit die Gesamtstrahlungsleistung gesteuert werden. Mit Hilfe eines Choppers kann der CW-Strahl für die Justage moduliert werden. Ein Shutter erlaubt es eine starke Wechselbeanspruchung zu simulieren. Die optische Kleinsignalanregung, welche im FTIR durch das Interferometer zustande kommt, wird im Testaufbau durch eine zweite unabhängige Quelle realisiert. Hierfür wurde eine analog modulierbare NIR-SLED (DenseLight Semiconductors, $\lambda = 1,6 \mu\text{m}$) ausgewählt. Ihre Strahlung wird über einen Faserkollimator und durch ein Loch im zweiten Spiegel eingekoppelt und auf den Detektor gerichtet. Der Detektor ist in einer (mit Peltierelementen) temperaturgeregelten Messfassung gehalten, und kann präzise über einen XY-Kreuztisch einjustiert werden. Die rauscharme Signalverstärkung erlaubt es Signal- und Rauschmessungen simultan auszuführen. Mit der zugehörigen Software können nahezu beliebige Messabläufe innerhalb der genannten Grenzen programmiert und automatisch abgefahren werden.

Abb.11 zeigt ein beispielhaftes Messergebnis. Hier wurde die Simulation aus Abb.7 im Experiment nachstellt, wobei die Ergebnisse im Wesentlichen bestätigt wurden. Zum Start des Versuchs ($t = 0 \text{ min}$) befindet sich der Prüfling auf Raumtemperatur (dunkelrote Kurve, $T = 23^\circ\text{C}$), die CW-Strahlung ist vollständig geblockt (grüne Kurve, $P = 0 \text{ mW}$). Auf den initialen Wert der Signalstärke (relative Empfindlichkeit, blaue Punkte) wird für den weiteren Verlauf normiert.

Im ersten Schritt des Experiments wird die Strahlungsleistung auf ein sehr niedriges

Grundniveau eingestellt und die Temperatur wird auf 30°C erhöht. Die Empfindlichkeit erhöht sich hier zunächst um einige Prozentpunkte. Im weiteren Verlauf wird nun die Bestrahlung schrittweise jeweils um einige Milliwatt erhöht. Die Empfindlichkeit verläuft im Wesentlichen flach mit einem sehr geringen Anstieg bis im Bereich von 30...40 mW ein Maximum erreicht wird.

Für den weiteren Verlauf des Experiments wurde die Bestrahlungsleistung bewusst deutlich über die übliche Anwendungsbedingungen erhöht, um das Verhalten des Detektors in diesem eigentlich unzulässigen Bereich zu analysieren und die Simulationsergebnisse zu verifizieren: Bei weiterer Erhöhung der Bestrahlung nimmt die Empfindlichkeit stark ab und bricht schließlich oberhalb von 55 mW komplett ein.

Wenn die Bestrahlung anschließend wieder zurückgesetzt wird, kann nach einiger Zeit eine Erholung beobachtet werden: Das Element kühlt wieder unter die Curietemperatur ab und das intrinsische Feld (Biasfeld) im DLaTGS-Material sorgt dafür, dass sich das Element wieder selbst polt. Dieser Prozess dauert einige Zeit und ist im gezeigten Experiment noch nicht vollständig abgeschlossen. Dieses Verhalten ist materialbedingt und lässt sich grundsätzlich nicht vollständig vermeiden. Vielmehr ist stets auf die richtige thermische Auslegung des Detektors und des Gesamtsystems zu achten, um die Entpolung des Sensorelements durch Überhitzung von vornherein auszuschließen.

Fazit

Moderne Technologien konsequent angewendet erlauben es, vermeintlich ausentwickelte Produkte (alte Hüte) auf ein neues, bisher nicht erreichtes Qualitätsniveau zu heben. FEM-Simulationen verkürzen den Entwicklungsprozess und erlauben sehr präzise Vorhersagen, welche sich mit entsprechender Messtechnik im Experiment bestätigen lassen. Moderne Methoden und Ausrüstungen für Prozessoptimierung und Überwachung lassen reproduzierbarere Ergebnisse und höhere Ausbeuten erzielen. Der vorliegende Beitrag lieferte ein paar punktuelle und ausgewählte Einblicke zur Untermauerung dieser Thesen und belegt dabei, dass der pyroelektrische DLaTGS-Detektor weiterhin eine wichtige Rolle in der IR-Messtechnik spielt.

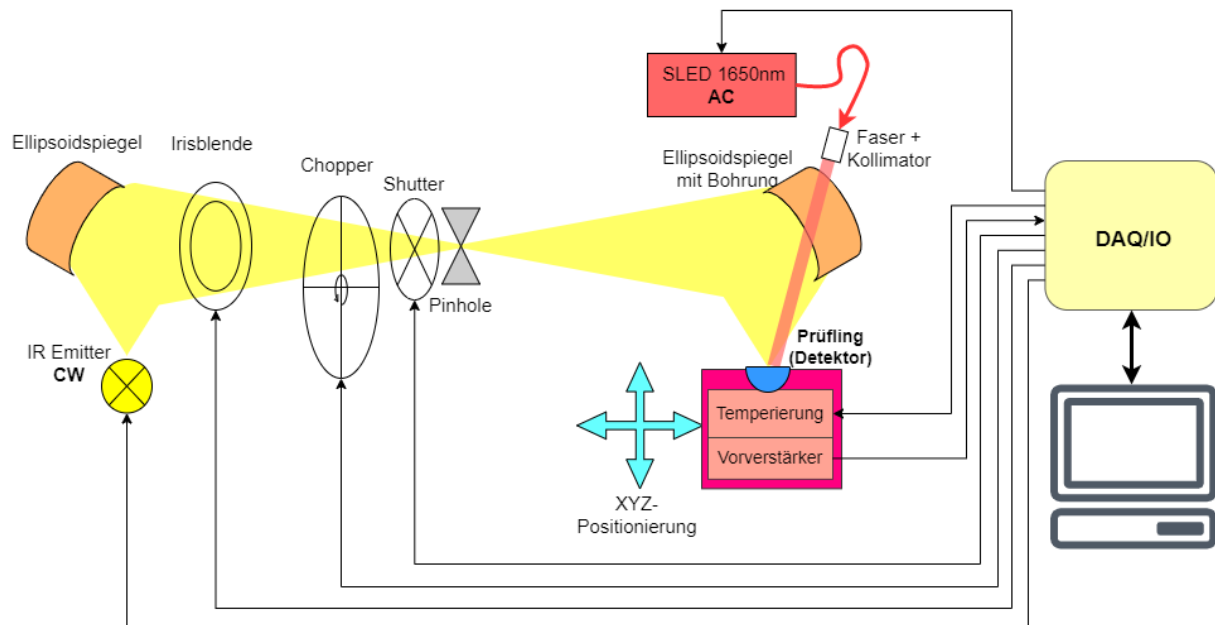


Abb. 10: Schema des Messstandes für Linearitäts-, Stabilitäts- und Lebensdauerprüfung an DLaTGS-Detektoren, Eigenentwicklung von InfraTec.

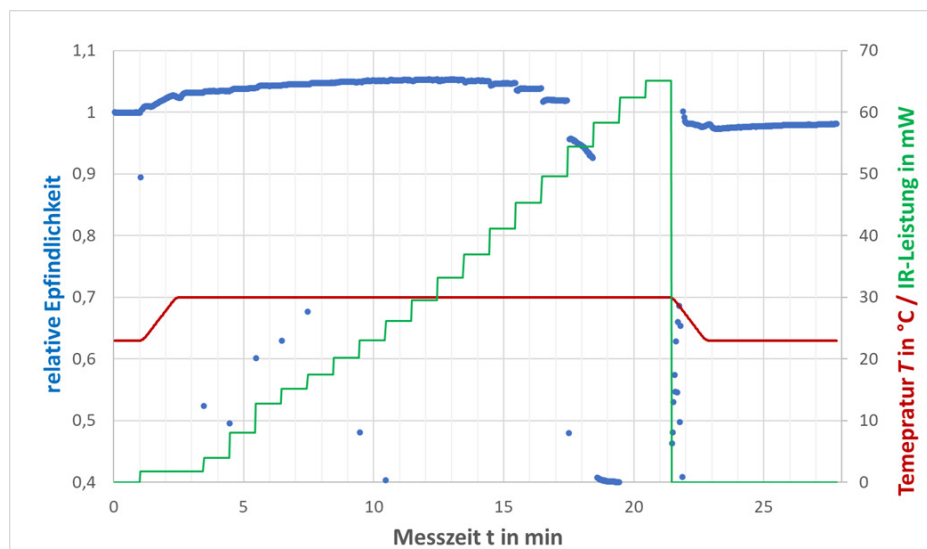


Abb. 11: Messdaten aus einem beispielhaften Experiment zur Untersuchung der Linearität und Stabilität eines DLaTGS-Detektors bei steigender CW-Bestrahlungsleistung mit dem in Abb. 10 gezeigten Messaufbau, vgl. Simulation aus Abb. 7.

Literaturnachweis

- [1] Whatmore, R. et al., *J. Appl. Phys.* 133, 080902 (2023); [doi: 10.1063/5.0141044](https://doi.org/10.1063/5.0141044)
- [2] Günther, A. et al., *Proceedings IRS² 2017*, S. 804–808 (2017), [doi: 10.5162/irs2017/i4.5](https://doi.org/10.5162/irs2017/i4.5)
- [3] Oettel, F. (1978): Pyroelectric detector having improved suspension means, [US4218620A](https://doi.org/10.1117/3.523499). G01J 1/00.
- [4] Saptari, V. (2003): Fourier-Transform Spectroscopy Instrumentation Engineering. Bellingham, Wash: SPIE Press (61). [doi: 10.1117/3.523499](https://doi.org/10.1117/3.523499), ISBN: 9780819451644
- [5] Neumann, N.: *Ferroelectrics* 142 (1), S. 83–92. (1993); [doi: 10.1080/00150199308237885](https://doi.org/10.1080/00150199308237885).
- [6] Kaplan, S.G. et al., *Appl Opt.* 1997 Dec 1;36(34):8896–908 (1997), [doi:10.1364/ao.36.008896](https://doi.org/10.1364/ao.36.008896)
- [7] Novotny, J. et al., *Cryst. Res. Technol.* 39, No. 12, 1089 – 1098 (2004); [doi: 10.1002/crat.200410294](https://doi.org/10.1002/crat.200410294)
- [8] Neuzner, A. et al., Pyroelectric detector for EE9 FORUM: design and characterization. In: Proc. SPIE 12777 (2022). 134. [doi: 10.1117/12.2690559](https://doi.org/10.1117/12.2690559)
- [9] Owton, D. et al., Current infra-red developments within Leonardo. In: Proc. SPIE 1268706 (2023) [doi: 10.1117/12.2677446](https://doi.org/10.1117/12.2677446)
- [10] Theocharous, E., *Applied optics* 47 (21), S. 3731–3736. (2008). [doi: 10.1364/AO.47.003731](https://doi.org/10.1364/AO.47.003731)

Tuned Sideband Diode Laser Absorption Spectroscopy

Kaigi Wu¹, Stefan Palzer¹

¹Professur Sensorik, TU Dortmund, Dortmund, 44227, Deutschland

Abstract

Tunable Diode Laser Absorption Spectroscopy allows for real-time, non-intrusive gas detection and concentration measurements. In conventional implementations it relies on determining the laser frequency during tuning, which hampers precise reconstruction of the absorption line due to non-linear effects. In this work, a method utilizing precisely controlled, frequency varying first order sidebands to probe an absorption line is demonstrated. A calibration procedure for determining the frequency modulation depth is introduced and implemented experimentally. The results demonstrate good agreement between simulation and measurement, highlighting the potential of sideband-sweeping for implementing accurate laser absorption spectroscopy.

Keywords: TDLAS, frequency modulation, sidebands sweeping, gas concentration

Introduction

Tunable Diode Laser Absorption Spectroscopy (TDLAS) is a well-established, non-invasive technique for detecting trace gases with high sensitivity and selectivity [1]. It uses the frequency tuning capabilities of a diode laser to probe molecular absorption features by adjusting either the injection current or the operating temperature of the laser. The resulting absorption signal is described by the Beer–Lambert–Bouguer law, which directly relates the attenuation of light intensity to the molecular number density along the optical path [2].

The accurate reconstruction of the absorption line is critical for determining gas concentration in TDLAS setups. However, the mechanisms of frequency tuning via injection current and thermal modulation are inherently complex, making precise determination of the laser frequency during tuning challenging [3]. One approach to address this issue involves using the line shape signal generated by frequency modulation (FM) as an absolute frequency reference [4]. In this method, the zero-crossing point of the line shape signal corresponds to the molecular resonance, while the two extrema on either side are symmetrically positioned at a distance equal to the modulation frequency from the resonance, effectively forming a frequency scale. Nevertheless, the resolution and accuracy of this calibration technique are limited by factors such as the characteristics of the phase shifter and nonlinearities introduced by intensity variations and it requires the use of photodiodes fast enough to resolve the sideband modulation.

Here, instead of directly tuning the laser frequency via the injection current the sideband frequencies of the laser are used to probe the absorption feature. These sidebands are generated through frequency modulation, producing two spectral components symmetrically positioned on either side of the carrier frequency [5]. The frequency offset of each sideband from the carrier is precisely equal to the modulation frequency. By varying this modulation frequency, the sidebands can be accurately tuned across the molecular absorption line without altering the carrier frequency itself. This method enables precise frequency control of the probing light and eliminates the need for high-speed photodiodes and heterodyne detection schemes. Instead, a standard photodiode is sufficient to measure the attenuation of the modulated light caused by sideband absorption.

Methodology

Laser frequency modulation has been extensively used and described for numerous applications [5]. When modulating the laser's central frequency ω , the resulting electric field of the emitted light can be expressed as $E = E_0 e^{i(\omega t + \beta \sin \Omega t)}$, where β is the modulation depth and Ω is the modulation frequency. This expression can be expanded using Bessel functions to:

$$E \approx E_0 [j_0(\beta) e^{i\omega t} + j_1(\beta) e^{i(\omega+\Omega)t} - j_1(\beta) e^{i(\omega-\Omega)t}], \quad (1)$$

with $j_0(\beta)$ and $j_1(\beta)$ being the zero order and first order Bessel functions, respectively, representing the amplitude contribution of the carrier and first sidebands for modulation

depths β sufficiently small to allow for disregarding higher order sidebands. When the modulation depth is small, the majority of the optical power is distributed among the carrier and the first-order sidebands, while the contribution from higher-order sidebands becomes negligible, i.e. $j_0^2(\beta) + 2j_1^2(\beta) \approx 1$. The method proposed here is demonstrated using a molecular oxygen absorption line as illustrated in Fig. 1 exhibiting a half width at half maximum (HWHM) width of 1.51 GHz at 1 bar and 296 K.

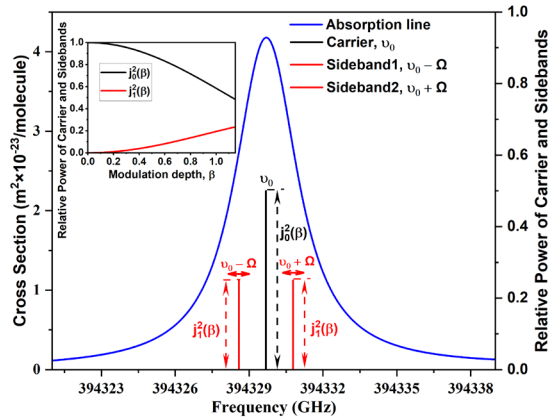


Fig. 1: Illustration of an absorption line (blue) at 1 bar and 296 K and the laser frequency components: the carrier is locked at the resonance center, while two first order sidebands are symmetrically positioned on either side of the carrier frequency.

As the modulation frequency increases, the sidebands are tuned away from the resonance, transitioning from regions of strong absorption to those with lower absorption. Consequently, the total absorption decreases, and the transmitted optical power increases. As the modulation depth increases, the relative power in the first order sidebands increases, while the relative power in the carrier decreases. This redistribution of power results in reduced total absorption, leading to an increase in transmitted optical power. The change in transmitted power is captured using a standard photodiode without the need for signal demodulation. Due to the precise frequency control of the probing sidebands, the detected photodiode signal can be analytically modeled using the Beer–Lambert–Bouguer law. This enables accurate reconstruction of the absorption line, facilitating precise quantification of the gas concentration.

Results

The calibration of the modulation depth can be performed by sweeping the modulation power while keeping the modulation frequency fixed and recording the corresponding photodiode signal. Increasing the modulation power results

in a higher photodiode signal in the presence of target gas molecules. To enable accurate calibration, the photodiode signal under unmodulated conditions must also be recorded. After calibration, the absorbance $\alpha(\nu_0)$ and $\alpha(\nu_0 - \Omega)$ corresponding to each swept modulation frequency can be determined. As shown in Fig. 2, the simulated absorbance and the experimentally measured values for different modulation frequencies exhibit good agreement. Notably, the determination of absorbance is independent of the total laser power P_0 , which eliminates the influence of power fluctuations caused by laser aging or contamination. This enhances the stability and accuracy of the proposed method further.

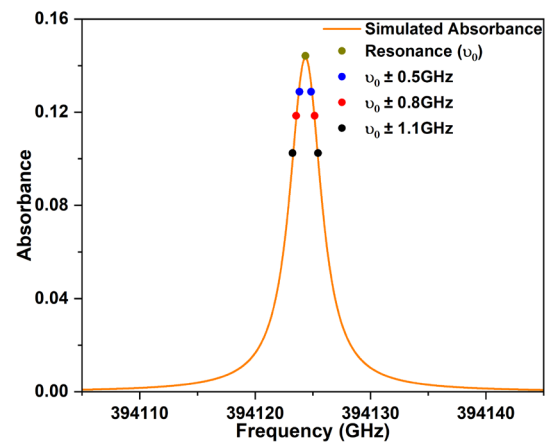


Fig. 2: Comparison between simulated and measured absorbance. The simulation is under known conditions: number density is 2.599×10^{25} molecules/m³, the effective absorption length is 1.05 m, temperature is 296 K and pressure is 1 bar.

References

- [1] P. Werle, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 54, no. 2, pp. 197–236, 1998, doi: 10.1016/S1386-1425(97)00227-8.
- [2] V. Shumakova and O. H. Heckl, *APL Photonics*, vol. 9, no. 1, 2024, doi: 10.1063/5.0167683.
- [3] K. Wunderle *et al.*, *Applied optics*, vol. 48, no. 4, B172-82, 2009, doi: 10.1364/ao.48.00b172.
- [4] K. Wu, M. Jakubowsky, and S. Palzer, *optics letters*, vol. 50, no. 11, pp. 3612–3615, 2025, doi: 10.1364/OL.557982.
- [5] G. C. Bjorklund, *optics letters*, vol. 5, no. 1, p. 15, 1980, doi: 10.1364/ol.5.000015.

Resonant-Akustischer Sensor für die gleichzeitige Detektion von Sauerstoff und Kohlenstoffdioxid im Lebensmittelbereich

Christian Weber^{1,2}, Katrin Schmitt^{1,2}, 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.

Zusammenfassung

Physikalische Verfahren wie die Messung der Schallgeschwindigkeit erlauben es in der Regel nur die Komponenten binärer Gasgemische voneinander unterscheiden zu können. In diesem Beitrag stellen wir eine neuartige Methode vor, bei der die photoakustische Schallerzeugung an der Wand eines akustischen Resonators genutzt wird, um sowohl eine zur Schallgeschwindigkeit proportionale Resonanzfrequenz als auch eine vom Gas abhängende Amplitude zu erhalten. Messungen haben gezeigt, dass sich Amplitude und Resonanzfrequenz einem Stickstoff-Sauerstoff-Gemisch anders zueinander verhalten als im Stickstoff-Kohlendioxid-Gemisch, womit eine Trennung der Komponenten im ternären Gasgemisch möglich wird. In Übereinstimmung mit der Theorie weist die Resonanzfrequenz eine sehr gute Linearität auf; bei der Amplitude ist dies ebenfalls für Sauerstoff der Fall, nicht jedoch für CO₂, was auf einen zusätzlichen Dämpfungseffekt durch CO₂ hindeutet. Die rauschäquivalenten Konzentrationen liegen für CO₂ und O₂ über die Resonanzfrequenz bei 72 ppm bzw. 140 ppm und über die Amplitude bei 0.08 vol.-% bzw. 0.16 vol.-%.

Keywords: Schallgeschwindigkeit, Resonanz, Gassensor, Sauerstoff, Kohlendioxid, Photoakustik

Motivation

In der Lebensmittelindustrie stellen die Konzentrationen von Kohlenstoffdioxid (CO₂) und Sauerstoff (O₂) bzw. deren Verhältnis in Stickstoff (N₂) einen der wichtigsten Parameter für die Alterung verpackter Lebensmittel dar [1]. Zudem lässt sich über das Verpacken der Lebensmittel unter definierten gehalten von CO₂ und O₂ (sogenannte Schutzatmosphäre) die Lebensmittelhaltbarkeit signifikant verlängern und damit die Lebensmittelsicherheit verbessern [2]. Somit ist für die Kontrolle der ordnungsgemäßen Verpackung von Lebensmitteln sowie für die stichprobenartige Kontrolle der Bedingungen innerhalb der Verpackung Sensorik erforderlich.

Die aktuellen Lösungen zur Messung der Konzentration basieren hauptsächlich auf filterphotometrischen Verfahren zur Messung von Kohlenstoffdioxid und von elektrochemischen Sensoren bzw. potentiometrischen Sensoren für die Sauerstoffmessung. Neben Nachteilen wie der begrenzten Lebensdauer der elektrochemischen Sensoren oder dem recht großen Probenvolumen, das für die Sensorkombination erforderlich ist, ist dieser Ansatz recht aufwändig und teuer. Dies erschwert den großflächigen Einsatz der Sensorik – beispielsweise zur kontinuierlichen Überwachung von

Großverpackungen und die Kontrolle der Verpackungsatmosphäre ist damit hauptsächlich auf Stichprobenmessungen im Labormaßstab limitiert. Um die Sensorik für die Bestimmung von CO₂ und O₂ zu vereinfachen, schlagen wir ein neues resonant-akustisches Messverfahren vor, welches auf einer außerordentlich linearen und störungsfreien Schallerzeugung mittels des photoakustischen Effekts basiert.

Messprinzip und Sensoraufbau

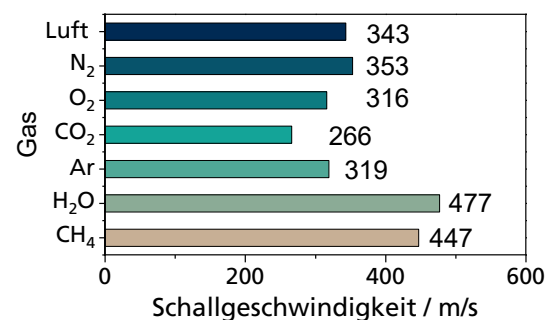


Abb. 1: Vergleich der Schallgeschwindigkeit in verschiedenen Gaskomponenten der Umgebungsluft

Die Schallgeschwindigkeit verschiedener Gaskomponenten unterscheidet sich, wie in Abbildung 1 dargestellt, teils erheblich. Mischungen der Gaskomponenten besitzen somit auch eine

individuelle Schallgeschwindigkeit, die vom Mischungsverhältnis abhängt. Sie errechnet sich allgemein über (1) [3]:

$$c = \sqrt{RT \cdot \frac{\sum_i^N x_i \gamma_i}{\sum_i^N x_i M_i}} \quad (1).$$

c ist dabei die Schallgeschwindigkeit der Mischung, T die Temperatur und R die universelle Gaskonstante. M_i sind die molaren Massen und γ_i die Isentropenexponenten der einzelnen Gas-komponenten. Ist die Schallgeschwindigkeit bekannt, so lässt sich somit auch auf das Mischungsverhältnis einer binären Gasmischung zurückschließen. Da die Schallgeschwindigkeit aber stets die Summe der Schallgeschwindigkeiten aller im Gemisch enthaltenen Gase ist, ist die Eindeutigkeit bei Mischungen mit drei oder mehr Gasen nicht mehr gegeben. Somit benötigt die Analyse von Gasgemischen mit mehreren relevanten Komponenten stets weitere Messmethoden, um die Schallgeschwindigkeit sinnvoll einsetzen zu können.

Für die Bestimmung der Schallgeschwindigkeit existieren eine Reihe an unterschiedlichen Verfahren. Dabei gibt es einerseits lauffzeitbasierte Verfahren, die die Schallgeschwindigkeit über den zeitlichen Versatz einer Schallwelle entlang eines Weges entweder über die Phasenlaufzeit eines modulierten Signals oder über die Gruppenlaufzeit eines Pulses bestimmen [4]. Andererseits gibt es auch resonatorbasierte Verfahren, bei denen die Schallgeschwindigkeit indirekt über die Frequenz eines akustischen Resonators bestimmt wird [5]. Eine der einfachsten Geometrien eines akustischen Resonators ist ein Hohlzylinder, der das zu analysierende Gas enthält.

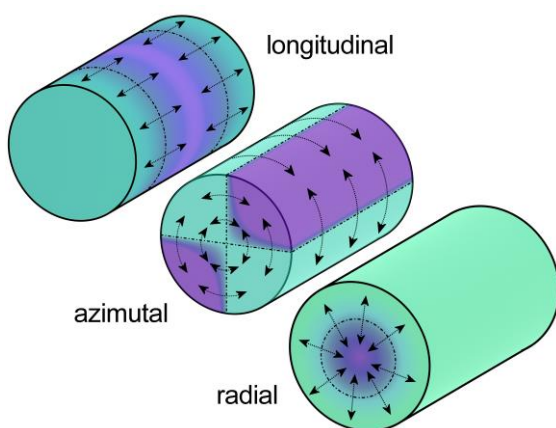


Abb. 2: Grafische Darstellung der möglichen resonanten Moden in einer Zylindergeometrie. Blau und Grün zeigen Druckmaxima bzw. -minima und die Pfeile zeigen die Auslenkungsrichtung der Schwingung an.

In einem Zylinder lassen sich verschiedene Zylindermoden anregen. Sie hängen bei einem Zylinder, dessen Gehäusematerial eine ausreichend hohe akustische Impedanz besitzt, in erster Näherung nur von der Geometrie und der Schallgeschwindigkeit der Gasfüllung ab. Die möglichen Moden, die im Zylinder anschwingen können, lassen sich in drei Kategorien unterteilen: longitudinale, azimuthale und radiale Moden. Diese sind in ihrer Grundform in Abbildung 2 grafisch dargestellt.

Longitudinale Moden schwingen in ihrer Auslenkung entlang der Zylinderachse, azimuthale Moden entlang der Mantelfläche und Radialmoden von der Mantelfläche zur Achse des Zylinders. Eine hohe Güte ist insbesondere bei der Radialmode erreichbar, weshalb sie sich besonders gut zur Messung der Schallgeschwindigkeit eignet.

Während in der Literatur oft piezoelektrische Elemente, Lautsprecher [4,6] oder auch andere elektromechanische Wandler [6] verwendet werden, kann der Schall auch photoakustisch erzeugt werden [7]. Dabei wird Licht auf einen absorbierenden Teil der Zelle geleitet, welches dort absorbiert wird, und dessen Energie sich in Wärme umwandelt. Die Wärme geht dann in die Gasfüllung über, und sorgt für eine Ausdehnung des Gases.

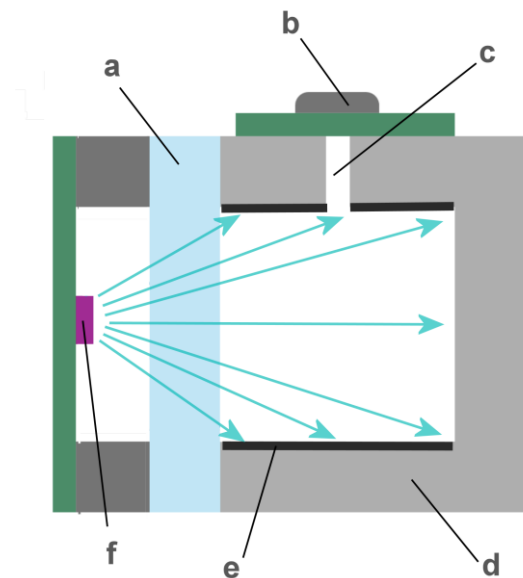


Abb. 3: Querschnitt durch das Sensorprinzip der photoakustischen Signalerzeugung kombiniert mit einem Zylinderresonator für eine Radialmode. Mit a) einem Fenster, b) MEMS-Mikrofon, c) Schallzugang (Bohrung), d) Aluminiumgehäuse mit zylindrischer Kavität, e) Farbschicht zur Erhöhung der Absorption, f) SMD-LED (450nm).

Bei gepulstem Licht treten die resultierenden wechselnden Druckschwankungen als Schallwellen auf, die den Resonator anregen kann. Ein großer Vorteil bei dieser Art der Schallerzeugung, liegt darin, dass der Schallwandler komplett frei von Eigenresonanzen ist. Das Verfahren ist linear im Bereich von wenigen Hertz bis weit in den Ultraschallbereich. Um mit diesem Verfahren einen Sensor aufzubauen, benötigt man also eine Lichtquelle, die auf einen Wandabschnitt des Zylinderresonators fokussiert werden kann. Besonders einfach ist dies für die Radialmode, da hier das Überlapp-Integral zwischen Anregung und Schallwelle auf der gesamten Innenmantelfläche besonders hoch ist. Im einfachsten Fall kann man damit eine LED, wie in Abbildung 3 gezeigt, frei in den Halbraum abstrahlend vor dem Zylinderresonator befestigt werden. Zur weiteren Erhöhung des Signals, wurde die Mantelfläche des Zylinders mit schwarzer Farbe (MTK 25, Edding group) beschichtet. Da die Absorption, welche auf der Stirnseite der Zylinderkavität stattfindet, das Überlappintegral verringern würde, wurde hier keine absorbierende Farbe aufgebracht. Minimale Verringerung durch die Absorption des blanken Aluminiums wurde dabei in Kauf genommen. Für den Durchmesser des Radialresonators wurden 10mm ausgewählt, mit dem die Frequenz der ersten Radialmode in etwa 42 kHz betragen sollte. Eine Länge von 8 mm sorgt dafür, dass die störenden Longitudinalmoden abseits der Radialmode liegen, sofern sie angeregt werden. Als LED kommt der Typ „GD VJLPL1.14-UNUO-34-1“ (ams-Osram AG, Österreich) zum Einsatz. Diese LED besitzt eine strahlende Fläche von 1.1 x 1.1 mm bei einer

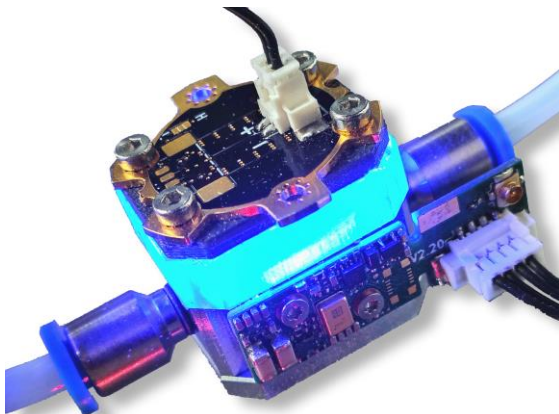


Abb. 4: Foto des fertig aufgebauten Sensors im Messbetrieb. Oben ist die Leiterplatte zu sehen, welche als Träger für die LED fungiert, und vorne ist die Leiterplatte mit dem Mikrofon zu sehen.

Zentralwellenlänge von 463 nm. Sie besitzt keine weitere package-interne Optik, was ein nahezu Lambertsches Strahlprofil zur Folge hat. In Abbildung 4 ist ein Foto des aufgebauten Sensors dargestellt. Die Detektion des Schalls erfolgt über ein MEMS-Mikrofon von Typ „IM73A135V01“ (Infineon AG, Deutschland). Das Mikrofonsignal wird Bandpassgefiltert mit einem Durchlassbereich zwischen 10 und 40 kHz und um den Faktor 22 verstärkt. Es sitzt auf einer dezidierten Analogplatine, welche in Abbildung 5 detailliert gezeigt ist. Das Mikrofonsignal wird anschließend mit einer vollständig differenziellen Konfiguration umgesetzt und über eine Kabelverbindung zu einem Embedded System, welches die Auswertung übernimmt. Dieses basiert auf einem PSoC6 Microcontroller und beinhaltet außerdem einen ADC, Power-Management für die LED-Spannung und die Analogversorgung, einen LED-Stromtreiber und ein USB-Interface, um die Daten über eine virtuelle serielle Schnittstelle an einen PC zu übertragen. Die Datenerfassung

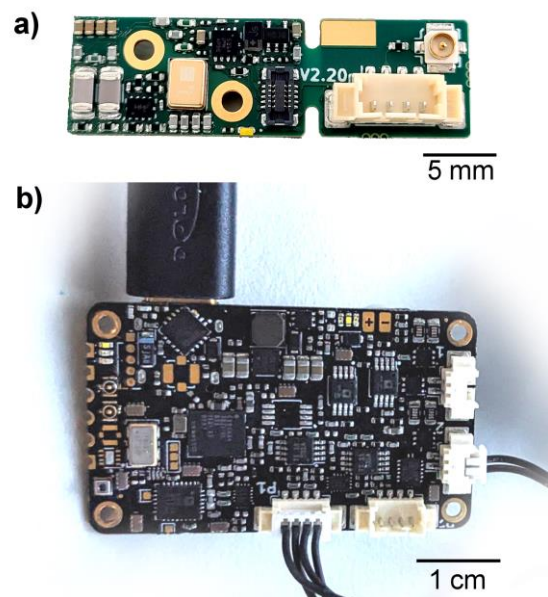


Abb. 5: a) Foto der analogen Vorverstärkerplatine mit MEMS-Mikrofon, einem Bandpassfilter und voll-differenzieller Treiberstufe. b) Embedded System zur Auswertung der Mikrofonsignale und zur Ansteuerung der LED.

findet mit 1024 kSa/s statt und mit der gleichen Frequenz erfolgt das Sampling der Ausgangskurvenform für den Stromtreiber. Die aufgenommenen Mikrofondaten werden von einem Sample-by-Sample Lock-In Filter, wie in [8] beschrieben, gefiltert.

Um die Resonanzfrequenz zu finden, sucht der Sensor iterativ nach der höchsten Amplitude und nimmt dann im Bereich der Resonanzfrequenz hochauflösende Daten auf, an die ein Resonatormodell angefügt wird [7]. Aus dem

Modell wird dann die Resonanzfrequenz errechnet und an einen Host-PC übertragen. Die Parameter des Suchalgorithmus wurde dabei so ausgewählt, dass das Finden der Resonanzfrequenz etwa 1100 ms benötigt. Als zusätzlicher Parameter wurde, anschließend auf der Resonanzfrequenz für 500 ms die Amplitude des akustischen Signals bestimmt. Zusätzlich wurden auch weitere Parameter des Fits ausgegeben. Messdaten standen somit in einem Intervall von 1.6 s zur Verfügung. Die Ansteuerung der LED erfolgte in beiden Messschritten sinusförmig mit 400 mA Peak-to-Peak.

Versuchsaufbau

Der wie zuvor beschriebene konfigurierte Sensor wurde anschließend am Gasmessplatz des Fraunhofer IPM vermessen. Dazu wurden M3 4mm IQS-Gasanschlüsse an den Sensor angebracht. Über Massenflussregler wurde ein relatives Verhältnis der einzelnen Gaskomponenten eingestellt (Prinzip in Abbildung 6 dargestellt).

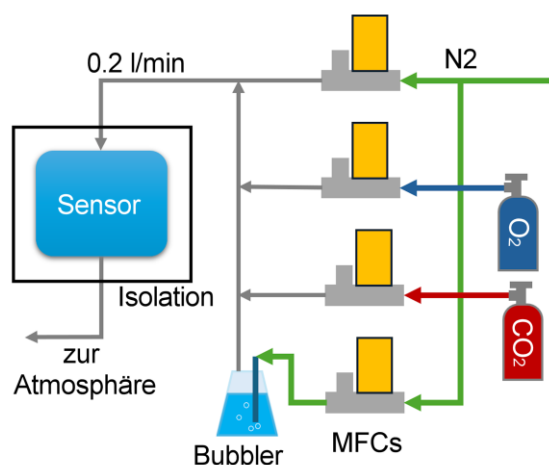


Abb. 6: Prinzipzeichnung des verwendeten Gasmessplatzes und Versuchsaufbau mit dem Sensor.

Die Gase CO_2 und O_2 wurden aus Druckgasflaschen bereitgestellt, die jeweils 100 % des jeweiligen Gases enthielten, während Feuchtigkeit über eine Bubbler-Flasche dem Stickstoffstrom zugemischt werden kann. Die Flüsse wurden während allen Versuchen so relativ zueinander geregelt, dass sich stets ein Summenfluss von 0.2 l/min einstellt. Da der Sensor eine geringe thermische Masse hat und sich im Betrieb erwärmt, wurde zudem ein Karton über dem Sensor platziert. Dieser soll eine allzu starke Anströmung mit Umgebungsluft verhindern, was zu starken Temperaturschwankungen führen könnte.

Messergebnisse und Diskussion

In einer ersten Messung wurde das Ansprechen des Sensors auf verschiedene Konzentrationen von CO_2 und Sauerstoff vermessen. Dazu wurden je drei Stufen von Sauerstoff und CO_2 in 10% Schritten gemischt und der Sensor damit beaufschlagt. Das Ergebnis dieser Messung ist in Abbildung 7 gezeigt. Sowohl die Amplitude bei akustischer Resonanz als auch die Resonanzfrequenz selbst zeigen eine deutliche Reaktion auf die Gasstufen. Wie erwartet, fällt die Resonanzfrequenz bei beiden Gasen mit zunehmender Konzentration ab. CO_2 verschiebt die Resonanzfrequenz um -0,22 %/Vol-% und Sauerstoff um -0,0088 %/Vol-%, was gut mit den theoretischen Vorhersagen übereinstimmt. Die Amplitude zeigt bei CO_2 ebenfalls ein fallendes Verhalten, bei Sauerstoff steigt die Amplitude jedoch an. Durch das unterschiedliche Vorzeichen sollten sich beide Gase voneinander unterscheiden lassen und auch die Bestimmung von tertiären Gasgemischen möglich

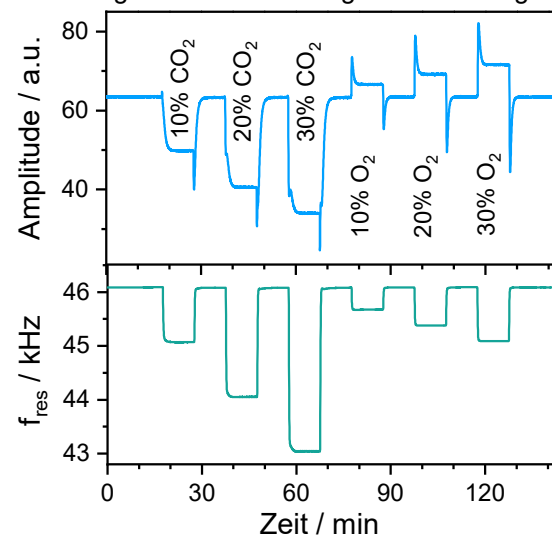


Abb. 7: Messergebnis bei Beaufschlagung von CO_2 und Sauerstoff. Abgebildet sind die bestimmte Resonanzfrequenz sowie die Amplitude des akustischen Signals bei der Resonanzfrequenz. Im oberen Panel sind die Konzentrations-Sollwerte von CO_2 und Sauerstoff gezeigt.

machen. Durch die direkte Platzierung des Sensors im Gasfluss und das geringe Volumen, ist die Reaktionszeit des Sensors sehr schnell. Die Zeitkonstante für die Reaktion der Resonanzfrequenz liegt bei 4.9 s bei der Reaktion auf CO_2 und bei 5.7 s bei der Reaktion auf Sauerstoff. Bei der photoakustischen Amplitude sind die Reaktionen auf die Gase komplexer (siehe Abbildung 8). Bei Sauerstoff reagiert die Amplitude sehr schnell auf die Gasänderung ($\tau = 6$ s), gefolgt von einer langsameren Rückreaktion geringerer Stärke ($\tau = 25$ s). Dies stammt vermutlich aus einer Überlagerung von zwei Effekten:

Einmal die Änderung der thermodynamischen Eigenschaften im Zylinderresonator selbst und mit einer langsameren Zeitkonstante, die Änderung der viskoelastischen Eigenschaften im Mikrofon selbst bzw. die akustische Kopplung des Mikrofons an das Gas. Diese Effekte sind auch bei CO_2 zu beobachten; in diesem Fall könnte jedoch auch die Dämpfung des Resonators selbst eine Rolle spielen, was zu einem

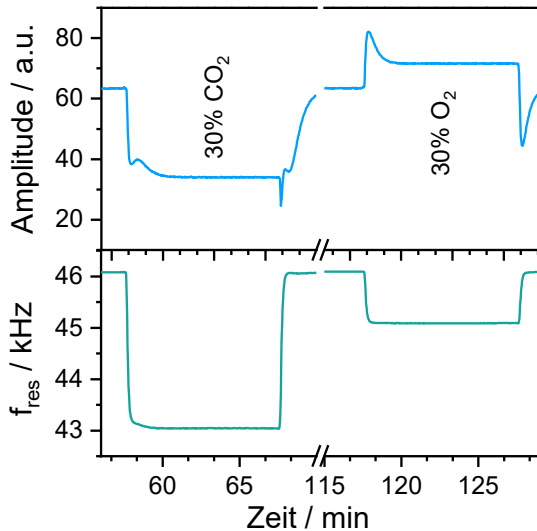


Abb. 8: Detail der Sensorreaktion auf die 30% Stufen von CO_2 und Sauerstoff. Die Amplitude der photoakustisch erzeugten Schallwelle zeigt ein komplexeres Verhalten auf den Gaswechsel als die Schallgeschwindigkeit, welches vermutlich durch eine Überlagerung verschiedener Effekte zustande kommt.

weiteren Effekt führt, der sich bei den schnellen Gaswechseln überlagert. Ohne weitere Kompensation limitieren diese Effekte die Ansprechzeiten des Sensors, da stets die längste Zeitkonstante im System die Zeit bis zu einem stabilen Messwert definiert. Da Werte wie die Dämpfung des Resonators und Mikrofonkopplung auch eine Frequenzabhängigkeit zeigen, könne unter Hinzunahme von weiteren akustisch anregbaren Moden in Zukunft eine Kompensation dieser Effekte erfolgen. Auch eine Unterscheidbarkeit quartärer oder höherer Gasgemische durch die saubere Trennung dieser Effekte wäre denkbar. In der nächsten Messung wurde die Linearität des Sensors bestimmt. Zu diesem Zweck wurde der Sensor mit Konzentrationen im Bereich von 0 bis 30 Vol.-% beaufschlagt, und die Resonanzfrequenz sowie die akustische Amplitude wurden erneut aufgezeichnet. Die Konzentrationen wurden in Schritten von 2 Vol.-% erhöht, wobei der Schritt bei 10 Vol.-% ausgelassen wurde. Bei Sauerstoff war eine 2 Vol.-% Stufe technisch nicht möglich. Die Ergebnisse sind in Abbildung 9 dargestellt.

Das Sensorsignal der Resonanzfrequenz deckt sich nahezu ideal der Theorie. Für beide

Signale sind die Abweichung vom linearen Zusammenhang kleiner als 0.01 % und damit weit geringer als die spezifizierte Linearität der Massenflussregler. Lediglich der reine Stickstoffwert zeigt bei beiden Gasen eine größere Abweichung, was vermutlich an einem Offsetverhalten des Massenflussreglers liegt.

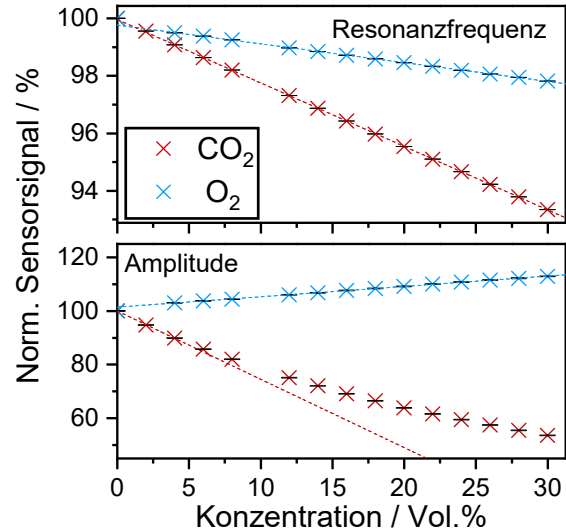


Abb. 9: Verlauf der Sensorsignale auf verschiedene Konzentrationen von O_2 und CO_2 . Die gestrichelten Linien geben den linearen Zusammenhang der ersten drei Punkte wieder, um die Linearität zu verdeutlichen.

Bei der Amplitude zeigt die Sensorreaktion auf Sauerstoff ebenfalls ein sehr lineares Verhalten mit Abweichungen kleiner der Standardabweichung der Messwerte. Die Reaktion auf CO_2 weist allerdings ein deutliches Absättigungsverhalten und damit eine starke Nichtlinearität auf. Auch dies deutet auf einen dritten überlagerten Effekt hin, der den Resonator dämpft.

Zuletzt wurde noch die Querempfindlichkeit des Sensors auf Wasser untersucht. Da das Gasgemisch aus Wasser, Stickstoff, Sauerstoff und CO_2 besteht – also ein quartäres Gasgemisch vorliegt, lässt sich die Wasserkonzentration nur mit der Amplitude und der Resonanzfrequenz allein nicht bestimmen und rechnerisch kompensieren. Daher muss die Wasserquerempfindlichkeit akzeptiert und die Konzentration entweder stabilisiert (Gastrockner o.Ä.) oder über einen zusätzlichen Feuchtesensor kompensiert werden. Für die Effektivität dieser Maßnahmen ist die Wasserquerempfindlichkeit entscheidend, da sie den Restfehler maßgeblich mitbestimmt, der nach einer Kompensation am Ende bestehen bleibt.

Die Messung erfolgt analog zu den vorangegangenen Messungen. Über eine Befeuchtung von Stickstoff (vergl. Abbildung 6) wurden 0-80% relative Feuchte bei 22 °C in 10 % Schritten generiert und auf den Sensor gegeben. Erneut wurden die Resonanzfrequenz und die

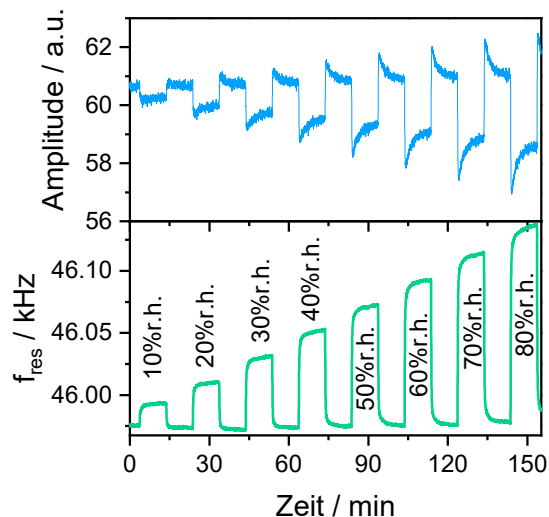


Abb. 10: Ergebnis der Querempfindlichkeitsmessung der Resonanzfrequenz und der Amplitude bei der Beaufschlagung des Sensors mit unterschiedlichen relativen Luftfeuchten bei 22 °C Umgebungstemperatur.

Amplitude aufgezeichnet. Das Ergebnis dieser Messung ist in Abbildung 10 dargestellt. Erneut deckt sich die Änderung der Schallgeschwindigkeit sehr gut mit den theoretischen Vorhersagen. Durch die geringe absolute Konzentration des Wassers ist der Effekt insgesamt gering. Er liegt bei etwa 2 Hz/%r.h. Dies entspricht einem Fehler von etwa 0.02 Vol.-%/%r.h. bei CO₂ und 0.05 Vol.-%/%r.h. bei Sauerstoff. Die Amplitude verhält sich ähnlich der Amplitudenänderung auf Sauerstoff mit der sichtbaren Überlagerung von zwei Effekten. Die zweite Zeitkonstante ist mit ($\tau = 118$ s) allerdings wesentlich länger, was durch die langsamere Bildung eines Gleichgewichts, mit dem die Oberflächen absättigenden Wasser erklärt werden kann. Auch bei der Amplitude sind die Effekte durch die kleinen absoluten Konzentrationen der Luftfeuchte gering. Sie liegt bei 0.08 a.u./%r.h. was Fehlern von 0.034 vol.-%/%r.h. bei CO₂ und 0.16 Vol.-%/%r.h. bei Sauerstoff entspricht. Eine Kompensation oder Stabilisierung der Luftfeuchte mit einer Genauigkeit von 5 %r.h. sollte also ausreichen, um den Fehler für Sauerstoff und CO₂ auf unter 1 Vol.-% zu reduzieren.

Zusammenfassung und Fazit

Zusammenfassend konnte gezeigt werden, dass der entwickelte resonant-akustische Sensor – der durch den photoakustischen Effekt Schall erzeugt – eine vielversprechende Alternative bisheriger Sensorik in der Lebensmittelindustrie darstellt.

Durch die Verwendung der akustischen Amplitude zusätzliche zur Resonanzfrequenz,

konnte ein zusätzlicher Parameter gefunden werden, der eine eindeutige Unterscheidung im tertiären Gasgemisch aus Stickstoff – Sauerstoff – Kohlenstoffdioxid möglich macht. In Gasmessungen konnten über die beiden Parameter eine rauschäquivalente Konzentration von 72 ppm bzw. 0.08 Vol.-% für CO₂ und 140 ppm bzw. 0.16 Vol.-% für Sauerstoff bei 1.6 s Messdauer bestimmt werden. Der Sensor weist eine inhärente Querempfindlichkeit zwischen Sauerstoff und Feuchte auf. Bei Raumtemperatur fallen die Fehler allerdings gering aus, so dass diese mit einem Polymer-Feuchtesensor kompensierbar sein sollten, ohne dass die Gesamtgenauigkeit zu stark beeinträchtigt wird. Dank seines kompakten und einfachen Designs sowie seiner simplen Signalauswertung lässt sich der Sensor auch unter beengten Platzverhältnissen, beispielsweise in handgehaltenen Messgeräten, einsetzen.

Literaturnachweis

- [1] Chaix, E., Couvert, O., Guillaume, C., Gontard, N. and Guillard, V. (2015), Predictive Microbiology Coupled with Gas (O₂/CO₂) Transfer in Food/Packaging Systems: How to Develop an Efficient Decision Support Tool for Food Packaging Dimensioning. COMPREHENSIVE REVIEWS IN FOOD SCIENCE AND FOOD SAFETY, 14: 1-21. <https://doi.org/10.1111/1541-4337.12117>
- [2] Genigeorgis, C. A. (1985). Microbial and safety implications of the use of modified atmospheres to extend the storage life of fresh meat and fish; International Journal of Food Microbiology, 1(5), 237-25;
- [3] Fomin, Y.D., Tsiok, E.N., Ryzhov, V.N. et al. Speed of Sound in Mixtures of Liquids: Noble Gases and Water-Alcohol Mixtures. Phys. Wave Phen. 33, 177–182 (2025). <https://doi.org/10.3103/S1541308X25700128>
- [4] J. S. Olfert, M. D. Checkel, C. R. Koch; Acoustic method for measuring the sound speed of gases over small path lengths. Rev. Sci. Instrum. 1 May 2007; 78 (5): 054901. <https://doi.org/10.1063/1.2736406>
- [5] Goodwin, A. R. H., and J. P. M. Trusler. "6 speed of sound." Measurement of the thermodynamic properties of single phases (2003): 237.
- [6] Segovia, J.J., Vega-Maza, D., Martín, M.C. et al. An Apparatus Based on a Spherical Resonator for Measuring the Speed of Sound in Gases and for Determining the Boltzmann Constant. Int J Thermophys 31, 1294–1309 (2010). <https://doi.org/10.1007/s10765-010-0746-4>
- [7] C. Weber, J. Kapp, J. Wöllenstein, K. Schmitt, Novel approach for efficient resonance tracking in photoacoustic gas sensor systems based on a light-induced wall signal, Photoacoustics, Volume 31, 2023, 100495, ISSN 2213-5979, <https://doi.org/10.1016/j.pacs.2023.100495>.
- [8] Weber, C. (2024). Miniaturized photoacoustic sensors for the detection of trace gases in ambient air.

Abbaubare Sensorknoten zur Überwachung der Bodenfeuchte im Smart Farming

Sebastian Böller¹, Benedikt Loepp¹, Thorben Greuter¹ und Anton Grabmaier²

¹Fraunhofer IMS, 47057 Duisburg, Germany

²Universität Duisburg-Essen, Fachgebiet Elektronische Bauelemente und Schaltungen

Zusammenfassung

Die präzise Bestimmung der Bodenfeuchte ist eine zentrale Voraussetzung für eine zielgerichtete Bewässerung in der modernen Landwirtschaft. Kabellose Sensorknoten ermöglichen dabei eine hohe zeitliche und räumliche Auflösung, verbleiben jedoch bei großflächigem Einsatz potenziell als Elektronikschrott im Feld. In diesem Beitrag wird ein abbaubarer Sensortransponder zur Bestimmung der Bodenfeuchte vorgestellt, der sich nach seiner Einsatzzeit umweltverträglich zersetzt und daher nicht wieder eingesammelt werden muss.

Das Konzept basiert auf einem passiven UHF-RFID-Sensortransponder, der ohne Batterie auskommt und somit einen besonders einfachen Aufbau aufweist. Als zentrale Komponente wurde eine abbaubare Leiterplatte aus Holz als Trägerschicht und Eisen als Leitermaterial entwickelt. Es konnte gezeigt werden, dass Leitungsbreiten bis 600 μm zuverlässig gefertigt werden können. Der kapazitive Bodenfeuchtesensor sowie die Antenne wurden simulationsbasiert ausgelegt und auf konventionellen Materialien aufgebaut und charakterisiert. Erste Messungen über einen Zeitraum von etwa fünf Tagen zeigen, dass der Transponder den Trocknungsvorgang im Boden qualitativ erfassen kann. Der genutzte Dynamikbereich von 4 pF bis 8 pF lässt Potenzial zur Optimierung des Sensordesigns erkennen. In weiteren Schritten soll die Beziehung zwischen Sensorkapazität und volumetrischem Wassergehalt mithilfe von Referenzsensoren ermittelt sowie der Transponder vollständig auf abbaubaren Materialien realisiert werden.

Keywords: Bodenfeuchte, abbaubare Elektronik, UHF-RFID, kapazitiver Sensor, Smart Farming

I Einleitung

Die moderne Landwirtschaft steht vor komplexen Herausforderungen. Durch den Klimawandel ändern sich die Bedingungen in diesem Wirtschaftszweig deutlich. Hitzewellen, Dürreperioden und Trockenzeiten können eine künstliche Bewässerung erforderlich machen, gleichzeitig ist das Wasser hierfür limitiert. Daher sollte nur zeitlich und räumlich zielgerichtet bewässert werden, wofür jedoch eine zeitlich und räumlich hoch aufgelöste Bestimmung der Bodenfeuchte erforderlich ist.

Zur Bestimmung der Bodenfeuchte kann zwischen drei Ansätzen unterschieden werden: 1) Es können Bodenproben entnommen und im Labor untersucht werden. Die Qualität der Messwerte ist hier am höchsten, der Arbeitsaufwand macht eine hohe Auflösung aber unpraktikabel. 2) Aufnahmen mit Satelliten oder Flugzeugen ermöglichen die Untersuchung einer großen Fläche. Die Bodenfeuchte ist hier aber nicht direkt als Messgröße zugänglich. Stattdessen muss auf andere Messgrößen zurückgegriffen werden, um die Bodenfeuchte zu schätzen. Darüber hinaus ist die Häufigkeit der Aufnahmen begrenzt. 3) Kabellose Sensorknoten ermöglichen eine Untersuchung mit hoher

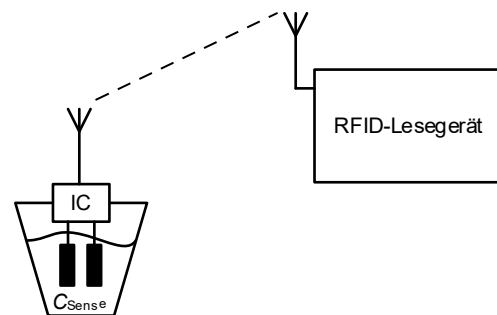


Abb. 1: Sensortransponder zur Bestimmung der Bodenfeuchte mit kapazitivem Sensor.

zeitlicher Auflösung. Für eine hohe räumliche Auflösung ist jedoch eine große Anzahl von Sensorknoten erforderlich.

Um unter Verwendung von Sensorknoten die räumliche Auflösung zu verbessern und die untersuchte Fläche zu vergrößern, muss dementsprechend die Zahl der Sensorknoten stark vergrößert werden. Damit alte Sensorknoten nicht als Elektronikschrott im Acker verbleiben, ist es dann jedoch erforderlich, ihre Position beim Ausbringen aufzuzeichnen und die Sensor-

knoten am Ende ihrer Lebenszeit wieder einzusammeln.

Sethi et al. haben eine neue Klasse von Sensorknoten vorgeschlagen, die vollständig aus biologisch abbaubaren Materialien bestehen [1]. Nach dem Ausbringen sollen sie zur Erfassung von Messwerten genutzt werden, die kabellos übertragen werden. Am Ende der Lebensdauer sollen sich diese Sensorknoten umweltverträglich zersetzen, sodass sie nicht wieder eingesammelt werden müssen und keine Schadstoffe in der Umwelt verbleiben. Sensorknoten mit dieser Eigenschaft bezeichnen wir im Folgenden als abbaubare Sensorknoten.

Alsultan et al. haben einen Sensortransponder vorgestellt, der ohne Batterie auskommt [4]. Dieser misst die Bodenfeuchte kapazitiv und stellt die Messwerte kabellos zur Verfügung. Allerdings ist der Sensortransponder nicht abbaubar und muss daher am Ende seiner Lebensdauer eingesammelt werden.

In diesem Beitrag verfolgen wir einen ähnlichen Ansatz und stellen einen batterielosen Sensortransponder vor, jedoch mit dem Ziel, diesen aus abbaubaren Materialien aufzubauen. In Abschnitt II beschreiben wir zunächst die Anforderungen an einen solchen abbaubaren Sensorknoten und stellen unser Konzept vor. In Abschnitt III beschreiben wir die Entwicklung einer abbaubaren Leiterplatte als zentrale Komponente eines abbaubaren Sensorknotens. In Abschnitt IV stellen wir das Design eines Sensortransponders zur Bestimmung der Bodenfeuchte vor, der zunächst auf konventionellen Materialien aufgebaut wurde. In Abschnitt V stellen wir erste Messergebnisse des Sensortransponders vor. Schließlich diskutieren wir die Ergebnisse in Abschnitt VI und ziehen ein Fazit.

II Abbaubare Sensorknoten: Anforderungen und Konzept

Abbaubarkeit kann unterschiedlich definiert werden und sowohl chemische, physikalische als auch biologische Mechanismen einschließen. Entscheidend für die Abbaubarkeit von Sensorknoten für die Landwirtschaft ist, dass sich diese nach einer bestimmten Einsatzzeit ausschließlich in Stoffe zersetzen, die für die Landwirtschaft im Speziellen und die Umwelt im Allgemeinen unbedenklich sind. Dabei sollte das Verhalten der Sensorknoten während der Einsatzzeit weitgehend unverändert bleiben, bevor sich diese nach der Einsatzzeit innerhalb einer überschaubaren Zeitspanne zersetzen. Vor allem die Messgenauigkeit sollte über die gesamte Einsatzzeit hinreichend hoch bleiben. Um einen Sensorknoten in einer abbaubaren Ausführung zu realisieren, müssen dementsprechend alle Komponenten, aus denen er besteht, auf Abbaubarkeit geprüft werden und ggf.

abbaubare Ausführungen der Komponenten entwickelt werden. Je mehr unterschiedliche Komponenten für einen Sensorknoten erforderlich sind, desto aufwändiger wird dieses Vorgehen. Für einen klassischen Sensorknoten umfasst es die Leiterplatte, die Batterie, den Sensor, die elektronischen Bauelemente, die Antenne und das Gehäuse. Der damit verbundene Aufwand, einen abbaubaren Sensorknoten zu entwickeln, ist entsprechend groß. Dieser Ansatz wurde in [2] verfolgt, wobei nur ein Teil der Komponenten in einer abbaubaren Ausführung realisiert wurde.

Sensortransponder basierend auf passiven Transpondern, wie sie für *Electronic Product Code Generation 2* (EPC Gen2) *Ultra High Frequency Radio-Frequency Identification* (UHF-RFID) [3] verwendet werden, zeichnen sich durch ihren simplen Aufbau aus. Als passive Transponder können sie ohne Batterie betrieben werden und bestehen üblicherweise nur aus einem *Integrated Circuit* (IC) und einer Antenne. Aus diesem Grund verfolgen wir zur Realisierung abbaubarer Sensoren den Ansatz eines solchen passiven Sensortransponders entsprechend Abb. 1.

Ein RFID-Transponder-IC wie das *EM4152* von EM Microelectronics [5] erlaubt das Auslesen einer Kapazität als Sensor. Diese Funktion dient der Messwerterfassung. Über die Antenne wird der Transponder vom Lesegerät kabellos mit Leistung versorgt, während die modulierte Rückstreuung die Übertragung der Messwerte zum Lesegerät erlaubt. Neben dem IC besteht ein solcher Sensortransponder nur aus einer Leiterplatte. Mit einem Gewicht von etwa 250 µg, von denen der überwiegende Anteil Silizium ist, wird die Schädlichkeit des ICs als vernachlässigbar eingestuft. Um einen abbaubaren Sensortransponder zu realisieren, muss daher eine abbaubare Leiterplatte entwickelt und geeignet strukturiert werden.

III Entwicklung einer abbaubaren Leiterplatte

Für den abbaubaren Sensortransponder ist eine abbaubare Leiterplatte erforderlich. Auf dieser sollen die Antenne und der kapazitive Sensor realisiert werden. Dazu ist eine Strukturierung der Leiterplatte erforderlich.

Unser Aufbau umfasst eine einlagige Leiterplatte aus Holz und Eisen. Das Holz dient dabei als Trägerschicht. Eisen als Leiter ist mit 5,63 % das vierthäufigste Element in der Erdkruste [6]. Es wird daher davon ausgegangen, dass es für einen abbaubaren Sensorknoten unbedenklich ist. Zur Verbindung beider Schichten wird der abbaubare Schmelzkleber Caremelt von TITK genutzt.

Um die Strukturierung der einlagigen Leiterplatte zu evaluieren, wurde eine Teststruktur entsprechend Abb. 2 (oben) designt. Auf dieser befinden sich Leitungen mit verschiedener Breite zwischen 200 μm und 700 μm . Die strukturierte Leiterplatte ist in Abb. 2 (unten) dargestellt. Für die Fertigung wurde ein ProtoMat Fräsbohrplotter von LPKF genutzt, womit Leitungen bis zu einer Breite von 600 μm zuverlässig auf der abbaubaren Leiterplatte realisiert werden konnten. Damit wird zwar nicht die Strukturgröße von Leiterplatten aus konventionellen Materialien erreicht, für die angestrebte Anwendung ist diese Strukturgröße jedoch ausreichend. Die Eigenschaft der Abbaubarkeit der Leiterplatte wird durch die gewählten Materialien erfüllt.

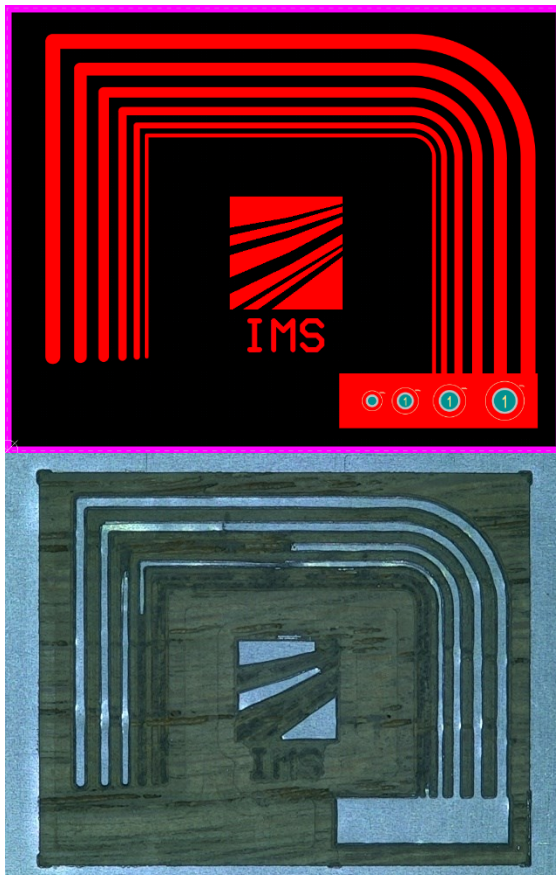


Abb. 2: Teststruktur auf abbaubarer Leiterplatte (Design oben, gefertigt unten): Die Teststruktur hat eine Größe von 20 mm $\times$ 25 mm. Die Breite der sechs Leitungen liegt in Schritten von 100 μm zwischen 200 μm und 700 μm . Die Löcher in der Teststruktur wurden in der dargestellten Version nicht gebohrt.

IV Design eines Sensortransponders zur Bestimmung der Bodenfeuchte

Der Sensortransponder basiert auf dem Transponder-IC *EM4152* von EM Microelectronics [5]. Dieser ist mit EPC Gen2 [3] kompatibel und kann zur Auslese eines kapazitiven Sensors von 0 pF bis 17 pF mit einer Auflösung von 7 Bit genutzt werden. Für den Sensortransponder muss das IC mit einer Antenne für den UHF-Bereich und einem kapazitiven Sensor zur Bestimmung der Bodenfeuchte verbunden werden.

Als Antenne wird ein mäandrierter Dipol genutzt. Dieser wurde in CST Studio designt und an die Impedanz des Transponder-ICs angepasst.

Der kapazitive Sensor wurde ebenfalls in CST Studio designt. Er besteht aus zwei rechteckigen Flächen. Wie in Abb. 3 zu sehen, erstreckt sich das elektrische Feld zwischen beiden Metallflächen auch über diese und ragt so deutlich in die Umgebung hinein. Wird der kapazitive Sensor in den Erdboden eingebracht, verläuft ein Teil des elektrischen Felds durch den Erdboden.

Die Gesamtkapazität des Sensors C_{Sense} wird als Parallelschaltung zweier Kapazitäten modelliert, wobei das Feld einer Kapazität C_0 innerhalb des Dielektrikums liegt und die andere Kapazität C_1 außerhalb des Sensors durch den Erdboden verläuft. Die erste Kapazität wird als konstant angenommen. Da Wasser eine deutlich größere relative Permittivität aufweist als trockene Erde, ändert sich die Kapazität des Sensors in Abhängigkeit der umgebenden Bodenfeuchte. In Abb. 4 ist die simulierte Änderung der Kapazität in Abhängigkeit der relativen Permittivität der Umgebung des Sensors dargestellt. Es zeigt sich ein linearer Verlauf von C_{Sense} über der relativen Permittivität der Umgebung des Sensors.

Um über C_{Sense} die Bodenfeuchte zu bestimmen, kann die Topp-Gleichung verwendet werden, die den Zusammenhang zwischen relativer Permittivität des Erdbodens und dessen volumetrischem Wassergehalt als rationale Funktion dritten Grades modelliert [8]. In Abb. 5 ist der Verlauf von C_{Sense} über dem volumetrischen Wassergehalt θ_{VWC} dargestellt. Der vom Transponder-IC auswertbare Bereich bis 17 pF ermöglicht demnach einen Messbereich von θ_{VWC} bis etwa 50 %. Dies entspricht dem Messbereich handelsüblicher Sensoren zur Bestimmung der Bodenfeuchte [7].

Der gefertigte Sensortransponder ist in Abb. 6 abgebildet. Er wurde derart gestaltet, dass ein einfaches Einbringen in den Boden möglich ist. Die Antenne am oberen Ende des Transponders bleibt im Einsatz über dem Boden. Der kapazitive Sensor befindet sich am unteren Ende des Transponders. Er ist mit einer schwarzen

Schutzhülle überzogen. Diese stellt einen Abstand zwischen den Metallflächen des Sensors und dem Boden her, der verhindert, dass der Sensor bei hoher Bodenfeuchte zu früh in Sättigung geht. Außerdem wird der kapazitive Sensor auf diese Weise vor einer leitfähigen Verbindung der Metallflächen durch den Erdboden geschützt.

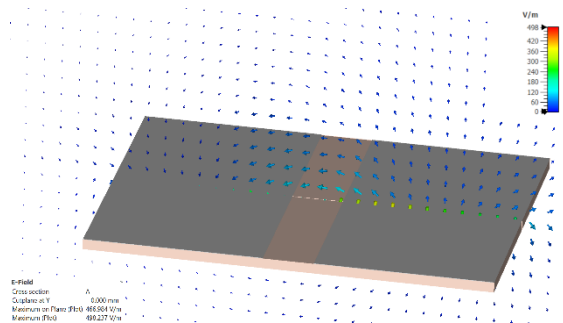


Abb. 3: Elektrisches Feld in der Umgebung des kapazitiven Sensors simuliert mit CST Studio.

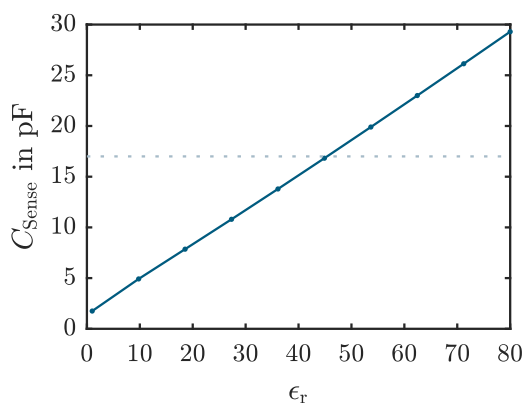


Abb. 4: Simulierte Änderung der Kapazität des Sensors mit der relativen Permittivität der Umgebung.

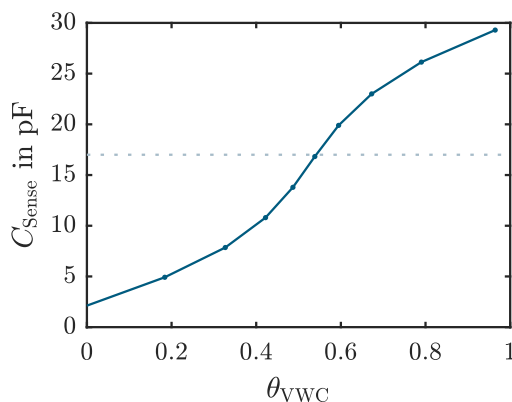


Abb. 5: Simulierte Änderung der Kapazität des Sensors mit dem volumetrischen Wassergehalt θ_{VWC} des Bodens.



Abb. 6: Sensortransponder gefertigt aus konventionellen Materialien: Das Transponder-IC und der kapazitive Sensor befinden sich unter der schwarzen Schutzhülle.

V Messergebnisse

Zur Untersuchung, inwieweit der entwickelte Sensortransponder zur Bestimmung der Bodenfeuchte geeignet ist, wurde dieser in einem kleinen Blumentopf von 11 cm Durchmesser etwa 5 cm tief in trockener Erde platziert. Ein RFID-Lesegerät wurde in einem Abstand von etwa 20 cm aufgestellt und mit einem Computer verbunden. Über diesen wurde der Sensorwert des Transponders mit einem Abtastintervall von 5 s ausgelesen. Nach 30 s wurde der Boden befeuchtet und die Trocknung über einen Zeitraum von etwa 5 Tagen aufgezeichnet.

Abb. 7 zeigt die aufgezeichneten Messwerte. Um der geringen Digitalisierungsauflösung (7 Bit) des Transponder-ICs entgegenzuwirken, wurde ein gleitender Mittelwert über jeweils 20 aufeinanderfolgende Datenpunkte gebildet. Die kleine Darstellung innerhalb von Abb. 7 zeigt die Rohdaten aus den ersten 5 Minuten ohne Filterung.

In den ersten 30 s vor der Zugabe des Wassers bleibt die Kapazität des Sensors konstant. Mit der Zugabe steigt der Wert schlagartig an, bevor er über den Zeitraum von etwa 1 min zügig abfällt. In diesem Zeitraum wurde das Wasser von der Erde aufgenommen. Danach fällt der Messwert in Übereinstimmung mit einer langsamen Trocknung ab. Jeweils gegen Mittag ist eine deutlichere Änderung des Messwerts zu erkennen, die vor allem an Tag 3 besonders ausgeprägt ist. In diesem Zeitraum kam es zu direkter Sonneneinstrahlung auf den Blumen-

topf, was die Trocknung vermutlich beschleunigte. Zum Ende der Messreihe erreichte der Sensor wieder einen Wert nahe dem des trockenen Bodens.

VI Diskussion und Ausblick

Die Messergebnisse zeigen, dass der Sensortransponder den Trocknungsvorgang im Boden erkennen kann. Der Verlauf der Messwerte ist plausibel, wie beispielsweise durch die starken Veränderungen zur Mittagszeit oder den Fakt, dass am Ende der Messung nach 5 Tagen wieder ungefähr der Startwert erreicht wurde, deutlich wird.

Mit einem Dynamikbereich von 4 pF bis 8 pF wird nur ein Teil der nutzbaren 17 pF des Transponder-ICs ausgenutzt. Verbunden mit der geringen Samplingtiefe des Transponder-ICs von 7 Bit ist die Auflösung des Sensors zunächst gering. Da die Messwerte zusätzlich statistischen Schwankungen unterliegen und die Messung in hoher Geschwindigkeit wiederholt werden kann, ermöglicht eine geeignete Filterung, die Auflösung der Messung zu verbessern.

Ob die Ausschläge gegen Mittag nur auf eine beschleunigte Trocknung zurückzuführen sind, oder der Sensortransponder auch eine Temperaturabhängigkeit aufweist, bedarf weiterer Untersuchung. Über die Topp-Gleichung und Messungen mit Referenzsensoren muss die Beziehung zwischen der Kapazität des Sensors und dem volumetrischen Wassergehalt des Bodens ermittelt werden. Darüber hinaus gilt es zu prüfen, inwieweit auch die Art des Bodens die Messung beeinflusst.

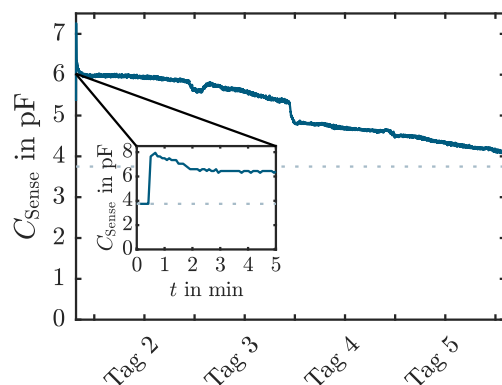


Abb. 7: Vom Sensortransponder mit einem Abtastintervall von 5 s erfasste Kapazität des Bodenfeuchtesensors: Für die Messreihe über etwa 5 Tage wurde der gleitende Mittelwert über jeweils 20 aufeinanderfolgende Datenpunkte berechnet. Die kleine Darstellung über 5 min zeigt ungefilterte Rohdaten.

In unseren bisherigen Versuchen und Simulationen hat sich herausgestellt, dass die Dicke der Schutzschicht einen wesentlichen Einfluss auf den Wertebereich des Sensors hat. Mit einer zu dünnen Schicht sättigt der Sensor bei hoher Bodenfeuchte. Mit einer zu dicken Schicht wird die Sensitivität eingeschränkt. Die Dicke der Schutzschicht muss daher entsprechend eingestellt werden. Zudem muss die Schwankung über verschiedene Transponder gering bleiben, um ein gleichbleibendes Verhalten sicherzustellen. Gleichzeitig ist die Schutzschicht in direktem Kontakt zur Umgebung. Für die Realisierung des abbaubaren Sensortransponders muss auch sie abbaubar sein, während sich die Dicke der Schicht über die Betriebsdauer des Sensortransponders hinweg jedoch nur geringfügig ändern sollte. Dies stellt eine zentrale Herausforderung für das Design eines abbaubaren kapazitiven Sensortransponders dar.

In einem nächsten Schritt soll der Sensortransponder auf der abbaubaren Leiterplatte gefertigt werden. Dadurch wird, verbunden mit einer geeigneten Schutzschicht, ein abbaubarer Sensorknoten zur Bestimmung der Bodenfeuchte entstehen. Um die Funktion über eine Pflanzperiode sicherzustellen, werden wir dabei verschiedene Materialien zur Verkapselung des Sensortransponders untersuchen.

Wie in Abschnitt II dargestellt, haben wir die Verwendung eines Transponder-ICs aufgrund der geringen Masse von etwa 250 µg und des hohen Anteils von Silizium als unbedenklich erachtet. Diese Annahme bedarf weiterer kritischer Reflexion, da in modernen CMOS-Prozessen neben dem Silizium diverse andere Stoffe zum Einsatz kommen, die in Teilen auch in kleiner Konzentration bedenklich sein können.

Um das Transponder-IC mit der abbaubaren Leiterplatte zu verbinden, kommen verschiedene Prozesse in Frage. Auch hier werden potenziell bedenkliche Substanzen, wenngleich in sehr geringen Mengen, eingesetzt. Für die Bewertung der Unbedenklichkeit gilt es daher, neben dem IC auch die Verbindungstechnik zu betrachten.

Die bisherigen Ergebnisse bilden eine vielversprechende Grundlage und bestätigen das Potenzial abbaubarer Sensortransponder mit kapazitivem Sensor zur Etablierung nachhaltiger Messtechnik in der Landwirtschaft.

Danksagung

Das diesem Bericht zugrunde liegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 031B1598 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Das Projekt FarmScan wurde im Rahmen des Green ERA-Hub durchgeführt, einer Koordinierungs- und Unterstützungsmaßnahme (CSA), die durch das Forschungs- und Innovationsprogramm „Horizont Europa“ der Europäischen Union unter der Fördervereinbarung Nr. 101056828 gefördert wurde.

Literaturnachweis

- [1] S. S. Sethi, M. Kovac, F. Wiesemüller, A. Miriyev und C. M. Boutry, „Biodegradable sensors are ready to transform autonomous ecological monitoring“, *Nature Ecology & Evolution* 6, 2022, doi: 10.1038/s41559-022-01824-w.
- [2] S. Voigt, A. Willert, W. Mende, T. Zschau, T. Oehme und R. Zichner, „Environmental friendly and low cost monitoring system for plant and agriculture fields“, *2023 Smart Systems Integration Conference and Exhibition (SSI)*, 2023.
- [3] GS1, „EPC Radio-Frequency Identity Generation-2 UHF RFID Standard“, Version 3.0, 2024.
- [4] M. A. Alsultan, J. Melià-Seguí, J. Parrón-Grados und S. López-Soriano, „A Battery-Less UHF RFID Sensor for Soil Moisture Monitoring“, *IEEE Journal of Radio Frequency Identification*, 2025, doi: 10.1109/JRFID.2025.3572843.
- [5] EM Microelectronic-Marin SA, „RAIN RFID TRANSPONDER IC WITH CAPACITIVE SENSOR INTERFACE“, März 2021, Version 4.2.
- [6] W. M. Haynes, „Abundance of Elements in the Earth's Crust and in the Sea“, *CRC Handbook of Chemistry and Physics*, 97th edition (2016–2017), sec. 14, pg. 17.
- [7] truebner, „SMT50 Bodenfeuchtesensor“, Januar 2018.
- [8] G. C. Topp, J. L. Davis und A. P. Annan, „Electromagnetic determination of soil water content: Measurements in coaxial transmission lines“, *Water Resour. Res.*, 1980, doi:10.1029/WR016i003p00574.

Inductive Harvester and Data Transmission on Medium-Voltage High Power Cable Shields for Autonomous Sensor Nodes

M. Lenzhofer¹, D. Holzmann¹, D. Sturm²

¹*Silicon Austria Labs GmbH, Europastrasse 12, 9524 Villach, Austria*

²*KNG-Kärnten Netz GmbH, Arnulfplatz 2, Postfach 200, 9020 Klagenfurt am Wörthersee, Austria*

Abstract

Medium-voltage (MV) cable systems represent a critical component of modern electrical infrastructure, where failures often occur at cable joints due to environmental stress and aging. Continuous monitoring of such joints is therefore highly desirable; however, their underground installation makes conventional communication approaches impractical. This paper presents an inductively coupled communication system that utilizes the existing cable infrastructure as a transmission medium. A low-power embedded system is developed, combining resonant inductive coupling with software-based signal detection using the Goertzel algorithm. Experimental results demonstrate reliable data transmission over several hundred meters under realistic load conditions. The influence of cable configuration, load current, and interference sources is analyzed, and design limitations are discussed. The results confirm the feasibility of the approach and provide guidelines for future system optimization.

Keywords: inductively coupled communication, power line communication, condition monitoring, medium-voltage cable, data transmission

Introduction

Medium-voltage (MV) distribution grids operating at approximately 20 kV are essential for reliable electrical energy delivery [1]. Due to their underground installation and long service life, these systems are exposed to environmental stress factors such as temperature variations, moisture, and corrosion, which can lead to degradation and eventual failure [1]. Cable joints, used to repair or extend cables, represent critical parts due to their structural and material discontinuities. Despite their importance, the condition of cable joints is typically unknown after installation. Embedded sensing systems offer a promising solution for monitoring parameters such as temperature, humidity, and corrosion state. However, their underground placement introduces significant challenges, particularly in terms of energy supply and data communication.

This work addresses the communication problem by utilizing the cable itself as a transmission medium. With the given coaxial structure of MV cables, the presence of grounded shields and the parallel laid ground conductor, a closed conductive loop is formed, refer to Figure 1. This topology enables inductive signal injection, making inductive coupling a suitable method for data

transmission.

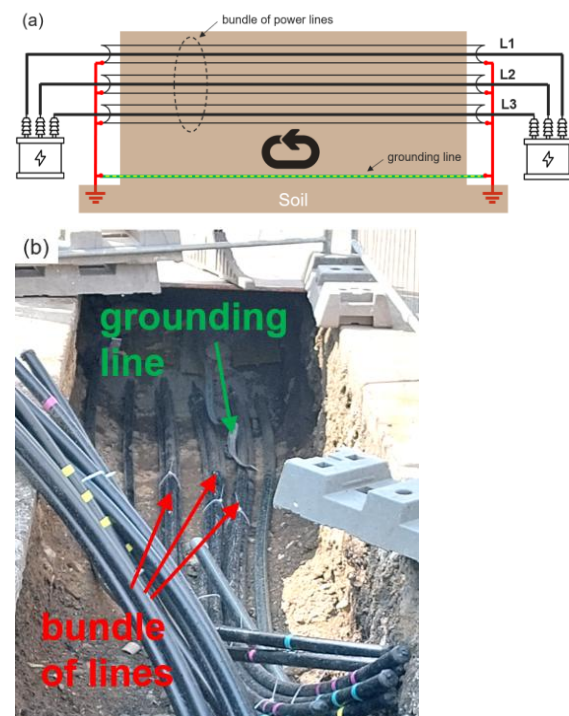


Fig. 1: (a) Schematic representation of a medium-voltage (MV) cable system, showing three-phase conductors with grounded shields and an additional ground conductor forming a closed loop. (b) Photograph of the cable installation in a real deployment environment.

Principle of Inductive Coupling

Inductive coupling relies on the generation of a time-varying magnetic field, which induces a voltage in a nearby conductor [2]. The induced voltage is governed by Faraday's law, Eq. (1):

$$u = -M \frac{di}{dt} \quad (1)$$

where M denotes the mutual inductance between the transmitting and receiving elements.

In contrast to conventional transformers, the presented system operates without a magnetic core between transmitter and receiver, resulting in reduced coupling efficiency [3]. The proposed approach extends this concept by using the cable shield and ground conductor as an intermediate coupling loop. The system can thus be interpreted as two cascaded inductive coupling stages: from transmitter coil to cable loop, and from cable loop to receiver coil, refer to Figure 2.

Although this double coupling reduces efficiency, it enables communication over significantly larger distances than in classical near-field systems such as RFID [3], based on inductively coupling mechanism through air.

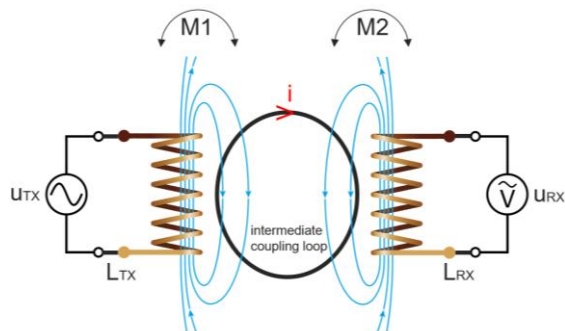


Fig. 2: Conceptual illustration of the two-stage inductive coupling mechanism utilizing an intermediate coupling loop formed by the cable shield and ground conductor.

System Design and Hardware Implementation

A. Electrical Characteristics and Frequency Selection

The cable shield and ground conductor form a low-resistance loop, which is essential for efficient coupling. However, discontinuities such as cable joints increase the loop resistance and reduce signal strength.

The operating frequency must balance data rate and transmission distance. Considering attenuation effects such as skin effect, a frequency range between 30 kHz and 60 kHz was selected. This ensures sufficient bandwidth

while maintaining acceptable propagation characteristics.

B. Transmitter and Receiver Design

The transmitter is based on the evaluation board NUCLEO-L412KB and uses an H-bridge driver (DRV8838) to excite a series resonant circuit at 30 kHz [4]. Data is modulated using On-Off Keying (OOK), where the carrier is switched according to the bit stream [5]. Figure 3 (a) shows the block diagram of the transmitter circuit.

The receiver employs a parallel resonant circuit tuned to the same frequency, followed by a notch filter at 50 Hz to suppress grid interference. Signal amplification is realized using the OPA2328DGKR [6], following standard low-noise analog front-end design principles [7, 8] and finally the signal is digitized via the microcontroller's ADC. Figure 3 (b) shows the block diagram of the receiver circuit.

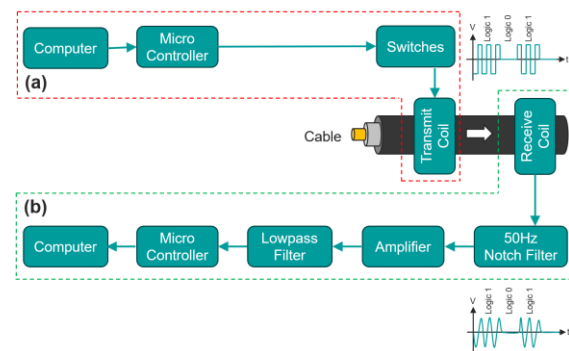


Fig. 3: Block diagrams of the system architecture: (a) transmitter unit and (b) receiver unit.

Figure 4 depicts the developed electronics for the transmitter and receiver unit with the inductive split core couplers.

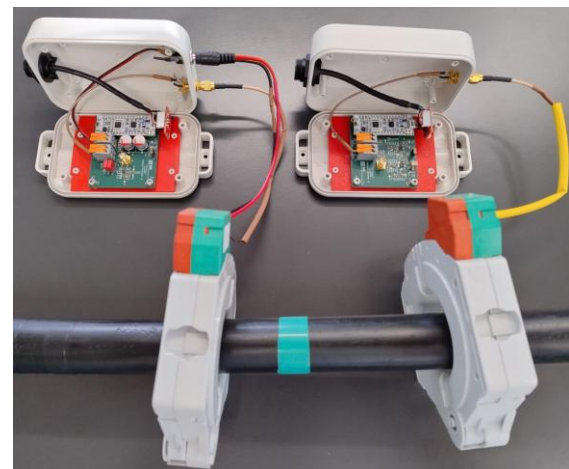


Fig. 4: Photograph of the developed transmitter and receiver prototypes, including inductive coupling elements mounted on a shielded MV cable.

Software Implementation

To minimize hardware complexity, signal detection is implemented entirely in software. The receiver continuously samples the input signal and evaluates the presence of the carrier frequency using the Goertzel algorithm [5]. The algorithm efficiently computes the spectral component at a predefined frequency of 30 kHz, making it well suited for low-power embedded systems. In contrast to a full discrete Fourier transform, the computational effort is significantly reduced, as only a single frequency bin is evaluated. Such digital signal processing approaches are widely used in communication systems and embedded implementations [5, 8].

A detection threshold is applied to the resulting magnitude of the frequency component to distinguish between logical '1' (carrier present) and '0' (carrier absent). To improve robustness against noise and interference, the decision is based on accumulated samples over a defined observation window rather than instantaneous values.

Communication is organized into frames consisting of a start bit, eight data bits, and a stop bit. Synchronization is achieved by detecting the start bit using a shorter sampling window, enabling rapid detection of transmission onset. After synchronization, the receiver switches to fixed-time sampling aligned with the known bit duration, ensuring reliable decoding of the subsequent data bits.

Adaptive thresholding is implemented to account for varying signal amplitudes due to cable length, load conditions, and environmental influences. For this purpose, the amplitude of the detected start bit is used as a reference to dynamically adjust the decision threshold for the following data bits. This approach improves detection reliability under changing channel conditions without requiring manual recalibration.

Figure 5 illustrates the flowcharts of the implemented software architecture for both system components. Subfigure (a) depicts the processing sequence within the transmitter unit, while subfigure (b) presents the corresponding algorithmic flow implemented in the receiver unit.

In addition, the continuous evaluation of the carrier component allows the receiver to operate in a passive listening mode, activating the decoding process only when a valid transmission is detected. This contributes to reduced computational load and supports

energy-efficient operation of the embedded system.

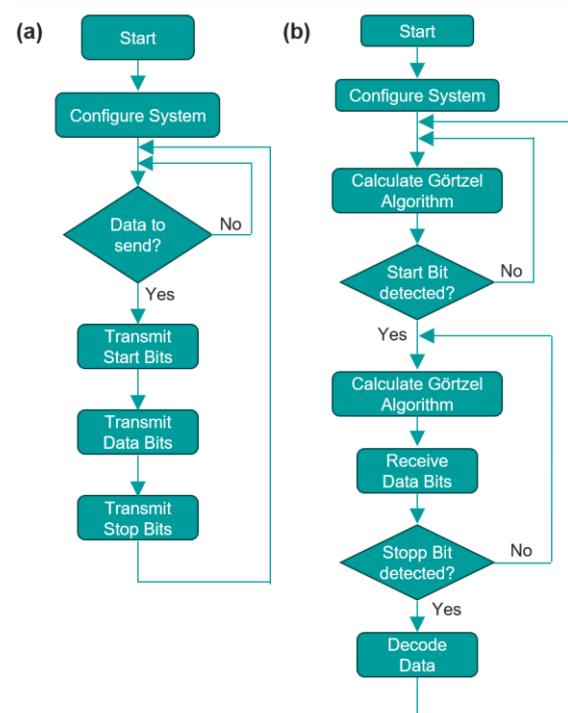


Fig. 5: Software flowcharts of the communication protocol implemented in (a) the transmitter unit and (b) the receiver unit.

Measurements and Results

A. Experimental Setup

A series of experiments were conducted to evaluate system performance under different conditions, including cable configuration, load current, generator influence, and coil integration, as well as different positions of the transmitter and receiver coils.

B. Transmission Performance

With a 200 m cable wound on a drum, reliable communication was achieved at currents up to 50 A. The increased inductance of the coiled cable improves coupling efficiency [3]. In contrast, for a 400 m unrolled cable, the maximum load current for reliable communication decreased to approximately 10 A.

The signal-to-noise ratio (SNR) was observed to decrease with increasing load current, from approximately 35 dB (no load) to 20 dB at 10 A.

C. Influence of External Disturbances

The presence of a generator significantly affected system performance. Even in idle state, the generator reduced the SNR from approximately 70 dB to 50 dB due to switching noise from internal power electronics [6].

D. Coil Interaction Effects

In addition to the inductively coupled communication system, the influence of an integrated energy harvesting module was investigated. The combined configuration consisted of a harvesting coil with approximately 1250 turns and a communication coil with 50 turns, both implemented on a shared magnetic core. As outlined previously, the communication signal is coupled via the cable shield, whereas energy harvesting is performed from the main conductor current within the cable core. The harvesting unit operates by charging an internal supercapacitor (470 mF), with a load switch enabling the supply of the electronic circuitry once a sufficient voltage level is reached.

However, the integration of both coils on a common core was found to significantly degrade the communication performance. From an electromagnetic perspective, this arrangement forms a multi-winding transformer, introducing unintended coupling between the communication and harvesting subsystems. In particular, parasitic inductive coupling and inter-winding capacitances create additional reactive loading paths, which adversely affect the high-frequency communication signal.

As a result, a portion of the injected communication signal is coupled into the harvesting coil. Although filtering is applied within the harvester circuit, the suppression of high-frequency components is incomplete. Consequently, this leads to an increased apparent impedance at the transmitter side and additional losses due to circulating currents within the harvesting branch. These effects reduce the effective signal amplitude available for reliable detection at the receiver.

Experimental results confirm this behavior, showing that the maximum permissible load current for reliable communication decreases to approximately 5 A when both coils are integrated on a single core, refer to figure 7. In contrast, a configuration employing physically separated cores for communication and energy harvesting allows operation up to approximately 10 A, as shown in figure 6.

These findings clearly demonstrate that physical and electromagnetic decoupling of the communication and energy harvesting subsystems is essential to maintain coupling efficiency and overall system performance.

E. Discussion

The results demonstrate that system performance is strongly influenced by:

- cable configuration,
- load current,
- external noise sources,

- magnetic coupling conditions.

Separating communication and harvesting coils onto different cores significantly improves performance.

Figure 6 presents the measurement results obtained for two distinct cable configurations. In configuration (a), a 200 m cable section was evaluated in conjunction with an equally long grounding conductor, forming a defined return path. In configuration (b), a total cable length of 400 m was investigated, where only the cable shields were interconnected.

Furthermore, the position of the transmitter unit was systematically varied with respect to the connection point, which represents a disturbance location introduced by a simplified junction realized through thin wire connections. Measurements were conducted with the transmitter placed both upstream and downstream of this point in order to assess the influence of discontinuities on signal propagation and coupling efficiency.

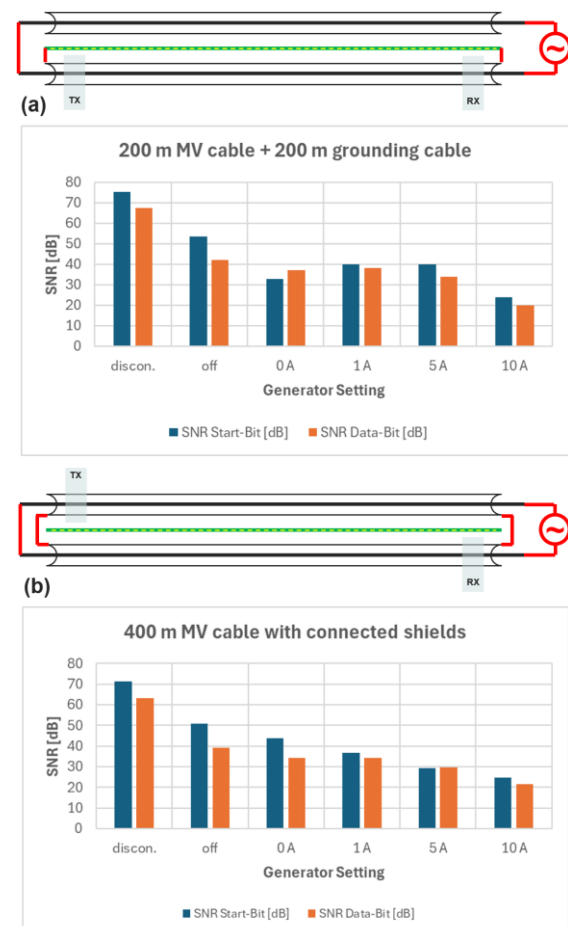


Fig. 6: Measured SNR as a function of load current for different cable configurations: (a) 200 m MV cable with an additional parallel 200 m grounding conductor, and (b) 400 m MV cable formed by two interconnected 200 m sections.

Finally, as illustrated in Fig. 7, the transmitter and receiver units are positioned in close proximity to each other, directly adjacent to the generator feeding point. In Fig. 7(a), the system operates using only split cores equipped with the communication coils, representing the baseline configuration. In contrast, Fig. 7(b) shows the results obtained when the transmitter core is replaced by the combined version integrating the energy harvesting unit.

In the latter case, communication is only established after a certain delay, corresponding to the time required for the harvester to charge its internal supercapacitor and subsequently activate the load switch supplying the electronics. Consequently, no communication results are observed when the generator is disconnected or switched off, as insufficient energy is available to power the transmitter circuitry.

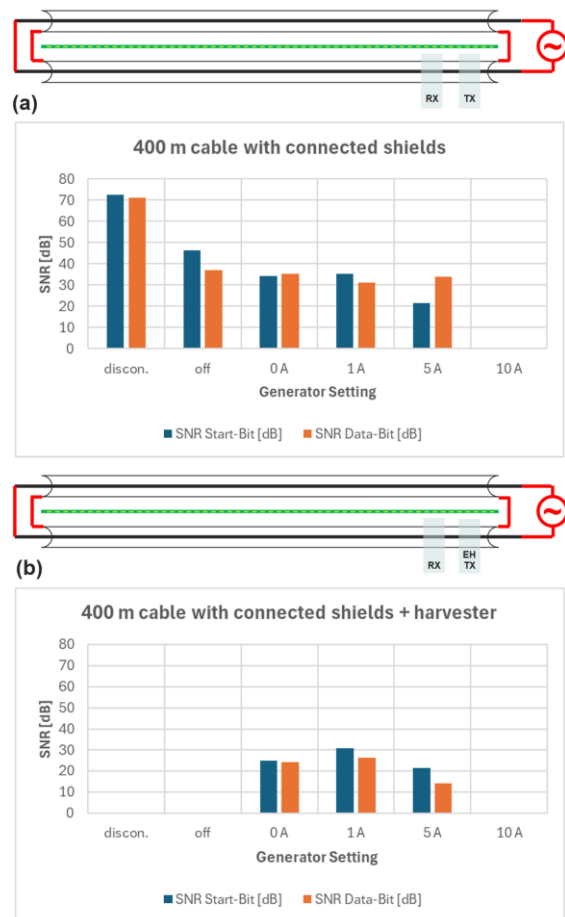


Fig. 7: Measured SNR as a function of load current for different system configurations: (a) 400 m cable with transmitter and receiver positioned in close proximity near the feeding point, and (b) 400 m cable with combined transmitter and energy harvesting coil on a shared core, evaluated against a separate receiver coil.

Conclusion and Outlook

This paper presented an inductively coupled communication system for embedded monitoring in medium-voltage cable networks. By leveraging the existing cable infrastructure as a transmission medium, the proposed approach enables wireless data transfer from buried sensor systems without additional communication lines.

The combination of resonant inductive coupling and software-based signal detection using the Goertzel algorithm results in a low-complexity and energy-efficient implementation. Experimental results demonstrate reliable communication over several hundred meters, even under realistic load conditions.

However, several limitations were identified, including sensitivity to external interference, reduced performance under load, and coupling losses in integrated coil configurations.

Future work will focus on improving system robustness and efficiency through: advanced modulation schemes (e.g., FSK or PSK), error detection and correction mechanisms, optimized coil and core designs, adaptive communication protocols for varying channel conditions.

These enhancements are expected to significantly improve transmission reliability and enable practical deployment in real-world MV grid monitoring applications.

Acknowledgement

This work has been jointly supported by the Industry partner KNG-Kärnten Netz GmbH and by Silicon Austria Labs (SAL), owned by the Republic of Austria, the Styrian Business Promotion Agency (SFG), the federal state of Carinthia, the Upper Austrian Research (UAR), and the Austrian Association for the Electric and Electronics Industry (FEEI).

References

- [1] M. Kuffel, W. S. Zaengl, J. Kuffel, High Voltage Engineering: Fundamentals, Butterworth-Heinemann, 2000.
- [2] S. Tumanski, "Induction Coil Sensors – A Review," Measurement Science and Technology, vol. 18, 2007.
- [3] K. Finkenzeller, RFID Handbook, 3rd ed., Wiley, 2010.
- [4] Texas Instruments, DRV8838 Datasheet, 2020.
- [5] S. Haykin, Communication Systems, 5th ed., Wiley, 2009.
- [6] Texas Instruments, OPA2328 Datasheet, 2018.
- [7] P. Horowitz, W. Hill, The Art of Electronics, 3rd ed., Cambridge University Press, 2015.
- [8] A. Sedra, K. Smith, Microelectronic Circuits, Oxford University Press, 2015.

Instrumentierung und Stabilisierung von Lichtbögen zur Plasma-Schmelzreduktion

Thomas Voglhuber-Brunnmaier¹, Benjamin Lehner¹, Lukas Ecker², Jakob Stanek¹, Erwin K. Reichel³

¹ *Institut für Mikroelektronik und Mikrosensorik, Johannes Kepler Universität, Linz, Altenberger Straße 69, 4040 Linz,*

² *Institut für Regelungstechnik, Johannes Kepler Universität, Linz, Altenberger Straße 69, 4040 Linz,*

³ *Metallurgisches Kompetenzzentrum K1-met, Stahlstraße 14, 4020 Linz*

Zusammenfassung

Der Wasserstoff-Plasma-Schmelzreduktionsprozess (HPSR) stellt eine vielversprechende Technologie zur CO₂-armen Stahlproduktion dar. Instabile Lichtbogenbedingungen limitieren jedoch Energieeffizienz, Wasserstoffausnutzung und Elektrodenverschleiß. Basierend auf experimentellen Untersuchungen wird ein Zusammenhang zwischen Lichtbogenlänge und elektrischen Kenngrößen im UI-Diagramm identifiziert und zur Stabilitätsbewertung genutzt. Zur aktiven Stabilisierung wird ein schneller Stromregler auf Basis eines Tiefsetzstellers entwickelt und implementiert. Ergänzend wird ein stereoskopisches Messsystem zur hochfrequenten Rekonstruktion der Lichtbogenlänge vorgestellt. Die Ergebnisse zeigen eine Korrelation der Lichtbogenlängen mit der gemessenen Lichtbogenspannung. Dies stellt den ersten Schritt in Richtung elektrisch kontrollierbarer Lichtbogendynamik dar.

Keywords: Buck-Converter, Stereoskopie, adaptiver Zustandsregler, Stahlerzeugung

1. Einführung

Der Wasserstoffplasma-Schmelzreduktionsprozess (engl. hydrogen-plasma smelting reduction, HPSR) ist einer der vielversprechendsten Ansätze in Richtung dekarbonisierter Stahlproduktion. Feinerze werden in einem Schritt reduziert und zur Abtrennung von Gangart und flüssigem Rohstahl aufgeschmolzen. Die Reduktion des Eisenerzes wird nicht durch Zugabe von Kohlenstoff, sondern durch im Lichtbogen ionisiertem Wasserstoff bewerkstelligt. Sowohl die Einbringung von Schmelzenergie als auch die Bereitstellung des Reduktionsmittels können auf diese Art vollständig durch erneuerbare Energieträger erbracht werden. Im Zielprozess werden durch hohle Graphitelektroden Eisenerz und ein Gemisch aus Argon und Wasserstoff zugeführt. Das Argon begünstigt hierbei die Plasmabildung. In der Literatur [1-3] wird ausgeführt, dass instabile Plasmabedingungen, das heißt, Fluktuationen der Lichtbogenlängen bis hin zu temporären Lichtbogenabbrissen, zu ungewöhnlich hohem Verschleiß der Graphitelektroden und zu reduzierten elektrischen Wirkungsgraden und einer schlechteren Wasserstoffausnutzung führen können. Besagte Plasmainstabilitäten ergeben sich zum einen aus der komplexen Verkopplung der thermodynamischen und

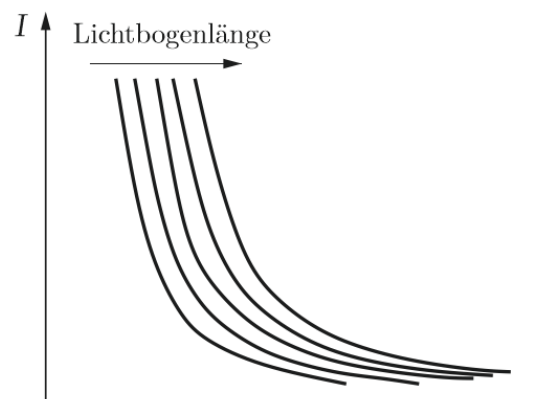


Abb. 1: Typische UI-Charakteristika von Lichtbögen unterschiedlicher Längen.

elektrischen Kräfte, aber auch aus dem Zusammenspiel der negativen differentiellen Leitfähigkeit des Lichtbogens und der Charakteristika des elektrischen Versorgungskreises. Diesem letzten Aspekt, der in der Literatur bisher weitgehend unberücksichtigt ist, wird in der Arbeit besonderes Augenmerk geschenkt. Ziel ist es verborgene Dynamiken des Lichtbogens zu erforschen und durch schnelle elektrische Intervention diesen zu stabilisieren.

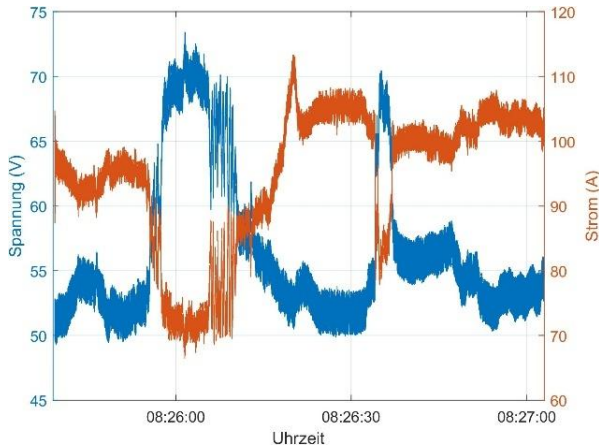


Abb. 2: Spannungs- und Strom-Rohdaten eines Lichtbogenversuchs.

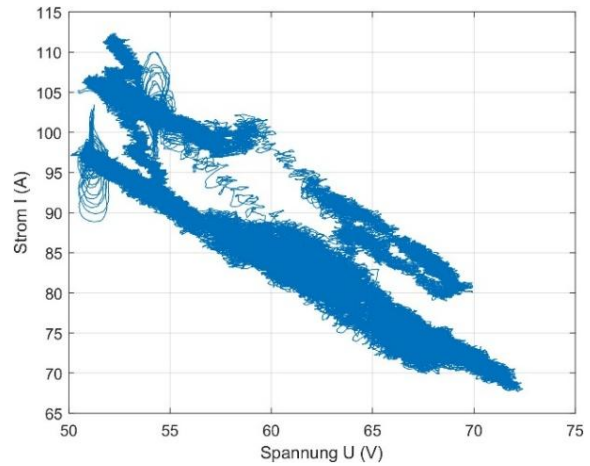


Abb. 3: UI-Darstellung der Daten in Abb. 2.

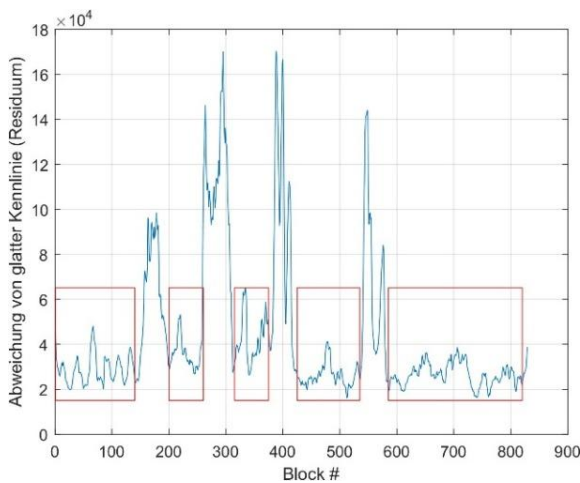


Abb. 4: Residuen der blockweisen Schätzungen von $\hat{r}_D$ und $\hat{I}_0$. (Fensterlänge 10^5 Werte = 1s, Blockversatz 1/10s). Rot umrandete Abschnitte werden ausgewertet.

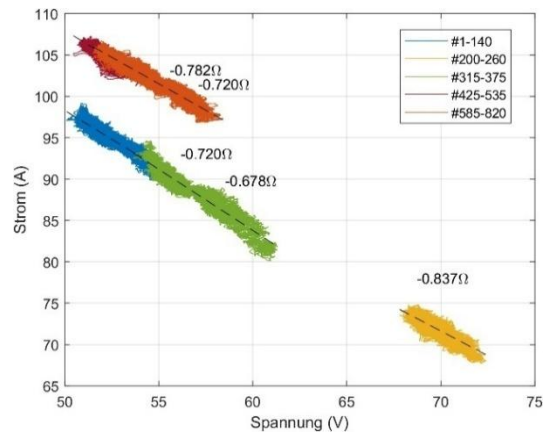


Abb. 5: Die selektierten Messdaten bilden die Lichtbogen-Kennlinien nach. Die lokalen Differenziellen Widerstände wurden individuell für die 5 in Abb. 4 gezeigten Blöcke berechnet.

Abb. 1 zeigt charakteristische Lichtbogenkennlinien unterschiedlicher Länge. Der negative differentielle Widerstand ist die Folge steigender Ionisation bei erhöhtem Leistungseintrag. Mit steigendem Strom steigt daher die Leitfähigkeit, wodurch sich der Spannungsabfall reduziert. Negative differentielle Widerstände führen potentiell zu Instabilitäten des Versorgungskreises sofern Zuleitungsinduktivitäten und Lichtbogenkapazitäten berücksichtigt werden [4]. Verlängerung des Lichtbogens bei konstant gehaltenem Strom bewirkt eine Erhöhung der Spannung. Dynamische Bogenlängenänderungen sind daher mit den UI-Daten korreliert. Abb. 2 zeigt Zeitreihen der gemessenen Spannung und des Stromes bei Versuchen an einem bestehenden Laborsystem zur Untersuchung metallurgischer Eigenschaften des Schmelzreduktionsprozesses. Die elektrischen Signale waren in

den originalen Untersuchungen nicht von Bedeutung. Der Versorgungskreis bestand aus einem thyristorgesteuerten Phasenanschnittsteller am Dreiphasennetz mit nachgeschaltetem Transformator und einem 6-Puls Gleichrichter. Der Arbeitspunkt wurde durch Einstellung des Phasenwinkels des Thyristorstellers festgelegt. Diese exemplarischen Spannungs- und Stromverläufe zeigen sich erratisch. Im zugehörigen UI-Diagramm in Abb. 3 deuten sich Abschnitte mit stabilen Lichtbögenlängen an. Zur Identifikation dieser Abschnitte wurde blockweise Schätzung des differentiellen Widerstands r_d und des (fiktiven) Kurzschlussstromes I_0 nach Gl. 1 bestimmt,

$$I = U/r_D + I_0, \quad (1)$$

$$\text{res} = \sum |I - U/\hat{r}_D - \hat{I}_0|, \quad (2)$$

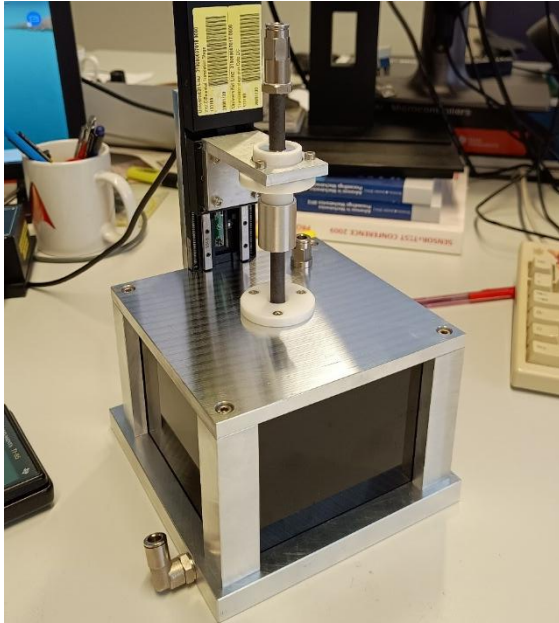


Abb. 6: Versuchsmesszelle für Lichtbogenversuche.

und die Residuen nach Gl. 2 berechnet. Bereiche mit geringen Werten deuten auf stabile Verhältnisse hin, wohingegen hohe Werte Lichtbogenübergänge anzeigen, die auf eine Änderung der Einschlagpunkte auf dem Schmelzbad oder einer Änderung des Arbeitspunktes hindeuten. Das UI-Diagramm der stabilen Abschnitte in Abb. 5 legt nahe, dass während eines Großteils des Betriebes der Lichtbogen zwischen zwei metastabilen Längen gesprungen ist und die Variationen der Zeitsignale hauptsächlich durch den instabilen Versorgungskreis hervorgerufen wurden.

1. Messzelle

Zur genauen Untersuchung der Eigenschaften des Lichtbogens wurde die Messkammer in Abb. 6 entworfen. Die vier Seiten der Kammer sind mit Schweißgläsern DIN6 der Größe 90x110mm² ausgestattet. Diese werden mittels Teflonband dicht in die Nuten der Aluminiumteile eingesetzt und bieten Schutz vor Blendung und UV-Strahlung. Die Deckelplatte beinhaltet eine wärmestabile Teflondurchführung für die Hohlelektrode die mittels Linearachse in vertikaler Richtung bewegt werden kann. Die Elektrode liegt gegen dem Aluminiumgehäuse auf negativem Potential und wird durch das Teflonelement zusätzlich elektrisch isoliert. Es ist entsprechend gestaltet, damit beim Aufziehen des Lichtbogens kein Überslag von Elektrode auf Deckelplatte entsteht. Das runde Aluminium-Element unterhalb des Elektrodenhalters in Abb. 6 ist ein Federelement. Dieses limitiert den Anpressdruck der Elektrode bei Kontakt-

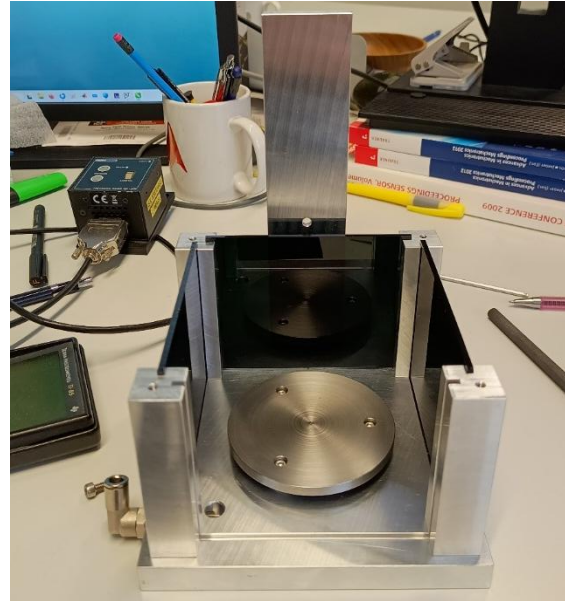
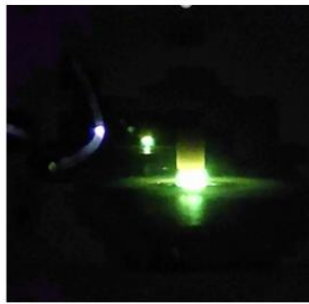


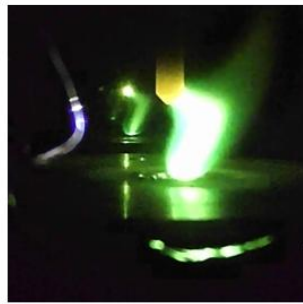
Abb. 7: Die geöffnete Messkammer.

zündung. Dies dient zum einen als mechanischer Schutz der Elektrode, wird aber auch dazu verwendet durch leichten Anpressdruck einen vergleichsweise hohen Kontaktwiderstand herzustellen der es erlaubt die Kontaktstelle vor der Plasmazündung vorzuheizen. Die Graphitelektrode hat einen Nenndurchmesser von 8mm und eine Länge von 200mm. Die Bohrung über die das Gas zugeführt werden kann, besitzt einen Durchmesser von 2mm. In den folgenden beschriebenen Versuchen wurden jedoch Vollelektroden verwendet. Bei diesen Versuchen wird das Gas über den Gasanschluss auf der Deckelplatte (links hinten in Abb. 6) oder über die Bodenplatte (rechts unten) zugeführt. Abb. 7. Zeigt die geöffnete Messkammer. Die runde Kontaktplatte besteht aus Baustahl St52 und ist mittels Keramikspacern thermisch von der Bodenplatte entkoppelt, aber über drei Schrauben M3 elektrisch mit dem Gehäuse verbunden. Erste Versuche wurden unter Verwendung des Labornetzgerätes HP6269B (Thyristorvorstufe mit nachgeschaltetem Linearregler, 45V, 50A) durchgeführt.

Das Etablieren des Lichtbogens erfolgt durch Kontaktzündung und nachfolgendem Aufziehen (siehe Abb. 8). Es zeigt sich, dass beim Verlassen der Kontaktbedingung, der zur Aufrechterhaltung des Stromes nötige sprungartige Spannungsanstieg nicht schnell genug bereitgestellt werden konnte, wodurch die Zündung nicht funktionierte.



Kurz nach Kontaktzündung



Lichtbogen tritt normal auf Elektrodenfläche aus



Längerer Lichtbogen -> höhere Leistung -> höherer Strömungsauftrieb -> Lichtbogen tritt normal auf Elektrode aber an kälterer Stelle aus (Der Strömungsauftrieb dominiert)



Frühphase der Plasmainstabilität



Plasmainstabilität kurz vor Plasmaabbriss -> Starkes Flattern an Elektrode

Abb 8: Kameraaufnahmen des Lichtbogens (allerdings ohne Argoneinblasung) bei unterschiedlichen Elektrodenabständen mit herkömmlichen Labornetzgerät (45V/50A) im Konstantstrombetrieb bei 20A. Längere Lichtbögen gehen mit höheren thermischen Energien einher, welche Instabilität begünstigen. Die Plasmafluktuationen kurz vor Abriss gehen mit der starken Oszillation des Versorgungskreises einher. Der schnelle optimierte Stromsteller eliminiert diese Instabilität und erlaubt das Erreichen höherer Lichtbogenleistungen.

Durch Hinzuschalten einer Induktivität von $L = 100\text{mH}$ mit einem Serienwiderstand von $R_s = 1\Omega$ konnte dieses Problem behoben werden. Jedoch kam es im höheren Leistungsbereich zu starken Oszillationen des Lichtbogens, die vom Netzteil induziert wurden.

moduliertes (PWM) Signal (Tastverhältnis $p = [0 \dots 1]$). Die Ausgangsspannung U_A beträgt bei Eingangsspannung U_B (unter Vernachlässigung der Diodenspannung und des Basis-Kollektor Abfalls des IGBT)

$$U_A = p U_B \quad (3)$$

2. Leistungselektronik

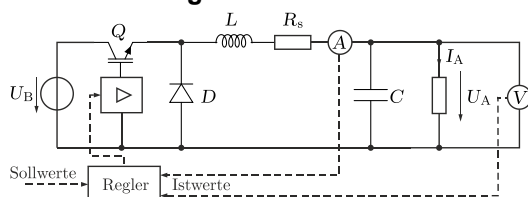


Abb. 9: Schaltbild des Tiefsetzstellers.

Wegen der oben angeführten Einschränkungen wurde ein auf einem Tiefsetzsteller (Buck-Converter) basierter Stromregler implementiert, der dem Labornetzteil nachgeschaltet wird. Diese Topologie (siehe Abb. 9) eignet sich besonders für diese Aufgabe, da die Ausgangsinduktivität den Ausgangsstrom bereits inhärent kurzzeitig stabilisiert. Die Eingangsgröße des Stellers ist ein Puls-Pausen

Dies bedeutet, dass ein überlagerter Stromregler den Wert p entsprechend der dynamischen Last und des Stromsollwerts nachregeln muss.

$$\Delta I_A = U_B (1 - p) p / (L f_{osc}) \quad (4)$$

$$\tau = L / R_{LB} \quad (5)$$

Rippelstrom ΔI_A und Zeitkonstante τ sind durch Gl. 4 und 5 gegeben, wobei R_{LB} den (positiven) Lichtbogenwiderstand beschreibt. Für geringe Rippelströme und Zeitkonstanten ist daher eine hohe Schaltfrequenz nötig. Der eingesetzte IGBT vom Typ IXXN100N65CH4H1 von IXYS eignet sich für Schaltfrequenzen bis zu 60 kHz, sodass eine Schaltfrequenz $f_{osc} = 50\text{ kHz}$ verwendet wird. Die Induktivität ist vom Typ Hammond 195C75 mit $L = 2.5\text{ mH}$ und $R_s = 8\text{ m}\Omega$.

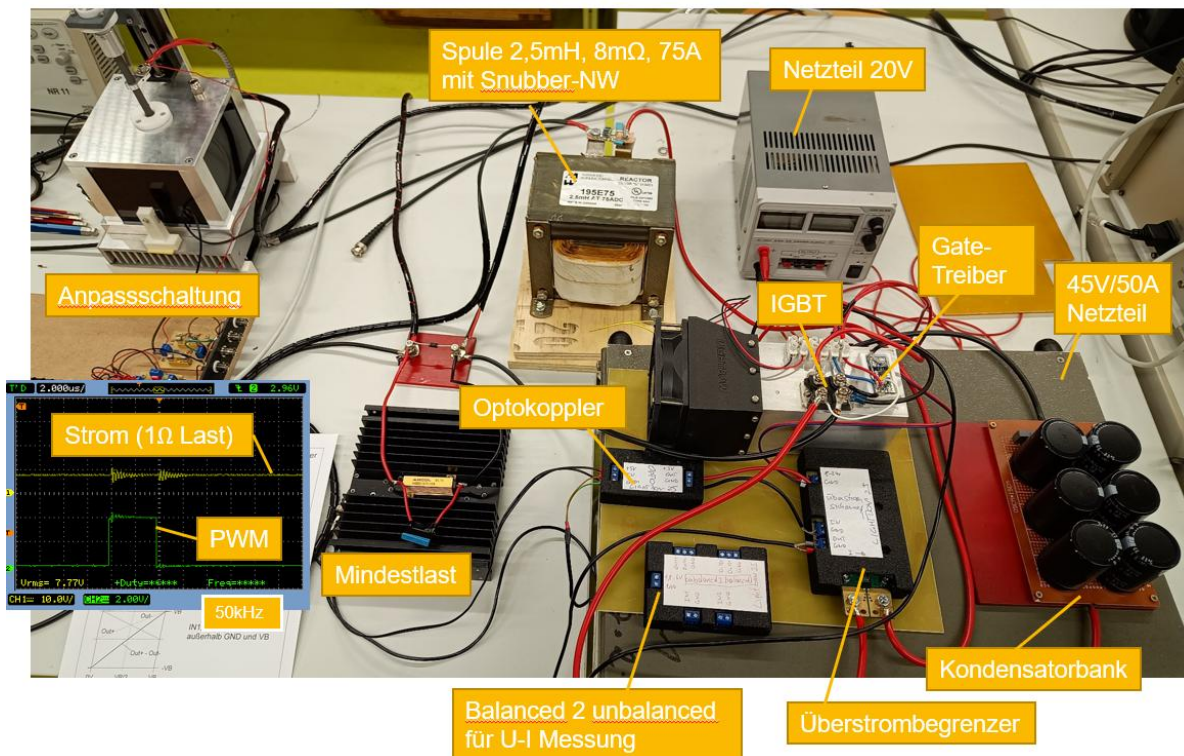


Abb. 11: Messaufbau mit elektronischen Komponenten. Das Regelsystem (nicht im Bild) misst den Strom mittels Flux-Gate Sensor und stellt die PWM des Tiefsetzstellers entsprechend einem Sollwert.

Abbildung 10 zeigt Spice-Simulationen des Regelverhaltens an einer konstanten Last von $R_{LB}=1\Omega$. Es wurde ein einfacher PI-Regler verwendet. Die Sollstromsprünge von 0A auf 5A, 10A, 20A, 30A, und 40A folgen mit der in Gl. 5 gegebenen Zeitkonstante $\tau = 2.5\text{ms}$. Stromänderungen auf höherem Stromniveau erfolgen daher naturgemäß langsamer. Die in der Simulation bestimmten Rippelströme $\Delta I_A = 40\text{mA}$, 65mA, 86mA, und 62mA stimmen mit den Werten, die durch Gl. 4 gegeben sind, überein. Der Maximalwert tritt bei $p = 0.5$ auf (20A). Abbildung 11 zeigt die elektrischen Komponenten des experimentellen Aufbaus. Die Kondensatorbank (60mF) wird verwendet um zu verhindern, dass beim Ausschalten des IGBTs Spannungsspitzen die Überspannungsabschaltung des Netzteils auslösen. Um den IGBT vor hohen Stromspitzen die bei Kurzschlussbedingungen durch die Kondensatoren hervorgerufen werden können, wurde eine Hysterese behaftete Überstromabschaltung basierend auf einem Hallstromsensor mit schneller Ansprechzeit ($<2\mu\text{s}$) implementiert.

Bei Kontaktzündung heizt sich die Graphitelektrode auf, was das Herauslösen der zur Erzeugung des Plasmas benötigten Elektronen erleichtert. Das Potential der

Graphitelektrode ist daher gegenüber der Bodenplatte negativ. Der Tiefsetzsteller in dieser Form liefert eine positive Spannung, weshalb zusätzlich galvanische Trenneinrichtungen für das Steuersignal und die gemessenen Größen eingesetzt werden.

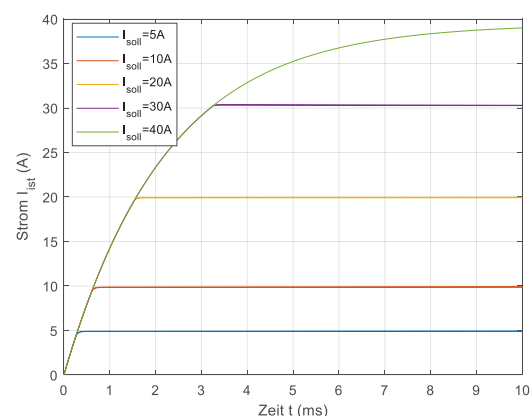


Abb. 10: Simuliertes Regelverhalten des Tiefsetzstellers bei verschiedenen Stromsollwerten.

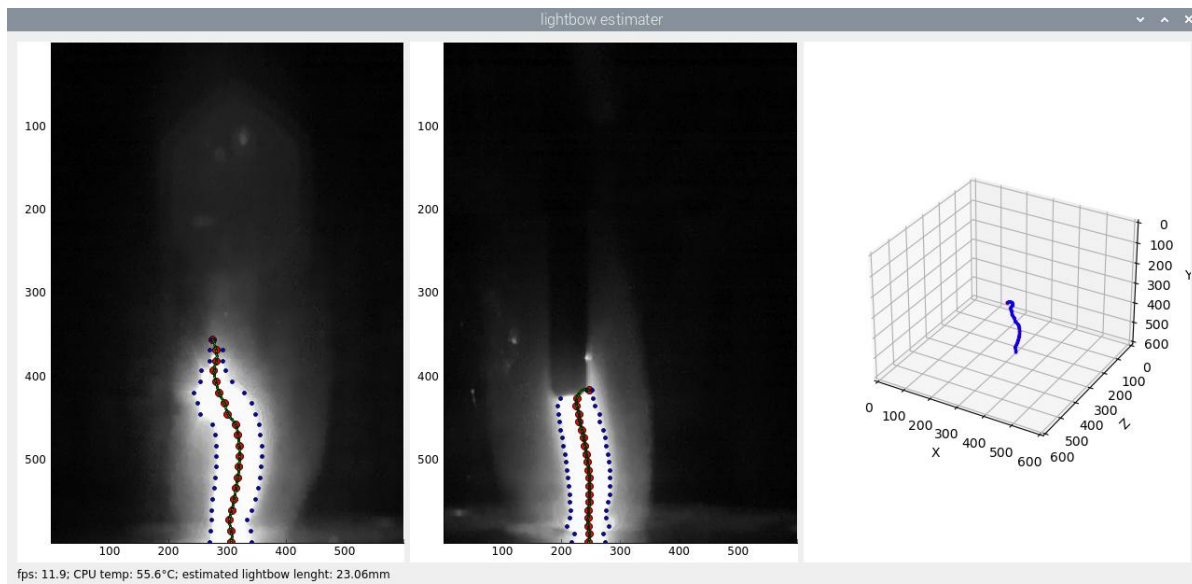
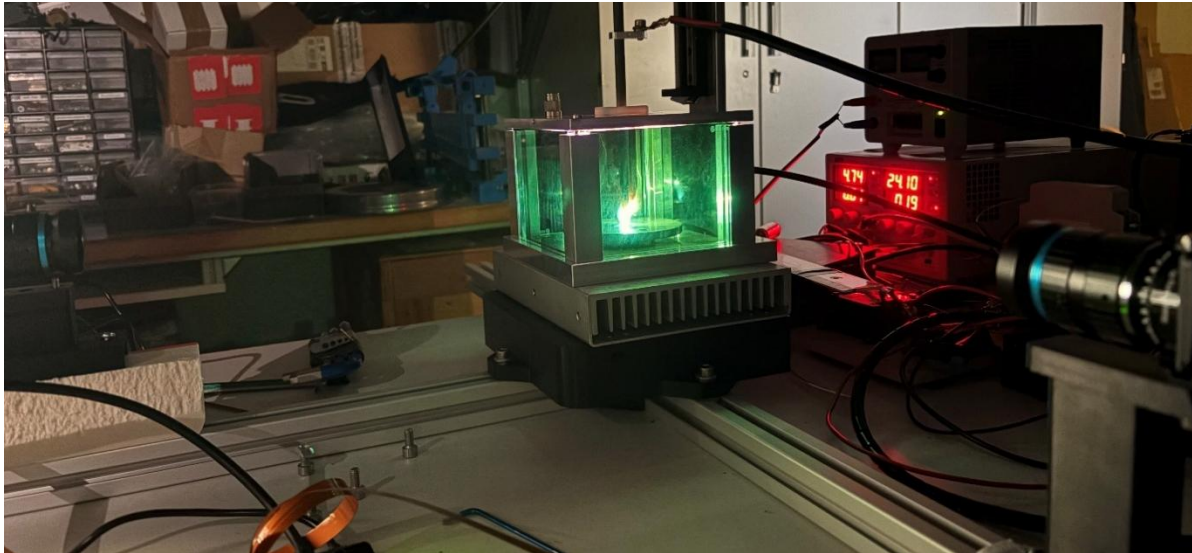


Abb. 12: Stereokamerasystem und Live-View der Lichtbogen-Rekonstruktion. Die Updaterate der geschätzten Längen und der Energien erhöht sich bei Deaktivierung der Visualisierung auf ca. 110Hz.

3. Regelungssystem und Implementierung

Zur Realisierung der Stromregelung wird ein industrieller Prozessrechner eingesetzt, auf welchem die Regelalgorithmen in der Echtzeitumgebung TwinCAT der Firma Beckhoff ausgeführt werden. Die Steuerung arbeitet mit einer festen Abtastzeit von 125 μ s, wodurch eine ausreichend hohe zeitliche Auflösung erreicht wird, um die schnellen dynamischen Effekte des Lichtbogens sowie der Leistungselektronik präzise zu erfassen und zu kompensieren. Die kurze Zykluszeit ermöglicht insbesondere die gezielte Dämpfung jener Oszillationen, die aus dem Zusammenspiel von negativer differentieller Leitfähigkeit des Lichtbogens und der

parasitären Energiespeicher im Versorgungspfad resultieren. Der implementierte Regler ist als adaptiver Zustandsregler mit Störgrößenkompensation realisiert. Ziel der Regelung ist die präzise Vorgabe des Lichtbogenstroms trotz stark variierender Lastcharakteristik. Der Lichtbogen stellt eine nichtlineare und zeitvariante Last dar, deren effektiver differentieller Widerstand sich sowohl mit der Bogenlänge als auch mit der lokalen Plasmatemperatur ändert. Zusätzlich wirkt die Lichtbogen Spannung als dominierende Störgröße, da sie direkt in die Dynamik der Spulenstromänderung eingeht jedoch durch eine entsprechende Vorsteuerung kompensiert werden kann. Die Kompensation verhindert, dass Änderungen der Lichtbogenlänge unmittelbar zu Stromabweichungen führen. Die Störgrößenkompensation verbessert somit

sowohl die Führungsdynamik als auch die Robustheit gegenüber abrupten Zustandsänderungen des Plasmas, beispielsweise beim Übergang zwischen metastabilen Bogenlängen oder bei kurzfristigen Bogenverlagerungen auf der Badoberfläche / Metallplatte. Der Regler passt seine Parameter adaptiv an Änderungen des differentiellen Lichtbogenwiderstands an, wodurch eine gleichbleibende Dynamik über einen weiten Arbeitsbereich von typischerweise 5 A bis 40 A erzielt werden kann. Die resultierende geschlossene Stromregelstrecke weist ein gut gedämpftes Verhalten mit kurzer Einschwingzeit auf und erlaubt eine stabile Einstellung auch bei stark schwankenden Plasmabedingungen. Die hardwarenahe Anbindung der Leistungselektronik erfolgt über eine Mikrocontrollerplattform auf Basis des Texas Instruments F28P65 (C2000-Serie). Der Mikrocontroller wird als EtherCAT-Slave betrieben und stellt die zeitkritischen Ein- und Ausgangseinheiten bereit. Dazu zählen insbesondere:

- Generierung des PWM-Signals für den IGBT des Tiefsetzstellers mit hoher zeitlicher Auflösung
- Erfassung der Lichtbogenspannung mittels schneller Analog-Digital-Wandlung
- Messung des Spulenstroms über einen Fluxgate- bzw. Hall-basierten Stromsensor
- deterministische Kommunikation mit dem übergeordneten Prozessrechner über EtherCAT

Durch die Auslagerung der zeitkritischen Signalverarbeitung auf den Mikrocontroller können sehr geringe Totzeiten erreicht werden. Gleichzeitig ermöglicht die EtherCAT-Kommunikation eine synchrone Einbindung in das Echtzeitsystem des Prozessrechners. Die resultierende Gesamtstruktur kombiniert die hohe Rechenleistung und Flexibilität der TwinCAT-Umgebung mit der schnellen Peripherieanbindung der C2000-Mikrocontrollerplattform. Die gewählte Architektur erlaubt zudem eine einfache Erweiterung um zusätzliche Regelungsfunktionen, beispielsweise eine überlagerte Leistungsregelung oder eine modellbasierte Schätzung der Lichtbogenlänge auf Basis der elektrischen Signale. Insgesamt zeigt sich, dass durch die Kombination aus schneller Echtzeitregelung, Störgrößenkompensation und leistungsfähiger leistungselektronischer Aktorik eine deutliche Verbesserung der Stromstabilität erreicht werden kann.

4. Instrumentierung

Die Lichtbogenstabilität wird als zeitliche Stabilität der Lichtbogenlänge definiert. In der Zielanwendung ist die optische Beobachtbarkeit des Lichtbogens in der Regel nicht gegeben, weshalb am Prüfstand untersucht wird inwieweit die elektrischen Signale mit der Lichtbogenlänge korrelieren. Dazu wurde ein Stereokamerasystem entworfen. Die verwendeten Kameras sind 1.6 Megapixel Global-Shutter mit Sony-IMX296 Sensor für Raspberry Pi mit 16mm Teleobjektiv. Der Global-Shutter bewirkt eine zeitgleiche Belichtung aller Bildpunkte des Sensors. Der Kamerateyp erlaubt die Festlegung einer Region-of-Interest (ROI), bei welcher nur ein Bildbereich ausgelesen werden muss, was die Bildrate wesentlich erhöht. Der Objektstand und die Blendeneinstellung wurden so gewählt, dass ein würfelförmiger Bereich mit 50mm Seitenlänge scharf abgebildet werden kann. Die Stereobilder werden auf einem Raspberry Pi 5 synchron aufgezeichnet. Der Lichtbogen wird in zwei orthogonalen Ebenen identifiziert und mittels Spline Funktionen interpoliert (siehe Abb. 12 unten). Die Länge wird aus der Mittellinie des Lichtbogens berechnet. Die Messmethode wurde mit gebogenen Drähten bekannter Länge kalibriert. Die Messdaten werden mit einer Updaterate von > 100Hz zum Prozessrechner übertragen.

5. Ergebnisse

Abbildung 13 zeigt gemessene Ströme, Spannungen und Lichtbogenlängen. Nach Kontaktzündung wird der Lichtbogen aufgezo-gen. Der Stromregler wurde auf Sollwerte im Bereich von 10-22A eingestellt. Strom-sollwertänderungen erfolgten konsistent mit einer Anstiegszeit (10-90%) von ca. 4ms. Ab 450s wurde der Lichtbogen bis zum Abriss aufgezo-gen. Die Länge und die Spannung korrelieren, bedingt durch die Abrisse, naturgemäß. Abb. 14 zeigt daher die Korrelation im Zeitbereich zwischen 285s bis 318s. In diesem Abschnitt war der Strom konstant 18A (Variationskoeffizient 0.03%). Die Spannung und die Länge fluktuieren ohne erkennbaren Trend. Trotzdem ist eine Korrelation zwischen den beiden Größen ersichtlich. Der Längeneinbruch bei 350s ist ein Messartefakt. Die Elektrode verdeckte den Lichtbogen in den Bildern der beiden rechtwinkelig angeordneten Kameras.

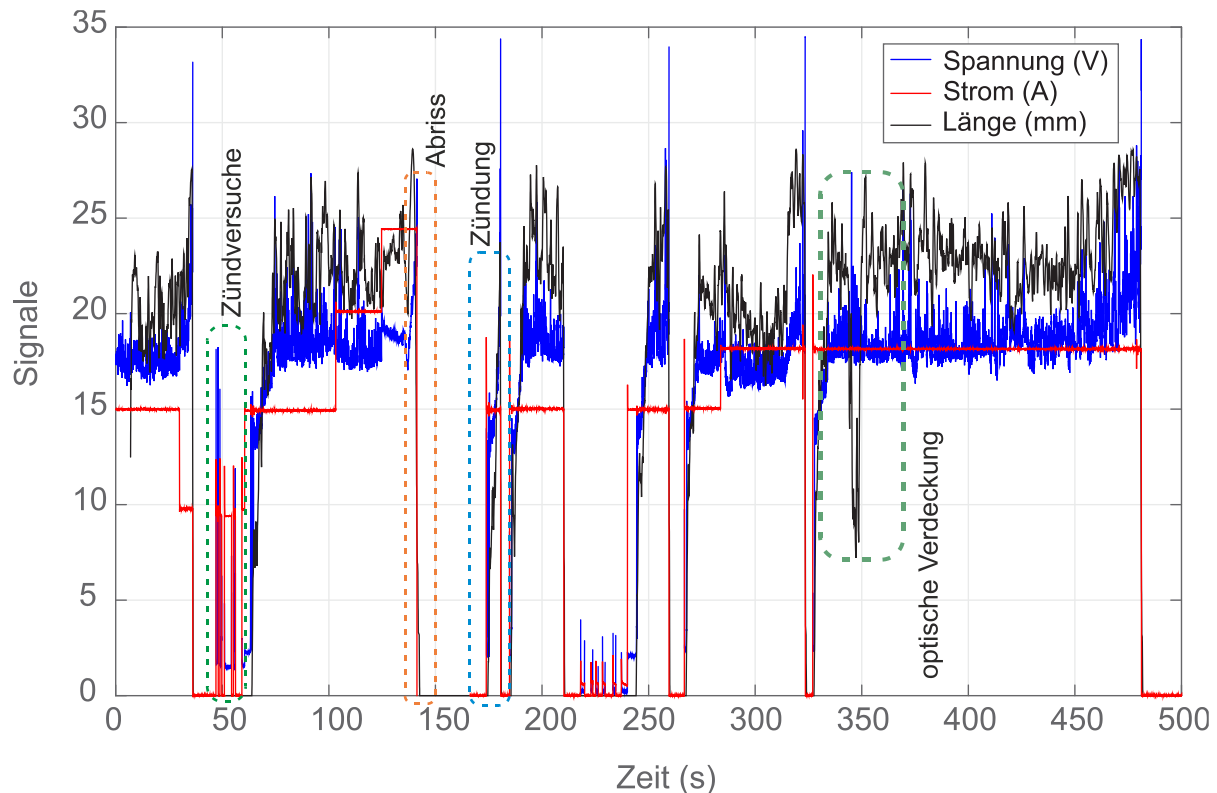


Abb. 13: Strom, Spannung und geschätzte Länge des Lichtbogens.

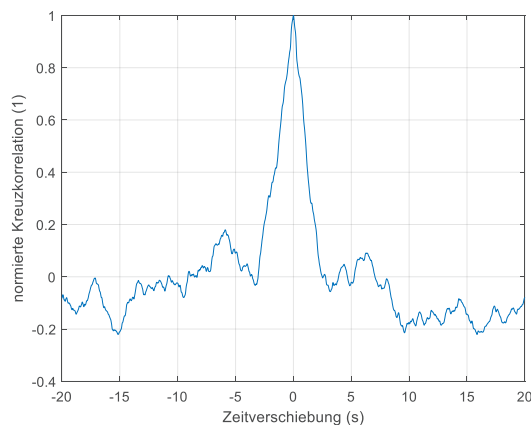


Abb. 14: Normierte Kreuzkorrelationsfunktion von Lichtbogenlänge und Spannung im Bereich 285s-318s.

Ausblick

Die größere Messkammer mit einer thermischen Nennleistung von 5KW wird gegenwärtig realisiert. Diese besitzt bereits eine ähnliche Dimension wie die metallurgische Versuchsanlage, sodass die hier erforschten Mess- und Regelstrategien übertragbar sind. Der Leistungsteil wird durch einen speziell entwickelten, beliebig parallelisierbaren, GaN-FET basierten Leistungssteller ersetzt. Die einzelnen Module werden dazu mit mehrphasig

mit einer Schaltfrequenz von bis zu 500kHz betrieben. Dies verspricht noch bessere Kontrolle über die elektrischen Charakteristika des elektro-thermischen Systems „Lichtbogen“.

Aknowledgement

Das Projekt LIGHTBOW wurde durch die Österreichische Forschungsförderungsgesellschaft (FFG) im Rahmen von Expedition Zukunft Wissenschaft gefördert.

Literaturnachweis

- [1] Zarl, M. A., Farkas, M. A., & Schenk, J. (2020). A study on the stability fields of arc plasma in the HPSR process. *Metals*, 10(10), 1394.
- [2] Leitner, E. M. (2023). Methodenentwicklung zur Stabilisation des Elektrischen Lichtbogens im Wasserstoff-Argon Schmelzreduktionsplasma.
- [3] Ernst, D., Zarl, M. A., Farkas, M. A., & Schenk, J. (2023). Effects of the Electrodes' Shape and Graphite Quality on the Arc Stability During Hydrogen Plasma Smelting Reduction of Iron Ores. *steel research international*, 94(7), 2200818.
- [4] Rieder, W. (1967). Plasma und Lichtbogen, Friedr. Vieweg & Sohn, Braunschweig.

CNN-based classification of scale deposits and damages in drainage pipes using a mobile inspection rover

*Veronika Putz¹, Christina Hofer¹, Sabrina Fleischanderl¹, Leander B. Hörmann¹,
Johannes Schröck¹, Wolfgang Hofer¹, Harald Hölzl², Tobias Schachinger², Vesna Micic Batka²*
¹Linz Center of Mechatronics GmbH, Altenberger Straße 69, A-4040 Linz
²ÖBB-Infrastruktur AG, Praterstern 3, A-1020 Wien

Abstract

A convolutional neural networks (CNN)-based approach for automated classification of scale deposits and structural damage in drainage pipes of railway tunnels is presented and validated using data collected by a mobile inspection rover - a waterproof, battery-powered robot equipped with cameras that record video data for offline analysis. For evaluation, two deep learning approaches are considered: image classification using VGG16 and object detection using YOLOv11. Both models were trained and validated using data acquired during field tests in Austrian railway tunnels. While VGG16 enables simple frame-wise classification, YOLOv11 provides detailed spatial information and allows parallel detection of defects and various severity levels of scaling per frame. Results showed reliable detection of severe defects, with some confusion between adjacent severity levels. To support planning of drainage pipes maintenance, a post-processing method was introduced that transformed frame-wise detections into a time-series based indicator implying the condition of pipe segments. The approach demonstrated the potential of combining autonomous inspection with machine learning in the future for condition-based and predictive maintenance of tunnel drainage systems.

Keywords: Predictive Maintenance, Fault Detection, Autonomous Inspection, Machine Learning, Tunnel Drainage, Scale Deposits

Introduction

In pressure-relieved railway tunnels, drainage systems are installed to protect tunnel structures against hydrostatic pressure caused by groundwater. Drainage is typically realized using slotted pipes that run parallel to each side of the tunnel. The long-term functionality of these drainage systems can be significantly impaired by the formation of deposits within the pipes: Especially at inner pipe surfaces and within the drainage slots, carbonate minerals, hereafter referred to as "scale deposits", can accumulate. In the drainage system, calcium carbonate originates primarily from mineralized groundwater as well as from leaching of cement used in the tunnel construction [1,2]. The rate at which the scale deposits form depends among other factors, on precipitation and the resulting variations in groundwater inflow. Under favorable conditions, deposits can rapidly accumulate within a short period of time. As a result, the effective cross-section of the drainage system reduces, and the drainage slots can be clogged by the accumulated deposits. To restore the functionality of the drainage pipes and to avoid damage, regular drainage cleaning is required. Such main-

tenance work is costly and usually requires closing the railway tunnel for traffic.

We are therefore developing an automated inspection system capable of autonomously inspecting the drainage and operating without tunnel closures. Regular automated inspections, combined with systematic classification and spatial mapping of scale deposits and potential damages within the piping system would allow monitoring growth rates of scale and hence would support not only the condition-based maintenance, but also predictive maintenance strategies.

The development of such an inspection system, capable of autonomously traversing tunnel drainage pipes, is associated with several technical challenges [3,4]. The used drainage pipes typically are made of thermoplastics, have a circular or bottom-flattened cross-section and have narrow diameters between 160 mm and 300 mm. These diameters may already be constricted by scale deposits, reducing the available cross-section for an inspection system. Lack of reliable radio communication within the pipes requires the system to move autonomously through the drainage despite pipe bends and branching points. It must be able to orientate, identify obstacles, and to collect meaningful data that

enables an assessment of the conditions inside the pipe. Furthermore, the drive of the inspection system must be able to cope with difficult conditions: Scale deposits at the bottom of the pipe can exhibit different consistencies, ranging from hard, consolidated layers to soft, sludge-like material [1], and can hinder the motion of an inspection system.

In this work, we present a demonstrator capable of moving through the drainage pipes and collecting meaningful data. An overview of feasible AI methods for classification of scale and damages in video data is provided, and two different models are implemented and validated using data collected in real railway tunnels. Finally, a method to post-process raw results, providing easy-to-read visualization for the state of the drainage system, is shown.

Autonomous inspection rover

The centerpiece of our approach is a field-proven demonstrator - a waterproof robot, that moves through drainage pipes autonomously. It is wireless, battery-powered, and equipped with sensors for obstacle detection essential for autonomous movement and recording of video data for offline evaluation. The demonstrator builds upon an initial concept, which was introduced in [3], but has been significantly improved based on experience from field tests.

Figure 1 above shows a technical drawing of the demonstrator. It consists of the rover, equipped with a crawler-track drive capable of moving forward and backward, and a sensor lance attached to its front. Figure 1 below shows the demonstrator, being immersed in a maintenance hole of the drainage system at the beginning of an inspection drive.

The sensor lance carries a red ring laser and an integrated camera, which are used for obstacle detection. At its tip, the lance contains a second integrated camera, which is used for recording video data from within the drainage pipe. An Arducam OV5647 camera module with a wide-angle lens (160° diagonal field of view), connected to a Raspberry Pi Zero 2W board, is used. A white-light LED source placed at the tip of the sensor lance provides the necessary background illumination for video recordings in the drainage pipe with no ambient light. During inspection, videos from this camera are captured at a resolution of 1920 x 1080 pixel and a frame rate of 30 frames per second (fps). Videos are stored locally on the demonstrator for offline analysis.

To allow building a more generalized database for defect detection and classification, instead of the sensor lance, a high-quality external camera (GoPro 10) can

be clamped onto the demonstrator, recording videos with a resolution of 3840 x 2160 pixel at a frame rate of 30 fps. Data acquired using these two cameras was used as input for training AI models for detecting scale deposits and defects like damaged pipe walls.

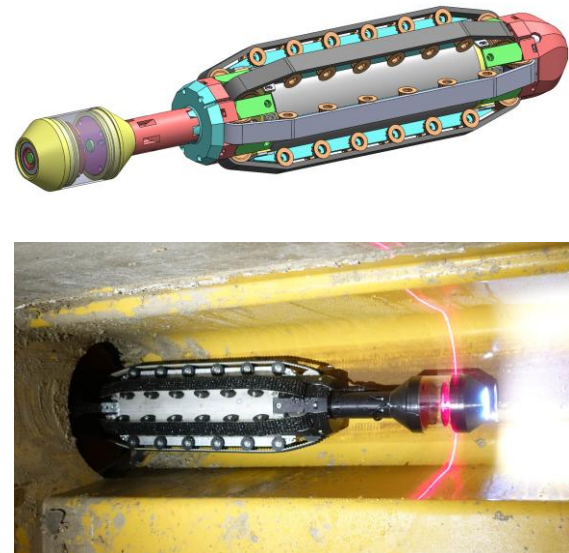


Figure 1: Technical drawing of the inspection rover (above), rover immersed in the drainage system (below)

During the inspection, the drive of the demonstrator is set to a constant velocity of approximately 0.3 m/s. However, its exact velocity may vary, since unknown drag, due to sludge-like sintering, low-traction sections and various obstacles, may reduce the set velocity. The true speed of the demonstrator is not yet measured, and the location, at which an individual video frame is recorded inside the drainage, can only roughly be reconstructed from its time stamp. In the next steps we plan to integrate detection of drainage branching points and maintenance holes as well as velocity estimation methods for improved location accuracy of the robot within the pipe.

Recorded data

The demonstrator was used to collect data in drainage systems of two operational railway tunnels in Austria during short maintenance closures:

- In the “Pummersdorfer Tunnel”, approx. 30 minutes of video data were collected using both cameras. Nine measurement runs were conducted with and against the flow of the water, starting at different entry points. A total length of approx. 550 m was covered. In this tunnel, only

slight scale and small isolated damages in the pipe walls occurred.

- In the “Sieberg Tunnel”, in four segments of the tunnel, 10 measurement runs were conducted with and against the flow of the water using both cameras, providing a total length of 37 minutes of video data. In the videos obtained, severe examples of both scaling and damage could be captured.

The structure and appearance of scale deposits can considerably vary, depending on their chemical composition [1]. Additionally, the color of the pipe material and the shape of the cross section of the pipe (circular, flattened) can differ in the field. The video material available from the experiments carried out covers various scenarios. Some examples are shown in Figures 2 to 5. Figure 2 shows a pipe made of blue thermoplastic material, partially filled with water, with scaling as well as a hole in the pipe wall visible on the right. Figure 3 shows a similar scene with more severe scaling in a pipe with reddish-brown walls.



Figure 2: Medium to severe scaling, visible hole in the pipe wall



Figure 3: Medium to severe scaling at the bottom, slight scaling at the side wall

A different scenario with severe scaling as well as formation of algae is shown in Figure 4, while Figure 5 contains isolated scaling occurring only at one side wall of the pipe. Note that Figures 4 and 5 were recorded using the high-quality external camera, while Figures 2 and 3 were recorded using the low-quality camera integrated in the sensor lance.



Figure 4: Medium to strong scaling at the bottom, formation of algae



Figure 5: Strong scaling with reddish color at the side wall of the pipe

Relevant approaches for detection and classification of scale and damage

Considering the strong variability in the appearance of scale and damages as well as possible variations in color and shape of the drainage pipes, convolutional neural networks (CNNs) provide a more robust and adaptive approach for classifying the state of the drainage system, compared to classical image analysis methods. Three fundamental approaches can be distinguished:

- **CNNs for image classification** like VGG16 [5], InceptionNet [6] and ResNet [7] assign a single class label to an entire image. Each video frame is treated as an independent image and receives one class prediction. Models only provide information about the presence of a class in the video frame but do not localize objects or indicate their position or size within the image.
- **CNNs for object detection and classification** like R-CNN [8], SSD [9,10], or YOLO [11], detect and classify multiple

objects within a single video frame. These models return bounding boxes that mark the position and size of each detected object along with the corresponding class label and a confidence score. Several instances from multiple classes can be detected within a single video frame. Additional information regarding the position and size of a detected object is available.

- **CNNs for image segmentation** provide pixel-level classification and therefore allow a more detailed analysis of objects. In *semantic segmentation* models such as U-Net [12, 13], each pixel in an image frame is assigned to a class, resulting in a segmentation map that distinguishes different object categories but does not separate individual objects of the same class. *Instance segmentation* models such as Mask R-CNN [14] detect individual objects and generate a separate segmentation mask for each instance, enabling precise localization of object boundaries.

Each of the three fundamental approaches provides a different level of information and is related to different requirements regarding training data volume, annotation effort and computational complexity: Image classification models require comparatively simple annotations - only a single label per video frame needs to be assigned. Object detection approaches require bounding box annotations for each object instance, while segmentation methods demand pixel-level labeling, which significantly increases the manual labeling effort as well as the amount of required training data. For the use-case in our study, defects such as severe damage or scale only rarely occur along the length of the pipes, leading to low availability of relevant training data for this class. Therefore, only two of the three fundamental approaches are considered for further use at this stage: VGG16 (due to its architectural simplicity, stable behavior, and ease of use) and YOLOv11 [15] (due to its ability for fast and accurate object detection).

Annotation of training data

Defect classes for scale deposits and damage were defined in consultation with field experts responsible for inspecting and maintaining tunnel drainage systems. According to their recommendation, the acquired dataset was annotated manually, using different severity levels for both damages and scale deposits. Due to the different working principles of

VGG16 and YOLOv11, annotation was handled individually for the two models.

For VGG16, each video frame can only have a single class label. It was decided to label frames according to the most severe visible error in the frame (e.g., if the same video frame contains instances of moderate and severe scaling, the label “severe scaling” is assigned). Since instances of scaling and damage may occur in the same video frame, two individual VGG16 models were used for classifying scaling and damages. Each model needed to be trained independently. Consequently, also two independent label sets were created for the training data. The following class labels were assigned:

- Label set 1 (damage): slight | moderate | severe
- Label set 2 (scaling): slight | moderate | severe

For object detection using YOLO, the release YOLOv11-OBb¹ was selected. It allows defining oriented bounding boxes, which - due to the given camera viewing angle - is highly beneficial in our use-case. For this model, if applicable, multiple defects were annotated in each individual frame by drawing bounding boxes around every visible defect. Size and orientation of the bounding boxes were adapted to best fit the visible size of the defect or scaling. The following labels were assigned to a single label set:

- Scaling: slight | moderate | severe
- Damage: slight | moderate | severe | very severe

Note that for object detection using YOLO, an additional class for ‘very severe’ damage was introduced. This class is only applied to rare cases in which e.g., a larger fragment of the pipe wall is damaged or even missing. Due to the lack of the ability to represent multiple classes within a single frame, this was omitted for VGG16; corresponding damages were instead included in the ‘severe’ class.

Classification of data using VGG16

Two individual VGG16 models for detecting damage and scale deposits were trained, using 2073 images for training, 445 images for validation, and 445 images (not part of the training data) were used to test the performance of the obtained models. Table 2 lists the total occurrence of each error for damage and scaling for the various error classes in the test dataset.

¹ <https://github.com/ultralytics/ultralytics>

Tab. 1: Total occurrence of each class in the test dataset used for VGG16

Error classes	Total occurrences	
	Damage	Scaling
none	86	86
slight	214	214
moderate	120	120
severe	25	25

The respective performance of the models for detection of damage and scaling is shown using normalized confusion matrices (see Figures 6 and 7). The classification result obtained from the models is good: For each class, more than 90% occurrences were detected correctly. Confusion between classes rarely occurs and is limited to confusion between neighboring severity levels.

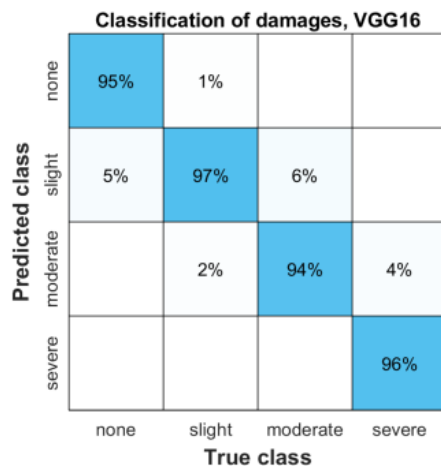


Figure 6: Normalized confusion matrix for classification of damages

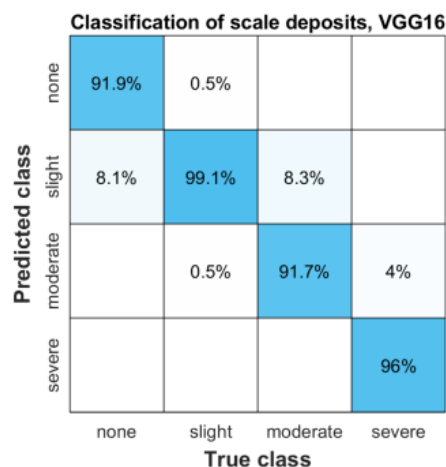


Figure 7: Normalized confusion matrix for classification of scaling

The obtained accuracy, which is calculated considering absolute occurrences of correct and false classifications, is 95.73% for damages and 95.51% for scaling.

Classification of data using YOLOv11

A single YOLOv11 model was trained using 609 video frames for training and 122 video frames for test. The datasets for training and test were selected to be mutually exclusive. Table 2 lists the total occurrence of each error class in the validation dataset in terms of video frames and in terms of instances. It is obvious that more than one instance of each class occurs in a single video frame.

Tab. 2: Total occurrence of each class in the test dataset used for YOLOv11

Error classes		Total occurrences	
		Frames	Instances
damage	slight	22	44
	medium	23	46
	severe	6	11
	very severe	3	6
scaling	slight	61	325
	medium	98	399
	severe	59	189

Figure 8 shows the obtained confusion matrix. Note that in Fig. 8, the abbreviation of row and column labels 'mod.' and 'v.s.' stand for 'moderate' and 'very severe', respectively. Confusion occurs between slight and moderate damages, as well as between slight, moderate and severe scaling. Severe and very severe damages were perfectly recognized (however, the total occurrence of these error classes in the test dataset is low). There is no confusion between damage and scaling.

The obtained accuracy, which is calculated considering absolute occurrences of correct and false classifications, is 70.3% and comparably low (mainly due to confusion between the slight and moderate classes). Combining slight and moderate damages and slight and moderate scaling would increase accuracy to 91.9%.

Figure 9 shows a sample video frame with the defects detected by the CNN marked as overlay. The benefit of the method, compared to VGG16, is obvious: The ability to detect multiple errors per video frame leads to an improved assessment of the overall area of the drainage pipe affected by scaling or damage within the image, which could not have been provided using the simple classification obtained from VGG16.

Predicted class	damages				scale deposits			
	slight	mod.	severe	v.s.	slight	mod.	severe	v.s.
damages	25%	17%						
damages	75%	83%						
damages			100%					
damages				100%				
scale deposits					67%	18%		
scale deposits					33%	75%	29%	
scale deposits						7%	71%	
	damages				scale deposits			
	slight	mod.	severe	v.s.	slight	mod.	severe	v.s.

Figure 8: Normalized confusion matrix for classification of damages and scaling using YOLOv11.



Figure 9: Video frame with detected artifacts from various classes

Metric for visualizing scaling and damage

For the planning of maintenance measures based on recorded videos, raw detections of scaling and damage (as visible in Figure 9 for a single video frame) need to be further processed to provide an easy-to-read indicator for maintenance planning. A metric that visualizes the condition of an inspected section of the drainage pipe is required.

For this purpose, the total number of detected error events (scale deposits and damages of different severity levels) per frame is treated as time-series signal, which can be post-processed to obtain a meaningful indicator for the condition of the drainage pipe. Figure 10 shows the total number of detections per frame of a selected video obtained using the trained YOLOv11-OBb model described above. The video has a total duration of approx. 25 seconds (~750 video frames). It was recorded in a pipe segment with significant

scaling and several damages. A smoothed yet representative result is obtained by block-wise calculation of the median value over 15 consecutive frames, resembling a time segment of 0.5 s (Fig. 10, green line). Blocks are built with 7 frames overlapping. To avoid flattening the visible peaks in the detections per frame – especially in segments with a high density of defects – instead of the median value, the 90% percentile value can be used (Fig. 10, green line).

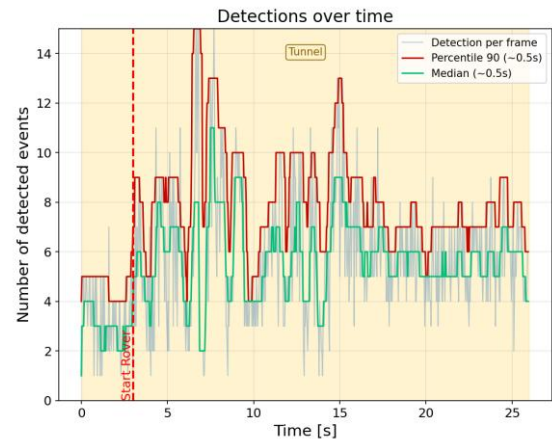


Figure 10: Total number of detections per frame vs. time stamp (raw detections and two indicators).

The obtained indicators are validated using an example image: In Figure 10, between 7 and 8 seconds after starting the experiment, a larger density of detected errors can be seen. Figure 11 shows a video frame recorded between these timestamps with the obtained detections of the CNN marked using color-coded bounding boxes, with many visible defects.

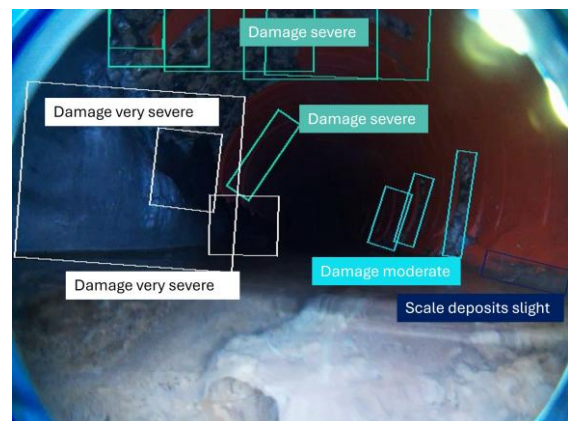


Figure 11: Sample video frame with many detections

Figure 12 shows a similar evaluation for a video recorded in a pipe segment, in which almost no errors (either scaling or damages) are visible. A sample video frame recorded approximately 37 seconds after starting the experiment is shown in Figure 13.

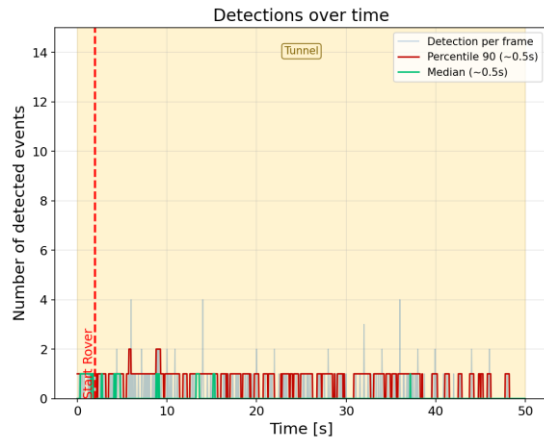


Figure 12: Total number of detections per frame vs. time stamp (raw detections and two indicators).

Comparing the evaluations in Figure 10 and Figure 12 as well as video frames shown in Figure 11 and Figure 13, the difference is obvious. By evaluating the frame-by-frame detections as time-series signal, an easy-to-read indicator for the state of a drainage segment can be obtained. Similar evaluations are feasible for each individual error class, which would allow different treatment, e.g., of severe damage and scaling and less severe error classes.

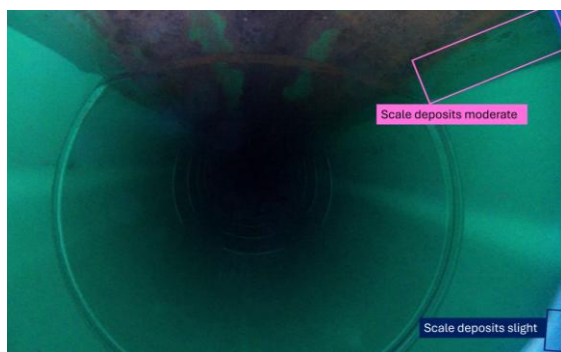


Figure 13: Sample video frame with a small number of detections

Summary and Outlook

Our next steps will focus on increasing data volumes by using further test runs in relevant scenarios, to increase reliability of the obtained results. To avoid confusion between slight and moderate scaling as well as between slight and moderate damage in the evaluation using

YOLOv11, we will furthermore consider including information obtained from the red line laser, used for obstacle detection, into the evaluation. Additionally, we will improve self-locating of the demonstrator within the drainage system using methods for velocity estimation and detection of maintenance holes and branching points.

Acknowledgement

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.

References

- [1] Eichinger, S. et al. (2020). „Assessment and formation mechanisms of scale deposits in tunnels of the ÖBB-Infrastruktur AG – A subproject of the Task Force Drainage”. *Geomechanics and Tunnelling*, 13(3), 273-285.
- [2] Lin, Z., et al. (2023). „Optimizations for tunnel drainage water system based on locations of crystallizations”. *Desalination and Water Treatment*, 299, 232–238.
- [3] Hofer, W. et al. (2023). „A robotic tunnel drainage inspection rover”. *Transportation Research Procedia*, 72, 2724-2731.
- [4] Micic Batka V. et al (2026). “Shaping the future of rail – R&D at ÖBB-Infrastruktur AG / Die Zukunft der Bahn mitgestalten – F&E bei der ÖBB-Infrastruktur AG”. *Railway Proceedings, Band 1 "Koralmittel – Technik im Fokus" / Verlag der TU Graz*
- [5] Simonyan, K. et al (2014). “Very deep convolutional networks for large-scale image recognition.” *arXiv preprint arXiv:1409.1556*.
- [6] Szegedy, Ch. et al. (2015). “Going deeper with convolutions.” *Proc. of the IEEE conference on computer vision and pattern recognition*.
- [7] He, K. et al (2016). “Deep residual learning for image recognition” In *Proc. of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- [8] Girshick, R. (2014). „Rich feature hierarchies for accurate object detection and semantic segmentation”. In *Proc. of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- [9] Liu, W. et al (2016). “SSD: Single shot multibox detector”. In *European conference on computer vision* (pp. 21-37).
- [10] Rudić, B. et al. (2021): “Robust Hierarchical Data Association for Multiple Object Tracking Under Abrupt Appearance Changes”. In *Proc. of the 2021 Int. Conf. on Electrical, Computer and Energy Technologies*
- [11] Redmon, J. et al. (2016). “You only look once: Unified, real-time object detection”. In *Proc. of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

- [12] Nguyễn, Đ. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". *Lecture Notes in Computer Science*.
- [13] Fleischanderl, S. et al. (2022). „CNN–based crack detection in oxide layers of hot rolled steel sheet samples for the validation of a pickling process model". In *Proc. of the 2022 European Symposium on Software Engineering*.
- [14] He, K. (2017). "Mask R-CNN". In *Proc. of the IEEE international conference on computer vision* (pp. 2961-2969).
- [15] Khanam, R. et al (2024). "Yolov11: An overview of the key architectural enhancements". *arXiv preprint arXiv:2410.17725*.

Additive Manufactured Sensitive Robotic Gripper Based on the Wedge-Hook Mechanism

Nikolai Hangst¹, Sebastian Wieber¹, Thomas M. Wendt¹, and Stefan J. Rupitsch²

¹ Offenburg University of Applied Sciences, Work-Life Robotics Institute, Offenburg, Germany

² University of Freiburg, Department of Microsystems Engineering, Laboratory for Electrical Instrumentation and Embedded Systems, Freiburg, Germany
 nikolai.hangst@hs-offenburg.de

Summary: In this contribution, we present an additive manufactured robotic gripper with sensitive force control. The gripper is based on a wedge-hook mechanism, resulting in a compact and cost-effective design. A key feature is an integrated force sensor located in one of the gripper jaws, which enables precise control of the gripping force directly at the gripped object. The control feature in form of accuracy, precision, and overshoot is analyzed. The sensor-integrated system is compared to a conventional approach. The results demonstrate that combining additive manufacturing with integrated sensing and controlled actuation enables a capable mechatronic system. Its accuracy exceeds that of current-based force estimation methods. This allows accurate and precise gripping without the need for complex traditional manufacturing processes.

Keywords: Sensitive robotic gripper, force control, wedge-hook mechanism, additive manufacturing, 3D printing, mechatronic system

Introduction

Additive manufacturing, also known as 3D printing, has made significant technological progress in recent years. In particular, its ability to fabricate complex, functionally integrated structures, and fully realized mechatronic systems represents a significant advancement in manufacturing technology [1, 2]. Especially in robotics, this opens up new potential for the development of innovative and complex systems [3–6]. Building on these advantages, additive manufactured robotic grippers with sensitive force control can be realized. Conventional mechanical grippers typically require extensive subtractive manufacturing processes, which are associated with high costs and long lead times [3, 7]. In contrast, additive manufacturing enables the direct fabrication of functional structures, including guiding elements. This allows the realization of modular and cost-efficient gripper designs that are particularly suited for sensitive handling tasks. A key aspect of such systems is the control of gripping forces directly at the gripped object. Industrial grippers commonly rely on motor current-based estimation methods, which are inaccurate at low force levels and under varying friction conditions [8, 9]. Integrated sensing approaches enable more precise force measurement and improved control performance, particularly in low-force applications.

Mechanical Design

The gripper consists of an additive manufactured base structure that integrates all the drive and electronic components. The sensitivity of the system is achieved through a dedicated sensi-

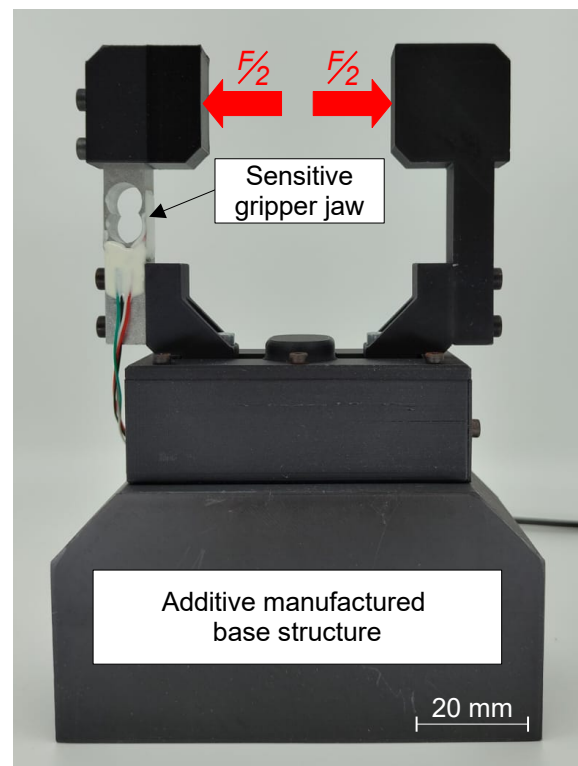


Fig. 1: Additive manufactured robotic gripper and its force directions on each gripper jaw.

tive gripper jaw in the form of a single-point load cell. The complete system, including the orientation of the gripping force, is shown in Fig. 1.

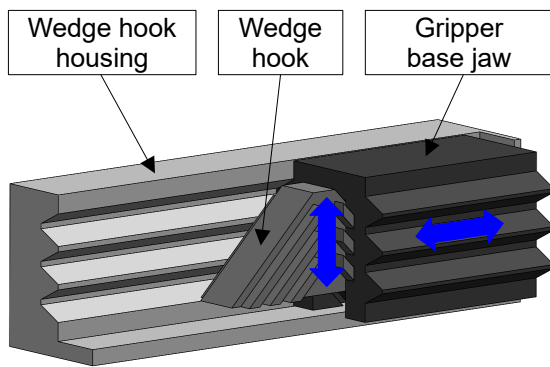


Fig. 2: Schematic of the applied wedge-hook mechanism and its mechanical principle.

The basic movement principle of the gripper jaws is based on a wedge-hook mechanism, as illustrated in Fig. 2. The rotational motion of a motor is transmitted via a worm gear and a spindle drive to a vertically moving wedge hook. This results in translational motion of the two symmetrically arranged gripper jaws. The implemented wedge-hook variant is based on a design approach for pneumatic grippers, which was optimized for a compact installation space [10].

Additive Manufacturing

The mechanical components of the base structure (see Fig. 1, lower part), with the exception of the motor, worm gear, spindle, and electronic components, are designed for additive manufacturing. The components are fabricated from carbon fiber reinforced polylactic acid (PLA, Polymaker) via fused layer modeling (FLM, Prusa MK3, nozzle diameter 0.4 mm).

Integration of Force Sensing

A standard resistive aluminum-based single-point load cell with a maximum load capacity of 30 N is used as the force sensor (see Fig. 1, upper part). The signal-conditioning unit consists of a 24-bit analog-to-digital converter (HX711) with a gain of 128 and a supply voltage of 5 V. The measurement principle is based on applying the gripping force to the load cell, as illustrated in Fig. 1. The four strain gauges mounted on the load cell are arranged in a Wheatstone full-bridge configuration. Their resistance changes according to the applied strain. The resulting differential bridge voltage is measured, amplified, and converted into a digital signal. The gripping force is then computed by a microcontroller (Arduino Nano ESP32).

Control Concept and Electronics

The microcontroller is utilized for both force signal processing and motor control. The force-control strategy is based on a combination of PID control of the PWM signal, a time-limited pulse actuation scheme, and a hysteresis band around

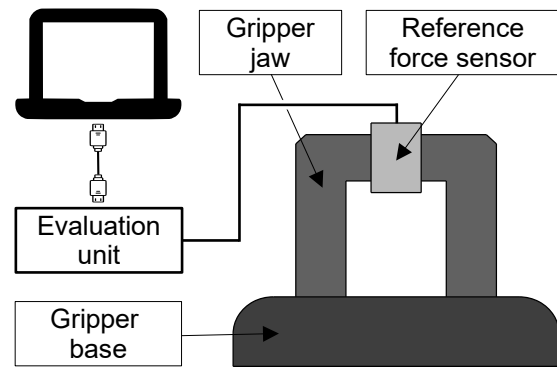


Fig. 3: Schematic of the experimental setup for evaluating the grippers' force control performance.

the setpoint. The PID controller determines the required control effort. The pulse-based actuation reduces continuous corrective actions and prevents overshooting of the system response. The hysteresis band avoids unnecessary control activity for small force deviations around the setpoint.

Experimental Setup

The gripper's performance is evaluated based on the force profiles obtained at nominal values of 5 N and 15 N, applying a reference force sensor (Type 8431-5100, Burster). For each nominal value, 10 measurements are performed to ensure statistical significance. As a benchmark against the state of the art, the gripping force profiles are compared with those of a conventional force-controlled gripper (CRG 200-85, Weiss Robotics). The reference system is evaluated at two distinct force levels representing the lower and upper regions of its operating range, specifically at 80 N and 180 N. The experimental setup is shown in Fig. 3.

Results and Discussion

Fig. 4 and 5 show the gripping force profiles for the applied nominal force values. Since only the force acting on a single gripper jaw is considered, the expected force values are halved (from 15 N to 7.5 N and from 5 N to 2.5 N). The corresponding results for both the proposed and reference gripper are summarized in Tab. 1.

The analysis of the gripping force profiles of the proposed gripper demonstrates that increasing the nominal value improves performance in terms of accuracy, precision, and overshoot. The most significant improvement is observed in accuracy, while precision and overshoot show only minor variations. The results confirm consistent behavior across repeated cycles. The observed overshoot indicates stable and well-damped control behavior. For a visual comparison of the two nominal force levels (5 N and 15 N), Fig. 6 illustrates the gripping of a kitchen sponge under the two different force settings.

Tab. 1: Evaluation results of gripping force measurements for the proposed gripper (nominal values 2.5 N and 7.5 N) and the conventional reference gripper (40 N and 90 N), considering the force on a single gripper jaw.

	Proposed force-control gripper		Conventional force-control gripper	
Nominal value (N)	2.5	7.5	40.0	90.0
Mean accuracy (% of nominal)	5.96	-1.41	6.70	-1.73
Precision (% RSD)	0.82	0.59	0.76	0.42
Mean overshoot (%)	2.09	1.24	4.23	2.03

Accuracy and precision evaluated in steady-state, overshoot calculated relative to the steady-state value.

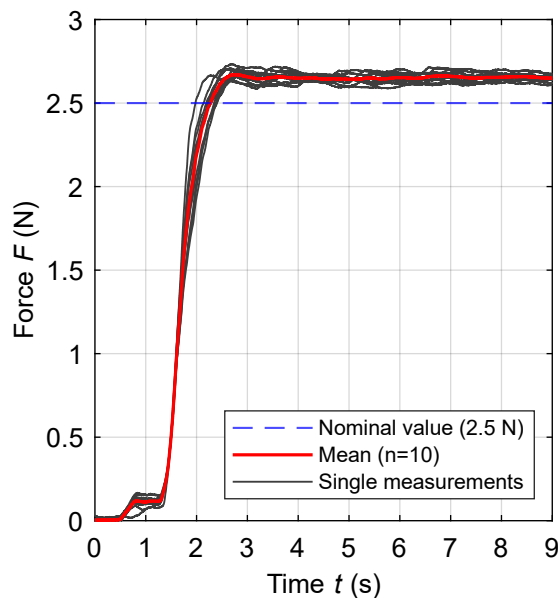


Fig. 4: Gripping force profile for a nominal force F of 5 N (single gripper jaw: 2.5 N).

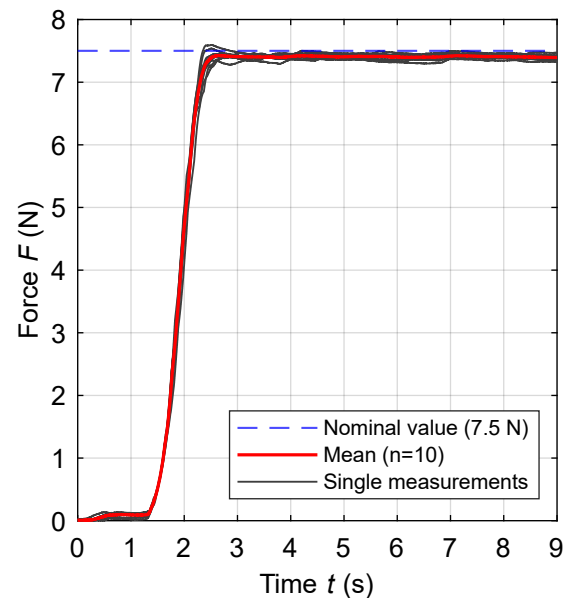


Fig. 5: Gripping force profile for a nominal force F of 15 N (single gripper jaw: 7.5 N).

A comparison between both systems reveals similar qualitative force profile characteristics. The additive manufactured gripper provides higher accuracy due to direct force measurement, particularly at low force levels. This improvement further suggests that the integrated force sensing and appropriate control strategies can compensate for manufacturing-induced tolerances inherent to additive manufacturing.

Conclusion and Outlook

In this contribution, we have proposed and experimentally investigated a novel additive manufactured sensitive robotic gripper based on a wedge-hook mechanism. The results demonstrate that the gripper reproducibly achieves defined gripping forces in the range of 5 N to 15 N. The comparison with a conventional ref-

erence gripper shows that the proposed prototype exhibits good sensitivity, particularly in the low-force range, with a maximum deviation of 5.96% of nominal $\pm$ 0.82% RSD. In addition, the results highlight the potential of integrated force sensing and control for improving robustness in additive manufactured grippers. This enables the realization of functional gripping systems without the need for complex conventional precision manufacturing. In summary, the gripper demonstrates the potential of additive manufacturing for functionally integrated robotic end-effectors.

In future work, the conventional force sensor will be replaced by a fully additive manufactured sensitive gripper jaw. This will enable a fully integrated system with individually adaptable parameters, including customized force ranges and gripper jaw geometries tailored to specific work-

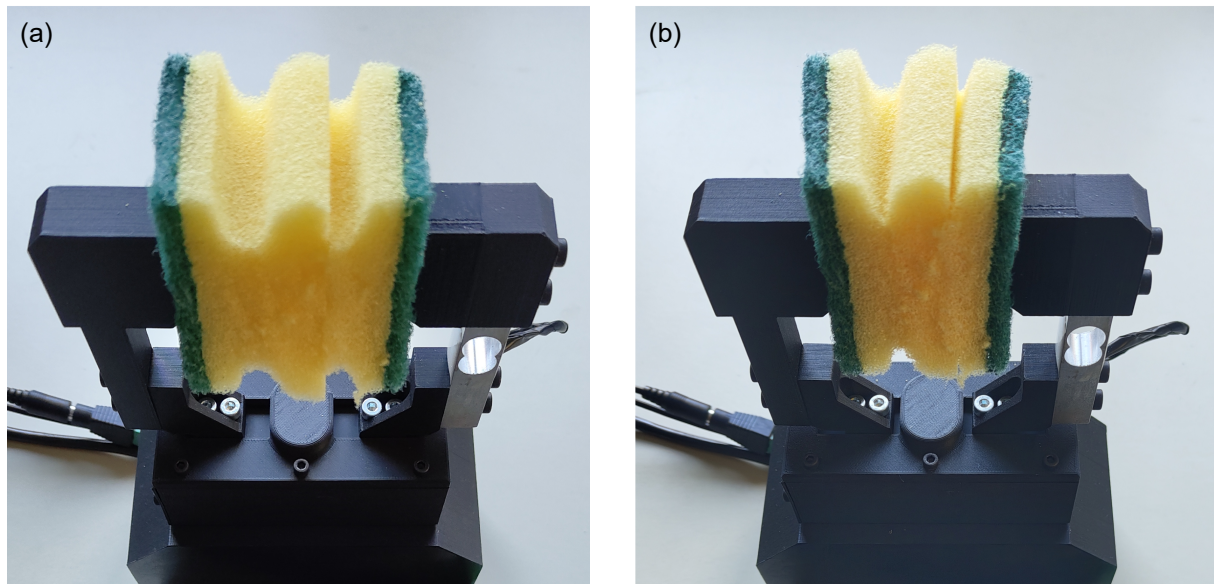


Fig. 6: Demonstration of grasping a kitchen sponge at nominal force levels of (a) 5 N and (b) 15 N.

pieces. The integration of sensing and structure within a single additive manufactured part also opens up the possibility of reducing the overall assembly complexity of the force sensing system.

References

- [1] D. Espalin, D. W. Muse, E. MacDonald, and R. B. Wicker, "3D printing multifunctionality: structures with electronics," *Int J Adv Manuf Technol*, vol. 72, 5-8, pp. 963–978, 2014, doi: 10.1007/s00170-014-5717-7.
- [2] Y. Song, R. A. Boekraad, L. Roussos, A. Kooijman, C. C. L. Wang, and J. M. P. Geraedts, "3D printed electronics: opportunities and challenges from case studies," in *Volume 1: 37th Computers and Information in Engineering Conference*, Cleveland, Ohio, USA, 2017.
- [3] M. Nagel, F. Giese, and R. Becker, "Flexible gripper design through additive manufacturing," in *Robotic Fabrication in Architecture, Art and Design 2016*, D. Reinhardt, R. Saunders, and J. Burry, Eds., Cham: Springer International Publishing, 2016, pp. 455–459.
- [4] C.-H. Liu, M.-C. Hsu, Y. Chen, W.-T. Chen, and T.-L. Chen, "A topology-optimized 3D printed compliant finger with flex sensor for adaptive grasping of unknown objects," in *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, Hong Kong, China, 72019, pp. 92–97.
- [5] N. Hangst, S. Junk, and T. Wendt, "Design of an additively manufactured customized gripper system for human robot collaboration," in *Industrializing additive manufacturing: Proceedings of AMPA2020*, M. Meboldt and C. Klahn, Eds., Cham, Switzerland: Springer, 2021, pp. 415–425.
- [6] G. L. Goh et al., "A 3D printing-enabled artificially innervated smart soft gripper with variable joint stiffness," *Adv Materials Technologies*, vol. 8, no. 24, 2023, doi: 10.1002/admt.202301426.
- [7] P. Falkowski, B. Wittels, Z. Pilat, and M. Smater, "Capabilities of the additive manufacturing in rapid prototyping of the grippers' precision jaws," in vol. 920, *Automation 2019: Progress in Automation, Robotics and Measurement Techniques*, R. Szweczyk, C. Zieliński, and M. Kaliczyńska, Eds., Cham: Springer International Publishing, 2020, pp. 379–387.
- [8] S. Hesse, *Greifertechnik: Effektoren für Roboter und Automaten*. München: Hanser, 2011.
- [9] A. Wolf and H. Schunk, *Greifer beim Robotereinsatz: Faszination der Automatisierung von Handhabungsaufgaben*, 3rd ed. München: Hanser, 2024.
- [10] Schunk GmbH & Co. KG, "Parallel gripper," *German Patent DE 100 13 022 C2*, Feb. 21, 2002.

Acknowledgement

The authors would like to thank the BurgerGroup for the collaboration and support within the joint research project.

Funding

This work was funded by the German Federal Ministry of Education and Research (BMBF) under grant number 03FHP127I. The responsibility for the content of this publication lies with the authors.

Robot-assisted test stand for calibrating 3-D-printed torque sensors

Lukas Stiglmeier¹, Thomas M. Wendt¹, and Stefan J. Rupitsch²,

¹ Hochschule Offenburg, Fakultät Wirtschaft, Work-Life Robotics Institute, Halle 101, Max-Planck-Straße 1, 77656 Offenburg, Germany

² Albert-Ludwigs-Universität Freiburg, Technische Fakultät, Institut für Mikrosystemtechnik Büro: Geb. 106, 4. OG, Raum 015, Georges-Köhler-Allee 106, 79110 Freiburg, Germany
lukas.stiglmeier@hs-offenburg.de

Summary: The metrological characterization of 3-D-printed torque sensors currently lacks standardized, automated procedures, which limits the comparability of research results. This contribution introduces a robot-assisted test stand designed for the automated and reproducible calibration of such sensors. The system employs a vertical articulated-arm robot to handle standard weights, effectively eliminating human-induced variability. In this setup, a reference torque sensor serves as the primary traceability standard, whereas the robot-controlled lever-mass system ensures a stable load application. A mathematical model based on the GUM (Guide to the Expression of Uncertainty in Measurement) is implemented to quantify the combined measurement uncertainty. Our analysis shows that the robot's positioning repeatability of 20 µm provides the necessary loading conditions for the reference instrumentation, thereby minimising parasitic forces. Consequently, the overall accuracy is primarily defined by the reference sensor itself. The presented setup enables an automated and traceable characterization of 3-D-printed torque sensors under reproducible conditions.

Keywords: test stand, torque sensor, 3-D printing, additive manufacturing, robot, calibration

Introduction

Measuring torque is essential in technical fields such as drive technology and research [1]. Several sensing principles can be employed for measuring torque, e.g. resistive, capacitive or inductive effects [1]. Currently, strain gauges are frequently employed in torque sensors [1]. However conventional force-torque sensors are typically expensive [2].

3-D printing offers new possibilities for the cost-effective design of application-specific sensors [3]. Various sensing principles can be exploited by creating functional structures using polymers such as PLA and conductive inks. [3]. Numerous publications describe 3-D-printed sensors in which the housings or force-bearing elements are 3-D-printed and then combined with conventional strain gauges or optical detectors as sensor elements [3]. Analysis of the calibration of these sensors reveals that they are often calibrated applying heterogeneous measurement setups involving test weights or reference sensors. Standardised procedures and complete documentation of sensor parameters are lacking in the current state of 3-D-printed torque sensors [3–7]. This hinders the comparability of research results.

This contribution addresses this gap by introducing a robot-assisted test stand for the automated calibration of 3-D-printed torque sensors. Equipped with a robotic arm, the system can apply defined torques and complex movement profiles with a high degree of repeatability. This minimises the impact of manual alignment errors during testing. The test stand allows 3-D-printed prototypes to be characterised under reproducible, automated conditions.

State of the art

Test stands are required for the metrological characterisation of torque sensors in order to determine measurement uncertainties, and ensure traceability to national standards [8]. The current state of the art is examined from two perspectives. Firstly, the measurement setups applied in the development of 3-D-printed torque sensors for calibration are considered. Secondly, established conventional and primary calibration test stands are examined.

Calibration setups for 3-D-printed torque sensors: Palli and Pirozzi [7] developed an optical torque sensor based on a flexible acrylonitrile butadiene styrene (ABS) frame, whose deformation is detected via the positional change of an LED relative to a photodetector. Calibration is performed using cable-guided linear motors with force transducers as a reference.

In [4], Aguilar-Segovia et al. present a 3-D-printed capacitive torque sensor. They employ a test stand with a BLDC motor as an actuator to compensate for the hysteresis and viscosity of thermoplastic polyurethane (TPU) structures employing real-time estimation algorithms.

Lee et al. [5] present a capacitive sensor structures based on printed circuit boards, which are calibrated in series with a reference sensor (ATI Nano 25).

Fu et al. [6] employ a system consisting of rotating discs and weights to characterise the static and dynamic properties of sensitive multi-axis sensors made of polyetheretherketone (PEEK). This system enables cross-axis crosstalk to be analysed.

Conventional and primary torque standards:

Conventional torque measurement technology relies either on motor-driven automated test stands or on primary standards that operate according to the dead-weight principle (lever-mass system) [9, 10]. The test stand, presented in [11], at the Physikalisch-Technische Bundesanstalt (PTB) serves as a reference system in the range of small torques up to 1 N·m [11]. The design of this primary standard is based on a horizontal measurement axis with an efficient lever-mass system [11]. The bearing is aerostatic via an air bearing, which reduces residual friction torque to less than 1 µN·m. The system features a weight-changing system that places masses into automated load cages to automate the calibration process. [11]. No vertical articulated-arm robot is used in this process. The counter drive is axially adjustable and protected against overload by a magnetic slip clutch [11]. In order to achieve the target expansion of the relative measurement uncertainty to 0.01%, the lever position is optically monitored and kept in continuous equilibrium via a counter-drive. [11].

A review of the literature on 3-D-printed torque sensors and conventional torque test stands revealed no test stand equipped with a vertical articulated-arm robot for applying test weights to a lever-mass system. This contribution addresses this gap with the aim of eliminating human-made influences on the measurement process through the employment of a robot arm.

Design of the test stand

The test stand comprises several components. Fig. 1 shows the CAD model of the setup. The base is a granite measuring plate. An aluminium slotted plate is mounted on top of this, to which the remaining components are attached. A distinctive feature of the test stand is its deployment of a vertical articulated-arm robot to place standard weights on a defined lever arm during the calibration process. Automating this process with the robotic arm ensures that the standard weights are handled in a reproducible manner without human intervention. This prevents potential errors caused by the operator. The TD110a 10Nm/VA reference sensor (ME-Messsysteme) is employed for measuring the torques that act during the calibration process. The measurement signals from the TD110a 10Nm/VA are processed by the GSV-2MSD-DI measurement amplifier, which is also supplied by ME-Messsysteme. This contribution presents an example of calibrating an inductive, 3-D-printed torque sensor. An LDC1614EVM evaluation board from Texas Instruments is used to process the measurement signals from the sensor. The test stand enables to apply both clockwise and counterclockwise torques to the calibrated sensor. Additionally, the test stand incorporates a temperature sensor via a sepa-

rate microcontroller to monitor temperature during calibration.

System control and software architecture

The test stand is centrally controlled through a PC workstation equipped with ethernet and USB-A interfaces. Communication with the industrial robot is established via an Ethernet/IP connection. The measurement amplifier (GSV-2MSD-DI) and microcontrollers, which acquire temperature data and sensor signals from the test specimen, are connected to the PC via a USB hub. A master software program written in Python serves as the central control unit. This program consists of several subroutines that are sequentially invoked during the calibration process. At the start of a calibration, the user configures the necessary parameters, such as the direction of torque rotation, via a graphical user interface (GUI). The TCP/IP connections to the robot and the measurement amplifier are then initialised. After manually activating the robot's power supply via the GUI, the automated process begins. Hereby, the software for acquiring data from the LDC1614EVM evaluation board is first launched automatically. This is done applying a 'BotModule', which uses image recognition to execute a predefined sequence of interactions to control the manufacturer-specific acquisition software and define the storage path. Next, the robot performs the mechanical calibration cycle, during which standard weights are placed on the lever arm in stages and then removed again. Once the motion sequence is complete, the robot sends a status signal to the master software. Data acquisition then terminates, the measured values are exported in CSV format and the main program closes. The CSV files can then be processed as needed. The system's modular design allows different robot programs to be loaded for specific calibration procedures. In addition to the automated handling of standard weights, this enables endurance tests to be performed, as well as specific measurement tasks for various sensor configurations.

Metrological characterization

The metrological characterisation of the test stand is based on [11]. In this article, Röske present a torque test stand with a lever-mass measurement system. They apply the guide to the expression of uncertainty in measurement (GUM) to evaluate the contributions to measurement uncertainty [11], as is also done in this contribution.

Torque generation: The torque is generated within the test stand based on the definition as the product of the lever arm and the force [11]. Mathematically, this can be expressed as the cross product of the position vector $\vec{r}$, which points from the reference point on the torque axis to the point of force application, and the

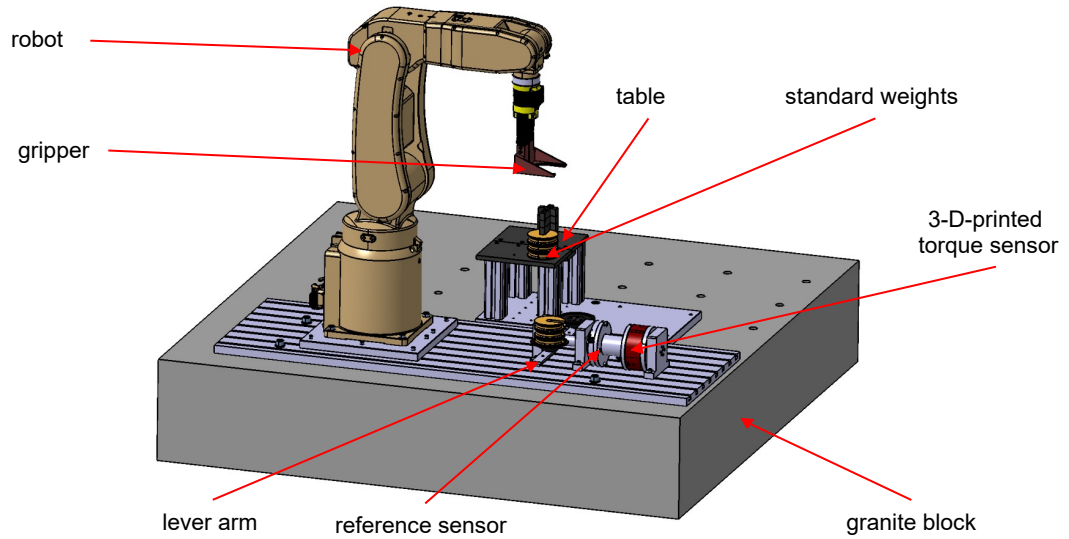


Fig. 1: CAD model of the robot-assisted torque test stand.

force vector $\vec{F}$ [11]:

$$\vec{M} = \vec{r} \times \vec{F}. \quad (1)$$

The sensor's calibration is supported by a lever-mass system that is suspended in several ball bearings. This prevents the applied load from generating an unwanted shear force or bending moment on the axis of rotation. The bearings balance the system's weight and provide the necessary counterforce. This enables the torque to be converted into a force pair and transmitted to the sensor.

Mathematical model: The test stand presented by Röske [11] is a classic lever-type measurement system. The torque applied is primarily determined by the configuration of the lever arm and the test weights. The lever arm and test weights are both traceable quantities and are considered the primary sources of torque and their value for the calibration. Röske [11] describes the mathematical model of his test stand and the generated torque M as follows [11]:

$$M = m \cdot g \cdot l \cdot \left(1 - \frac{\rho_L}{\rho_m}\right) \cdot \cos(\alpha) + \sum_{i=1}^n \Delta M_i \quad (2)$$

In this equation, m represents the total mass of the standard weights and g represents the local gravitational acceleration at the system's location, and l represents the length of the lever arm. The term $(1 - \rho_L/\rho_m)$ corrects for the physically induced air buoyancy, where ρ_L denotes the air density and ρ_m denotes the material density of the masses applied. The angle of inclination of the lever arm relative to the horizontal is denoted by α . Additionally, the term $\sum \Delta M_i$ accounts for up to n different

factors that can influence the total torque [11].

In contrast to the classical lever-mass system described in [11], in which the torque is defined by mechanical parameters, only the reference sensor (TD110a 10Nm/VA) serves as the feedback variable in the robot-assisted test stand. During calibration, the values indicated by the reference sensor are employed. The masses and the lever arm then merely generate the necessary load. The tolerances of the lever arm, reference weights and angle of inclination are no longer directly incorporated into the reference torque calculation. The actual torque is read directly from the reference sensor. As a result, the mathematical model can be significantly simplified:

$$M = M_{ref} + \sum_{i=1}^n \Delta M_i \quad (3)$$

Here, M_{ref} is the value indicated by the reference sensor, and the term $\sum \Delta M_i$ accounts for up to n different factors that can influence the total torque. These include, among other things, the uncertainty of the measurement amplifier and the repeatability of the robot.

Uncertainty Sources

The following section examines the key factors that influence measurement uncertainty of the robot-assisted test stand. Fig. 2 provides an overview of these factors in the form of an ishikawa diagram.

Geometrical characteristics: Alignment errors resulting from the manufacturing tolerances of the individual components, as well as deviations that occur during the assembly of the test stand,

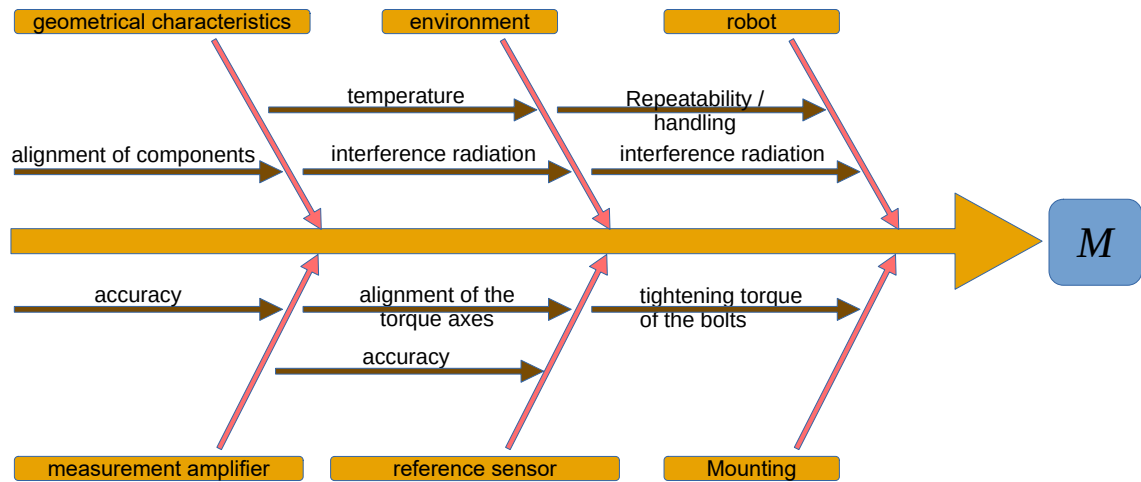


Fig. 2: Ishikawa diagram showing the acting influence factors for the torque.

are referred to here. For instance, the rotational axes are unlikely to be perfectly aligned, resulting in slight discrepancies in the actual torque acting on the sensor. Alignment errors in the rotational axes can also cause parasitic forces and moments.

Environment: The application of physical torque on the test stand is affected by environmental conditions. The primary factor to consider here is temperature. Fluctuations in temperature can cause the metal components of the test stand to expand thermally. Furthermore, temperature affects both the reference sensor and the sensor being calibrated. To minimise measurement uncertainty, ambient temperatures are recorded and incorporated into the mathematical model used to determine the torque.

In addition to temperature, electromagnetic interference fields can affect the measurement. This applies to both the reference sensor and the sensor being calibrated.

Robot: The employment of an industrial robot to automate the calibration process reduces human intervention, allowing standard weights to be placed on the lever arm in a reproducible manner. The robot's repeatability (20 μm [12]) is crucial for positioning the standard weights on the lever arm, since deviations in position lead to variations in the effective length of the lever arm. Additionally, electromagnetic interference from the robot's drives can affect the measurement signal chain. In order to eliminate dynamic effects such as vibrations or acceleration forces, measurements are only taken after a stable static state has been achieved following the placement process.

Reference sensor: The actual torque is measured employing a TD110a 10Nm/VA

torque sensor from ME-Messsysteme as the reference sensor [13]. The accuracy of this sensor is a key factor in the uncertainty assessment of the test stand. In addition, the uncertainty of the measurement amplifier must also be considered.

Mounting: The tightening torque during the installation of the test specimen is a critical factor. This can subject the calibrated sensor to undesirable mechanical stress. Insufficient tightening torque reduces contact stiffness at the connection point. This can lead to micro-movements and nonlinearities in the force flow, resulting in increased hysteresis and reduced repeatability. Conversely, excessive or asymmetrical tightening torque can create parasitic stress states within the sensor structure. These stresses can shift the zero point and impair the sensor's sensitivity by altering the structural stiffness. Furthermore, uneven torque distribution often results in parasitic load components, such as shear forces or bending moments, that disrupt ideal force transmission.

Uncertainty budgets

In order to evaluate the accuracy of the torque measurements, the combined measurement uncertainty is determined in compliance with the GUM. In this setup, the reference torque sensor (TD110a 10Nm/VA) is defined as the primary reference. Therefore, the mathematical model for the reference torque simplifies to (3). According to the law of uncertainty propagation, the combined standard uncertainty $u_c(M)$ is calculated from the standard uncertainties of the individual input quantities $u(x_i)$ and their respective sensitivity coefficients c_i [11, 14]:

$$c_i = \partial M / \partial x_i \quad (4)$$

$$u_c(M) = \sqrt{\sum_{i=1}^n \left(\frac{\partial M}{\partial x_i} \cdot u(x_i) \right)^2} \quad (5)$$

As the uncertainty according to equation (3) is primarily determined by the reference sensor, the contributions of the sensor-related uncertainties are included in the measurement uncertainty budget with equal weighting.

Uncertainty components: The following uncertainty components are considered based on the calibration approach:

- **Reference sensor (u_{ref}):** The uncertainty of the TD110a 10Nm/VA sensor (Class 0.1), including hysteresis and non-linearity. This is the primary source of uncertainty.
- **Measurement amplifier (u_{amp}):** The uncertainty of the measurement amplifier GSV-2.
- **Zero point drift TD110 10a (u_{drift_s}):** Potential drift of the sensor signal before or during the measurement cycle.
- **Zero point drift GSV-2 (u_{drift_a}):** Potential drift of the signal before or during the measurement cycle.
- **Repeatability (u_{rep}):** This component accounts for the variability in measurements, indirectly capturing mechanical influences such as the robot's positioning repeatability, lever arm variations, and angular alignment.

Sensitivity coefficients: All components of measurement uncertainty are equally weighted and are included in the measurement uncertainty budget with a sensitivity coefficient of $c_i = 1$. Other mechanical factors (mass, lever arm, inclination) are no longer treated as direct additive terms in the torque equation. Table 1 summarizes the estimated values and distributions for a nominal torque of 10 N·m. The reason for this is that the accuracy of the reference sensor is specified relative to 10 N·m. The expanded uncertainty is calculated as

$$U = k \cdot u_c(M) \quad (6)$$

with a coverage factor of $k = 2$, providing a confidence level of approximately 95%.

Results

Analysis of the Budget: The metrological analysis reveals that the reference sensor (TD110a 10Nm/VA) is the dominant factor, accounting for the largest share of the uncertainty budget. The systematic errors of the mechanical lever-mass system (e.g. precise lever length or local gravity) do not directly increase the measurement uncertainty of the reference value when the sensor is employed as the reference.

Instead, the industrial robot's precision of 20 μ m ensures reproducible loading conditions. This reduces the "Repeatability" component of the sensor signal to a minimum. The integration

of the robotic system effectively eliminates manual alignment errors and parasitic forces, ensuring that the reference sensor operates under sufficient, axial loading conditions. This result underscores the suitability of the robotic system for automated calibration, as it ensures that the uncertainty is limited by the reference instrumentation rather than the handling process.

Realized test stand: Fig. 3 shows the assembled test stand. The T-slot plate was manufactured by an external contractor. All other metal components, such as bearings, adapters and shafts, were made in the university's metal workshop. During assembly, care was taken to ensure the robot was properly grounded.

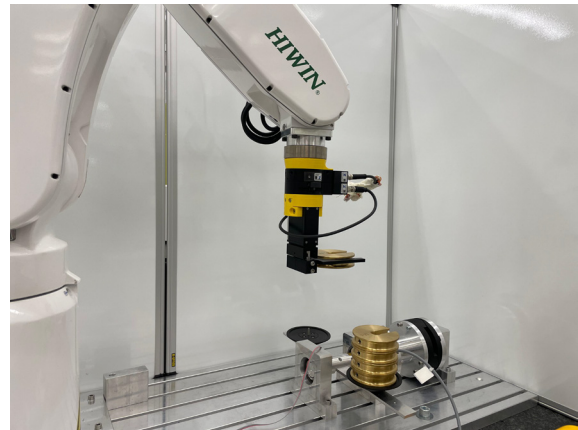


Fig. 3: Photo of the realized test stand.

Measurement: A 3-D-printed inductive torque sensor is used to demonstrate the functionality of the test stand described here. Fig. 4 shows the signals from the 3-D-printed sensor and the reference sensor under stepwise loading. It can be seen that the load can be applied reproducibly employing the test stand. Based on the measured values, the typical parameters of the sensor to be calibrated can be determined.

The next step is to evaluate the test stands' ability to apply consistent and repeatable load levels to the sensor being calibrated. This is achieved by analysing the measured values from the reference sensor across multiple calibration runs and calculating the standard deviation. The load steps are determined automatically and the average value for each level is calculated. This value is then employed as a representative figure for each load step. Clockwise and counterclockwise torque are analysed separately. A total of 24 measurement runs are taken for the test. The results are shown in Table 2. The standard deviations of the load levels are of similar magnitude. The largest standard deviation was observed at level $M_{step} = 2.5$ N·m and

Tab. 1: Measurement uncertainty budget at 10 N·m.

Component x_i	Value x_i	$u(x_i)$	Distribution	Contribution $u_i(M)$
Reference Sensor (TD110a 10Nm/VA)	10 N·m	0.01 N·m	Rectangular	0.00577 N·m
Amplifier Accuracy (GSV-2)	10 N·m	0.005 N·m	Rectangular	0.00289 N·m
Zero Point Drift (Sensor)	-	0.0005 N·m	Rectangular	0.000289 N·m
Zero Point Drift (GSV-2)	-	0.001 N·m	Rectangular	0.000577 N·m
Repeatability (Robot-assisted)	10 N·m	0.002 N·m	Rectangular	0.001155 N·m
Combined $u_c(M)$				0.0066 N·m
Expanded $U(k=2)$				0.0132 N·m

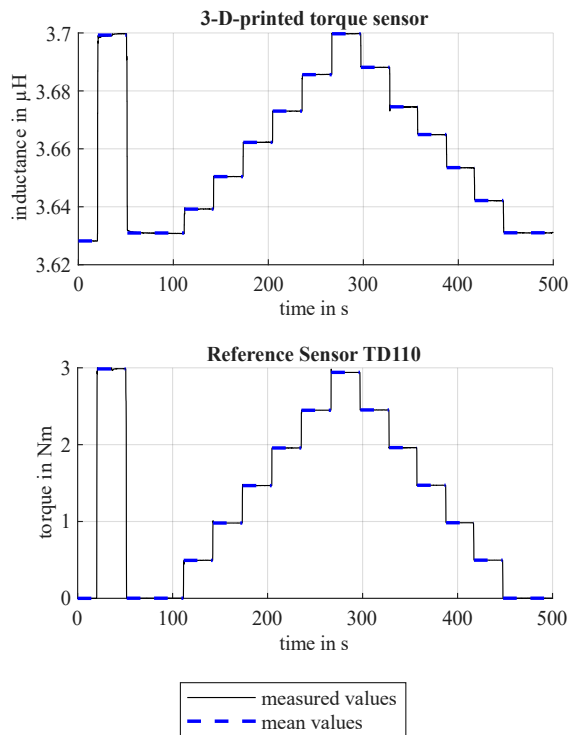


Fig. 4: Measurement values from the reference sensor and the sensor to be calibrated during stepwise loading.

$M_{step} = 2.0 \text{ N·m}$ (clockwise rotation). Overall, the results show that the robot system can be employed to apply load levels reproducibly and perform calibration.

Conclusion

This contribution presents a robot-assisted test stand designed for the automated calibration of 3-D-printed torque sensors. The vertical articulated-arm robot can handle standard weights, effectively eliminating human-induced variability and manual alignment errors. Metrological characterization in accordance with GUM guidelines confirms the system's performance. Defining the reference torque sensor as

Tab. 2: Repeatability of measured values for clockwise and counterclockwise ($n=24$ cycles).

Torque step M_{step} (N·m)	Clockwise		Counterclockwise	
	M_{avg} (N·m)	σ (mN·m)	M_{avg} (N·m)	σ (mN·m)
0.0	-0.01	2.9	0.0	3.0
0.5	-0.51	2.5	0.50	3.2
1.0	-1.03	2.5	0.98	3.3
1.5	-1.54	2.6	1.47	3.3
2.0	-2.06	2.6	1.96	3.4
2.5	-2.58	2.6	2.45	3.5
3.0	-3.10	2.7	2.94	3.5
2.5	-2.58	4.2	2.45	3.2
2.0	-2.06	4.2	1.96	3.2
1.5	-1.54	4.1	1.47	3.2
1.0	-1.03	3.8	0.99	3.4
0.5	-0.52	3.5	0.50	3.3
0.0	-0.01	4.0	0.0	3.3

the primary traceability standard means that the measurement uncertainty is no longer directly influenced by the mechanical tolerances of the lever-mass system. Instead, the robot's positioning repeatability of $20 \mu\text{m}$ ensures that the reference sensor operates under reproducible and sufficient loading conditions, minimizing parasitic forces and off-axis effects. Consequently, the absolute accuracy of the test stand is primarily defined by the accuracy class of the reference sensor (TD110a 10Nm/VA), whereas the robotic handling ensures that this accuracy is reached consistently across all measurement cycles.

Acknowledgment

One student made a significant contribution through his work, although he was not directly involved in writing the article: Andreas Schäuble designed and tested the software architecture for controlling the test stand as part of his bachelor's thesis at Offenburg University of Applied Sciences.

Language polishing and grammatical enhancements were performed using Gemini and DeepL.

References

- [1] Rainer Schicker and Georg Wegener. *Drehmoment richtig messen*. Hottinger-Baldwin-Messtechnik, Darmstadt, 2002. ISBN: [3000090150](#).
- [2] Max Yiye Cao, Stephen Laws, and Ferdinando Rodriguez y. Baena. Six-Axis Force/Torque Sensors for Robotics Applications: A Review. *IEEE Sensors Journal*, 21(24):27238–27251, 2021. doi: [10.1109/JSEN.2021.3123638](#).
- [3] Lukas Stiglmeier, Thomas M. Wendt, and Stefan J. Rupitsch. 3-D-Printed Torque Sensors: A Review. *IEEE Sensors Journal*, 24(12):18740–18761, 2024. doi: [10.1109/JSEN.2024.3397919](#).
- [4] Jose Eduardo Aguilar-Segovia, Maxime Manzano, Sylvain Guégan, Ronan Le Breton, Alice Farhi-Rivasseau, Sylvain Lefebvre, and Marie Babel. Multi-Material Torque Sensor Embedding One-Shot 3D-Printed Deformable Capacitive Structures. *IEEE Sensors Letters*, 8(9):1–4, 2024. doi: [10.1109/LESENS.2024.3440195](#).
- [5] Dong-Hyuk Lee, Uikyum Kim, Hosang Jung, and Hyouk Ryeol Choi. A Capacitive-Type Novel Six-Axis Force/Torque Sensor for Robotic Applications. *IEEE Sensors Journal*, 16(8):2290–2299, 2016. doi: [10.1109/JSEN.2015.2504267](#).
- [6] Liyue Fu, Aiguo Song, and Dapeng Chen. A Polyetheretherketone Six-Axis Force/Torque Sensor. *IEEE Access*, 7:105391–105401, 2019. doi: [10.1109/ACCESS.2019.2932387](#).
- [7] G. Palli and Salvatore Pirozzi. An Optical Torque Sensor for Robotic Applications. *International Journal of Optomechatronics*, 7(4):263–282, 2013. doi: [10.1080/15599612.2013.879500](#).
- [8] Diedert Peschel and Dirk Röske. Messunsicherheiten bei der Drehmomentmessung. *tm - Technisches Messen*, 72(5):308–317, 2005. doi: [10.1524/teme.72.5.308.64796](#).
- [9] Hans Jörg Fraiss, Lioba Stenner, and Dirk Röske. Development of a New 400 kN m Torque Calibration Machine. pages 1–4, August 30–September 4 2015. Accessed via IMEKO Repository.
- [10] Mikołaj Woźniak and Dirk Röske. Investigation of the Calibration and Measurement Capabilities of the New 5 kN m Torque Calibration Machine at GUM. pages 1–4, August 30–September 4 2015. Accessed via IMEKO Repository.
- [11] Dirk Röske. Metrological Characterization of a 1 Nm Torque Standard Machine at PTB, Germany. *Metrologia*, 51(1):87–96, 2014. doi: [10.1088/0026-1394/51/1/87](#).
- [12] Hiwin GmbH. Montageanleitung RA605-710-SL-A, 08 2001. Dokumentationsnummer: RA605-01/2-DE-2001-MA.
- [13] ME-Meßsysteme GmbH. *Datenblatt TD110a 10Nm/VA*. ME-Meßsysteme GmbH, Hennigsdorf, Deutschland, 05 2026. Dokumentationsnummer: 6480.
- [14] Joint Committee for Guides in Metrology (JCGM). JCGM 100:2008. Evaluation of measurement data – Guide to the expression of uncertainty in measurement. Guide, BIPM, 2008. GUM 1995 with minor corrections; corrected version 2010.

Shared Control in Assistive Robotics: Characteristics Influencing User Acceptance

Anke Fischer-Janzen¹, Thomas M. Wendt¹

*Work-Life Robotics Institute, Offenburg University of Applied Science
Max-Planck-Str. 1, 77656 Offenburg*

Abstract

Robots used in care and daily life offer a new approach to addressing labor shortages in nursing professions. In this context, various types of robots are already available, or are being developed, to address different ongoing challenges. In this context, shared controls can be realized with different robotic systems and control strategies. This work provides an overview of the characteristics that influence current issues and the ways in which implementing these characteristics into robotic systems can be beneficial.

Keywords: Robotics, Health Care, Shared control, human-robot interaction, robot acceptance.

Introduction

The shortage of skilled workers in the nursing sector will significantly worsen in the coming decades. Forecasts show that by 2049 there could be a shortage of between 280,000 and 690,000 nursing staff [1]. This growing shortage highlights the urgent need for innovative solutions to maintain the quality and continuity of care. In this context, robotic systems and assistive technologies are gaining increasing importance, as they can support caregivers in physically demanding or repetitive tasks, enhance efficiency, and allow staff to focus more on direct human interaction [2].

With the rapid advancement of AI, new approaches to designing robotic controls are emerging. This work presents characteristics influencing robot acceptance in relation to shared controls. We discuss challenges by providing an overview of ongoing developments in the field.

Examples for varying robot application

The diversity of robotic systems reflects the wide variety of tasks and environments in which they can support humans. In this work, we provide characteristics describing articulated robots for people with physical disabilities, logistic robots in healthcare, and social robots. These robots differ greatly in their capabilities and reduce the workload of care staff in different ways.

Articulated robots, also known as wheelchair-mounted robotic arms or assistive robotic arms, directly assist users by enabling them to

perform everyday tasks. This capability provides users with greater independence in daily life, leading to a higher quality of life [3]. In contrast, logistic robots such as mobile robots are used in hospitals to transport bedding, equipment, and medical supplies. Their main purpose is to relieve care staff so that they can spend more time with patients. Social robots, such as Nao, Pepper, and Nao, are designed to interact with people. In care facilities, they motivate elderly individuals to exercise, play games, or engage in conversations [4]. This is particularly important as the number of lonely elderly people continues to rise.

Characteristics

The characteristics that influence the acceptance of these systems vary depending on the type of robot application. In previous work [5], we investigated factors influencing a gaze-driven control for articulated robots. Our evaluation showed that robot acceptance is a complex and multifaceted topic. It is strongly influenced by the following groups of characteristics: level of autonomy, abilities, user preferences, and other contextual factors.

Level of autonomy:

In shared control, the level of autonomy defines the degree to which the robot performs tasks independently. It can range from manually controlled to fully autonomous [6]. Major challenges arise when automating robots to perform a wide range of daily tasks in cluttered, unstructured environments [7], which

must be addressed as prerequisites for the application of such robots.

Depending on the system's performance, the user's sense of agency can vary significantly. Trust in the system and a stronger sense of agency develop when the robot behaves as expected, suggesting that increased autonomy does not necessarily lead to higher user acceptance [8].

Error rates can be reduced by shifting more control to the user. Experiments have shown that even when tasks take longer, users are generally more satisfied with the results [8]. However, other studies found that high levels of automation can lead to acceptance when the robot accurately predicts, recognizes, and respects human intent [3, 9]. Furthermore, the sense of agency is influenced by transparency [10]. In highly autonomous modes, it can be improved through artificial feedback. For example, Baptista et al. integrated mixed reality into their robotic system to provide users with visual feedback indicating whether their inputs were correctly registered [11]. The generated natural feel of the control can reduce cognitive load, making the system less tedious to operate [10, 12].

Abilities: The relevant characteristics differ among robot types. While the goal of articulated robots is to perform a wide range of everyday tasks sequentially [13], logistic and social robots are usually specialized for narrower domains. Performing these tasks robustly and reliably is crucial to maintaining safety standards. In shared control systems with varying levels of autonomy, grasping and trajectory planning can differ in realization. Studies have shown that even small details such as the way an object like a book is handed over can affect user acceptance [14].

User preferences: In assistive contexts, especially for people with disabilities, multimodal control adapted to the user's abilities is highly relevant for wheelchair-mounted robotic systems. Moreover, adapting the system to individual user preferences is a key factor in robot acceptance. Parameters such as robot velocity [8] and the level of automation can be adjusted using policy-blending formalisms [15]. Other aspects such as living arrangements, building architecture, and familiarity with technology can also influence acceptance [7].

Other factors: Finally, ethical considerations and regulatory requirements must be respected. In hospitals and private homes, user privacy must be protected, particularly when handling sensitive personal information [16]. Safety must be ensured, as humans and robots will share the same workspace. ISO

15066:2016 provides guidance for human-robot shared workspaces, while ISO 13482 defines safety requirements for personal care robots. Additionally, appearance, initial and maintenance costs, and everyday maintenance all influence acceptance among stakeholders who may not interact directly with the robot.

Conclusion

Established characteristics must be evaluated alongside new ones that have emerged from ongoing developments in AI and human-robot interaction (HRI). In this work, we provided an overview of the key characteristics influencing robot acceptance that should be considered during system design to enhance marketability and user satisfaction.

References

- [1] Deutscher Pflegekongress 2024, *Zahlen und Fakten zur Pflege in Deutschland* (2024), online.
- [2] M. Kyranini et al., *Technologies*, 9(8), 1–26 (2021); doi:10.3390/technologies9010008
- [3] A. Nanavati, *Annual Review of Control, Robotics, and Autonomous Systems*, 7, 123–147 (2024); doi: 10.1146/annurev-control-062823-024352
- [4] A. Fischer-Janzen et al., *HCI in Business, Government and Organizations. HCII 2024*, 14720, doi: 10.1007/978-3-031-61315-9_20
- [5] A. Fischer-Janzen, T.M. Wendt, D. Görlich, K. Van Laerhoven, *arXiv:2505.23147 [cs.HC]*; doi: 10.48550/arXiv.2505.23147
- [6] J.M. Beer, A.D. Fisk, W.A. Rogers, *J Hum Robot Interact.* 3(2):74–99 (2014); doi: 10.5898/JHRI.3.2.Beer.
- [7] C. A. Stanger et al., *IEEE Transactions on Rehabilitation Engineering*, 2(4), 256–265 (1994); doi: 10.1109/86.340872
- [8] T. Bhattacharjee et al., *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*, 181–190 (2020); doi: 10.1145/3319502.3374818
- [9] Losey et al., *Applied Mechanics Reviews*, 70(1), 1–19 (2018); doi: 10.1115/1.4039145
- [10] C. Mazzola et al., *arXiv:2508.13976[cs.RO]* (2025); doi: 10.48550/arXiv.2508.13976
- [11] R.R. Baptista et al., *arXiv:2505.03929[cs.RO]* (2025), doi: 10.48550/arXiv.2505.03929
- [12] M.E. Loke et al., *International Journal of Conservation Science* (2023); doi: 10.36868/IJCS.2023.03.02
- [13] A. Hagengruber et al., *Scientific Reports*, 15(1) (2025); doi: 10.1038/s41598-025-89405-2.
- [14] M. Cakmak et al., *IEEE IROS 2011*, 1986–1993 (2011); doi: 10.1109/IROS.2011.6094735
- [15] A.D. Dragan, S. Srinivasa, *The International Journal of Robotics Research*, 32(7) (2013); doi: 10.1177/0278364913490324
- [16] L. Riek, D. Howard, *A Code of Ethics for the Human-Robot Interaction Profession, Proceedings of We Robot* (2014).

Additively Manufactured Calibration Target for Intensity-Based LiDAR Calibration in Industrial Robotics

Thomas M. Wendt¹, Sinan Sme¹, Andi Schnebel¹, Stefan J. Rupitsch²

¹ Offenburg University of Applied Sciences, Badstr. 24, 77656 Offenburg, Germany

² Laboratory for Electrical Instrumentation and Embedded Systems, Department of Microsystems Engineering, University of Freiburg, Georges-Khler-Allee 106, 79110 Freiburg, Germany
sinan.sueme@hs-offenburg.de

Summary: Accurate extrinsic calibration of robot-mounted LiDAR sensors requires the extraction of unambiguous geometric features from the sensor data. Intensity-based corner detection is an established strategy, but it relies on a calibration target that provides a stable, high-contrast reflectivity pattern in the near-infrared spectrum. Conventional targets based on printed or painted checkerboards are subject to surface degradation and are restricted to planar geometries. The presented contribution proposes a calibration target fabricated by multi-material additive manufacturing, combining Reflect-O-Lay (a retroreflective PLA composite) and Fiberlogy PA12+CF15 (a carbon-fibre-reinforced polyamide) to produce a five by five checkerboard with intrinsic near-infrared reflectivity contrast at 940 nanometres. Due to the contrast embedded in the material composition, the target is inherently robust against abrasion, moisture, and ultraviolet exposure, and the additive manufacturing route permits arbitrary geometric modifications without tooling cost. Corner detection accuracy is evaluated experimentally on a RoboSense E1R solid-state LiDAR sensor mounted on a KUKA LBR iisy 11 R1300 manipulator. Across ten robot configurations, the detected inner corner positions agree with the nominal grid geometry to a Procrustes RMSE of 7.6 millimetres with a systematic spacing bias below half a percent, consistent with the ranging precision of the sensor. A comparison against reference targets reported in the literature places the proposed AM target in the same accuracy regime as established solutions while offering superior material stability and design flexibility.

Keywords: LiDAR calibration, additive manufacturing, multi-material FDM, intensity-based feature extraction, industrial robotics

Introduction

Three-dimensional LiDAR sensors are increasingly employed in robotic manufacturing cells for illumination-independent spatial perception tasks such as quality inspection, bin picking, and dimensional verification. Integrating such sensors into a robotic system requires that the sensor coordinate frame be precisely related to the robot kinematic chain, a problem known as *eye-in-hand extrinsic calibration*. The unknown flange-to-sensor transformation $\mathbf{X} \in SE(3)$ satisfies the classical hand-eye equation $\mathbf{A}\mathbf{X} = \mathbf{X}\mathbf{B}$, first established by Tsai and Lenz [1] and Shiu and Ahmad [2].

Target-based calibration methods commonly extract features by fitting the checkerboard as a dominant plane in the point cloud and deriving bounding edges as pose constraints [3, 4]. Such geometric approaches are fundamentally limited by the discrete angular sampling of the LiDAR beam pattern, which introduces quantization errors at target boundaries [5]. Intensity-based corner detection overcomes this limitation by exploiting the per-point reflectivity contrast between adjacent target regions [6, 7, 8], but presupposes a target with stable, high-contrast reflectivity in the operating wavelength of the sensor. Conventional painted or printed checkerboards rely on surface coatings whose optical properties are known to degrade under abrasion, humidity, and UV exposure [9], and alternative target geometries

such as 3D towered checkerboards [10] or boards with circular cut-outs [11] require custom machining and remain restricted to a fixed geometry.

The present work proposes a multi-material AM calibration target in which the reflectivity contrast is intrinsic to the material composition. The contributions are:

- A multi-material AM target combining Reflect-O-Lay and PA12+CF15 that provides long-term reflectivity stability and geometric design freedom beyond the planar paradigm
- The characterization of the near-infrared reflectivity contrast at 940 nm obtained through the material combination
- An experimental evaluation of the corner detection accuracy achievable with this target on a solid-state LiDAR sensor, including a comparison against reference targets reported in the literature

Additively Manufactured Calibration Target Material Selection

The target exploits the opposing near-infrared reflectivity of two commercially available FDM filaments at the RoboSense E1R operating wavelength of $\lambda = 940$ nm.

High-reflectivity constituent—Reflect-O-Lay (Lay Filaments, Germany) [12] is a semi-flexible PLA-based composite loaded

with retroreflective microsphere pigments. The retroreflective mechanism returns incident laser radiation towards its source with minimal angular deviation, independent of the angle of incidence, ensuring consistent high-intensity LiDAR returns across the full $\pm 40^\circ$ tilt range. Processing parameters: extrusion temperature 210°C , nozzle diameter $\geq 0.4\text{ mm}$.

Low-reflectivity constituent—Fiberlogy PA12+CF15 (Fiberlab S.A., Poland) [13] is a polyamide-12 matrix reinforced with 15 % short carbon fibres. Carbon-fibre-filled polymers are established near-infrared absorbers at 940 nm [14], yielding consistently low LiDAR return intensities. The material additionally provides high dimensional stability and thermal resistance up to 160°C . Processing parameters: extrusion temperature $255\text{--}270^\circ\text{C}$, heated bed $90\text{--}110^\circ\text{C}$, hardened-steel nozzle.

The reflectivity contrast ΔR is thereby embedded within the material composition, eliminating degradation mechanisms associated with abrasion, moisture absorption, and UV exposure that affect surface coatings.

Target Geometry

The target comprises a 5×5 checkerboard pattern with a field size of $45\text{ mm} \times 45\text{ mm}$, yielding $4 \times 4 = 16$ inner corner points as calibration landmarks. Including the surrounding border (4.5 mm width), the total plate dimensions are $234\text{ mm} \times 234\text{ mm}$, as displayed in Fig. 1.

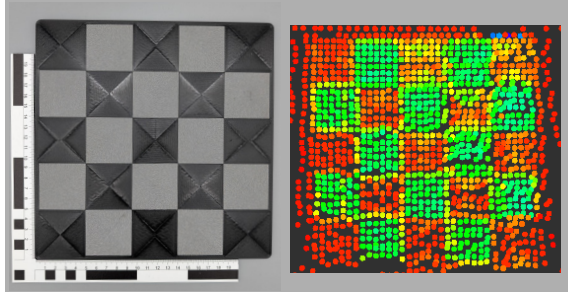


Fig. 1: AM multi-material calibration target. Light fields: Reflect-O-Lay. Dark fields: PA12+CF15 (left). E1R intensity image showing the reflectivity contrast ΔR at 940 nm (right).

At the design working distance of $0.4\text{--}1.2\text{ m}$, the E1R angular resolution of 0.625° yields approximately 4–6 beam footprints per field boundary, sufficient for stable intensity-gradient computation across each material interface. The AM fabrication approach allows field size, board dimensions, and aspect ratio to be adapted to any sensor–distance combination by modifying the CAD model, without incurring tooling cost. Non-planar surface elements or mounting features may be incorporated within the same print job.

Calibration Framework

System Configuration and Kinematic Model

The experimental platform consists of a KUKA LBR iisy 11 R1300 six-axis collaborative manipulator with a RoboSense E1R solid-state LiDAR sensor attached to the flange.

Four coordinate frames are involved (Fig. 2):

- Robot base frame $\{B\}$
- Flange frame $\{F\}$ with pose $\mathbf{T}_{F,i} \in SE(3)$ from the KUKA controller
- Sensor frame $\{S\}$ related to $\{F\}$ by the unknown offset $\mathbf{X} \in SE(3)$
- Stationary target frame $\{P\}$.

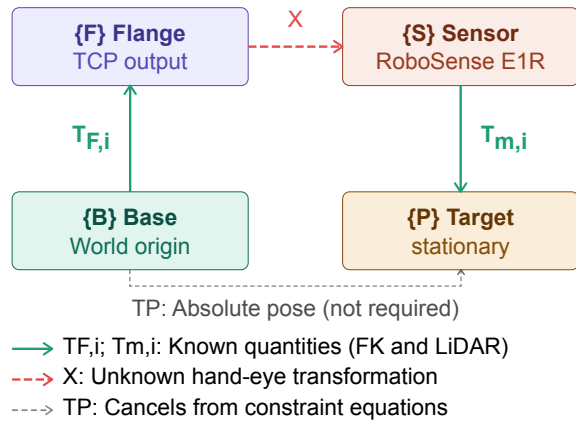


Fig. 2: Coordinate frames. Known: flange pose $\mathbf{T}_{F,i}$ (green). Unknown: hand-eye offset $\mathbf{X}$ (red). The absolute target pose $\mathbf{T}_P$ cancels from the estimation problem.

Since the target is stationary, the kinematic chain $\mathbf{T}_{F,i} \cdot \mathbf{X} \cdot \mathbf{T}_{m,i} = \mathbf{T}_P = \text{const}$ holds for all configurations i . Equating two configurations i, j and eliminating $\mathbf{T}_P$ yields the classical hand-eye equation:

$$\underbrace{\mathbf{T}_{F,j}^{-1} \mathbf{T}_{F,i}}_{\mathbf{A}_{ij}} \cdot \mathbf{X} = \mathbf{X} \cdot \underbrace{\mathbf{T}_{m,j} \mathbf{T}_{m,i}^{-1}}_{\mathbf{B}_{ij}}, \quad (1)$$

where $\mathbf{A}_{ij}$ is the relative flange motion from forward kinematics and $\mathbf{B}_{ij}$ the relative target displacement observed by the sensor. The absolute target pose cancels out and is neither measured nor required.

Intensity-Based Corner Detection

At each configuration, the E1R acquires a point cloud of the target. Each point carries a range value (resolution 5 mm) and an 8-bit intensity value ($0\text{--}255$), where Reflect-O-Lay fields produce high and PA12+CF15 fields produce low return intensities. The processing pipeline operates as follows:

1. **Intensity image formation.** Per-point intensity values are projected onto a regular angular grid (0.625° spacing) using the known sensor beam geometry, yielding a 197×143 pixel intensity image.
2. **Region-of-interest extraction.** The target region is segmented by intensity thresholding (fraction 0.30 of the image maximum) followed by morphological filtering.
3. **Corner detection.** The intensity image is up-scaled by a factor of 8 (bicubic interpolation), binarised (Otsu thresholding), and processed with `findChessboardCornersSB` [15] with sub-pixel refinement. Multiple preprocessing variants (Gaussian blur kernels, polarity inversion) are attempted.
4. **3D back-projection.** Each detected 2D corner is mapped back to 3D by intensity-weighted averaging of the neighbouring point cloud returns within a $\pm 1.2^\circ$ angular window.

The result is a set of 16 corner positions $\mathbf{c}_k^{(i)} \in \mathbb{R}^3$ ($k = 1, \dots, 16$) per detected pose, expressed in the sensor coordinate frame.

Parameter Estimation

Rather than constructing intermediate target-pose matrices $\mathbf{T}_{m,i}$ from the detected corners, a direct scatter minimisation approach is employed. For a candidate $\mathbf{X} \in SE(3)$, each corner is projected to the base frame:

$$\mathbf{p}_k^{(i)} = \mathbf{T}_{F,i} \cdot \mathbf{X} \cdot \mathbf{c}_k^{(i)}. \quad (2)$$

Since the target is stationary, $\mathbf{p}_k^{(i)}$ should be identical across all poses i for each corner k . The optimal $\hat{\mathbf{X}}$ is obtained by minimizing total scatter:

$$\hat{\mathbf{X}} = \arg \min_{\mathbf{X} \in SE(3)} \sum_{k=1}^{16} \sum_{i=1}^{N_{\text{eval}}} \left\| \mathbf{p}_k^{(i)} - \bar{\mathbf{p}}_k \right\|^2, \quad (3)$$

where $\bar{\mathbf{p}}_k = \frac{1}{N_{\text{eval}}} \sum_i \mathbf{p}_k^{(i)}$ is the mean position of corner k across all N_{eval} evaluated poses. This formulation uses all $16 \times N_{\text{eval}}$ corner observations equally and avoids the error-prone construction of 6-DoF target-pose matrices from subsets of corners. The optimisation is solved by Levenberg–Marquardt initialised from the identity.

Experimental Results

A total of $N_{\text{total}} = 20$ robot configurations are collected, distributed to ensure variation in all six degrees of freedom with translational offsets of $0.4 \text{ m} - 1.0 \text{ m}$ and angular tilts up to $\pm 40^\circ$ about the target normal. For each configuration, a static point cloud is acquired. After applying the processing pipeline, $N_{\text{eval}} = 10$ poses covering the full intended working envelope (sensor-to-target distance $0.46 \text{ m} - 1.01 \text{ m}$) are retained for the accuracy analysis. The remaining configurations include cases at extreme viewing angles

or low point density and are excluded from the quantitative evaluation; extending the approach to such limiting conditions is discussed in the Conclusions.

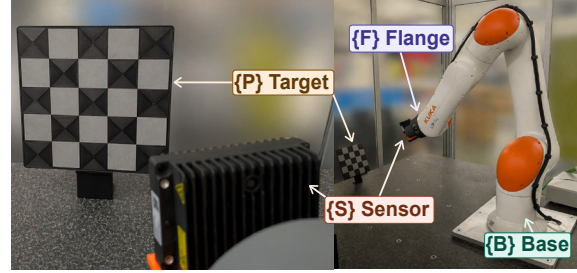


Fig. 3: Experimental setup. KUKA LBR iisy 11 R1300 with RoboSense E1R on the flange and AM target at a fixed position.

Corner Detection Accuracy

The primary metric is the Procrustes fit: for each detected pose, the 16 corner positions in sensor coordinates are rigidly aligned (rotation and translation, no scaling) to the nominal 4×4 grid (45 mm spacing). Tab. 1 reports the per-pose residuals.

Tab. 1: Corner detection accuracy per pose. Procrustes RMSE and maximum per-corner error (rigid alignment to nominal 4×4 grid with 45 mm spacing, no scaling).

Pose	d (m)	RMSE (mm)	Max. (mm)
1	0.73	6.3	11.4
2	1.01	8.6	13.5
3	0.83	7.3	13.6
4	0.64	9.4	15.7
5	0.49	8.1	14.8
6	0.66	8.1	14.8
7	0.52	8.2	13.8
8	0.82	6.6	12.9
9	0.46	6.1	12.4
10	0.61	7.1	16.2
Mean	0.68	7.6	13.9
Std.	0.18	1.0	1.4

Across all ten evaluated poses (sensor-to-target distance $0.46 \text{ m} - 1.01 \text{ m}$), the Procrustes RMSE is $7.6 \pm 1.0 \text{ mm}$ with a maximum per-corner error of $13.9 \pm 1.4 \text{ mm}$.

Grid Geometry Consistency

As a complementary metric independent of rigid-body fitting, the inter-corner spacings are compared directly to the nominal value. Across all 240 adjacent-corner distances (horizontal and vertical) from the ten evaluated poses, the measured mean spacing is 45.18 mm (nominal:

Tab. 2: Comparison of reference calibration targets for LiDAR. Reported accuracy values refer to corner- or feature-localization RMSE on the target surface; values are approximated from the cited references and depend on sensor and distance.

Reference	Fabrication	Contrast source	Accuracy (mm)
Pandey et al. [3]	Printed paper on board	Ink reflectance	~10–20
Zhou et al. [4]	Printed checkerboard	Ink reflectance	~10–15
Beltrán et al. [11]	Milled planar board with cut-outs	Geometric (depth edges)	~15
Ren et al. [10]	Stacked planar boards (3D tower)	Printed pattern	~8–12
Kaasalainen et al. [9]	Natural reference samples	Material reflectance	n/a (intensity ref.)
This work	Multi-material AM checkerboard	Material reflectance	7.6 ± 1.0

45.0 mm), corresponding to a systematic bias of $+0.18$ mm ($+0.4\%$) with a standard deviation of 5.8 mm. The planarity of the detected grid (distance of each corner from the best-fit plane through the 16 points) yields an RMSE of 5.2 ± 1.1 mm. Both the spacing deviation and the out-of-plane scatter are consistent with the specified ranging precision of the E1R sensor (± 50 mm at 1σ), confirming that the observed corner position errors are dominated by sensor noise rather than systematic target artefacts.

Multi-Pose Consistency

In order to assess the repeatability of the cross-pose corner extraction, the detected corners from all ten evaluated poses are transformed into a common reference frame using the measured flange pose and the estimated flange-to-sensor offset $\hat{X}$ from Eq. (3). For each of the 16 corners, the scatter of the ten observations around their mean position is evaluated. The resulting overall RMSE is 10.9 mm across all $16 \times 10 = 160$ projected corner positions, which is consistent with the single-pose Procrustes RMSE combined with the specified ranging precision of the E1R sensor and confirms that corner extraction is reproducible across the workspace.

Comparison with Target References

Tab. 2 compares the proposed AM target against a representative selection of calibration targets reported in the recent literature [3, 4, 11, 10, 9]. The comparison considers the target fabrication route, the source of the reflectivity contrast, and the reported corner or feature localization accuracy. The comparison is intended to situate the proposed target within the established accuracy regime rather than to identify a single best performer, since absolute accuracy figures are influenced by the LiDAR angular resolution and working distance.

The proposed AM target achieves corner localization within the accuracy range reported for established printed or machined targets [3, 4, 11, 10], while offering two distinctive features. First, the reflectivity contrast originates from the bulk material rather than from a surface coating, following the rationale of material-based radiomet-

ric reference targets [9] and thereby avoiding the coating-degradation mechanisms identified as a limitation of printed checkerboards. Second, the AM fabrication route allows the target geometry, field size, aspect ratio, or out-of-plane features to be reconfigured through the CAD model without any tooling-related cost, a flexibility that machined planar boards [11] or stacked 3D targets [10] cannot provide.

Conclusions

A multi-material additively manufactured calibration target has been proposed and experimentally evaluated for intensity-based LiDAR corner detection in an eye-in-hand robotic configuration. Through the combination of Reflect-O-Lay and PA12+CF15 as retroreflective and absorptive constituents, a stable near-infrared intensity contrast is embedded within the material composition, eliminating reliance on surface coatings and the associated degradation mechanisms. The 5×5 checkerboard with 45 mm field size provides 16 inner calibration corners that are automatically detectable by standard algorithms applied to the LiDAR intensity channel.

Experimental evaluation on a RoboSense E1R sensor ($\lambda = 940$ nm, angular resolution 0.625°) mounted on a KUKA LBR iisy 11 R1300 yields a corner detection accuracy of 7.6 ± 1.0 mm (Procrustes RMSE) with a spacing bias below 0.5% and a multi-pose repeatability of 10.9 mm RMSE. The observed accuracy is consistent with the specified ranging precision of the sensor and places the proposed AM target within the accuracy range reported for established printed, milled, and stacked calibration targets, while offering superior material stability and geometric design freedom.

Future work will exploit the design freedom of additive manufacturing to produce non-planar target geometries, including tilted, stepped, or multi-facet arrangements, that are expected to improve feature observability at extreme viewing angles and reduce the number of required robot configurations. Higher-resolution LiDAR sensors are also expected to lower the corner localization error below the current sensor-noise-limited regime.

References

- [1] R. Y. Tsai and R. K. Lenz, A new technique for fully autonomous and efficient 3D robotics hand/eye calibration, *IEEE Trans. Robot. Autom.* 5, 345–358 (1989); doi: 10.1109/70.34770
- [2] Y. C. Shiu and S. Ahmad, Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form $AX=XB$, *IEEE Trans. Robot. Autom.* 5, 16–29 (1989); doi: 10.1109/70.88014
- [3] G. Pandey, J. R. McBride, S. Savarese, and R. M. Eustice, Extrinsic calibration of a 3D laser scanner and an omnidirectional camera, *IFAC Proc. Vol.* 43, 336–341 (2010); doi: 10.3182/20100906-3-IT-2019.00059
- [4] L. Zhou, Z. Li, and M. Kaess, Automatic extrinsic calibration of a camera and a 3D LiDAR using line and plane correspondences, *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 5562–5569 (2018); doi: 10.1109/IROS.2018.8593660
- [5] Q. Zhang and R. Pless, Extrinsic calibration of a camera and laser range finder (improves camera calibration), *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2301–2306 (2004); doi: 10.1109/IROS.2004.1389752
- [6] J. Levinson and S. Thrun, Automatic on-line calibration of cameras and lasers, *Robotics: Science and Systems IX* (2013); doi: 10.15607/RSS.2013.IX.029
- [7] M. Mattamala, J. Frey, P. Libera, N. Chebrolu, G. Martius, C. Cadena, M. Hutter, and M. Fallon, Extrinsic calibration of camera to LiDAR using a differentiable checkerboard model, *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 10404–10411 (2023); doi: 10.1109/IROS55552.2023.10341383
- [8] W. Wang, K. Sakurada, and N. Kawaguchi, Reflectance intensity assisted automatic and accurate extrinsic calibration of 3D LiDAR and panoramic camera using a printed chessboard, *Remote Sensing* 9, 851 (2017); doi: 10.3390/rs9080851
- [9] S. Kaasalainen et al., Radiometric calibration of LIDAR intensity with commercially available reference targets, *IEEE Trans. Geosci. Remote Sens.* 47, 588–598 (2009); doi: 10.1109/TGRS.2008.2003351
- [10] F. Ren et al., Extrinsic calibration of camera and LiDAR systems with three-dimensional towered checkerboards, *Int. J. Intell. Syst.* 2024, 2478715 (2024); doi: 10.1155/2024/2478715
- [11] J. Beltrán, C. Guindel, A. de la Escalera, and F. García, Automatic extrinsic calibration method for LiDAR and camera sensor setups, *IEEE Trans. Intell. Transp. Syst.* 23, 17677–17689 (2022); doi: 10.1109/TITS.2022.3155228
- [12] K. Parthy, Reflect-O-Lay: Retroreflective FDM filament—product description and print parameters, Lay Filaments, Germany (2015). Available: <https://filament2print.com/en/pla/690-reflect-o-lay.html>
- [13] Fiberlab S.A., Fiberlogy Nylon PA12+CF15: Technical datasheet, Pruszków, Poland (2023). Available: <https://fiberlogy.com/en/Nylon-PA12CF15-Filament>
- [14] C. Amendola et al., Optical characterization of 3D-printed PLA and ABS filaments for diffuse optics applications, *PLOS ONE* 16, e0253181 (2021); doi: 10.1371/journal.pone.0253181
- [15] G. Bradski, The OpenCV Library, *Dr. Dobb's Journal of Software Tools* (2000). Available: <https://opencv.org>

Analyse der UAV-basierten Methanferndetektion unter variierenden Reflexionseigenschaften von Oberflächen

*Eric Maier^{1,2}, Johannes Herbst², Sven Rademacher², Thomas Strahl^{1,2}, Dominik M. Spale³,
Stefan J. Rupitsch³, Jürgen Wöllenstein^{1,2}, Katrin Schmitt^{1,2}*

¹ Universität Freiburg, IMTEK – Gassensoren, Georges-Köhler-Allee 102, 79110 Freiburg

² Fraunhofer IPM, Georges-Köhler-Allee 301, 79110 Freiburg

³ Universität Freiburg, IMTEK – Elektrische Messtechnik und Eingebettete Systeme, Georges-Köhler-Allee 106, 79110 Freiburg

Zusammenfassung

In dieser Arbeit wird ein kompaktes, drohnenbasiertes Sensorsystem zur Methanferndetektion vorgestellt. Das Sensorsystem basiert auf einer Wellenlängenmodulationsspektroskopie (WMS) mit einem DFB-Laser bei 1651 nm. Die Bestimmung der Methan-Säulenkonzentrationen wird über eine WMS-2f/1f-Demodulation ermöglicht. Der Sensor wurde hinsichtlich Gewicht, Leistungsaufnahme und Messreichweite für den Einsatz auf einem Quadropter optimiert. Feldmessungen zeigen, dass die Messergebnisse stark von der Stärke des rückgestreuten Laserlichts abhängen. Nicht-kooperative Rückstreu hintergründe und große Messentfernungen führen zu erhöhtem Messrauschen und können systematische Überschätzungen der Methan-Säulenkonzentration verursachen. Es konnte eine zuverlässige Methanferndetektion demonstriert werden, unter anderem durch Überflüge über einen methangefüllten Gassack als simuliertes Gasleck. Die Ergebnisse verdeutlichen das Potenzial des Systems für die drohnenbasierte Leckageerkennung an gasführender Infrastruktur und zeigen zugleich Ansatzpunkte für zukünftige Verbesserungen auf, insbesondere hinsichtlich der optischen Empfangseinheit und der Signalverarbeitung.

Keywords: Methandetektion, Gassensorik, Laserabsorptionsspektroskopie, Drohnenbasierte Messsysteme, Optische Ferndetektion

Motivation und Einführung

Die Reduktion von Methan-Emissionen (CH_4) ist eine der effektivsten Maßnahmen zur kurzfristigen Eindämmung des Klimawandels, da Methan über Jahrzehnte hinweg eine starke Treibhauswirkung entfaltet. Aktuelle globale Bewertungen unterstreichen, dass eine rasche Reduktion von CH_4 die Erderwärmung in den kommenden Jahrzehnten messbar verlangsamen kann [1]. Studien zeigen zudem, dass selbst moderate Leckagen entlang der Erdgas-Infrastruktur die erwarteten Klimavorteile gegenüber anderen fossilen Brennstoffen zunichtemachen können. Eine präzise Detektion und Kontrolle diffuser Emissionen ist daher unerlässlich [2].

Neben dem Klimaschutz erfordern auch sicherheitstechnische Aspekte – etwa nach einem Erdbeben oder anderen Katastrophen im Bereich gasführender Anlagen – eine schnelle und flächendeckende Erkennung von Gefahrenzonen. Um Einsatzkräfte keinen unnötigen

Risiken auszusetzen, bieten sich berührungslose, drohnenbasierte Messsysteme an. Diese können aus geringer Höhe auch komplexes, teils unzugängliches Gelände absuchen und liefern georeferenzierte, quantitative Ergebnisse [3, 4]. Insbesondere Multikopter eignen sich als Trägerplattformen, da sie präzise Suchflüge in geringer Höhe sowie punktuell Schweben (Hovering) zur gezielten Verifikation über potenziellen Leckagen ermöglichen. Niedrige Arbeitshöhen verkürzen den optischen Pfad, verbessern die Rückstreugeometrie und erhöhen so das Signal-zu-Rausch-Verhältnis [3–5]. Kombiniert mit einer Signalverarbeitung an Bord lassen sich so direkt im Feld Frühwarnungen generieren, die bei Einsätzen zur Entscheidungsfindung beitragen [4].

Der in dieser Arbeit verfolgte Ansatz basiert auf der durchstimmbaren Laserspektroskopie (Tunable Laser Spectroscopy - TLS) in einer Standoff-Geometrie, bei der nicht-kooperative Ziele beleuchtet und das diffus rückgestreute Licht detektiert wird. Um eine rein optische Fernmessung zu realisieren, wird die Wellenlängenmodulationsspektroskopie (WMS) im nahen Infrarot bei etwa 1651 nm eingesetzt. Dies

ermöglicht kompakte, energieeffiziente und schnelle Messungen der CH_4 -Säulenkonzentration (in ppm·m), die sich ideal für die UAV-gestützte Gefahrenkartierung eignen [3].

Grundlagen der WMS zur Gaskonzentrationsbestimmung

Die abstimmbare Laserspektroskopie im nahen und mittleren Infrarot (NIR/MIR) eignet sich aufgrund ihrer hohen spektralen Selektivität und Empfindlichkeit besonders gut zur quantitativen Bestimmung gasförmiger Moleküle. Zu den wichtigsten Verfahren zählen die direkte Absorptionsspektroskopie (DAS) und die Wellenlängenmodulationsspektroskopie, die sich hinsichtlich Signalentstehung und Auswertverfahren unterscheiden.

Bei der DAS wird die Emissionswellenlänge eines monochromatischen Lasers über eine Absorptionslinie des Zielmoleküls abgestimmt. Die dabei auftretende spektrale (wellenzahlabhängige) Abschwächung der Intensität wird gemäß dem Beer-Lambert-Gesetz beschrieben:

$$I(\tilde{\nu}) = I_0(\tilde{\nu})e^{-\alpha(\tilde{\nu})\ell}. \quad (1)$$

Dabei bezeichnet $I_0(\tilde{\nu})$ die spektrale Ausgangsintensität des Lasers, $\alpha(\tilde{\nu})$ den spektralen Absorptionskoeffizienten des Zielgases und ℓ die optische Weglänge durch das Gas. Hierbei ist es üblich die Wellenzahl $\tilde{\nu}=1/\lambda$ (in cm^{-1}) für die mathematische Beschreibung und Darstellung anstatt der Wellenlänge (z.B. $\lambda = c/\nu$ in nm bzw. optische Frequenz ν in Hz) zu verwenden.

In der WMS wird die Emissionswellenlänge des Lasers mit einer harmonischen Modulation überlagert, sodass es zu einer periodischen Absorption des Laserlichts kommt. Diese resultierende periodische Intensitätsmodulation (IM) lässt sich mit einem Photodetektor messen. Während die DAS ein unverändertes Absorptionsprofil liefert, verschiebt die Wellenlängenmodulationsspektroskopie das Absorptionssignal in höhere Harmonische der Modulationsfrequenz. Diese Signaltransformation ermöglicht eine schnelle, ressourcenschonende, robuste Detektion mittels Lock-In-Techniken und macht die WMS besonders interessant für drohnenbasierte Gasferndetektion.

Ausgehend von Gl. (1) wird das Produkt $\alpha(\tilde{\nu})\ell$ auch spektrale Absorbanz (A) genannt, welche über den negativen natürlichen Logarithmus in Verbindung mit der Transmission (τ) steht. Es gilt $A(\tilde{\nu}) = \alpha(\tilde{\nu})\ell = -\ln(\tau)$, sodass Beer-Lambert auf unterschiedlichen Wegen eingeführt oder auch ausgedrückt werden kann. Des Weiteren lässt sich der spektrale Absorptionskoeffizient durch die Teilchenzahldichte (n), die Linienstärke ($S(T)$) und entsprechende

Linienform (ϕ) des Zielmoleküls in Bezug setzen ($\alpha(\tilde{\nu}) = nS(T)\phi(\tilde{\nu})$). Die Teilchendichte des Zielgases ist unter Berücksichtigung des idealen Gasgesetzes direkt proportional zur Gaskonzentration (χ) des Zielgases, sodass die Absorbanz wie folgt beschrieben werden kann:

$$A(\tilde{\nu}) = \alpha(\tilde{\nu})\ell = \chi \frac{P}{k_B T} S\ell\phi(\tilde{\nu}). \quad (2)$$

Die Linienform ($\phi = \phi(\tilde{\nu}) = \phi(\tilde{\nu}, T, P)$), auch Linienprofil genannt, ist normiert, sodass $\int d\tilde{\nu}\phi(\tilde{\nu}) = 1$ gilt. Sie berücksichtigt auch die beiden vorherrschenden Verbreiterungsmechanismen der Linienbreite wie Stoß- und Dopplerverbreiterung, welche druck- (p) bzw. temperaturabhängig (T) sind. Mathematisch wird die Linienform durch ein Voigtprofil oder bei Standardbedingungen in guter Näherung durch ein Lorentzprofil beschrieben. Zur Berechnung oder Simulation der Absorptionslinie(n) werden die spektroskopischen Parameter aus der HITRAN-Datenbank verwendet [6]. Für eine Methankonzentration von 1 ppm bei einer optischen Weglänge von 1 m unter Standardbedingungen (296 K und 1013 mbar) ergibt sich ein Maximum der Absorbanz von 5×10^{-5} bei 1651 nm bzw. 6057 cm^{-1} (s. Abb. 1)

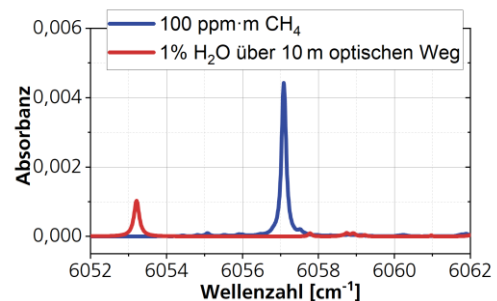


Abb. 1: Die auf HITRAN basierende Simulation zeigt die Absorption von CH_4 und Wasserdampf (1% absolute Feuchte) im Bereich der Messwellenlänge von 1651 nm.

Das statische elektrooptische Verhalten eines durchstimmbaren Lasers kann anhand des Injektionsstroms (i in mA) und Lasertemperatur (T in $^{\circ}\text{C}$) für einen spezifischen Arbeitspunkt mit der Wellenzahl $\tilde{\nu}_p = \tilde{\nu}(T_p, i_p)$ bzw. Intensität $I_0^p = I_0(T_p, i_p)$ charakterisiert werden. Bei diesen Lasern verhalten sich die Intensität und Wellenlänge näherungsweise linear zum Laserstrom, d.h. eine kleine Stromänderung über die Zeit führt zu einer Leistungs- bzw. Intensitäts- ($P_L \sim I_0$) und Wellenlängenmodulation des Lasers (kurz IM und WM).

Folglich führt eine kleine sinusförmige Modulation des Laserstroms um i_p zu einer kleinen WM bzw. IM um die Arbeitspunkte:

$$i(t) = i_p + \Delta i \sin(2\pi f_m t) \quad (3a)$$

$$\tilde{v}(t) = \tilde{v}_p + \Delta \tilde{v} \sin(2\pi f_m t) \quad (3b)$$

$$I_0(t) = I_0^p + \Delta I_0 \sin(2\pi f_m t) \quad (3c)$$

Typischerweise befindet sich die Modulationsfrequenz f_m im kHz-Bereich und die Stromamplitude Δi ist in der Regel signifikant kleiner als der Strom am Arbeitspunkt i_p . In der Praxis folgen die Emissionswellenlänge bzw. -wellenzahl und Intensität der kleinen Strommodulation. Je nach Laserstruktur und Modulationsfrequenz ist jedoch eine Phasenverschiebung zwischen IM und WM zu beobachten, welche an dieser Stelle vernachlässigt wird.

Das Prinzip der WMS wird zum besseren Verständnis unter Vernachlässigung der zeitabhängigen IM des Lasers eingeführt (vgl. Gl. (3c)), es gilt $I_0(t) = I_0^p$. Unter Berücksichtigung der WM gemäß Gl. (3b) lässt sich das detektierte Signal gemäß dem Beer–Lambert-Gesetz für eine kleine Absorptionen wie folgt ausdrücken:

$$I(t) \approx I_0^p \left[1 - \chi \frac{P}{k_B T} S \ell \phi(\tilde{v}(t)) \right]. \quad (4)$$

Da $\tilde{v}(t)$ periodisch mit der Modulationsfrequenz f_m variiert, lässt sich das detektierte Signal $S = I(t)$ in einer Fourier-Reihe entwickeln. Es gilt

$$S = \sum_{n=0}^{\infty} S_n \cdot \cos(n \cdot \theta) \quad (5)$$

wobei S_n Fourier-Koeffizienten (bzw. Harmonischen) sind und $\theta = 2\pi f_m t$ einer Substitution entspricht. Unter Vernachlässigung der IM des Lasers bzw. konstanten Intensität ($I_0(t) = I_0^p$) in Kombination mit einem Lorentzprofil als Linienform können die Harmonischen analytisch berechnet werden. Das Lorentzprofil ist durch über die zentrale Absorptionswellenzahl $\tilde{v}_0$ und die halbe Halbwertsbreite (HWHM, γ) der Absorptionslinie beschrieben:

$$\phi(\tilde{v}) = \frac{1}{\pi \gamma} \left[1 + \left(\frac{\tilde{v} - \tilde{v}_0}{\gamma} \right)^2 \right]^{-1} \quad (6)$$

Ausgehend von Arndt et al. [7] lassen sich die Signalkomponenten in Form der n-ten Harmonischen unter Berücksichtigung der WM gemäß Gl. (3b), (4) und (5) wie folgt beschreiben:

$$S_n = I_0^p \left[J_n(0) - \chi \frac{PS\ell}{k_B T} \frac{1}{\pi \gamma} H_n(x, m) \right] \quad (7)$$

Hierbei ist $J_n(0)$ die Besselfunktion n-ter Ordnung und

$$H_n(x, m) = \frac{\varepsilon_n i^n}{2 m^n} \frac{\left[\sqrt{(1-ix)^2 + m^2} - (1-ix) \right]^n}{\sqrt{(1-ix)^2 + m^2}} + c.c.$$

Hierbei ist $\varepsilon_0 = 1$ bzw. $\varepsilon_n = 2$ für $n > 0$, c.c. das komplex Konjugierte, $m = \Delta \tilde{v} / \gamma$ der Modulationsindex und $x = (\tilde{v} - \tilde{v}_0) / \gamma$.

Ausgehend von dieser Beschreibung kann gezeigt werden, dass für kleine m die Fourier-Koeffizienten S_n proportional zur n-ten Ableitung des Linienprofils $\partial^n \phi / \partial \tilde{v}^n$ der Lorentzfunktion sind. Folglich ist eine WM rund um die zentrale Absorptionswellenzahl ideal geeignet zur Gasanalyse mit Hilfe der 2. Harmonischen (S_2). Diese ist proportional zur Krümmung der Absorptionslinie und damit zur Konzentration des Zielgases. In der Praxis lässt sich der Realteil von S_2 über ein Lock-in-Verfahren bei der doppelten Modulationsfrequenz ($2f$) ermitteln.

S_2 ist nahezu hintergrundfrei (kein Offset), was die Detektion kleinster Konzentrationen ermöglicht. Trotzdem skaliert es linear mit der absoluten Lichtintensität. Eine Abschwächung des Laserlichts führt fälschlicherweise zu einem geringeren Konzentrationswert. Eine Normierung auf die erste Harmonische S_1 löst diese Problematik. Dies ist insbesondere für die Ferndetektion mit stark schwankenden Lichtintensitäten bedingt durch unterschiedliche Rückstreuoberflächen zwingend notwendig. Die Messgröße für die Konzentration ist somit das Verhältnis zwischen zweiter und erster Harmonischen:

$$\chi \propto \frac{S_2}{S_1} \quad (8)$$

Diese Methode verspricht eine erhöhte Robustheit in der Ferndetektion. Bei geringen Absorptionen ist das Verhältnis dieser Harmonischen ($2f/1f$) weitestgehend nur noch von den molekularen Parametern und der gewählten Modulationsamplitude abhängig.

Sensorkonzept

Das Messsystem basiert auf der durchstimmbaren Laserabsorptionsspektroskopie und arbeitet wie folgt: Ein Distributed-Feedback-Laser (DFB-Laser) der Firma Eblana Photonics wird über eine spezifische Absorptionslinie von Methan abgestimmt (hier: 1651 nm, vgl. Abb. 1). Die Durchstimmung erfolgt über den

Lasertreiber der Auswerteelektronik. Hierzu wird mittels eines Digital-Analog-Wandlers (DAC) ein sinusförmiger Modulationsstrom mit einer Frequenz von etwa 1 kHz ausgegeben und dem Laser zugeführt. Der Laser wandelt diesen Strom in optische Strahlung um, wobei die emittierte Wellenlänge in Abhängigkeit vom Laserstrom variiert.

Die Einstellung und Stabilisierung der Lasertemperatur erfolgt über das integrierte thermoelektrische Kühlsystem (TEC). Über die Lasertemperatur wird die zentrale Emissionswellenlänge des Lasers eingestellt. Diese wird so gewählt, dass sich die Methan-Absorptionslinie an den Wendepunkten der sinusförmigen Wellenlängenmodulation befindet.

Das modulierte Laserlicht wird mithilfe einer Asphärenlinse (Thorlabs C330TMD-C) nahezu kollimiert und auf einen halben Öffnungswinkel von etwa 3 mrad eingestellt. Der Laserstrahl wird nach unten emittiert, dort am Untergrund reflektiert und diffus gestreut. Befindet sich Methan innerhalb der optischen Wegstrecke, so wird ein Teil des Laserlichts im Bereich der Absorptionslinie absorbiert. Die absorbierte Lichtmenge ist dabei proportional zur Methankonzentration entlang des Strahlengangs.

Zur Detektion wird ein Teil des rückgestreuten Lichts mittels einer Fresnel-Linse (Durchmesser 2“, Brennweite 32 mm) eingesammelt. Nach spektraler Filterung durch einen optischen Hochpassfilter (Thorlabs FELH1600) wird das Licht auf eine InGaAs-Photodiode (Thorlabs FGA21) fokussiert. Die Photodiode wandelt das modulierte optische Signal in ein entsprechendes sinusförmiges Stromsignal um. Die nachgeschaltete Vorverstärkereinheit besteht aus einem Transimpedanzverstärker (TIA) sowie weiteren Verstärker- und Filterstufen, welche das Stromsignal in ein Spannungssignal umwandeln und für die digitale Weiterverarbeitung aufbereiten.

Das aufbereitete Signal wird über einen 16-Bit-Analog-Digital-Wandler (ADC) erfasst und von einem Mikrocontroller auf der Auswerteelektronik digital verarbeitet. Zur spektralen Auswertung wird ein Görtzel-Algorithmus eingesetzt, mit dem gezielt die Amplituden der ersten (1f) und zweiten Harmonischen (2f) der Modulationsfrequenz bestimmt werden. Das 2f-Signal enthält dabei die absorptionsbedingte Gasinformation, während das 1f-Signal überwiegend die Reflektivität des rückstreuenden Hintergrunds repräsentiert. Durch Normierung des 2f-Signals mit dem 1f-Signal und Multiplikation mit einem Kalibrierfaktor k ($\frac{2f}{1f} \cdot k$) wird die Methan-Säulenkonzentration in der Einheit ppm·m berechnet. Eine schematische

Darstellung des Messprinzips ist in Abb. 2 gezeigt.

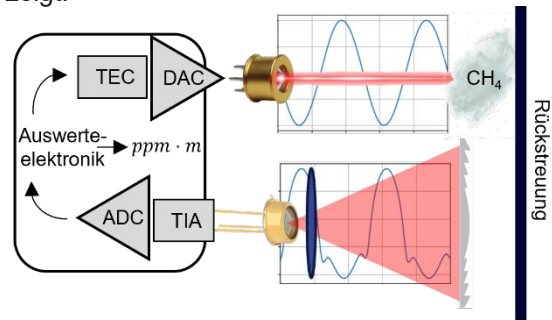


Abb. 2: Schematischer Aufbau und Funktionsprinzip des TDLAS-Messsystems.

Der Sensor dient der Detektion von Erdgasleckagen und soll frühzeitig vor Methankonzentrationen im Bereich der unteren Explosionsgrenze (UEG; 4,4 Vol.-% CH₄ in Luft) warnen. Zu diesem Zweck wird der Sensor auf einer Drohne (Quadrocopter) installiert. Die Drohne überfliegt das zu untersuchende Gelände, beispielsweise entlang von Pipelines oder anderen gasführenden Infrastrukturen.

Um für den Einsatz auf Drohnen geeignet zu sein, muss das Sensorsystem spezifische Anforderungen erfüllen. Insbesondere sind ein geringes Gewicht sowie eine niedrige Leistungsaufnahme erforderlich, da andernfalls die Flugeigenschaften und die verfügbare Akkulaufzeit der Drohne erheblich beeinträchtigt werden. Das Zielgewicht des Sensorsystems liegt dabei idealerweise unter 500 g.

Das realisierte Sensorsystem weist Außenabmessungen von 95 × 70 × 60 mm³ auf und kommt auf eine Gesamtmasse von 314 g. Die gesamte elektrische Leistungsaufnahme beträgt etwa 2 W. Mit dem Sensor lassen sich Messdistanzen von bis zu 60 m erreichen, auch auf nicht-kooperative Hintergrundziele, beispielsweise eine weiß folierte Tür. Eine fotografische Darstellung des Sensors ist in Abb. 3 gezeigt.

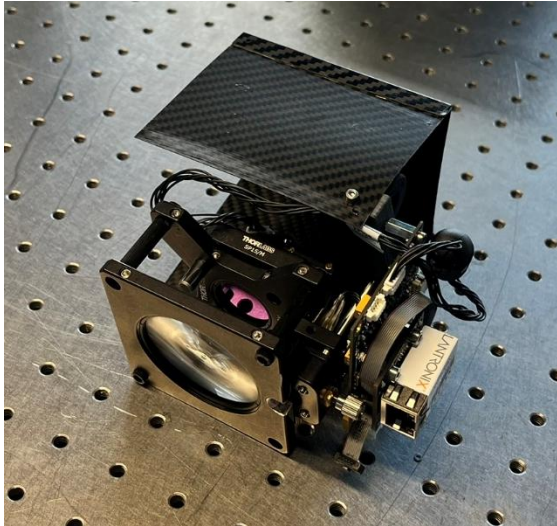


Abb. 3: Sensorsystemaufbau, erkennbar sind die Fresnel-Linse zur Sammlung des rückgestreuten Laserlichts, der optische Filter mit charakteristischer pinker Färbung, die optomechanische Baugruppe sowie die Auswerteelektronik.

Für Feldmessungen wurde der Sensor auf einem Quadroptor montiert. Als Trägerplattform kam eine Drohne auf Basis des Typs PremiumModellbau PM Q-685 zum Einsatz. Diese ist mit einem Pixhawk 2.1 Cube Orange+ Flightcontroller ausgestattet, der als Open-Source-System eine flexible und anwendungsspezifische Konfiguration ermöglicht. Der Multicopter besitzt ein Eigengewicht von etwa 1600 g und erreicht eine maximale Flugzeit von rund 50 Minuten. Bei einer Nutzlast von 500 g beträgt die Flugzeit circa 40 Minuten.

Die Positionsstabilisierung der Drohne erfolgt mittels GPS in Kombination mit Real-Time-Kinematic-Verfahren (RTK). Die Basisreferenzierung wird dabei über den RTK-Dienst SAPOS realisiert. Die Übertragung der Messdaten vom Sensor zur Bodenstation erfolgt drahtlos über WLAN.

Im Feldbetrieb auf der Drohne treten mehrere messrelevante Herausforderungen auf. Zum einen nimmt die Intensität des eingesammelten rückgestreuten Laserlichts mit zunehmender Messdistanz deutlich ab, was zu einem Anstieg des Messrauschens führt. Darüber hinaus verursacht das Hintergrundlicht durch Sonneneinstrahlung einen zusätzlichen Offset im Detektorsignal. Aufgrund dieses Offsets muss die Vorverstärkerschaltung mit TIA so ausgelegt werden, dass die erste Verstärkerstufe nur eine begrenzte Verstärkung aufweist, um eine Sättigung durch das Hintergrundsignal zu vermeiden. Nach der ersten Verstärkerstufe wird das niederfrequente Hintergrundsignal mittels eines

geeigneten Hochpassfilters unterdrückt. Erst anschließend erfolgt eine stärkere Verstärkung, um die sehr schwachen optischen Messsignale des rückgestreuten Laserlichts weiter aufzubereiten. Die optische Leistung des gesammelten Messsignals beträgt dabei, abhängig von der Messdistanz, lediglich wenige Nanowatt.

Eine weitere Herausforderung stellen Vibrationen sowie lagebedingte Stabilisierungsbewegungen der Drohne dar. Da in der vorliegenden Messkonfiguration auf den Einsatz eines Gimbals verzichtet wird, folgt der optische Messstrahl den Bewegungen des Multicopters. Dies führt dazu, dass der Laserstrahl insbesondere im Schwebeflug nicht kontinuierlich auf denselben Rückstreu hintergrund gerichtet ist. Dieser Effekt verstärkt sich mit zunehmender Flughöhe.

Messungen

Zur Bewertung der Messperformance des Sensors unter realen Feldbedingungen wurden gezielte Flug- und Messkampagnen durchgeführt. Ziel dieser Untersuchungen war es, den Einfluss zentraler Randbedingungen des Drohnenbetriebs auf die Messqualität systematisch zu analysieren.

Einfluss von Rückstreu hintergrund und Messentfernung

Zur Untersuchung des Einflusses der Hintergrundbeschaffenheit sowie der Messentfernung auf das Sensorsignal wurde eine Messreihe durchgeführt, bei der sowohl der Rückstreu hintergrund als auch die Flughöhe der Drohne variiert wurden. Hierzu wurde die Drohne nacheinander über verschiedene, hinsichtlich ihrer optischen Eigenschaften unterschiedliche Rückstreu hintergründe positioniert. Für jeden Hintergrund wurde die Drohne stabil im Schwebeflug gehalten, während die Flughöhe schrittweise auf 2 m, 5 m und 10 m eingestellt wurde. Jede Höhenstufe wurde für eine Dauer von jeweils etwas mehr als 1 Minute beibehalten, um eine statistisch aussagekräftige Erfassung der Messdaten zu gewährleisten.

Während des Schwebeflugs wurde kontinuierlich die vom Sensor bestimmte Methan-Säulenkonzentration aufgezeichnet. Als Rückstreu hintergründe wurden folgende typische Feldoberflächen untersucht: nasser Asphalt, Pflastersteine, eine bewachsene Wiesenfläche sowie das Dach eines Fahrzeugs. Diese Auswahl deckt ein breites Spektrum unterschiedlicher Reflexions- und Streueigenschaften ab und erlaubt somit eine realistische Bewertung der Sensorperformance in praxisrelevanten Szenarien.

Die Ergebnisse dieser Messreihe sind in Abb. 4 dargestellt und werden im Kapitel

Diskussion detailliert analysiert und diskutiert, insbesondere im Hinblick auf Signal-Rausch-Verhältnis, Hintergrund-abhängigkeit und Entfernungseinfluss.

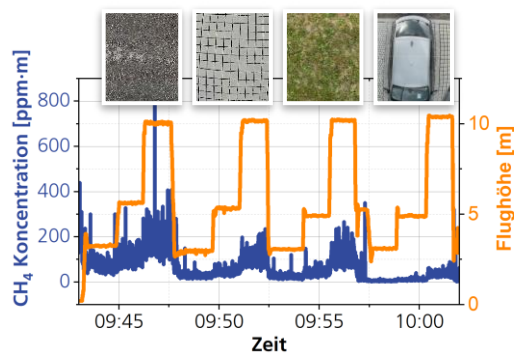


Abb. 4: Methanmessungen aus 2, 5 und 10 m Höhe über Asphalt, Pflaster, Wiese und einem Autodach.

Detektion eines simulierten Gaslecks

Zusätzlich wurde eine Messung zur Detektion eines definierten, lokal begrenzten Methanvorkommens durchgeführt. Hierzu wurde ein mit 1 Vol.-% Methan in Stickstoff gefüllter Gassack verwendet, der ein virtuelles Gasleck simulieren sollte. Der Gassack besitzt Abmessungen von $40 \times 40 \text{ cm}^2$ bei einer mittleren Höhe von etwa 20 cm und ist für Licht bei der Laser-Wellenlänge durchlässig. Um eine stabile Positionierung zu gewährleisten, wurde der Gassack mit einem KG-Rohr (Kanalgrundrohr) auf einer Asphaltfläche fixiert.

Die Messung erfolgte bei einer konstanten Flughöhe von 5 m über Grund (vgl. Abb. 5). Die Drohne überflog den Gassack mit einer horizontalen Fluggeschwindigkeit von etwa 1–2 m/s. Die Flugbahn wurde manuell so gesteuert, dass der Laserstrahl während des Überfluges möglichst auf den Bereich des Gassacks gerichtet war. Während des gesamten Überfluges wurde die Methan-Säulenkonzentration kontinuierlich aufgezeichnet.

Die Messergebnisse dieser Gaslecksimulation sind in Abb. 6 dargestellt und werden ebenfalls im Kapitel Diskussion im Hinblick auf Detektionsfähigkeit und praktische Einsatzgrenzen des Sensors diskutiert.



Abb. 5: Überflug der Drohne über ein simuliertes Gasleck. Der Gassack ist mit 1 Vol.-% Methan gefüllt und mit einem KG-Rohr am Boden fixiert.

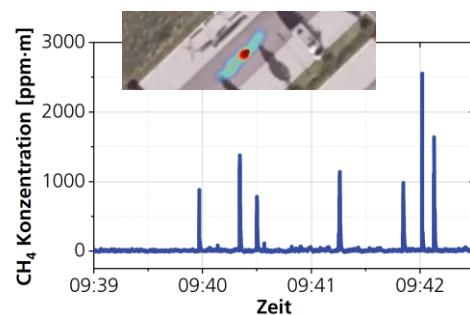


Abb. 6: Der Überflug über das simulierte Gasleck zeigt ausgeprägte Konzentrationspeaks. Mithilfe der GPS-Positionsdaten der Drohne lassen sich diese Peaks räumlich zuordnen und in einer Karte eindeutig als Leck georeferenzieren.

Diskussion

Einfluss von Rückstreu hintergrund und Messentfernung

Unter idealisierten Bedingungen ist zu erwarten, dass die gemessene Methan-Säulenkonzentration mit zunehmender Flughöhe linear ansteigt. Bei einer Methan-Hintergrundkonzentration von etwa 1,8 ppm ergeben sich theoretisch Säulenkonzentrationen von 3,6 ppm·m bei 2 m, 9 ppm·m bei 5 m sowie 18 ppm·m bei 10 m Flughöhe, sofern eine homogene

Methanverteilung entlang des optischen Pfades vorliegt.

In den durchgeführten Messungen (vgl. Abb. 4) werden diese erwarteten Werte jedoch nicht erreicht. Stattdessen zeigen sich eine deutliche Abhängigkeit der gemessenen Konzentrationen sowie der Rauschamplitude vom Rückstreu hintergrund und von der Flughöhe. Besonders über nassem Asphalt ist das Signalverhalten stark eingeschränkt. Zwar steigt die berechnete Methan-Säulenkonzentration mit zunehmender Flughöhe grundsätzlich an, die gemessenen Werte liegen jedoch deutlich über den physikalisch zu erwartenden Größenordnungen.

Diese systematische Überschätzung ist auf eine Schwäche des eingesetzten Auswertalgorithmus auf Basis einer Görtzel-FFT zurückzuführen. Bei rückstreu schwachen Hintergründen sind sowohl 1f- als auch 2f-Signalamplituden gering, sodass das absorptionspezifische 2f-Signal häufig nicht eindeutig aus dem Rauschen hervortritt. Die Amplitudenbestimmung erfolgt über die euklidische Norm der komplexen Signalanteile, die bei reinem Rauschen einen positiven Offset aufweist und dadurch zu einer Überschätzung des 2f-Signals führt. Durch die anschließende Normierung mit dem ebenfalls kleinen 1f-Signal wird dieser Effekt weiter verstärkt.

Die Stärke dieses Effekts hängt maßgeblich von der Rückstreuintensität ab. Über dem Rückstreu hintergrund "Autodach", der eine hohe Reflektivität aufweist, werden die erwarteten Säulenkonzentrationen bis zu einer Flughöhe von etwa 5 m gut reproduziert, da sich das 2f-Signal deutlich vom Rauschuntergrund abhebt.

Insgesamt zeigen die Ergebnisse, dass unter realen Feldbedingungen mit nicht-kooperativen Rückstreu hintergründen derzeit nur größere Leckagen zuverlässig detektiert werden können. Die praxisnahe Nachweisgrenze liegt bei Säulenkonzentrationen oberhalb von etwa 500 ppm·m. Die Messungen verdeutlichen zugleich, dass eine Erhöhung der eingesammelten rückgestreuten optischen Leistung, etwa durch größere Empfangsöffnungen oder eine optimierte optische Kopplung, ein zentrales Potenzial zur Verbesserung der Sensorperformance darstellt.

Detektion eines simulierten Gaslecks

Bei dieser Messung wurde ein mit 1 Vol.-% Methan gefüllter Gassack mehrfach manuell aus einer Flughöhe von 5 m über Grund überflogen. Die horizontale Fluggeschwindigkeit variierte dabei und betrug bis zu 2 m/s. Der Gassack weist eine Höhe von etwa 20 cm auf, was einem einfachen optischen Weg von 0,2 m entspricht. Daraus ergibt sich für einen idealen Überflug eine erwartete Methan-Säulenkonzentration

von etwa 2000 ppm·m. Zusätzlich ist ein statischer Offset von circa 9 ppm·m zu erwarten, der auf die Methankonzentration der Umgebungsluft zurückzuführen ist.

Wie in Abb. 6 zu erkennen ist, treten die Konzentrationspeaks nicht regelmäßig auf. Dies zeigt, dass der Gassack nicht bei jedem Überflug exakt vom Laserstrahl erfasst wird. Ursache hierfür sind lagebedingte Stabilisierungsbewegungen der Drohne, die insbesondere bei manuellem Flug ohne Gimbal zu Abweichungen der Strahlposition führen. Wird der Gassack jedoch vom Messstrahl getroffen, sind deutliche Konzentrationspeaks im Bereich zwischen 800 und 2600 ppm·m zu beobachten.

Die Abweichung der gemessenen Maximalwerte von der theoretisch erwarteten Säulenkonzentration ist vor allem auf die Relativbewegung zwischen Drohne und Gassack zurückzuführen. Durch die endliche Messzeit und die zeitliche Mittelung im Auswertalgorithmus wird der Laserstrahl nur für einen Teil des Überflugs vollständig durch das Gasvolumen geführt, wodurch die Spitzenwerte abgeschwächt werden. Da sich das Methan innerhalb des Gassacks befindet, ist es von äußeren Strömungseinflüssen, insbesondere vom Propeller-Downwash der Drohne, nicht betroffen.

Zusammenfassend lässt sich festhalten, dass der mit Methan gefüllte Gassack trotz der genannten dynamischen Effekte zuverlässig detektiert werden konnte. Die Messung bestätigt somit die grundsätzliche Fähigkeit des Systems, lokal begrenzte Methanansammlungen unter realistischen Feldbedingungen zu erkennen.

Zusammenfassung und Ausblick

In dieser Arbeit wurde ein drohnenbasiertes Sensorsystem zur gasfernen Detektion von Methan mittels durchstimmbarer Laserabsorptionsspektroskopie (TLS) vorgestellt, das unter realen Feldbedingungen charakterisiert wurde. Das kompakte Sensorsystem kombiniert einen DFB-Laser bei 1651 nm mit einer wellenlängenmodulierten Anregung und einer 2f/1f-Auswertung zur Bestimmung der Methan-Säulenkonzentration. Durch die Integration auf einen Quadropter können potenzielle Erdgasleckagen aus der Luft erfasst werden, wobei das System bei einer Gesamtmasse von 314 g und einer Leistungsaufnahme von etwa 2 W drohnenkompatibel ausgelegt ist.

Die Feldmessungen zeigen, dass die Sensorperformance maßgeblich von der Stärke des rückgestreuten Laserlichts abhängt. Insbesondere nicht-kooperative Rückstreu hintergründe und große Messdistanzen führen zu einem reduzierten Signal-zu-Rausch-Verhältnis und zu systematischen Überschätzungen der

Methan-Säulenkonzentration, die auf algorithmische Effekte bei der Amplitudenbestimmung zurückzuführen sind. Dennoch konnte die Detektion größerer Methanfreisetzungen zuverlässig demonstriert werden, wie Messungen an einem simulierten Gasleck mithilfe eines mit Methan gefüllten Gassacks zeigen.

Zukünftige Arbeiten sollen darauf abzielen, die Nachweisgrenze und Messstabilität weiter zu verbessern. Ein zentraler Ansatzpunkt ist die Erhöhung der eingesammelten rückgestreuten optischen Leistung, beispielsweise durch den Einsatz einer größeren Empfangslinse oder eines stärkeren Lasers. Darüber hinaus bieten alternative Auswertestrategien Potenzial zur Reduktion von Rauscheinflüssen. Zusätzlich könnten eine aktive Strahlstabilisierung, der Einsatz eines Gimbals oder verbesserte Flugregelungsstrategien die Reproduzierbarkeit der Messungen unter realen Einsatzbedingungen weiter erhöhen.

Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft (DFG) für die Förderung dieser Arbeit im Rahmen des Schwerpunktprogramms „Messtechnik auf fliegenden Plattformen“ (SPP 2433), Projekt ZeMeRGe (Projekt-nummer 541028991).

Literaturnachweis

- [1] IPCC, 2023: *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]*. IPCC, Geneva, Switzerland, (2023); doi: 10.59327/IPCC/AR6-9789291691647.
- [2] R. A. Alvarez *et al.*, *Greater focus needed on methane leakage from natural gas infrastructure*, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 17, pp. 6435–6440, (2012); doi: 10.1073/pnas.1202407109.
- [3] J. Li *et al.*, *Standoff Chemical Detection Using Laser Absorption Spectroscopy: A Review*, *Remote Sensing*, vol. 12, no. 17, p. 2771, (2020); doi: 10.3390/rs12172771.
- [4] J. T. Shaw *et al.*, *Methods for quantifying methane emissions using unmanned aerial vehicles: a review*, *Philos Trans A Math Phys Eng Sci*, vol. 379, no. 2210, p. 20200450, (2021); doi: 10.1098/rsta.2020.0450.
- [5] W. Ye *et al.*, *Mid-infrared dual-gas sensor for simultaneous detection of methane and ethane using a single continuous-wave interband cascade laser*, *Opt. Express*, OE, vol. 24, no. 15, pp. 16973–16985, (2016); doi: 10.1364/OE.24.016973.
- [6] I. E. Gordon *et al.*, *The HITRAN2024 molecular spectroscopic database*, *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 353, p. 109807, (2026); doi: 10.1016/j.jqsrt.2026.109807.
- [7] R. Arndt, *Analytical Line Shapes for Lorentzian Signals Broadened by Modulation*, *Journal of Applied Physics*, vol. 36, no. 8, pp. 2522–2524, (1965); doi: 10.1063/1.1714523.

Mikrobolometerplattform mit anpassbaren Metamaterialabsorbern zur multispektralen Bildgebung im Infrarotbereich

Alexander Litke¹, Sascha Weyers¹, Anna Lena Schall-Giesecke^{1,2,3}

¹Fraunhofer IMS, Finkenstraße 61, 47057 Duisburg

²Universität Duisburg-Essen, Forsthausweg 2, 47057 Duisburg

³Center for Nanointegration Duisburg-Essen (CENIDE), Carl-Benz-Straße 199, 47057 Duisburg

Zusammenfassung

In dieser Arbeit stellen wir einen ungekühlten Bildsensor für die multispektrale Bildgebung im mittelwelligen und langwelligen Infrarotbereich vor, der aus einem Mikrobolometerarray mit anpassbaren plasmonischen Metamaterialabsorbern (PMA) und integrierter digitaler Ausleseschaltung besteht. Der Sensor verfügt über eine Auflösung von 320×240 Pixeln. Die spektrale Empfindlichkeit der einzelnen Pixel kann durch Anpassung der Geometrie der oberen Metalllage der PMAs auf die jeweilige Anwendung zugeschnitten werden. Die zu diesem Zweck verwendeten Nanotube-Mikrobolometer des Fraunhofer IMS weisen einen Pixel-Pitch von $17 \mu\text{m}$ auf. Der vorgestellte Herstellungsprozess erfolgt auf 200-mm-CMOS-Wafern mit integrierter Ausleseschaltung. Eine mit dem Bildsensor angefertigte multispektrale Aufnahme demonstriert die erfolgreiche Umsetzung.

Keywords: Metamaterial, Mikrobolometer, multispektral, Infrarot, post-CMOS

Einführung

Die multispektrale Bildgebung kombiniert Informationen aus verschiedenen Bereichen des elektromagnetischen Spektrums. Mögliche Anwendungen im Infrarotbereich (IR) sind dabei die optische bildgebende Gasdetektion [1], die Detektion von Flammen [2] oder die Objekterkennung in sicherheitsrelevanten Szenarien [3]. Bestehende multispektrale IR-Kamerasysteme basieren entweder auf gekühlten photonischen Detektoren oder verwenden breitbandige thermische Detektoren in Kombination mit spektralen Filtertechniken auf Systemebene. Beide sind technisch komplex, kostspielig und hinsichtlich der Systemminiaturisierung begrenzt.

Eine Möglichkeit, die spektrale Selektivität auf Pixelebene in einem ungekühlten Detektor zu implementieren, ist die Integration von plasmonischen Metamaterialabsorbern (PMA) in Mikrobolometerarrays. [4, 5] Diese Absorber bestehen aus Metall-Isolator-Metall-Stapeln (MIM) mit einer strukturierten oberen Metallschicht und ermöglichen eine schmalbandige Absorption in einem einstellbaren Wellenlängenbereich im mittelwelligen (MWIR, $3 \mu\text{m}$ – $5 \mu\text{m}$) bis hin zum langwelligen Infrarotbereich (LWIR, $8 \mu\text{m}$ – $14 \mu\text{m}$). [6]

Konzept

In dieser Arbeit stellen wir eine technologische Plattform für die multispektrale Bildgebung im infraroten Wellenlängenbereich vor (Abb. 1). Die Basis bilden die Nanotube-Mikrobolometer des Fraunhofer IMS [7] mit einem Rastermaß von $17 \mu\text{m}$.

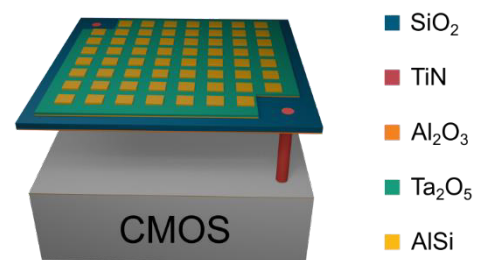


Abb. 1: Gerendertes Bild eines Nanotube-Mikrobolometers mit integriertem PMA. Seine Absorptionswellenlänge wird durch die Größe der Patches in der obersten Metallschicht bestimmt. Der bildgebende Sensor besteht aus einem Array aus 320×240 Mikrobolometern. Das Array wird auf einem CMOS-Wafer mit integrierter Ausleseschaltung hergestellt.

Sie sind ausgeführt als Focal Plane Array (FPA) bestehend aus 320×240 Mikrobolometern mit einer integrierten Ausleseschaltung. Die spektrale Empfindlichkeit der einzelnen Pixel wird durch Anpassung der Geometrie und der Materialien des Metamaterials eingestellt. Das Konzept ist in Abb. 1 schematisch dargestellt. Der Herstellungsprozess ist so konzipiert, dass Absorber mit unterschiedlichsten auf die konkrete Anwendung zugeschnittenen Eigenschaften in Hinblick auf Absorptionswellenlänge, Bandbreite oder Polarisationsabhängigkeit integriert werden können.

Mikrosystemtechnischer Herstellungsprozess

Der Herstellungsprozess wird auf 200-mm-Wafern mit bereits gefertigter CMOS-Ausleseschaltung durchgeführt. Diese werden zunächst durch chemisch-mechanisches Polieren (CMP) planarisiert. Dann werden Vias zur oberen Metalllage hergestellt, verfüllt und das überschüssige Metall erneut durch CMP entfernt. Hierauf wird eine zusätzliche Metalllage abgeschieden, strukturiert und passiviert, die als MEMS-Interface dient.

Eine schematische Darstellung des darauffolgenden vereinfachten Prozessablaufs der mikrosystemtechnischen Herstellung ist in Abb. 2 zu sehen. Zunächst wird mit reaktivem Ionenätzen (RIE) die Passivierung der oberen Metalllage geöffnet (Abb. 2a). Anschließend wird amorphes Silizium (a-Si) durch plasmaunterstützte chemische Gasphasenabscheidung (PECVD) als Opferschicht abgeschieden. Eine Passivierungsschicht aus Aluminiumoxid (Al_2O_3) und erneut a-Si als Sensorschicht werden mittels Atomlagenabscheidung (ALD) bzw. PECVD abgeschieden (Abb. 2b). Die Durchkontaktierungen respektive späteren Nanotubes werden mittels RIE und DRIE geätzt, gefolgt von der Abscheidung der Titannitrid-Kontakte (TiN) durch ALD und physikalische Gasphasenabscheidung (PVD) sowie ihrer Strukturierung (Abb. 2c). Die Sensorschicht wird anschließend mit Siliziumdioxid (SiO_2) passiviert und die Nanotubes mit a-Si überdeckt, um Rückstände innerhalb der Nanotubes aus nachfolgenden Prozessschritten zu vermeiden und so die thermische Isolation der Membran nicht zu beeinträchtigen (Abb. 2d). Danach wird der MIM-Stapel durch PVD und ALD abgeschieden und die Strukturierung der oberen Metalllage durchgeführt, wodurch die Absorptionseigenschaften der einzelnen Pixel definiert werden (Abb. 2e).

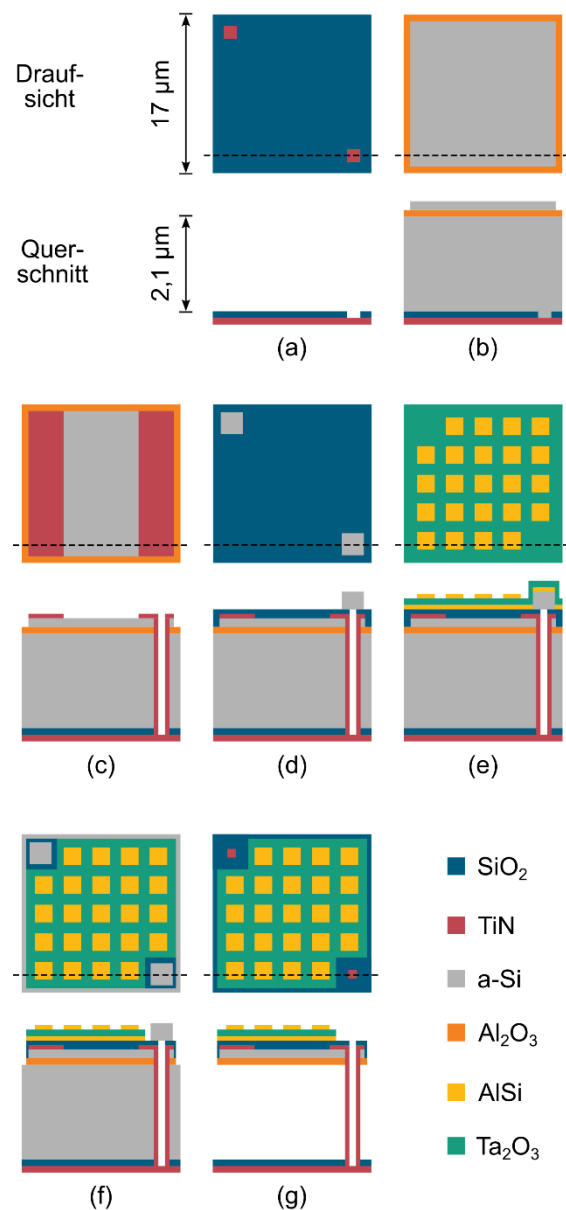


Abb. 2: Übersicht über den vereinfachten Prozessablauf. Dargestellt sind Draufsicht und Querschnitt entlang der gestrichelten Linie. (a) Freilegen der Kontakte der oberen Metallschicht des ROIC; (b) Abscheidung der Opferschicht, der Passivierung und der Sensorschicht; (c) Ätzen der Durchkontaktierungen und Herstellung der Kontakte; (d) Einkapseln der Sensormembran und Abdecken der Durchkontaktierungen; (e) Abscheidung des Absorberschichtstapels und Strukturierung seines oberen Metalls; (f) Freilegung der Abdeckungen und Trennung der Mikrobolometer; (g) Entfernung der Abdeckungen und der Opferschicht. Abb. adaptiert aus unserer Arbeit [5].

Es folgt die Öffnung der Nanotube-Abdeckungen und die anschließende Trennung der einzelnen Mikrobolometer voneinander (Abb. 2f). Schließlich werden die Opferschicht und die Abdeckungen durch isotropes Gasphasenätzen (VPE) entfernt (Abb. 2g).

Das Mikrobolometerarray ist in einem Vakuumgehäuse untergebracht, um die durch Konvektion bedingte Wärmeleitung zu minimieren. Dies wird durch das Chip-Scale-Wafer-Level-Gehäuse des Fraunhofer IMS [8] realisiert. IR-transparente Abdeckungen werden im Flip-Chip-Verfahren auf den ROIC aufgebracht. Es können Abdeckungstypen verwendet werden, die für eine hohe Durchlässigkeit im MWIR- oder LWIR-Bereich optimiert sind. Im Inneren des Gehäuses wird ein Getter-Material aufgebracht, um Restgasatome und -moleküle einzufangen und so den Innendruck des Gehäuses zu senken.

Ergebnisse der Fertigung

In Abb. 3 ist eine mittels Rasterelektronenmikroskop (REM) angefertigte Aufnahme eines planarisierten Wafers im Querschnitt zu sehen. Erkennbar sind mehrere Metalllagen der CMOS-Ausleseschaltung, einschließlich der gekennzeichneten obersten. Darüber ist in hellem Grau die planarisierte Passivierung zu sehen. Als Inset ist eine zweite Aufnahme in einem späteren Bearbeitungsstand eingefügt. Darin sind die bereits gefertigten und verfüllten Vias zu sehen, die den elektrischen Kontakt zur Ausleseschaltung herstellen. Darüber ist die als MEMS-Interface dienende zusätzliche Metalllage zu erkennen.

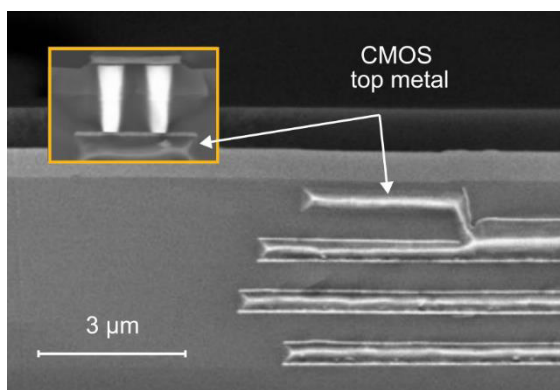


Abb. 3: REM-Aufnahmen eines planarisierten CMOS-Wafers im Querschnitt. Gekennzeichnet ist die oberste Metalllage. Im Inset ist eine zweite Aufnahme eingefügt, die die bereits gefertigten Vias sowie die zusätzliche als MEMS-Interface fungierende Metalllage zeigt.

Abb. 4 zeigt REM-Aufnahmen eines Ausschnitts des gefertigten Mikrobolometerarrays nach dem letzten Fertigungsschritt, dem Entfernen der Opferschicht (vgl. Abb. 2g). Die Membranen sind hier freistehend, jeweils auf zwei hohlen Nanotubes mit einem Durchmesser von ca. 450 nm und einer Wandstärke von ca. 10 nm. Fehlende oder verkippte Mikrobolometer sind nicht ersichtlich.

In der Vergrößerung in Abb. 4 unten lassen sich eindeutig die unterschiedlichen Geometrien der Absorber der vier Pixeltypen erkennen. Eine geringe Wölbung der Membranen ist unabhängig von der Absorbergeometrie in dieser Ansicht feststellbar.

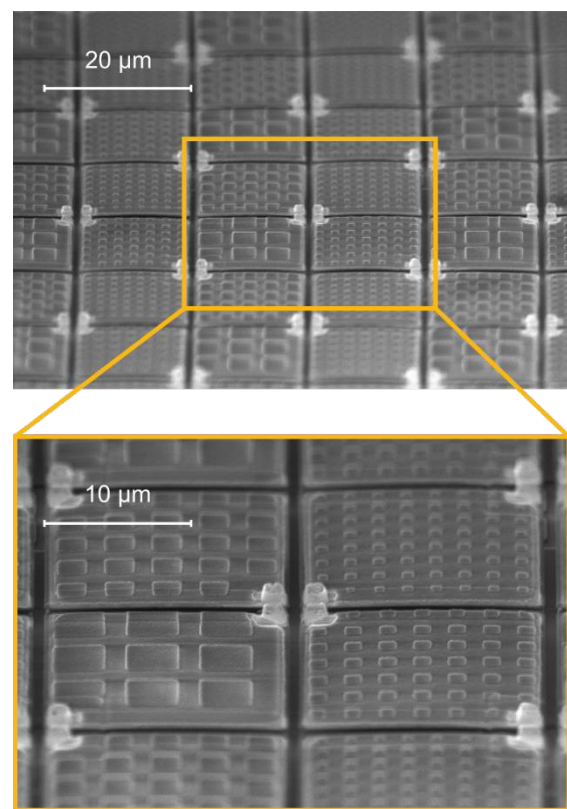


Abb. 4: REM-Aufnahmen eines Mikrobolometerarrays nach dem Entfernen der Opferschicht in der Schrägansicht. Unten sind vergrößert die unterschiedlichen Geometrien der vier für verschiedene Wellenlängen ausgelegten Absorber der Pixeltypen A, B, C und D erkennbar.

In den Ecken der einzelnen Bolometer oberhalb der Nanotubes sind unregelmäßige dünne Strukturen erkennbar. Bei diesen handelt es sich um redeponiertes Material an den Flanken der zwischenzeitlich aufgetragenen Nanotube-Verschlüsse (vgl. Abb. 2d-2f). Weil es sich dabei um Oxide handelt, ist nicht von einem

Einfluss auf die Absorptionseigenschaften der Mikrobolometer auszugehen.

Elektro-optischer Funktionstest

Die elektrische Funktionalität wird nach Abschluss des Packagings durch einen elektro-optischen Test geprüft. Hierbei wird der Bildsensor ohne weitere Optik gegenüber einem Schwarzkörperstrahler bei 20 °C positioniert. Es werden fünfzig Einzelaufnahmen angefertigt und daraus für jedes Mikrobolometer ein arithmetisches Mittel des digitalen Ausgabewerts errechnet.

Die Ergebnisse für ein Array sind in Abb. 5 gezeigt. Aufgetragen ist jeweils der digitale Ausgabewert in *least significant bits* (LSB) für jedes der 76800 Mikrobolometer im Array, gruppiert nach den vier Pixeltypen A, B, C und D. Grün gekennzeichnet ist das Intervall, in dem ein Mikrobolometer für gut befunden wird.

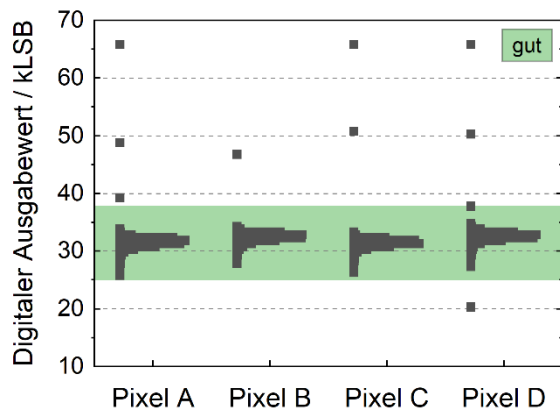


Abb. 5: Ergebnisse des elektro-optischen Funktionstests eines Bildsensors. Gezeigt ist der digitale Ausgabewert eines jeden der 76800 Mikrobolometer im Array bei Bestrahlung mit einem Schwarzkörperstrahler bei 20 °C, gruppiert nach Pixeltypen. Nur vereinzelte Mikrobolometer sind nicht gut getestet. Zwischen den Pixeltypen zeigen sich keine signifikanten Unterschiede.

Nur vereinzelte Mikrobolometer sind nicht „gut“ getestet. Ist das Mikrobolometer vollständig elektrisch defekt, beträgt die Ausgabe $2^{16} - 1 = 65535$ LSB. Dies ist für drei Mikrobolometer im gesamten Array der Fall. Zwischen den Pixeltypen zeigen sich keine signifikanten Unterschiede.

Anwendungsbeispiel

Eine naheliegende Anwendung multispektraler Bildgebung im Infrarotbereich ist die

Thermographie. Hierbei macht man sich zunutze, dass unterschiedliche Materialien unterschiedliche wellenlängenabhängige optische Eigenschaften aufweisen. Ein fusioniertes Bild aus Einzelbildern aus verschiedenen spektralen Bändern kann so mehr Information enthalten als jedes der Einzelbilder.

So weist beispielsweise Silikatglas im LWIR-Bereich einen geringen Transmissionsgrad auf, die bei Wellenlängen über 7,5 µm praktisch null erreicht.

In typischen Thermografieanwendungen sind daher Alltagsgegenstände wie Fenster, Trinkgläser oder Brillen nicht transparent [9]. Abb. 6 zeigt eine mit dem multispektralen Bildsensor angefertigte Abbildung eines ca. 300 °C heißen Lötkolbens in einem Becherglas neben einer mit heißem Wasser gefüllten Tasse. Das Bild wurde aus zwei Kanälen zusammengefügt. Grün wurde für den LWIR- und Rot für den MWIR-Kanal gewählt.

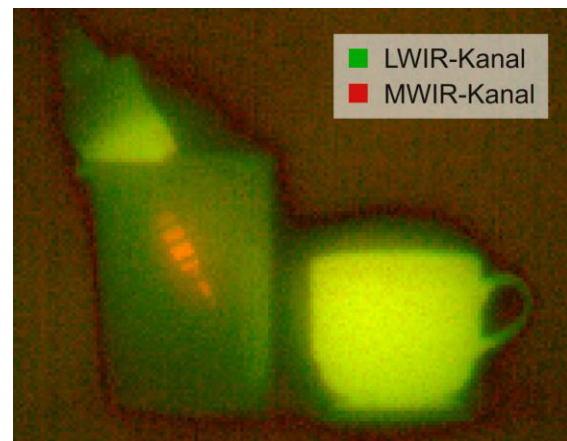


Abb. 6: Mit dem multispektralen Bildsensor angefertigte Abbildung eines heißen Lötkolbens in einem Becherglas neben einer Tasse. Das Bild wurde aus zwei Kanälen zusammengefügt. Grün wurde für den LWIR- und Rot für den MWIR-Kanal gewählt. Weil Glas im LWIR-Bereich intransparent ist, ist die Spitze des Lötkolbens nur im MWIR-Kanal sichtbar. Sie erscheint in der Abbildung daher rötlich.

Weil Glas im LWIR-Bereich intransparent ist, ist die Spitze des Lötkolbens nur im MWIR-Kanal sichtbar. Sie erscheint in der Abbildung daher rötlich. Zugleich sind die kühleren Elemente, wie etwa der Henkel der Tasse und der Griff des Lötkolbens, deutlich grün. Dies entspricht den Erwartungen für Planck'sche Strahler, die ihr Emissionsmaximum bei Raumtemperatur bei einer Wellenlänge von ca. 9,6 µm und damit im LWIR-Bereich aufweisen. Auf diese Weise

enthält das überlagerte Bild mehr Information als jedes der Einzelbilder.

Zusammenfassung

In dieser Arbeit wird ein anpassbarer, ungekühlter Bildsensor für die multispektrale Bildgebung im mittelwelligen und langwelligen Infrarotbereich vorgestellt, der aus Mikrobolometerarrays mit plasmonischen Metamaterialabsorbern und integrierter digitaler Ausleseschaltung besteht. Die Grundlage bilden die Nanotube-Mikrobolometer des Fraunhofer IMS mit einem Pixel-Pitch von 17 µm. Das Array besteht aus 320 × 240 Einzelbolometern. Dabei kann die spektrale Empfindlichkeit jedes Pixels durch Anpassung der Geometrie der oberen Metalllage der Absorber auf die jeweilige Anwendung zugeschnitten werden. Der vorgestellte Herstellungsprozess erfolgt auf 200-mm-CMOS-Wafern mit integrierter Ausleseschaltung, die dank eines Prozesses zur Herstellung eines MEMS-Interfaces beliebigen Ursprungs sein können. REM-Aufnahmen des gefertigten Arrays und ein elektro-optischer Test belegen die erfolgreiche Umsetzung der Herstellung. Schließlich demonstriert eine mit dem Bildsensor angefertigte multispektrale Aufnahme die Funktionalität und eine mögliche Anwendung des Sensors.

Zukünftige Arbeiten können sich die Anpassbarkeit der Absorber zunutze machen und weitere Anwendungen in den Fokus nehmen. Ein zentrales Thema ist dabei die bildgebende Gasdetektion, für die die spektrale Empfindlichkeit der Pixel auf Absorptionsbanden von Gasen abgestimmt wird.

Förderung

Das dieser Veröffentlichung zugrunde liegende Vorhaben wurde im Rahmen der APECS-Pilotlinie durch Chips Joint Undertaking, gefördert durch Horizon Europe (ID: 101183307) und Digital Europe (ID: 101182906) Programme, und durch nationale Förderungen von Belgien, Deutschland, Finnland, Frankreich, Griechenland, Österreich, Portugal und Spanien kofinanziert.



**Co-funded by
the European Union**

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Chips Joint

Undertaking. Neither the European Union nor the granting authority can be held responsible for them.

Literaturnachweis

- [1] Cui, F.-X., Zhao, Y., Wang, A.-J., Ma, F.-X., Wu, J., Li, Y.-Y., Li, D.-C. and Dong, W.-C., Optics 3(4), 473–482 (2022). DOI: <https://dx.doi.org/10.3390/opt3040040>
- [2] Wing Au, K., Larsen, C., Cole, B. and Venkatesha, S., Proc. SPIE 9836, 983605 (2016). DOI: <https://dx.doi.org/10.1117/12.2224295>
- [3] Karali, A. O., Cakir, S. and Aytac, T., Appl. Opt. 54(13), 4172 (2015). DOI: <https://dx.doi.org/10.1364/AO.54.004172>
- [4] Maier, T. and Brückl, H., Opt. Lett. 34(19), 3012–3014 (2009). DOI: <https://dx.doi.org/10.1364/OL.34.003012>
- [5] Litke, A., Zakizade, E., Weyers, S., Schall-Giesecke, A. L., Opt. Lett. 50, 5767–5770 (2025). DOI: <https://dx.doi.org/10.1364/OL.573586>
- [6] Litke, A., Zakizade, E., Michel, M., Weyers, S. and Schall-Giesecke, A. L., Micro Nano Eng. 23(1), 100262 (2024). DOI: <https://dx.doi.org/10.1016/j.mne.2024.100262>
- [7] Michel, M., Blaeser, S., Litke, A., Zakizade, E., Weyers, S. and Weiler, D., Proc. SPIE 12534, 1253419 (2023). DOI: <https://dx.doi.org/10.1117/12.2663466>
- [8] Bauer, J., Weiler, D., Ruß, M., Heß, J., Yang, P., Voß, J., Arnold, N. and Vogt, H., Proc. SPIE 7834, 78340S (2010). DOI: <https://dx.doi.org/10.1117/12.865019>
- [9] Lane, D. W., J. Phys. D: Appl. Phys. 23(12), 1727–1734 (1990). DOI: <https://dx.doi.org/10.1088/0022-3727/23/12/037>

Anwendung konfokaler Abstandssensoren zur Spaltemessung in einem aktiven Spalthaltemechanismus in einem elektrischen Fantriebwerk

Niels Herter¹, Andreas Hupfer¹

¹ Institut für Aeronautical Engineering, Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85579 Neubiberg

Zusammenfassung

Die präzise Erfassung des Schaufelspitzenspalt stellt eine wesentliche Voraussetzung für die Bewertung und Optimierung des Betriebsverhaltens von Fantriebwerken dar. Insbesondere elektrisch angetriebene Antriebssysteme mit hohen Umfangsgeschwindigkeiten und geringen Spaltweiten erfordern Messverfahren, die eine hochaufgelöste und berührungslose Abstandserfassung auch unter beengten Einbaubedingungen ermöglichen. Ziel der vorliegenden Arbeit ist die experimentelle Bewertung konfokal-chromatischer Abstandssensoren zur schaufelindividuellen Erfassung des Schaufelspitzenspalt unter realitätsnahen Betriebsbedingungen.

Hierzu wurden Untersuchungen an elektrisch angetriebenen Fantriebwerken mit Rotoren aus kohlefaserverstärktem Kunststoff (CFK) sowie integralen Aluminiumblisks durchgeführt. Neben der Messstabilität wurden insbesondere die Einflüsse hoher Drehzahlen, unterschiedlicher Oberflächeneigenschaften und begrenzter Einbauräume auf die Signalqualität untersucht. Die Ergebnisse zeigen, dass eine reproduzierbare und schaufelindividuelle Abstandsmessung auch bei Drehzahlen bis $24\,000\text{ min}^{-1}$ möglich ist. Gleichzeitig wird deutlich, dass die optischen Eigenschaften der Messoberfläche einen wesentlichen Einfluss auf die Qualität der Messsignale besitzen.

Die Untersuchungen bestätigen die grundsätzliche Eignung konfokal-chromatischer Abstandssensoren für die hochaufgelöste Schaufelspaltemessung in elektrisch angetriebenen Fantriebwerken. Die erzielten Ergebnisse bilden zugleich die messtechnische Grundlage für die Integration der Sensorik in ein aktives Spalthaltemechanismus auf Basis auxetischer Strukturen.

Keywords: konfokal-chromatische Abstandssensorik, Schaufelspitzenspalt, aktive Spaltemessung, elektrisches Fantriebwerk, auxetische Strukturen.

1 Einleitung und Motivation

Der Schaufelspitzenspalt zwischen Rotor und Gehäuse zählt zu den wesentlichen Einflussgrößen für den Wirkungsgrad axialer Turbomaschinen. Bereits geringe Änderungen der Spaltweite führen zu erhöhten Leckageströmungen und damit zu aerodynamischen Verlusten. Besonders betroffen sind Fan-, Verdichter- und Turbinenstufen moderner Flugtriebwerke, bei denen hohe Umfangsgeschwindigkeiten und geringe Spaltweiten unmittelbar die Leistungsfähigkeit beeinflussen [1], [2].

Die im Betrieb auftretenden Spaltänderungen werden durch zahlreiche Einflüsse verursacht. Hierzu zählen unter anderem Fertigungs- und Montageabweichungen, thermische Dehnungen, Fliehkräfte, aerodynamische Belastungen sowie dynamische Rotorbewegungen. Abhängig von den jeweiligen Betriebsbedingungen entstehen unterschiedliche Formen der Spaltänderung, wie sie

beispielhaft in Abb. 1 dargestellt sind. Die Auslegung eines möglichst kleinen Montagepalt stellt daher eine anspruchsvolle Optimierungsaufgabe dar, bei der sowohl Fertigungstoleranzen als auch betriebsbedingte Verformungen berücksichtigt werden müssen [3].

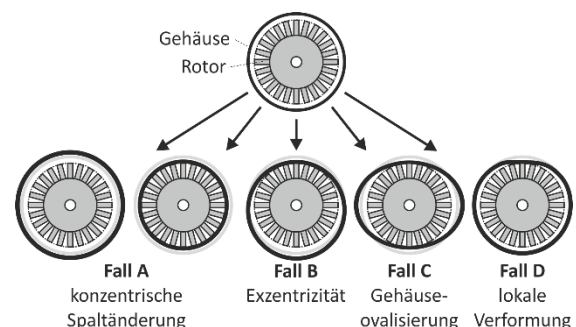


Abb. 1: Ausprägungen von Spaltänderungen [3]

Mit der Entwicklung elektrisch angetriebener Fantriebwerke gewinnen minimale Schaufel-spitzenspalte zusätzlich an Bedeutung. Die kompakte Bauweise elektrischer Antriebssysteme sowie hohe Drehzahlen und dynamische Lastwechsel führen zu erhöhten Anforderungen an die Rotor-Gehäuse-Interaktion. Zur Ausschöpfung des aerodynamischen Potenzials werden daher zunehmend aktive Spalthaltungskonzepte untersucht, die eine kontinuierliche Anpassung des Gehäusedurchmessers an den jeweiligen Betriebszustand ermöglichen.

Neue konstruktive Möglichkeiten ergeben sich hierbei durch additive Fertigungsverfahren, mit denen sich unter anderem auxetische Strukturen herstellen lassen. Aufgrund ihrer negativen Querkontraktionszahl können diese Strukturen gezielt zur Durchmesseränderung ringförmiger Bauteile eingesetzt werden und eignen sich damit für innovative Konzepte einer aktiven Spalthaltung. Ein entsprechender Ansatz für elektrisch angetriebene Fantriebwerke wurde bereits patentiert [4], siehe Abb. 2.

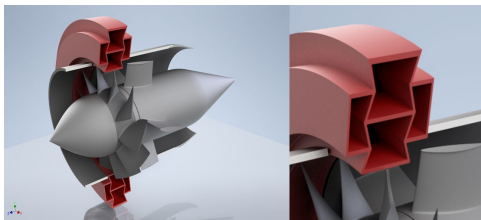


Abb. 2: Patentiertes Konzept einer aktiven Spalthaltung für elektrisch angetriebene Fantriebwerke [4]

Für die Realisierung einer aktiven Spalthaltung ist neben der Verstellmechanik insbesondere die kontinuierliche Erfassung des aktuellen Schaufelspitzenspalts erforderlich. Hierfür wird ein Messverfahren benötigt, das auch unter hohen Drehzahlen, beengten Einbaubedingungen und im elektromagnetischen Umfeld elektrischer Antriebssysteme zuverlässig arbeitet. Ziel der vorliegenden Arbeit ist daher die experimentelle Untersuchung konfokal-chromatischer Abstandssensoren zur schaufel-individuellen Schaufelspaltemessung in elektrisch angetriebenen Fantriebwerken sowie die Bewertung ihrer Eignung als Grundlage zukünftiger aktiver Spalthaltungssysteme.

2. Konfokal-chromatisches Messprinzip

Aus den beschriebenen Anforderungen ergibt sich die Notwendigkeit eines Messverfahrens, das kleine Schaufelspitzenspalte berührungslos, hochaufgelöst und unter beengten Einbaubedingungen erfassen kann. Bildgebende Verfahren ermöglichen zwar eine qualitative

Beurteilung des Spaltverhaltens, stoßen jedoch bei einer präzisen Quantifizierung unter hochdynamischen Betriebsbedingungen an Grenzen. Ursächlich hierfür sind unter anderem eingeschränkte Sichtbedingungen im Strömungskanal, optische Verzerrungen, hohe Anforderungen an die Fokussierung sowie der vergleichsweise hohe Auswerteaufwand. Darüber hinaus ist die Integration kamera-basierter Systeme häufig nur mit erheblichem konstruktivem Aufwand möglich und kann die Strömungsführung beeinflussen.

Im Rahmen von Voruntersuchungen wurden außerdem kapazitive Messverfahren betrachtet. Für den Einsatz in elektrisch angetriebenen Fantriebwerken erwiesen sich diese aufgrund elektromagnetischer Einflüsse des Antriebssystems und der Leistungselektronik jedoch als nur eingeschränkt geeignet. Vor diesem Hintergrund wurde die konfokal-chromatische Abstandssensorik als Messverfahren ausgewählt.

Das Messprinzip basiert auf polychromatischem Licht, das durch ein optisches System mit definierter chromatischer Aberration entlang der Messachse fokussiert wird. Jeder Wellenlänge ist dabei eine eindeutige Fokusslage zugeordnet. Liegt die Oberfläche des Messobjekts in einer dieser Fokusslagen, wird die entsprechende Wellenlänge bevorzugt reflektiert und spektral ausgewertet. Aus der Lage des reflektierten Spektrums wird der Abstand zwischen Sensor und Messoberfläche bestimmt.

Konfokal-chromatische Abstandssensoren zeichnen sich durch eine hohe Ortsauflösung, eine kompakte Bauform sowie eine hohe Unempfindlichkeit gegenüber elektromagnetischen Störfeldern aus und eignen sich daher besonders für den Einsatz in elektrisch angetriebenen Fantriebwerken. Die Qualität der Messsignale hängt jedoch wesentlich von den optischen Eigenschaften der Messoberfläche ab. Insbesondere Reflektivität, Oberflächenrauheit und lokale Inhomogenitäten beeinflussen die Intensität und Stabilität des detektierten Signals [5]. Diese Randbedingungen sind bei der Sensorintegration und der Auslegung des Messaufbaus zu berücksichtigen.

3. Messaufgabe und Versuchsaufbau

Zur experimentellen Bewertung konfokal-chromatischer Abstandssensoren wurden Untersuchungen an mehreren elektrisch angetriebenen Fantriebwerken im Rahmen des Forschungsprojekts ELAPSED (Electrified Aircraft Propulsion – safe, efficient, digitally linked) am Zentrum für Digitalisierungs- und Technologieforschung der Bundeswehr

(dtec.bw) durchgeführt. Ziel der Messungen ist die schaufelindividuelle Bestimmung des Schaufelspitzenspalt beziehungsweise der radialen Längenänderung der Rotorschaukeln in Abhängigkeit von der Drehzahl. Die Ergebnisse dienen sowohl der Validierung der zugrunde liegenden Auslegungs- und Entwurfsmethodik [6] als auch der Bewertung der Eignung konfokal-chromatischer Abstandssensoren für den Einsatz in elektrisch angetriebenen Fantriebwerken.

Die Untersuchungen wurden an drei unterschiedlichen Versuchsträgern durchgeführt, um sowohl verschiedene Werkstoffe als auch unterschiedliche geometrische und betriebliche Randbedingungen abzudecken.

Der erste Versuchsaufbau besteht aus einer konventionellen Rotor-Stator-Stufe mit Rotorschaukeln aus kohlefaserverstärktem Kunststoff (CFK). Aufgrund des anisotropen Werkstoffverhaltens und des gewählten Laminataufbaus weisen diese Schaufeln eine drehzahlabhängige radiale Längenänderung auf [7]. Die experimentelle Bestimmung dieser Verformungen ist von besonderer Bedeutung, da sie den Schaufelspitzenspalt und damit die aerodynamische Leistungsfähigkeit unmittelbar beeinflussen.

Im weiteren Projektverlauf wurde der CFK-Rotor durch einen Rotor in integraler Aluminiumbliskbauweise ersetzt. Die wesentlichen Betriebsparameter der Rotor-Stator-Stufe blieben unverändert, sodass ein direkter Vergleich beider Rotorvarianten unter weitgehend identischen Randbedingungen möglich war. Neben den unterschiedlichen strukturellen Eigenschaften unterscheiden sich beide Rotoren insbesondere hinsichtlich ihrer Oberflächencharakteristik, wodurch sich gleichzeitig der Einfluss der Messoberfläche auf die konfokale Abstandsmessung untersuchen lässt.

Als dritte Versuchskonfiguration wurde ein gegenläufig rotierendes elektrisch angetriebenes Fantriebwerk aufgebaut. Dieses System besitzt zwei integrale Aluminiumblisks mit 11 beziehungsweise 13 Schaufeln und erreicht maximale Drehzahlen von bis zu $24\,000\text{ min}^{-1}$. Dadurch ergeben sich gegenüber den vorhergehenden Versuchsaufbauten deutlich höhere Anforderungen an die zeitliche Auflösung und die Stabilität der Abstandsmessung.

Für sämtliche Untersuchungen wurde der konfokal-chromatische Abstandssensor confocalDT IFS 2403-1,5 der Firma Micro-Epsilon verwendet. Die wesentlichen technischen Kenndaten des Sensors sind in Tab. 1 zusammengefasst.

Tab. 1: Technische Daten des verwendeten Abstandssensors [5]

Messbereich	1,5 mm
Messbereichsanfang	8 mm
Lichtpunktdurchmesser	15 μm
statische Auflösung	<30 nm
dynamische Auflösung	<186 nm
Außendurchmesser	8 mm

Die Integration des Sensors erfolgte jeweils im Fangehäuse, sodass der radiale Abstand zwischen Schaufelspitze und Gehäuseinnenwand direkt erfasst werden kann (Abb. 3 und Abb. 4).

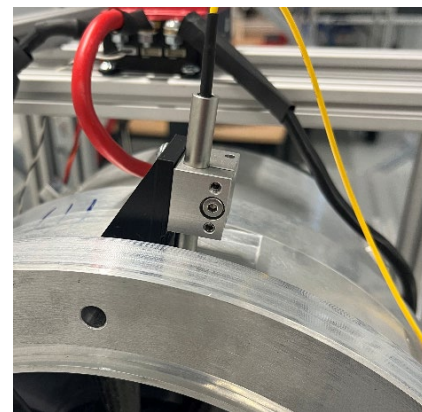


Abb. 3: Fangehäuse mit integriertem Abstandssensor IFS 2403-1,5



Abb. 4: Innenseite Fangehäuse

Aufgrund des begrenzten Bauraums im Strömungskanal war eine kompakte Sensorbauform erforderlich. Gleichzeitig mussten Einbauposition und Messrichtung so gewählt werden, dass eine zuverlässige Erfassung sämtlicher Schaufelpassagen während des Betriebs gewährleistet ist.

Die wesentlichen Anforderungen an das Messsystem ergeben sich aus den hohen Rotordrehzahlen, der schaufelindividuellen Erfassung der Messsignale sowie den begrenzten Einbaumöglichkeiten im Triebwerk. Im Gegensatz zu thermisch hochbelasteten Gasturbinen spielt eine temperaturbedingte Beanspruchung des Sensors im vorliegenden

Anwendungsfall eine untergeordnete Rolle. Entscheidend sind vielmehr die erreichbare zeitliche Auflösung und eine stabile Signaldetektion während des laufenden Betriebs.

Der verwendete Controller ermöglicht Messfrequenzen von bis zu 30 kHz. Neben der Messfrequenz stellt insbesondere die Belichtungszeit einen wesentlichen Parameter dar, da sie die Intensität und Qualität des detektierten Lichtsignals bestimmt. Aufgrund einer festen internen Signalverarbeitungszeit von $3,5\text{ }\mu\text{s}$ verkürzt sich mit steigender Messfrequenz die maximal verfügbare Belichtungszeit. Für alle Messungen wurden Messfrequenz und Belichtungszeit so gewählt, dass eine möglichst hohe Signalqualität bei gleichzeitig ausreichender zeitlicher Auflösung erreicht wird.

Eine weitere wesentliche Einflussgröße stellt die Oberflächenbeschaffenheit der Schaufelspitzen dar. Da die konfokale Abstandsmessung auf der Reflexion des Lichtsignals basiert, wurden die Oberkanten der CFK-Schaufeln mit einer reflektierenden Beschichtung versehen (Abb. 5).



Abb. 5: beschichtete CFK-Schaufeloberkanten

Hierdurch konnte die Signalintensität deutlich erhöht und die Stabilität der Messungen verbessert werden. Die metallischen Oberflächen der Aluminiumblisks bieten demgegenüber bereits günstige Voraussetzungen für eine zuverlässige Abstandsmessung.

Die Messdatenerfassung erfolgte schaufelaufgelöst über mehrere Sekunden Messdauer. Dadurch können sowohl die drehzahlabhängige radiale Längenänderung als auch schaufelindividuelle Unterschiede im Verformungsverhalten der Rotoren bestimmt werden. Gleichzeitig ermöglicht der Versuchsaufbau die Bewertung des Messverfahrens unter unterschiedlichen werkstofftechnischen, geometrischen und betrieblichen Randbedingungen.

4 Experimentelle Ergebnisse

Ziel der experimentellen Untersuchungen ist die Bewertung der Eignung konfokal-chromatischer Abstandssensoren zur schaufelindividuellen Erfassung des Schaufelspitzenspalts unter unterschiedlichen werkstofftechnischen und betrieblichen Randbedingungen. Im Mittelpunkt stehen dabei die Reproduzierbarkeit der Messungen, die erreichbare zeitliche Auflösung sowie der Einfluss der Oberflächenbeschaffenheit auf die Signalqualität.

Aluminiumblisk

Zur Beurteilung der Messstabilität wurden zunächst Referenzmessungen an einer einzelnen Schaufel einer Aluminiumblisk bei einer Drehzahl von etwa 30 min^{-1} durchgeführt. Untersucht wurden insbesondere die Einflüsse von Messfrequenz und Belichtungszeit auf die Reproduzierbarkeit der Messwerte.

Die ermittelten Abweichungen vom jeweiligen Mittelwert liegen bei etwa $5\text{ }\mu\text{m}$ und belegen eine stabile und reproduzierbare Abstandserfassung unter stationären Bedingungen. Gleichzeitig zeigen die Untersuchungen, dass sich durch eine geeignete Wahl von Messfrequenz und Belichtungszeit eine hohe Signalqualität erreichen lässt.

Aufbauend auf den Referenzmessungen wurden alle elf Schaufeln der Aluminiumblisk über zehn Umdrehungen vermessen. Die gemittelten Ergebnisse sind in Abb. 6 dargestellt.

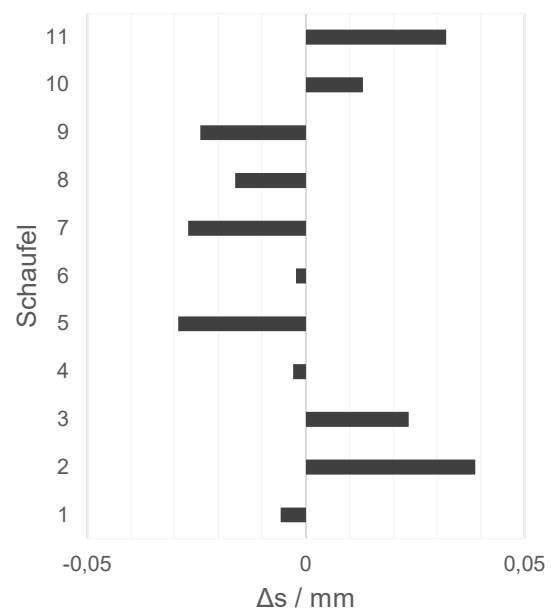


Abb. 6: Schaufellängenabweichung im Vergleich zum Mittelwert

Mit zunehmender Drehzahl nimmt die radiale Längenänderung aller Schaufeln kontinuierlich zu. Dieses Verhalten entspricht der erwarteten strukturellen Reaktion auf die zunehmende Zentrifugalbelastung. Die gemessenen maximalen radialen Längenänderungen liegen bei etwa 0,08 mm.

Im oberen Drehzahlbereich ist eine zunehmende Streuung zwischen den einzelnen Schaufeln erkennbar. Als Ursache kommen geometrische Unterschiede, fertigungsbedingte Toleranzen oder schwingungsbedingte Einflüsse in Betracht. Abb. 7 zeigt die Längenabweichung der einzelnen Schaufeln vom Mittelwert bei 80 % der maximalen Rotordrehzahl.

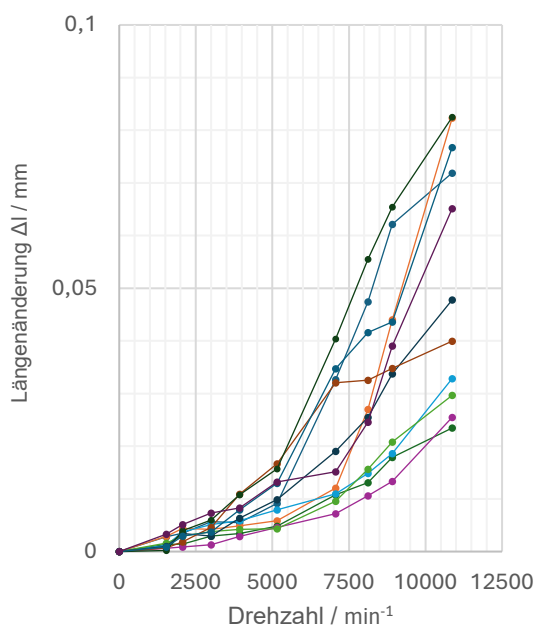


Abb. 7: Radiale Längenänderung der 11 Aluminiumschaufeln in Abhängigkeit der Drehzahl

Insgesamt zeigt die Aluminiumblisk jedoch ein weitgehend homogenes Verformungsverhalten mit vergleichbaren Dehnungsverläufen aller Schaufeln.

CFK-Rotor

Für die CFK-Schaufeln besitzt die experimentelle Bestimmung der radialen Längenänderung besondere Bedeutung, da deren anisotroper Laminataufbau ein deutlich komplexeres Verformungsverhalten erwarten lässt. Erste Untersuchungen erfolgten mithilfe einer High-Speed-Kamera, siehe Abb. 8. Aufgrund optischer Verzerrungen und der eingeschränkten Quantifizierbarkeit wurde dieses Verfahren anschließend durch die konfokal-chromatische Abstandsmessung ersetzt.

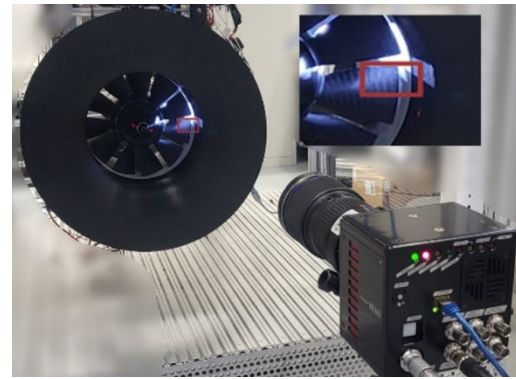


Abb. 8: Experimenteller Aufbau zur Spaltmessung mittels High-Speed Kamera

Zur Verbesserung der Signalqualität wurden die Schaufeloberkanten mit einer reflektierenden Beschichtung versehen. Dadurch konnten Messungen bis zu etwa 85 % der maximalen Drehzahl von $14\,500\text{ min}^{-1}$ durchgeführt werden. Abb. 9 zeigt die Ergebnisse beider Messverfahren.

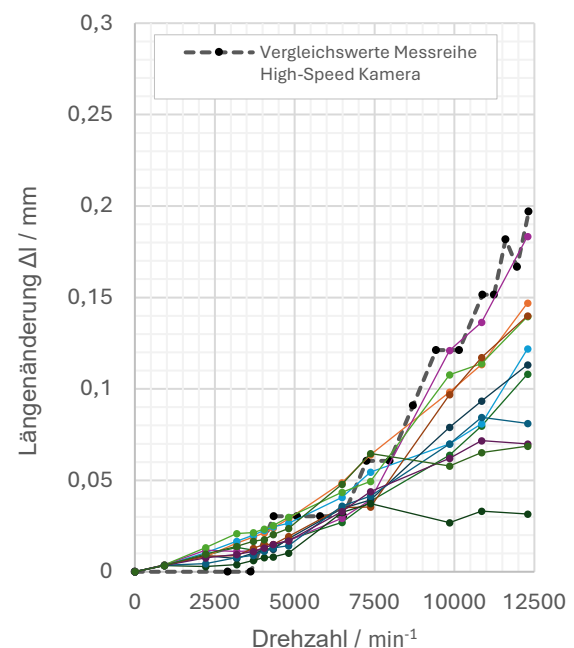


Abb. 9: Radiale Längenänderung der 11 CFK-Schaufeln in Abhängigkeit der Drehzahl

Im Gegensatz zur Aluminiumblisk weisen die CFK-Schaufeln deutlich größere Unterschiede in der radialen Längenänderung auf. Dieses Verhalten stimmt mit den in [7] beschriebenen Untersuchungen überein und ist auf fertigungsbedingte Streuungen infolge des manuellen Laminierprozesses sowie Unterschiede im Laminataufbau und in der Faserorientierung zurückzuführen.

Die Ergebnisse zeigen gleichzeitig, dass auch bei CFK-Schaufeln eine schaufelindividuelle Erfassung der radialen Längenänderung möglich ist. Voraussetzung hierfür ist jedoch eine ausreichend reflektierende Messoberfläche.

Gegenläufiges Fantriebwerk

Zur Untersuchung der Anwendbarkeit unter erweiterten Betriebsbedingungen wurde das Sensorsystem anschließend an einem gegenläufig rotierenden Fantriebwerk mit zwei Aluminiumblisks eingesetzt, siehe Abb. 12. Im Vergleich zu den vorherigen Untersuchungen weist dieses System kleinere Schaufelhöhen sowie Drehzahlen bis $24\,000\text{ min}^{-1}$ auf.

Die Messergebnisse für beide Rotoren sind in Abb. 10 und Abb. 11 dargestellt. Mit zunehmender Drehzahl steigt die radiale Längenänderung aller Schaufeln kontinuierlich an. Die gemessenen maximalen Werte liegen bei etwa $0,1\text{ mm}$.

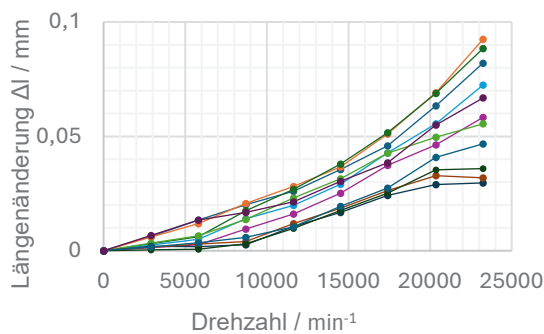


Abb. 10: Radiale Längenänderung der 11 Aluminiumschaufeln in Abhängigkeit der Drehzahl

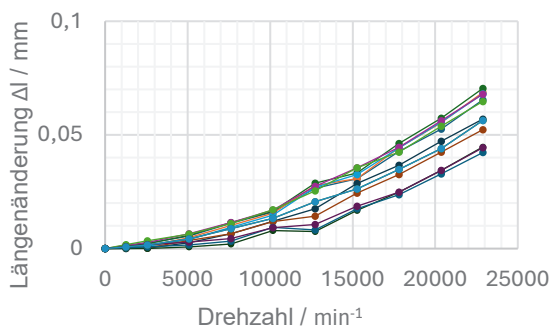


Abb. 11: Radiale Längenänderung der 13 Aluminiumschaufeln in Abhängigkeit der Drehzahl

Während der vordere Rotor eine erkennbare Streuung zwischen den einzelnen Schaufeln aufweist, zeigt der hintere Rotor ein homogeneres Verformungsverhalten. Eine mögliche Ursache hierfür ist die kompaktere

Schaufelgeometrie, die eine gleichmäßigere Lastverteilung begünstigt.

Trotz der erhöhten Drehzahlen und der größeren Anzahl zu erfassender Schaufeln konnten sämtliche Schaufelpassagen eindeutig detektiert und die gemessenen Spaltweiten den jeweiligen Schaufeln zuverlässig zugeordnet werden.

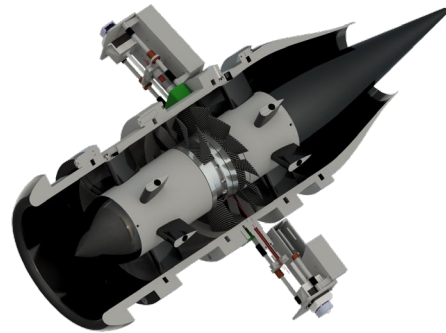


Abb. 12: gegenläufiges Fantriebwerk

Bewertung des Messverfahrens

Die experimentellen Untersuchungen bestätigen die grundsätzliche Eignung konfokal-chromatischer Abstandssensoren für die hochaufgelöste Schaufelspaltmessung in elektrisch angetriebenen Fantriebwerken.

Für Aluminiumrotoren konnte eine reproduzierbare Abstandserfassung mit homogener Signalqualität nachgewiesen werden. Die Untersuchungen an CFK-Schaufeln zeigen darüber hinaus, dass auch Faserverbundwerkstoffe zuverlässig vermessen werden können, sofern die optischen Eigenschaften der Messoberfläche berücksichtigt werden.

Besonders hervorzuheben ist die zuverlässige schaufelindividuelle Erfassung der Messsignale auch bei Drehzahlen bis $24\,000\text{ min}^{-1}$. Dadurch lassen sich nicht nur globale Änderungen des Schaufelspitzenspalts, sondern auch Unterschiede im Verformungsverhalten einzelner Schaufeln analysieren. Die Ergebnisse belegen somit, dass konfokal-chromatische Abstandssensoren die messtechnischen Anforderungen zukünftiger elektrisch angetriebener Fantriebwerke erfüllen und eine geeignete Grundlage für aktive Spalthaltungssysteme darstellen.

5 Integration der Abstandssensorik in ein aktives Spalthaltungskonzept

Die experimentellen Untersuchungen zeigen, dass konfokal-chromatische Abstandssensoren eine schaufelindividuelle und reproduzierbare Erfassung des Schaufelspitzenspalts auch unter den Betriebsbedingungen elektrisch angetriebener Fantriebwerke ermöglichen. Damit ist eine wesentliche Voraussetzung für

die Realisierung aktiver Spalthaltungssysteme geschaffen.

Im Rahmen des Forschungsprojekts ELAPSED wird hierzu ein Spalthaltungskonzept auf Basis auxetischer Strukturen verfolgt. Das zugrunde liegende Prinzip ist in Abb. 13 dargestellt.

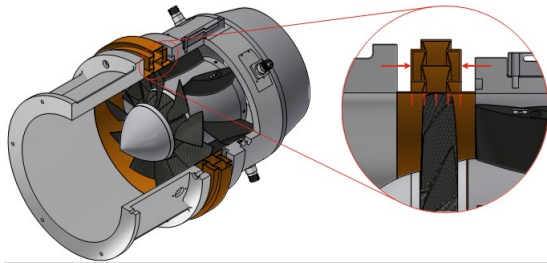


Abb. 13: Patentiertes Konzept einer aktiven Spalthaltung auf Basis auxetischer Strukturen [4]

Die Besonderheit des bereits patentierten Konzepts besteht darin, dass durch eine axiale Belastung einer auxetischen Ringstruktur eine gezielte radiale Durchmesseränderung des Gehäuses erzeugt werden kann [4]. Hierdurch lässt sich der Schaufelspitzenpalt aktiv an unterschiedliche Betriebszustände anpassen. Für die praktische Umsetzung eines solchen Systems ist neben der Verstellmechanik insbesondere die kontinuierliche Erfassung des aktuellen Schaufelspitzenpalts erforderlich. Aus diesem Grund ist vorgesehen, den konfokal-chromatischen Abstandssensor direkt in die verstellbare Gehäusestruktur zu integrieren.

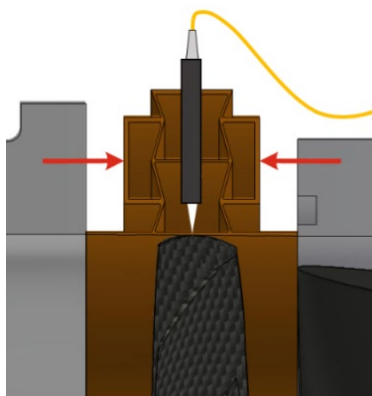


Abb. 14: Prinzipielle Integration des Abstandssensors in die auxetische Struktur

Die prinzipielle Sensorintegration ist in Abb. 14 dargestellt. Durch die direkte Einbindung des Sensors in die bewegte Struktur bleibt dessen relative Position zur Gehäuseinnenwand während der Verstellung erhalten. Dadurch

können zusätzliche Messfehler infolge der Strukturbewegung minimiert werden und gleichzeitig die Voraussetzungen für eine kontinuierliche Erfassung des Schaufelspitzenpalts geschaffen werden.

Die im Rahmen dieser Arbeit nachgewiesene schaufelindividuelle Abstandserfassung bildet damit die messtechnische Grundlage für die Entwicklung eines geschlossenen Regelkreises zur aktiven Spalthaltung.

6 Schlussfolgerungen

Ziel der vorliegenden Arbeit war die experimentelle Bewertung konfokal-chromatischer Abstandssensoren zur schaufelindividuellen Erfassung des Schaufelspitzenpalts in elektrisch angetriebenen Fantriebwerken.

Hierzu wurden Untersuchungen an Rotoren aus kohlefaserverstärktem Kunststoff, an integralen Aluminiumblisks sowie an einem gegenläufigen elektrisch angetriebenen Fantriebwerk durchgeführt. Die Messungen erfolgten unter unterschiedlichen werkstofftechnischen und betrieblichen Randbedingungen bis zu maximalen Drehzahlen von $24\,000\text{ min}^{-1}$.

Die Ergebnisse zeigen, dass eine reproduzierbare und hochaufgelöste Abstandsmessung auch unter diesen Betriebsbedingungen möglich ist. Darüber hinaus konnte nachgewiesen werden, dass eine schaufelindividuelle Erfassung sämtlicher Rotorschaukeln selbst bei hohen Drehzahlen zuverlässig möglich ist. Für Aluminiumblisks wurde ein weitgehend homogenes Verformungsverhalten beobachtet, während CFK-Schaukeln aufgrund ihres Laminataufbaus größere schaufelindividuelle Unterschiede aufweisen. Gleichzeitig konnte gezeigt werden, dass die Signalqualität wesentlich von den optischen Eigenschaften der Messoberfläche beeinflusst wird.

Insgesamt bestätigen die Untersuchungen die grundsätzliche Eignung konfokal-chromatischer Abstandssensoren für die hochaufgelöste Schaufelspaltmessung in elektrisch angetriebenen Fantriebwerken. Die Sensorik ermöglicht sowohl die Analyse globaler Änderungen des Schaufelspitzenpalts als auch die Untersuchung schaufelindividueller Verformungen und stellt damit eine geeignete Grundlage für zukünftige aktive Spalthaltungssysteme dar. Zukünftige Arbeiten konzentrieren sich auf die experimentelle Integration der Sensorik in auxetische Gehäusestrukturen sowie auf die Entwicklung eines geschlossenen Regelkreises zur aktiven Anpassung des Schaufelspitzenpalts an unterschiedliche Betriebszustände.

Literaturnachweis

- [1] J. W. Bisset, D.C. Howe, *Energy Efficient Engine Flight Propulsion System Preliminary Analysis and Design Report*, Final Update Report, NASA-CR-174701, Lewis Research Center, Cleveland, Ohio, 1983.
- [2] S. B. Lattime, , B. M. Steinetz, *Turbine Engine Clearance Control Systems: Current Practices and Future Directions*, AIAA-2002-3790, NASA/TM-2002-211794, 38th Joint Propulsion Conference and Exhibit, July 7–10, 2002.
- [3] A. Hupfer, *Spalte in Fluggasturbinen – Ursachen, Folgen und Optimierungsstrategien am Beispiel des Turboverdichters*, Habilitationsschrift, Verlag Dr. Hut, München, 2019.
- [4] A. Hupfer, N. Herter, *Element und Verfahren zum Ausgleichen von Toleranzen und/oder Spaltweiten einer Vorrichtung und Triebwerk, Montageverfahren für eine Welle-Nabe-Verbindung*, DE 10 2021 126 424.3.
- [5] Micro-Epsilon, *confocalDT Produktkatalog*, verfügbar unter: <https://www.micro-epsilon.de/fileadmin/download/products/cat--confocalDT--de.pdf>, zuletzt aufgerufen am 16.05.2026.
- [6] S. Hawner, A. Hupfer, *Entwicklung eines CFD-Auslegungstools für ein gegenläufiges, elektrisches Fantriebwerk und Anwendung für eine Parameterstudie im Designpunkt*, Deutsche Gesellschaft für Luft- und Raumfahrt - Lilienthal-Oberth e.V., Deutscher Luft- und Raumfahrtkongress, Dresden, 2022
- [7] S. Hawner, L. Rduch, H. Baumhöfener, A. Hupfer, *Development of Carbon Composite Blades within the Context of the Experimental Validation of a CFD-Based Design Tool for Contra-Rotating, Electric Fan Engines*, Aerospace 11, 527 (2024); doi: 10.3390/aerospace11070527.

Enhancing Radar-Based Pedestrian Detection with Artificial Intelligence for Multipath Mitigation

Ingo Hoyer¹, Sebastian Böller¹, Karsten Seidl^{1,2}

¹*Fraunhofer IMS, 47057 Duisburg, Germany*

²*Department of Electronic Components and Circuits, University of Duisburg-Essen*

Abstract

Pedestrian safety in automated road traffic requires detection systems that function under complex environmental conditions. Camera-based systems may fail in fog, darkness, or occlusion, making radar an attractive alternative. This work investigates passive radar transponders integrated into pedestrian clothing or footwear, detected by a vehicle-mounted radar, to enable robust pedestrian localization. A key challenge is multipath propagation, which distorts the backscattered signal and degrades distance and angle estimation accuracy. To address this, an AI-based approach is proposed: a convolutional neural network (CNN) for distance estimation and a fully connected feedforward neural network (FC-FNN) for angle estimation. Both models are trained and evaluated on synthetically generated radar datasets covering distances up to 10 m and angles up to 90°, with scenarios ranging from a single reflection up to 50 multipath reflections, including a realistic case with a random number of reflections and added noise at 23 dB SNR. The AI-based approach is benchmarked against conventional signal processing algorithms. In the scenario with random number of reflections, distance and noise, the proposed models achieve a mean absolute error (MAE) of 0.29 m for distance and 7.50° for angle estimation, compared to 0.40 m and 8.19° for the conventional method—improvements of approximately 28 % and 9 %, respectively. In most tested multipath conditions, the AI-based approach outperforms the conventional baseline. Future work will focus on real radar measurement validation, hyperparameter optimization, and extension to higher distances and moving objects.

Keywords: Radar, AI, Multipath Mitigation, Signal Processing, Pedestrian Detection.

Introduction

In a future with fully automated road traffic, protecting pedestrians is crucial, as they are the most vulnerable traffic participants. Today, low visibility increases the risk of pedestrians being involved in a crash [1]. Detection and localization systems must therefore operate reliably under all weather and visibility conditions, including scenarios in which conventional camera-based artificial intelligence (AI) systems may fail due to fog, darkness, or optical obstructions. Radar is an attractive technology in this context because its sensing capabilities remain robust even under challenging environmental influences. A promising approach to enhance pedestrian safety involves integrating passive radar transponders based on [2] into clothing or shoes. These transponders can be detected by a vehicle-mounted radar and enable the reliable localization of pedestrians even in visually occluded situations, thereby supporting timely

identification of potentially dangerous events and improving overall traffic safety.

The transponder receives energy via an electromagnetic transmission link from the vehicle and returns a modulated backscatter signal. While clutter originating from static background reflections can largely be suppressed due to this modulation [2], multipath propagation remains a major challenge, as reflections from the environment may distort the backscattered signal and thereby reduce the accuracy of distance and angle estimation. AI techniques have recently shown promising results in radar signal processing tasks [3].

AI-based Approach

To address this challenge, in this paper the use of AI techniques to improve estimation robustness under multipath conditions is investigated. A convolutional neural network (CNN) is employed for distance estimation, whereas a fully connected feedforward neural

network (FC-FNN) is implemented for angle estimation. The two-stage pipeline architecture—CNN followed by FC-FNN—reflects the inherent structure of the estimation problem. Distance estimation requires extracting spectral peak information from high-dimensional FFT data containing hundreds of frequency bins, which motivates the use of convolutional layers for automatic feature extraction and dimensionality reduction. Angle estimation, in contrast, relies primarily on the phase relationship between the two antenna channels at a single spectral location, yielding a low-dimensional input for which a small fully connected network is sufficient. By conditioning the angle estimator on the predicted distance rather than the ground truth, the full pipeline can be deployed end-to-end at inference time without requiring prior knowledge of the target range. This design also means that errors in the distance estimate propagate into the angle estimation. However, the results in Tab. 1 demonstrate that the CNN achieves sufficient distance accuracy for the subsequent angle estimation to remain effective.

As proof of concept, a dataset with only one reflection is generated first. Synthetic radar signals for two receiving antennas are computed using a mathematical model implemented in a Python script, covering distances up to 10 m and angles up to 90°. The dataset is split into 80 % for training, 10 % for validation, and 10 % for testing. Each antenna signal is preprocessed using a Fast Fourier Transform (FFT).

The CNN for distance estimation comprises three 1D convolutional layers with 6, 10, and 14 filters and kernel sizes of 9, 7, and 5, respectively. Each convolutional layer uses same-padding and ReLU activation. The first two convolutional layers are each followed by a MaxPooling layer with pool size 2, which progressively reduces the temporal dimension and encourages the network to learn increasingly abstract spectral features. The third convolutional layer is followed by a Global Average Pooling (GAP) layer, which collapses the remaining temporal dimension into a single feature vector, thereby reducing the number of trainable parameters and mitigating the risk of overfitting. The subsequent dense layers consist of 12 neurons with ReLU activation and a single linear output neuron. The target distance is normalized by a fixed scale of 10 m prior to training, so the network output lies approximately in the range [0, 1]. The Adam optimizer is used with a learning rate of 0.001, and early stopping with a patience of 5 epochs

monitors the validation loss to prevent overfitting. Training is conducted for 25 epochs with a batch size of 64. Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used for evaluation.

The FC-FNN for angle estimation operates on a compact three-dimensional feature vector derived from the radar signals and the CNN output. First, the trained CNN predicts the normalized target distance, which determines the expected beat frequency. Around this beat frequency bin, the inter-antenna phase difference is computed using a magnitude-weighted circular mean over a window of ± 8 frequency bins. The sine and cosine of the resulting mean phase difference, together with the normalized predicted distance, form the input feature vector $[\sin(\Delta\varphi), \cos(\Delta\varphi), d_1]$. The FC-FNN itself consists of two hidden layers with 16 and 8 neurons, both using ReLU activation and L2 regularization ($\lambda = 10^{-4}$) to constrain the weights. The output is a single linear neuron predicting the normalized angle. Training uses the Adam optimizer with a learning rate of 0.001 for up to 30 epochs with a batch size of 64, with early stopping at a patience of 6 epochs. The architectures of both networks were based on preliminary experiments and their hyperparameters tuned empirically.

Evaluation and Comparison to Conventional Methods

To assess the robustness of the approach, additional datasets with different numbers of reflections (specifically 2, 5, 10, 20, and 50) are generated. In each evaluation step, the previously trained CNN and FC-FNN are retrained to improve their performance under increased multipath conditions. As in real-world scenarios the number of reflections is unknown, an additional dataset with a random number of reflections between 2 and 50 is created. Finally, additive noise with a signal-to-noise ratio (SNR) of 23 dB is applied.

For comparison, a conventional methodology for both distance and angle estimation is also implemented, serving as a non-learned baseline to quantify the performance gain achieved by the proposed AI-based estimators. For distance estimation, spectral peak detection is applied on the preprocessed beat signal spectrum. The beat frequency $f_b f_j$ is identified as the frequency corresponding to the maximum amplitude in the magnitude

spectrum, and the target range is obtained via the standard FMCW range equation

$$d = \frac{f_b c_0 T_{\text{FMCW}}}{2 B_{\text{FMCW}}}, \quad (1)$$

where c_0 denotes the speed of light, T_{FMCW} the chirp duration, and B_{FMCW} the sweep bandwidth. The parameters used in this work are

$$B_{\text{FMCW}} = 100 \text{ MHz} \quad \text{and}$$

$$T_{\text{FMCW}} = 10 \text{ ms}.$$

The theoretical minimum resolvable range therefore is

$$\Delta d = \frac{c_0}{2 B_{\text{FMCW}}} \approx 1.5 \text{ m}. \quad (2)$$

For angle estimation, the monopulse phase-comparison principle is applied. The inter-antenna phase difference is computed at the dominant beat frequency bin and a magnitude-weighted circular mean is taken over a window of ± 5 bins around the peak to improve robustness. The target angle is subsequently recovered by inverting the phase-angle relationship. Since the effective antenna baseline is approximately one half wavelength at 24 GHz, the maximum phase difference spans nearly one full cycle of 2π , introducing inherent ambiguity that limits angular accuracy.

Results and Discussion

The results, summarized in Tab. 1, indicate strong performance across all tested scenarios. While the models perform best with only one reflection, their estimations for datasets with a random number of reflections (between 1 and 50) and added noise remain sufficiently accurate for the intended application. In the noisy scenario, the models achieve an MAE of 0.29 m for distance estimation and 7.50° for angle estimation.

Fig. 1 and Fig. 2 illustrate the prediction accuracy for the simplest scenario with a single reflection and no noise. For distance estimation (Fig. 1), the AI-based CNN predictions closely follow the ideal $y = x$ line across the full range from 0 to 10 m, with only minimal scatter. The conventional peak-detection method, by contrast, exhibits a visibly stepped pattern attributable to the limited range resolution of $\Delta d \approx 1.5 \text{ m}$, which prevents it from resolving closely spaced distance values. This fundamental limitation explains the large baseline MAE of

0.34 m even in the absence of multipath interference. For angle estimation (Fig. 2), both methods perform well in the single-reflection case, with the AI-based FC-FNN achieving an MAE of 0.06° compared to 3.18° for the phase-comparison method.

The more realistic scenario with a random number of reflections and additive noise is depicted in Fig. 3 and Fig. 4. For distance estimation (Fig. 3), the AI-based CNN maintains a concentrated scatter around the ideal line, whereas the conventional method shows broader dispersion, particularly at shorter distances where multipath reflections can shift the dominant spectral peak. For angle estimation (Fig. 4), both methods exhibit increased scatter compared to the single-reflection case, which is expected given the higher complexity introduced by multipath propagation and noise. Nevertheless, the AI-based FC-FNN produces a tighter distribution around the diagonal than the conventional phase-comparison approach, especially in the mid-angle range between 20° and 60° .

A comparison between the AI-based approach and the conventional baseline (Tab. 2) reveals that the proposed models consistently outperform the traditional methods across nearly all multipath scenarios. For distance estimation, the AI-based CNN achieves substantially lower MAE already from a single reflection (0.06 m vs. 0.34 m), a difference that is attributable to the fundamental range resolution limit of $\Delta d_1 \approx 1.5 \text{ m}$ inherent to the spectral peak detection method. Notably, the conventional distance estimator remains relatively stable across increasing reflection counts, as the peak-detection error is dominated by the resolution limit rather than by multipath distortion. The CNN, while starting from a much lower error, shows a gradual increase in MAE with more reflections, reaching 0.61 m at 50 reflections—a regime where the spectral shape is heavily distorted by overlapping multipath components.

For angle estimation, the conventional phase-comparison method performs comparably to the FC-FNN at low reflection counts, but degrades more rapidly with increasing multipath interference, consistent with the phase ambiguity introduced by the 2π constraint. In the most realistic scenario—random multipath reflections with added noise—the AI-based models achieve MAE improvements of approximately 28 % for distance (0.29 m vs. 0.40 m) and 9 % for angle (7.50° vs. 8.19°), demonstrating the practical advantage of the

learned approach under real-world-like conditions. The improvement observed when adding noise to the random-reflections dataset (MAE decreasing from 0.37 m to 0.29 m for distance and from 8.29° to 7.50° for angle) suggests that the models benefit from training on the larger combined dataset, which effectively provides a form of data augmentation that improves generalization.

Conclusion and Outlook

Notably, the neural network weights were not reset between experiments involving increasing numbers of reflections. The improvement when adding noise suggests that the models benefit from training with a larger dataset. Future work will focus on validating the approach using real-world radar measurement to determine whether the findings can be transferred to practical traffic situations.

Further analysis will be considered for the approach, e.g. how the final model performs for low numbers of reflections, how further hyperparameter tuning might impact the model, how further restricting the angle improves performance, how the performance develops for higher distances relevant to road traffic scenarios and how moving objects may impact the model. Finally, a validation with real data is planned.

For real-world scenarios, fast computation is critical. Therefore an implementation may require hardware accelerators. Generating hardware accelerators using a high level synthesis tool (HLS) such as tinyHLS is one feasible approach [4].

Acknowledgement

This manuscript was edited with the assistance of a generative AI model to enhance clarity, readability, and consistency. All suggestions were critically reviewed and accepted or modified by the authors, who take full responsibility for the final content. No AI tools were used to create, alter, or manipulate original research data or results.

This research has been funded in the project *SafeMe* by the German ministry for research, technology and space under grant no. 16ME0887.

References

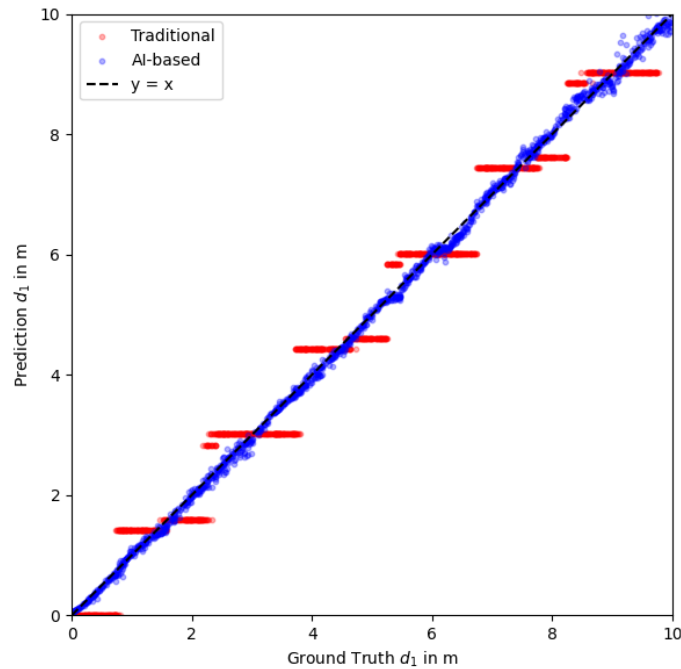
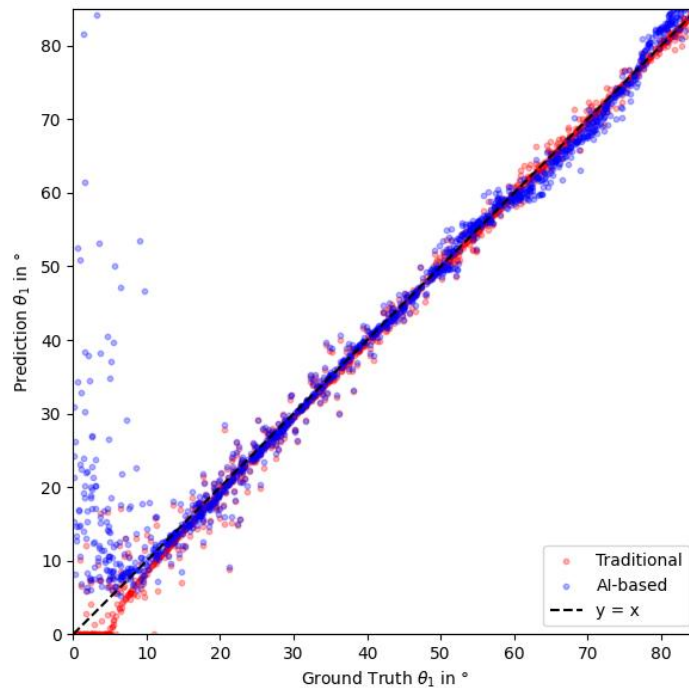
- [1] Campbell, Bob J., et al. "A review of pedestrian safety research in the United States and abroad." (2003).
- [2] S. Böller, T. Greuter and A. Grabmaier, "FMCW Based Positioning Using Multiple SHF RFID Transponders," *2023 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nuremberg, Germany, 2023, pp. 1-5
- [3] Feng, Weike, Xiaowei Hu, and Xingyu He. "Artificial Intelligence (AI)-Based Radar Signal Processing and Radar Imaging." *Electronics* 13.21 (2024): 4251.
- [4] Hoyer, Ingo, et al. "tinyHLS: a novel open source high level synthesis tool targeting hardware accelerators for artificial neural network inference." *Physiological Measurement* 46.1 (2025): 015002.

Tab. 1: Test MAE for different numbers of reflections using the proposed AI-based approach.

Reflections	1	2	5	10	20	50	Random	Noise
Distance in m	0.0643	0.1432	0.3188	0.2137	0.4963	0.6117	0.3745	0.2869
Angle in °	0.0600	2.31	3.79	5.83	6.93	10.97	8.29	7.50

Tab. 2: Test MAE for different numbers of reflections using traditional algorithms.

Reflections	1	2	5	10	20	50	Random	Noise
Distance in m	0.3448	0.3451	0.3400	0.3594	0.3673	0.5871	0.4276	0.4019
Angle in °	3.18	2.16	4.28	6.86	8.02	12.28	8.35	8.19

**Fig. 1:** Distance prediction vs. ground truth on the test set for the ai-based method (blue) and conventional method. The dataset features only one reflection and without noise.**Fig. 2:** Angle prediction results vs. ground truth on the test set for AI-based method (blue) and conventional method (red). The dataset features only one reflection without noise.

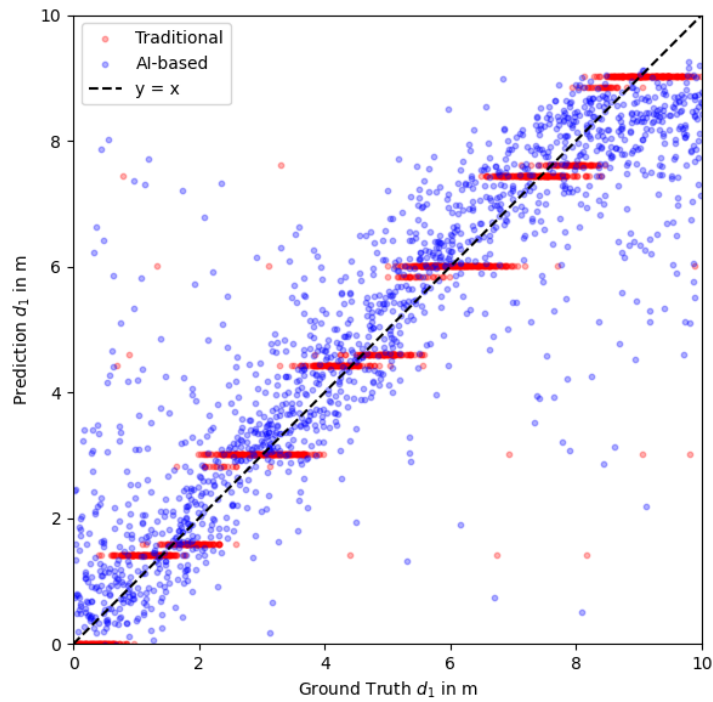


Fig. 3: Distance prediction results from the test dataset with random number of reflections and noise of a standard deviation of 0.05 for the traditional and AI-based methodology.

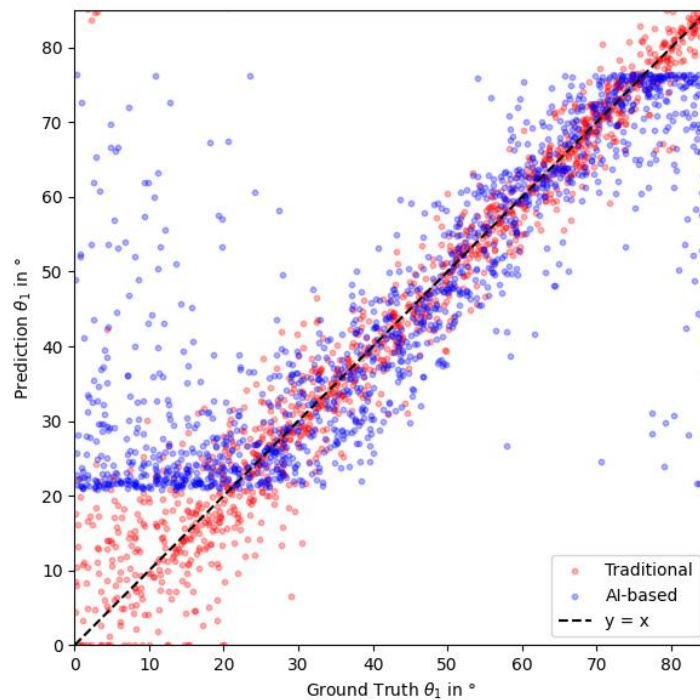


Fig. 4: Angle prediction results from the test dataset with random number of reflections and noise of a standard deviation of 0.05 for the traditional and AI-based methodology.

Metamodel-based Optimization of Coil Units of Axial-Flux Machines and their Validation Using a Test Bench

Folke Schwinning¹, Jana Ihrens¹, Thorsten A. Kern¹

*¹Institute for Mechatronics in Mechanics, Hamburg University of Technology,
Eißendorfer Str. 38, 21073 Hamburg, Germany*

Abstract

The demand for high power density in electrical machines has led to the increasing adoption of axial-flux topologies. Due to the strict spatial and gravimetric constraints of these systems, optimizing the coil unit – which constitutes a primary source of thermal and magnetic losses – is a critical design objective. While metamodel-based methodologies enable rapid, multiphysical design exploration, their predictive accuracy relies heavily on empirical validation. Conventional simulation tools require significant computational effort to accurately resolve the coupled effects of manufacturing tolerances, process-induced material degradation, and complex thermal pathways. Consequently, high-quality empirical data is necessary to validate simulation models and train reliable metamodels. However, existing testing frameworks for axial-flux architectures are typically fragmented, evaluating physical domains in isolation and thereby limiting the analysis of coupled multiphysical interactions.

This paper presents a novel, integrated multiphysical test bench designed for the comprehensive validation of axial-flux coil units. Operating at up to 1000 V or 600 A, the platform enables simultaneous thermal, magnetic, and mechanical characterization of a single specimen without requiring a spinning rotor. Key innovations include active field emulation for manipulating the air gap flux density distribution, a 6-DoF hexapod with custom actuators to emulate manufacturing tolerances and measure resultant multidimensional forces, and an automated 6-DoF kinematic 3D Hall sensor system for precise air gap flux and stray flux mapping. The system architecture is coupled with a Digital Twin for asynchronous experiment planning and safe operation. By establishing this infrastructure, the proposed test bench provides the high-fidelity, cross-domain empirical data necessary to transition advanced axial-flux concepts from simulation to validated prototypes.

Keywords: axial-flux, testbench, multiphysics, magnetic measurements, thermal characterization

Introduction and Motivation

The demand for high-performance electrical machines for transport and industry is driving research and innovation, especially in terms of power density, efficiency and thermal management. A particular focus is placed on axial-flux machines. These topologies are highly attractive not only due to their exceptional gravimetric power density – with recent industry benchmarks reaching up to 59 kW/kg [1] – but also regarding other critical high-performance metrics such as overall efficiency, favorable thermal time constants, and scalable production and assembly costs. They are increasingly being adopted beyond automotive applications, from aerospace to auxiliary drives like hydraulic pumps for off-highway applications or coolant pumps in data centers. To push those performance boundaries, continuous design optimization and the exploration of novel materials and cooling concepts are necessary.

Traditionally, the design and optimization of electrical machines is conducted with Finite Element Analysis (FEA) and Computational Fluid Dynamics (CFD). While FEA and CFD are highly capable, incorporating real-world influences such as manufacturing tolerances, process-induced material changes, and coupled temperature dependencies results in significant computational effort. For novel concepts like direct-liquid-cooled litz wire hollow conductors, validated simulation models or methodologies are often not yet established.

To explore the design space and find an optimal design with minimal computational effort, the optimization of multiphysical systems is increasingly shifting towards metamodel-based methodologies [2]. Metamodels – typically in the form of Functional Mock-up Units (FMUs) – enable the rapid approximation of key performance indicators for complex designs. However, the prediction accuracy is constrained

by the quality and quantity of training data used to train the model.

Acquiring this training data presents a critical challenge, as there is a significant validation gap in the methodology: unlike radial-flux machines, there are no equivalent testing standards for the axial-flux architecture. Existing test procedures are typically fragmented, focusing on a single physical domain (e.g. magnetic or thermal), limiting the ability to evaluate complex, coupled multiphysical interactions.

To bridge this validation gap, a versatile testbench was developed at Hamburg University of Technology (TUHH) specifically for the simultaneous multiphysical characterization of axial-flux coil units (see Figure 1). This paper presents the key features and capabilities, and discusses measurement methods and strategies as well as preliminary results obtained from thermal experiments. Specimens up to a cross-section of 50x50 mm and axial lengths of 100 mm can be accommodated, as well as entire machines up to a diameter of the magnetic circuit of 160 mm. Simultaneous thermal, magnetic, and mechanical test procedures and measurements can be conducted on a specimen in a single setting without the need to transfer it between different test setups.

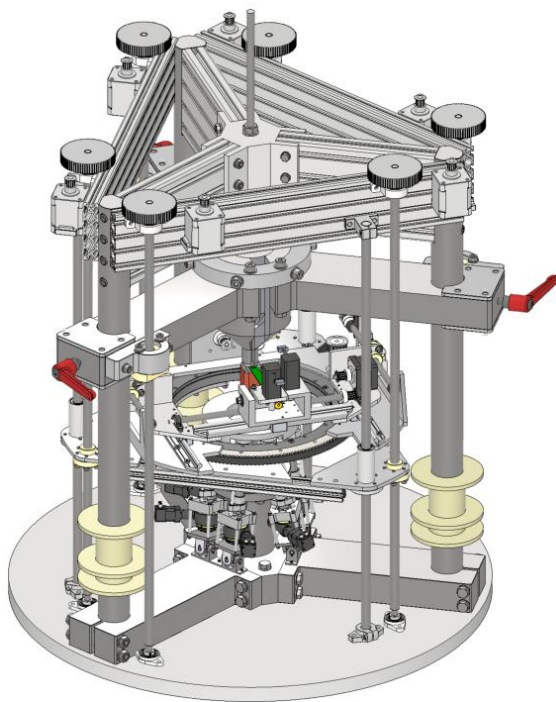


Figure 1: CAD assembly overview of the multiphysical test bench

The Optimization Target: The Coil Unit

Axial-flux machines are subject to strict space and weight constraints, driven by the demanding requirements of the automotive and aerospace sectors. In these highly compact architectures, the stator coil unit represents the primary bottleneck, but it also exhibits the highest potential for optimization [3], primarily due to its unique geometric freedom and the absence of a traditional stator yoke and end windings. While rotor optimization is relatively straightforward and achieved by adjusting permanent magnet grades, thicknesses, and relying on basic rotational forced convection for cooling, the coil unit is subject to concentrated high-frequency losses and complex thermal pathways. Just as stators remain fundamentally similar across synchronous, induction, or reluctance machines despite varying rotor designs, the axial-flux coil unit is a universally critical component that heavily dictates overall machine performance.

In case of the Yokeless and Segmented Armature (YASA) topology [4], the coil unit consists of a magnetically soft core and pole shoes as well as a concentrated single-tooth winding (see Figure 2). Consequently, research is increasingly focused on next-generation materials and topologies for the coil unit. Conductor innovations include the use of direct liquid cooling topologies such as hollow conductors [5], flooded stator cooling [6], aluminium as a conductor material [7] and high thermal conductivity resin-potted coils. Simultaneously, advanced core materials like Soft Magnetic Composites (SMC), nanocrystalline materials, grain-oriented steel and FeCo alloys are being investigated.

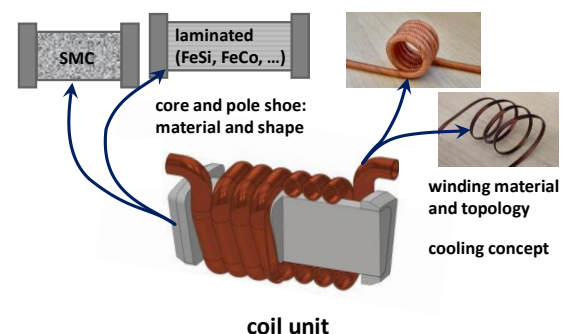


Figure 2: Coil Unit with different parameters that can be varied: material, shape, conductor topology, winding material, cooling concept etc.

Evaluating these novel concepts requires addressing a significant gap of the impact of manufacturing processes on theoretical material properties in the current literature. Unlike radial-flux machines, axial-flux

architectures have a shorter history of extensive research and development, especially given the wide variety of topologically different approaches (e.g., yokeless vs. yoked; single vs. dual rotor or stator). Various manufacturing processes are employed today – such as stacking, baking, grinding, etching, and wire electrical discharge machining (EDM) – and each exerts a different, complex influence on the magnetic material [8]. Furthermore, the cores and pole shoes of axial-flux coil units inherently exhibit complex 3D flux distributions. Validating these degraded, 3D multiphysical states necessitates dedicated empirical testing.

Multiphysical Test Bench Architecture

The test bench is conceptualized as a highly modular, multi-domain architecture. The core hardware comprises a main frame integrating a custom 6-Degree-of-Freedom (6-DoF) hexapod structure and a central platform designed to securely hold the specimen. A dedicated measurement PC is responsible for the synchronous recording of all data and the execution of real-time control algorithms.

The test bench is deeply integrated with the Power-Hardware-in-the-Loop laboratory (PHiLsLab) infrastructure at TUHH [9]. This integration provides critical high-voltage, high-current and high-frequency power supplies, high-speed measurement recorders, auxiliary diagnostic equipment, and fluid thermal interfaces.

A key advantage of this architecture is its dynamic real-time emulation capability. Realistic operational scenarios, including the application of arbitrary voltage curves, can be emulated and controlled in real time. This positions the platform as an exceptional prototype testing infrastructure, enabling not only the isolated testing of the coil unit but also its interaction with peripheral hardware. For example, the system can analyze a coil's thermal and magnetic response to high-frequency ripple voltages when supplied by an external integrated inverter, or it can emulate inverters and other dynamic loads via Power-Hardware-in-the-Loop (PHiL). The overarching capabilities and boundary parameters of the test bench are summarized in Table 1.

Table 1: Test bench specifications and maximum parameters for dimensions, voltage, current, and frequency.

Property	Value	Unit
number of coil units	1...3	-
maximum diameter	50	mm
maximum length	100	mm
maximum current	600	A
maximum voltage (DC)	1000	V
maximum peak voltage (AC / arbitrary waveform)	424	V
max. frequency	30	kHz

Measurement Subsystems & Methodology

1. Digital Twin Integration & Safe Operation

The physical test bench is coupled with a Digital Twin, which is utilized offline and asynchronously to guide the experimental process. The workflow is iterative: initial, approximate coil parameters are derived from the simulation for experiment planning. Following the physical measurements, the empirical data is used to refine the parameters within the Digital Twin, continuously improving model fidelity for subsequent test iterations.

To ensure safe operation during these rigorous tests, the Digital Twin is utilized to estimate expected threshold values for multidimensional forces, temperatures, and induced voltages. If physical sensors detect values approaching these simulated thresholds, the system issues operator warnings. Experiments are strictly supervised, and operating parameters are raised gradually to safely approach physical edge cases. Furthermore, the Digital Twin employs a numerical optimization routine in MATLAB, controlling a COMSOL model to calculate the exact manipulation coil voltages required to achieve desired, inhomogeneous flux density distributions in the air gap for active field emulation.

2. Emulation of Manufacturing Tolerances

Manufacturing and assembly tolerances – such as discrepancies in iron core lengths, pole shoe assembly positions, or air gap widths – create unintended and unbalanced magnetic forces under load. Specifically, an uneven air gap directly results in asymmetrical magnetic forces that significantly accelerates bearing wear. Under high-frequency excitation, these static and dynamic forces translate into high-frequency vibrations that can ultimately cause structural machine failure. The severity of these vibrations is heavily dependent on the chosen manufacturing process and material of the

machine casing, whether it be milled aluminum, carbon fiber reinforced polymers (CFRP), or injection-molded reinforced thermoplasts. The test bench is designed to determine which tolerance classes are acceptable to prevent these failure modes, thereby informing the viability of specific manufacturing processes.

To empirically emulate these mechanical tolerances, the test bench utilizes a 6-DoF hexapod mechanism (see Figure 4). A critical feature of this system is the custom design of its linear actuators, developed specifically to withstand high magnetic forces (up to 2.5 kN) while exhibiting minimal structural deflection over a precise travel range. Off-the-shelf linear actuators capable of meeting these strict requirements are prohibitively expensive and typically overengineered; thus, the custom design provides a necessary and cost-effective solution for accurate tolerance emulation.

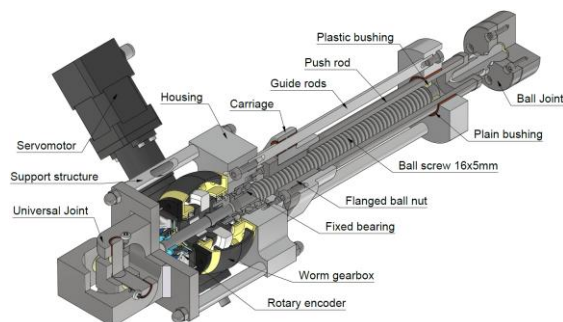


Figure 3: Custom-designed linear actuator for the hexapod

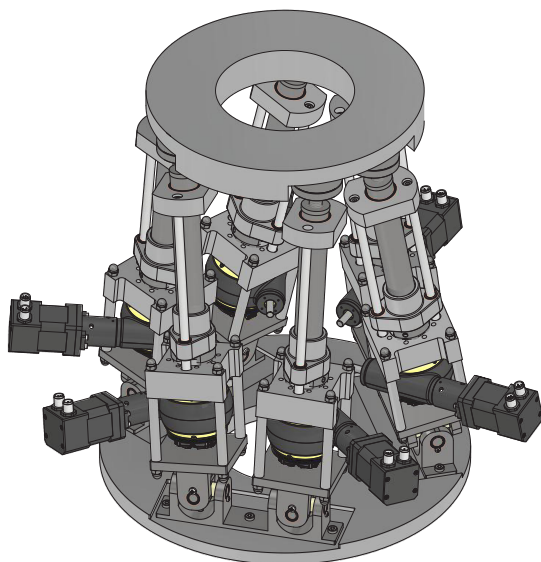


Figure 4: 6-DoF hexapod mechanism with custom linear actuators for manipulating the pose of the specimen between the two test bench air gaps

3. Multidimensional Force Measurement

Force and torque measurements are integrated directly into the custom linear actuators of the hexapod, enabling precise 6-DoF mechanical characterization without the need for separate, bulky external load cells. To achieve high fidelity across varied operating conditions, a dual-sensor concept is implemented. Piezoelectric sensors are utilized to capture high-dynamic, high-frequency vibrations, where their fast response times excel. In tandem, strain gauges are deployed to measure static loads and long-term baseline forces, effectively compensating for the inherent temporal drift of piezoelectric materials and being able to measure remanent magnetic forces.

4. Fundamental Electrical Measurements

High-precision electrical measurement forms the foundation for all subsequent magnetic and thermal characterizations. The core measurement equipment consists of two coupled ZES Zimmer LMG671 power analyzers. While standard experiments run at moderate sampling rates, the system features an extreme maximum bandwidth of 1.2 MS/s to capture highly dynamic edge cases, e.g. for high-frequency experiments or non-sinusoidal waveforms. The measurement tolerances and low measurement uncertainty of this device are critical capabilities of the overall platform. Temperature data is synchronized continuously, utilizing either an external NI USB-6210 DAQ or directly through the LMG's Process Signal Interface (PSI) at a rate of 100 S/s.

A strict 4-wire measurement topology is utilized: current is measured through high-cross-section supply cables, while voltage is probed directly across the specimen itself. This isolates the true impedance and power losses of the coil unit, eliminating parasitic errors caused by the electrical resistance and thermal heating of the supply infrastructure.

The power supply infrastructure is highly versatile. DC characterizations are powered by precision laboratory sources, while AC scenarios utilize a three-phase Spitzenberger & Spies APS7500 system capable of generating arbitrary signal waveforms. Combined with an additional high-current transformer and an optional B6 rectifier setup, currents of up to 1800 A can be supplied. This configuration was successfully proven during the testing of litz wire hollow conductors [10], although the maximum current is realistically limited to 600A.

by the current sensors and supply wires. Crucially, the Spitzenberger units are linear power sources rather than Switched-Mode Power Supplies (SMPS). This provides an exceptionally pure, noise-free signal, which is mathematically vital for ensuring accuracy when integrating induced voltage $u_{ind} = -\frac{d\Phi}{dt}$ to calculate the flux density B for magnetic characterization.

5. Active Field Emulation

To evaluate the specimen under realistic dynamic conditions without requiring a physical spinning rotor, an active field emulation mechanism is utilized. The specimen is positioned between two pole shoes, where the top pole shoe is separated into three magnetically isolated iron return paths (see Figure 5). Each path is equipped with an independent manipulation coil acting as a secondary electromagnet. These manipulation coils generate an assisting or opposing magnetic voltage to actively manipulate the primary flux generated by the specimen.

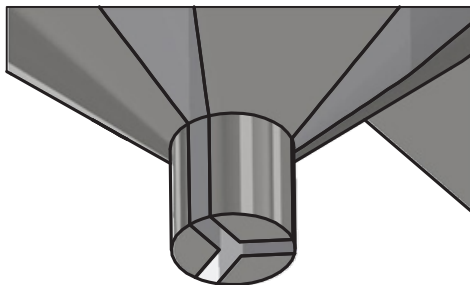


Figure 5: Separated top pole shoe with magnetically isolated iron return paths

This setup treats the air gap boundary as a black box, allowing for the emulation of arbitrary surface flux density distributions. Different scenarios such as the transient passage of a permanent magnet pole or shifting reluctance paths towards the shortest flux path are possible. Compared to state-of-the-art methods that rely on passive measurement or mechanical alteration of the air gap [11], this active electronic shaping provides vastly superior experimental flexibility. While the ultimate goal is dynamic real-time emulation, initial testing focuses on applying static DC currents to emulate discrete "snapshots" of various rotor positions, as this approach is computationally highly reliable to optimize within the COMSOL Digital Twin. The results of the active field emulation in the Digital Twin can be seen in Figure 6.

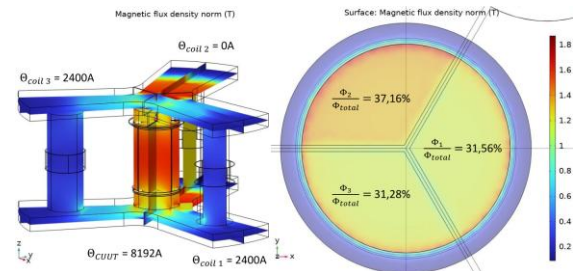


Figure 6: Electromagnetic FEM Digital Twin and manipulated air gap flux density distribution (here: shown with a coil unit originally planned to be significantly larger)

6. Magnetic Characterization

The magnetic characterization of the specimen's soft magnetic material is based on the principles of the DIN EN IEC 62044-3 standard. However, direct application of the norm is impossible due to the complex, mixed-material nature of the test rig, which combines the solid iron of the test bench with the laminated or SMC materials of the specimen. To capture the induced voltage safely and accurately, dedicated measurement coils – either a single extended coil on the lower pole shoe or separate coils on the three upper iron return paths – are used. This prevents the exposure of measurement equipment to dangerously high voltages that can occur in the primary windings at high flux densities and frequencies.

This infrastructure enables the mathematical reconstruction of B-H curves from the induced voltage and current measurements. Simulation studies have proven this reconstruction is highly viable even with noisy data with a Signal-to-Noise Ratio of 40dB (see Figure 7).

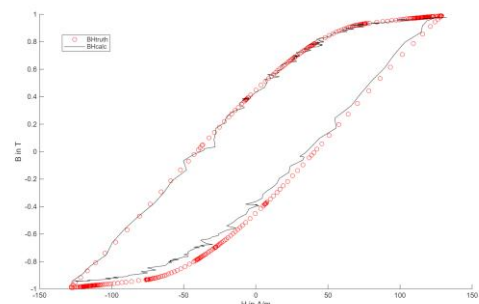


Figure 7: Reconstruction of a B-H-curve (black) from noisy measurement with an SNR of 40dB; red: ground truth B-H-curve in Simulink

To isolate the specimen's true core losses from the inherent parasitic losses of the test bench, a step-by-step empirical substitution method is applied:

- Small-signal DC and AC impedance baselines (up to 10 kHz) are established for the coil both within the test bench and in free air.
- Total losses are recorded with closed pole shoes across varying frequencies and amplitudes.
- The exact measurements are repeated using a solid cylinder "dummy" specimen (constructed from the test bench material, matching the pole shoe dimensions) across varying air gap lengths.
- Data from these steps is used to construct a precise approximation model of the test bench's inherent losses.
- During actual specimen testing, the modeled bench losses are subtracted from the total measured losses, isolating the empirical core loss of the prototype.

7. 3D Air Gap Flux Mapping

To validate the simulated air gap flux density, a 6-DoF sensor kinematic system is employed (see Figure 8), equipped with a SENM3Dx 3D Hall sensor [12]. The sensor is embedded in a specifically designed PCB measuring less than 1 mm in thickness, allowing insertion into minimal air gaps. The sensor's Flux Sensitive Volume (FSV) is located approximately 0.35 mm below the cap surface, enabling the measurement point to reach within ~0.5 mm of the specimen surface.

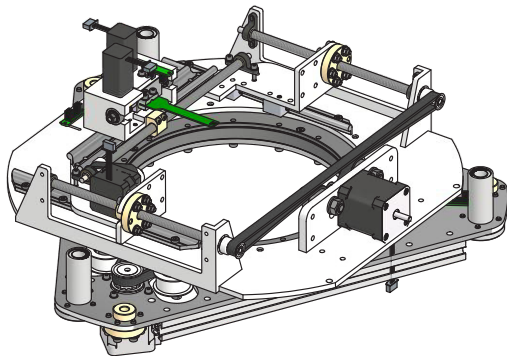


Figure 8: 6-DoF kinematic for mapping the flux density in and around the air gap (isolated view)

To eliminate manual positioning errors and expedite mapping, the sensor's movement is automated using ROS2. The path planning algorithm utilizes a simplified CAD representation of the coil unit to execute precise spatial sweeps, see Figure 9. This integration also features essential collision detection to protect the fragile sensor PCB during automated operation.

At the same time, the test bench was designed to accommodate a rotor, stator or even an entire machine to measure the flux density of surface permanent magnets, the flux density at the stator surface or even in the air gap during static operation (without rotation).

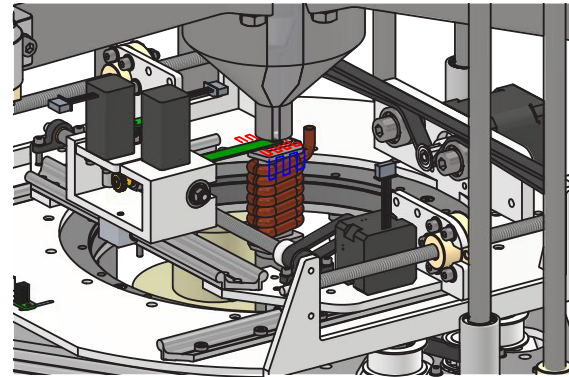


Figure 9: Possible measurement path in the air gap (red) and around the pole shoe for stray flux measurement

8. Thermal & Hydraulic Characterization

Thermal modeling is a primary focus of the test bench, supported by a comprehensive array of temperature and flow sensors to capture complex thermal gradients. The primary internal temperature of the copper winding is calculated indirectly from its electrical impedance. Because the baseline small-signal DC impedance is established prior to testing, real-time shifts in the active impedance during operation directly correlate to the average internal conductor temperature. While the raw electrical signal requires numerical filtering within the power analyzer to mitigate inherent noise, this methodology provides highly accurate internal thermal data without embedding destructive physical sensors into the winding structure. This impedance-based approach is supplemented by an extensive network of external sensors: multiple surface thermocouples (Type K) monitor localized thermal boundaries, two high-precision Pt100 sensors track the coolant fluid inlet and outlet temperatures, and a volumetric fluid flow meter enables full calorimetric power loss measurements. Additionally, a thermographic camera is utilized for the manual validation of surface temperature distributions. Together, this multi-domain thermal data provides the empirical foundation required for the creation and parameter identification of advanced thermal models.

Beyond thermal data, the test bench empirically verifies the hydraulic pressure drop across the specimen's cooling system (e.g., direct-liquid-cooled hollow conductors). Contrasting these

empirical measurements with theoretical calculations (such as the Gnielinski correlation [13]) is critical to investigate novel cooling concepts for their suitability (with respect to the necessary pump pressure) and validate simulation models. By sweeping through different volumetric flow rates, coolant types, and simulated power losses, the platform enables the comprehensive benchmarking, model verification and model calibration of novel cooling concepts.

Preliminary Experimental Results

To demonstrate the data acquisition capabilities of the test bench, preliminary thermal evaluation experiments were conducted under load, utilizing a novel litz wire hollow conductor prototype as the test specimen.

The primary objective of these experiments was to rigorously investigate the performance limits of the novel conductor topology, evaluating its capability to sustain unprecedented, record-breaking continuous current densities under extreme thermal loads. Concurrently, these experiments demonstrated the immense value of the test bench's comprehensive thermal monitoring capabilities. The system successfully captured and contrasted the distinct thermal gradients between the coolant stream temperature, the external surface temperature, and the internal impedance-derived conductor temperature, see Figure 10.

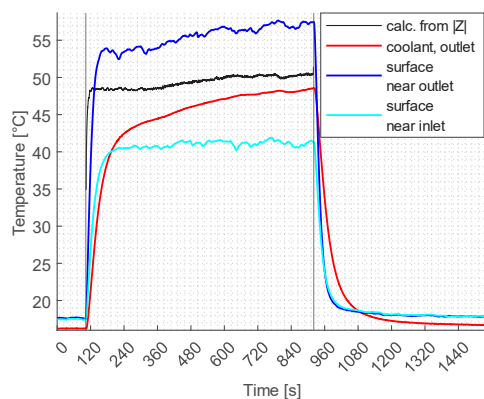


Figure 10: Transient temperature curves comparing internal impedance, surface, and coolant temperatures over time ($j = 100 \text{ A/mm}^2$)

Capturing these distinct thermal gradients and their respective temporal response delays is crucial. This multidimensional, high-resolution thermal data provides the exact empirical foundation required for the accurate parameter identification of advanced thermal models, laying the groundwork for training Lumped-Parameter Thermal Networks (LPTNs) or the Thermal Neural Networks (TNNs) discussed in the introduction.

Conclusion and Outlook

The conceptual and infrastructural foundation for a fully integrated, multiphysical test bench has been successfully established at TUHH. Key systems, including the high-precision electrical measurement infrastructure, versatile power emulation, and core thermal evaluation concepts, are already operational. Furthermore, foundational simulations regarding the Digital Twin, magnetic characterization methodologies, and active field manipulation have been fully developed and proven theoretically. Even in its current build phase, this architecture clearly demonstrates the capacity to capture the complex, multi-domain data required to finally bridge the validation gap for metamodel-based axial-flux machine design.

Immediate next steps focus on manufacturing of the magnetic circuit, final dimensioning and assembly of the custom hexapod actuators with their integrated force sensors, and the finalization of the ROS2-driven 6-DoF Hall sensor kinematics. Upgrades to specific hydraulic sensors have already been done (pressure and volumetric flow rate), with results from repeated experiments expected in Q2 2026.

Once physical commissioning is complete, research will shift toward validating the simulated magnetic characterization and active field manipulation models on real hardware. Ultimately, the acquired measurement data and the resulting validated simulation models will be utilized to optimize axial-flux machines across a wide variety of combinations involving different materials, geometric shapes, conductor topologies, and novel cooling concepts. This high-fidelity empirical foundation will directly support the overarching research methodology: generating highly accurate metamodels and Functional Mock-up Units (FMUs) to rapidly explore the design space and accelerate future motor development.

Bibliography

- [1] "YASA smashes own unofficial power density world record pushing state-of-the-art electric motor to staggering new 59kW/kg benchmark," YASA Limited. Accessed: May 16, 2026. [Online]. Available: <https://yasa.com/news/yasa-smashes-own-unofficial-power-density-world-record-pushing-state-of-the-art-electric-motor-to-staggering-new-59kw/kg-benchmark/>
- [2] F. Schwinning and T. A. Kern, "Metamodel-Based Design Exploration and Optimization for Multiphysical Designs of Axial-Flux Synchronous Machines," in *2024 IEEE Design*

- Methodologies Conference (DMC)*, Nov. 2024, pp. 1–6. doi: 10.1109/DMC62632.2024.10812148.
- [3] P. Lindh *et al.*, “Direct Liquid Cooling Method Verified With an Axial-Flux Permanent-Magnet Traction Machine Prototype,” *IEEE Trans. Ind. Electron.*, vol. 64, no. 8, pp. 6086–6095, Aug. 2017, doi: 10.1109/TIE.2017.2681975.
 - [4] T. J. Woolmer and M. D. McCulloch, “Analysis of the Yokeless And Segmented Armature Machine,” in *2007 IEEE International Electric Machines & Drives Conference*, May 2007, pp. 704–708. doi: 10.1109/IEMDC.2007.382753.
 - [5] M. Numberger and J. Fleming, “The Turbocharger for Electric Motors,” *MTZ Worldw.*, vol. 87, no. 1, pp. 32–39, Jan. 2026, doi: 10.1007/s38313-025-2142-2.
 - [6] R. Camilleri, D. A. Howey, and M. D. McCulloch, “Predicting the Temperature and Flow Distribution in a Direct Oil-Cooled Electrical Machine With Segmented Stator,” *IEEE Trans. Ind. Electron.*, vol. 63, no. 1, pp. 82–91, Jan. 2016, doi: 10.1109/TIE.2015.2465902.
 - [7] C. R. Sullivan, “Aluminum Windings and Other Strategies for High-Frequency Magnetics Design in an Era of High Copper and Energy Costs,” *IEEE Trans. Power Electron.*, vol. 23, no. 4, pp. 2044–2051, Jul. 2008, doi: 10.1109/TPEL.2008.925434.
 - [8] N. Gadiyar, J. Van Verdegheem, and E. L. Severson, “A Review of Axial Flux Permanent Magnet Machine Technology,” *IEEE Trans. Ind. Appl.*, vol. 59, no. 4, pp. 3920–3933, Jul. 2023, doi: 10.1109/TIA.2023.3258933.
 - [9] J. Ihrens, S. Möws, L. Wilkening, T. A. Kern, and C. Becker, “The Impact of Time Delays for Power Hardware-in-the-Loop Investigations,” *Energies*, vol. 14, no. 11, p. 3154, May 2021, doi: 10.3390/en14113154.
 - [10] F. Schwinning, J. Ihrens, and T. A. Kern, “Thermal Investigation and Validation of Litz Wire Hollow Conductors for Electrical Machines,” in *2026 International Conference on Electrical Machines (ICEM)*, Sep. 2026, in press.
 - [11] H. Vansompel, P. Sergeant, L. Dupré, and A. Van den Bossche, “Axial-Flux PM Machines With Variable Air Gap,” *IEEE Trans. Ind. Electron.*, vol. 61, no. 2, pp. 730–737, Feb. 2014, doi: 10.1109/TIE.2013.2253068.
 - [12] D. Popovic Renella *et al.*, “Revealing the potential of a new 3D Hall sensor in advanced inspection robotics,” *Acta IMEKO*, vol. 13, no. 4, pp. 1–5, Nov. 2024, doi: 10.21014/actaimeko.v13i4.1752.
 - [13] P. Stephan, S. Kabelac, M. Kind, D. Mewes, K. Schaber, and T. Wetzel, Eds., *VDI-Wärmeatlas: Fachlicher Träger VDI-Gesellschaft Verfahrenstechnik und Chemieingenieurwesen*. in Springer Reference Technik. Berlin, Heidelberg: Springer, 2019. doi: 10.1007/978-3-662-52989-8.

Ein integriertes LSPIV Fernerkundungssystem, Sensordatenfusion und Bildverarbeitung für Assistenzsysteme im regulären Einsatz der Fährschifffahrt

Nick Rüssmeier^{1,2}, Anael Ngome¹

¹Jade Hochschule, Fachbereich Ingenieurwissenschaften, Wilhelmshaven, Deutschland

²Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI), Marine Perception, Oldenburg, Deutschland

Zusammenfassung

Zentrale Herausforderungen der operativen Schifffahrt sind der nachhaltige und optimale Schiffsbetrieb sowie die Reduktion von Emissionen. Insbesondere in Strömungs- und Tidengewässern lassen sich Treibstoffeinsparungen erzielen, wenn Umwelteinflüsse für eine sparsame Fahrweise ausgenutzt werden. Im Rahmen der Forschung an Assistenzsystemen zur Erleichterung und Effizienzsteigerung des Schiffsbetriebes wird dazu ein Umweltsensor-System für lokale live Strömungsmessungen für den Betrieb an Bord einer Fähre entwickelt. Auf Basis dieser Sensordaten errechnet ein KI-basiertes Assistenzsystem die zum Fahrzeitpunkt manövertechnisch effizienteste Route im regulären Fährbetrieb. In diesem Zusammenhang präsentieren wir ein integriertes LSPIV (Large Scale Particle Imaging Velocimetry) Fernerkundungssystem, welches Langzeitmessreihen in Form ortsveränderlicher Vektorfelder oberflächennaher Strömungsverhältnisse mit einem relativ großen räumlichen Abdeckungsbereich bspw. Strömungen von Flüssen in realen Studiumgebungen im Fahrtgebiet über Flüsse erfasst und analysiert. Mit der aktuellen Forschungsarbeit werden Funktionalitäten für einen einfachen, automatisierten und kontinuierlichen mobilen Einsatz implementiert und getestet. Neben der automatischen extrinsischen Kalibrierung der eingesetzten Thermalkamera zur Kompensation der sich ständig verändernden variablen Ausrichtung an Bord der Fähre unterwegs, geben wir einen Einblick in die Architektur und integrierte Datenstromverarbeitung des optischen Sensorsystems. Die Studie schließt mit den Ergebnissen der frühzeitigen Verifikation des Fernerkundungssystems im Strömungskanal ab, und gibt einen Ausblick auf die weitere Integration auf einem Fährschiff mit dem Einsatz eines KI-basierten Assistenzsystems für den regulären Fährbetrieb.

Keywords: Fernerkundung, optische Strömungsmessverfahren, automatisierte großräumige Partikelbild-Geschwindigkeitsmessung (LSPIV), Fährschifffahrt, Assistenzsystem Energieeinsparung.

Einleitung und Motivation

Im Rahmen der Forschung an Assistenzsystemen zur Erleichterung des Schiffsbetriebes ist es eine Herausforderung, die nautischen Erfahrungen in ein automatisches Assistenzsystem zu überführen. Für die Navigation auf Wasserstraßen ist heutzutage noch eine sehr genaue Kenntnis über das zu befahrene Revier notwendig. Um das Schiff unter ständig wechselnden Umwelteinflüssen, wie Wind und Strömung, sicher ans Ziel zu bringen, wird die vorherrschende Umwelt- und Verkehrslage in aktueller Situation durch die schiffführende Besatzung an Bord durch unterschiedliche Informationsquellen aufgenommen und bewertet. In der See- und Binnenschifffahrt kann hierfür beispielsweise auf Kurs- und

Bahnregler zurückgegriffen werden. Ein System zur Manöveroptimierung entlang vorgegebener Routen [1] kann den Kurs eines Schiffes, entsprechend der vorliegenden Umweltsituation somit ermitteln und für den Schiffsführer als Entscheidungsunterstützung für die Navigation zur Verfügung stellen.

Mittels dem LSPIV (Large Scale Particle Imaging) Messverfahrens könnten Vektorfelder von oberflächennahen Strömungsverhältnissen mit einem relativ großen räumlichen Abdeckungsbereich (bspw. Strömungen von Flüssen in realen Studiumgebungen) berechnet und in einem aktuellen situativen Lagebild ausgegeben werden. Dieses Messverfahren bietet somit eine Informationsbasis für weitere Navigationslösungen sowie Assistenzsysteme zur Erleichterung des Schiffsbetriebes.

Mit diesem Beitrag stellen wir die Entwicklung eines automatisierten und integrierten LSPIV Fernerkundungssystems zur mobilen optischen Strömungsmessung vor. Der Aufbau der Arbeit widmet sich eingehend einer kurzen Vorstellung des zu Grunde liegenden Messverfahrens sowie dem Stand der Wissenschaft und Technik mit vorangegangenen Arbeiten, hierauf folgt eine Darstellung der Sensorsystemarchitektur und Datenstromverarbeitung sowie Ergebnisse der frühzeitigen Verifikation im Strömungskanal, und schließt mit einer Zusammenfassung und Ausblick der weiteren Integration auf einem Fährschiff mit dem Einsatz eines KI-basierten Assistenzsystems für den regulären Fährbetrieb.

Stand der Wissenschaft und Technik, vorangegangener Arbeiten

Bei dem PIV-Messverfahren handelt es sich um ein bildauswertendes Messverfahren. Mittels LSPIV wird das PIV-Messverfahren auf einen größeren räumlichen Abdeckungsbereich appliziert. Insgesamt bieten (passive/aktive) optische Sensoren durch deren unterschiedlichen Sensor-Charakteristika ein weites elektromagnetisches Spektrum für das Anwendungsfeld in maritimen Umgebungen oder Binnenwasserstraßen. Als Grundlage für die LSPIV Berechnung dienen Merkmale in einem aufgenommenen Bild in Form von natürlich vorkommenden Tracern, wobei je nach Spektralbereich (visuell oder infrarot) entweder Schaumkronen/Whitecaps oder Signaturen auf dem Wasser, welche einen von ihrer Umgebung unterschiedlichen Emissionskoeffizienten aufweisen, gut genutzt werden können [2].

Die grundlegende Funktionsweise der zeitaufgelösten PIV Analyse basiert auf der Kreuzkorrelation zweier aufeinanderfolgenden Bildsequenzen (Frame), welche in einem definierten Zeitabstand Δt aufgenommen wurden. Die zu analysierenden Bildpaare werden hierzu in ortsbezogene Abfragesegmente unterteilt und der jeweilige Korrelationskoeffizient als statistische Verteilung von Tracer-Verschiebung in Subpixelbereich von x - und y - Koordinaten im Abfragebereich gebildet. Nach der Analyse einer Verschiebung von x - und y - Koordinaten werden Vektoren gebildet $[px/Frame]$ und darauffolgend ein Orthorektifikationskoeffizient $[mm/px]$ auf die Vektormatrix angewendet, somit Bildkoordinaten in Weltkoordinaten mittels extrinsischer Korrekturparameter überführt. Weitere Forschungsanwendungen, wie bspw. in „LSPIV implementation for

environmental flow in various laboratory and field cases“ von Kantoush et al. [3] oder in „Large-scale particle image velocimetry for measurements in riverine environments“ von M. Muse et al. [4] oder auch Z.J. Taylor et al. [5], zeigen bereits vielfältige Möglichkeiten im Einsatz stationärer Beobachtungssysteme, bei der dieses Messverfahren auf Basis aufgezeichneter Videosequenzen zum Einsatz kam und plausible Ergebnisse bei schwierig zu messenden Umgebungssituationen erzielt wurden. Vielfältige dieser Anwendungen verwenden visuelle Kamerasysteme, welche in der Studien- oder Laborumgebung teilweise auf künstliche Tracer zurückgreifen und die Systemkalibrierung zur Orthorektifikation mittels Kontrollpunkte vor Ort mit Spezialgeräten vermessen wird. Da die großräumige Partikelbild-Geschwindigkeitsmessung empfindlich auf die vom Anwender festgelegten Parameter und Berechnungsoptionen reagiert, stellen manuelle Erfassung und Verarbeitung von Video- und Datenmaterial eine Fehlerquelle dar [6]. Es existieren weiterhin Radarsensoren zur Strömungsmessung, welche allerdings von Schiffen nicht unbedingt ökonomisch betrieben werden können [7, 8]. Und im Langzeitbetrieb von in situ hydroakustik Unterwassermesstechniken [9, 10] besteht, bspw. durch Biofouling ein erheblicher Wartungsaufwand.

Vorangegangene LSPIV-Offshore-Anwendungsstudien zeigen diesbezüglich bei stationären Langzeitbeobachtungen, dass selbst stark ungleichmäßige Strömungsverhältnisse mit Rückströmungen, starken Querströmungen usw., räumlich und zeitlich hochaufgelöst gemessen und visualisiert werden können. Von der Umgebungsbeleuchtung unabhängige Beobachtungen wurden hierbei mittels Infrarotkamera ermöglicht. Auf Basis dessen wurden nützliche Informationen für weitere Ausbaustufen und Algorithmen für ein integriertes LSPIV Fernerkundungssystem gesammelt [11, 2].

Für Anwendungen im regulären Fährbetrieb wird deutlich, dass Erweiterungen zur Automatisierung der Orthorektifikation in Bezug auf den Eigenvektor mit Translation und Rotation des an Bord befindlichen, mobilen LSPIV-Sensorsystems, sowie Funktionen zur Echtzeitanalyse und Datenfusion, notwendig und Gegenstand unserer aktuellen Forschung sind, auf die wir in diesem Beitrag eingehen. Nachfolgend wird eine Systemarchitektur sowie Datenstromverarbeitung und -fusion mit den erweiterten, implementierten Ansätzen vorgestellt, auf deren Basis eine nachrüstbare

und flexibel einsetzbare Lösung zur Strömungsmessung für unterschiedliche Anwendungsszenarien geschaffen wird.

Systemarchitektur und Funktionsweise

Mit der Architektur des LSPIV-Sensorsystems wurden flexible und erweiterbare Lösungen für den mobilen und stationären Einsatz zur Echtzeitdatenanalyse berücksichtigt und in zwei übergreifende funktionale Einheiten unterteilt: ein kamerabasiertes Sensormodul (Abb. 1, oben) mit integrierter Referenzmesstechnik zur Orts- und Lagebestimmung als inertiales Navigationssystem (Abb. 1, rechts), sowie einer dezentralen Verarbeitungseinheit (Abb. 1, unten). Um einen zuverlässigen Betrieb des Systems unter harschen und variierenden Umweltbedingungen sicherzustellen, ist die Sensortechnik in einem Outdoor-Schutzgehäuse mit Belüftung, Begleitheizung und Energieversorgung integriert. Das Sensormodul beinhaltet eine radiometrische Infrarotkamera (2), welche über Software Development Kit (SDK) angesteuert wird und kontinuierlich eine Temperaturmatrix aus dem beobachteten Sichtbereich bereitstellt. Die Temperaturmatrix wird im weiteren als Bilddatenstrom verarbeitet. Die integrierte Referenzmesstechnik (3) umfasst mehrere Sensoren zur Bestimmung der räumlichen Orientierung und Position. Über eine Inertial Measurement Unit (IMU) (3b) werden Orientierungsinformationen in Form von Roll-, Pitch- und Yaw-Winkel gemessen. Darüber hinaus stellt die IMU zusätzliche Messgrößen wie lineare Beschleunigungen und magnetische Orientierungsinformationen bereit. Über ein GNSS-Modul (3a) in Verbindung mit einer externen Antenne, werden darüber hinaus absolute Positions- und Zeitinformationen aufgezeichnet. Ergänzend dazu erfasst ein barometrischer Drucksensor (3c) den vorherrschenden Umgebungsdruck und ermöglicht die Ableitung von Höheninformationen des Sensorsystems.

Diese Referenzdaten werden mit einsprechenden Kameraparametern kontinuierlich in einer Schleife zur extrinsischen Korrektur mittels einer inversen Homographietransformation verrechnet, und bilden die Grundlage zur Orthorektifikation des Bilddatenstroms.

Eine dezentrale Sensorbox (5) bildet die Schnittstelle zum Sensormodul samt Energieversorgung und embedded PC zur Datenstromverarbeitung, und stellt über Netzwerkverbindung die ausgewerteten Strömungsinformationen für ein übergeordnetes KI-basiertes Assistenzsystem

zur optimierten Navigation auf der Fähre bereit. Spezifikationen einzelner Komponenten sind in Tab. 1 zusammengefasst.

Insbesondere für die dynamischen Einsatzbedingungen, bieten diese Datenströme somit eine Grundlage zur automatisierten Orthorektifikation in Bezug auf den Eigenvektor mit Translation und Rotation, des an Bord der Fähre befindlichen, mobilen LSPIV-Sensorsystems.

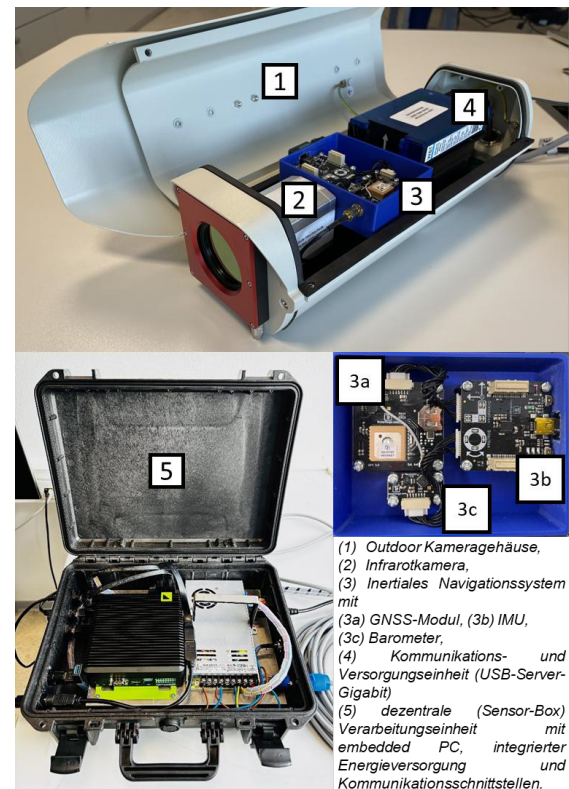


Abb. 1: Systemaufbau des integrierten LSPIV-Sensorsystems und dessen Komponenten.

Tab. 1: Übersicht der Hauptkomponenten und technischen Spezifikationen

Komponente	Typ	Spezifikationen
Radiometrische Kamera	Optris PI 640i (Mikrobolometer)	640×480 px; bis 32 Hz; 17 µm Detektor; < 40 mK NETD; ±2 °C Genauigkeit; 8–14 µm Wellenlänge
Optiken	Kalibrierte Wechselobjektive	15° x 11° / f = 41,5 mm 33° x 25° / f = 18,7 mm 60° x 45° / f = 10,5 mm 90° x 64° / f = 7,7 mm
Recheneinheit	NVIDIA Jetson Orin NX (Embedded AI Computing Modul)	8-Core ARM; 1024 CUDA Cores; AI bis 100 TOPS; 16 GB RAM; Linux OS; 10–25 W Leistung
Inertiales Navigationssystem	Tinkerforge IMU Brick 2.0 GPS 3.0 Barometer 3.0	9-DOF IMU; GNSS ±2.5 m; Barometer 300–1100 hPa; Updaterate 10 - 100 Hz

Datenstromverarbeitung und -fusion

Für die Datenstromverarbeitung und Implementierung der vielfältigen Funktionen des mobilen LSPIV-Sensorsystems wurde als Middleware das Robot Operating System (ROS 2) eingesetzt [12]. Die Open-Source-Software hat sich bereits vielfältig bei anderen Anwendungen zur Entwicklung autonomer Systeme mit Kinematik-Korrekturen, sowie der Umwelt- und Verkehrslagebilderstellung bewährt [13]. Zu Grunde liegt hierbei eine modulare Systemarchitektur mit Sensor-, Middleware- und Anwendungsebene um eine unmittelbare, leistungsfähige und echtzeitnahe Verarbeitung der Sensordaten am Einsatzort zu gewährleisten.

Durch die modulare Integration der Sensoren und Verarbeitungskomponenten innerhalb der ROS2-basierten Systemarchitektur können einzelne Komponenten und deren datenabhängige Softwaremodule unabhängig voneinander erweitert oder ersetzt werden, ohne dass sich die grundlegende Systemstruktur verändert. Eine reproduzierbare Integration wird darüber hinaus durch die Nutzung von Docker-Images unterstützt. Eine dezentrale, fehlertolerante und skalierbare Kommunikation zwischen den Systemkomponenten erfolgt über den Data Distribution Service (DDS) als standardisiertes Middleware-Protokoll. Diese Intra-Prozess-Kommunikation erlaubt einen optimierten Datenaustausch innerhalb desselben Prozesses, wodurch sich die Latenz reduziert und die Effizienz der Datenverarbeitung erhöht.

Abbildung 2 zeigt die entwickelte Datenstromverarbeitungs-pipeline mit den Verarbeitungsschritten auf Basis des UNIX-Zeitstempels. Zu Beginn erfolgt eine automatische Initialisierung der Sensoren (a). Während des Betriebs werden die Sensordaten zur Bestimmung von Position, Orientierung, Höhe und Zeit synchron erfasst und bereitgestellt. Parallel dazu wird die Temperaturmatrix des beobachteten Sichtbereichs kontinuierlich erfasst und mittels Temperatur-Mapping verarbeitet. Hierbei wird die Temperaturmatrix in ein radiometrisches 16-Bit-Bildformat überführt indem die physikalischen Temperaturmesswerte (in°C/K) auf einen Intensitätsbereich von 0 bis 65535 skaliert werden, so dass auch geringe Temperaturunterschiede zuverlässig erfasst werden. Dies ist insbesondere für die LSPIV-Analyse von Vorteil, da stabile Intensitätsgradienten die Qualität der Kreuzkorrelation und somit die Bestimmung der Strömungsvektoren verbessern.

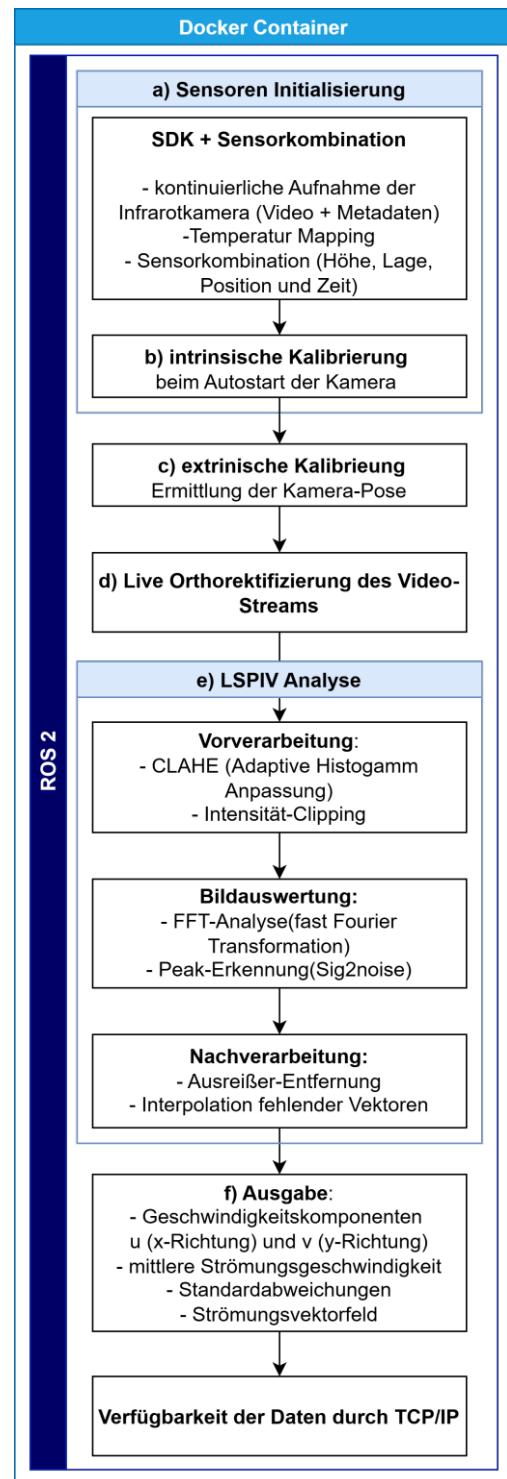


Abb. 2: Funktionsübersicht der integrierten Datenstromverarbeitung zur Echtzeitdatenverarbeitung von Sensorinitialisierung und Kalibrierung über die Auswertalgorithmik bis zur Bereitstellung der LSPIV-Strömungsmessdaten via TCP/IP in einem Netzwerk für die KI-Anwendung des Routen-Assistenzsystems auf dem Fährschiff.

Im nachfolgenden Schritt erfolgt die intrinsische Kalibrierung (b) und die automatische extrinsische Kalibrierung der Kamera zur Ermittlung der Kamerapose (c) und Orthorektifizierung der Bilddatenströme (d). Die extrinsische Kalibrierung umfasst die Bestimmung der Position sowie der Orientierung (Rotation und Translation) der Kamera im Weltkoordinatensystem. Sie bildet mit der intrinsischen Kalibrierung die Grundlage für eine Orthorektifikation. Das Ziel der extrinsischen Kalibrierung besteht in der Korrektur projektiver Verzeichnungen der Bilddaten hin zu einer maßstabsgerechten orthogonalen Ansicht. Darüber hinaus bietet sich auch die Möglichkeit ein einheitliches Weltkoordinatenmodell für multi-kamera-Beobachtungen zur Verfügung zu stellen, die bei weiteren Anwendungen [14] genutzt werden können. Über die automatische extrinsische Kalibrierung werden Position und Orientierung der Kamera in ein ENU (East, North, Up) Weltkoordinatensystem überführt [15], und bildet zusammen mit der intrinsischen Kalibrierung die Grundlage der Orthorektifikation, ohne die Notwendigkeit spezifisch festgelegte oder zuvor vermessene Kontrollpunkte aus der beobachteten Szene für die LSPIV-Analyse anzugeben. Hierbei wird mittels einer inversen Rotations- und Homographie-transformations-Matrix [16, 17] ein diskretes Pixelraster eines Bildes auf die planare Bodenebene transformiert, womit eine maßstabsgerechte orthogonale Darstellung berechnet, sowie ein einheitlicher Arbeitsbereich als Ground Bounding Box für die folgende LSPIV-Echtzeitanalyse zur Verfügung gestellt wird.

Die folgende LSPIV-Analyse (e) basiert auf dem orthorektifizierten Bilddatenstrom. Die Kernfunktionen werden hierbei über OpenPIV v. 0.25.3 [18], einer freien und quelloffener Software, via Python abgebildet. Bei der LSPIV-Analyse erfolgt zunächst eine Bilddatenvorverarbeitung, mit dem Ziel den Bildkontrast von relevanten Tracern durch eine adaptive Histogrammanpassung zu verbessern und gleichermaßen störende Intensitätseffekte zu reduzieren. Im nächsten Schritt werden zwei aufeinanderfolgende, zeitlich korrelierende Bildpaare mittels Kreuzkorrelation in Ortsbezogenen Such- und Abfragesegmente unterteilt und ausgewertet. Dabei kommt die schnelle Fourier-Transformation zum Einsatz. Die resultierende Korrelation wird hinsichtlich des Signal-Rausch-Verhältnisses bewertet, und valide Verschiebung in Subpixelbereich von x- und y-Koordinaten im Abfragebereich zur Vektorbildung verwendet. Mit der

anschließenden qualitätsorientierten Nachverarbeitung werden fehlerhafte Vektoren identifiziert und entfernt sowie ggf. fehlende Werte über Interpolationsverfahren ergänzt. Das Resultat ist ein konsistentes Strömungsvektorfeld, auf dessen Basis statistische Kenngrößen zur mittleren Strömungsgeschwindigkeit als auch -Richtung sowie die zugehörige Standardabweichung kontinuierlich mit Zeitstempel in einer Zeitmessreihe zur Verfügung stehen. Die berechneten Geschwindigkeiten und dazugehörigen statistischen Kenngrößen werden abschließend in einem standardisierten Datenformat über eine TCP/IP-basierte Schnittstelle an das übergeordnete KI-basierte Assistenzsystem für die Anwendung im Fährbetrieb bereitgestellt.

Verifikation des Systems

Im Rahmen der experimentellen Verifikation des entwickelten LSPIV-Systems wurde am Institut für Chemie und Biologie des Meeres (ICBM, Wilhelmshaven) eine Messkampagne unter kontrollierten Laborbedingungen durchgeführt. Das Ziel bestand darin, die gesamte Systemfunktionalität zur Erfassung von Strömungsvektorfeldern zu beurteilen und LSPIV-Programmparameter zur Berechnung, Verarbeitung und Analyse der Strömungsvektorfelder frühzeitig vor dem Einsatz an Bord zu optimieren. Der Versuchsaufbau ist in Abb. 3 dargestellt.



Abb. 3: Übersicht experimentelle Verifikation des entwickelten LSPIV-Systems. (A) kamerabasiertes Sensormodul oberhalb des überdachten Strömungsbeckens (B); (C) Übersicht des Messaufbaus mit ADCP in situ Referenzmessung; (D) dezentrale Verarbeitungseinheit innerhalb der Containeranlage.

Darüber hinaus sollte der Einfluss geometrischer Parameter (Kamerahöhe, Kamerapose, Objektivwechsel) auf die Qualität der Orthorektifizierung beurteilt, und die so erhaltenen LSPIV-Strömungsmessdaten mit in situ Referenzmessdaten verglichen werden.

Die Kamera wurde oberhalb eines Strömungsbeckens in ca. 4,3 m Höhe montiert, wodurch die Wasseroberfläche aus definierter Höhe und Blickrichtung beobachtet werden konnte. Die Aufnahme und Verarbeitung der Bilddatenströme erfolgte über die dezentrale Recheneinheit, welche innerhalb der Containeranlage betrieben wurde. Simultan wurden in situ Referenzmessungen im Strömungsbecken ca. 20cm unterhalb der Wasseroberfläche mit einem kompakten ADCP (Acoustic Doppler Current Profiler, Aquadopp2, Nortek Group) erhoben und aufgezeichnet.

Die resultierenden Bilddatenströme wurden über die zuvor beschriebene LSPIV-Verarbeitungspipeline (Abb. 2) verarbeitet und entsprechende Strömungsvektorfelder berechnet (vgl. Abb. 4). Zu Beginn des Experiments wiesen die Vektorfelder eine hohe Anzahl an falsch-positiven Daten auf (Abb. 2 A), welche dann über die nachgelagerten Filterprozesse schrittweise bereinigt werden (Abb. 2 B). Die gewählten LSPIV-Filtereinstellungen waren ($\Delta t = 0,035$ s) mit Signal-Rausch-Verhältnis ($\text{sig2noise} = 1,4$), adaptive Kontrastanpassung (CLAHE-Clipping = 11.000–15.000 bei 16 Bit), Vektor-Median-Filter (Median-Threshold = 2,0), sowie Größe korrespondierender Abfragebereiche (Interrogation Areas = 38 px.). Nach der Parameteroptimierung zeigt das Endergebnis (C) ein konsistentes und physikalisch plausibles Strömungsfeld mit ausschließlich validen Strömungsvektoren.

Im Rahmen der experimentellen Verifikation wurden ebenfalls ergänzende Tests bei variierenden geometrischen Randbedingungen (Kamerahöhe, Blickwinkel, Objektivwahl) durchgeführt, um den Einfluss auf die Qualität der Strömungsmessung systematisch zu untersuchen. In diesem Zusammenhang wurden Messungen mit unterschiedlichen Objektiven (15°, 33°, 60° und 90° Sichtfeld) durchgeführt, um der Zusammenhang zwischen dem Öffnungswinkel und der Bildabdeckung auf die resultierende Messgenauigkeit zu analysieren. Zusätzlich wurde ein Nadir-Test mit dem 90°-Objektiv durchgeführt. Die Ergebnisse zeigen, dass bei einer Kamerahöhe von 1,60 m das 90°-Objektiv die besten Resultate liefert. Dies ist insbesondere auf die größere abgedeckte Fläche sowie die verbesserte Erfassung von Partikelbewegungen innerhalb des Sichtfeldes

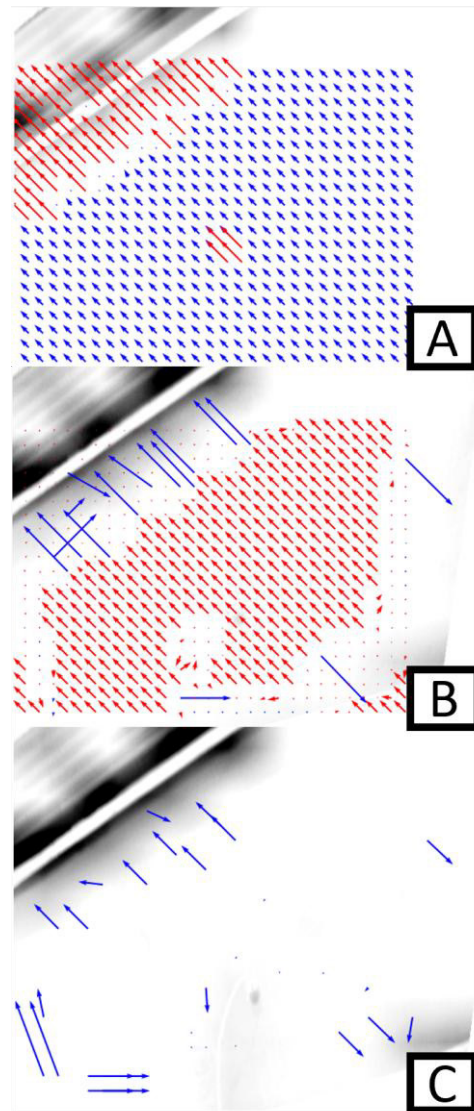


Abb. 4: Beispielübersicht einzelner Verarbeitungsstufen und Ergebnisse einer LSPIV-Vektorfeldberechnung (in vergrößerter Darstellung, als Auszug des beobachteten Bereiches im Strömungsbecken). (A) Vektorfeld mit zahlreichen falsch-positiven Vektoranteilen. (B) Darstellung gültiger (valider) Vektoren im Vergleich zu verworfenen (invaliden) Vektoren und als Ergebnis der Nachbearbeitung das gefilterte Vektorfeld (C).

zurückzuführen. Die Wahl der geeigneten Optik erweist sich somit als entscheidender Faktor für die Leistungsfähigkeit des Systems und steht in direktem Zusammenhang mit den gewählten geometrischen Randbedingungen. Zur Bewertung der Leistungsfähigkeit des entwickelten LSPIV Fernerkundungssystems wurde der Mittelwert aller im Sichtbereich

vorkommenden Oberflächenströmungsgeschwindigkeiten gebildet und mit in situ Referenzdaten des ADCP-Sensors verglichen. Die Darstellung in Abbildung 5 zeigt den Vergleich der aus dem LSPIV-Sensorsystem (rot) und dem ADCP Referenzsensor (blau) erhobenen Strömungsdaten nach der LSPIV-Parameteroptimierung. Die beispielhaft dargestellte Zeitreihe der Strömungsgeschwindigkeiten (siehe Abbildung 5 A) zeigt im Vergleich, dass die LSPIV Messwerte stellenweise höhere Amplituden gegenüber den in situ ADCP Messdaten aufweisen. Diese lokale Überschätzung der LSPIV Messdaten spiegelt sich in der mittleren Abweichung von $MAD=0,016\text{ m/s}$ wieder, und weist auf eine geringfügige systematische Überbewertung des LSPIV-Verfahrens hin. Die Regression mit $RMSE = 0,086\text{ m/s}$ wurde über den beispielhaft betrachteten Zeitraum quantifiziert, und die mittlere Abweichung zwischen beider Messdatenreihen entspricht je nach Strömungsgeschwindigkeitsniveau einem relativen Fehler von etwa $\delta = 8\text{--}17\%$. Dieser Wert kann für LSPIV-Anwendungen im Feldeinsatz als akzeptabel bis gut bewertet werden, und entspricht den Erwartungen bei den hier im Vergleich gesetzten oberflächennahen Strömungen und den Strömungsverhältnissen in der Wassersäule. Eine weitgehende Übereinstimmung der vorherrschenden Geschwindigkeitsverteilung aus beiden Datenreihen ist bei der Überschneidung der Histogramme evident (siehe Abbildung 5B). Das LSPIV-Histogramm weist jedoch leicht verbreiterte Werteverteilungen ($> 0,3\text{ m/s}$) auf, die vor allem durch unvollständige Partikelverfolgung, sowie kamerabedingtes Rauschen verursacht werden können. Diese Beobachtung steht im Einklang mit der zuvor berechneten RMSE, und die Verteilung bestätigt eine geringe mittlere Abweichung. Bei der Betrachtung des Zeitverlaufs der Strömungsrichtung (siehe Abbildung 5C) zeigen die Werte beider Datenreihen eine deutliche Streuung ohne erkennbaren Zusammenhang. Dies war bei dem hier verwendeten kleinen geschlossenen Strömungsbecken zu erwarten, bei dem keine definierte Strömungsrichtung eingestellt werden konnte. In der Versuchsumgebung entstehen hierdurch, bspw. durch Wellenrücklauf und lokal turbulente Oberflächenstrukturen, instationäre Strömungsbedingungen. Insgesamt zeigen die durchgeführten Untersuchungen, dass die LSPIV-Ergebnisse quantitativ mit der Referenzmessung überein-

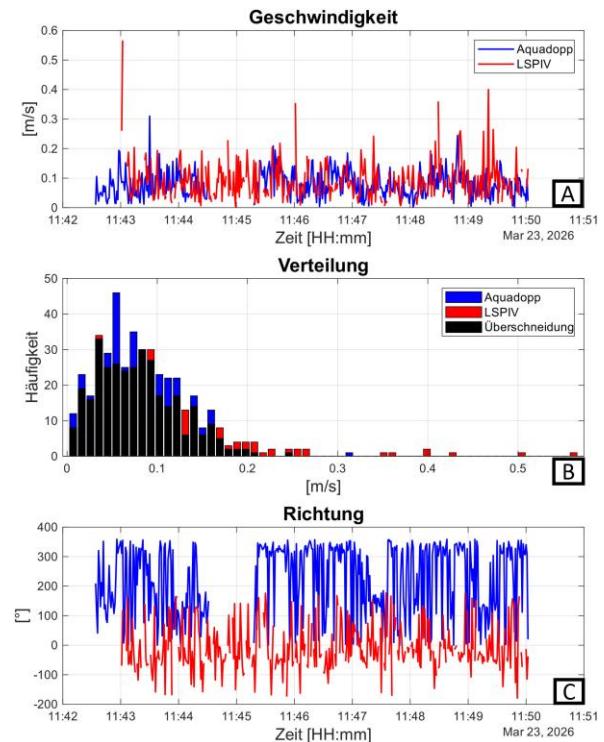


Abb. 5: Strömungsgeschwindigkeit (A) und Histogramm (B) zur Geschwindigkeitsverteilung sowie Strömungsrichtung (C) von LSPIV und in situ ADCP Referenzmessdaten über einen beispielhaften Zeitraum.

stimmen, während die Richtungsmessung aufgrund lokaler Turbulenzeinflüsse weniger stabil ist. Die Messkampagne lieferte somit zentrale Erkenntnisse über die Leistungsfähigkeit des entwickelten LSPIV Fernerkundungssystems. Es konnte gezeigt werden, dass die integrierten Funktionen und Analyseverfahren grundsätzlich in der Lage sind, physikalisch konsistente Strömungsvektorfelder zuverlässig zu messen und diese als mittlere Strömungsgeschwindigkeit und -richtung über einen Zeitverlauf auszugeben. Zudem wurde deutlich, dass sowohl die geometrischen Randbedingungen (Kamerahöhe, Blickwinkel, Objektivwahl) als auch die Sensitivität gegenüber der LSPIV-, Bildvorverarbeitungs- und Filterparameter entscheidend für die Qualität der resultierenden Strömungsvektorfelder sind.

Zusammenfassung und Ausblick

Die Kombination aus präziser Orthorektifikation und robuster LSPIV-Analyse bildet die Grundlage für weiterführende Validierungen im Feldeinsatz und mobilen Betrieb unterwegs auf einer Fähre unter realen Umweltbedingungen und eröffnet gleichzeitig ein großes Optimierungspotenzial mit vielversprechenden

Ergebnissen. Die Erprobung im Feld unterwegs auf dem Fluss Weser ist hierzu über einen Erprobungszeitraum von etwa 10 Monaten parallel zu einer Referenzmessstelle vorgesehen. Mit dem Einsatz im regulären Fährbetrieb sind weiterführende Automatisierungsansätze und Optimierungen insbesondere hinsichtlich variierender Umweltbedingungen zur robusten Bilddatenvorverarbeitung und Fehlertoleranz geplant um eine nachrüstbare und flexibel einsetzbare Lösung zur Strömungsmessung in mobilen und stationären Anwendungsszenarien zu ermöglichen.

Beitrag der Autoren und Danksagungen

NR war für Hauptkonzept und Architektur des Systems verantwortlich und entwarf und schrieb die Arbeit, zu der die weiteren Autoren beitrugen. AN war für die Implementierung, Systemaufbau sowie Verifizierung verantwortlich. Die Autoren bedanken sich bei der Arbeitsgruppe Prozesse und Sensorik mariner Grenzflächen am ICBM, Universität Oldenburg für die Unterstützung bei der Messkampagne am Strömungskanal. Die Jade Hochschule dankt der Deutschen Bundesstiftung Umwelt für die finanzielle Unterstützung im Rahmen der Projektförderung „Eco-Crossing“ (DBU Projekt 39313/01)

Literaturnachweis

- [1] Y. Ender, et al., Optimising ship manoeuvring time during port approach using a decision support system: a case study in Turkey. *Ships and Offshore Structures* 18.12, 1713-1725 (2023); doi: 10.1080/17445302.2022.2140522
- [2] N. Ruessmeier, O. Zielinski, Ocean current observations by infrared and visual Large Scale Particle Image Velocimetry (LSPIV). *OCEANS 2021: San Diego-Porto*, 1-10 (2021); doi: 10.23919/OCEANS44145.2021.9705808
- [3] S. A. Kantoush, et al., LSPIV implementation for environmental flow in various laboratory and field cases. *Journal of Hydro-environment Research* 5(4), 263-276 (2011); doi: 10.1016/j.jher.2011.07.002
- [4] M. Muste, et al., Large-scale particle image velocimetry for measurements in riverine environments. *Water resources research* 44.4 (2008); doi: 10.1029/2008WR006950
- [5] Z.J. Taylor, R. Gurka, G.A. Kopp, A. Liberzon, Long-Duration Time-Resolved PIV to Study Unsteady Aerodynamics, Instrumentation and Measurement, *IEEE Transactions* 59 (12), 3262-3269 (2010); doi: 10.1109/TIM.2010.2047149
- [6] G. Bodart, et al., Quantifying and reducing the operator effect in LSPIV discharge measurements. *Water Resources Research* 60.2 (2024); doi: 10.1029/2023WR034740
- [7] K. G. Hessner, et al., Nautical radar measurements in Europe: Applications of WaMoS II as a sensor for sea state, current and bathymetry. In *Remote Sensing of the European Seas* 435-446 (2008); doi: doi.org/10.1007/978-1-4020-6772-3_33
- [8] K. Reichert, et al., WaMoS II: A radar based wave and current monitoring system. *ISOPE International Ocean and Polar Engineering Conference*. ISOPE-I-99-246 (1999)
- [9] H. Rubio-Arias, et al., ADCP, multi-frequency analysis for flow measurements in rivers. *Acta Scientiarum Polonorum. Formatio Circumiectus*, 18.1, 101-112 (2019); doi: 10.15576/ASP.FC/2019.18.1.101
- [10] A. Vergne, et al., Acoustic backscatter and attenuation due to river fine sediments: Experimental evaluation of models and inversion methods. *Water Resources Research* 57.9 (2021); doi: 10.1029/2021WR029589
- [11] N. Rüssmeier, et al., Ocean surface water currents by large-scale particle image velocimetry technique. In *OCEANS 2017- Aberdeen*, 1-10 (2017); doi: 10.1109/OCEANSE.2017.8084843
- [12] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, W. Woodall, Robot Operating System 2: Design, architecture, and uses in the wild, *Science Robotics* vol. 7, (2022); doi: 10.1126/scirobotics.abm6074
- [13] A. Lejman, N. Rüssmeier, O. Ferdinand, & F. Stahl, Eine integrierte Datenstromverarbeitung zur Erfassung von Umweltlagebildern. *Tagungsband: 22. ITG/GMA-Fachtagung Sensoren und Messsysteme*, 499-504 (2024); doi:10.5162/sensoren2024/P20_499-504.
- [14] R. Tsubaki, Multi-camera large-scale particle image velocimetry. *Measurement Science and Technology*, 31(8) 084004 (2020); doi: 10.1088/1361-6501/ab85d5
- [15] ASAM e.V., OpenSCENARIO DSL – Coordinate Systems, 2026. [Online]. Verfügbar unter [8.3 Coordinate systems :: ASAM OpenSCENARIO](#). [Letzter Zugriff: 14. April 2026]
- [16] D. J. Justs, R. Novickis, K. Ozols and M. Greitāns, Bird's-eye view image acquisition from simulated scenes using geometric inverse perspective mapping, *17th Biennial Baltic Electronics Conference (BEC) Tallinn, Estonia*, 1-6 (2020); doi: 10.1109/BEC49624.2020.9277042
- [17] J. Liebowitz, A. Zisserman, Metric rectification for perspective images of planes. In *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 482-488 (1998); doi: 10.1109/CVPR.1998.698649
- [18] A. Liberzon, R. Gurka, and Z. Taylor, OpenPIV Home Page, (2009) [Online]. Verfügbar unter: <http://www.openpiv.net> [Letzter Zugriff: 14. April 2026]

Enable Transferability of Interpretable ML Models for Bearing Diagnostic in Changing Operating Conditions

Julian Schauer^{1,2}, Houssam El Moutaouakil^{1,2}, Andreas Schütze²

¹*Centre for Mechatronics and Automations Technology, Saarbrücken, Germany*

²*Lab for Measurement Technology, Saarland University, Saarbrücken, Germany*

(julian.schauer),(h.moutaouakil)@zema.de, schuetze@lmt.uni-saarland.de

Abstract

In industrial machine learning applications, changing environmental and operating conditions significantly reduce the robustness and predictive quality of machine learning models, leading to performance drops under domain shifts. Interpretable machine learning algorithms based on physically meaningful features allow us to build models that are more robust to data distribution shifts. However, for significant domain shifts, the model should be able to adapt to the novel domain to avoid a performance decrease. Especially in industrial settings, the influences of external sources are hard to control. This study utilizes a novel approach to represent interpretable machine learning algorithms as deep neural networks, enabling transfer learning to adapt to novel domains. In this study, an interpretable machine learning algorithm comprising feature extraction, feature selection, and classification was developed to predict the condition of a roller bearing using accelerometer data. The novel approach demonstrates the ability to adapt the model via transfer learning to a new, unseen rotational speed with only 32 new transfer samples. While keeping information from the source domain out, the novel approach increases the model's accuracy from 77.6% to 100% in the unseen operating condition.

Keywords: Smart Sensors, Interpretable Machine Learning, Transfer Learning, Domain Shifts

1. Introduction

Reliable bearing diagnostics are essential for ensuring the safety, efficiency, and longevity of rotating machinery in a wide range of industrial applications. The degradation of roller bearings directly affects overall system performance and maintenance costs [1]. The introduction of Predictive Maintenance (PM) and Condition Monitoring (CM) based on Machine Learning (ML), along with appropriate sensors, enables efficient and reliable diagnosis of bearing conditions [2]. However, data-driven approaches like ML are sensitive to variations in the underlying data, which may result from changing operational or environmental conditions [3]. This is particularly relevant for roller bearings, which often operate under varying loads, speeds, and unpredictable environmental conditions. Developing diagnostic models that can adapt to these new domains with minimal additional data is therefore a key challenge for achieving robust, scalable PM. An approach to building robust and interpretable ML models is to use Feature Extraction (FE), Feature Selection (FS), and Classification (C) algorithms (FESC) [4]. These algorithms are based on physically interpretable features and have proven to achieve robust predictions in several applications [3]. Especially

on unseen data, the FESC model shows reliable performance. Nevertheless, when FESC models face significant domain shifts, they lack transferability to new, unseen domains, without building a new model from scratch with data from the novel domain. This results in a high-effort collection of new datasets from the target domain to build a novel trained model.

In the field of Deep Neural Networks (DNNs), Transfer Learning (TL) has emerged as a powerful strategy to address this challenge [5]. TL enables knowledge learned from an initial domain to be effectively reused in a related target domain by fine-tuning the initial model. However, in industrial environments interpretable and trustworthy ML models are required to avoid unwanted false predictions. The “Black-Box” nature of DNNs with very limited interpretability and their limited robustness often prevent their use in industrial environments [6].

For this purpose, a novel approach to represent the inference of interpretable FESC models as DNN (IDNNRep) was recently introduced [7]. Primarily, this approach enables the implementation of FESC inference after offline training on generic AI accelerators, thereby reducing inference time and leading to a more energy-efficient implementation [8]. Furthermore, this

approach allows combining the interpretable and robust characteristics of FESC models with recently available TL techniques from the field of DNNs to adapt models to new and unseen domains with a small number of transfer samples.

This study investigates the retraining of specific parts of an FESC model using IDNNRep to adapt to new, unseen operating conditions of a roller bearing, while preserving the interpretable nature and physical meaning of the engineered features. For this purpose, data from a roller bearing testbed are investigated by varying the testbed's rotational speed to create domain shifts.

2. Methods

2.1 Data

The paper investigates a roller bearing testbed designed to assess the robustness of ML models. Figure 1 shows the testbed to collect the dataset. The dataset [9], [10] contains data from a triaxial accelerometer collected from damaged and undamaged roller bearings at a sampling frequency of 20 kHz. Three roller bearings were tested at four different mounting positions within three separate runs. Each run contains four force levels, each operated at four distinct operational speeds. Thus, the complete dataset contains 1151 measurement files, each containing a one-minute measurement from the three-axis accelerometer for a specific combination of covariates. In this study, only the accelerometer axis vertical to the shaft is used. The covariate variations are shown in Table 1.

Table 1: Covariates for the testbed included in the dataset

Parameter	Quantity	Values
Bearing	3	B10, B20, B30
Damage State	2	No Damage, Small Damage
Run	3	R1, R2, R3
Position	4	PA, PB, PC, PD
Force Level [N]	4	0, 1600, 2500, 3000
Rotational Speed [rpm]	4	85, 392, 706, 969

Table 1 also highlights the importance of dataset variability, when investigating ML models in the context of domain shifts in industrial applications. By trying to capture most of the variability of a complex system, the dataset and the experimental design become significantly larger, often resulting in incomplete coverage of that variability. This study focuses on the changing rotational speed, which represents one possible change in the operating condition. The introduced dataset allows keeping variations as small as possible by changing only the rotational speed. This preserves the influence of different variations without disturbing the model, allowing it to focus on predicting damage at the defined operating rotational speed.

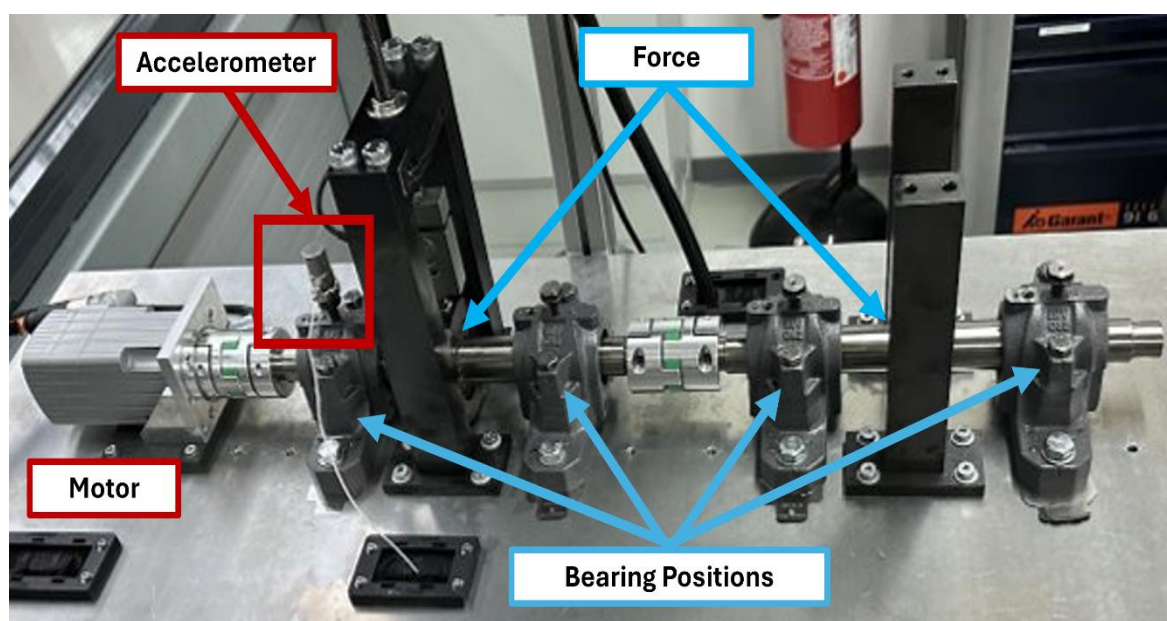


Figure 1: Roller bearing testbed with triaxial accelerometer

2.2 IDNNRep

The IDNNRep represents the inference equation of a trained FESC algorithm using DNN layers. For that purpose, the inference equation of the FESC algorithm is broken down into basic mathematical operations, such as addition, subtraction, multiplication, and division, as well as DNN-specific operations, such as convolution and pooling. The FESC algorithm, which is used to predict the roller bearing condition based on the data of an accelerometer, consists of the following parts and is included in the LMT Toolbox [11]:

- **Best Fourier Coefficient (BFC):** The BFC algorithm calculates the amplitudes and phases of the spectrum for each input time series and returns the 10% of amplitudes with the highest average absolute value and their corresponding phases. The 10% represent a suitable trade-off between a low approximation error and a low number of features [12].
- **Pearson Correlation:** The Pearson correlation algorithm calculates the linear relationship between the extracted features

and the target. The number of features to keep was defined to 20[4].

- **Linear Discriminant Analysis and Mahalanobis distance classifier (LDAMahal):** The LDAMahal algorithm first performs a linear projection of the feature space into the lower-dimensional LDA [13] subspace. After performing the LDA, the Mahalanobis distance [14] to each class center is calculated. The algorithm then classifies the input data based on the smallest calculated Mahalanobis distance.

Each part of the described FESC algorithm can be represented as a DNN; the detailed description is provided in [7]. The final FESC algorithm is then simply a concatenation of the layer graphs for each part. As shown in Figure 2, the final IDNNRep can be represented with the basic mathematical layers. This IDNNRep lays the foundation for the following TL to adapt the model to a novel domain.

2.3 Transfer Learning

Especially in industrial applications, environmental and operating conditions can change,

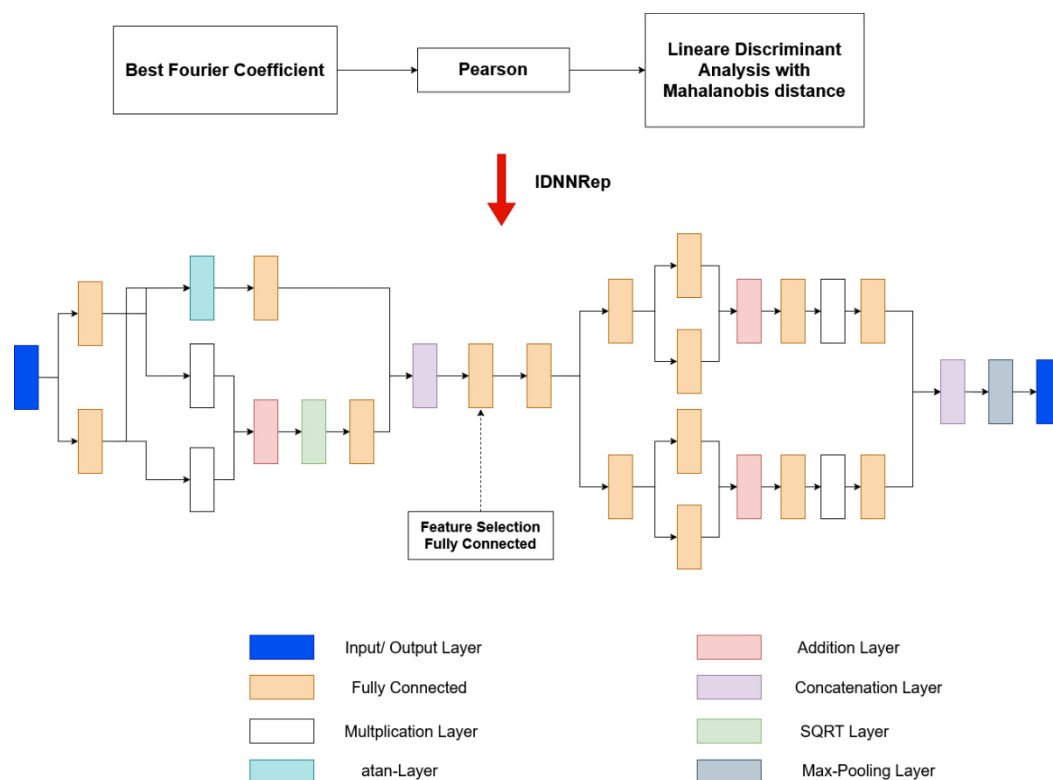


Figure 2: IDNNRep of the FESC algorithm

resulting in significant domain shifts. To avoid a decrease in model performance, it is necessary to retrain the underlying ML model to the novel domain. Besides data-based transfer techniques, especially TL, have shown significant benefits when domain shifts occur. This study focuses on the change in the operating condition of a roller bearing. When the rotational speed changes, the ML model that predicts the roller bearing's damage state experiences a significant domain shift. The goal is to retrain the IDNNRep of the FESR model using a small number of samples from the new domain to adapt to it, as shown in Figure 3.

For this purpose, an initial model was trained on accelerometer data at 706 rpm and 969 rpm and subsequently tested on 50% of the 396 rpm data, which were held out as validation data. To minimize data variation, only one roller bearing was used. Also, the other covariates, run (R1), position (PA), and the force level (0 N), were kept constant. After training the initial model, it is converted to an IDNNRep, and only the Feature Selection IDNNRep (one fully connected layer) part is retrained on a data subset from the target domain at 396 rpm. This allows us to retain information from the initial domain and fine-tune only a small number of parameters for the new domain.

3. Results and Discussion

To demonstrate the significance of the domain shift, a short investigation of the data is shown.

For that purpose, a Principal Component Analysis (PCA) is calculated based on the extracted features of the BFC. Figure 4 demonstrates the clear separation between the three rotational speeds. It also shows a clear separation between the damaged and undamaged data. As this study focuses on transferring an interpretable ML model to a new domain, this data shows good characteristics for such a problem.

To test the retraining of the IDNNRep, 50% of the data at 396 rpm is left out during training. These data are used to test the initial model and compare this with the retrained IDNNRep on the new domain. For retraining in the new domain, the amount of transfer samples was subsequently increased. The number of transfer samples investigated in this study was 0 (initial model), 8, 16, and 32. For retraining, all learnable parameters of the IDNNRep were frozen, except the weights and biases of the fully connected layer, which represent the FS. This layer was trained with an initial learning rate of 10^{-3} and a learning rate drop of 0.9 for 10 epochs. Table 2 lists the testing results of the initial model and the TL of the IDNNRep, with N denoting the number of samples used for finetuning. The initial model still achieves acceptable performance on an unseen operating speed with an accuracy of 77.6%, again demonstrating the robustness of FESC models under domain shifts. However, in most industrial applications an accuracy of 77.6% would be too low for reliable bearing diagnostic.

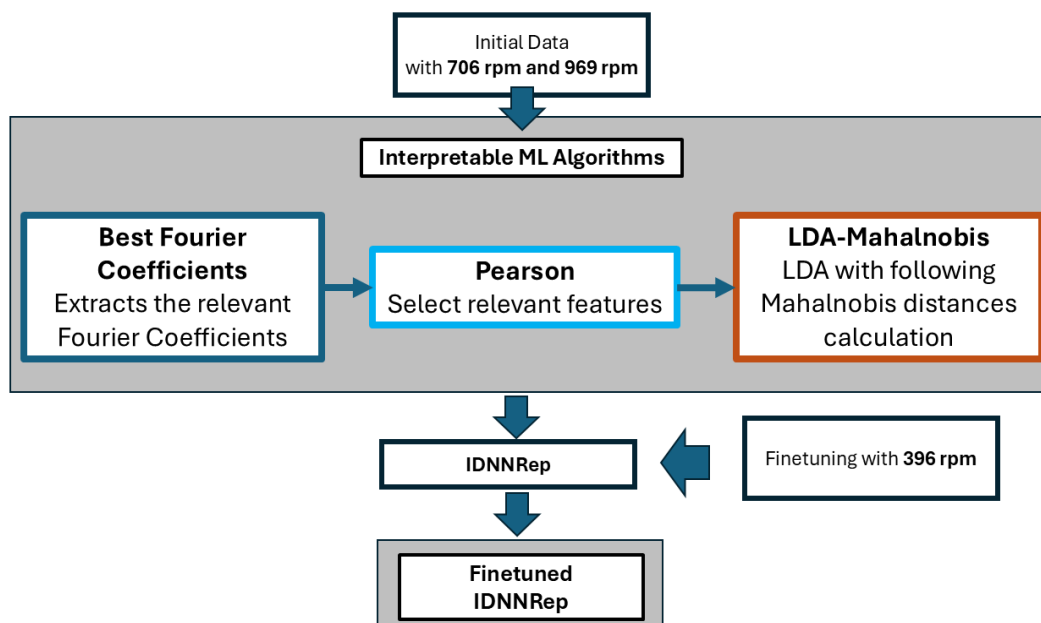


Figure 3: Design of Experiment for retraining the IDNNRep

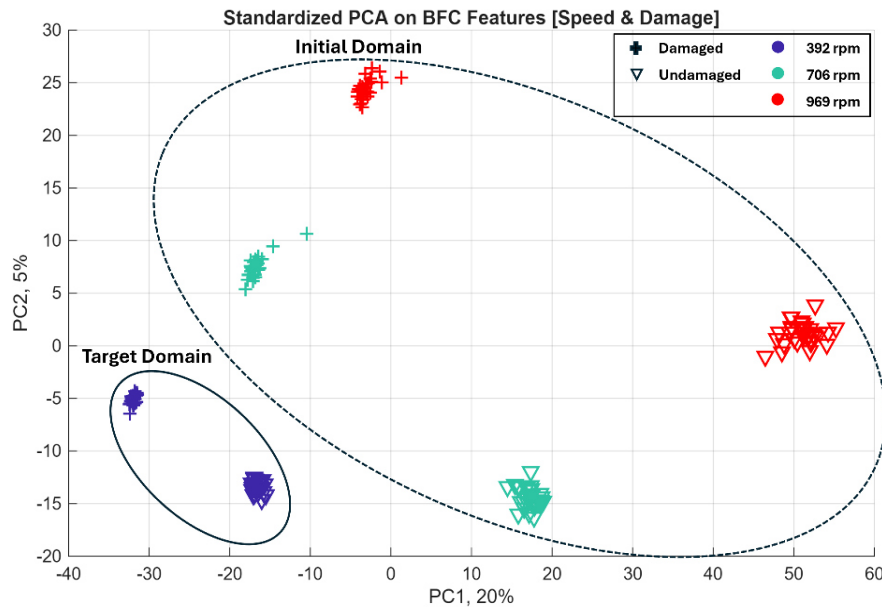


Figure 4: Standardized PCA of BFC Features

Retraining the initial model as an IDNNRep with 8 transfer samples already increases accuracy by 15.7% to 93.3%. This demonstrates a significant improvement compared to the initial model. Increasing from 8 transfer samples to 16 achieves an accuracy to 95.4%, an increase of only 2.1% compared to the retrained IDNNRep with 8 transfer samples. Using 32 transfer samples, the IDNNRep achieves 100% accuracy and yields the best possible result under the tested conditions.

Table 2: Validation Accuracy

	Accuracy (Testing)
Initial Model (N = 0)	77.6 %
TL (N=8)	93.3 %
TL (N=16)	95.4 %
TL (N=32)	100 %

4. Conclusion and Future Work

This study demonstrates a novel approach for model-based transferability for interpretable ML models. With the introduction of the IDNNRep, it is possible to retrain an initial interpretable ML after a significant domain shift. Even with only 8 transfer samples, TL increases the initial model's accuracy from 77.6% to 93.3%. By including 32 transfer samples, the model achieves 100%

accuracy, demonstrating the significant improvement of the new approach.

This demonstrates that the presented approach can be used to adapt interpretable ML models to novel domains, which presents an important challenge in the field of Machine Learning. Since this study focuses on a relatively simple domain shift, future investigations should include more complex domain shifts, also including additional covariates from the dataset. Additionally, more effective TL strategies and additional hyperparameter tuning offer the potential for further improvements.

Acknowledgement

This paper was written in the context of the VProSaar project which is funded by the Ministry of Economic Affairs, Innovation, Digital and Energy (MWIDE) and the European Funds for Regional Development (EFRE). The project is located in the federal state of Saarland and is executed by ZeMA – Centre for Mechatronics and Automation Technology gGmbH.

References

- [1] R. V. Daniel, S. A. Siddhappa, S. B. Gajanan, S. V. Philip, and P. S. Paul, "Effect of Bearings on Vibration in Rotating Machinery," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 225, no. 1, p. 012264, Aug. 2017, doi: 10.1088/1757-899X/225/1/012264.
- [2] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mechanical Systems and Signal Processing*, vol. 25, no. 2,

- pp. 485–520, Feb. 2011, doi: 10.1016/j.ymssp.2010.07.017.
- [3] P. Goodarzi, A. Schütze, and T. Schneider, “Domain shifts in industrial condition monitoring: a comparative analysis of automated machine learning models,” *J. Sens. Syst.*, no. 14, pp. 119–132, 2025.
 - [4] T. Schneider, N. Helwig, and A. Schütze, “Automatic feature extraction and selection for classification of cyclical time series data,” *tm-Technisches Messen*, vol. 84, no. 3, pp. 198–206, 2017, doi: 10.1515/teme-2016-0072.
 - [5] K. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning,” *J Big Data*, vol. 3, no. 1, p. 9, May 2016, doi: 10.1186/s40537-016-0043-6.
 - [6] V. Buhrmester, D. Münch, and M. Arens, “Analysis of explainers of black box deep neural networks for computer vision: A survey,” *Machine Learning and Knowledge Extraction*, vol. 3, no. 4, pp. 966–989, 2021, doi: 10.3390/make3040048.
 - [7] J. Schauer, P. Goodarzi, and A. Schütze, “Efficient Hardware Implementation of Interpretable Machine Learning Based on Deep Neural Network Representations for Sensor Data Processing,” *Journal of Sensors and Sensor Systems*, 2025, doi: 10.5162/sensoren2024/A1.4.
 - [8] J. Schauer, P. Goodarzi, J. Morsch, and A. Schütze, “A Performance Study of Deep Neural Network Representations of Interpretable ML on Edge Devices with AI Accelerators,” *Sensors*, vol. 25, no. 18, p. 5681, 2025, doi: 10.3390/s2518568.
 - [9] C. Schnur, P. Goodarzi, Y. Robin, J. Schauer, and A. Schütze, “A Machine Learning Dataset of Artificial Inner Ring Damage on Cylindrical Roller Bearings Measured Under Varying Cross-Influences,” *Data*, vol. 10, no. 5, p. 77, May 2025, doi: 10.3390/data10050077.
 - [10] C. Schnur *et al.*, “A Machine Learning Dataset of Artificial Inner Ring Damages on Cylindrical Roller Bearings Measured Under Varying Cross-Influences.” Zenodo, Apr. 15, 2025. doi: 10.5281/zenodo.11108503.
 - [11] ZeMA-gGmbH, “LMT-ML-Toolbox,” *GitHub repository*. GitHub, 2017. [Online]. Available: <https://github.com/ZEMA-gGmbH/LMT-ML-Toolbox>
 - [12] F. Mörchén, “Time series feature extraction for data mining using DWT and DFT.” Citeseer, 2003.
 - [13] R. H. Riffenburgh, “Linear discriminant analysis,” PhD Thesis, Virginia Polytechnic Institute, 1957. [Online]. Available: <http://hdl.handle.net/10919/80187>
 - [14] G. J. McLachlan, “Mahalanobis distance,” *Resonance*, vol. 4, no. 6, pp. 20–26, 1999, doi: 10.1007/BF02834632.

Fabrication and Characterization of a Sapphire Fabry–Pérot Fiber Sensor for High-Temperature Measurement up to 1000 °C

Pritesh Soni¹, Alexander Roehr², Fabian Buchfellner², and Johannes Roths

^{1,2}Photonics Laboratory, Munich University of Applied Sciences, Lothstrasse 34, 80335 Munich, Germany

Summary

We present the fabrication and characterization of a sapphire-based Fabry–Pérot interferometric sensor produced by optimized filament splicing, demonstrating a compact and reliable device capable of accurate and robust temperature measurement up to 1000 °C.

Keywords: High-temperature sensing, Sapphire fiber, Fabry–Pérot interferometer, Optical fiber sensor, Interferometric sensing

Introduction

Fiber-optic sensors are widely used for high-temperature monitoring due to their compact size, electromagnetic interference (EMI) immunity, and wavelength encoding. These features make them highly suitable for harsh environments such as aircraft engines, energy systems, and industrial processes [1–3]. However, high-temperature fiber sensors in conventional silica fibers, including regenerated Fiber Bragg Gratings (RFBGs), Type-I, and Type-II sensors, exhibit signal drift or degradation above 700 °C due to structural relaxation and viscoelastic behavior of silica [4,5]. This limits their stability and long-term reliability in extreme conditions. To overcome this, sapphire fibers (SF), with melting point above 2000 °C, can serve as a stable high-temperature sensing element due to their excellent thermal robustness, chemical stability, and wide optical transparency window [6,7]. At the same time, in our approach, standard silica fibers continue to provide efficient light guidance and compatibility with interrogation systems, enabling a hybrid sensing configuration.

Existing sapphire-based sensors such as multimode FBGs and extrinsic Fabry–Pérot interferometers (FPIs) often require complex designs, precise alignment, or femtosecond laser processing, which limits their practicality and scalability [8–13]. To offer a more practical solution, we implement a sapphire fiber based Fabry–Pérot interferometer (SFPI) formed by splicing a short sapphire fiber piece (~2 mm) onto a standard single-mode fiber (SMF), thereby creating a compact intrinsic

interferometric cavity defined by the silica–sapphire interface and the sapphire–air end-face. This configuration enables wavelength-encoded sensing based on interference between reflected signals. In this study, we present the detailed fabrication procedure and complete characterization of the SFPI, demonstrating stable temperature measurement up to 1000 °C with good fringe visibility, repeatability, and calibration performance.

Working principle and FPI theory

A commercially available single-mode interrogator (SM125, Micron Optics, USA) was used for signal generation and detection in reflection mode; thus, only a single delivery fiber was required. The optical signal propagates through the standard silica fiber (refractive index $n_1 = 1.4478$) toward the silica–sapphire interface, where a partial Fresnel reflection ($R_1 = 0.87\%$) occurs due to the refractive index mismatch between silica and sapphire ($n_2 = 1.7464$). The remaining optical power is transmitted into the sapphire fiber, where it propagates and undergoes a second partial reflection at the sapphire–air interface ($n_0 = 1$), with reflectivity $R_2 = 7.39\%$. These two reflecting interfaces form a low-finesse intrinsic Fabry–Pérot interferometer (SFPI) with cavity length $L \approx 2.2 \text{ mm}$ (see Fig. 1).

Due to the relatively low reflectivities ($R_1 \ll 1\%$ and $R_2 < 10\%$), the cavity operates in the low-finesse regime, where multiple internal reflections can be neglected and the interference can be described using a two-beam interference model.

The resulting interference intensity is given by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos\left(\frac{2\pi}{\lambda} \text{OPD}(T) + \phi_0\right), \quad (1)$$

where I_1 and I_2 correspond to the reflected intensities associated with R_1 and R_2 , respectively. The interference term contains the phase component, which is governed by the optical path difference (OPD) of the cavity

$$\text{OPD}(T) = 2n(T)L(T), \quad (2)$$

which depends on both the temperature-dependent refractive index $n(T)$ (thermo-optic effect) and the thermally induced expansion of the cavity length $L(T)$. Constructive interference (resonance condition) occurs when,

$$2n(T)L(T) = m\lambda, \quad (3)$$

where m is the interference order. Variations in temperature modify both $n(T)$ and $L(T)$, resulting in a shift of the resonant wavelengths (fringe positions). By tracking this wavelength shift, the temperature can be determined.

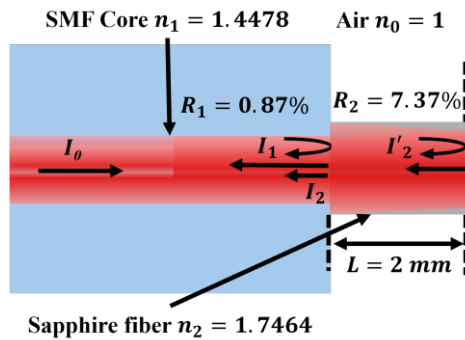


Fig. 1: SFPI schematic with Fresnel reflections (R_1 , R_2), refractive indices (n_1 , n_2 , n_0), and cavity length L .

Fabrication of SFPI

The SFPI sensor fabrication begins with the preparation of a standard single-mode fiber (SMF) and a single-crystal sapphire fiber (SF). The SMF was first stripped of its protective coating and then cleaved using a precision cleaver to obtain a flat and perpendicular end-face (Fig. 2A), ensuring efficient optical coupling at the splice interface.

The sapphire fiber was initially available as a longer segment with a rough and irregular end-face (Fig. 2B), along with its side profile as shown in Fig. 2C. Due to its anisotropic crystal structure and high mechanical hardness, the sapphire fiber cannot be cleaved reliably. Therefore, the end-face was mechanically polished using progressively finer abrasive films (6 μm to 1 μm) to achieve a smooth and optically flat and perpendicular surface (Fig.

2E), which is essential for controlled reflection at the sapphire–air interface.

Following end-face preparation, both fibers were positioned for splicing with careful core-to-core alignment and a controlled pre-gap (Fig. 2D). This pre-gap is critical to prevent unwanted physical contact during the heating phase, accounting for thermal elongation of the sapphire fiber (~24 μm observed at 52 W filament power). Since direct temperature control is not available in the filament splicer, the splice parameters were empirically optimized to compensate for thermal expansion during the heating stage and subsequent contraction upon cooling, ensuring stable interface formation.

The fibers were then fused using a filament-based fusion splicing system (LSF4000, Vytran, UK), where controlled heating softens the silica while maintaining the structural integrity of the sapphire. The final splice region (Fig. 1F) exhibits a mechanically robust joint with precise core alignment, forming a stable optical interface suitable for interferometric operation.

After splicing, the sapphire fiber was intentionally shortened by controlled breaking at approximately 2–3 mm from the splice point to define the cavity length. The newly created sapphire end-face was again polished using the same procedure, resulting in two well-defined, parallel reflecting surfaces forming the intrinsic Fabry–Pérot cavity. This sequence ensures reproducible SFPI fabrication with controlled geometry and reliable optical performance.

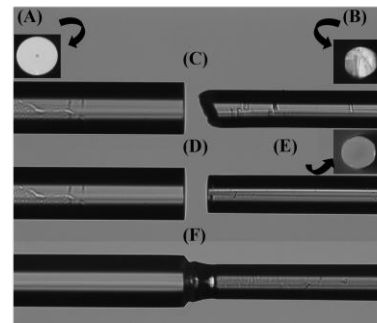


Fig. 2: SFPI fabrication: (A) SMF cleaved end-face, (B) unpolished SF end-face, (C) alignment, (E) polished SF, (D) pre-splice positioning, (F) final splice

Following the splice formation, the reflected signal evolves with cavity formation and end-face quality. For a long sapphire fiber (~30 cm), no interference is observed and the signal remains nearly constant (Fig. 4A), due to weak back-coupling and multimode propagation. After trimming to ~2.5 mm, partial modulation

appears (Fig. 4B), but fringe visibility remains low due to the unpolished sapphire–air interface causing scattering and reduced reflectivity. After polishing the second end-face, a stable Fabry–Pérot interference pattern with high visibility ($\sim 55\%$) is obtained (Fig. 4C), confirming the formation of a well-defined two-beam cavity.

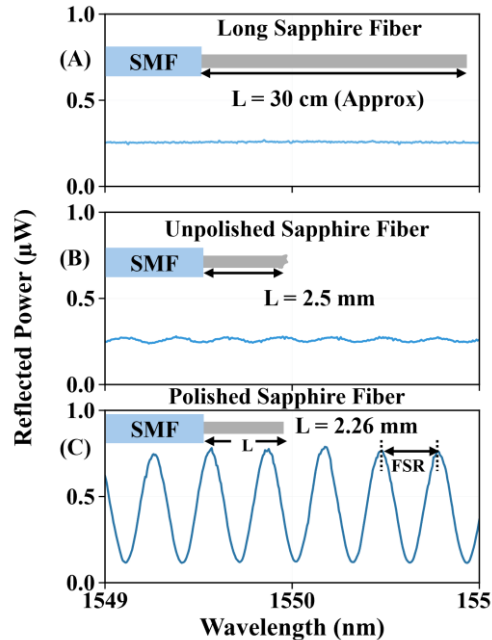


Fig. 4: Reflected signal evolution after splice: (A) long SF (no interference), (B) trimmed unpolished SF (low visibility), (C) polished SFPI (~ 2.26 mm) with strong-visibility of Fabry–Pérot fringes ($FSR \approx 0.304$ nm)

The measured spectra confirm this behavior, and the free spectral range ($FSR \approx 0.304$ nm) validates the cavity length using Eq. (4),

$$L = \frac{\lambda^2}{2n_2 \cdot FSR} \approx 2.26. \quad (4)$$

After obtaining a stable high-visibility SFPI signal and a well-defined cavity structure, the sensor was encapsulated in a capillary for mechanical protection and environmental isolation, making it ready for high-temperature testing and calibration.

High-temperature measurements:

After fabrication, the sensor was placed inside a calibration furnace (Pegasus Plus 1200S, Isothermal Technology Ltd., UK) together with a Type-R thermocouple, as shown schematically in Fig. 3. The SFPI and thermocouple were positioned in close proximity to ensure identical thermal conditions. The reflected spectra were recorded using an optical interrogator (1510–1590 nm, 5 pm resolution) at a sampling rate of

1 Hz. An in-house LabVIEW program processed each spectrum in real time by fitting a second-order polynomial to the upper 50% of the tracked fringe peak, enabling wavelength tracking.

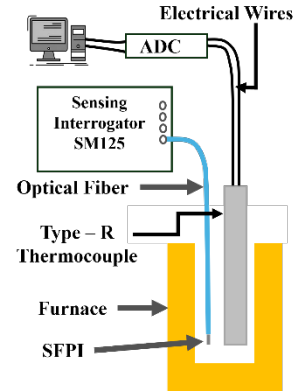


Fig. 3: Schematic of the high-temperature calibration setup showing the SFPI sensor and Type-R thermocouple co-located inside a furnace, with optical interrogation via a broadband interrogator.

The measured oven temperature and the tracked fringe wavelength are presented in Fig. 5, where the left y-axis corresponds to oven temperature (set and measured), and the right y-axis corresponds to the fringe wavelength. The temperature profile consisted of stepwise heating from room temperature (~ 20 °C) to 1000 °C followed by cooling, with a dwell time of 2 h at each step to ensure thermal equilibrium. The tracked fringe wavelength follows the thermocouple temperature closely over the entire cycle, demonstrating stable operation of the SFPI. Over this range, the

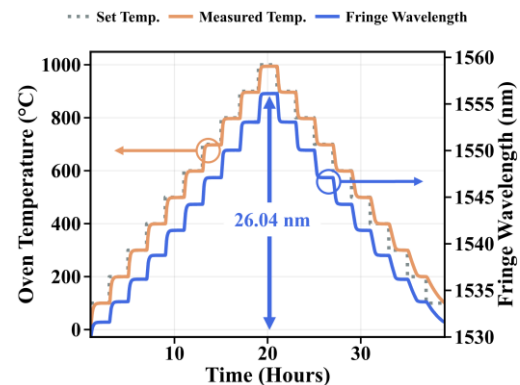


Fig. 5: Set and Measured temperature steps (left y-axis) and corresponding fringe wavelength response (right y-axis); the tracked fringe follows the measured temperature steps, showing a total wavelength shift of ~ 26.04 nm.

selected fringe exhibits a total wavelength shift of approximately 26.04 nm, corresponding to a temperature change from ~21 °C to ~994 °C.

The calibration curve is shown in Fig. 6, where the fringe shift $\Delta\lambda(T)$, referenced to the wavelength at 200 °C ($\lambda_{ref} = 1533.785$), is plotted as a function of temperature. The fringe shift values were obtained from steady-state averages over a 10 min window near the end of each dwell period. A fifth-order polynomial was fitted to the combined heating and cooling dataset, yielding an excellent agreement. The residuals remain below ± 0.02 nm across the full temperature range, confirming strong consistency between heating and cooling branches and negligible hysteresis.

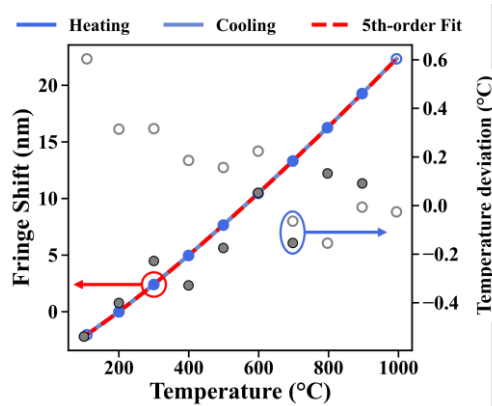


Fig. 6: Fringe shift versus temperature referenced to 200 °C; heating (filled) and cooling (hollow) data with a fifth-order polynomial fit (dashed, red). Left y-axis: fringe shift (nm); right y-axis: temperature deviation (°C), remain within ± 0.6 °C.

The fitted response exhibits a nonlinear increase in sensitivity with temperature. The sensitivity rises from approximately 23.18 pm/°C at 200 °C to 31.63 pm/°C at 1000 °C, reflecting the combined temperature dependence of the thermo-optic coefficient and thermal expansion of sapphire.

Using the polynomial calibration fit parameters, the temperature was reconstructed from the measured fringe wavelength. The temperature deviation between the SFPI-reconstructed temperature and the measured thermocouple temperature,

$$\text{Temperature Deviation} = T_{SFPI} - T_{TC}, \quad (5)$$

is shown in Fig. 6 (right y-axis). The resulting temperature deviation remains within ± 0.6 °C over the full range. These results demonstrate high repeatability of the SFPI sensor over the investigated temperature range.

Summary

An intrinsic sapphire fiber Fabry–Perot interferometer (SFPI) was demonstrated for high-temperature sensing up to 1000 °C. The sensor exhibits a wavelength shift of ~26.04 nm and a high sensitivity reaching 31.6 pm/°C at 1000 °C. The reconstructed temperature shows excellent agreement with a reference Type-R thermocouple, with a maximum deviation below 0.6 °C, confirming stable operation over the full range.

The proposed approach enables efficient light guidance using a standard silica single-mode fiber (SMF) integrated with a compact sapphire sensing element (~2.2 mm), allowing the use of a simple, commercially available interrogation system. The silica–sapphire splice remains mechanically robust and optically stable under high-temperature conditions, demonstrating reliable sensor integrity. Compared to regenerated Fiber Bragg Gratings (RFBGs), the SFPI exhibits significantly higher sensitivity, highlighting the advantage of interferometric sensing [4]. Overall, the sensor offers a compact, fabrication-efficient, and high-performance solution for harsh-environment temperature sensing.

References

- [1] Stadler, A., Zeisberger, A., Buchfellner, F., Roehrl, A., Koch, A. W., et al., "Accurate high-temperature profile sensing with dense multipoint arrays of regenerated fiber Bragg gratings," *Results in Physics*, Vol. 65, 1 Jan. 2024, p. 107970. doi: 10.1016/j.rinp.2024.107970.
- [2] Dutz, F. J., Boje, S., Orth, U., Koch, A. W., and Roths, J., "High-Temperature Profile Monitoring in Gas Turbine Exhaust-Gas Diffusers with Six-Point Fiber-Optic Sensor Array," *International Journal of Turbomachinery, Propulsion and Power*, Vol. 5, No. 4, 1 Jan. 2020, p. 25. doi: 10.3390/ijtp5040025.
- [3] Dutz, F. J., Heinrich, A., Bank, R., Koch, A. W., and Roths, J., "Fiber-Optic Multipoint Sensor System with Low Drift for the Long-Term Monitoring of High-Temperature Distributions in Chemical Reactors," *Sensors (Basel, Switzerland)*, Vol. 19, No. 24, 1 Jan. 2019. doi: 10.3390/s19245476.
- [4] Buchfellner, F., Stadler, A., Bian, Q., Hennesen, M., Zeisberger, A., et al., "Generalized and wavelength-dependent temperature calibration function for multipoint regenerated fiber Bragg grating sensors," *Optics express*, Vol. 30, No. 25, 1 Jan. 2022, pp. 44769–44784. doi: 10.1364/OE.470093.
- [5] Lindner, M., Bernard, D., Heilmeier, F., Jakobi, M., Volk, W., et al., "Transition from purely elastic to viscoelastic behavior of silica optical fibers at high temperatures

- characterized using regenerated Bragg gratings," *Optics express*, Vol. 28, No. 5, 1 Jan. 2020, pp. 7323–7340.
doi: 10.1364/OE.384402.
- [6] A. Roehrl, A. Stadler, F. Buchfellner, S. Zehetmair, and J. Roths, "Ultra-high Temperature Sensor Utilizing an Intrinsic Sapphire Fabry-Pérot Interferometer: in Advanced Photonics Congress 2024, Technical Digest Series," 1 Jan. 2024.
 - [7] Pritesh Soni, "Fiber-coupled Sapphire Fiber Fabry-Pérot Interferometry for High-temperature Measurement," Master Project Study, 1 Oct. 2025.
 - [8] Wang, B., Niu, Y., Qin, X., Yin, Y., and Ding, M., "Review of high temperature measurement technology based on sapphire optical fiber," *Measurement*, Vol. 184, 1 Jan. 2021, p. 109868.
doi: 10.1016/j.measurement.2021.109868.
 - [9] A. Wang, S. Gollapudi, R. G. May, K. A. Murphy, and R. O. Claus, "Advances in sapphire-fiber-based intrinsic interferometric sensors," *Opt. Lett.*, Vol. 17, No. 21, 1 Jan. 1992, pp. 1544–1546.
doi: 10.1364/OL.17.001544.
 - [10] Shen, Y., Jiang, Y., Xie, S., Sui, G., and Zhang, X., "Cascaded Fabry-Perot cavity and fiber Bragg grating on sapphire fibers for high-temperature strain sensing," *Optics and Lasers in Engineering*, Vol. 184, 1 Jan. 2025, p. 108674.
doi: 10.1016/j.optlaseng.2024.108674.
 - [11] Grobncic, D., Mihailov, S. J., Smelser, C. W., and Ding, H., "Sapphire Fiber Bragg Grating Sensor Made Using Femtosecond Laser Radiation for Ultrahigh Temperature Applications," *IEEE Photonics Technology Letters*, Vol. 16, No. 11, 1 Jan. 2004, pp. 2505–2507.
doi: 10.1109/LPT.2004.834920.
 - [12] Ma, S., Xu, Y., Pang, Y., Zhao, X., Li, Y., et al., "Optical Fiber Sensors for High-Temperature Monitoring: A Review," *Sensors (Basel, Switzerland)*, Vol. 22, No. 15, 1 Jan. 2022.
doi: 10.3390/s22155722.
 - [13] Wang, B., Niu, Y., Zheng, S., Yin, Y., and Ding, M., "A High Temperature Sensor Based on Sapphire Fiber Fabry-Perot Interferometer," *IEEE Photonics Technology Letters*, Vol. 32, No. 2, 1 Jan. 2020, pp. 89–92.
doi: 10.1109/LPT.2019.2957917.

Hyper-fast Gas Chromatography Coupled with Ion Mobility Spectrometry

*J. A. Wittwer, A. Nitschke, M. Hitzemann, J. Winkelholz, A. T. Kirk, C. Schaefer,
T. Kobelt, C. Thoben, M. Lippmann, S. Zimmermann*
Leibniz University Hannover, Institute of Electrical Engineering and Measurement Technology,
Appelstr. 9A, 30167 Hannover (Germany)

Abstract

Ion mobility spectrometers (IMS) are frequently coupled with gas chromatographs (GC), particularly when analyzing complex samples. However, analysis time remains a critical parameter in gas chromatography. Therefore, a hyper-fast GC represents a promising approach for GC-IMS, allowing for rapid two-dimensional separation in less than 60 s and highest sensitivity. To simultaneously detect both ion polarities, a new IMS with two drift tubes in an axial arrangement has been developed, specifically designed to resolve even hyper-fast GC peaks.

Keywords: Ion mobility spectrometry, hyperfast gas chromatography, process monitoring

Introduction

For several reasons, such as reducing chemical cross sensitivities caused e.g. by competing ionization processes and varying cluster size of reactant ions with humidity or adding a second dimension of separation, ion mobility spectrometers (IMS) are often coupled with gas chromatographs (GC). The combination of GC and IMS is used to tackle complex backgrounds, particularly when analyzing complex sample mixtures. However, this contradicts one major advantage of IMS, which is rapid detection of compounds present at ppbv and even pptv levels in less than one second measurement time, while already offering decent resolving power. Coupling IMS with GC usually extends measurement time to at least several minutes if not tens of minutes. Therefore, a hyper-fast GC is a very promising approach for GC-IMS to combine highest sensitivity of IMS with two dimensional separation while allowing for rapid analysis within the double-digit second range [1].

Instrumentation and Results

This study presents a novel drift tube IMS, specifically designed to resolve even hyper-fast GC peaks [2]. Key is the directed sample gas flow through the ionization region, resulting in a small effective detector volume matching the volume of peaks eluting from the hyper-fast GC, so that the GC peaks do not merge in the IMS reaction region [3]. Figure 1 shows an

exemplary measurement containing a mixture of 4-ketones within a short measurement time of only 30 seconds.

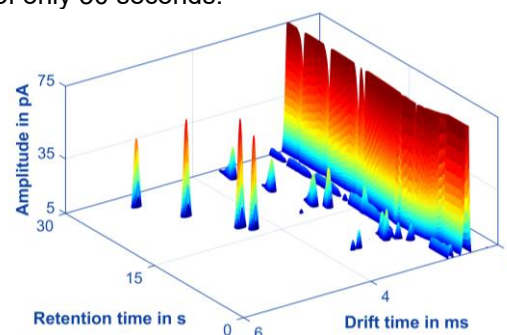


Fig. 1: 3D-GC-IMS chromatogram of an exemplary ketone mix containing 4-heptanone, 2-heptanone, 2-octanone and 2-nonanone.

To minimize condensation of analytes, the IMS is isothermally heated to 120 °C. A drift tube length of only $L_D = 41$ mm with a drift voltage of $U_D = 2.7$ kV leads to high resolving power of $R_p = 60$ and an IMS repetition rate of 100 Hz, even for ions with low reduced ion mobility down to $0.43 \text{ cm}^2 \text{V}^{-1} \text{s}^{-1}$. This setup can resolve very short GC peaks with full width at half maximum of only 140 ms. It is demonstrated that this system enables the rapid analysis of complex mixtures, such as those composed of hops or chemical warfare agents, within a time frame of 60 seconds.

References

- [1] A. Nitschke, M. Hitzemann, J. Winkelholz, A.T. Kirk, M. Lippmann, C. Thoben, J.A. Wittwer, S. Zimmermann, *J. Chromatogr. A* 1736, 465376 (2024); doi: 10.1016/j.chroma.2024.465376.
- [2] A. Nitschke, M. Hitzemann, J. Winkelholz, A.T. Kirk, C. Schaefer, T. Kobelt, C. Thoben, M. Lippmann, J.A. Wittwer, S. Zimmermann, *J. Chromatogr. A* 1750, 465923 (2025); doi: 10.1016/j.chroma.2025.465923.
- [3] V. Maynard, E. Grushka, *Anal. Chem.* 44, 1427–1434 (1972); doi: 10.1021/ac60316a005.

Evaluation of Time Synchronization Mechanisms in Legal Metrology

Dr.-Ing. Marko Esche¹, Dipl.-Ing. Felix Reichmuth¹, Dipl.-Inf. Reinhard Meyer¹, Martin Nischwitz¹, M.Sc.

¹ Physikalisch-Technische Bundesanstalt, Fachbereich 8.5 „Metrologische IT“, Abbestraße 2-12, 10587 Berlin

Abstract

To ensure correct billing in case of dynamic tariffs, measuring instruments subject to legal control must be able to correctly obtain legal time from an authenticated source. Typically, each EU member state designates one institution for distributing legal time locally. Several technical implementations for time synchronization are introduced and discussed in this paper. They are evaluated using a harmonized risk assessment approach based on ISO 18045 vulnerability analysis. Results indicate that protection for high-risk scenarios, e.g., critical infrastructure, can be ensured via state-of-the-art cryptographic protocols or by combining different synchronization mechanisms. Based on the results of the assessment, recommendations for the practical usage of the different implementations are made.

Keywords: legal metrology, network time protocol, dynamic tariffs, electric vehicle supply equipment, software risk assessment.

Introduction

In the European Union, certain types of measuring instruments used for commercial transactions are subject to legal control under the Measuring Instruments Directive (MID) [1]. These include widely used instruments such as active electricity meters, gas meters, taximeters, length measuring devices and automatic weighing instruments. The central aim of the MID is to ensure consumer protection through provision of fair and traceable measurements.

Since 2024 the EU legislative bodies have been working towards an amendment of the MID to extend its scope to electric vehicle supply equipment (EVSE) and dispensers for compressed gases [2]. Both are seen as instruments crucial for facilitating the transition from fossil fuels to green energy sources. Given the volatile nature of renewable electrical energy in particular, many EU member states are already experiencing the advent of new and dynamic, i.e., time-dependent, pricing schemes for electrical energy supply, including EVSEs, replacing classical fixed-price schemes known from fuel dispensers. While these pricing schemes will allow users to profit from cheaper energy when solar or wind power are available in abundance, they also pose new challenges for the measuring instruments manufacturer,

with the necessity to create a direct link between measured quantity and applicable price at a given time [3]. Such time-dependent tariffs necessitate that measuring instruments have access to a trustworthy and reliable source of legal time to ensure correct application of varying prices.

Given that MID Annex I requires measuring instruments, including their software, parameters and data, to be adequately secured and protected against manipulation [1], protection means for these assets now have to be extended to cover mechanisms for time synchronization as well. This extension poses new challenges for the legal metrology community since time information is inherently dynamic as opposed to static binary information typically communicated between components of a measuring system.

In this paper, the prevalent time synchronization mechanisms are investigated and assessed regarding their suitability for fulfilling the essential requirements of MID [1]. Based on the outcome of the assessment, recommendations are made to ensure MID-compliant implementation and usage of time synchronization mechanisms for legally regulated measuring instruments.

The remainder of the paper is structured as follows: In the next Section, a short analysis of

the legal requirements for time synchronization in measuring instruments is given. Afterwards, the available time synchronization mechanisms for measuring instruments are introduced and discussed, including several remarks regarding their practical usage and associated security aspects. Subsequently, the risk assessment methodology used in this paper is presented and applied to the different time synchronization mechanisms. Based on the outcome of the analysis, the final Section draws a number of conclusions and makes recommendations regarding the legally relevant usage of time synchronization.

Requirements Overview

As mentioned in the introduction, MID Annex I lays down generic essential requirements applicable to a large range of measuring instruments. By means of the annex, MID aims to establish a common level of consumer protection and trust in measurements conducted for commercial transactions. Amongst the requirements from Annex I, number 8.4 directly addresses IT-related assets, "Measurement data, software that is critical for measurement characteristics and metrologically important parameters stored or transmitted shall be adequately protected against accidental or intentional corruption." [1] From this essential requirement it can be concluded that measurement data must be secured and protected against attacks, whereas adequacy of the securing and protection means can be demonstrated by means of a mandatory risk assessment, see MID Annex II.

Whenever measurement data are combined with time information, e.g., for assigning different tariffs to measured quantities, the essential requirements need to be extended to the time information, too. This is typically achieved by requiring the measuring instrument's internal clock to have a maximum drift compared to legal time, e.g., 3% of a billing interval, corresponding to 9 seconds for quarter-hourly tariff changes. In addition, setting of the clock must be adequately protected and secured, see clause 6.2.7.1 in [4]. Since time information, as opposed to most other measurement data, must be processed immediately, a bidirectional setup between time source and measuring instrument is typically used, in which all changes to the setup are protected and secured to provide evidence of intervention for later checks by market or user surveillance authorities.

Time Synchronization Mechanisms

Since 1985 the network time protocol (NTP) [5] has been used to automatically synchronize network-connected devices all around the world. The protocol uses time-stamped UDP messages to allow NTP clients to calculate round-trip times and individual time offsets compared to an NTP server. Because of the reliance on UDP and due to the lack of an authentication mechanism within that protocol, it is potentially susceptible to a man-in-the-middle attack (MITM) which could endanger compliance with the essential requirements from MID. The international organization of legal metrology (OIML) addressed this issue in its latest version of the OIML Document D 31 "General requirements for software-controlled measuring instruments", see clause 6.2.7 in [4]. Notably, the NTP protocol does not provide any proof of identity or authentication mechanism for the time source and relies on a low-latency bidirectional communication channel to ensure proper functioning of the synchronization mechanism.

In 2020, the network time security protocol (NTS) [6] was published, extending NTP with cryptographic security for client-server communication. The protocol consists of two main phases: During the key establishment phase, an encrypted communication channel with a trusted key server is established via TLS and TCP. During this stage, so-called NTS-cookies are shared with the client to ensure that both parties can properly authenticate each other during the second phase by means of the authenticated encryption with associated data (AEAD) mechanism. During the second stage, NTP with a cryptographic extension in the NTS cookie placeholder extension field is used. This field is added to the NTP packets, allowing clients and servers to check each other's identity. At the same time, clients can request new cookies with this extension, whereas servers can monitor the state of the synchronization and issue new NTS cookies when needed. In theory, the NTS key establishment only must be carried out once. It should also be noted that key server and NTP server can be operated separately, putting lower strain on network resources. Practical aspects of the protocol will be discussed during risk assessment and subsequent analysis. When compared to simple NTP, NTS poses additional demands on client devices in particular: For establishment of the TLS connection, cryptographic certificates and keys must be properly handled. Additionally, mechanisms for the correct usage of the

aforementioned NTS cookies have to be implemented in each client.

In Germany, a historical alternative is still the DCF77 transmitter which is physically located in Frankfurt/Main, Germany and continually transmits an analog time signal at a frequency of 77.5 kHz [12]. The base frequency of the signal acts as a frequency normal and is directly fed via a fiber-optical link from the official atomic clocks at the Physikalisch-Technische Bundesanstalt (PTB) Braunschweig, Germany's national metrology institute, which is tasked with distributing legal time in Germany. In addition, phase- and amplitude-modulated information is contained in the signal to enable receivers to decode minute, hour, calendar day, number of the month and year. Given the unidirectional nature of the analog signal, it does not contain any authentication or integrity protection mechanisms. How manipulation of the analog signal can be mitigated nonetheless will be discussed during the following risk assessment.

Risk Assessment

To enable conformity assessment bodies and manufactures of measuring instruments to evaluate the suitability of the implemented securing and protection mechanisms in a measuring instrument, the European Cooperation in Legal Metrology has harmonized a risk assessment methodology in WELMEC Guide 7.6 [8]. The method is based on an original publication by PTB from 2015 [8] and consists of three main stages:

During risk identification potential threats to assets are identified. In the scenario discussed here, such a threat could be formulated in the following manner: An attacker manages to manipulate the internal clock of a measuring instrument without evidence of intervention.

During risk estimation, technical steps needed to realize a threat are investigated and assigned specific point scores making use of ISO 18045 vulnerability analysis [11]. The point scores aim to quantify the resources necessary to implement an attack vector and are divided into the following categories: time required (0 for less than a day to 19 for at least half a year), necessary expertise (0 for a laymen to 9 for an expert in multiple fields), required knowledge of the target of evaluation (0 for public information to 11 for critical data), window of opportunity (between 0 for unlimited or unnecessary access to 10 for severely restricted access) and equipment needed (0 for standard to 9 for multiple bespoke equipment). After having

assigned these scores to an attack vector, a sum score can be calculated which in turn is mapped to a probability score according to Table 1.

Tab. 1: Mapping between sum score, resistance to attacks and probability score. [10]

Sum of point scores	Resistance to attacks	Probability score
0-9	no rating	5
10-13	basic	4
14-19	enhanced basic	3
20-24	moderate	2
>24	high	1

The probability score is then multiplied with the assigned impact ($\frac{1}{3}$ for attacks on single measurements, 1 for attacks on multiple measurements) to yield the final risk score.

Finally, the risk score is put into the context of the instruments usage during risk evaluation.

The method was later extended in [9] to allow for the inclusion of attacker motivation in risk assessment. To deal with complex attacks requiring multiple steps, attack probability trees were introduced in [10]. These attack probability trees are graphical representations of an attack and potential subgoals signified by a root node with linked child nodes. A visual example of an attack probability tree with one parent node and two child nodes may be found in Figure 3. In [10] it was proven that any attack probability tree can be represented as a binary tree which enables assessors to assign point scores to the leaf nodes alone. Following rules defined in [10], these point scores can then be propagated up the tree to calculate the risk of the root node and its associated threat.

The results of such a risk assessment can then be used to identify potential countermeasures to said attack vectors. In this paper, a generic architecture of an EVSE and its environment is used to identify potential attack vectors and influence factors for time synchronization, see Figure 1. The EVSE is assumed to contain an EVSE measuring system (EVSE-MS) which is subject to legal control. Typically, the EVSE-MS contains a separately certified active electrical energy meter which calculates measurement data, provides time stamps for beginning and end of a transaction and cryptographically signs the resulting charge data record. Other software

modules or hardware components, both legally relevant or legally non-relevant, may also be part of the EVSE-MS.

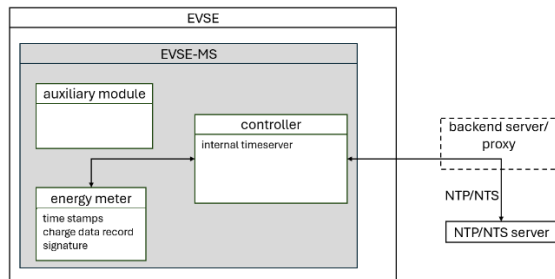


Fig. 1: Generic architecture of an EVSE and its environment. The measuring system of the EVSE (EVSE-MS) creates time-stamped, cryptographically signed charge data records, while the charge point management system (CPMS) acts as a transparent tunnel to the network time security (NTS) server used for distributing legal time.

To obtain valid time information, the EVSE-MS communicates with a pre-established time source, e.g., time servers provided by PTB, and sets the energy meter's internal clock accordingly. If NTP or NTS are used, setting does not occur abruptly but is done in incremental steps to ensure proper functioning of the meter's operating system. In the following, all three time-synchronization mechanisms are examined separately, and the applicable attack vectors are assigned their proper point scores.

DCF77

If the analog DCF77 is used, two potential attack paths appear to be viable: Either an attacker blocks the time signal altogether, which will eventually lead to uncontrollable drift of the internal clock or a sufficiently strong fake signal is provided. Since arbitrary drift of the clock is unlikely to profit a potential attacker, only the latter scenario is examined: Potentially, one could record the time signal and reuse it to artificially increase the duration of a charging session or to assign daytime session to more costly nighttime prices. The equipment needed to capture and retransmit such a signal will have to be specifically built for the purpose (point score of 7) and will require an electronics engineer (point score of 6 for expertise) at least a few days (point score of 2 for time). At the same time, all necessary data can be found readily online, and the attack itself does not require physical access to the EVSE itself.

Instead, the fake time signal generator simply must be put in the vicinity of an EVSE, thus potentially also being able to manipulate multiple instruments and charging sessions simultaneously (impact score of 1). The resulting attack probability tree consisting of a single root node is given in Figure 2.

A01: False time source DCF77	
Ti: 2	WO: 0
Ex: 6	Eq: 7
Kn: 0	Σ : 15

Fig. 2: Attack probability tree consisting of a single root node addressing manipulation of measuring instruments using DCF77 time synchronization. The node is assigned scores for time (Ti), expertise (Ex), knowledge of the device (Kn), needed equipment (Eq) as well as the window of opportunity (WO). Additionally, the sum score (Σ) for the attack is given.

The sum score of 15 results in enhanced basic protection (risk score of 3) for the configuration, see Table 1. To ensure adequate protection against manipulation, a secondary time source like NTP could be used to mitigate the attack and reduce the risk of manipulation even further.

NTP

If the network time protocol is used, a trusted source must be configured during initial installation of the EVSE. Since this process is subject to a separate conformity assessment module, one can assume that the NTP server address is configured correctly at initial startup and that a correct time signal is received at least once. During use of the instrument, however, this will not necessarily hold true. To install a false time source (see Figure 3), an attacker would need to set up a false NTP server (child node A12 in Figure 3) and redirect NTP traffic via a manipulated proxy (child node A13).

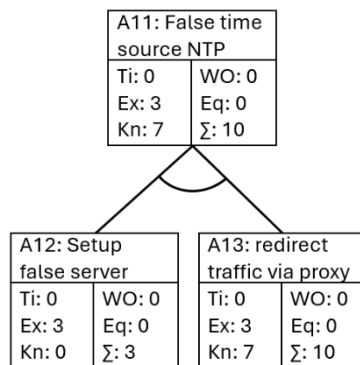


Fig. 3: Attack probability tree for an attack on an NTP-synchronized measuring instrument by setting up a fake time server (A12) and manipulating NTP traffic via a proxy (A13).

Once these scores have been propagated up the tree to the root node A11, a sum score of 10 is obtained, which is equivalent to a probability score of 4. Since the manipulated NTP server would simultaneously affect all measurement results of the attacked EVSE (impact score of 1), the overall risk score has a value 4 as well. Compared with the usage of the analog DCF77 signal, it should be clear that a higher risk of manipulation exists in this scenario.

NTS

To mitigate the shortcomings of NTP, the NTS protocol was introduced by IETF in 2020. With NTS, protection of time information is reliant on two factors: the strength of the cryptographic primitives used for TLS and AEAD in the two protocol phases and the protection of the corresponding cryptographic material used in the measuring instrument. Since TLS uses state-of-the-art RSA or elliptic curve cryptography (ECC) with typical key lengths of 4096 bits for RSA and 256 bits for ECC, attacks on this part of the NTS protocol would realistically take several years on the world's best supercomputers, rendering an attack virtually impossible. According to [6], server and client can negotiate the encryption algorithm used for AEAD, but it is typically AES with a minimum key length of 128 bit. This is again state of the art and extremely difficult to break. Thus, attacks on the protection of the instrument and its cryptographic certificates are much more likely and will be examined here.

To falsify an NTS time source, an attacker would first need to set up an alternative NTS server, see node A22 in Figure 4. Since new certificates would be necessary for such a false

server, these need to be created and introduced into the instrument to be attacked, see node A23. Since setup of NTP and NTS servers are almost identical tasks, typically involving the same open-source Linux package chrony, and redirection of traffic can be realized similarly in both scenarios, the child nodes of A22 are identical to those of node A11 in Figure 3. To introduce false cryptographic certificates into the target measuring instrument, an attacker first needs to generate a false certificate (node 26) and install it in the instrument (node 27). Since certificates are complex configuration files but do not contain any executable code, they are usually interpreted as device-specific parameters of a measuring instrument. This implies that they are protected and secured against exchange and that any modification of the registered certificates is logged in an audit trail.

Since the certificate of the NTS server typically changes on a yearly basis, the corresponding root certificate is usually installed in the measuring instrument instead. This means that for a successful attack, an expert (score of 6) would need to set up their own certificate chain, starting with a fake root, which will take at least a day (time score of 1). Since this does not require any access to the targeted device and can be done with freely available software like OpenSSL, all other scores can be set to 0, see node 26 in Figure 4. The more challenging part for any attacker would be installing the fake certificate on the target device, see node 27. To achieve this on a device with a properly configured operating system, an attacker would likely need to find a loophole in the device's firewall and then use privilege escalation via a vulnerability in one of the operating system's modules. Again, this will require a software expert at least a few days (time score of 1) and will also require sensitive knowledge of the target system (knowledge score of 7). Moreover, the attacker will likely have to develop their own attack tool, making it a bespoke equipment (score of 7) from the point of view of ISO 18045 [11]. Since all parts of the attack on NTS can be done remotely and would affect all future measurements performed by one or more devices, the window of opportunity is unlimited (score of 0) and the impact would be set to 1. Once the scores for the leaf nodes have been propagated up the tree, this results in a sum score of 21 for the root node, an attack probability score of 2 (see Table 1) and an equivalent risk score of 2. This means that the method offers at least moderate protection and

fulfils requirements from WELMEC Guide 7.6 for all risk classes.

Analysis and Conclusions

Based on the results of the risk assessment, recommendations can be made regarding the secure and MID-compliant configuration of time synchronization mechanisms for EVSEs. As can be seen from the total risk scores for the three different synchronization mechanisms, DCF77 offers medium protection against manipulation simply by making the attack itself hardware resource intensive.

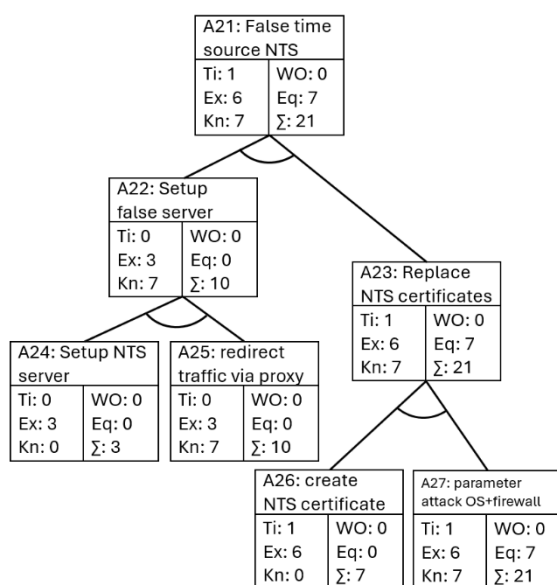


Fig. 4: Attack probability tree for an attack on a NTS-synchronized measuring instrument by setting up a fake time server (A22) and modifying the installed certificates on the device accordingly (A23). Both child nodes are further subdivided.

If combined with a second unauthenticated synchronization mechanism for plausibility checking, DCF77 will likely remain an adequate technical means for the foreseeable future for all WELMEC risk classes. However, compared to the other two network-based solutions, it does require the manufacturer to realize a specific hardware configuration in their device (temperature-controlled real time clock, analog

radio receiver). These are likely to stop most parties from pursuing this solution.

NTP, by comparison, is already available in most common embedded devices via their operating system. It does not require an expensive RTC on the client side, if synchronizations with the server occur sufficiently often. However, it does not offer sufficient protection against manipulation and attacks could potentially affect many devices and measurements.

NTS should be an acceptable alternative for most use cases. While the cryptographic strength of the protocol and its different phases is well known, this has to be supported by proper configuration of the client's operating system to ensure adequate protection of the associated cryptographic material.

In this regard, OIML D 31 and other requirement documents still appear to provide insufficient guidance on the handling of cryptographic material needed for usage of time synchronization mechanisms, specifically in the context of MID requirements for prevention of inadmissible modifications.

It is anticipated that the evaluated time synchronization mechanisms should cover a significant portion of EVSE models currently in use. For other types of measuring instruments that use legal time, e.g., smart utility meters for electrical energy, gas, water, or thermal energy, requirements are typically less strict since the supplied customer does not change frequently, and billing does not need to be done in real time. Subsequently, a combination of DCF77 and NTP as a standalone mechanism should be acceptable in this scenario. In Germany, the Federal Office for Information Security (BSI) prescribes the usage of NTS for such meters. How to safely commission NTS client devices without the need for manual checks of the installed certificates remains a task for future work.

References

- [1] EC, "Directive 2014/32/EU of the European Parliament and of the Council of 26 February 2014 on the harmonisation of the laws of the Member States relating to the making available on the market of measuring instruments," European Union, Council of the European Union; European Parliament, Directive, February 2014

- [2] EC, "Measuring instruments", [Online]. Available: "https://single-market-economy.ec.europa.eu/single-market/goods/building-blocks/legal-metrology/measuring-instruments_en" [Accessed 21 November 2025].
- [3] "OIML G22 Electric Vehicle Supply Equipment (EVSE)," International Organization of Legal Metrology, Paris, France, Tech. Rep., 2022
- [5] D. Mills, U. Delaware, E. J. Martin, J. Burbank and W. Kasch, "RFC 5905 - Network Time Protocol Version 4: Protocol and Algorithm Specification", [Online]. Available: <https://tools.ietf.org/html/rfc5905>. [Accessed 4 May 2020].
- [4] "OIML D31 General requirements for software-controlled measuring instruments," International Organization of Legal Metrology, Paris, France, Tech. Rep., 2023.
- [6] D. Franke, D. Sibold, K. Teichel, M. Dansarie and R. Sundblad, "RFC 8915 - Network Time Security for the Network Time Protocol", [Online]. Available: <https://tools.ietf.org/html/rfc8915>. [Accessed 13 February 2025].
- [8] "WELMEC Guide 7.6: Software Risk Assessment for Measuring Instruments", WELMEC, 2021.
- [8] M. Esche and F. Thiel, "Software risk assessment for measuring instruments in legal metrology," in Proceedings of the Federated Conference on Computer Science and Information Systems, Lodz, Poland, September 2015, pp. 1113–1123
- [9] M. Esche, F. Thiel „Incorporating a Measure for Attacker Motivation into Software Risk Assessment for Measuring Instruments in Legal Metrology“, 18. GMA/ITG-Fachtagung Sensoren und Messsysteme, Nuremberg, Germany, May 2016, doi: 10.5162/sensoren2016/P7.4
- [10] M. Esche, F. Grasso Toro, and F. Thiel, "Representation of attacker motivation in software risk assessment using attack probability trees," in Proceedings of the Federated Conference on Computer Science and Information Systems, Prague, Czech Republic, September 2017. doi: 10.15439/2017F112 pp. 763–771.
- [11] "ISO/IEC 18045:2022 Information security, cybersecurity and privacy protection - Evaluation criteria for IT security - Methodology for IT security evaluation", ISO, Geneva, August 2022
- [12] D. Piester, "DCF77", [Online]. Available: <https://www.ptb.de/cms/en/ptb/fachabteilungen/abt4/fb-44/ag-442/dissemination-of-legal-time/dcf77.html> [Accessed 10 February 2026]

Strukturmonitoring mit passivem RFID-Dehnungssensor: Flexible Lösung für Industrie und Infrastruktur

*Thomas Frank¹, Stefan Hermann¹, André Grün¹, Manuel Kermann¹, Michael Hintz¹, Andrea Cyriax¹,
Klaus Ettrich¹, Peter Peitsch², Reinhard Jurisch², Sebastian Kahl²*

*¹CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Germany,
tfrank@cismst.de*

²Micro-Sensys GmbH, In der Hochstedter Ecke 2, 99098 Erfurt, Germany

Zusammenfassung

Die Zustandsüberwachung von Bauwerken und Anlagen erfordert zuverlässige Systeme zur Früherkennung struktureller Veränderungen. Der vorgestellte passive RFID-Dehnungssensor kombiniert hochempfindliche Silizium-Dehnungsmessstreifen (Si-DMS) mit einer Wheatstone'schen Messbrücke, einen Temperatursensor zur Kompensation sowie drahtlose Energieversorgung und Datenübertragung. Das System wird direkt auf Bauwerke montiert und ermöglicht kontaktlose Messungen ohne Batterie oder Verkabelung. Die Kalibrierung erfolgt mittels FEM-Simulation und experimenteller Prüfstandversuche. Die erreichte Empfindlichkeit beträgt $90 \mu\text{V}/(\text{V}\mu\epsilon)$. Die digitale Auflösung beträgt über 13 Bit. Derzeit ist noch eine Einzelkalibrierung erforderlich; angestrebt wird eine Messspannenvarianz von $\pm 1\%$ innerhalb einer Charge. Das wartungsarme System eignet sich besonders für vorausschauende Instandhaltung und Langzeitmonitoring schwer zugänglicher Strukturen.

Keywords: Strukturmonitoring, Dehnungssensor, RFID/NFC-Transponder, Si-DMS, Temperaturkompensation.

Einleitung

Die Zustandsüberwachung von Bauwerken, Maschinen und Anlagen ist ein wesentlicher Bestandteil moderner Instandhaltungs- und Sicherheitskonzepte. In Industrie und kritischer Infrastruktur steigt der Bedarf an zuverlässigen Verfahren, mit denen sich strukturelle Veränderungen frühzeitig erkennen lassen, bevor es zu Funktionsstörungen, Schäden oder ungeplanten Ausfällen kommt. Besonders bei tragenden Bauteilen ist es wichtig, mechanische Belastungen, Dehnungen und mögliche Veränderungen des Tragverhaltens kontinuierlich oder zumindest in definierten Intervallen zu erfassen. Auf diese Weise können Instandhaltungsmaßnahmen gezielt geplant und Risiken für Betrieb und Sicherheit reduziert werden.

Ein zentrales Ziel der Strukturüberwachung besteht darin, nicht nur offensichtliche Schäden wie Risse oder Brüche zu detektieren, sondern auch frühzeitige Vorstufen wie plastische Verformungen, lokale Überlastungen oder schleichende Materialermüdung zu erkennen. Gerade bei komplexen Bauwerken oder schwer zugänglichen Anlagen ist dies mit herkömmlichen visuellen Inspektionen nur

eingeschränkt möglich. Messbasierte Verfahren bieten hier einen entscheidenden Vorteil, da sie objektive und reproduzierbare Daten liefern und Entwicklungen über längere Zeiträume dokumentieren können.

Vor diesem Hintergrund gewinnen drahtlose Sensorsysteme zunehmend an Bedeutung. Sie ermöglichen eine direkte Erfassung mechanischer Größen an den relevanten Belastungspunkten, ohne dass aufwendige Verkabelungen oder regelmäßige manuelle Eingriffe erforderlich sind. Dies ist insbesondere dort von Vorteil, wo Nachrüstungen schwierig, Wartungszyklen teuer oder Umgebungsbedingungen rau sind. Ein robustes, wartungsarmes System zur Dehnungsmessung kann daher einen wichtigen Beitrag zur vorausschauenden Instandhaltung leisten und die Lebensdauer von Bauwerken und Anlagen verlängern.

Die in diesem Artikel beschriebene Lösung basiert auf einem passiven, drahtlos auslesbaren Sensorsystem mit integrierter Dehnungsmessung. Es kombiniert die Vorteile einer präzisen mechanischen Messung mit einer kompakten, energieautarken und für den industriellen Einsatz geeigneten Transpondertechnik. Dadurch entsteht ein Ansatz, der sowohl für

die einmalige Zustandsbewertung als auch für das periodische Langzeitmonitoring genutzt werden kann.

Systemdesign

Das beschriebene Sensorsystem kombiniert die Vorteile objektiver Messdaten mit einer flexiblen und wartungsarmen drahtlosen Technik. Im Mittelpunkt stehen hochempfindliche Silizium-DMS-Dehnungssensoren (Si-DMS), die auf einem passiven, hermetisch verkapselten RFID/NFC-Transponder kombiniert sind. Der Transponder enthält neben dem piezoresistiven Sensor und Speicher für Mess- und Kalibrierdaten auch einen Temperatursensor zur Kompensation von Umgebungseinflüssen. Die Energie erhält die Messeinheit ausschließlich drahtlos vom Lesegerät; eine Batterie ist nicht notwendig.

Der Transponder wird direkt am Bauwerk installiert und kann durch Punktschweißen oder Kleben befestigt werden. Dadurch ist der Einbau einfach und auch bei schwer zugänglichen Konstruktionen möglich. Während der Inspektion liest das Lesegerät die gespeicherten Daten kontaktlos aus. Alternativ kann das System auch für periodisches Langzeitmonitoring eingesetzt werden.

Mechanische Struktur

Die mechanische Struktur des Sensors basiert auf einem Si-DMS, der als hochempfindliches Messelement in den mechanischen Aufbau integriert ist. Entscheidend ist dabei, dass die Dehnung der Trägerstruktur möglichst verlustarm auf den Sensor übertragen wird, damit ein reproduzierbares und empfindliches Messsignal entsteht. Die mechanische Kopplung zwischen Bauwerk, Montageträger und Si-DMS bestimmt wesentlich die Qualität der Messung. Die Si-DMS ermöglicht den Aufbau neuer leistungsfähigerer Messsysteme. Der Si-DMS beinhaltet 4 zu einer Wheatstone'schen Messbrücke kombinierte Messwiderstände. Dies hat den Vorteil, dass Störeinflüsse, die auf alle vier Messwiderstände gleichermaßen wirken, das Messergebnis nicht verfälschen. Voraussetzung dafür sind eine symmetrische Brückenschaltung sowie eine identische Einwirkung äußerer Einflüsse auf alle Messwiderstände.

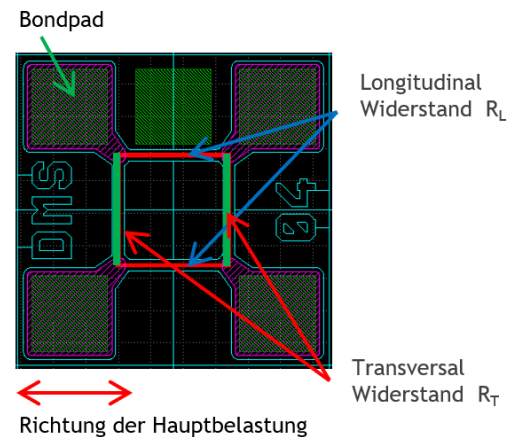


Abb. 1: Vollbrücke auf dem Si-DMS-Chip

In der Abb. 1 ist das Layout des Si-DMS mit integrierter Vollbrücke dargestellt. Basis ist n-Silizium mit implantierten Messwiderständen aus p-Silizium. Bei der dargestellten Hauptbelastung werden die roten Messwiderstände longitudinal und die grünen transversal belastet. Die Abb. 2 zeigt die resultierende Schaltung.

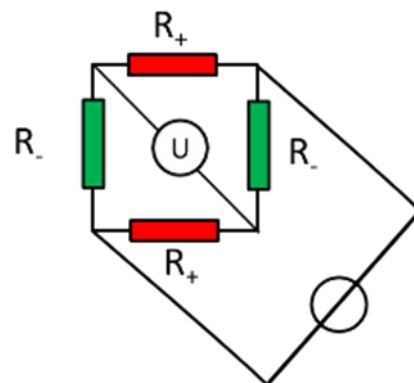
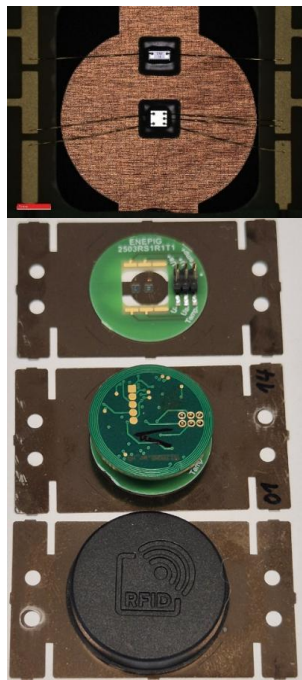


Abb. 2: Wheatstone'sche Messbrücke

Der Si-DMS ist auf einem Montageträger befestigt, der die auftretende Dehnung aufnimmt und an das Sensorelement weitergibt. Für das Fügen des Si-DMS wird eine Glasfritte verwendet, eine stabile Verbindungsart für dauerhafte und mechanisch belastbare Verbindung. In der Abb. 3 ist der Aufbau des Sensor-TAGs dargestellt. Die thermische Dehnung des Montageträgers ist an den Werkstoff des Bauelementes angepasst, Federstahl 1.4310, bzw. C-Stahl 1.1274.



Dehnungssensor, Si-DMS Träger: Si-DMS, Temperaturdiode auf einem flachen Stahlträger, Detail

Dehnungssensor, Si-DMS Träger mit Kontakt-PCB

RFID-Frontend Modul: HF- RFID- Frontend, Antenne, einen Mikrocontroller (μC), Powermanagement, und einer Messelektronik inkl. High End ADC

Si-DMS Sensormodul mit Gehäuse zum Einsatz am Bauwerk

Abb. 3: Aufbau des Sensor-TAGs

Kalibrierung des Sensor-TAGs

Die Abb. 4 zeigt den verwendeten Montageträger mit dem Si-DMS. Die Bohrungen sind die Kraftangriffspunkte für den Zugversuch in der Prüfmaschine.

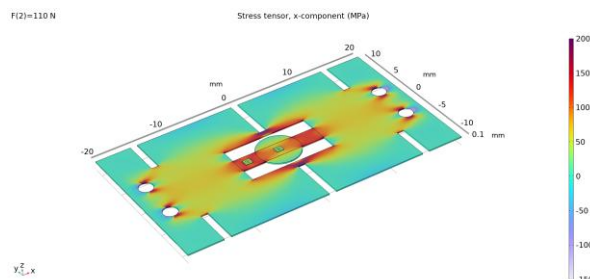


Abb. 4: FEM-Simulation des Montageträger mit gefügtem Si-DMS. Messkraft 10 N und 110 N
Signal Si-DMS = 40 mV/V
Dehnung = 455 $\mu\epsilon$
Stress σ_x = 92 MPa

In der Mitte des Montageträgers sind der Si-DMS und ein Temperatursensor platziert. Optional kann der Träger in der Mitte mit einem Steg versehen werden. Der Steg dient zur mechanischen Entkopplung von Längs- und Querdehnung. Er stellt sicher, dass vorwiegend die Dehnung in Hauptbelastungsrichtung auf den Si-DMS übertragen wird, während Querdehnungseffekte (Poisson-Effekt) unterdrückt werden. Dies erhöht die Messgenauigkeit bei einachsiger Belastung und reduziert Querempfindlichkeiten.

Aus der Simulation geht die Dehnung bei gegebener Belastung hervor. Diese Belastung, die zugehörige Dehnung und das Signal des Si-DMS werden für die Kalibrierung verwendet. Anschließend wird der Sensor-TAG auf einen Testkörper gefügt, um die Dehnungsübertragung vom Bauwerk auf den Sensor zu validieren. Es wird davon ausgegangen, dass die Dehnung vom Testkörper nur mit geringen Verlusten auf den Sensor-TAG übertragen wird. Zum Nachweis wird die FEM-Simulation in der Abb. 5 herangezogen.

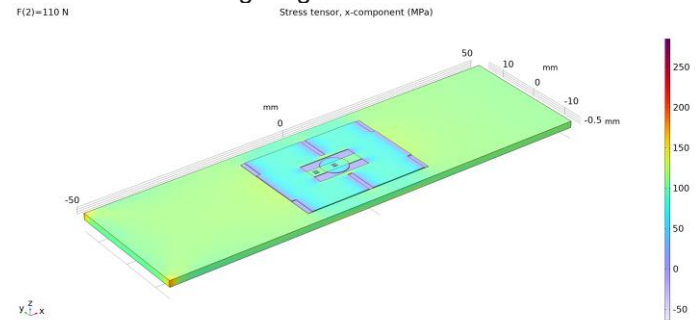


Abb. 5: Montageträger gefügt mit Epoxyklebstoff auf den Testkörper, Signal = 41 mV/V
Dehnung = 471 $\mu\epsilon$
Stress σ_x = 92 MPa

Vor dem Fügen wurden die Anschlussbohrungen für den Kraftangriff gekappt. Die Belastung wird so gewählt, dass der Stress an der betreffenden Stelle, Position des Si-DMS, des ungefügten Montageträgers gleich dem Stress im Testkörper ist.

Die Simulation zeigt, dass die Dehnung und das Signal vergleichbar sind. Dies ermöglicht die Aufnahme der Kraft-Signalkennlinie am Sensor-TAG und die Übertragung auf das zu untersuchende Bauwerk.

Evaluierung des Systems

Zur Evaluierung wurden zu Beginn die Montageträger analog zu Abb. 4 geprüft.

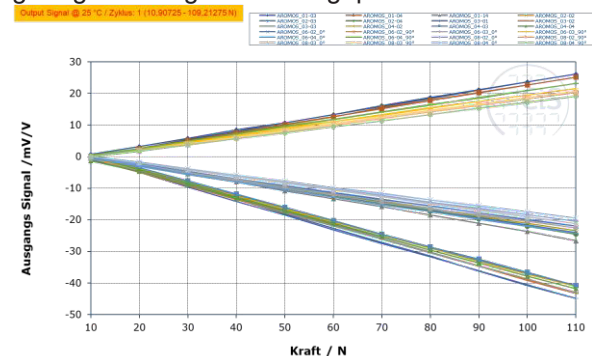


Abb. 6: Kraft-Signalkennlinie 3 verschiedener Sensor-TAGs

In der Abb. 6 sind die Kraft- Signalkennlinien 3 verschiedener Sensor-TAGs dargestellt. Die Streuung der verschiedenen Sensor-TAGs erfordert zurzeit noch eine Einzelkalibrierung. Angestrebt wird eine Variation der Messspanne von $\pm 1\%$ über die gesamte Charge, damit kann auf eine 100% Kalibrierung verzichtet werden.

Aus der Messung wird die Sensitivität S_0 der Sensor-TAGs bestimmt. Anschließend werden die Träger auf die Testkörper gefügt. Die Verfahren sind Verschraubung über die Kraftanriffspunkte, Kleben und Punktschweißen.

In diesem Zustand wird zuerst die Temperaturabhängigkeit des Nullpunktes, TKN, bestimmt, der TKN beträgt im Mittel:

- $TKN=50 \mu V/(V \cdot K)$

Die Ursache für die Temperaturabhängigkeit des Nullpunktes ist der Materialmix des Aufbaus am Bauwerk infolge der verschiedenen CTE (Coefficient of Thermal Expansion, Wärmeausdehnungskoeffizienten) der Materialien und diese Wirkung auf den Si-DMS.

Anschließend werden die Testkörper analog zu Abb. 5 geprüft. Das Klebeverfahren erreichte die beste Übereinstimmung der Dehnung im Montageträger zu der Dehnung im Testkörper. Weitere Tests wurden mit dem Klebeverfahren durchgeführt.

Für den Prototypen sind die Werte für die Kalibrierung:

- Empfindlichkeit $S_0 = 90 \mu V/(V \cdot \mu\epsilon)$ bezogen auf Stahl
- $TKN=50 \mu V/(V \cdot K)$

Um die mechanischen Spannungen durch die Montage zu kompensieren wird nach der Montage, bei Start der Messung der Signalwert

- Sig_0 =Signal bei Start der Messung und für die Temperaturkompensation die Starttemperatur t_{Start} gespeichert

- t_{Start} = Temperatur bei Start der Messung

Berechnung der Temperaturkorrektur und des Messwertes:

- $Sig_{Korr} = (t_{Start} - t_{ist}) \cdot TKN$
- $Messwert = (Sig_{ist} - Sig_{Korr} - Sig_0) \cdot S_0$

Signalverarbeitung

Die Elektronikinheit besteht aus einem ADU mit Verstärker, einem Mikrocontroller, einem EEPROM dem RFID-Frontend sowie der RFID-Antenne, vgl. Abb. 7. Mittels dieser Komponenten werden die Sensoren angesteuert und die analogen Signale des Si-DMS und des Temperatursensors ausgelesen. Die Digitale Auflösung ist besser als 13 Bit. Die Reichweite beträgt maximal 50 mm. Zur Evaluierung wurden die Aufbauten mit einer Prüf-

maschine belastet und das Signal des Sensor-TAGs gemessen.

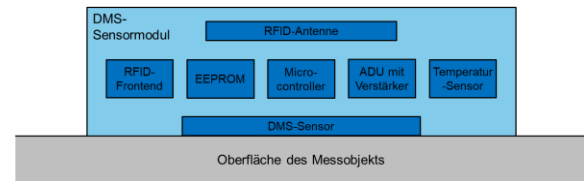


Abb. 7: Blockschaftbild des Sensorsystems

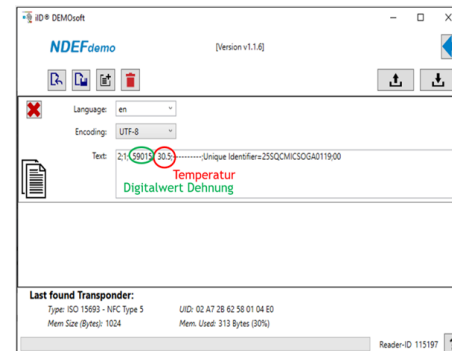


Abb. 8: Auswertung: Messwert in $\mu V/V$ und Temperatur in $^{\circ}C$

Die Abb. 8 zeigt das GUI der Testsoftware. Die Ergebnisse sind in der Abb. 9 dargestellt.

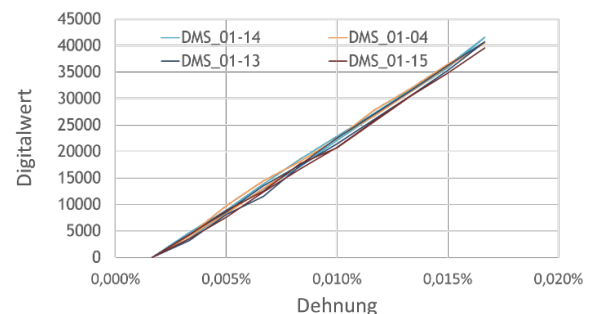


Abb. 9: Ergebnis der Evaluierung von 4 Sensor-TAGs

Literaturnachweis

- [1] Medienresistenter Absolutdrucksensor ohne Ölfüllung und Frontbündigem Medienkontakt, Thomas Frank, et al., 23. ITG/GMA-Fachtagung Sensoren und Messsysteme 2026

Danksagung

Diese Arbeit wurde im Rahmen des Forschungsprojekts AROMOS (Autarkes RF-basiertes Oberflächen-Montierbares-Dehnungsmesssystem) durchgeführt und vom BMW unter dem Förderkennzeichen KK5085709RH4 finanziell unterstützt.

Medienresistenter Absolutdrucksensor ohne Ölfüllung und frontbündigem Medienkontakt

Thomas Frank, Stefan Hermann, André Grün, Manuel Kermann, Andrea Cyriax, CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Germany tfrank@cismst.de

Zusammenfassung

In diesem Beitrag wird die Entwicklung absoluter Drucksensoren mit frontbündigem Medienkontakt, die für extrem kalte und hochmedienbeständige Anwendungen (z. B. Flüssigwasserstoff in der Raumfahrt) geeignet sind beschrieben. Statt herkömmlicher piezoresistiver Drucksensoren, die bei Kryotemperaturen mit Ölfüllung und komplexem Aufbau unhandlich sind, werden Silizium-Dehnungsmessstreifen (Si-DMS) in einer kompakteren, hochauflösenden Schaltung eingesetzt. Die Si-DMS werden per angepasster Glasfritte auf eine Frontplatte integriert und über Laserschweißverbindungen hermetisch mit einem Edelstahlgehäuse verbunden. Durch gezielte Anordnung der Sensorwiderstände in definierten Spannungsfeldern und Verwendung querdehnungsunempfindlicher Si-DMS mit geeigneter Kristallorientierung wird eine hohe Messgenauigkeit bei gleichzeitig robustem, mediengetrenntem Aufbau erreicht. Zur Bewertung werden die Temperaturkoeffizienten des Brückenwiderstands, des Nullpunktes und der Empfindlichkeit untersucht. Durch optimale Prozessparameter konnten diese kompensiert und minimiert werden. Drei Varianten für 100 bar relativ und 10 bar absolut in Edelstahl oder Titan wurden realisiert. Die Druckkennlinien zeigen geringe Hysterese ohne Offset bei 0 bar, die sich durch alternierende Belastung reduzieren lassen. Nach einer Lagerung bei 4 K über 2 h bleiben elektrische und mechanische Parameter unverändert, was die Zuverlässigkeit des Sensorkonzepts für luft- und raumfahrttechnische Anwendungen mit Flüssigwasserstoff unterstreicht.

Keywords: Medienkontakt, medienresistent, frondbündig, Silizium, Dehnungsmessstreifen, piezoresistiv, Drucksensorik

Einleitung

Im Rahmen der Druckmesstechnik gewinnt die zuverlässige Erfassung von Drücken unter extremen Einsatzbedingungen zunehmend an Bedeutung, insbesondere in kryogenen Anwendungen und im Umgang mit Wasserstoff. In industriellen und prozesstechnischen Anwendungen haben sich hierfür piezoresistive Drucksensoren mit Dünn- oder Dickschichttechnologie auf Stahl- oder Keramikkörpern etabliert. Sie weisen jedoch konstruktiv bedingte Einschränkungen bei frontbündigen und differenzdruckfähigen Ausführungen auf.

MEMS-Drucksensoren sind grundsätzlich geeignet, benötigen aber vielfach einen komplexen Aufbau mit Trennmembran und Ölfüllung. Für kryogene Temperaturen bei gleichzeitiger hoher Medienbeständigkeit, insbesondere mit Ölfüllung, ist dieser Aufbau nicht geeignet.

Siliziumbasierte Dehnungsmessstreifen (Si-DMS) eröffnen hier zusätzliche Freiheitsgrade im Sensordesign und ermöglichen kompakte, hochauflösende Messanordnungen mit anpassbarer mechanischer Integration.

Der in dieser Arbeit vorgestellte Ansatz beschreibt die Entwicklung eines Absolutdrucksensors mit frontbündigem Medienkontakt auf Basis chemikalien- und tieftemperaturbeständigen Edelstählen. Auf diese Weise sind sie für Einsatztemperaturen bis 40 K und Lagertemperaturen bis hinunter zu 4 K ausgelegt. Kern des Aufbaus ist eine Frontplatte, auf der Si-DMS mittels Glasfritte gefügt und zu einer Vollbrücke verschaltet werden. Die Frontplatte ist über eine lasergeschweißte Verbindung hermetisch mit dem Sensorkörper gekoppelt. Durch die gezielte Anordnung der Messwiderstände in Bereichen definierter mechanischer Spannungen sowie den Einsatz querdehnungsunempfindlicher Si-DMS mit geeigneter Kristallorientierung wird eine hohe Messgenauigkeit bei gleichzeitig robustem, mediengetrenntem Aufbau angestrebt.

Das entwickelte Sensorkonzept adressiert insbesondere Anwendungen in der Luft- und Raumfahrt, in denen Flüssigwasserstoff als Arbeits- oder Treibmedium eingesetzt wird und Absolutdruckmessungen zur Detektion von Phasenübergängen und zur Überwachung variierender Umgebungsdrücke erforderlich sind.

Dabei steht neben der Druckmessgenauigkeit insbesondere das Verhalten der Si-DMS im kryogenen Temperaturbereich bis hin zum Ausfrieren der Ladungsträger sowie die Langzeitstabilität des Verbunds aus Silizium, Glasfritte und Edelstahl im Vordergrund.

Dehnungssensoren aus Silizium

Dehnungsmessstreifen (DMS) sind Bauelemente zur elektrischen Erfassung mechanischer Spannungen. In Kombination mit geeigneten Verformungskörpern lassen sich damit physikalische Größen wie Kraft, Druck oder Drehmoment messen.

Im Vergleich zu metallischen DMS bieten Silizium-Dehnungsmessstreifen (Si-DMS) entscheidende Vorteile: Sie besitzen eine bis zu 40-fach höhere Empfindlichkeit, zeigen kaum Kriecheffekte und ermöglichen dadurch eine deutlich präzisere Messung. Die verwendeten Si-DMS besitzen eine Empfindlichkeit von 0,4 - 0,5 mV/(V MPa), diese entspricht einem k-Faktor von 80 - 100.

Zudem sind sie in der Fertigung kosteneffizient, während die Herstellung herkömmlicher Metall-DMS weiterhin aufwendig bleibt. Ein weiterer Vorteil ergibt sich aus der Montage mittels Glasfritte. Infolge dessen Einsatztemperaturen bis 300 °C erreicht werden.

Durch diese Eigenschaften setzen Si-DMS neue Maßstäbe in der Sensortechnik und ermöglichen leistungsfähigere Messsysteme. Zur Evaluierung neuer Aufbautechnologien dienen klassische Drucksensorkörper aus verschiedenen Stahlsorten und Keramiken als Referenz.

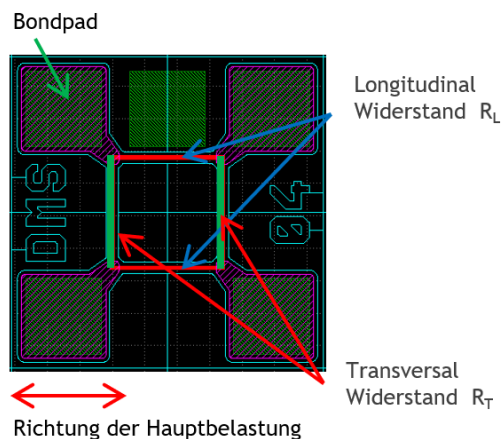


Abb. 1 Vollbrücke auf dem Si-DMS- Chip

In Sensorsystemen werden die Si-DMS in der Regel als Wheatstone'sche Messbrücke verschaltet. Dies hat den Vorteil, dass Störeinflüsse, die auf alle vier Messwiderstände gleichermaßen wirken, das Messergebnis nicht verfälschen. Voraussetzung dafür sind eine

symmetrische Brückenschaltung sowie eine identische Einwirkung äußerer Einflüsse auf alle Messwiderstände.

In der Abb. 1 ist das Layout des Si-DMS mit integrierter Vollbrücke dargestellt. Basis ist n-Silizium mit implantierten Messwiderständen aus p-Silizium. Bei der dargestellten Hauptbelastung werden die roten Messwiderstände longitudinal und die grünen transversal belastet. Die Abb. 2 zeigt die resultierende Schaltung. In der Abb. 3 und Abb. 5 sind zur Illustration eine einfache Simulation einer druckbelasteten Biegeplatte dargestellt.

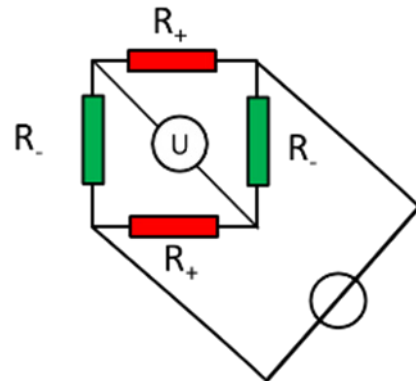


Abb. 2: Wheatstone'sche Messbrücke

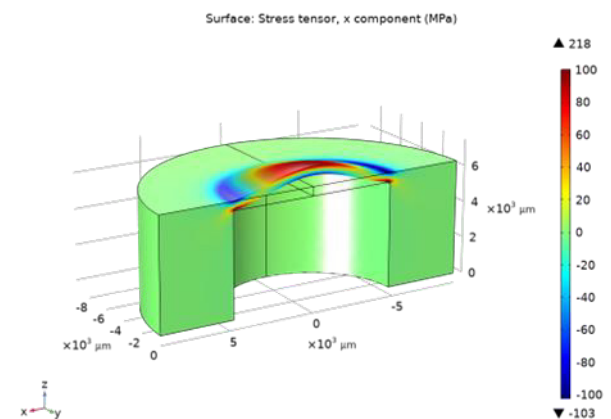


Abb. 3 Simulation einer Biegeplatte eines Drucksensors [1]

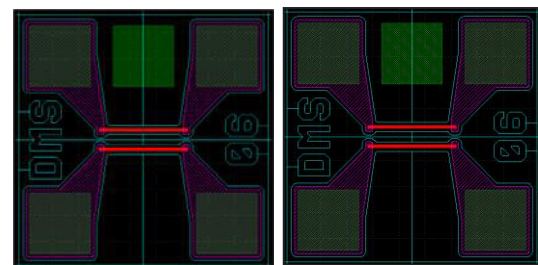


Abb. 4: 2 Paar Einzelwiderstände, Vollbrücke aus 2 Chips im Druckspannungsbereich und Zugspannungsbereich

Aus der Abb. 5 kann die jeweilige Position der verschiedenen Si-DMS-Varianten auf der Platte abgelesen werden, c) für den Si-DMS mit Vollbrücke. Der Si-DMS mit integrierter Vollbrücke ist optimal am Rand platziert, maximaler Stress in der Hauptrichtung solid.sx und minimaler Stress der Nebenrichtung solid.sy. In der Mitte der Biegeplatte arbeitet dieser Chip nicht, da solid.sx und solid.sy den gleichen Betrag besitzen und alle 4 Messwiderstände identisch belastet werden. Die Brückenschaltung und die gleichzeitige longitudinale und transversale Belastung führt zu einer fast vollständigen Kompensation des Messsignals.

In der Abb. 6 sind Piezoeffizienten in Abhängigkeit von der Dotierung und der Kristallorientierung dargestellt. Besonders wichtig ist dies für die Positionierung der Si-DMS in der Mitte der Biegeplatte. An dieser Position sind Messwiderstände notwendig, welche eine hohe Querdehnungsunempfindlichkeit besitzen. Das bedeutet, keinen Transversaleffekt. Messwiderstände aus p-Silizium auf <110> Material in Richtung <110>.

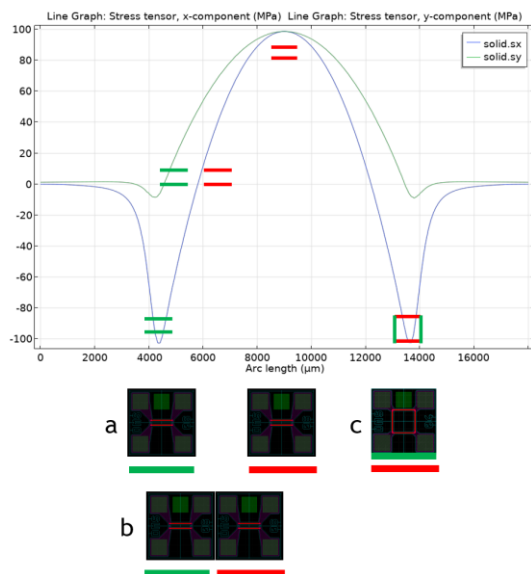


Abb. 5: Stress in der Biegeplatte, Position des Si-DMS: a) und b) 2 DMS mit je 2 Messwiderständen im Druckspannungsbereich und Zugspannungsbereich, vorzugsweise nur longitudinal belastet, beides auf der Biegeplatte eines Drucksensors angeordnet und als Vollbrücke verschaltet, c) 1 Chip mit einer Vollbrücke, 2 Transversalwirkende und 2 Longitudinalwirkende Messwiderstände.

In dieser Position ist der Transversaleffekt minimal, bei hinreichend großem Longitudinaleffekt. Aufgrund des fehlenden Transversaleffektes sind mit diesem Material sind nur

Einzelwiderstände, keine integrierten Vollbrücken, möglich. Diese werden als Paar eingesetzt, siehe Abb. 4.

In der Abb. 5 sind 2 Möglichkeiten, a) und b), mit diesen Einzelwiderständen dargestellt. 2 Si-DMS mit je 2 Messwiderständen im Druckspannungsbereich und Zugspannungsbereich, vorzugsweise nur longitudinal belastet. Die Position ist besonders wichtig da die maximale mechanische Spannung in der Mitte der Platte ausgenutzt werden kann. Dies ist nur möglich, wenn der Messwiderstand eine hohe Querdehnungsunempfindlichkeit besitzt, keinen Transversaleffekt.

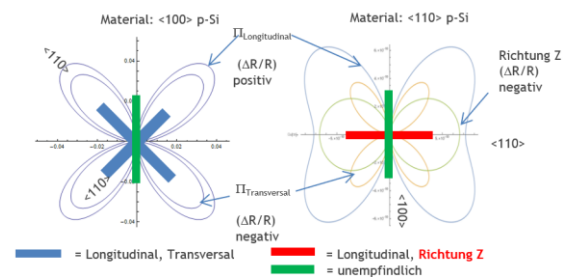


Abb. 6: Piezoeffizienten in Kristallrichtungen

$$\text{pSi } \langle 100 \rangle: \quad \Pi_L = 7,2 \cdot 10^{-10} \text{ Pa}^{-1}, \quad \Pi_T = -6,1 \cdot 10^{-10} \text{ Pa}^{-1}$$

$$\text{pSi } \langle 110 \rangle: \quad \Pi_L = 7,2 \cdot 10^{-10} \text{ Pa}^{-1}, \quad \Pi_T = -1,1 \cdot 10^{-11} \text{ Pa}^{-1}$$

Gefertigt werden diese auf 150 mm großen Siliziumwafern in einem MEMS-Prozess, einzeln durch einen „Dicing Before Grinding“ Prozess (DBG). Die Dicke beträgt nach dem Grinding 15 - 20 μm .



Abb. 7: Drahtgebondeter Si-DMS, gefügt mit Glasfritte auf Stahlrunde (1.4542), Prozesstemperatur ca. 420 °C [2]

Diese Chips werden mit einer angepassten Glasfritte auf die Stahlkörper/-ronden gefügt. Die Glasfritte liegt als druckfähige Paste vor und wird mit Schablonendruck auf die Ronden aufgetragen.

Durch Drahtbonden wird der Chip mit der Signalauswertung verbunden. Die Abb. 8 zeigt ein fertiges Testmodul.

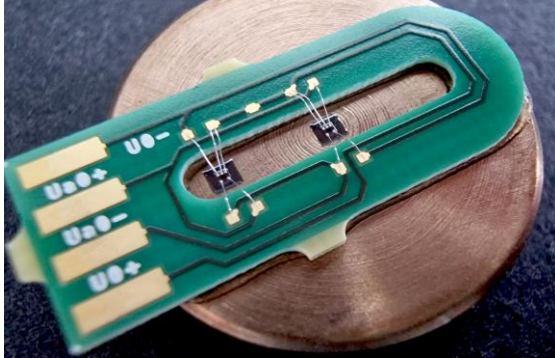


Abb. 8: Testaufbau mit 2 Si-DMS auf einer Stahlrunde (1.4542) und PCB

Der gesamte Fügeprozess ist nur auf einer ebenen Fläche wirtschaftlich durchzuführen. In Kavitäten ist das Fügen nur mit großem Aufwand möglich.

Ölfreier Absolutdrucksensor mit frontbündigem Medienkontakt

Für einen ölfreien Absolutdrucksensor mit frontbündigem Medienkontakt wurden folgende Randbedingungen abgeleitet.

- Ebene Biegeplatte zum Fügen der Si-DMS
- Si-DMS und die elektrische Kontaktierung befindet sich im Referenzdruckvolumen
- Medienberührende Materialien, beständiger Stahl
- Eignung für kryogene Temperaturen

In der Abb. 9 ist der Aufbau vereinfacht dargestellt. Basis bildet die Biegeplatte, der Medienkontakt.

Die Si-DMS werden auf diese gefügt und mittels eines Flexleiters elektrisch angeschlossen, siehe Abb. 10a. Diese Platte wird mit dem Gehäuse durch Laserschweißen hermetisch dicht gefügt. Der Flexleiter wird mit dem Abschlusssockel verbunden. Final wird der Abschlusssockel unter Vakuum mit dem Gehäuse hermetisch dicht verschweißt.

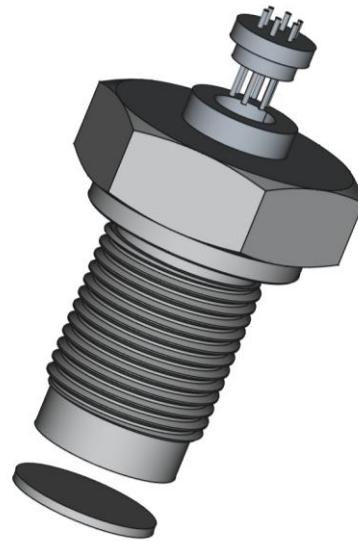
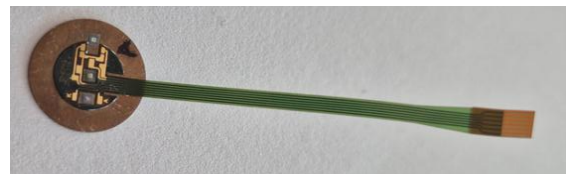


Abb. 9: Aufbau Ölfreier Absolutdrucksensor mit frontbündigem Medienkontakt, Biegeplatte, Gehäuse mit Anschlussgewinde, Abschlusssockel mit hermetisch dichten Glasdurchführungen

Durch die innenliegenden Si-DMS sind die Position a) und c) aus Abb. 5 nicht möglich, da die erforderliche Stressverteilung auf der Innenseite nicht vorhanden ist. Die Mitte der Platte ist, aufgrund der hohen Empfindlichkeit ideal, aber nur für die querdehnungsunempfindliche Variante. Den Gegenpart bildet der erste Chip aus Position b). Die Verteilung ist nicht ideal symmetrisch, zeichnet sich aber durch eine sehr gute Empfindlichkeit aus.



a)



b)

Abb. 10: a) Platte Ø 16 mm, Dicke 0,25 mm, b) gefügte Platte und Abschlusssockel.

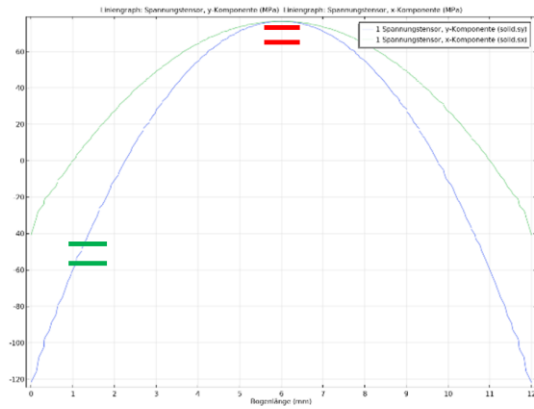


Abb. 11: Position der Si-DMS auf der Unterseite der Platte.

Bewertung der Sensoren

Zwei wichtige Eigenschaften bei Drucksensoren sind eine hohe Druckabhängigkeit und eine geringe Temperaturabhängigkeit des Signals.

Zur Bewertung des Temperatureinflusses werden mehrere Koeffizienten bestimmt:

- Temperaturkoeffizient des Brückenwiderstandes: TKR
- Temperaturkoeffizient des Nullpunktes: TKN
- Temperaturkoeffizient der Empfindlichkeit, Sensitivität: TKS

Der vorliegende Aufbau zeigt, dass diese Koeffizienten von einer Reihe von Parametern abhängen. Die wesentlichen Temperatureinflüsse sind:

- Die temperaturbedingte Widerstandsänderung der Messwiderstände im Nullpunkt: Die Wirkung des TKR auf das Messsignal wird nahezu vollständig durch die Brückenschaltung kompensiert. Der Brückenwiderstand ändert sich mit der Temperatur, der Nullpunkt bleibt konstant. Der TKR kann durch Variation der Prozessparameter im Bereich von 0,5 bis ca. 3,3 %/10K angepasst werden. Das Optimierungskriterium ist in der gewählten Schaltung die Temperaturkompensation des TKS.
- Die thermisch induzierten mechanischen Spannungen durch die verschiedenen Temperaturkoeffizienten (CTE) der Materialien bewirken durch den piezoresistiven Effekt eine Änderung der Widerstände der Messwiderstände. Im Idealfall, bei einem perfekten Glasfritprozess, wirken identische mechanische Spannungen auf alle Messwiderstände. Dadurch ist auch die Widerstandsänderung identisch und dies wird durch die Messbrücke kompensiert. Der TKN ist bei guter Prozessführung, Chipfertigung und Fügetechnik minimal.

- Die Piezokoeffizienten sind von der Temperatur abhängig. Bei der gewählten Technologie fallen diese mit steigender Temperatur, bezogen auf den ungefügen Si-DMS. Der TKS ist so eine Eigenschaft der Messwiderstände im Silizium. Insbesondere der TKR ist bei der Kompensation des TKS wichtig.

Ein hoher TKN deutet auf einen nicht optimalen Aufbau hin, einen hohen Offset im Nacktchip und ungünstige Lage in der Glasfritte. Der dadurch verursachte unterschiedliche Stress in den Messwiderständen führt zu unterschiedlichen Änderungen der Messwiderstände und folglich zu einer temperaturabhängigen Änderung des Nullpunktes.

Es wurden drei Varianten nach der beschriebenen Technologie aufgebaut.

- 100 bar relativ, Stahl 1.4404 (4 Stück)
- 10 bar absolut, Stahl 1.4404 (1 Stück)
- 10 bar absolut, Titan Grade 5 (2 Stück)

Die aufgebauten Drucksensoren weisen ein typisches Verhalten auf.

Das Diagramm, Abb. 12 zeigt die Druck-Kennlinie bei 30°C bis in den Überlastbereich von 200 bar. Auffällig ist die leichte Druckhysterese, Abb. 13. Es wird die Differenz im Signalverlauf vom Nullpunkt zum Nenndruck und zurück aufgetragen. Wichtig ist, dass es am Nullpunkt, 0 bar, kein Gap gibt. Dies würde auf eine starke plastische Verformung bei der Belastung hinweisen. Durch eine alternierende Belastung bis 100% Überlast, 200 bar, konnte die Hysterese bereits reduziert werden.

Die Diagramme Abb. 14 bis Abb. 16 zeigen die Kennwerte der Absolutdrucksensoren.

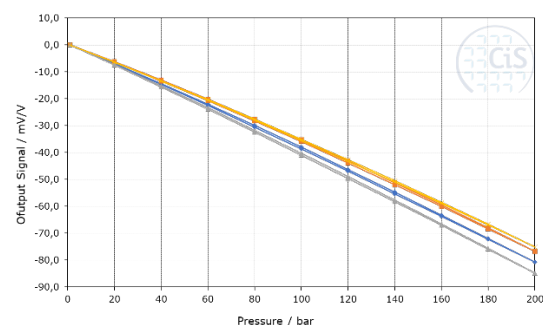


Abb. 12: Im Nullpunkt korrigierte Kennlinie, Nenndruck 100 bar, Überlast 200 bar, Temperatur: 30°C, 100 bar relativ, Stahl 1.4404

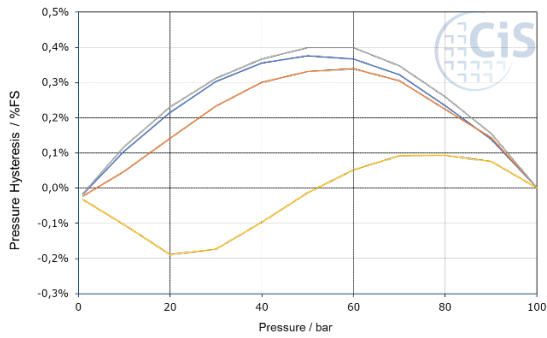


Abb. 13: Druckhysterese 0 bar bis 100 bar, Temperatur: 30°C, 100 bar relativ, Stahl 1.4404

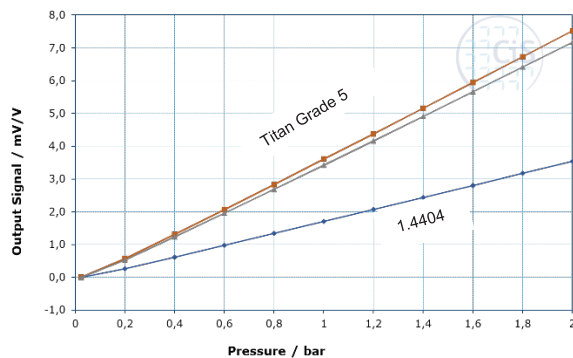


Abb. 14: Im Nullpunkt korrigierte Kennlinie, Nenndruck 10 bar, Teilbereich 2 bar, Temperatur: 30°C.

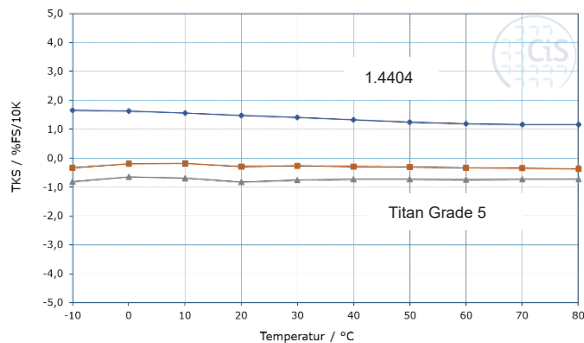


Abb. 15: Temperaturkoeffizient der Sensitivität, TKS

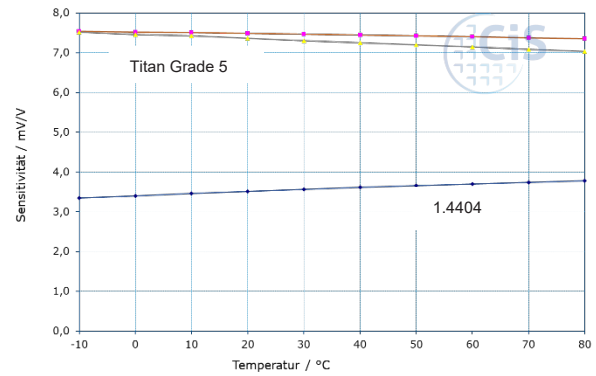


Abb. 16: Sensitivität bei 2 bar, Temperatur -10°C bis 80°C

Anschließend wurde eine Lagerung bei 4 K über 2 Stunden durchgeführt. Dies haben die Sensoren gut überstanden. Die Parameter der Sensoren vor und nach der Lagerung waren gleich. Die mitgelockten elektrischen Parameter zeigten ebenfalls keine Ausfälle oder kurzzeitige Aussetzer. Zur Funktionalität bei unter 50 K konnten noch keine Untersuchungen durchgeführt werden.

Literaturnachweis

- [1] Thomas Frank, André Grün, Christian Maier, Andrea Cyriax, M. T. Asghar, Michael Hintz, U. Krieger, Thomas Ortlepp, "Siliziumdehnungssensoren zur Spannungsanalyse", MikroSystemTechnik Kongress, Ludwigsburg, Deutschland, 08.-10.11.2021
- [2] André Grün, "Innovative Fügetechnologien zum Aufbau von feuchte- und temperaturstabilen Kraftsensordsystemen basierend auf piezoresistiven Silizium-Dehnmessstreifen", Deutsche IMAPS-Konferenz 2021, München, Deutschland, 21.-22.10.2021.

Monocular Inter-UAV Distance Estimation Using Machine Learning and Geometric Scaling

Robert Pommeranz¹, Kevin Tebbe¹, Jakob Fischer¹, Ralf Heynicke¹, Gerd Scholl¹

*¹Electrical Measurement Engineering
Helmut-Schmidt-University, Hamburg, Germany
robert.pommeranz@hsu-hh.de*

Summary

Reliable Inter-UAV distance measurement is a prerequisite for formation control and collision avoidance in cooperative Unmanned Aerial Vehicle (UAV) swarms. This paper presents the system integration and experimental characterization of a monocular distance estimation pipeline that combines a fine-tuned YOLOv8s object detector with pinhole-model geometric scaling to convert bounding-box pixel widths into real-time depth estimates. The pipeline runs on an NVIDIA Jetson Orin NX companion computer via DeepStream SDK. The camera is intrinsically calibrated, and the effective projected width of each target UAV is determined from short-range hover reference measurements rather than nominal dimensions. An uncertainty propagation model is derived that combines camera calibration uncertainty, target width variability, and detector bounding-box jitter into a depth confidence interval. Two independent ground truth sources are used for validation: Laser Range Finder (LRF) measurements at thirteen waypoints (5 m – 50 m) with a DJI Matrice 4T target, and dual-rover RTK GNSS during a dynamic reference flight with a Holybro X650 target. A mean absolute error (MAE) below 2 m is achieved for ranges up to 35 m, with the DJI Matrice 4T. The pipeline requires only a low-cost camera and runs on embedded hardware, making it directly applicable to vision-based swarm coordination on weight- and power-constrained platforms.

Keywords: UAV, Object Detection, Distance Estimation, Machine Learning, Monocular Vision

Introduction

The rapid expansion of commercial and custom-built Unmanned Aerial Vehicles (UAVs) has fundamentally altered both civil airspace and the modern battlefield [1]. Swarm-based Counter Unmanned Aerial System (C-UAS) architectures have enabled multiple cooperative interceptors to engage simultaneous threats through coordinated behavior. A prerequisite for such swarms is continuous, online knowledge of the range to each neighboring UAV to fly in formation, avoid collisions, or coordinate based on mission requirements [2], [3]. Passive monocular ranging is equally relevant when an unknown UAV is detected by an active sensor such as radar, as the camera has a higher update rate and can provide continuous range tracking without generating further detectable emissions.

High accuracy ranging sensors like Laser Range Finder (LRF), stereo cameras or Ultra-Wideband (UWB) radio, achieve centimeter-level precision, but their cost, weight, and power budget can be prohibitive on smaller platforms [3]. A monocular camera, by contrast, is already present on most aerial vehicles. When the physical dimensions of the tracked target are known, the pinhole projection model yields a

depth estimate from a single frame without any cooperative signal [4].

Recent advances in real-time object detection [5], [6] and hardware-accelerated inference on companion computers like Nvidia Jetson now enable detection, tracking and distance estimation online and with low latency. Nevertheless, the accuracy of the geometric approach degrades quadratically with range, and its practical limits have not been systematically characterized at inter-UAV operational ranges. This paper addresses the gap by presenting the system integration and experimental characterization of a monocular inter-UAV distance estimation pipeline deployed on an embedded Jetson Orin NX mounted on a Holybro X650 quadcopter [7]. The distance estimation itself follows the pinhole projection model, and the contribution lies in the complete system realization and field validation. This work integrates a Camera Serial Interface (CSI) camera, YOLOv8s inference, object tracking, and geometric distance estimation into a single hardware-accelerated pipeline. Furthermore, the uncertainties of the model are quantified following the Guide to the Expression of Uncertainty in Measurement (GUM), combining camera calibration

uncertainty, target-width variability, and empirical detector pixel jitter [8]. Experimental validation characterizes the practical limits of this approach.

Related Work

Detecting smaller UAVs in real time on embedded hardware is a challenging task because computational resources are limited. One-Stage detectors are the state of the art for real-time applications due to their efficiency. This approach is established for example with YOLO [5]. Open-source datasets are available for training and benchmarking on UAV detection. A popular dataset is the Anti-UAV, which focuses on single-object tracking and comprises 160 video sequences of infrared (IR, 640 px × 512 px) and electro-optical (EO, 1920 px × 1080 px) imagery, of which approximately 100 annotated sequences are provided as training data [9]. The Drone-vs-Bird Challenge provides another dataset which consists of 77 annotated video sequences or 106.466 single images in different environments and small UAVs [10]. The focus of this dataset is that flying birds are as background clutter at larger distances.

Estimation of depth from a single image is a vision problem which has been solved by two main approaches: geometric and learning-based. The geometric approach relies on the pinhole-model scaling, which provides a solution for depth estimation when the target dimensions are known, or a true depth value is available at least once [4], [11]. For instance, Pang et al. [12] utilized a known-size marker to measure distance, while Bui et al. [13] extended this method to simultaneously measure distance and angle. Idohen et al. [14] achieved sub-millimeter errors at short range through curve fitting. However, these studies primarily evaluate their methods at short, controlled ranges, and the accuracy of geometric approaches at operational UAV engagement distances, where the quadratic depth-error growth dominates, has received less attention. Learning-based approaches, on the other hand, rely on large, annotated datasets with depth labels for training. Vajgl et al. [15] integrated a distance regression head into YOLO (Dist-YOLO), achieving 11 % mean relative error up to 150 m on the KITTI dataset [16]. However, this dataset is not representative of aerial UAV

scenarios, and the features in the images are different from those in the KITTI dataset.

Identifying the same object across frames even when the detection is partially lost is crucial for maintaining a constant bounding box around the target. Methods based on correlation filters, such as Discriminative Correlation Filter (DCF), are computationally efficient and can run in real time on embedded hardware [17]. Deep association trackers like DeepSORT leverage appearance features for improved tracking performance, but at a higher computational cost [18]. NVIDIA's DeepStream SDK provides GPU-accelerated implementations of such algorithms (NvDCF, NvSORT, NvDeepSORT) [19].

Methodology

This section describes the algorithmic pipeline for monocular UAV distance estimation, deployed on a Holybro X650 observer platform [7]. The pipeline proceeds in three stages: image capture, hardware-accelerated detection and tracking on the companion computer, and geometric distance estimation with uncertainty propagation.

The camera is calibrated using Zhang's method, which is implemented in OpenCV, recovering the intrinsic matrix K (focal length f_x , f_y and principal point c_x , c_y) and distortion coefficients [20]. Prior to distance estimation, each detected bounding box is undistorted using these coefficients so that the pixel width w_{px} is consistent with ideal pinhole geometry.

The detection and tracking pipeline runs on the Nvidia Jetson Orin NX using DeepStream SDK 7.1 for hardware-accelerated video processing. A YOLOv8s model [6] is fine-tuned on the Drone-vs-Bird Detection Challenge dataset [10]. The dataset is partitioned into training (80 %), validation (10 %), and test (10 %) subsets. Training uses 1280 px × 1280 px input resolution, batch size 32, and early stopping after 50 epochs without validation improvement (max 100 epochs). The trained weights are exported to ONNX and compiled into a TensorRT engine (FP16) using DeepStream-YOLO¹. For tracking, the NvDCF algorithm was selected for its high accuracy and efficiency [19].

The YOLO detection configuration is set so that the detection happens on every second frame to reduce the computational load, while the

¹ <https://github.com/marcoslucianops/DeepStream-Yolo> (accessed 23.03.2026)

tracker maintains the bounding box on the skipped frames and holds the object for up to ten frames without a new detection. This configuration was chosen to reduce the GPU load while maintaining a high bounding-box update rate.

With the calibrated focal length f_x and the known physical width W of the target UAV, the distance along the optical axis is estimated from the pixel width w_{px} using the pinhole camera model:

$$Z = \frac{f_x \cdot W}{w_{px}} \quad (1)$$

Substituting $w_{px} = f_x W / Z$ into the derivative $\frac{\partial Z}{\partial w_{px}} = -f_x W / w_{px}^2$, the following error approximation is obtained:

$$\Delta Z \approx -\frac{Z^2}{f_x W} \Delta w_{px} \quad (2)$$

showing that the depth error grows quadratically with distance.

The combined relative depth uncertainty is:

$$\frac{u_c(Z)}{Z} = \sqrt{\left(\frac{u(f_x)}{f_x}\right)^2 + \left(\frac{u(W)}{W}\right)^2 + \left(\frac{u(w_{px})}{w_{px}}\right)^2} \quad (3)$$

In Eq. (3) the only variable term is $u(w_{px})/w_{px}$, because bounding box decreases with distance and the object width and focal length stay constant in this setup. The baseline calibration records the pixel-width empirical standard deviation $s(w_{px})_{5m}$ over 1000 frames at five meters as a Type A evaluation; the standard uncertainty is set equal to this value $s(w_{px})_{5m} = u(w_{px})_{5m}$. As the target moves further away and its projected pixel width decreases, the absolute pixel-width jitter will also decrease. Therefore, the uncertainty $u(w_{px})$ is modeled as a proportional function (Eq. (4)) of the detected width, where $\rho = u(w_{px})_{5m} / w_{px,5m}$ is the relative jitter ratio, and a minimum jitter of 1.5 px is imposed.

$$u(w_{px}) = \max(1.5, \rho \cdot w_{px}) \quad (4)$$

The uncertainty of the target width is modeled as a Type B evaluation using a uniform distribution between the motor-to-motor and tip-to-tip propeller spans.

$$u(W) = \frac{W_{max} - W_{min}}{\sqrt{12}} \quad (5)$$

Two RTK GNSS rovers provide the 3D baseline between the observer and target platforms. Applying the Law of Propagation of Uncertainty $D = \sqrt{\Delta E^2 + \Delta N^2 + \Delta U^2}$ yields the UAV distance uncertainty:

$$u(D_{RTK}) = \frac{\sqrt{2}}{D} \sqrt{u_h^2(\Delta E^2 + \Delta N^2) + u_v^2 \Delta U^2} \quad (6)$$

where u_h and u_v are the per-rover horizontal and vertical position uncertainties. The factor $\sqrt{2}$ arises because the baseline is the difference of two independent rover positions, doubling the variance. Horizontal and vertical components are propagated separately since RTK receivers specify distinct accuracies for each axis.

The DJI Matrice 4T LRF specifies a range-dependent accuracy of $\pm(a + bD)$ [m] for a distance $D > 3$ m. Modelling this as a uniform distribution over the stated error interval and converting it to standard deviation gives Eq. (7), where a is a constant offset and b (dimensionless) is the range-dependent coefficient:

$$u(D_{LRF}) = \frac{a + bD}{\sqrt{3}} \quad (7)$$

To validate the system, the residual $\Delta = Z - D_{GT}$ between the camera estimate and the ground truth (reference) is evaluated using $u_c(\Delta) = \sqrt{u^2(Z) + u^2(D_{GT})}$. The measurement is considered acceptable if the absolute residual falls within the expanded uncertainty bounds ($|\Delta| \leq U(\Delta)$), where $U(\Delta) = k \cdot u_c(\Delta)$ using a coverage factor of $k = 2$.

Experimental Setup

The observer, shown in Fig. 1, is a Holybro X650 quadcopter equipped with a Jetson Orin NX companion computer. A Raspberry Pi Camera Module V2 (Sony IMX219, 8 MP, rolling shutter, F/2.0, 3.04 mm focal length, 62.2° HFOV) is used to capture the images. It operates at 1280 px × 720 px @60fps in a forward-facing mount, pitched upward from horizontal to keep the target in the field of view during level flight [21]. The intrinsic camera calibration yields $f_x = 1291.9$ px with a Type A standard uncertainty $u(f_x) = 0.47$ px. The RTK GNSS antenna is mounted on top of the frame; the camera optical center is offset from the

antenna by approximately $d_{cam} \approx (0.30, 0, -0.15)$ in the body frame (forward, right, down) in meter. Because the target adjusts its altitude to remain near the image center during the trials, the optical-axis depth closely approximates the Euclidean range, and no additional pitch correction is applied.



Fig. 1: UAV equipped with camera, RTK GNSS antenna, and Nvidia Jetson, to estimate the distance.

Two UAVs are used as targets: a Holybro X650 (motor-to-motor span: 0.65 m; tip-to-tip including propellers: 1.03 m) and a DJI Matrice 4T (motor-to-motor span: 0.39 m; tip-to-tip including propellers: 0.66 m). Based on Eq. (5) the mean and uncertainty values for both UAV dimensions (width) are shown in Tab. 1.

Tab. 1: Target width and uncertainty values

Target	$\bar{W}$ [m]	$u(W)$ [mm]
DJI M4T	0.53	78.0
X650	0.84	110.0

The uncertainty of the reference systems is derived from manufacturer specifications and error propagation. Both X650 UAVs carry a u-blox ZED-F9P dual-frequency GNSS receiver; a third unit serves as RTK base station, streaming RTCM3 corrections [22]. In RTK Fixed mode the manufacturer specifies the position accuracy as: $\leq 0.01 \text{ m} + 1 \text{ ppm}$ (CEP horizontal; 50th-percentile vertical). Converting to standard uncertainty gives $u_h = 8.5 \text{ mm}$ and $u_v = 15 \text{ mm}$ per rover. Since the baseline is the difference between two independent rover positions, the combined Euclidean distance uncertainty ranges from $\sqrt{2}u_h \approx 12 \text{ mm}$ for a purely horizontal baseline to $\sqrt{2}u_v \approx 21 \text{ mm}$ when the two platforms differ in altitude.

According to the manufacturer datasheet, the DJI Matrice 4T integrated LRF achieves a range-dependent accuracy of $0.2 + 0.0015D$ for more than three meters. Using the equation (7), the corresponding uncertainty ranges from 0.12 m at 5 m to 0.14 m at 50 m.

Results and Discussion

A precision of 0.984, recall of 0.970, mAP@0.5 of 98.8 %, and mAP@0.5:0.95 of 78.4 % are achieved by the fine-tuned YOLOv8s model on the test set. The mAP@0.5 is sufficient to provide reliable bounding-box information for the distance estimation in Eq. (1). However, the mAP@0.5:0.95 value indicates imprecise localization at stricter Intersection over Union (IoU) thresholds, which translates directly into bounding-box width jitter. Fig. 2 shows both UAVs at 5 m distance. These images are chosen to visualize the fact that the bounding box does not always cover the UAV correctly.

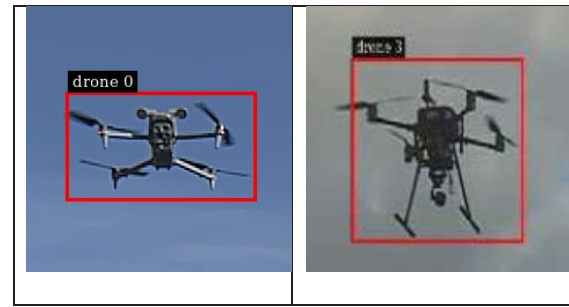


Fig. 2: Bounding-Box at 5m of the DJI M4T (left) and the Holybro X650 (right)

The DJI Matrice 4T was tracked at thirteen hovering waypoints in the range of 5 m – 50 m using the target width $W = 0.53 \text{ m}$, $u(W) = 78.0 \text{ mm}$ from Tab. 1. Beyond 50 m, detection became unreliable and the target was lost. The per-waypoint measurement and uncertainty values are summarized in Tab. 2. Across the complete static-hover measurement, the estimator achieves a Mean Absolute Error (MAE) between 0.24 m and 7.86 m, depending on the distance, and a mean signed error (bias) of -1.3 m. This systematic bias, which grows with range according to Eq. (2), is fully covered within the uncertainty bound at every distance. Fig. 3 shows the boxplot of every waypoint and the absolute error is below 2 m up to 35 m.

Tab. 2: Measurement and uncertainty values with DJI M4T and LRF

Dist. [m]	$\bar{w}_{bbox}$ [px]	$U(k=2)$ [m]	MAE [m]
5	164.0	1.58	0.24
7	120.3	2.17	0.15
9	92.3	2.77	0.31
11	76.1	3.38	0.33
13	63.7	3.98	0.50
15	55.0	4.59	0.59
20	42.5	6.22	0.34
25	34.7	7.94	0.64
30	28.8	9.78	0.90
35	25.0	11.74	1.10
40	23.0	13.84	2.85
45	21.2	16.10	4.56
50	20.3	18.53	7.86

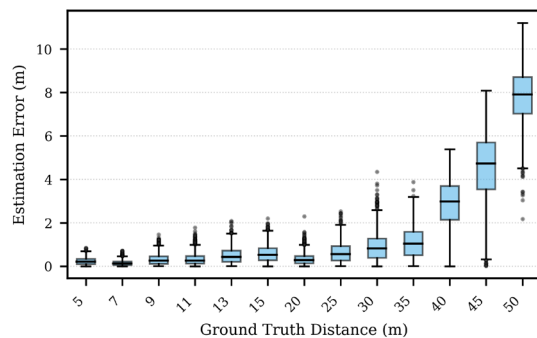


Fig. 3: Absolute error of the distance estimation across all thirteen waypoints

The second measurement involves two Holybro X650 UAVs, each equipped with a u-blox ZED-F9P RTK receiver. One UAV executed a predefined reference flight, departing from approximately 5 m at a constant velocity of 2 m/s to a maximum distance of 95 m and returning to the starting point. The target width $W = 0.84$ m, $u(W) = 110.0$ mm from Tab. 1 was used directly. The residual plot in Fig. 4 reveals a systematic negative bias which grows over distance. The MAE during the flight is 3.12 m, and the bias is -3.01 m.

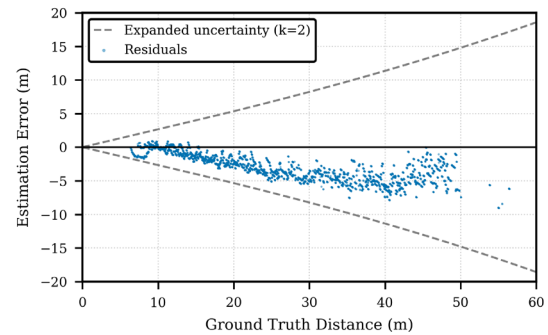


Fig. 4: Residual of the raw camera estimates across the reference flight (RTK GNSS)

Conclusion

This paper presents the system integration and experimental characterization of a monocular inter-UAV distance estimation pipeline combining a fine-tuned YOLOv8s object detector with pinhole-model geometric scaling, deployed entirely on an embedded NVIDIA Jetson Orin NX companion computer. An uncertainty propagation model was derived to predict the per-frame depth confidence from camera calibration uncertainty, target UAV width variability, and detector bounding-box jitter. The pipeline was validated against two independent ground truth sources.

In the LRF reference measurement trails the per-waypoint MAE remained below 1.1 m out to 35 m, rising to 4.56 m at 45 m and 7.86 m at 50 m. Therefore, all waypoints satisfy the predicted uncertainty.

With the RTK GNSS reference flight, the estimator achieved a MAE of 3.12 m, and detected the UAV up to 50 m. As with the previous measurement, the measurement values lie in the predicted uncertainty range.

This experiment has shown that a target UAV with a width of at least 0.6 m can be ranged out to approximately 50 m, where the bounding-box size decreases to less than 20 px, which is the limit of the detector. The pipeline constitutes a lightweight, low-power ranging solution that requires only a single calibrated camera and runs fully onboard.

Acknowledgement

This research is funded by dtcc.bw – Digitalization and Technology Research Center of the Bundeswehr. dtcc.bw is funded by the European Union – NextGenerationEU (project “Digital Sensor-2-Cloud Campus Platform” (DS2CCP), <https://dtccbw.de/home/forschung/hsu/projektds2ccp>).

References

- [1] K. Tebbe, R. Pommeranz, R. Heynicke, and G. Scholl, “From An Engineering Perspective: Small Drones Affecting The Course Of Warfare,” Jul. 2025, doi: 10.5281/ZENODO.15794252.
- [2] J. N. Yasin, S. A. S. Mohamed, M.-H. Hagbayan, J. Heikkinen, H. Tenhunen, and J. Plosila, “Unmanned Aerial Vehicles (UAVs): Collision Avoidance Systems and Approaches,” *IEEE Access*, vol. 8, pp. 105139–105155, 2020, doi: 10.1109/ACCESS.2020.3000064.
- [3] N. Gageik, P. Benz, and S. Montenegro, “Obstacle Detection and Collision Avoidance for a UAV With Complementary Low-Cost Sensors,” *IEEE Access*, vol. 3, pp. 599–609, 2015, doi: 10.1109/ACCESS.2015.2432455.
- [4] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, 2. edition, 17. printing. Cambridge: Cambridge Univ. Press, 2018.
- [5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 779–788. doi: 10.1109/CVPR.2016.91.
- [6] Jocher, Glenn and Chaurasia, Ayush and Qiu, Jing et al., *Ultralytics YOLOv8*. (2023). Ultralytics. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [7] R. Pommeranz, K. Tebbe, R. Heynicke, and G. Scholl, “A Modular and Scalable System Architecture for Heterogeneous UAV Swarms Using ROS 2 and PX4-Autopilot.” doi: 10.48550/arXiv.2510.27327.
- [8] “Evaluation of measurement data — Guide to the expression of uncertainty in measurement.” Accessed: Apr. 15, 2026. [Online]. Available: <https://www.bipm.org/doi/10.59161/JCGM100-2008E>
- [9] N. Jiang et al., “Anti-UAV: A Large-Scale Benchmark for Vision-Based UAV Tracking,” *IEEE Trans. Multimed.*, vol. 25, pp. 486–500, 2023, doi: 10.1109/TMM.2021.3128047.
- [10] A. Coluccia, A. Fascista, L. Sommer, A. Schumann, A. Dimou, and D. Zarpalas, “The Drone-vs-Bird Detection Grand Challenge at ICASSP 2023: A Review of Methods and Results,” *IEEE Open J. Signal Process.*, vol. 5, pp. 766–779, 2024, doi: 10.1109/OJSP.2024.3379073.
- [11] A. Torralba, P. Isola, and W. T. Freeman, *Foundations of computer vision*. in Adaptive computation and machine learning series. Cambridge, Massachusetts London, England: The MIT Press, 2024.
- [12] Y. Pang, Y. Zhao, J. Chen, S. Wang, and H. Chen, “Viewing distance measurement using a single camera,” in *2014 IEEE 7th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China: IEEE, Dec. 2014, pp. 512–515. doi: 10.1109/ITAIC.2014.7065103.
- [13] M. T. Bui, R. Doskocil, and V. Krivanek, “Distance and angle measurement using monocular vision,” in *2018 18th International Conference on Mechatronics - Mechatronika (ME)*, IEEE, Dec. 2018, pp. 1–6. [Online]. Available: <https://ieeexplore.ieee.org/document/8624694>
- [14] E. E. Idohien, O. M. Afolayan, and U. A. Umar, “Object distance estimation using a monovision camera,” *IAES Int. J. Robot. Autom. IJRA*, vol. 12, no. 4, p. 325, Dec. 2023, doi: 10.11591/ijra.v12i4.pp325-331.
- [15] M. Vajgl, P. Hurtik, and T. Nejezchleba, “Dist-YOLO: Fast Object Detection with Distance Estimation,” *Appl. Sci.*, vol. 12, no. 3, p. 1354, Jan. 2022, doi: 10.3390/app12031354.
- [16] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The KITTI dataset,” *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013, doi: 10.1177/0278364913491297.
- [17] A. Lukežič, T. Vojříř, L. Čehovin Zajc, J. Matas, and M. Kristan, “Discriminative Correlation Filter Tracker with Channel and Spatial Reliability,” *Int. J. Comput. Vis.*, vol. 126, no. 7, pp. 671–688, Jul. 2018, doi: 10.1007/s11263-017-1061-3.
- [18] N. Wojke, A. Bewley, and D. Paulus, “Simple online and realtime tracking with a deep association metric,” in *2017 IEEE International Conference on Image Processing (ICIP)*, Beijing: IEEE, Sep. 2017, pp. 3645–3649. doi: 10.1109/ICIP.2017.8296962.
- [19] “NVIDIA DeepStream 7.1 SDK Documentation: Gst-nvtracker.” Accessed: Mar. 25, 2026. [Online]. Available: https://docs.nvidia.com/metropolis/deepstream/7.1/text/DS_plugin_gst-nvtracker.html
- [20] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000, doi: 10.1109/34.888718.
- [21] “Camera - Raspberry Pi Documentation.” Accessed: Apr. 02, 2026. [Online]. Available: <https://www.raspberrypi.com/documentation/accessories/camera.html>
- [22] u-blox, “ZED-F9P-04B Data sheet.” [Online]. Available: https://content.u-blox.com/sites/default/files/ZED-F9P-04B_DataSheet_UBX-21044850.pdf

A Data-Driven Multi-Configuration Parameter based Characterization of Eddy Current Inductive Sensors' Signals using Machine Learning

Aimal Khan , Alexander Hetznecker, Maximilian Kagerer, Thomas Greiner

Institute for Smart Systems and Services, Pforzheim University, Germany

{aimal.khan, alexander.hetznecker, maximilian.kagerer, thomas.greiner}@hs-pforzheim.de

Abstract—Eddy current based inductive sensors are widely used in industrial systems for non-contact material characterization, thickness measurement, and defect detection. However, their deployment typically requires extensive manual calibration of hardware and software parameters, resulting in time-consuming and error-prone configuration procedures. This paper presents a generalized data-driven characterization framework using machine learning, in which raw sensor signals are jointly processed with configuration parameters. A deep neural network learns nonlinear dependencies between sensor settings, electromagnetic responses, and target variables. Experimental validation using an industrial eddy current sensor demonstrates significant improvements in material identification and zinc layer thickness classification accuracy, enabling adaptive and task independent inductive sensing.

I. INTRODUCTION

Eddy current inductive sensors play a central role in industrial automation, quality assurance, and non-destructive testing due to their robustness, contactless operation, and suitability for harsh industrial environments. They are widely deployed in applications such as material identification, coating and layer thickness estimation, crack and defect detection, as well as position and displacement measurement. Compared to optical or mechanical sensing principles, eddy current sensors are largely insensitive to contaminants such as oil, dust, or moisture, which makes them particularly attractive for inline industrial measurements. The underlying operating principle is based on electromagnetic induction. A time-varying magnetic field generated by an excitation coil induces eddy currents in electrically conductive targets. These eddy currents, in turn, generate secondary magnetic fields that interact with the primary field, leading to measurable changes in coil impedance or induced voltage. The resulting sensor signal is strongly influenced by the electrical conductivity, magnetic permeability, geometry, and surface condition of the target material. In practical industrial systems, however, the measured sensor signal is influenced not only by the intrinsic properties of the target object but also by a large number of sensor configuration and signal processing parameters. These include excitation current amplitude and slope, pulse duration, amplification gain, signal offset, sampling rate, oversampling strategy, and digital filtering or preprocessing steps. The combined effect of these parameters introduces complex, nonlinear distortions into the measured signal. As a result, even small changes in configura-

tion can significantly alter the sensor response, making reliable interpretation challenging. Conventional calibration and characterization approach typically rely on manual parameter tuning supported by expert knowledge and empirical testing. Such procedures are time-consuming, task-specific, and difficult to scale across different measurement scenarios. In industrial environments where sensors must be rapidly reconfigured for new materials, geometries, or production conditions, this dependence on manual calibration becomes a major limitation. Furthermore, analytical modeling approaches often fail to capture higher-order nonlinearities, cross-parameter interactions, and real-world disturbances present in industrial settings. To address these challenges, data-driven methods based on machine learning have gained increasing attention in recent years. Various studies have demonstrated the potential of supervised learning algorithms, including support vector machines, random forests, and neural networks, to improve classification accuracy and robustness in eddy current sensing applications. Deep learning approaches, in particular, enable automatic feature extraction directly from raw time-domain or frequency-domain sensor signals. Despite these advances, most existing machine learning based approaches assume a fixed sensor configuration during training and deployment. Changes in excitation parameters or signal conditioning typically require retraining the model or repeating the calibration process. Consequently, generalization across varying sensor configurations remains limited, restricting the applicability of such methods in flexible industrial environments. This paper addresses this gap by proposing a generalized machine learning framework that explicitly incorporates sensor configuration parameters into the learning process. By jointly processing raw sensor signals and configuration metadata, the proposed approach enables the learning model to compensate for parameter-induced signal variations and to generalize across a wide range of operating conditions. This significantly reduces the need for manual recalibration and enables adaptive, task independent inductive sensing suitable for modern industrial applications.

II. RELATED WORK

Inductive [1] and capacitive [2] sensors are used for material identification and material deformation analysis. The inductive sensors work on the principle of electromagnetic induction. It applies a time-varying eddy current pulse to the target

material and performs the task of identification based on the response of the material. However, these sensors are affected by multiple parameters. In order to achieve accurate identification results, calibrations must be carried out for these parameters. Traditional eddy current signal analysis is predominantly based on analytical and semi-analytical electromagnetic models derived from Maxwell's equations. Classical approaches describe the interaction between excitation coils and conductive targets by analyzing variations in impedance, phase, or induced voltage as functions of material properties, geometry, and lift-off distance. Well-established formulations exist for simplified geometries such as infinite plates, cylindrical conductors, and layered structures, enabling thickness measurement, crack detection, and material characterization. For example, simplified inductance and impedance models for single-turn eddy current sensors were developed in [1], providing physically interpretable relations between coil parameters and conductive targets. Such analytical descriptions form the basis of many industrial inductive sensing systems. Machine learning has also been widely applied to sensor calibration problems. Neural-network-based self-calibration algorithms have been proposed for pressure sensors [3], while random forest models have been used to improve calibration accuracy in low-cost air-quality monitoring systems [4]. Data-driven calibration strategies have similarly been applied to optical and light sensors [5]. These studies demonstrate the potential of machine learning to compensate for nonlinearities, drift, and environmental effects in sensor systems. Related research in capacitive and proximity sensing further highlights the trend toward integrating intelligent data processing with physical sensor models [2], [6].

Recent advances in eddy current testing (ECT) have increasingly focused on integrating machine learning techniques to improve automation and accuracy. A comprehensive trend toward intelligent defect detection systems has been observed, where machine learning models are used to replace manual interpretation of eddy current signals. For instance, Sahu et al. demonstrated a random forest-based approach for automated flaw detection in heat exchanger tubes, achieving very high classification accuracy and highlighting the effectiveness of data-driven methods for large-scale inspection tasks [7]. Such approaches significantly reduce human dependency and enable real-time analysis.

In addition to classical machine learning methods, deep learning has gained significant attention in recent years. Convolutional neural networks (CNNs) have been successfully applied to eddy current data by treating signals or impedance trajectories as images. Niu et al. proposed a CNN-based defect recognition framework for weld inspection using ECT data, demonstrating robust performance in distinguishing defective and non-defective regions [8]. Similarly, deep learning architectures have been used for solving inverse problems in ECT, such as defect reconstruction and depth estimation, where nonlinear relationships between signals and physical properties are learned directly from data.

Further developments include CNN-based inversion and

uncertainty-aware models for eddy current testing. For example, a deep learning-based inversion model incorporating uncertainty estimation was proposed in [9], showing improved defect detection performance by combining region-of-interest extraction with CNN classification. In another study, complex-valued CNNs were utilized to reconstruct defect geometries from eddy current signals, demonstrating the ability of deep learning to extract spatial and structural information beyond traditional feature-based approaches [10].

Feature-based machine learning methods also remain relevant, particularly when interpretability is required. A notable approach is the extraction of geometric features from Lissajous impedance curves, followed by classification using multiple machine learning models. Such methods provide a balance between physical interpretability and data-driven modeling, achieving high detection accuracy while maintaining insight into signal characteristics [11].

Moreover, recent studies have emphasized the importance of reliability and evaluation frameworks when applying AI to safety-critical Non-Destructive Testing (NDT) applications. For example, sequence-to-sequence deep learning models and probabilistic evaluation frameworks have been proposed for rail inspection using eddy current testing, addressing the need for trustworthy and certifiable AI systems in industrial environments [12]. These works highlight that, beyond accuracy, robustness and explainability are critical for real-world deployment.

Despite these significant advancements, a common limitation across most existing approaches is the assumption of a fixed sensor configuration. The majority of machine learning models are trained and validated under specific excitation conditions, coil parameters, and signal processing settings. Consequently, their generalization capability across varying configurations remains limited. Even deep learning models, while powerful, typically learn implicit relationships within a fixed parameter space and require retraining when the configuration changes.

In contrast to these approaches, only limited work has explored the explicit incorporation of sensor configuration parameters into the learning process. While multi-modal learning strategies have been investigated in other domains, their application to eddy current sensing remains largely unexplored. This gap motivates the proposed work, which introduces a unified data-driven framework that jointly models sensor signals and configuration parameters, enabling improved generalization across varying operating conditions.

III. SENSOR SYSTEM AND DATA ACQUISITION

The experimental setup is based on an industrial eddy current inductive sensor system provided by Picosense Forum GmbH, designed for high-precision material characterization and thickness measurement in industrial environments. The system consists of an excitation and sensing coil unit, analog front-end electronics, and a configurable digital signal processing chain. Its modular architecture allows flexible adjustment of both hardware related and software related parameters,

making it particularly suitable for systematic investigation of parameter dependent signal behavior. A wide range of sensor configuration parameters can be adjusted through the system interface. These include the excitation current characteristics, such as amplitude and current slope, pulse length of the excitation signal, analog output signal amplification, scan depth, signal offset, and the digital oversampling factor. Each of these parameters influences the temporal shape and amplitude of the measured sensor response, either directly through electromagnetic coupling or indirectly via signal conditioning and digitization. The interaction between these parameters is highly nonlinear, motivating a data-driven analysis approach. Measurements were conducted on a set of metallic test samples representative of typical industrial use cases. The material classification dataset includes aluminum, copper, and iron samples, each with comparable geometric properties and surface conditions. To further demonstrate the capability and limits to the classification, samples with and without geometric modifications, such as holes, were included. In addition, another data set uses zinc-coated steel sheets with nominal coating thicknesses of $10\text{ }\mu\text{m}$, $21\text{ }\mu\text{m}$, and $32\text{ }\mu\text{m}$ were used to evaluate the classification of coating thickness performance. These thickness values were selected to represent realistic industrial tolerances and to challenge the sensor's sensitivity. To ensure coverage of a broad operating space, sensor configuration parameters were varied systematically within their permissible ranges. Rather than relying on a single fixed operating point, parameter combinations were selected to reflect realistic variations encountered during industrial commissioning and operation. This approach enables the acquisition of a diverse dataset that captures both material-induced and configuration-induced signal variations. Each individual measurement produces a time-domain sensor signal consisting of variable length (up to 924 sampled values, for shorter signals zero padding is applied) depending on the configuration parameters, corresponding to the electromagnetic response following excitation. The raw signal samples were directly used as input to the learning algorithm. To incorporate configuration dependency, the corresponding sensor parameter values were appended to the signal samples, forming a high-dimensional combined feature vector. This joint representation allows the learning model to explicitly associate signal characteristics with specific parameter settings. The complete dataset was divided into mutually exclusive training, validation, and test subsets. The training set was used for model parameter optimization, while the validation set supported hyperparameter tuning. The test set was reserved exclusively for final performance evaluation to ensure unbiased assessment of generalization capability.

IV. MACHINE LEARNING METHODOLOGY

To model the nonlinear relationship between raw sensor signals, sensor configuration parameters, and the corresponding target classes, a supervised deep learning approach was employed. Given the high dimensionality of the input data and the complex interactions between electromagnetic effects and configuration dependent signal conditioning, deep neural

networks were selected as a suitable modeling framework. Fully connected feed-forward architectures were chosen due to their flexibility and their proven effectiveness in handling heterogeneous input features without requiring explicit spatial or temporal assumptions.

The applied machine learning methodology is depicted in Fig. 1. To carry out experiments with all possible combinations of the 6 parameters require significant amount of time. Therefore, to simplify the setup, the configuration parameters are split to 8 unique sets and for each set 420 responses are generated as dataset for training and testing. We have chosen only 8 unique parameter sets to have quasi-random behavior for the parameters. For the sake of training, the input to the Deep Neural Network (DNN) consists of three complementary components consisting of $M = 930 + 1$ features: the time-domain sensor signal samples (response), the associated sensor configuration parameters, and the label. The raw sensor signal, comprising up to maximum 924 sampled values per measurement represents the transient electromagnetic response of the target material. The configuration parameters encode the operating point of the sensor system (6 parameters in total), including excitation, amplification, and sampling settings. Prior to training, all input features were normalized to ensure numerical stability and to prevent dominance of individual parameters due to differing value ranges. After the learning is performed, the network only get the test data consisting of the response and the configuration parameters. The network architecture is composed of multiple fully connected hidden layers with decreasing numbers of neurons, forming a bottleneck-like structure. This design enables hierarchical feature extraction, where early layers learn low-level signal characteristics and later layers capture more abstract, task-relevant representations. Rectified Linear Unit (ReLU) activation functions were employed in the hidden layers to introduce non-linearity and to mitigate vanishing gradient issues during training. The output layer uses a softmax activation function to produce normalized class probabilities. Two independent models were trained to address distinct classification tasks. The first model performs material classification, distinguishing between different metallic materials and geometric variations. The second model addresses zinc coating thickness classification, where discrete thickness classes are predicted. Model training was conducted using categorical cross-entropy as the loss function, which is well suited for multi-class classification problems. The Adam optimizer was employed due to its adaptive learning rate capabilities and robust convergence behavior. To reduce the risk of over-fitting, early stopping was applied based on validation loss, terminating training when no further improvement was observed. In addition, validation data were used to monitor generalization performance and to guide architectural and hyperparameter selection. This learning strategy enables the network to jointly model signal characteristics and configuration-induced variations, allowing it to learn compensation mechanisms that would be difficult to capture with analytical models or conventional feature-based approaches.

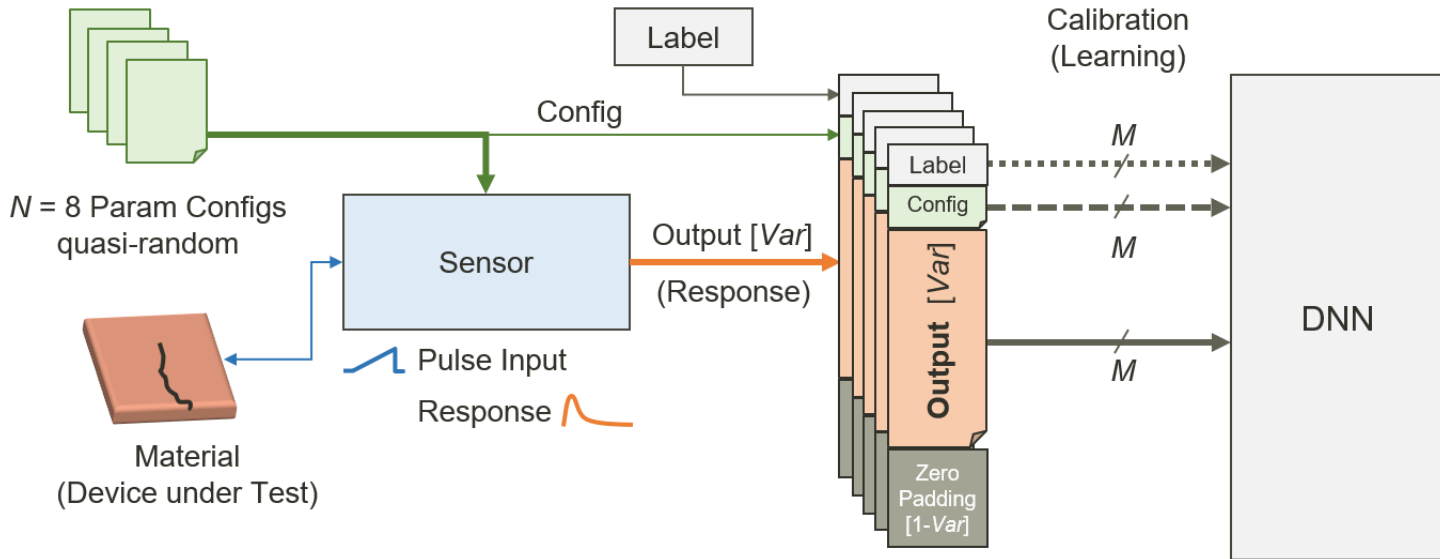


Fig. 1: Machine learning methodology for the calibration learning

V. EXPERIMENTAL RESULTS

A. Material Classification

The performance of the proposed framework was evaluated using two representative classification tasks: material identification and zinc coating thickness classification. These tasks were selected to assess both the sensor's ability to discriminate between materials with different electromagnetic properties and its sensitivity to subtle variations in coating thickness. The proposed neural network architecture is a fully connected feed-forward model designed for multi-class classification. The network accepts an input vector consisting of 930 (924 raw data samples and 6 values for parameters) features that represent the processed signal samples used for classification. The six parameters included as input to the machine learning algorithm are the pulse length (sample count), current slope (slope of the excitation pulse), offset (reference), gain (amplification of the feedback resistor, oversampling rate and, a range extending the value of the second capacitor (1 bit) used for long pulses. The input layer is followed by a sequence of three dense hidden layers with 32, 16, and 8 neurons, respectively. Each hidden layer employs the Rectified Linear Unit (ReLU) activation function to introduce non-linearity and enable efficient learning of complex feature representations while maintaining computational efficiency. This progressively decreasing layer size forms a compact architecture that reduces dimensionality and encourages the extraction of discriminative features. The final layer is a dense output layer with a number of neurons equal to the total number of target classes and uses the softmax activation function to produce normalized probability distributions over all classes. Model performance is evaluated using test accuracy, which measures the proportion of correctly classified samples on an unseen test dataset and provides an overall indication of the generalization capability of the proposed network.

The first evaluation task focused on material classification, and the respective results are provided in Fig. 2 and Fig. 3. Six distinct classes were defined, corresponding to three metallic materials: aluminum, copper, and iron, each measured in two geometric variants with and without geometric modifications. This setup introduces both material-related and geometry-related signal variations by making holes in the materials. The class with holes is denoted by appending H, e.g. for copper without holes we have Cu and with holes we have CuH. Measurements are taken at the target material at defect-free spots and at spots with holes. When the neural network was trained using only raw sensor signals as input features, the model achieved an average classification accuracy of approximately 83% on the test dataset as shown in Fig. 2. Analysis of the confusion matrix revealed that most misclassifications occurred between materials with similar electromagnetic characteristics, particularly in cases where geometric modifications altered the effective eddy current distribution. These results indicate that raw signal information alone is insufficient to fully disentangle material-specific features from configuration-induced distortions. When sensor configuration parameters were included as additional input features, classification performance improved significantly as shown in Fig. 3. The average classification accuracy increased to approximately 94%, representing a substantial relative improvement. Confusion matrix analysis showed a marked reduction in systematic misclassification, particularly for material pairs previously affected by overlapping signal characteristics. This demonstrates that the neural network successfully learned to compensate for variations introduced by different sensor configurations, enabling more reliable material discrimination across multiple operating points.

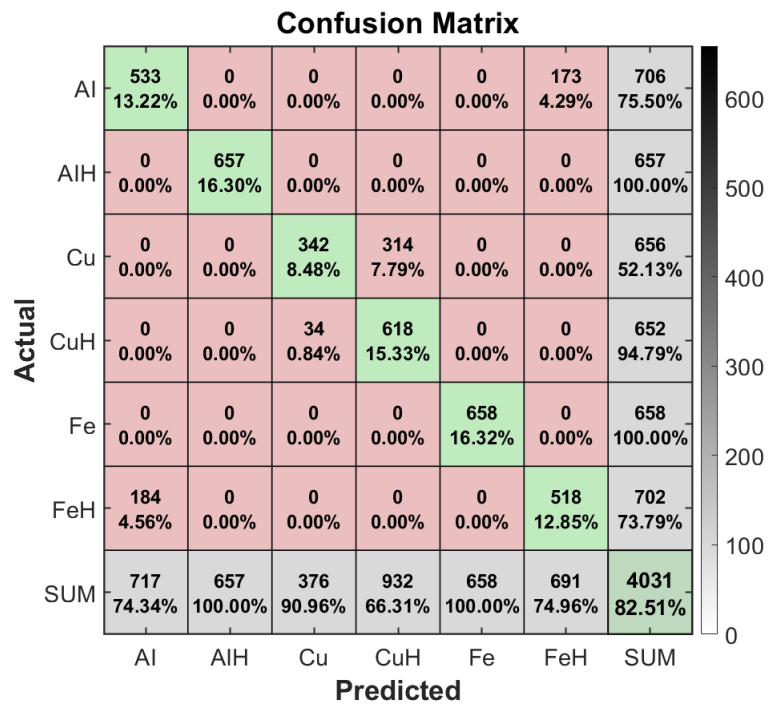


Fig. 2: Confusion matrix for the classification of the metallic materials based on raw signals without inclusion of the calibration parameters in the learning process

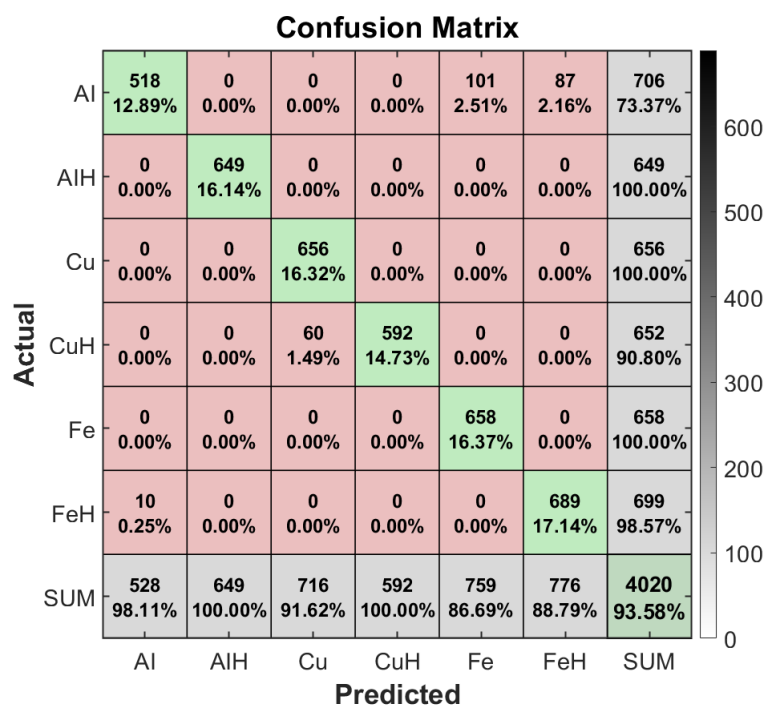


Fig. 3: Confusion matrix for the classification of the metallic materials based on raw signals with inclusion of the calibration parameters in the learning process

B. Zinc Coating Thickness Classification

The second evaluation task addressed zinc coating thickness classification for three nominal thickness values of 10 μm , 21 μm , and 32 μm . This task evaluates the sensor's sensitivity to comparatively small variations in coating thickness, which are of high practical relevance in industrial quality control. Using only raw sensor signals, the neural network already achieved a high classification accuracy exceeding 99% on the test dataset as shown in Fig. 4. This indicates that the eddy current sensor provides strong discriminatory information for thickness estimation under controlled conditions. Nevertheless, minor classification inconsistencies were observed for measurements acquired under differing configuration settings. By incorporating sensor configuration parameters into the input representation, classification robustness was further improved. The resulting model achieved 100% classification accuracy on the test dataset, with no observable misclassifications as shown in Fig. 5. This highlights the benefit of explicitly modeling configuration dependent effects, even in scenarios where baseline performance is already high.

C. Discussion and Results

Overall, the experimental results clearly demonstrate that incorporating sensor configuration metadata significantly enhances model generalization across varying sensor settings. While raw sensor signals contain rich information about material properties, the inclusion of configuration parameters enables the learning model to disentangle material-induced signal characteristics from configuration-induced variations. This leads to more stable and reliable performance across a wide range of operating conditions, which is essential for deployment in flexible industrial measurement systems

VI. CONCLUSION

This paper presented a generalized data-driven characterization framework for eddy current inductive sensors. By integrating raw sensor signals with configuration parameters, a machine learning model was able to achieve robust and task-independent classification performance. Experimental results demonstrated significant improvements in material identification and coating thickness classification accuracy.

VII. ACKNOWLEDGMENT

This work is funded by Carl-Zeiss-Stiftung under project KISA "Hybrid AI for smart sensors and actors in production".

REFERENCES

- [1] U. Tröltzsch, F. Wendler, and O. Kanoun, "Simplified Analytical Inductance Model for a Single Turn Eddy Current Sensor," *Sensors and Actuators A: Physical*, vol. 191, pp. 11–21, 2013.
- [2] R. Moheimani, P. Hosseini, S. Mohammadi, and H. Dalir, "Recent Advances on Capacitive Proximity Sensors: From Design and Materials to Creative Applications," *Carbon Research*, vol. 8, 2022.
- [3] A. M. M. Almassri, W. Z. Wan Hasan, S. A. Ahmad, S. Shafie, and K. Horio, "Self-Calibration Algorithm for a Pressure Sensor with a Real-Time Approach Based on ANN," *Sensors*, vol. 18, 2018.
- [4] N. Zimmerman, A. A. Presto, J. Gu, A. Hauryliuk, E. S. Robinson, and R. Subramanian, "A Machine Learning Calibration Model using Random Forests to improve Sensor Performance for Lower-Cost Air Quality Monitoring," *Atmospheric Measurement Techniques*, vol. 11, no. 1, 2018.
- [5] Y. Zhang, L. O. H. Wijeratne, S. Talebi, and D. J. Lary, "Machine Learning for Light Sensor Calibration," *Sensors*, vol. 21, no. 18, 2021.
- [6] Y. Ding, H. Kisner, T. Kong, and U. Thomas, "Using Machine Learning for Material Detection with Capacitive Proximity Sensors," in *International Conference on Intelligent Robots and Systems (IROS)*, 2020.
- [7] T. K. Sahu, S. Thirunavukkarasu, and A. Kumar, "Intelligent Flaw Detection in Eddy Current Inspection Data Through Machine Learning Model," *Journal of Nondestructive Evaluation*, vol. 44, p. 88, 2025.
- [8] G. Niu, Y. Ji, W. Tang, and B. Zhang, "Defect Recognition for Eddy Current Testing Using Convolutional Neural Network," in *IEEE Industrial Electronics Society Annual Online Conference (ONCON)*, pp. 1–6, IEEE, 2023.
- [9] P. Zhu, Y. Cheng, P. Banerjee, A. Tamburrino, and Y. Deng, "A Novel Machine Learning Model for Eddy Current Testing with Uncertainty," *NDT & E International*, vol. 101, pp. 104–112, 2019.
- [10] B. Chen and T. Yu, "Complex-Valued CNN-Based Defect Reconstruction of Carbon Steel from Eddy Current Signals," *Applied Sciences*, vol. 15, no. 12, p. 6599, 2025.
- [11] X. Zhang *et al.*, "A Novel Feature Extraction Method of Eddy Current Testing for Defect Detection Based on Machine Learning," *NDT & E International*, vol. 107, p. 102108, 2019.
- [12] D. Kanzler, G. Olm, A. Friedrich, and M. Selch, "Evaluation of AI for Eddy Current Testing According to the Reliability Framework for Rail Inspection," *NDT & E International*, vol. 159, p. 103611, 2026.

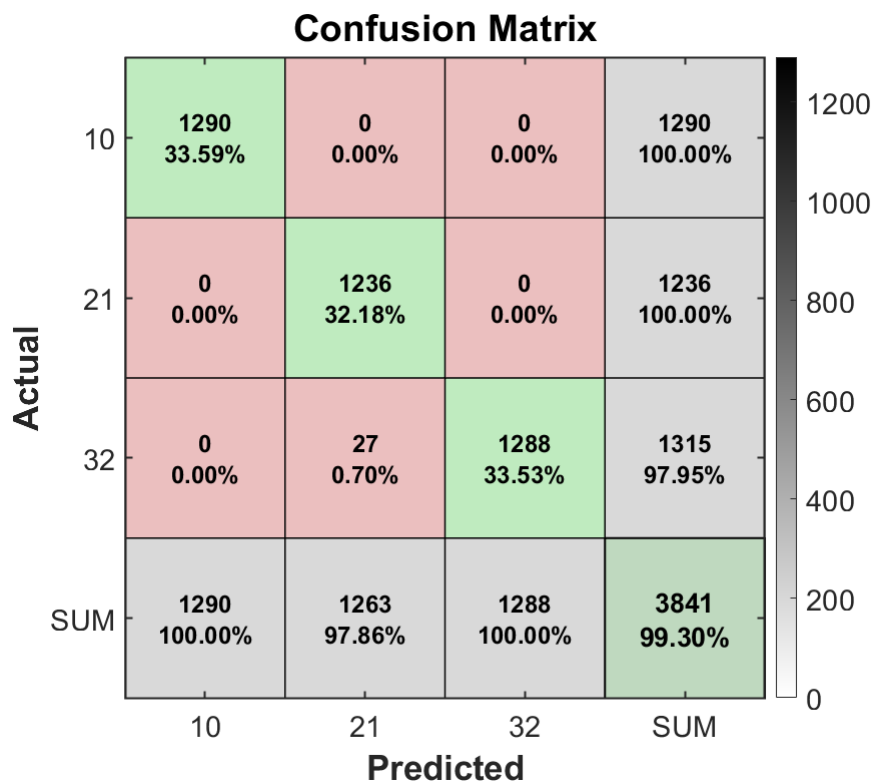


Fig. 4: Confusion matrix for the classification of zinc coating thickness based on raw signals without inclusion of the calibration parameters in the learning process

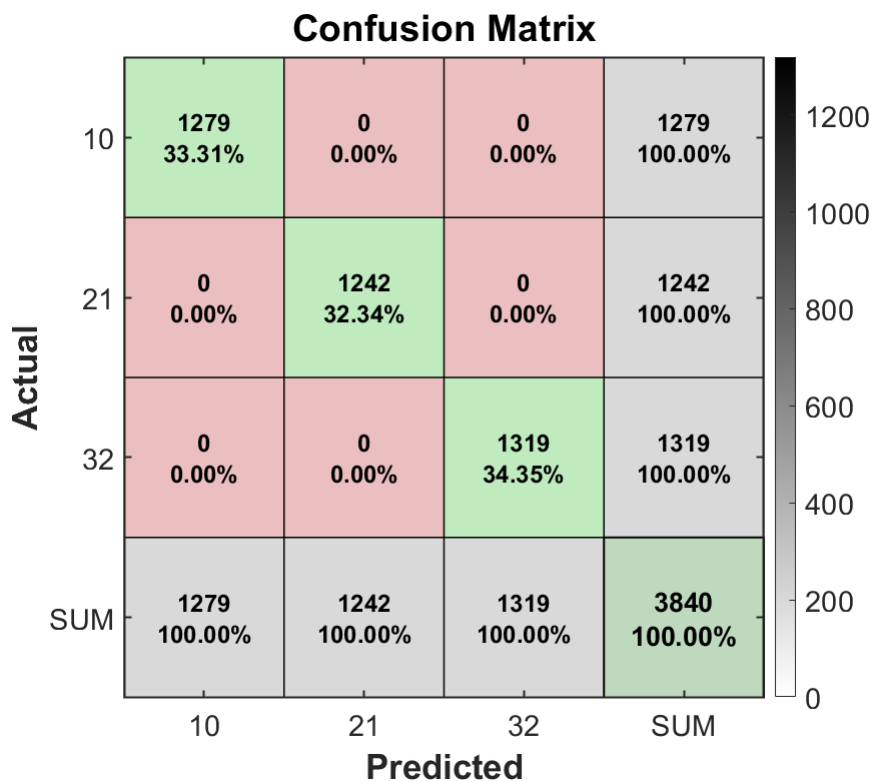


Fig. 5: Confusion matrix for the classification of zinc coating thickness based on raw signals with inclusion of the calibration parameters in the learning process

A Dataset for Condition Monitoring of Screwdriver Bits based on a Piezoelectric Accelerometer for Machine Learning

Houssam El Moutaouakil^{1,2}, Julian Schauer^{1,2}, Andreas Schütze¹

¹*Centre for Mechatronics and Automations Technology, Saarbrücken, Germany*

²*Lab for Measurement Technology, Saarland University, Saarbrücken, Germany*

(h.moutaouakil), (julian.schauer)@zema.de, schuetze@imt.uni-saarland.de

Abstract

To ensure high-precision industrial operations, monitoring assembly processes and tool health using machine learning has become critical. In automated screw assembly, screw bits are vulnerable to progressive wear and sudden damage, which can lead to workpiece defects and costly production downtime if not identified early. Conventional monitoring techniques often rely on torque or angle analysis; however, these typically provide post-process diagnostics rather than real-time insights.

This paper presents a high-resolution vibration dataset designed to advance research in real-time tool wear detection and process monitoring. The data was captured using a piezoelectric accelerometer mounted on an electric screwdriver, which recorded vibrations during both tightening and loosening cycles. To simulate industrial degradation, controlled wear was artificially induced on the bits across multiple incremental stages. The resulting dataset includes three independent experimental series, comprising varied screw hole manufacturing processes and randomized sequences to ensure model robustness and prevent overfitting. By providing well-annotated time-series data in MATLAB format, this resource enables the development of interpretable machine learning models that predict wear stages and recommend proactive tool replacement. Initial validation using time-frequency feature extraction and principal component analysis demonstrates clear signal separability, confirming the dataset's utility for both classification and regression-based condition monitoring.

Keywords: Assembly Process, Condition Monitoring, Interpretable Machine Learning, Process Monitoring, Screwdriver, Screw Bit, Wear Detection, Piezoelectric Accelerometer.

Introduction

Monitoring assembly processes and tool conditions using machine learning has become increasingly crucial to ensure the proper execution of industrial operations [1]. In screw assembly processes, which primarily involve tightening screws with electric screwdrivers, screw bits are subject to wear and damage over time [2]. Detecting their wear stage is essential to prevent workpiece damage and production delays or downtime. Previous studies have primarily focused on analyzing angle or torque signals to assess screwdriving performance [3]. While such approaches can detect bit damage, they only provide a diagnosis after the process is complete, making them inefficient for real-time decision-making [4].

This paper introduces a high-resolution vibration dataset collected from screwdriving processes to support research in tool wear detection and process monitoring. The dataset

enables the development of machine learning (ML) models that predict the wear stage of screw bits and recommend bit replacement before functional failure, thereby improving time efficiency and reducing resource use [5]. The dataset contains well-annotated measurements of both tightening and loosening processes. Controlled wear has been artificially induced on screw bits to represent incremental wear stages.

Experimental Setup

To collect the experimental data, a setup consisting of a screwing testbed, an electric screwdriver, and a data acquisition (DAQ) system is used (Fig. 1) [6]. A piezoelectric accelerometer is mounted non-destructively onto the screwdriver to capture high-frequency vibrations near the bit holder (Fig. 2). These vibration signals are essential to identifying the bit's condition and its level of wear [7]. To process the acquired data, a MATLAB-based ML toolbox developed at the Lab of

Measurement Technology (LMT) is used [8]. These models are employed to evaluate the dataset and confirm that early wear detection is indeed feasible with this approach.

The testbed comprises 16 screw holes, each manufactured through different processes, labeled A to D, to introduce signal variability to make the ML models more robust. Specifically, A and B are hand-cut and hand-formed threads, while C and D are machine-cut and machine-formed [9]. We excluded the four central holes because the shock absorber interferes with the screwing process (Fig. 3). The entire testbed slides along pins, which, combined with the shock absorber, help minimize counterforces during the tightening and loosening cycles [9].

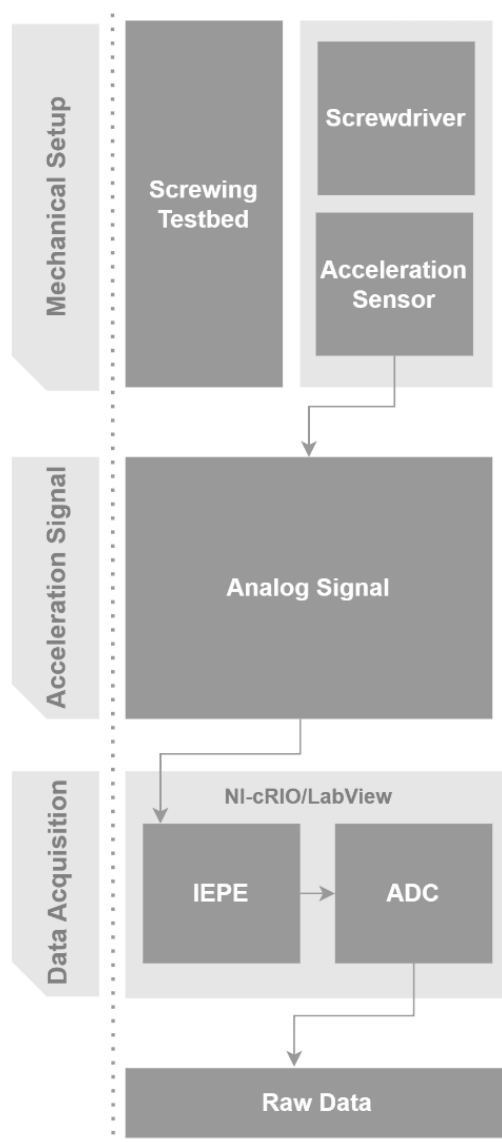


Fig. 1: Schematic representation of the data acquisition pipeline, illustrating the flow from analog signal generation at the sensor level to the digitized raw data formatted for signal processing.



Fig. 2: Handheld electric screwdriver with a piezoelectric accelerometer non-destructively mounted near the bit holder to capture process-related vibrations.

To execute the screwdriving operations, an electric angle- and torque-driven screwdriver (Rexroth, type DAH72736) was employed (Fig. 2). The device was operated in manual mode.

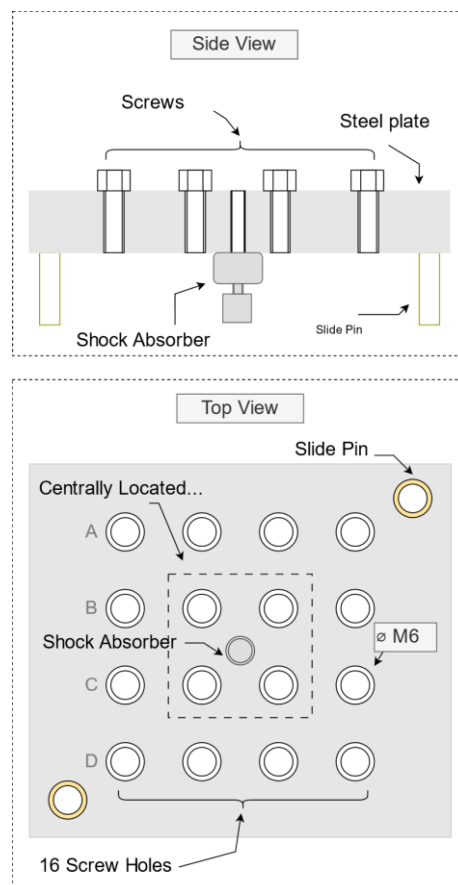


Fig. 3: Screwing testbed utilized for dataset generation. The four centrally located holes are excluded from the analysis due to signal interference from the shock absorber.

Vibrations are captured using a three-axis piezoelectric accelerometer (Dytran 3233A), as shown in Fig. 2. The sensor provides a sensitivity of 1000 mV/g and a frequency response range from 0.4 Hz to 6 kHz. For this study, only the z-axis data is analyzed, as the x- and y-axes are susceptible to coupling artifacts that could compromise signal integrity [10].

The screw bits investigated in this study feature a hexagonal socket geometry with dimensions of 25 mm in length and 5 mm in width, designed to interface with the M6 threaded holes of the testbed (Fig. 3). Three separate bits were evaluated. Initially, the baseline mass of each bit was recorded. Wear was then artificially induced through a controlled grinding process. Each wear stage corresponds to a uniform material reduction of approximately 10 mg (Table 1).

The generated acceleration signals were recorded using a data acquisition (DAQ) system equipped with an NI 9232 module (National Instruments) and an Integrated Electronics Piezo-Electric (IEPE) interface. This interface facilitates the low-noise transfer of high-dynamic signals by providing a constant current source to power the sensor [11]. Signal digitization is performed by an integrated analog-to-digital converter, which quantizes the signal at 24-bit resolution. The sampling rate was set to 20 kHz to provide comprehensive coverage of the high-frequency components present in the vibration data.

A cRIO-9040 control unit (National Instruments) collects the digitized data and stores it in the Technical Data Management Streaming (TDMS) format, a LabVIEW-native format optimized for large, continuous data streams. To ensure high transparency and reproducibility, each measurement is accompanied by detailed metadata specifying the testbed configuration, wear stage, and damage type (Fig. 4).

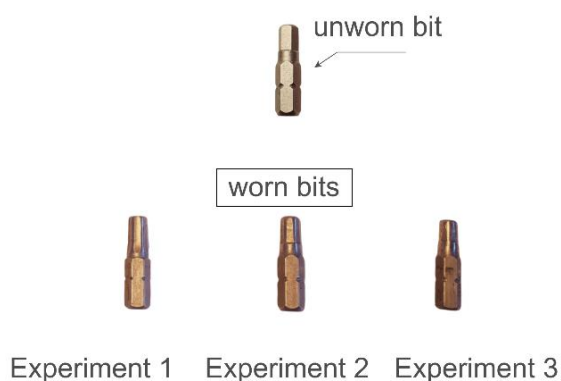


Fig. 4: Comparison between the unworn and worn bits

Table 3: Bits' weight and the corresponding wear stage.

Wear Stage	Bit 1 Weight (mg)	Bit 2 Weight (mg)	Bit 3 Weight (mg)
0	5568	5646	5728
1	5562	5635	5713
2	5556	5627	5699
3	5550	5610	5682
4	5544	5600	5667
5	5538	5580	5652
6	5527	-	-
7	5515	-	-

The presented dataset comprises time series data from three independent experiments conducted with three separate bits:

- Experiment 1 includes 8 wear stages (W_S).
- Experiment 2 and 3 each consist of 6 wear stages. The third experiment utilizes randomized screw hole sequences to enhance model robustness.

For each wear stage, two complete runs (R) are performed. Each run includes 24 individual screwing processes (S_P), divided into 12 tightening and 12 loosening operations across 12 screw holes (S_H), resulting in a final dataset of 960 time series:

$$12 S_H \times 2 S_P \times 2 R \times [8 + 6 + 6] W_S = 960$$

All time series data, along with the corresponding experimental profiles and labeling information, are provided in MATLAB .mat format.

Machine Learning Methods

This section details the methodologies applied to the acquired vibration data and outlines the general processing pipeline, including an exemplary workflow for wear detection (Fig. 5).

The raw acceleration data undergo automated segmentation into 2.5-second intervals to isolate individual tightening and loosening operations. Given the sampling frequency of 20 kHz, each interval comprises 50,000 samples. To enable real-time model responsiveness, these intervals are further subdivided into ten segments.

Due to the high-frequency nature of the acceleration signals, time-frequency analysis is employed for feature extraction. This approach ensures the interpretability of the features and the subsequent model [12]. The machine learning architecture incorporates a time-frequency feature extraction (TFEx) stage, illustrating the dataset's utility for interpretable machine learning [13].

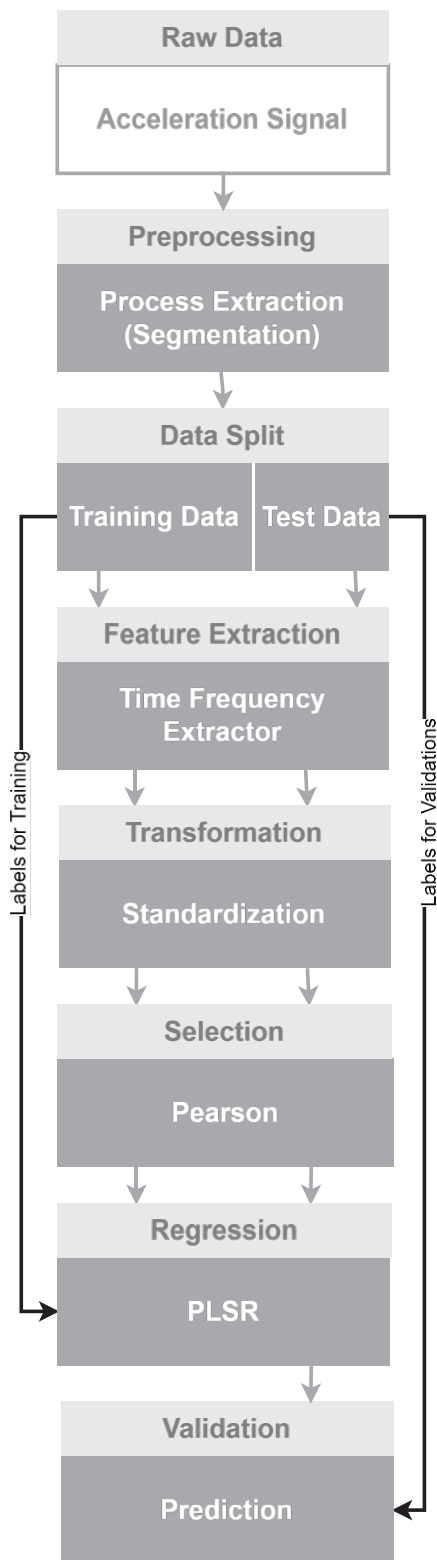


Fig. 5: Schematic of the ML architecture, illustrating the data processing pipeline from raw acceleration signals and segmentation to frequency-domain feature extraction and PLSR-based wear prediction.

Within the TFEx module, the previously defined intervals are partitioned into 20 sub-intervals,

each of which is transformed into the frequency domain via the Fast Fourier Transform (FFT) [14]. Subsequently, each FFT-transformed signal is divided into 40 equal bins for final analysis [15].

Finally, statistical descriptors are calculated to construct the feature vector. These include kurtosis, linear slope, maximum and minimum values, the position of the maximum, root-mean-square (RMS), peak-to-RMS ratio, skewness, and variance.

To visually evaluate the quality of the extracted features, a Principal Component Analysis (PCA) was performed [16]. The resulting scores demonstrate clear cluster separability between the healthy state (indicated by blue markers) and the wear-affected states during the tightening process (Fig. 6, Fig. 7, and Fig. 8) [17]. Furthermore, a Pearson correlation-based selector was used to identify the most representative features [18]. To mitigate the risk of overfitting, this selector was constrained to the five most significant features.

Subsequently, a Partial Least Squares Regression (PLSR) model was trained against the wear labels to perform wear prediction, as is suitable for continuous prediction ranges as well as for noisy data [19]. The dataset architecture supports a robust validation strategy to mitigate systematic error. This strategy employs Leave-One-Group-Out Cross-Validation (LOGO-CV), where groups are defined by wear stages or screw hole types [20]. This workflow is referred to as the Feature Extraction, Selection, and Regression (FESR) stack [15].

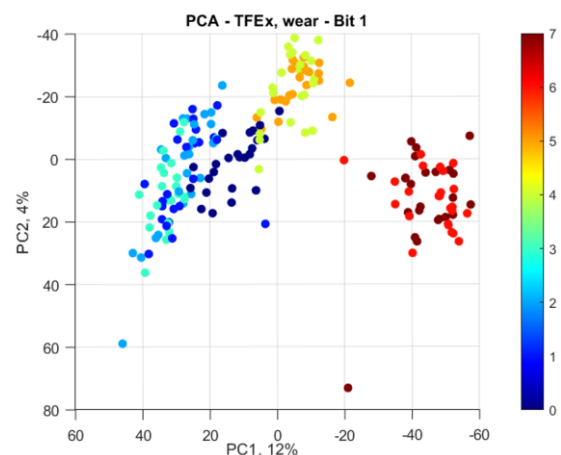


Fig. 6: PCA score plot of the time-frequency features (TFEx) extracted during the tightening process for Experiment 1. The projection onto the first two principal components (PC1 and PC2) illustrates the clear cluster separability according to the progressive wear stages of the bit.

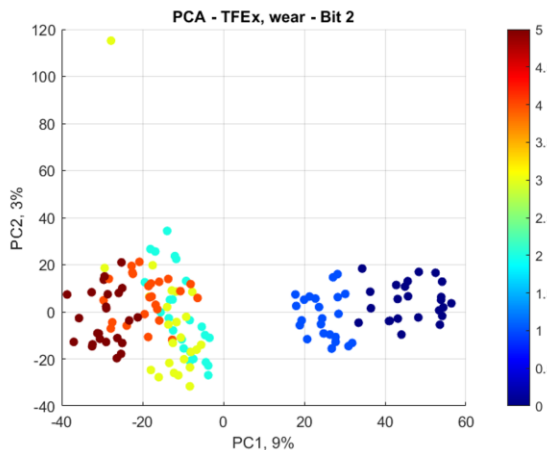


Fig. 7: PCA for Experiment 2. As for Experiment 1, the first two principal components capture the variance associated with bit degradation.

Results

Applying the processing pipeline (Fig. 5) yields a prediction error of 1.25 for the first experiment, 0.98 for the second, and 1.19 for the third. Since the model employs regression for prediction, the first and last wear stages of each experiment were excluded from the validation set, preventing the model from extrapolating beyond the defined training interval [21].

The validation results for the first experiment (stages 1 to 6) and the second experiment (stages 1 to 4) are presented in Figs. 8 and 9, respectively. These results demonstrate the feasibility of real-time wear stage monitoring. Consequently, a maintenance alert for bit replacement can be triggered as soon as wear stage 4 is detected.

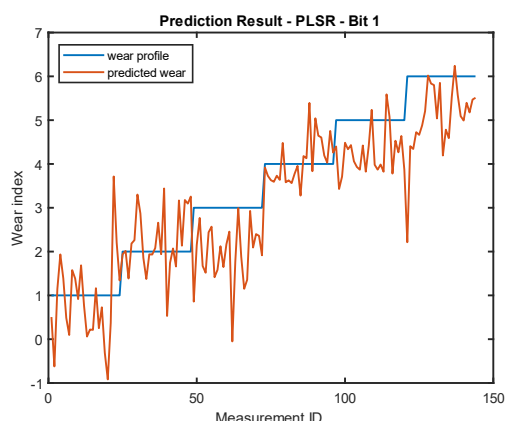


Fig. 8: Validation results for Experiment 1 using the PLSR model. The plot compares the predicted wear stages against the ground truth for stages 1 through 6.

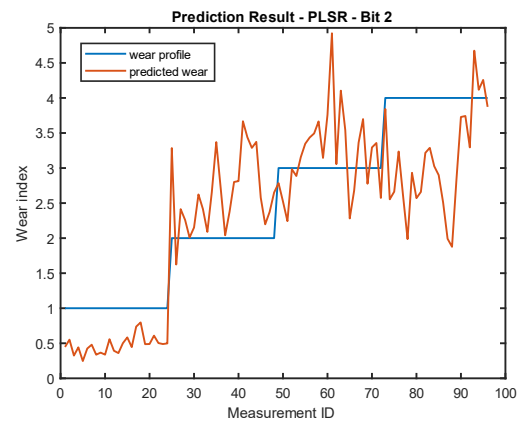


Fig. 9: Validation results for Experiment 2 using the PLSR-based regression. The plot compares the predicted wear stages against the ground truth for stages 1 through 4.

Conclusion

This paper presents a dataset for the condition monitoring of screwdriver bits utilized with a handheld electric screwdriver. Due to the high-frequency nature of the vibration signals generated during the assembly process, data acquisition was performed using a piezoelectric accelerometer. Furthermore, a machine learning model based on an FESR stack was developed and validated to demonstrate the feasibility of detecting tool wear. For the first experiment, the model achieved an average prediction error of 1.25 stages, representing the deviation from the ground truth. In the second experiment, the model reached a prediction error of 0.98 and remained consistent at 1.19 for the third.

While these results are promising, future work should focus on enhancing and validating model robustness through additional experimental trials. Furthermore, the model's transferability across different bit geometries and screwdriver types remains an area for further exploration.

Acknowledgment

This paper was written in the context of the VProSaar project which is funded by the Ministry of Economic Affairs, Innovation, Digital and Energy (MWIDE) and the European Fonds for Regional Development (EFRE). The project is located in the federal state of Saarland and is executed by ZeMA – Centre for Mechatronics and Automation Technology gGmbH.

References

- [1] E. Ungvári, I. G. Gyurika, and T. Csiszér, "Evaluation of the failure effects of a screwing station using a new approached FMEA," *MethodsX*, vol. 11, p. 102278, Dec. 2023, doi: 10.1016/j.mex.2023.102278.
- [2] H. N. Teixeira, I. Lopes, A. C. Braga, P. Delgado, and C. Martins, "Screwing process analysis using multivariate statistical process control," *Procedia Manufacturing*, vol. 38, pp. 932–939, 2019, doi: 10.1016/j.promfg.2020.01.176.
- [3] K. Althoefer, B. Lara, and L. D. Seneviratne, "Monitoring of self-tapping screw fastenings using artificial neural networks," *J. Manuf. Sci. Eng.*, vol. 127, no. 1, pp. 236–243, Feb. 2005, doi: 10.1115/1.1831286.
- [4] A. Al Assadi, D. Holtz, F. Nägele, C. Nitsche, W. Kraus, and M. F. Huber, "Machine learning based screw drive state detection for unfastening screw connections," *J. Manuf. Syst.*, vol. 65, pp. 19–32, Oct. 2022, doi: 10.1016/j.jmsy.2022.07.013.
- [5] R. M. Aronson *et al.*, "Data-driven classification of screwdriving operations," in *Experimental Robotics (Springer Proceedings in Advanced Robotics)*, vol. 1, D. Kulić, Y. Nakamura, O. Khatib, and G. Venture, Eds. Cham, Switzerland: Springer, 2017, pp. 244–253, doi: 10.1007/978-3-319-50115-4_22.
- [6] H. El Moutaouakil, J. Schauer, and A. Schütze, "An interpretable machine learning approach for condition monitoring of screwdriver bits based on time–frequency features and vibration analysis," in *Proc. 2026 IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Glasgow, Scotland, May 2026, to be published.
- [7] V. R. Miranda and J. L. Jr., "Comparison of the signal characteristics measured by a MEMS and a piezoelectric accelerometers," *IJAERS*, vol. 5, no. 11, pp. 148–152, 2018, doi: 10.22161/ijaers.5.11.21.
- [8] T. Dorst, Y. Robin, T. Schneider, and A. Schütze, "Automated ML toolbox for cyclic sensor data," presented at the Joint Virtual Workshop of ENBIS and MATHMET on Mathematical and Statistical Methods for Metrology (MSMM), 2021. [Online]. Available: <https://www.mathmet.org>
- [9] C. Schnur, "Methodisches Vorgehen zur Realisierung von maschinellen Lernprojekten im Mittelstand," Ph.D. dissertation, Universität des Saarlandes, Saarbrücken, Germany, 2024, doi: 10.22028/D291-45080.
- [10] M. Gaitan, I. M. López Bautista, and J. Geist, "Reduction of calibration uncertainty due to mounting of three-axis accelerometers using the intrinsic properties model," *Metrologia*, vol. 58, no. 3, p. 035006, Jun. 2021, doi: 10.1088/1681-7575/abeccf.
- [11] C. Jonscher, B. Hofmeister, T. Gießmann, and R. Rolfes, "Very low frequency IEPE accelerometer calibration and application to a wind energy structure," *Wind Energ. Sci.*, vol. 7, no. 3, pp. 1053–1067, May 2022, doi: 10.5194/wes-7-1053-2022.
- [12] C. Fuchs, S. Klein, J. Schauer, A. Schütze, and T. Schneider, "Eine Methode zur erklärbaren Merkmalsextraktion aus dem Zeit- und Frequenzbereich für Condition Monitoring," in *Proc. Sensoren und Messsysteme 2024*, Wunstorf, Germany, 2024, pp. 556–562, doi: 10.5162/sensoren2024/P36.
- [13] S. Klein, Y. Wilhelm, A. Schütze, and T. Schneider, "Combination of generic novelty detection and supervised classification pipelines for industrial condition monitoring," *tm - Technisches Messen*, vol. 91, no. 9, pp. 454–465, Sep. 2024, doi: 10.1515/teme-2024-0016.
- [14] H. J. Nussbaumer, *Fast Fourier Transform and Convolution Algorithms*, 2nd ed. Berlin, Germany: Springer-Verlag, 1982, doi: 10.1007/978-3-642-81897-4.
- [15] P. Goodarzi, S. Klein, A. Schütze, and T. Schneider, "Comparing different feature extraction methods in condition monitoring applications," in *Proc. 2023 IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Kuala Lumpur, Malaysia, May 2023, pp. 1–6, doi: 10.1109/I2MTC53148.2023.10176106.
- [16] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, no. 1–3, pp. 37–52, Aug. 1987, doi: 10.1016/0169-7439(87)80084-9.
- [17] L. Ferré, "Selection of components in principal component analysis: A comparison of methods," *Comput. Stat. Data Anal.*, vol. 19, no. 6, pp. 669–682, Jun. 1995, doi: 10.1016/0167-9473(94)00020-J.
- [18] M. A. Hall, "Correlation-based feature selection for machine learning," Ph.D. dissertation, Dept. Comput. Sci., Univ. Waikato, Hamilton, New Zealand, 1999.
- [19] S. Wold, M. Sjöström, and L. Eriksson, "PLS-regression: a basic tool of chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 58, no. 2, pp. 109–130, Oct. 2001, doi: 10.1016/S0169-7439(01)00155-1.
- [20] P. Goodarzi, A. Schütze, and T. Schneider, "Domain shifts in industrial condition monitoring: a comparative analysis of automated machine learning models," *J. Sens. Sens. Syst.*, vol. 14, no. 2, pp. 119–132, Jul. 2025, doi: 10.5194/jsss-14-119-2025.
- [21] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*, 5th ed. Hoboken, NJ: Wiley, 2012.

Analysis of the Intrinsic Noise Spectrum of MEMS Microphones

Emanuel Wallesch¹, Gabriel Rodriguez Gutierrez¹, Stefan Palzer¹

¹*Technical University Dortmund, Professorship for Sensors, Otto-Hahn-Str. 8, 44227 Dortmund*

Abstract

MEMS microphones are silicon-based devices widely used in consumer electronics due to their small size, low power consumption, and low cost. Their intrinsic noise spectrum depends on environmental parameters such as pressure, temperature, and gas composition. In this study, the noise spectrum of a MEMS microphone was measured from 5–95 kHz in mixtures of nitrogen (N_2), carbon dioxide (CO_2), and methane (CH_4). Principal component analysis (PCA) shows that CO_2 and CH_4 concentrations can be distinguished, indicating the potential for low-cost environmental sensing using noise analysis.

Keywords: Noise analysis, gas sensors, MEMS

Introduction

Micro-electro-mechanical system (MEMS) microphones have evolved into a standard component for consumer electronics products including e.g. mobile phones, smart watches, and tablets with several billion units produced every year. A central goal of manufacturers is to improve signal to noise ratio and sensitivity in the audible range up to 20 kHz [1]. However, the intrinsic noise spectrum of MEMS microphones shows complex dependencies on environmental parameters including temperature and gas composition [2]. This offers means to make use of MEMS microphones as environmental sensors via spectral analysis of its intrinsic noise spectrum, in particular in the non-audible frequency range. While the prominent resonant feature near 20 kHz has been modelled and studied in the past [2], the analysis of the complete spectrum is bound to yield a more reliable classification of the gas composition.

To this end, this contribution presents results on intrinsic noise analysis of a MEMS microphone in a gas matrix composed of three components, i.e. nitrogen (N_2), carbon dioxide (CO_2), and methane (CH_4) by means of a principal component analysis (PCA). The approach could enable low-cost sensing solutions for applications requiring distributed monitoring, for example the detection of methane and carbon dioxide emissions from landfills [3].

Experimental

The experimental setup is shown in Figure 1. To determine the intrinsic noise spectrum the signal of a IDK InvenSense Inc. ICS-40720 microphone that is powered by a supply voltage of $V_{cc} = 3.3$ V is amplified by a factor of 10,000 and then fed to a FSV3013 spectrum analyzer (Rohde-Schwarz GmbH&Co. KG). It determines the signal in a range from 5 kHz – 95 kHz and uses a total of 1001 linearly spaced frequency values. The microphone is embedded into an apparatus to simulate real-world conditions in the laboratory [4] that is able to control the gas composition and temperature.

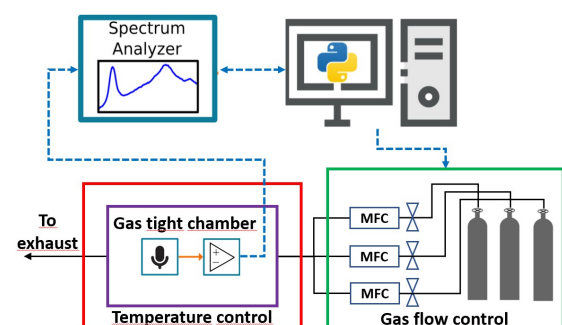


Figure 1: The experimental setup. The microphone and the amplifier sit inside a temperature controlled gas tight chamber. Software controls the Mass Flow Controllers to set a specific atmosphere inside the measurement chamber. The spectral data is sampled by a spectrum analyzer. [2]

Figure 2 shows three exemplary spectra recorded under different conditions.

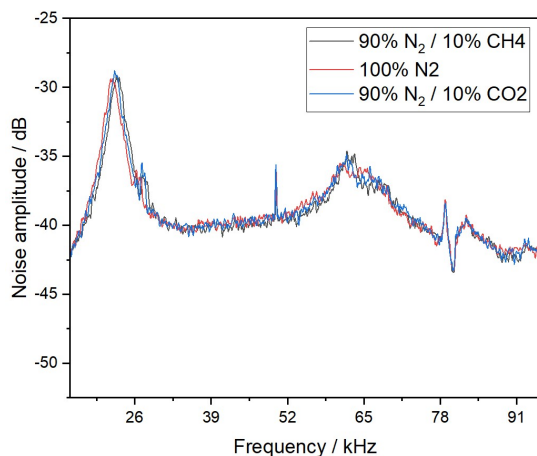


Figure 2: Intrinsic noise spectrum in the frequency range from 15 kHz – 95 kHz under different gas compositions at 23°C and 1 bar pressure.

To analyze the spectra a principal component analysis (PCA) has been employed using 889 frequency values as input features. To analyze each gas composition 140 spectra of each gas composition step were taken to not include spectra where the target gas composition was not reached already. Each of those spectra was then normalized and mean-centered. The PCA was then performed in MATLAB on the spectra between 15 kHz and 95 kHz.

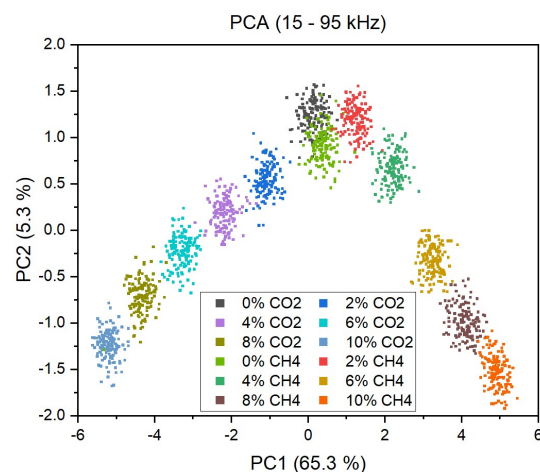


Figure 3: The results of the PCA show a separation of CO₂ and CH₄.

The 0% CO₂ and 0% CH₄ step can be seen almost overlapping in the top center of the figure. From this point it can be seen that with rising concentrations the points move away from the center. CO₂ moves to the third quadrant while CH₄ is moving into the fourth quadrant with rising concentrations.

Conclusion

Gas sensing using intrinsic noise analysis of MEMS microphones, which are mass-produced, inexpensive, low-power, and already integrated into many electronic devices, is feasible. The concept may be used to deploy large-area networks of sensors across landfill sites. Such distributed systems could provide continuous, spatially resolved monitoring of methane and carbon dioxide emissions at significantly lower cost than conventional gas analyzers. Therefore, the results of support the concept that intrinsic noise analysis of MEMS microphones could form the basis of scalable, low-cost landfill gas monitoring systems, capable of detecting methane and carbon dioxide variations in nitrogen-dominated environments.

Future work could extend this approach to measure environmental parameters such as humidity, temperature values using only the intrinsic noise spectrum of MEMS microphones. Such developments could enable compact, low-cost and non-invasive monitoring solutions for atmospheric or industrial applications.

References

- [1] P. R. Scheeper, W. Olthuis, and P. Bergveld, "Improvement of the performance of microphones with a silicon nitride diaphragm and backplate," *Sensors and Actuators A: Physical*, vol. 40, no. 3, pp. 179–186, 1994, doi: 10.1016/0924-4247(94)87003-9.
- [2] G. Rodriguez Gutierrez, E. Wallesch, A. Ortiz Perez, and S. Palzer, "Temperature and chemical sensing via intrinsic noise analysis of MEMS microphones," *Sensors and Actuators B: Chemical*, vol. 450, p. 139199, 2026, doi: 10.1016/j.snb.2025.139199.
- [3] D. H. Cusworth, R. M. Duren, A. K. Ayasse, R. Jiorle, K. Howell, A. Aubrey, R. O. Green, M. L. Eastwood, J. W. Chapman, A. K. Thorpe, J. Heckler, G. P. Asner, M. L. Smith, E. Thoma, M. J. Krause, D. Heins, and S. Thorneoloe, "Quantifying methane emissions from United States landfills," *Science*, vol. 383, no. 6690, pp. 1499–1504, Mar. 2024, doi: 10.1126/science.adi7735.
- [4] J. Kneer, A. Eberhardt, P. Walden, A. Ortiz Pérez, J. Wöllenstein, and S. Palzer, "Apparatus to characterize gas sensor response under real-world conditions in the lab," *The Review of scientific instruments*, vol. 85, no. 5, p. 55006, 2014, doi: 10.1063/1.4878717.

Energieautarkes LoRaWAN-Sensornetzwerk und integrierte IoT-Datenmanagementplattform für Smart-City- und Umweltmonitoring – Pilotprojekt „IoT-Lab“ der TU Bergakademie Freiberg

Pal Arki¹, Yvonne Joseph¹, Thomas Schumann², Jörn Kunze³, Knut Voigtländer³

*¹Institut für Nanoskalige und Biobasierte Materialien, TU Bergakademie Freiberg,
Gustav-Zeuner Str. 3, 09599, Freiberg, Deutschland*

*²Institut für Informatik, TU Bergakademie Freiberg,
Bernhard-von-Cotta-Straße 2, 09599 Freiberg, Deutschland*

*³Advanced Data Processing Dresden GmbH,
Breitscheidstraße, 78/Haus C1, 01237 Dresden, Deutschland*

Zusammenfassung

Die vollflächige Integration von Multisensoren für Smart-City- und Umweltmonitoring-Anwendungen setzt eine autarke, kabellose Netzwerkinfrastruktur voraus. Aufgrund seiner hohen Reichweite und seines geringen Energieverbrauchs eignet sich das funkbasierte Long Range Wide Area Network (LoRaWAN) besonders für die Übertragung von Multisensordaten. Als Low-Power-Wireless-Technologie bietet LoRaWAN gegenüber dem LTE-Mobilfunkstandard sowie lokalen WLAN-Technologien zahlreiche Vorteile und eine höhere Flexibilität. Der in Europa lizenzfrei nutzbare Frequenzbereich von 868 MHz ermöglicht zudem einen kostengünstigen Betrieb. Beim Einsatz von Hunderten Internet-of-Things- (IoT-) Geräten fallen die Kosten im Vergleich zu kommerziellen, mobilfunkbasierten Sensornetzwerken deutlich geringer aus. Auf Basis dieser Infrastruktur wurde ein cloudbasiertes IoT-Datenmanagementsystem entwickelt, das Messwerte von Sensoren in den Bereichen Smart City und Smart Campus, einschließlich Energie-, Gebäude- und Umweltmonitoring, speichert und auswertet. Es umfasst eine skalierbare Datenbank, einen MQTT-Server, eine Dashboard-Anwendung sowie automatisierte Berichtsfunktionen. Die Plattform „IoT-Lab“ der TU Bergakademie Freiberg ermöglicht die Integration eigenentwickelter und kommerzieller Sensor-Nodes sowie die Speicherung, mobile Anzeige, historische Abfrage und Analyse von Sensordaten. Das System stellt aktuelle Messwerte über eine Applikation bereit. Gespeicherte Daten können als Verlaufsdigramme visualisiert werden; zudem werden täglich automatisierte E-Mail-Berichte generiert. Für weiterführende Analysen sind direkte Datenbankabfragen möglich. Mithilfe der Hauptkomponentenanalyse (Principal Component Analysis, PCA) auf Basis der Singulärwertzerlegung (Singular Value Decomposition, SVD) können Zusammenhänge, Trends und Anomalien erkannt werden. Ein kontinuierliches Monitoring von Empfangssignalstärke (RSSI), Signal-to-Noise Ratio (SNR) und Batteriestatus unterstützt außerdem die frühzeitige Erkennung von Ausfällen und Kommunikationsproblemen.

Keywords: LoRaWAN-Sensornetzwerk, Smart- City und Umweltmonitoring, IoT-Datenmanagement, energieautarke Sensorstationen, Principal Component Analysis

Einleitung

Die intelligente Vernetzung von Multisensoren bildet eine wesentliche Grundlage für moderne Städte und Universitäten der Zukunft. Für die Realisierung digitaler Smart-City- und Umweltmonitoring-Lösungen werden energie- und kosteneffiziente sowie robuste und langlebige Sensoren und Aktoren benötigt. Diese müssen über zuverlässige drahtlose Verbindungen mit entsprechenden Applikationsservern kommunizieren, um

Messdaten, Zustände oder Sollwerte auszutauschen.

Eine effiziente Möglichkeit, Geräte über große Distanzen bei geringem Energieverbrauch drahtlos miteinander zu vernetzen, bietet die LoRa-Funktechnologie. Darauf aufbauend wurde das Netzwerkprotokoll LoRaWAN (Long Range Wide Area Network) entwickelt, das die Datenübertragung von den Sensoren über Gateways bis hin zum Applikationsserver organisiert. Sowohl kommerzielle als auch frei

zugängliche LoRaWAN-Netze sind verfügbar, weisen jedoch regional eine stark unterschiedliche Abdeckung und Ausprägung auf.

Insbesondere in Klein- und Mittelstädten befindet sich der Ausbau von LoRaWAN-Netzwerken auf regionaler, nationaler und globaler Ebene weiterhin im Aufbau. In der Universitätsstadt Freiberg (Sachsen), zwischen Dresden und Chemnitz gelegen, war LoRaWAN bis vor Kurzem noch nicht verfügbar.

Ziel war es, in Freiberg ein energieautarkes LoRaWAN-Sensornetzwerk sowie eine integrierte IoT-Datenmanagementplattform im Rahmen des Pilotprojekts „IoT-Lab“ zu etablieren. Das Projekt bezieht sowohl den Campus als auch das Stadtgebiet ein. Universität, Stadt, Unternehmen sowie Bürgerinnen und Bürger sollen dadurch die freie IoT-Plattform gemeinsam nutzen können. Im Mittelpunkt stehen der Aufbau einer LoRaWAN-Infrastruktur sowie die einfache und offene Implementierung geeigneter Sensorik. Darüber hinaus sollen nachhaltige IoT-Lösungen in den Bereichen Energie, Mobilität, Umweltmonitoring und Wohnungswirtschaft getestet, evaluiert und perspektivisch skaliert werden. Eine Applikation zur Anzeige aktueller Sensormesswerte, zur Visualisierung historischer Daten sowie zur täglichen Generierung von Reports per E-Mail wird erstellt.

LoRaWAN-Technologie

LoRaWAN ist ein Funkstandard für drahtlose Netzwerke, der kleine Datenmengen über große Distanzen bei sehr geringem Energieverbrauch überträgt und vor allem im Bereich des Internets der Dinge (IoT) Anwendung findet. Die Technologie ermöglicht Reichweiten von etwa 2 bis 20 Kilometern und eignet sich ideal für batteriebetriebene Sensoren und Aktoren, deren Batterien oft bis zu zehn Jahre halten können. Geräte und Gateways sind kostengünstig, der Standard ist offen, und pro Gateway können sehr viele Geräte betrieben werden. Weitere Vorteile eines LoRaWAN-Netzwerks sind die gute Durchdringung von Gebäuden und Kellern, die hohe Sicherheit durch Ende-zu-Ende-Verschlüsselung sowie die geringe Strahlenbelastung.

Ausbau der LoRaWAN-Infrastruktur

Mit Gateways und leistungsfähigen Antennen wurde in Freiberg eine solide LoRaWAN-Infrastruktur aufgebaut. Mithilfe eines autark über Solarpanels betriebenen LoRaWAN-

Gateways (Abb. 1) konnte bereits eine gute Funkabdeckung in Freiberg und der Umgebung erzielt werden (Abb. 2). Gleichzeitig wurden bestehende „Funklöcher“, insbesondere im Innenstadtbereich, identifiziert.



Abb. 1: Autarkes LoRaWAN-Gateway (Milesight SG50) mit leistungsstarker 9-dBi-Stubantenne (INBM – TUBAF Freiberg)

Das Gateway Milesight SG50 [1] ist auf einen vollständig autarken Betrieb ausgelegt und kann ohne lokale Netzwerk- und Energieinfrastruktur betrieben werden. Es kombiniert ein kompaktes LoRaWAN-Gateway mit LTE-Anbindung, Akku und Solarpanel und ermöglicht damit die Erweiterung des LoRaWAN-Netzes auch an abgelegene oder infrastrukturell schwer erschließbare Orte. Der integrierte Akku erlaubt zudem einen mehrtägigen Betrieb auch bei fehlender Sonneneinstrahlung.

Das Gateway wurde über einen Zeitraum von sechs Monaten unter realen Außenbedingungen betrieben. Auf diese Weise konnten sowohl die Eignung für den autarken Dauerbetrieb als auch die Robustheit des Systems im praktischen Einsatz nachgewiesen werden. Lediglich im Januar kam es bei länger anhaltender fehlender Sonneneinstrahlung zu Ausfällen von ein bis zwei Tagen. Die Sensoren senden sonst zuverlässig Daten aus dem Stadtgebiet sowie aus den Randbereichen an das Gateway.

Dabei können auch Entfernungen von mehreren Kilometern überbrückt werden (Abb.3). In der dicht bebauten Innenstadt zeigen sich jedoch vereinzelt Versorgungslücken.

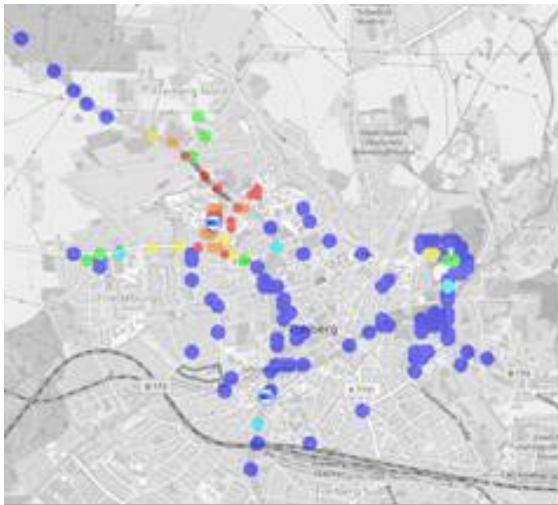


Abb. 2: Karte (Heatmap) mit Gateways in Freiberg und Umgebung sowie der der Empfangsstärke

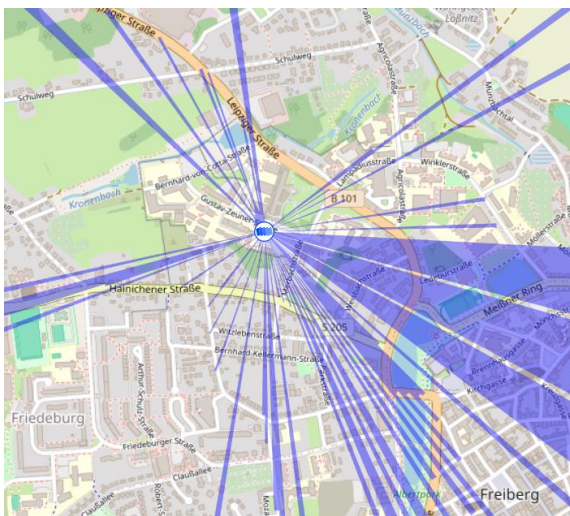


Abb. 3: Visualisierung des Datenaustauschs zwischen Sensoren und Gateway

Durch die Installation weiterer Gateways an ausgewählten, exponierten Standorten oder markanten Gebäuden, beispielsweise auf dem ca. 40 m hohen Bibliotheksturm, kann die Funkabdeckung und Redundanz weiter verbessert werden. Perspektivisch wird dadurch eine nahezu lückenlose Integration von Sensoren in Freiberg und Umgebung ermöglicht.

Integrierter Multisensor-Lösungen

Eine integrierte Multisensor-Infrastruktur zur Übertragung, Speicherung, Visualisierung und Analyse von Sensordaten wurde entwickelt (Abb.4). Als technische Grundlage wurde MQTT als IoT-Messagesystem ausgewählt. Dafür wurden eine Topic-Struktur zur Datenzuordnung und Sensorsteuerung sowie eine JSON-Dokumentstruktur für Sensordaten definiert.

Die Lösung wurde in einer Cloud-Umgebung auf einem IONOS-Server umgesetzt. IONOS ist eine Europäische Gesellschaft (SE) und Anbieter für Webhosting- und Cloud-Lösungen. Anschließend wurden ein MQTT-Server mit Mosquitto sowie eine Datenbank mit MS SQL Server konfiguriert und in Betrieb genommen. Darüber hinaus wurde eine Applikation erstellt, die aktuelle Sensormesswerte anzeigt, historische Daten visualisiert und täglich Reports per E-Mail (Abb.5) generiert. Zur Validierung der implementierten Lösung wurde die Integration verschiedener Sensoren und Sensorsysteme demonstriert.

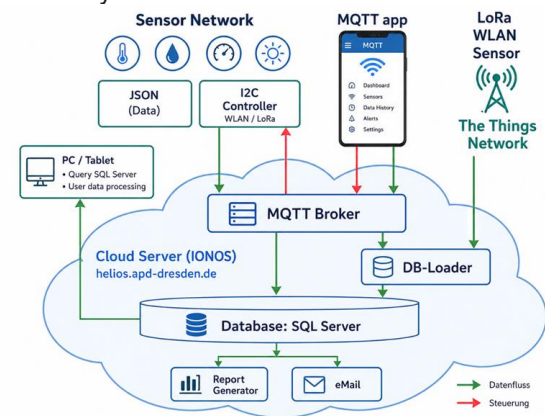


Abb. 4: Integrierten Multisensor-Infrastruktur

Das Kernstück der entwickelten Lösung ist einerseits eine Datenbank, in der alle Dokumente der Sensoren zur Datenerhebung und Sensorsteuerung persistiert werden und andererseits das MQTT Message System, welches die gezielte Verteilung und Zustellung der Dokumente mittels eines „publish-subscribe“ Verfahrens gewährleistet. Jede Komponente (Sensoren, Smartphon App, PC-Analyse Applikation, Überwachungs- und Alarmsystem, DatenLogger, etc) kann sich am MQTT-Broker anmelden und sich für die jeweiligen Datenströme und Control-Kommandos unter den entsprechenden Topics selektiv registrieren. Der DB-Loader registriert sich („subscribe“) für alle Topics und stellt so sicher, dass alle über das Netzwerk verbreiteten Datenpakete historisiert und für einen Abruf zur Analyse und Visualisierung historischer Daten verfügbar sind.

Um das entwickelte System hinsichtlich seiner Leistungsfähigkeit unter vielfältigen Nutzungsbedingungen zu entwickeln und zu testen wurden in großem Umfang und in großer Variabilität eigenentwickelte und kommerzielle Sensoren integriert. Bis zum April 2026 sind in der Datenbank 160 Mio. Datensätze mit einem Gesamtspeicherplatz von 64GB eingelaufen. Für effiziente Abfragen großer Datenmengen wurden Indizes auf Kombinationen aus Zeitstempel und Topic-Struktur eingerichtet. Insgesamt entstand damit eine skalierbare Plattform, über die unterschiedliche Sensoren und Sensorsysteme integriert, gesteuert, historisiert, mobil angezeigt und gezielt ausgewertet werden können.

Sensoren

Kommerzielle Sensoren wurden sowohl in den Laboren des Instituts für Nanoskalige und Biobasierte Materialien als auch auf dem Campus der TUBAF installiert. Sie erfassen unter anderem Temperatur, Luftfeuchtigkeit, CO₂-Konzentration, Lichtintensität und Luftdruck.

Darüber hinaus wurden selbst entwickelte Multisensoren als autarke Sensorboxen für den mobilen Einsatz konzipiert. Diese dienen als flexible Messstationen zur Erfassung verschiedener Umweltdaten. Neben Temperatur und Luftfeuchtigkeit wurden auch Halbleiter-Gassensor-Arrays sowie GPS-Sensoren erfolgreich integriert (Abb. 5).



Abb. 5: Autarkes mobiles Multisensorbox

Erfassung und Darstellung der Messwerte

Die Sensordaten wurden über den MQTT-Server von The Things Network für LoRaWAN erfasst. Pro Tag werden 72 Dokumente generiert, da alle 20 Minuten ein Datenpaket erzeugt und archiviert wird. Die Sensoren senden seit mehreren Monaten stabil Messdaten an die Sensorplattform. Von dort aus werden täglich Reports mit Messwertdiagrammen erstellt, archiviert und per E-Mail versendet (Abb.6-7).

Abb.6 zeigt beispielhaft einen automatisch generierten Tagesreport, der per E-Mail versendet wird und den Verlauf der Datenaufzeichnung über die jeweils letzten sieben Tage darstellt. Neben den Messwerten des Drucksensors, der seit April 2024 bereits fast drei Millionen Datensätze geliefert hat, sind auch die Messwerte von zwei unterschiedlichen Temperatursensoren dargestellt. Alle erfassten Sensordaten können in beliebiger Kombination und für verschiedene Zeiträume frei als Report konfiguriert werden.

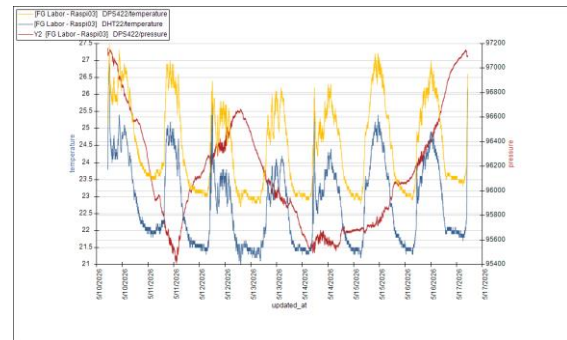


Abb. 6: Täglich fortlaufender Email-Report-Chart der Temperatur- und Luftdruckwerte

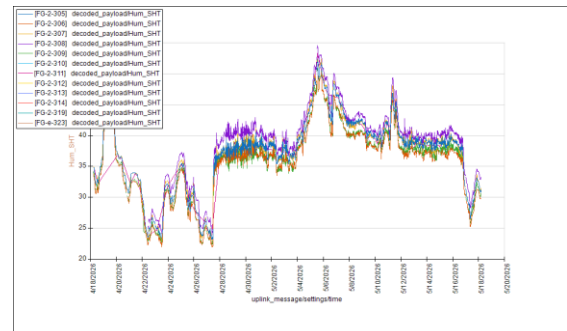


Abb. 7: Täglich fortlaufender Email-Report-Chart der Luftfeuchtigkeit in Laboren

Diese Darstellungen verdeutlichen die Funktionsweise des entwickelten Gesamtsystems. Der Sensor wird unter Nutzung der empfohlenen Technologien und Spezifikationen über LoRaWAN integriert. Ab diesem Zeitpunkt sind alle Datensätze online auf mobilen Geräten verfügbar, werden in der Cloud gespeichert, als Verlaufsdiagramme dargestellt und können in zyklisch, beispielsweise täglich oder wöchentlich, automatisch generierte E-Mail-Reports aufgenommen werden. Darüber hinaus stehen sie für individuelle Analysen und Weiterverarbeitungen per Datenbankabfrage zur Verfügung.

Auswertung der Messwerte

Aus den JSON-Dokumenten werden stündliche Mittelwerte für Temperatur, Luftfeuchtigkeit, RSSI (Received Signal Strength Indicator) und SNR (Signal-to-Noise Ratio) extrahiert.

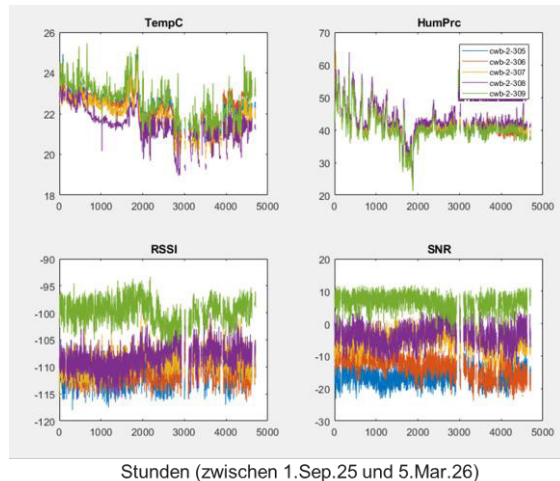


Abb. 8: stündliche Mittelwerte für Temperatur, Luftfeuchtigkeit, RSSI (Received Signal Strength Indicator) und SNR (Signal-to-Noise Ratio) zwischen 1.Sept 25 und 5.März 26

Die Abb.9 zeigt, wie die erfassten Sensordaten für die weitere Analyse in einer vierdimensionalen Matrix, also einem Tensor, strukturiert werden.

Datenrepräsentation als 4-D-Matrix (Tensor)

- 4-D-Tensor(5,24,7,28) für Temp, Luftfeuchtigkeit, rssi, snr
- enthält teilweise fehlende Werte (2.309 / 23.520) ~ 10%



Abb. 9: Datenrepräsentation als 4-D-Matrix (Tensor) mit den Dimensionen Sensor, Tagesstunde, Wochentag und Kalenderwoche.

Für fünf Sensoren, 24 Stunden, sieben Wochentage und 28 Kalenderwochen ergibt sich eine strukturierte Datenbasis mit 23.520 möglichen Datenpunkten pro Messgröße. Diese Struktur wird für Temperatur, Luftfeuchtigkeit, RSSI und SNR verwendet. Aufgrund fehlender oder nicht übertragener Datenpakete enthält der Datensatz teilweise Lücken; im Beispiel fehlen 2.309 Werte, was etwa 10 % entspricht.

Die Varianzkomponentenanalyse des Tensors mit vier Dimensionen zeigt, dass Umweltgrößen wie Temperatur und Luftfeuchtigkeit hauptsächlich durch langfristige wöchentliche Veränderungen geprägt sind, während Kommunikationsgrößen wie RSSI und SNR vor allem durch sensor- bzw. standortabhängige Unterschiede bestimmt werden.

Tab. 1: Wichtigste Einflussdimensionen der einzelnen Messgrößen aus der Varianzkomponentenanalyse

Messgröße	wichtigste Dimension	Interpretation
TempC	Kalenderwo.	saisonalen Verlauf / langfristige Temperaturänderung
HumPrc	Kalenderwo	saisonale oder wetterbedingte Feuchteänderung
RSSI	Sensor	standortabhängige Signalstärke
SNR	Sensor	starke Unterschiede der Signalqualität zwischen Sensoren

Die Methode Principal Component Analysis (PCA) mittels Singular Value Decomposition (SVD) wurde verwendet, um die Messgrößen des Sensorgitters in ihre wichtigsten Muster zu zerlegen (Abb 10 und 11).

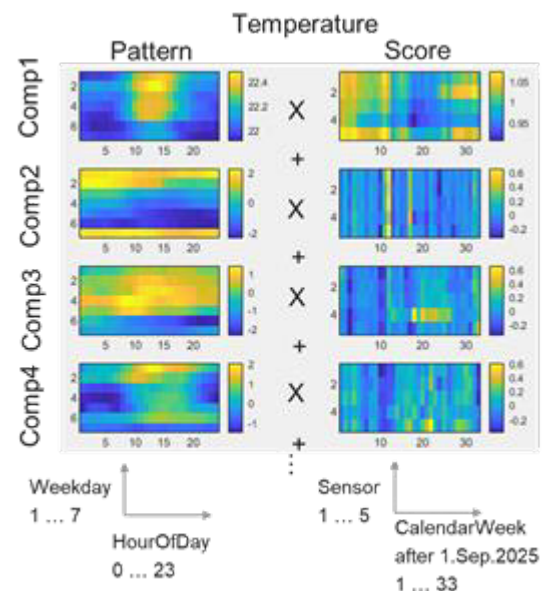


Abb. 10: PCA/SVD-Zerlegung der Temperaturdaten eines Sensorgitters bestehend aus 5 Sensoren.

Die Messdaten hängen hier von Stunde des Tages: 0–23 ; Wochentag: 1–7 ;Sensor: 1–5 Kalenderwoche: 1–33 nach dem 1. September 2025 ab.

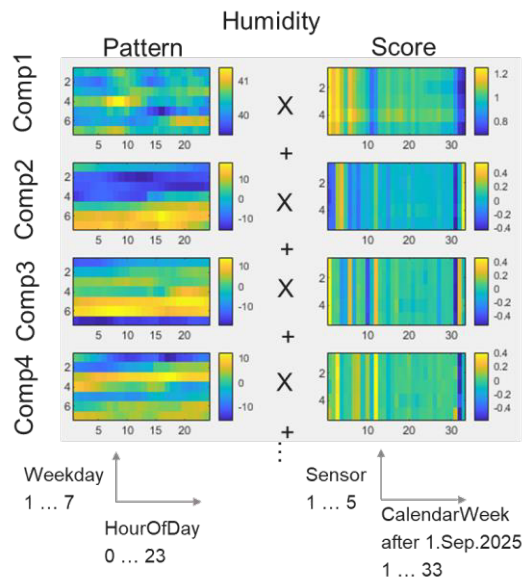


Abb. 11: PCA/SVD-Zerlegung der Luftfeuchtigkeitsdaten eines Sensorgitters.

Die PCA versucht, diese komplexen Daten in wenige einfache Bausteine zu zerlegen. Für jede Hauptkomponente (Comp1–Comp4) wird ein zeitliches Muster („Pattern“) über die Stunde des Tages und den Wochentag mit einem räumlich-zeitlichen Gewicht („Score“) über Sensoren und Kalenderwochen kombiniert. Bei den Temperaturdaten zeigt sich ein klarer Tagesrhythmus. Insgesamt wird die Temperaturvariabilität stark durch den Tag-Nacht-Rhythmus, bestimmt.

Bei der Luftfeuchtigkeit (Abb.11) zeigt sich ein anderes Bild. Die erste Hauptkomponente beschreibt ein relativ stabiles Grundniveau der Feuchte, jedoch ohne stark ausgeprägten Tagesgang. Auffällig ist, dass die Luftfeuchtigkeit stärker von langfristigen Wetterlagen, saisonalen Effekten und episodischen Ereignissen beeinflusst wird als von der Tageszeit. Im Gegensatz zur Temperatur ist der diurnale Zyklus hier deutlich schwächer ausgeprägt.

Die Kommunikationsparameter RSSI (Empfangssignalstärke in Dezibel) (Abb. 12 und SNR (Signal-Rausch-Verhältnis) (Abb. 13.) werden hauptsächlich durch sensorstandortspezifische Unterschiede bestimmt. Dies weist darauf hin, dass Faktoren wie Sensorposition, Distanz zum Gateway, Abschattung durch Wände oder lokale Interferenzen die Verbindungsqualität dominieren, während zeitliche Effekte vergleichsweise schwach ausgeprägt sind.

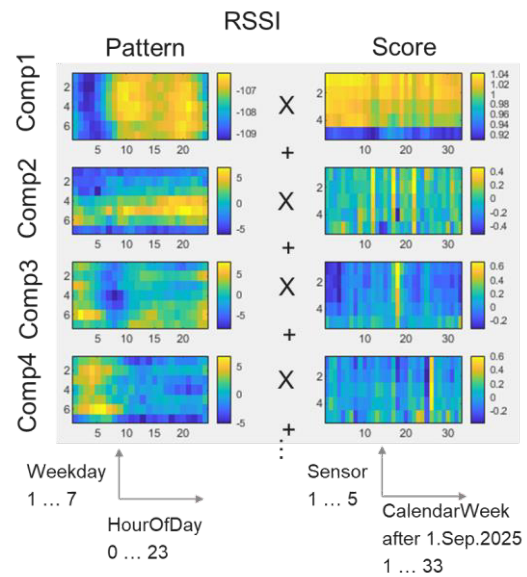


Abb. 12: PCA/SVD-Zerlegung der Empfangssignalstärke (RSSI)

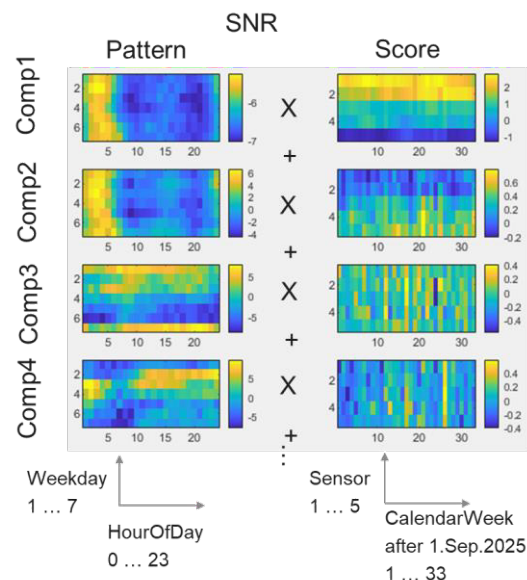


Abb. 13: PCA/SVD-Zerlegung des Signal-Rausch-Verhältnisses

Die Methode (PCA) mittels (SVD) ermöglicht es, dominante Muster zu identifizieren, Unterschiede zwischen Sensoren sichtbar zu machen und saisonale/wetterbedingte und räumliche/sensortechnische Veränderungen kompakt darzustellen. Dadurch können komplexe Sensordaten effizient interpretiert und charakteristische Umweltprozesse erkannt werden.

Literaturnachweis

- [1] A. Möcker, Milesight SG50: Solarbetriebenes LoRaWAN-Gateway im Test, c't 21/2024 S. 94

Ultra-Low-Power Wireless Temperature Measurement for High-Temperature Rotating Equipment

Leander B. Hörmann¹, Bernhard Uhl¹, Greta Minelgaite², Jon Barrenetxea-Arando², Noelia Montes², Veronika Putz¹

¹ Linz Center of Mechatronics GmbH, Altenberger Straße 69, 4040 Linz, Austria

² Fundación Inatec (OTUA), San Blas 35, 01170 Legutio, Spain

Abstract

This paper presents a wireless, energy-autonomous temperature measurement system for a rotary aluminium melting furnace. The system is designed for temperature acquisition at locations where wired sensors are difficult to install due to rotation, thermal load, and limited accessibility. It consists of photovoltaic-powered wireless sensor nodes, BLE-based communication, and a Linux-based edge gateway for data processing and integration into downstream systems.

The sensor nodes use thermocouple measurement electronics and supercapacitor-based energy storage. Stainless-steel fixtures and thermal shielding enable mounting on the rotating furnace while limiting the thermal load on the electronics. Industrial measurements demonstrate long-term temperature acquisition at the furnace core, outer surface, and exhaust gas path. The results demonstrate the feasibility of cable-free temperature monitoring on rotating high-temperature equipment and identify energy availability as a key factor for reliable long-term operation. The acquired temperature data provides a basis for process monitoring, forecasting, and digital twin modelling of the melting process.

Keywords: Wireless Temperature Measurement, Energy-Autonomous Sensor Nodes, Rotating Machinery Monitoring, Energy Harvesting, Ultra-Low-Power Industrial IoT

Introduction

Temperature is a key process variable in high-temperature industrial applications. Accurate knowledge of the current process temperature is required to monitor and control the state of the system. This is particularly relevant for metal melting processes, where the temperature directly determines the physical state and quality of the molten material. Precise temperature information enables the identification of optimal process conditions, ensuring material quality while simultaneously minimizing excessive heating. As a result, reliable temperature measurement contributes not only to product quality but also to improved energy efficiency by reducing unnecessary heating time and associated energy consumption.

In many industrial applications, metal melting is performed in rotary furnaces. The continuous rotation enables a homogeneous distribution of temperature and material composition, leading to improved process stability and uniformity of the molten material. However, this rotation introduces significant challenges for temperature measurement. Sensors must operate under high thermal loads, mechanical motion, and limited accessibility. Conventional

wired measurement systems are often unsuitable due to mechanical wear, cable fatigue, and the difficulty of routing and maintaining connections on rotating components. Wireless sensor nodes provide a viable alternative for such demanding measurement scenarios. They enable the acquisition of physical quantities at locations where wired solutions are impractical or infeasible [1], [2]. When combined with energy harvesting, these systems can operate without any wired infrastructure for power or communication, allowing fully autonomous deployment [3], [4]. However, the available harvested energy in industrial environments is typically limited and often discontinuous [5], [6]. This imposes strict constraints on system design, requiring ultra-low-power hardware architectures, efficient sensing strategies, and energy-aware communication protocols and software implementation to ensure reliable operation [7], [8].

This paper presents the application of wireless, energy-autonomous sensor nodes for temperature measurement in an aluminium melting process using a rotary furnace. The measured temperature is used to estimate the optimal scrap melting state and time-to-tap approximation, enabling improved process

control and reduced energy consumption. The work describes the overall system architecture of the wireless sensor network, the mechanical integration of the sensor nodes on the rotating furnace including thermal shielding concepts, and the energy harvesting approach. Furthermore, first experimental results are presented, including the performance of the energy harvesting system under realistic operating conditions and temperature measurements at different locations on the rotary furnace and within its exhaust gas stream.

Application Scenario and Requirements

The wireless temperature sensor solution was implemented at the premises of a prominent European metal recycling company, dedicated to secondary aluminium smelting, and located in Spain. The refinery's main activities are focused on two production lines, which include high-capacity smelting rotary furnaces (30 t), holding furnaces, and ingot production lines. The company is dedicated to recycling a wide range of waste containing aluminium, predominantly of post-consumption origins. Prior to industrial smelting process, the received aluminium scrap is characterized, and its quality, as well as the obtained yield are evaluated utilizing refinery's pilot scale rotary furnace of approximately 10 kg capacity which is shown in Figure 1. This scrap characterization, quality and yield evaluation step is crucial to estimate an optimal scrap charging mix for the industrial furnaces. Refinery's pilot rotary furnace is retrofitted with a number of different sensors, such as energy meter, gas totalizer, door sensors, etc. (EU project REVaMP, H2020 No 869882) for process control and optimization purposes. Nevertheless, the molten aluminium scrap state evaluation and time-to-tap decisions at current are based on furnace operator's experience – expert knowledge, as no temperature sensors for continuous measurements were previously installed. Thus, the pilot scale rotary furnace was selected as an asset for wireless temperature measurement solution application.

Rotating furnace temperature monitoring is challenging, as sensors must operate under 1) high thermal loads – typical rotary furnace operating temperatures range from 650 °C to 800 °C [9]-[11], 2) mechanical motion, such as inclination, rotation, and 3) limited accessibility – any cable installation may obstruct the free movement of the furnace. During refinery's pilot rotary furnace operation, the target molten aluminium tapping temperature is around 650 °C, while external backside surface



Figure 1: Application scenario showing the rotary furnace with 10 kg capacity for scrap characterization.

temperatures in a range of 82 °C – 95 °C can be reached (measured by an infrared laser thermometer; HIBOK-77 (DT-8016H/-50 °C + 1600 °C), Ref. 430, Útiles y Máquinas Industriales, S.A.). Regarding the exhaust gas temperature, some measures are periodically carried out in the corresponding filtering systems. However, they are located in a distance from direct exhaust gas flow, and therefore, the reference values may be somewhat lower, in a range of 69 °C – 80 °C (refinery's internal reports data).

For such demanding temperature measurement scenarios, wireless sensor nodes can serve as a viable alternative, which was proposed and applied for the selected asset (the rotary furnace) of the present study.

System Architecture of the Wireless Temperature Measurement System

The wireless temperature measurement system is composed of three main components: distributed wireless sensor nodes, a transceiver unit, and a gateway for data aggregation and processing. The system is designed to capture temperature information at multiple critical locations within the rotary furnace process.

Temperature measurements are performed at three distinct positions: on the outer surface of the rotary furnace, inside the fireclay insulation of the rotary furnace, and at the exhaust gas collection funnel. Accordingly, three wireless sensor nodes are deployed. Two nodes are mounted directly on the rotating furnace, while a third node is installed on the stationary support structure of the exhaust gas funnel,

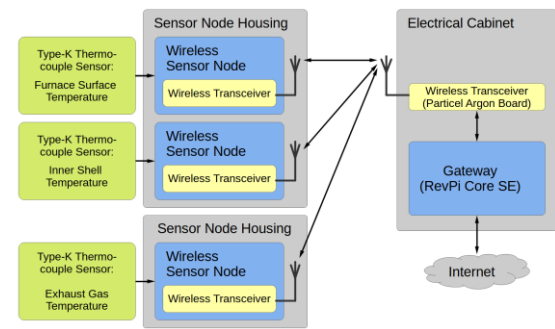


Figure 2: Block diagram of the system architecture showing the three wireless sensor nodes, the gateway with its transceiver.

located several meters away from the furnace. Each sensor node is based on a Nordic nRF52840 microcontroller, which integrates both processing capabilities and a wireless transceiver, enabling compact and energy-efficient operation. The nodes communicate wirelessly with a dedicated transceiver unit implemented using a Particle Argon platform. This transceiver serves as an interface between the sensor network and the higher-level system.

The gateway is realized using an industrial-grade Revolution Pi (RevPi), which performs data collection, preprocessing, and forwarding to external systems. This architecture separates low-power sensing and communication tasks from higher-level data handling, ensuring scalability and robustness in industrial environments. A block diagram of the overall system architecture is shown in Figure 2.

Wireless Sensor Node and Analog Circuitry

Each sensor node is designed as an energy-autonomous embedded system integrating sensing, signal conditioning, processing, communication, and power management within a compact architecture. The node is based on a Nordic nRF52840 microcontroller, which combines an ARM Cortex-M processor with an integrated 2.4GHz transceiver and other peripherals, enabling energy-efficient wireless communication and local data processing.

The temperature measurement frontend is designed to support multiple sensor types, including Pt100, Pt1000, and thermocouples. In the presented use case, type-K thermocouples are employed due to their wide temperature range and robustness in harsh environments. The measurement electronics follow a low-power design strategy, where analog components are only activated during measurement phases to minimize energy consumption. The frontend includes a

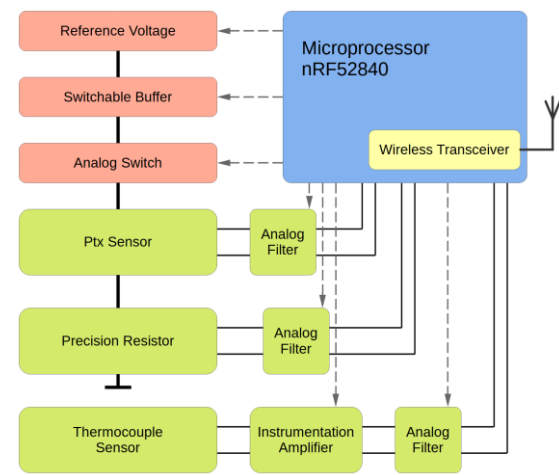


Figure 3: Block diagram of the sensor node's analogue front-end.

precision voltage reference (LTC6652, 2.5 V), signal conditioning, precision resistors, and the internal 12-bit analog-to-digital converter (ADC) of the nRF52840 configured in 16-fold oversampling to reach 14-bit accuracy, enabling low power data acquisition.

For thermocouple measurements, multiple acquisition steps are performed. First, the internal ADC is initialized, and its internal offset calibration is performed. Then, a separate temperature measurement for the cold junction temperature compensation is performed. Therefore, a Pt1000 resistive sensor element is integrated inside the connector of the thermocouple. For each Pt100 or Pt1000 temperature measurement, the reference voltage, the voltage of the reference resistor and the voltage of the Pt100 or Pt1000 resistor are measured. These values are used to calculate the cold junction temperature. After this, the thermocouple voltage is measured, and the corresponding temperature is derived using standardized polynomial approximations. Finally, cold-junction compensation is done using the cold junction temperature measured beforehand. A block diagram of the analogue front-end is shown in Figure 3.

Wireless Communication

Wireless communication is implemented using Bluetooth Low Energy (BLE), which serves both for configuration of the sensor nodes and for periodic transmission of measurement data and status information, including internal supply and harvesting-related parameters. The sensor node firmware is implemented in C/C++ on the Zephyr RTOS and runs on the Nordic nRF52840 platform. It controls the analog measurement circuitry, manages the acquisition of status parameters such as supercapacitor voltage, photovoltaic voltage,

and internal chip temperature, and handles data transmission.

A lightweight application protocol is implemented on top of BLE, supporting both control commands and data exchange. Commands include starting and stopping measurements, reading configuration parameters, and triggering system-level actions such as reboot. Measurement data and status information are transmitted using structured payloads tailored to different sensor types. These payloads are structured into a TLV (Type-Length-Value) scheme and base64-encoded for transport efficiency and compatibility.

The combination of ultra-low-power sensing, duty-cycled operation, and energy-aware communication enables reliable system operation under the strict and fluctuating energy constraints imposed by photovoltaic harvesting in industrial environments.

Gateway

On the gateway side, data acquisition and processing are handled by a Linux-based edge device (RevPi Core SE). The gateway software is implemented in Python and serves as the central interface between the wireless sensor network and the cloud-based data processing systems. It manages BLE scanning and maintains connections to multiple sensor nodes via BLE.

Incoming data frames are received from the sensor nodes, decoded from base64, and parsed according to the implemented TLV-based application protocol. The gateway performs initial data processing by converting raw sensor values into engineering units. This includes the transformation of resistive measurements (e.g., Pt100/Pt1000) into temperature values and the conversion of thermocouple voltages into temperature using standardized characteristics for type K thermocouples.

In addition to data decoding and preprocessing, the gateway provides interfaces for data forwarding and system integration. Measurement data can be transmitted to external platforms via REST-based interfaces for example. Furthermore, time-series data is stored remotely using InfluxDB, enabling continuous monitoring, visualization, and further analysis within the overall data collection architecture.

Energy Harvesting and Power Management

The energy supply is realized through a photovoltaic cell (SM111K09L, IXYS), which converts ambient light into electrical energy. The harvested energy is buffered using a supercapacitor (40F), allowing operation

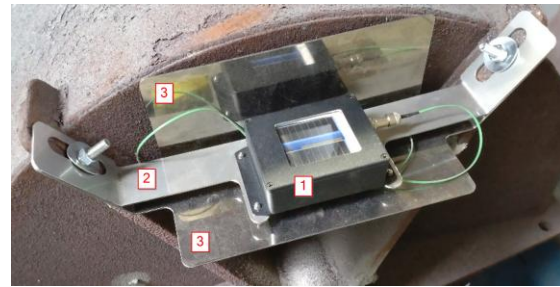


Figure 4: Wireless sensor node housing (1), fixture (2), and heat shield (3) mounted on the rotary furnace.

under intermittently available illumination conditions. A dedicated power management unit, based on a maximum power point tracking (MPPT) concept, ensures efficient energy extraction and regulates the supply voltages for the sensor node electronics. This approach enables fully autonomous operation without any wired power source.

Mechanical Integration and Thermal Protection for the Rotating Furnace

Two wireless sensor nodes were mounted on the backside of the rotary furnace. Both nodes were integrated into a dedicated metal housing featuring a glass window to enable photovoltaic energy harvesting. This housing provides mechanical robustness while allowing sufficient optical access for the integrated solar cells.

The sensor node housings were attached to the furnace using a custom-designed fixture made of stainless steel. Stainless steel was selected due to its relatively low thermal conductivity compared to other metals, reducing conductive heat transfer from the furnace surface to the sensor node. The fixture itself was mounted directly onto the outer surface of the rotating furnace.

To further mitigate thermal exposure, an additional heat shield, also made of stainless steel, was implemented. The heat shield has been bent at a right angle to shield the heat from the two main directions. The fixture and the heat shield are designed as two separate components, creating an air gap between the furnace surface and the shield, as well as between the shield and the sensor node housing. These gaps enable convective cooling and significantly reduce radiative heat transfer from the hot furnace surface to the sensor electronics. The combined effect of conduction reduction, radiative shielding, and convective cooling ensures that the thermal load on the sensor nodes remains within acceptable limits. The complete fixture design is shown in Figure 4.

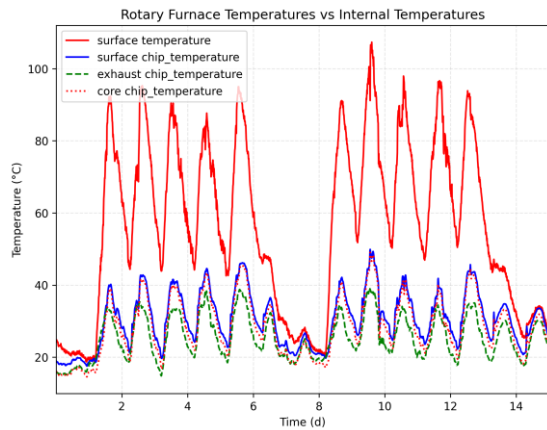


Figure 5: Rotary furnace surface temperature compared to the internal (chip) temperature traces of all three sensor nodes during two of the hottest weeks in the year 2025 (15th to 30th of June).

Figure 5 presents the measured temperature of the furnace surface and the internal temperatures of the three sensor node electronics within the housing. While the furnace surface reaches temperatures of up to 114.4 °C on the hottest day of the measurement period, the internal temperature of the sensor nodes remains limited to 50.25 °C. This reduction demonstrates the effectiveness of the thermal protection concept. The internal temperature is influenced by ambient environmental conditions, as reflected by a third sensor node installed at a distance from the furnace for exhaust gas measurements. This node operates close to ambient temperature and serves as a reference for environmental conditions within the industrial setting. The estimated environmental temperature reached on the same day 39.25 °C. Maintaining a limited internal temperature is critical for reliable system operation. In particular, the employed supercapacitors are specified for a maximum operating temperature of 85 °C. Operating significantly below this limit increases reliability and extends the lifetime of both the energy storage components and the overall sensor node electronics.

Energy Availability and Storage Behaviour

The wireless sensor nodes were originally developed for a different industrial use case. Therefore, it was not known in advance whether the available photovoltaic energy under the illumination conditions of the rotary furnace environment would be sufficient for perpetual operation. Figure 6 shows the supercapacitor voltages (solid) and the solar

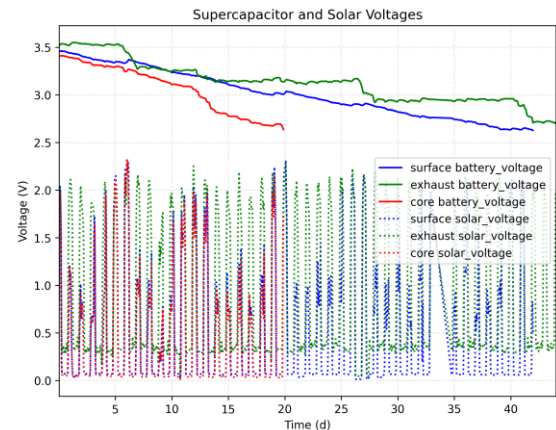


Figure 6: Voltages of the sensor nodes' supercapacitors (solid) and solar cell voltages (dotted) during a six-week period beginning on 21st of October 2024 with insufficient illumination.

cell voltages (dotted) of all three sensor nodes over a six-week period from the end of October to the beginning of December 2024. The supercapacitor voltage traces indicate that the available illumination was insufficient for energy-neutral operation although they receive environmental power every day as can be seen by the solar cell voltage traces. Two of the three sensor nodes discharged progressively and stopped operating after their storage voltage fell below the undervoltage threshold of approximately 2.65 V.

The insufficient energy availability is caused by two main factors. First, the sensor nodes were not originally designed for the low illumination levels present near the rotary furnace, resulting in a negative long-term energy balance. Second, furnace rotation causes the photovoltaic cells to alternate between illuminated and shaded orientations. This strongly time-varying input power can bias the MPPT operating point when open-circuit voltage sampling occurs during low-illumination phases, further reducing harvesting efficiency. As explained in more detail in [12], this effect can reduce the harvested energy by approximately 15 % under comparable periodically varying illumination conditions. Combined with the already limited illumination level in the industrial environment, this additional loss explains the progressive discharge of the supercapacitors and the shutdown of two sensor nodes.

Thus, an artificial light source was installed to illuminate the wireless sensor nodes. During weekend stops, the rotary furnace is positioned so that the sensor nodes face the artificial light source. This allows the nodes mounted on the rotary furnace to recharge during

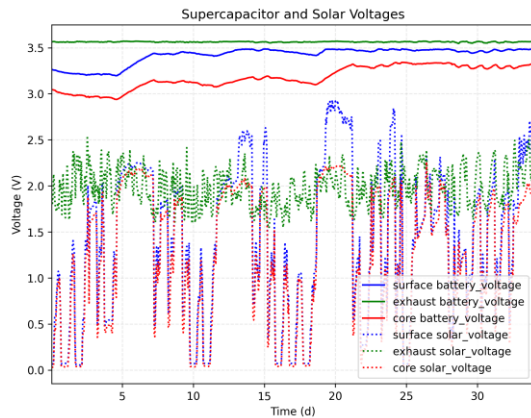


Figure 7: Voltages of the sensor nodes' supercapacitors (solid) and solar cell voltages (dotted) during a five-week period beginning on 26th of January 2026 with artificial illumination during the weekends.

non-operational periods. The stationary sensor node used for exhaust gas temperature measurement is also illuminated by an artificial light source.

Figure 7 shows the supercapacitor voltages (solid lines) and solar cell voltages (dotted lines) of all three sensor nodes over a five-week period starting on 26th of January 2026. The results indicate that artificial illumination during weekend stops is sufficient to enable perpetual operation of the sensor nodes mounted on the rotary furnace. The stationary sensor node receives continuous illumination and therefore remains fully charged throughout the observation period.

However, the harvested power of the furnace-mounted sensor nodes still depends on the exact stopping position of the rotary furnace. If the photovoltaic cells are not optimally aligned

with the artificial light source, the charging performance is reduced.

Temperature Measurement Results

The primary purpose of the wireless sensor nodes is the continuous measurement of process temperatures at three different locations within the rotary furnace system. Figure 8 shows the recorded temperatures over a long-term measurement period starting on 16 Nov. 2025 and ending on 26 Apr. 2026.

The measurements clearly reveal the cyclic operating behaviour of the melting process. During active production phases, the temperature of the rotary furnace repeatedly reaches values above 600 °C, while shutdown periods are characterized by a rapid decrease towards ambient temperature. The weekly production cycle is clearly visible, with regular interruptions during weekends. Furthermore, an extended period of inactivity can be observed around the turn of the year, corresponding to seasonal production holidays. The different measurement locations also show distinct thermal characteristics. The highest temperatures are observed directly at the rotary furnace, while the remaining measurement positions exhibit lower temperatures and slower thermal dynamics. The long-term measurements demonstrate stable operation of the wireless sensor network over several months and confirm the capability of the system to continuously monitor the thermal behaviour of the industrial melting process.

Figure 9 shows the same temperature traces during one representative week of continuous operation. The cyclic behaviour of the melting process can be clearly identified. At the beginning of each production day, the rotary furnace is pre-heated before the aluminium

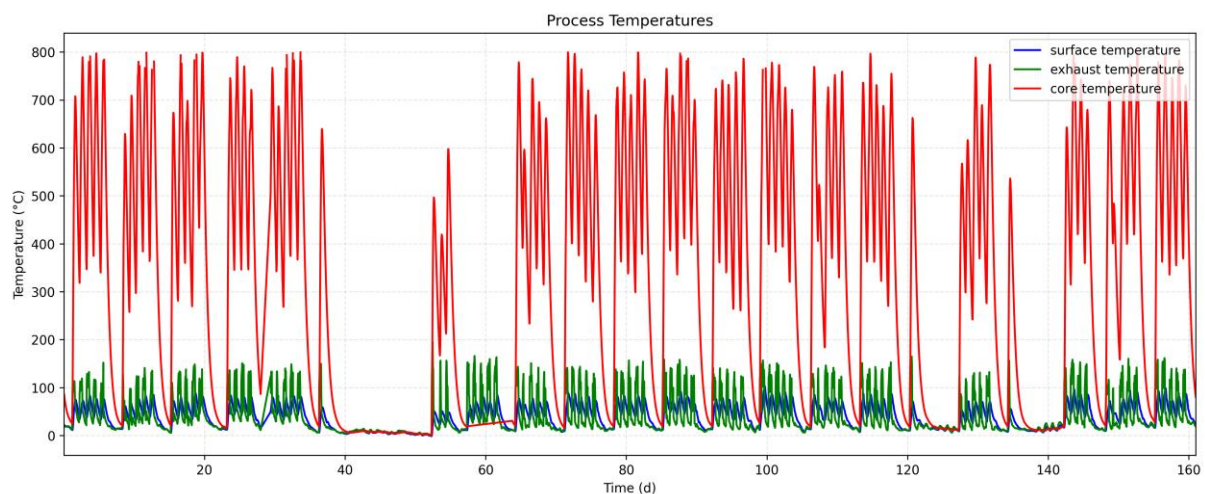


Figure 8: Long-term temperature measurements (furnace core, outer furnace surface, and exhaust gas system) of the rotary furnace process from November 2025 to April 2026 showing weekly production cycles and seasonal shutdown periods.

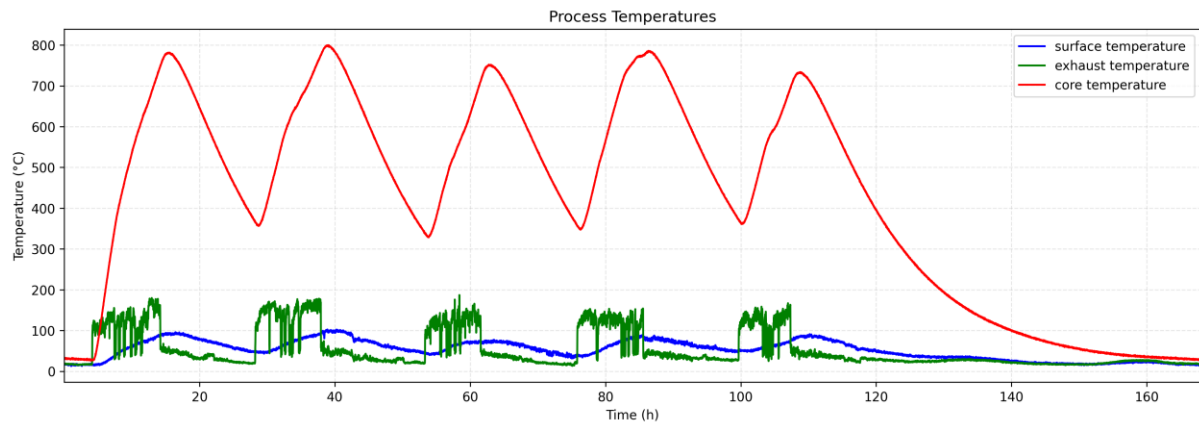


Figure 9: Temperature profiles during one representative production week showing cyclic heating and cooling behaviour of the furnace system.

scrap material is introduced and melted. This heating phase is visible through the rapid increase of the exhaust gas temperature, which reacts first due to its direct exposure to the combustion process and hot gas flow.

The temperature measured inside the furnace core increases with a noticeable delay. This behaviour results from the sensor placement within the fireclay lining rather than direct contact with the molten material itself. Consequently, the measured temperature reflects the thermal inertia of the refractory material. The outer surface temperature of the rotary furnace exhibits the slowest response and lowest absolute temperature due to the thermal insulation provided by the furnace structure.

During operation, the furnace core temperature reaches values above 700 °C, while the outer surface temperature remains below approximately 100 °C. The exhaust gas temperature shows rapid fluctuations caused by changing burner operation and process dynamics. After daily operation ends, all temperatures decrease gradually due to the thermal mass of the furnace system. At the end of the working week, a continuous cooldown towards ambient temperature can be observed over several days, illustrating the long thermal time constants of the rotary furnace and its refractory lining.

Figure 10 shows the temperature profiles during one representative production day in the middle of the working week. In contrast to the beginning of the week, the furnace core temperature is already elevated at the start of the measurement period due to the remaining thermal energy stored in the refractory lining from previous operation cycles.

The detailed temperature traces show that the rotary furnace operates in alternating heating and standby phases rather than being heated

continuously. During active heating, the exhaust gas temperature reacts immediately and fluctuates due to burner operation, while the furnace core and outer surface temperatures respond more slowly because of the thermal inertia of the fireclay lining and furnace structure. During standby phases, no melting is performed and the temperatures stabilize or decrease slowly, indicating that the furnace retains substantial thermal energy for subsequent process cycles.

The presented temperature measurements provide a basis not only for process monitoring, but also for forecasting and digital twin modelling of the rotary furnace. The combination of exhaust gas, furnace core, and outer surface temperatures enables identification of heating, melting, standby, and cooldown phases. Like digital twin approaches for industrial furnace systems, these data can be used to synchronize thermal models with the real process, estimate non-directly measurable thermal states, and predict future

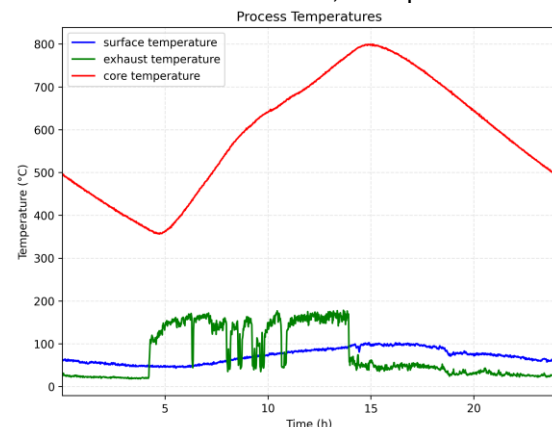


Figure 10: Detailed temperature traces during one production day illustrating alternating heating and standby phases of the rotary furnace.

temperature evolution. This supports model-based process optimization and reduction of energy consumption [13].

Conclusion and Outlook

This paper presented a wireless, energy-autonomous temperature measurement system for a rotary aluminium melting furnace. The results demonstrate reliable cable-free temperature monitoring at multiple furnace locations and confirm that the mechanical fixture and thermal shielding keep the sensor electronics within acceptable temperature limits. Energy availability was identified as the main limitation for long-term autonomous operation, especially under low and dynamically varying illumination caused by furnace rotation. However, additional artificial illumination enabled stable operation over extended periods. The measured temperature data further provides a basis for process monitoring, forecasting, and digital twin modelling. Future work will focus on improved energy harvesting, further power reduction, and the use of the measured temperature data for forecasting and digital twin applications.

Acknowledgements

This work has been supported in part 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, and the Trineflex project. Trineflex (trineflex.eu) has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101058174. Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, the Austrian federal government or the federal state of Upper Austria. Neither the European Union nor the granting authority can be held responsible for them.

References

- [1] G. Bernard, L. Budde, R. Hänggi, and T. Friedli, "Driving continuous improvement with industry 4.0 technologies: Lessons from multiple use case analysis," *Applied Sciences*, vol. 15, no. 4, 2025. Available: <https://www.mdpi.com/2076-3417/15/4/2191>
- [2] D. T. Matt, G. Pedrini, A. Bonfanti, and G. Orzes, "Industrial digitalization. A systematic literature review and research agenda," *European Management Journal*, vol. 41, no. 1, pp. 47–78, 2023. Available: <https://www.sciencedirect.com/science/article/pii/S0263237322000019>
- [3] A. Kansal, D. Potter, and M. B. Srivastava, "Performance aware tasking for environmentally powered sensor networks," *SIGMETRICS Perform. Eval. Rev.*, vol. 32, no. 1, pp. 223–234, jun 2004; doi: 10.1145/1012888.1005714
- [4] L. B. Hörmann, T. Hölzl, C. Kastl, P. Priller, H.-P. Bernhard, P. Peterseil, and A. Springer, "Evaluation of Solar-based Energy Harvesting for Indoor IoT Applications," in *European Test and Telemetry Conference (ettc2022)*, 2022, pp. 190–198; doi: 10.5162/ettc2020/1.2
- [5] L. B. Hörmann, A. Berger, A. Pötsch, P. Priller, and A. Springer, "Estimation of the harvestable power on wireless sensor nodes," in *2015 IEEE International Workshop on Measurements & Networking (M&N)*, 2015, pp. 1–6; doi: 10.1109/IWMN.2015.7322964
- [6] A. Janek, C. Trummer, C. Steger, R. Weiss, J. Preishuber-Pfluegl, and M. Pistauer, "Simulation based verification of energy storage architectures for higher class tags supported by energy harvesting devices," *Microprocessors and Microsystems*, vol. 32, no. 5, pp. 330–339, 2008; doi: 10.1016/j.micpro.2008.03.009
- [7] L. B. Hörmann, P. M. Glatz, C. Steger and R. Weiss, "Designing of efficient energy harvesting systems for autonomous WSNs using a tier model," *2011 18th International Conference on Telecommunications, Ayia Napa, Cyprus*, 2011, pp. 174–179, doi: 10.1109/CTS.2011.5898913.
- [8] L. B. Hörmann, T. Buchegger and C. Steger, "Optimizing the energy supply of autonomous wireless sensor nodes," *2014 Microelectronic Systems Symposium (MESS)*, 2014, pp. 1–6, doi: 10.1109/MESS.2014.7010257.
- [9] B. Zhou, Y. Yang, M.A. Reuter, U.M.J. Boin, "Modelling of aluminium scrap melting in a rotary furnace," *Minerals Engineering*, Volume 19, Issue 3, 2006, 299–308, ISSN 0892-6875; doi: 10.1016/j.mineng.2005.07.017
- [10] O. Majidi, S.G. Shabestari, M.R. Aboutalebi, "Study of fluxing temperature in molten aluminum refining process," *Journal of Materials Processing Technology*, Volume 182, Issues 1–3, 2007, 450–455, ISSN 0924-0136; doi: 10.1016/j.jmatprotec.2006.09.000.
- [11] Dai, J.; Liu, H.; Zhang, Y.; Shi, H.; Ling, P. An Optimization Design of Energy Consumption for Aluminum Smelting Based on a Multi-Objective Artificial Vulture Algorithm. *Metals* 2025, 15, 105.; doi: 10.3390/met15020105
- [12] L.B. Hörmann, L. Springer and V. Putz, "Impact of Periodically Varying Illumination on Photovoltaic-Based Energy Harvesting Performance", In *Proceedings of the 13th International Workshop on Energy Harvesting and Energy-Neutral Sensing Systems*, accepted, to be published.
- [13] M. Mayr and G. C. Chasparis, "Causally-Guided Pairwise Transformer - Towards Foundational Digital Twins in Process Industry", *Procedia Computer Science*, Volume 277, 2026, pp. 175–184; doi: 10.1016/j.procs.2026.02.059

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

Platzhalter

M3-3.3 HiL-Echtzeitsimulation zur Nachbildung und Untersuchung eines energieautarken Sensor-Netzwerks

A Comparative Study on Measurement Accuracy: Wireless Sensor Networks vs. Conventional Systems for the Measurement of Large-Scale Geometric Structures

Lars-Michel Bretthauer¹, Ralf Heynicke¹, Gerd Scholl¹

¹Helmut Schmidt University, Holstenhofweg 85, Hamburg (Germany)

Abstract

Fast and highly precise geometric shape acquisition for in-process control poses a significant challenge for all kind of measurement systems, particularly in the context of large-scale structures. Conventional systems like laser trackers or photogrammetric setups are commonly employed for this purpose. However, these instruments do not provide real-time measurements, and their limitations become increasingly pronounced as the object size grows. Real-time capability, in contrast, offers substantial advantages for closed control loop forming processes in industries such as aerospace and wind energy. This requirement is addressed by our approach employing a dense network of wireless sensor probes that determine the shape of the object through the measurement of surface normal relative probe positions. Within this architecture, all measurement data is aggregated and processed in a central unit at a fixed sampling interval that remains independent of the network size. As a result, the overall measurement time stays constant, even for large-scale structures. In this work, we compare the measurement accuracy and repeatability of this novel wireless sensor network with the conventional approaches referenced above, following the procedures defined in DIN ISO 22514.

Keywords: Sensor Network, Wireless, Uncertainty Prediction

Introduction

The evaluation of new measurement systems is curial to assess the performance and measurement accuracy to existing systems. Therefore, in this work we compare the measurement accuracy and repeatability of a wireless sensor network for measuring shapes of large structures with the conventional approaches, following the procedures defined in DIN ISO 22514 and VDA Vol. 5 [1], [2]. For this purpose, an appropriate experimental setup was developed, as shown in Fig.1. The measurement object is a fiber-reinforced composite panel with rough dimensions of 1.8 m x 2.6 m. The goal is to capture the shape of the measurement object with a sensor network, which consists of 22 equally distributed sensor probes and compare it to laser tracker and photogrammetry systems. In order to obtain a variety of measurements in accordance with the above-mentioned standards, a controlled deformation is introduced using two linear motors, primarily to assess the repeatability of the measurement under varying deformation states. In addition, positional uncertainties of the sensor probes are compared with analytically and simulated uncertainty values, which were obtained in accordance with the “Guide to the expression of uncertainty in measurement” (GUM) framework [3], [4], [5].

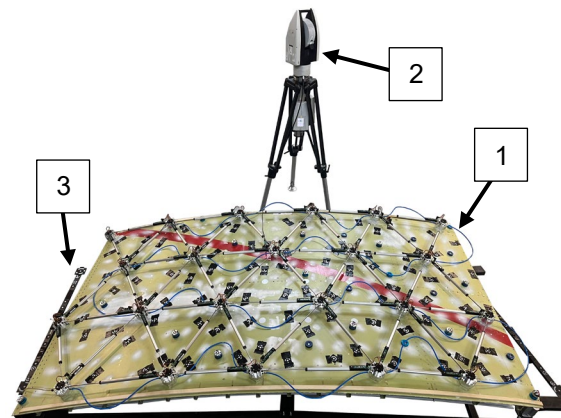


Fig. 1: Comparative measurement setup panel with applied sensor network (1) and the reference systems laser-tracker (2) and photogrammetric system (3).

Conventional Measurement Approaches for Large-Scale Structures

Laser trackers and photogrammetry have established themselves as common measurement systems in industry for measuring large components with outer dimensions greater than 0.5 m, for example in the aerospace and wind power industries [6]. A laser tracker uses a laser interferometer to measure distances with a retroreflective sphere, generating a point cloud (PC) along the surface

of a measurement object, which is often done manually. The resulting offset from the sphere must be compensated along the surface normal of the object. A photogrammetric system reconstructs a 3D point cloud from multiple photographs with different viewing angles on the measurement object. In each photograph, coded markers are detected and then converted into a point cloud using bundle adjustment. In most cases, reference scales are placed around the measurement object in order to achieve accurate results. The Leica AT 960-LR laser tracker is used for comparative measurements with the sensor network as well as the Aicon DPA Pro photogrammetry system. The performance values of the two systems are presented in Tab. 1. In order to combine the measurement data from both systems in one coordinate system, several measuring points are used, which are detected by both systems. This allows the point clouds to be seamlessly aligned with each other.

Tab. 1: Specifications of the conventional measurement systems

	Laser tracker AT 960-LR	Photogrammetry Aicon DPA Pro
Measuring range	80m	30m
Measurement accuracy	$\pm(15\mu\text{m} + 6\mu\text{m}/\text{m})$	$\pm 10\mu\text{m}/\text{m}$

Wireless Sensor Network Measurement System

The sensor network (SN) consists of distributed wireless sensor probes interconnected via mechanical links, forming a scalable architecture. Each probe operates within its own local coordinate system and provides measurements of surface stress and deformation in its immediate vicinity. These measurements are obtained by evaluating data from sensors integrated within the probe, as well as from its directly connected neighboring probes. To determine the surface stress and deformation across the entire component, the locally acquired information must be transformed into a global component coordinate system. This global coordinate system is defined by a set of reference probes with precisely known positions, denoted as S_{Ref}^C , where the superscript indicates the coordinate system in which the position is expressed. In this notation, the superscript "C" refers to the global component coordinate system. The position of a sensor j with respect to the local coordinate system of sensor i is denoted as S_j^i , as illustrated in Fig. 2. This position, expressed in the local coordinate system of probe i , can be transformed into the global coordinate system

as S_j^C by applying the transformation from the local coordinate system of probe i to the global frame, i.e., by mapping S_j^i into the global coordinate system. Consequently, the positions of all sensor probes in the global component coordinate system are obtained through successive transformations of positions defined in local coordinate systems.

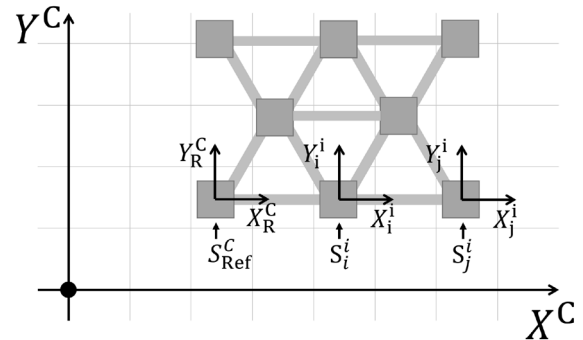


Fig. 2: Definition of the coordinate systems.

To determine the local position S_j^i , each sensor probe estimates the surface normal at its location using a 3D inclinometer. Furthermore, the mechanical links between adjacent probes are instrumented with inclinometers to measure the inclination δ_{i-j} , while the distance k_{i-j} between probes is obtained using linear displacement sensors. All inclinometer measurements are realized using MEMS-based acceleration sensors. The mechanical connections between probes are implemented using spherical ball bearings, enabling rotational flexibility while preserving mechanical stability. By combining the measured inclination angles and inter-probe distances, and considering the underlying geometric constraints, the positions of all probes within the local coordinate system can be reconstructed [5], [7].

Measurement System Analysis (MSA)

A measurement system analysis (MSA) is required to evaluate the suitability of a measurement system for industrial applications. This analysis provides information on the reliability, accuracy, and repeatability of the system with respect to a specified tolerance T . Two closely related approaches are commonly applied: VDA Volume 5 and DIN ISO 22514-7, both of which focus on the evaluation of measurement uncertainty and statistical testing in accordance with the Guide to the Expression of Uncertainty in Measurement [1], [2]. The classical MSA approach further defines the short-term capability of a measurement system using the indices C_g and C_{gk} , where the measurement system is required to utilize no

more than 20% of the tolerance range T_R up with its measurement uncertainty. The C_g factor primarily refers to the variance of the measurement system in relation to the tolerance range, while the C_{gk} value also accounts for systematic deviation in the measurement. These indices can be determined using the tolerance range T_R , the coverage factor k , the sample standard deviation s and sample mean $\hat{\mu}$ from repeated measurements against a normal or reference value with

$$C_g = \frac{0.2 \cdot T_R}{k \cdot s}, \quad (1)$$

$$C_{gk} = \frac{0.1 \cdot T_R - |\hat{\mu} - T_M|}{0.5 \cdot k \cdot s}. \quad (2)$$

Values greater than 1.33 are generally considered acceptable for both indices, values between 1.00 and 1.33 are regarded as conditionally acceptable, and values below 1.00 indicate that the measurement system is not suitable for the intended application. Within the norm approach, a distinction is made between the capability of a measurement system and that of a measurement process. In a first step, the uncertainty of the measurement system must be evaluated with a normal or reference measurement based on

$$u_{ms} = \sqrt{u_{cal}^2 + u_{lin}^2 + u_{bi}^2 + u_{ev}^2}. \quad (3)$$

- u_{cal}^2 = Calibration of the standard or workpiece
- u_{lin}^2 = Deviations from linearity
- u_{bi}^2 = Bias offset
- u_{ev}^2 = Repeatability on standards or workpiece

From this, the uncertainty of the measurement process can be determined based on the evaluated measurement system uncertainty and a repeated measurement in an operating environment with multiple operators with

$$u_{mp} = \sqrt{u_{ms}^2 + u_{gv}^2 + u_{stab}^2 + u_{av}^2}. \quad (4)$$

- u_{ms}^2 = Measurement system
- u_{gv}^2 = Reproducibility of the measuring system
- u_{stab}^2 = Reproducibility over time
- u_{av}^2 = Reproducibility of the operator

Based on the measurement system and process uncertainties, a capability index can then be calculated accordingly

$$C_{ms} = \frac{0.2 \cdot T_R}{k \cdot u_{ms}}, \quad (5)$$

$$C_{mp} = \frac{0.4 \cdot T_R}{k \cdot u_{mp}}. \quad (6)$$

For these indices, the same threshold values as for short-term capability must be achieved in order to be considered suitable. However, the measurement process uncertainty is permitted to utilize up to 40% of the tolerance range, as it accounts for additional statistical influences affecting the measurements. If a measurement system meets these criteria, it can be used reliably in an industrial environment.

Experimental Setup and Evaluation Method

For the measurement setup, the panel is fixed within a steel frame to provide a rigid substructure. The short side of the component is connected to the outer structure by a fixed clamp, while the opposite side is connected to two linear motors, as shown in Fig. 3.

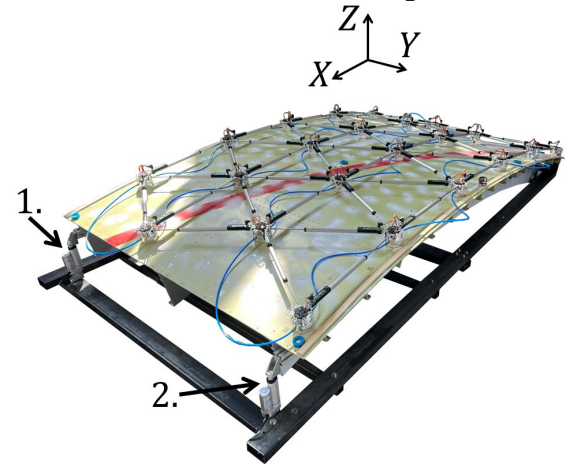


Fig. 3: Measurement setup of the panel with applied sensor network and linear motors on the left (1) and right (2) side.

The motors are designed to deform the component under multidimensional stress, with a maximum extension length of approximately 50 mm. The objective is to demonstrate that the sensor network is capable of detecting this type of deformation. According to ISO 22514-7, multiple measurement objects, operators, and repeated measurements are typically required. Since the present study is conducted on a single panel, the different deformation states are treated as individual measurement objects. These states are generated using the full extension of the linear motors and are summarized in Table 2. The poses are cycled

sequentially from 1 to 3 to obtain repeatable measurements from the same panel, while the fourth pose serves as a control measurement. A reliable assessment of system repeatability requires more than 30 repeated measurements [2]. This criterion is fulfilled in the present study, with a total of 58 measurements conducted. Furthermore, to showcase the systems measurement stability, the measurement takes place over a two-day duration.

Tab. 2: Linear motor positions for poses 1 to 4, retracted = 0, extended = 1

	Left Linear Motor Position	Right Linear Motor Position	Number of Repeated Measurements
Pose 1	0	0	20
Pose 2	1	0	19
Pose 3	0	1	17
Pose 4	1	1	2
	Sum		58

For the comparison of the point clouds, each measurement dataset is transformed into a global coordinate system. In this case, the origin of the global coordinate system is defined by the position of sensor probe #22. To achieve an initial alignment, the point clouds of the reference systems are registered using the previously measured positions of probes #22 and #11, as depicted in Fig. 4. Also, the sensor network is adjusted to the surface of the component using the reference systems to eliminate offsets from the setup of the sensor network.

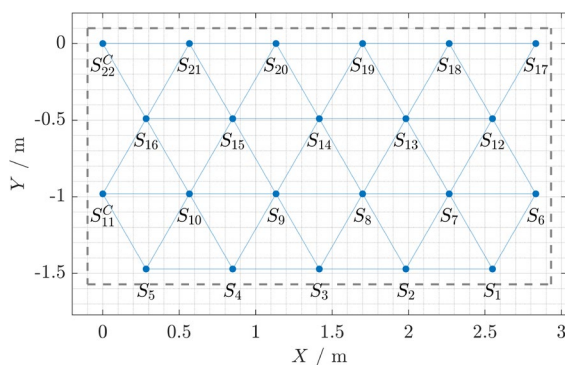


Fig. 4: Network structure of the SN with the reference probes #11 and #22.

During the subsequent fine alignment, a surface fit is generated from the point clouds to provide additional support points and, if necessary, to reduce the influence of outliers. The support points derived from the fitted surface are then used for fine alignment based on the Iterative Closest Point (ICP) algorithm [8]. The overall process for point cloud alignment and evaluation is illustrated in Fig. 5.

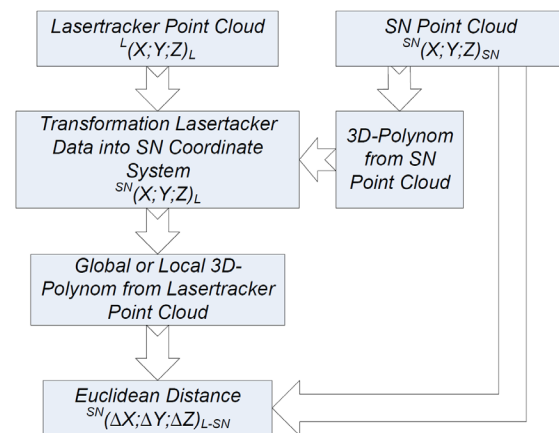


Fig. 5: Alignment procedure of the point clouds and evaluation.

The alignment process enables the calculation of deviations between the sensor network and the aligned reference data. Since the positions of the sensor probes cannot be measured directly with the reference systems, but only points in their vicinity are captured, interpolation between these points is required. Two interpolation approaches are applied. The first approach is based on a global surface fit using all measurement data, taking advantage of the regular geometry of the object. This method is suitable for evaluating the overall shape profile and is comparable to a direct comparison with a CAD model, but it does not account for local deformations. The second approach uses a local surface fit, where only points within a defined radius around each measurement probe are considered. This method allows the assessment of the waviness profile of the object up to a certain wavelength. In both cases, the surfaces are determined using weighted least-squares fitting. The two approaches are exemplarily presented for the laser tracker measurement in pose 1 at sensor probe #1 in Fig. 6.

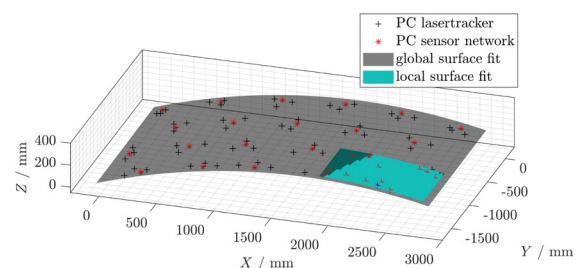


Fig. 6: Comparison of the laser tracker and sensor network point clouds (PC) with the global and local surface fit for probe #1.

To further enhance measurement accuracy, a surface fit can also be applied to the sensor network PC, enabling the extraction of the

underlying surface shape and correct the PC. This method is only applicable on the global evaluation, since it requires from the data to be globally representable with a polynomial expression. For the evaluation of the sensor network, the mean deviation is used to assess systematic errors (bias), while the standard deviation is considered to evaluate repeatability. Furthermore, these results are incorporated into the MSA to assess the capability of the system for deployment in an industrial environment. Since, this measurement setup resembles more a measurement process an expansion of the allowable usage of the tolerance band is 30%. For the tolerance T a value of ± 1.5 mm is used, since this is a commonly used tolerance in the large structure industry.

Uncertainty Estimation of the Sensor Network

As a general statement on the measurement uncertainty of the sensor network is not feasible. Since, it strongly depends on the specific measurement task, its determination in advance is essential to assess whether the achievable uncertainty is suitable for the intended application. An estimation of the measurement uncertainty can be performed in accordance with the Guide to the Expression of Uncertainty in Measurement (GUM) using either analytical or simulation-based approaches. Due to the complex structure of the sensor network, it is advantageous to decompose the system into subsystems and evaluate them individually, as described in [4]. Based on this approach, an overall uncertainty for the positions of the measurement probes can be derived. In the present work, a Monte Carlo method is employed to estimate these uncertainties, using a number of trials of $M = 10^5$ in order to obtain statistically robust results in accordance with the GUM framework [9]. The resulting absolute measurement uncertainty for each sensor is shown in Fig. 7, alongside the correct data from the surface fit.

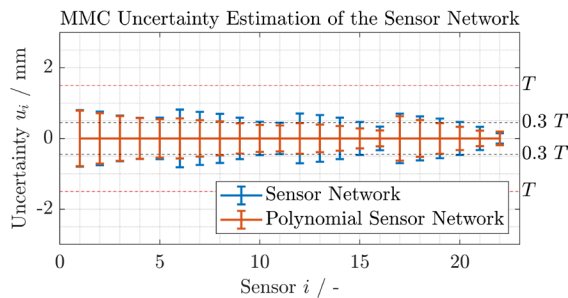


Fig. 7: Estimated uncertainty of each sensor probes position with corrected data from the surface fit correction

Since the position uncertainty propagates through the network and increases proportionally to the square root of the number of connections,

$$u_i \propto \sqrt{M_i}. \quad (5)$$

Thus, the uncertainties for the distant measurement probes relative to the reference probe are greater. This effect can be seen in the stair-step-like rise and fall of the measurement uncertainty, with reference probe #22 exhibiting the smallest uncertainty, which is primarily caused by the reference system. The probe #1 has the largest uncertainty with an uncertainty of $u_1 = 0.9$ mm with a distance $M_1 = 6$. The surface fit corrected measurement points showed a decrease of up to 15% in the measurement uncertainty. This method showed close to no improvement for the sensor probes on the outer edge of the panel.

Results

The local comparison should be evaluated as a first result, since this comparison gives a statement of the systems performance for non-regular shapes. The mean deviation between the sensor network the reference systems with a 1σ error bar is illustrated in Fig. 8. A slight bias is visible in all measurement probes, but it is significantly greater in probes #22 and #11. This is primarily due to the poor coverage of the measurement points at the outer locations by the reference systems. Generally, the sensor network performed better in the comparison to the laser tracker measurements. This can also be attributed to the system's higher sampling rate at the measuring probes.

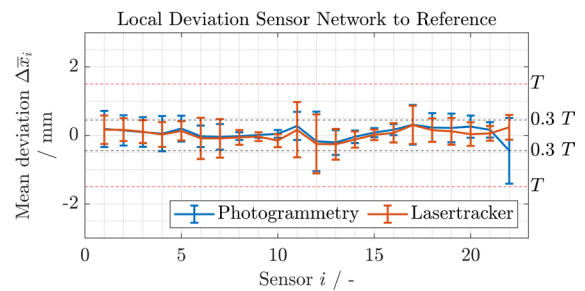


Fig. 8: Local deviation of the sensor network to the two reference measurement systems with error bars of 1σ .

In the global comparison the sensor network performed similarly to the bias deviation from the local comparison. This indicates, that the measurement object can be represented by globally fitted surface polynomial. In Fig. 9 the global deviation from the sensor network to the laser tracker is presented with the surface fit

corrected network. The corrected data showed a great improvement by reducing the average standard deviation $\bar{s}_G$ by 45%.

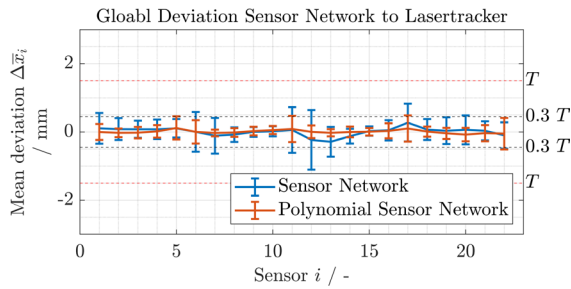


Fig. 9: Global deviation of the sensor network to the laser tracker data with error bars of 1σ .

Figure 10 presents the global deviations of the reference systems with respect to the sensor network. The shape of these curves closely resembles that shown in Fig. 9, indicating that both reference systems capture these deviations consistently. In both systems, outliers such as sensor probes #12 and #17 exhibit increased standard deviation and mean deviation. This behavior can be attributed to their large distance from the reference probe as well as their external positioning. However, the deviations of these probes can be significantly reduced by applying a surface fit correction.

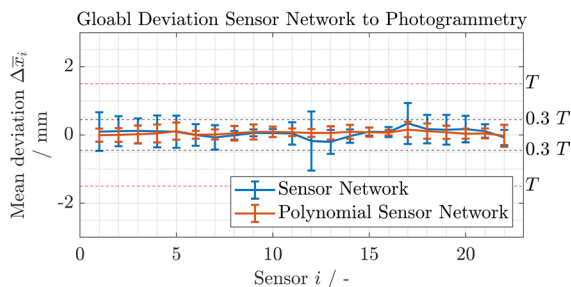


Fig. 10: Global deviation of the sensor network to the photogrammetry data with error bars of 1σ .

In order to provide a better comparison of the methods and systems, the mean deviation and standard deviation of the measurement probes are presented in Table 3. It can be observed that the system shows comparable values for the mean standard deviations $\bar{s}_L$ and $\bar{s}_G$ for both the local and global evaluation methods. In addition, a slight reduction in the mean deviation from the local to the global method is visible. The most significant improvement is observed in the corrected sensor network data, with an enhancement of approximately 45%, accompanied by a slight increase in the average deviation across all sensor probes.

Tab. 3: Linear motor positions for poses 1 to 4, retracted = 0, extended = 1

		Laser tracker		Photogrammetry	
		SN	Poly SN	SN	Poly SN
Local	$\hat{\mu}_L$ / mm	0.032	-	0.067	-
	$\bar{s}_L$ / mm	0.383	-	0.379	-
Global	$\hat{\mu}_G$ / mm	0.006	0.006	0.056	0.056
	$\bar{s}_G$ / mm	0.371	0.211	0.346	0.189

Compared to the calculated uncertainties, the network showed lower values, indicating that the individual sources of uncertainty had been overestimated. However, the outliers in measurement probes #11 and #17 show good agreement with the MMC. Therefore, the simulated uncertainties can be used to predict the factors in future applications. For the previously specified tolerance T , the capability factors obtained from the MSA are presented in Table 4. Using the local evaluation method for the probe positions, only 15% of the laser tracker measurements and approximately 23% of the photogrammetry measurements meet the requirements for the short-term capability index C_g . In both cases, around 35% of the probes satisfy the C_g criterion. These results indicate that a local evaluation approach is only feasible for larger tolerances in accordance with the relevant standards. Slightly higher C_g indices were observed for the global evaluation. In fact, more than 55% of the measurement points exhibited no significant bias with respect to the tolerance T . After correction, this proportion increased to over 70% of the indices. In particular, the bias-related index C_{gk} exceeded

Tab. 4: Capability indices from the MSA of the sensor network in the global and local evaluation method.

		Laser tracker		Photogrammetry	
		Local			
		SN	Poly SN	SN	Poly SN
C_g	1.00 – 1.33	9.1%	–	18.2%	–
	1.33 <	4.5%	–	4.5%	–
C_{gk}	1.00 – 1.33	13.6%	–	9.1%	–
	1.33 <	22.7%	–	27.3%	–
C_{MP}	1.00 – 1.33	3.4%	–	3.4%	–
	1.33 <	13.8%	–	13.8%	–
		Global			
		SN	Poly SN	SN	Poly SN
C_g	1.00 – 1.33	9.1%	18.2%	13.6%	36.4%
	1.33 <	13.6%	50.0%	22.7%	40.9%
C_{gk}	1.00 – 1.33	13.6%	18.2%	13.6%	0%
	1.33 <	50.0%	77.3%	40.9%	100%
C_{MP}	1.00 – 1.33	22.7%	9.1%	0.0%	18.2%
	1.33 <	22.7%	68.2%	36.4%	77.3%

95% compliance in both systems. Also, the measurement process index is met with 80-90% for the corrected data. By applying corrected measurement data to a regular surface, the results can be significantly improved for industrial applications.

Conclusion

In this work, the sensor network is compared with conventionally used industrial measurement systems, namely the laser tracker and photogrammetry, based on a dedicated experimental setup. The study is further designed to evaluate the performance of the network in accordance with relevant standards for a specified tolerance. In this context, multiple evaluation methods are applied to compare the reference systems with the sensor network. The correction of the sensor network data using a surface fit leads to a significant improvement, with more than 70% of the indices meeting the required capability criteria. These results demonstrate the feasibility of the system for deployment in an industrial environment, particularly for regularly shaped components.

Acknowledgements

The authors would like to acknowledge Mr. Michael Geist and Mr. Dominik Hack for their highly appreciable support.

This research is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU (project “Digital Sensor-2-Cloud Campus Platform” (DS2CCP), <https://dtecbw.de/home/forschung/hsu/projektds2ccp>).

References

- [1] DIN ISO 22514-7: Statistische Verfahren im Prozessmanagement – Fähigkeit und Leistung – Teil 7: Fähigkeit von Messprozessen, Berlin., Mar. 2025.
- [2] S. Bracke, Technische Zuverlässigkeit: Datenanalytik, Modellierung, Risikoprognose, 1. Aufl. 2022. Berlin, Heidelberg: Springer Berlin Heidelberg, 2022. doi: 10.1007/978-3-662-65015-8.
- [3] BIPM, “Evaluation of measurement data — Guide to the expression of uncertainty in measurement,” Sep. 2008, doi: 10.59161/JCGM100-2008E.
- [4] L.-M. Bretthauer, R. Heynicke, and G. Scholl, “Uncertainty Modeling of a Cyber-Physical Finite Element Sensor Network for Surface Measurements,” vol. 22. GMA/ITG-Fachtagung Sensoren und Messsysteme 2024, pp. 380–387, Jun. 2024, doi: 10.5162/sensoren2024/D5.1.
- [5] L.-M. Bretthauer, R. Heynicke, and G. Scholl, “Design methodology and uncertainty estimation of a wireless sensor network for surface strain and shape measurements,” *tm - Technisches Messen*, Apr. 2025, doi: 10.1515/teme-2025-0008.
- [6] C. P. Keferstein and M. Marxer, *Fertigungsmesstechnik: Praxisorientierte Grundlagen, moderne Messverfahren*. Wiesbaden: Springer Fachmedien, 2015. doi: 10.1007/978-3-8348-2583-4.
- [7] L.-M. Bretthauer, R. Heynicke, and G. Scholl, “Proposal of a Cyber Physical Finite Element Sensor Network for Surface Measurements,” in 2024 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), May 2024, pp. 1–6. doi: 10.1109/I2MTC60896.2024.10560668.
- [8] A. Segal, D. Haehnel, and S. Thrun, “Generalized-ICP,” in *Robotics: Science and Systems V, Robotics: Science and Systems Foundation*, Jun. 2009. doi: 10.15607/RSS.2009.V.021.
- [9] BIPM et al., “Evaluation of measurement data — Supplement 2 to the ‘Guide to the expression of uncertainty in measurement’ — Extension to any number of output quantities,” Oct. 2011, doi: doi.org/10.59161/JCGM102-2011.

Direct 3-D Printing of Conductive Polymers on Metal Sheets: A Novel Electrical Connection Strategy

Steffen Schröder¹, Thomas M. Wendt¹ and Stefan J. Rupitsch²

¹ Offenburg University of Applied Sciences, Work-Life Robotics Institute, 77652 Offenburg, Germany

² University of Freiburg, IMTEK - Department of Microsystems Engineering, 79110 Freiburg, Germany
steffen.schroeder@hs-offenburg.de

Summary: This study presents a novel electrical connection strategy for additive manufactured electronics and sensors. The conductive polymer layer was printed directly on a metal surface using fused filament fabrication. This eliminating intermediate bonding steps and post processing effort for electrical connecting printed structures. Contact resistance was characterized through electrical measurement of different 3-D-printed samples. This direct deposition strategy establishes robust interfacial connections between conductive polymers and metal substrates while eliminating printing process interruptions and post-processing complexity. The approach demonstrates significant potential for advancing additively manufactured electronics, printed sensor systems, and integrated electronic devices.

Keywords: Additive manufacturing, electrical connection, printed electronics, conductive polymers, contact resistance.

Motivation

The integration of 3-D-printed conductive elements with conventional electronics raises significant challenges. The bulk resistance is typically low enough to be neglected under low-current conditions for metals. However, the contact between a conducting polymer and a metal is expected to be significantly more resistive [1]. Current methods for establishing electrical connections between additive manufactured parts and traditional electronic components vary widely in their approaches.

One approach is to rely on mechanical pressure to establish contact between the polymer and a metal component, such as a screw or an alligator clip. A stainless steel screw has been used in [1], but such contacts can suffer from mechanical instabilities that cause variations in electrical contact resistance, leading to increased noise or irreproducible measurements. Surface roughness further compromises mechanical electrical contact. Conductive silver ink can be applied to the printed structures to establish a better electrical contact surface [2]. Another method with silver ink is to glue copper pads next to the printed conductive polymer. Connecting wires to the copper pads. Then, the copper and polymer pads are coated with conductive silver ink creating the electrical connection [3], [4], [5]. The manual application causes variability in thickness and coverage, which compromises reproducibility. Furthermore, application to intricate geometries remains challenging.

The surface roughness can also be improved using an ironing. This involves a hot nozzle going over the printed surface with finer line spacing to melt the surface and fill in the gaps [6]. Also manual surface heat treatment creating a

smoother finish [7]. This still depends on post printing connections.

Melting the polymer to connect it with the metal increases the effective contact area and lowers the electrical contact resistance. Embedding a copper wire into the electrode body during 3-D printing represents an implementation of this approach [8], [9]. This method necessitates manual wire insertion during printing interruption, introducing assembly variability that affects reproducibility. Post-printing soldering of components onto 3-D-printed conductive substrates is feasible as shown in [10]. This technique exhibits low reproducibility and is recommended exclusively for rapid prototyping of individual structures.

Electrical connection Strategy

In this study, we introduce a novel electrical connection strategy that leverages direct deposition of conductive polymer onto metal substrates during fused filament fabrication (FFF), eliminating additional connecting methods.

The conductive polymer layer is printed directly on a metal surface using FFF. The process is automated by the 3-D printer and no manual intervention is necessary. This eliminates intermediate bonding steps and post processing effort for electrical connecting printed structures. Only a previous adjustment to the printing bed is necessary. The metal plates must be secured to the printing bed. This was done using magnets, but other detachable connection are viable. The resulting offset in z-direction due to the metal thickness can be corrected by the 3-D printer, through print bed calibration before each print. Otherwise, a manual correction is needed before printing.

The first layer has the smoothest surface, resulting in a larger contact area. Molten polymer fills substrate surface irregularities. In addition, preheating of the print bed further enhances interfacial adhesion compared to mid-process plate insertion.

Material and Methods

Test samples were designed to validate the concept (see Fig. 1). The shape consists of two inclined legs forming an A-frame, connected by a horizontal top bar. The 60° angle enables production without support structures. Each leg ends on a flat rectangular base plate at the bottom, for improved surface area and better bonding to the print bed. The geometric dimensions of the base plates are 7 mm in width, 7 mm in length, and 0.4 mm in thickness, respectively. The cross-section of the inclined legs is rectangular with 9 mm² area size. The samples are printed with conductive PLA (Proto-pasta). They were manufactured using a Prusa MK4S printer with 0.2 mm layer height and 100 % infill. The print bed was covered by two copper plates. The z-direction offset resulting from metal thickness was automatically compensated via print bed leveling prior to each print job. This eliminates potential print bed calibration errors associated with manual processing. A small gap between the plates provides electrical isolation between them.

The print bed temperature for the first layer was increased from 60 °C to 100 °C. At 60 °C, the material exhibited poor adhesion to the metal substrate, whereas 100 °C ensured robust bonding.

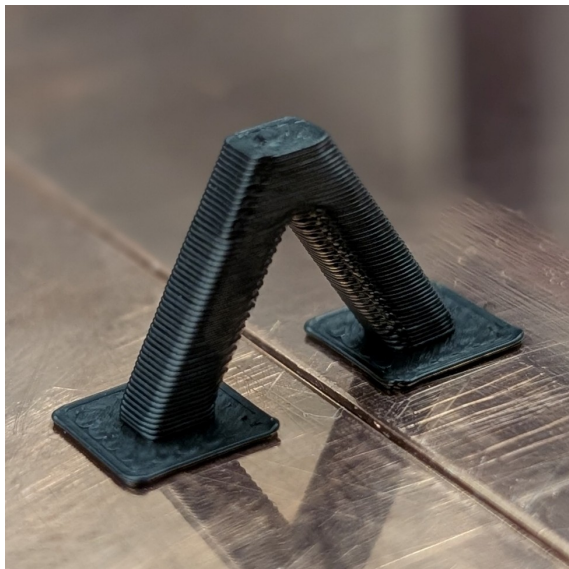


Fig. 1: Direct 3-D-printed resistor sample: conductive polymer deposited directly onto a metal substrate to form an electrically connected structure.

Experimental setup

We used a four-wire measurement setup to determine the samples resistance. In this configuration two force wires deliver a constant current through the samples. The other two wires measure the voltage alongside the sample. This separation eliminates errors from the wire and contact resistance of the voltage sensing wires. For a statistically significant result, 500 measurements were taken with an Keithley 2450 SourceMeter, and the mean values were calculated. Three different measurements were conducted where the sensing wires had various measuring positions. The measuring positions are shown in Fig. 2. The force wires were connected to the copper plates at position 0 and 5. The total resistance of a sample were measured at position 1 and 4, further referred to as R_{14} . To calculate the contact resistance R_{12} and R_{34} between the copper and the 3-D-printed part the resistance R_{13} and R_{24} were measured and subtracted from R_{14} .

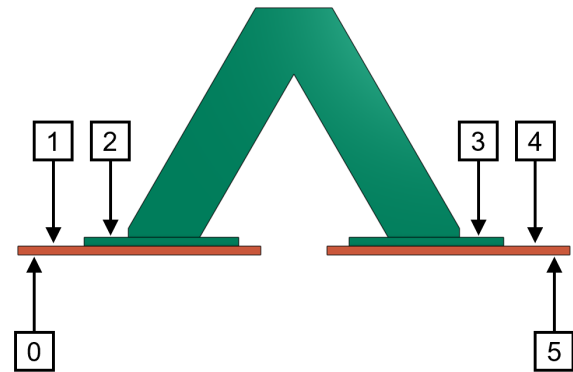


Fig. 2: Sample showing marked positions for electrical measurements.

Results and Discussion

We tested the samples in an experimental setup. The standard deviation of the measured values ranges from 5 mΩ to 17 mΩ. The electrical resistance among the five samples is similar but not the same, as Fig. 3 shows. The five samples R_{14} value exhibits a range 23.09 Ω with a mean resistance of 487.96 Ω. The resistance range compared to the mean resistance is 4.73 %. This behavior could depend on a not uniform electrical resistance in the raw filament before being melted together. The inhomogeneity of the conductive particle in the filament creating irregular conductive paths. This is also reported in other literature [3], [11].

Figure 4 shows the contact resistance R_{12} for the five samples. The values span from 0.21 Ω (Sample 3) to 1.19 Ω (Sample 2), with a mean of 0.76 Ω. The range of 0.98 Ω reflects consistent contact quality, comparable to R_{34} measurements despite slightly higher variability.

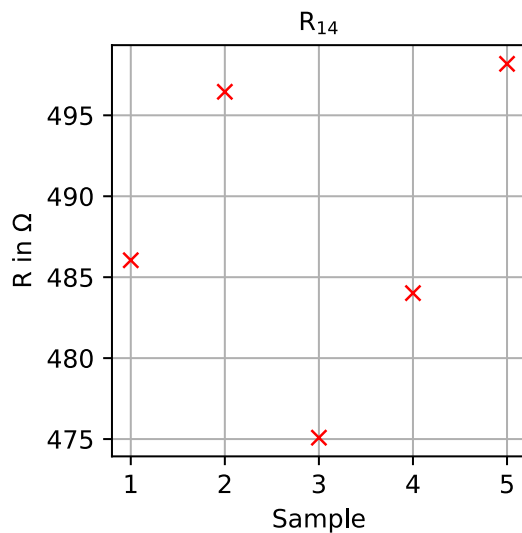


Fig. 3: Measured total mean resistance R_{14} of five samples.

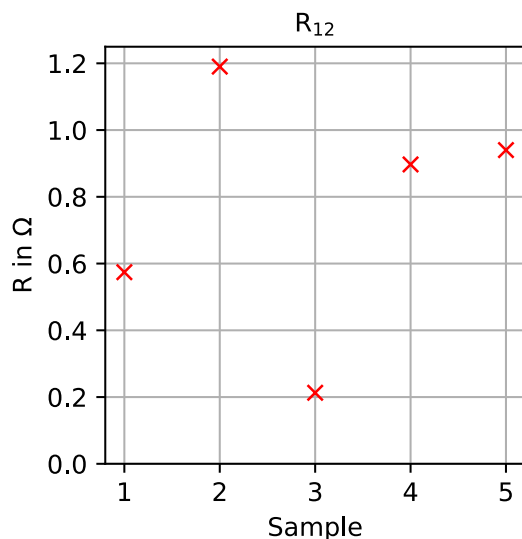


Fig. 4: Mean resistance R_{12} of five samples.

Figure 5 presents the resistance R_{34} for the five samples. The values span from 0.13Ω (Sample 3) to 0.71Ω (Sample 4), with a mean of 0.54Ω . The majority of samples cluster tightly between 0.57Ω to 0.71Ω , while Sample 3 exhibits notably lower resistance, potentially attributable to superior electrode-surface bonding or printing uniformity.

The resistance R_{12} and R_{34} consists of the contact resistance between copper and polymer as well as the resistance of the 3-D-printed base plate part. It is not possible to determine the contact resistance exactly, but it should be lower than 1.19Ω . The low range of 0.98Ω for R_{12} and 0.59Ω for R_{34} across the samples indicates

reproducibility of the 3-D-printed electrode contacts, consistent with the overall resistance measurement. The difference may due to the two copper plates geometrical variation or surface quality.

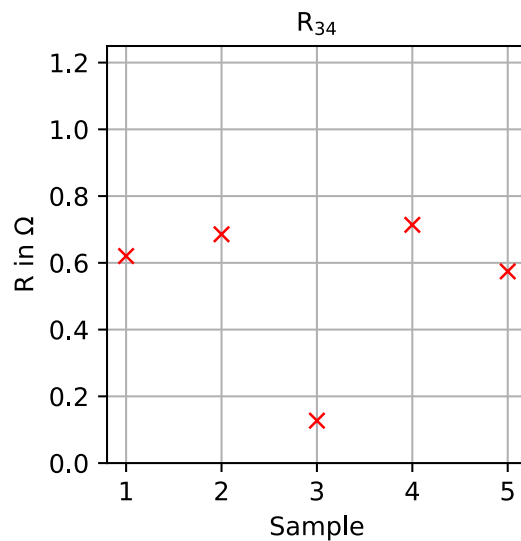


Fig. 5: Mean resistance R_{34} of five samples.

Conclusion

We demonstrated a novel connection strategy for 3-D-printed electronics. Fully 3-D-printed resistors have been fabricated using this new approach and evaluated. This method, using a conductive metal sheet as print beds in combination with printed conductive polymers, opens up new design possibilities and enables fully additive one-shot electronics manufacturing, without mid- or post-process connection. The concept can be easily scaled and adapted for a wide range of printed electronics. The measurement results indicate a reproducible contact resistance lower than 1.19Ω .

Future work

The influence of printing parameters on the electrical connection quality, along with the thermal behavior of various print bed materials, could be investigated in future research. The full integration of the proposed connection strategy into the FFF-printing process represents a key direction. Specifically, a modified print bed configuration, featuring cut-outs to securely position the copper plates is anticipated (see Fig. 6). This adaptation enables seamless process integration and supports the scalability of the approach for industrial additive manufacturing applications. Such enhancements are expected to yield a fully practical connection methodology compatible with standard FFF workflows.

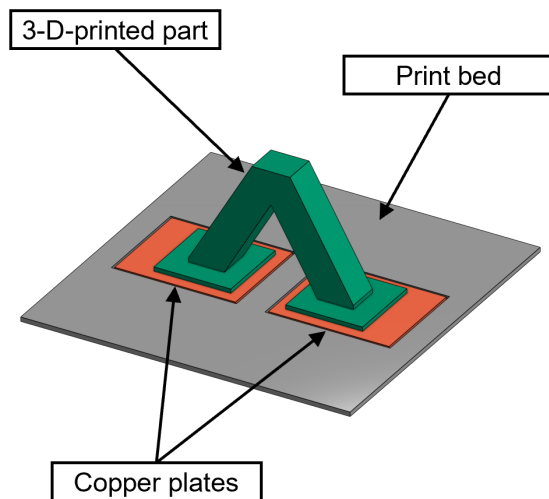


Fig. 6: Future integration in FFF-printing process with modified print bed.

References

- [1] I. Verlangieri, T. G. de Oliveira, F. S. Lopes, I. G. R. Gut, L. Angnes, and C. L. do Lago, "Handling electric connections in 3D-printed electrodes and sensors," *Microchimica Acta*, vol. 192, p. 265, Mar. 2025.
- [2] T. Blachowicz, G. Ehrmann, and A. Ehrmann, "Recent Developments in Additive Manufacturing of Conductive Polymer Composites," *Macromolecular Materials and Engineering*, vol. 308, no. 7, p. 2200692, 2023.
- [3] N. Hangst, T. M. Wendt, and S. J. Rupitsch, "Fully 3D-printed gripper jaw with embedded sensitive sensor structures for robotic applications," *Sensors and Actuators A: Physical*, vol. 399, p. 117334, Mar. 2026.
- [4] P. Gawron, T. M. Wendt, and S. J. Rupitsch, "Additively Manufactured Multimaterial Cantilever for Electromagnetic Vibration Harvesting," *IEEE Sensors Journal*, vol. 25, pp. 6593–6605, Feb. 2025.
- [5] R. Waltersbacher, L. Stiglmeier, and T. M. Wendt, "Additive Manufacturing for Inductive Power Transfer in Robotic Applications," in *2025 The 16th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT)*, pp. 143–150, May 2025. ISSN: 2694-3182.
- [6] B. Neuhaus, M. K. Idris, P. Naderi, Y. El-Hajj, and G. Grau, "Low-Roughness 3D-Printed Surfaces by Ironing for the Integration with Printed Electronics," *Advanced Engineering Materials*, vol. 26, no. 3, p. 2301711, 2024.
- [7] L. Stiglmeier, T. M. Wendt, N. Hangst, P. Gawron, and S. J. Rupitsch, "A guideline for the fabrication of fully 3D-printed torque sensor elements - demonstrated based on a real example," *Vorträge*, pp. 213–220, June 2024.
- [8] S. Schröder, T. M. Wendt, and S. J. Rupitsch, "Conceptual design of additive manufactured capacitive displacement sensors for adaptive pin array grippers," *Journal of Sensors and Sensor Systems*, vol. 14, pp. 265–273, Nov. 2025. Publisher: Copernicus GmbH.
- [9] S. Schröder, T. M. Wendt, and S. J. Rupitsch, "Additive Manufactured Capacitive Displacement Sensor for Pin Array Grippers: One-Shot and Stacked Fabrication Approaches," in *2025 IEEE SENSORS*, pp. 1–4, Oct. 2025. ISSN: 2168-9229.
- [10] B. Podsiadły, A. Skalski, and M. Słoma, "Soldering of Electronics Components on 3D-Printed Conductive Substrates," *Materials*, vol. 14, p. 3850, Jan. 2021. Publisher: Multidisciplinary Digital Publishing Institute.
- [11] J. H. Ahn, D. Kim, S. A. Park, J. S. Hong, and K. H. Ahn, "Variation of the Electrical Conductivity of PLA-Based Composites with a Hybrid of Graphene and Carbon Black During 3D Printing," *Macromolecular Materials and Engineering*, vol. 307, no. 9, p. 2200145, 2022.

Photoakustisches Messsystem zur Produktüberwachung

*Heike Wünscher, Thomas Frank, Christoph Heinze, Toni Schildhauer
CiS Forschungsinstitut für Mikrosensorik GmbH, Erfurt, 99099, Deutschland*

Zusammenfassung

Um den Restammoniakgehalt im Gas am Ausgang eines Crackers zu bestimmen, wurde ein nicht-resonantes photoakustisches Messsystem aufgebaut, das zunächst mit Labormesstechnik betrieben wird und für die Messung im Durchfluss vorgesehen ist. Es basiert auf einem innovativen thermischen Emitter und arbeitet ohne Chopper. Es wurde die Abhängigkeit des akustischen Signals vom Design der Detektionskammer studiert. Erste Ergebnisse werden vorgestellt. Die Versuche mit unterschiedlichen Gasfüllungen zeigen, dass ein starkes Hintergrundsignal auftritt, das überwiegend auf die thermomechanischen Wechselwirkungen des IR-Lichts mit den Komponenten des Messsystems zurückzuführen ist. Elektronische Einflüsse, die aus dem Betrieb des Strahlers oder der Kabel resultieren, erweisen sich als weitaus geringer.

Keywords: Gassensorik, optische Messverfahren, MEMS-Mikrofon, Emitter

Motivation

Beim Cracken von Ammoniak entsteht ein Gasgemisch aus Wasserstoff und Stickstoff. Dabei kann stets ein Restgehalt an nicht umgesetztem Ammoniak enthalten sein. Um die Einspeisung dieses Gases in einen Verbraucher (z. B. eine Brennstoffzelle) zu ermöglichen, muss der Ammoniakgehalt unterhalb eines definierten Limits gehalten werden. Dazu ist ein Sensor erforderlich, der kontinuierlich oder in kurzen Zeitintervallen messen kann und für den Einbau in eine Rohrleitung oder einen Bypass geeignet ist. Für diese Anwendung wird eine niedrige Nachweisgrenze für Ammoniak und eine hohe Ansprechgeschwindigkeit gebraucht.

Um beides zu erreichen, wurde ein photoakustischer Ansatz gewählt. In der Regel erfolgt die Detektion resonant, entweder mit gepulsten Laserquellen oder mit thermischen Emittoren unter Verwendung eines Blendenrads (Chopper) [1–2]. Bei Lasern besteht der wesentliche Nachteil in der Notwendigkeit einer Frequenzstabilisierung, was sich preislich niederschlägt. Die Nutzung eines thermischen Strahlers und Choppers reduziert dagegen die mechanische Robustheit. In dieser Arbeit wird daher ein nicht-resonanter Ansatz unter Verwendung eines innovativen thermischen Emitters vorgestellt.

Experimenteller Ansatz

Im Gas am Ausgang des Crackers ist nicht umgesetztes Ammoniakgas die einzige

infrarotaktive Komponente, sodass dessen Absorptionen zur Bestimmung genutzt werden können. Der Lösungsansatz basiert auf dem photoakustischen Effekt, bei dem die elektromagnetische Strahlung vom Gas absorbiert und akustisch detektiert wird.

Die Absorption verursacht eine lokale Erwärmung, die eine thermische Ausdehnung bewirkt und somit eine Druckwelle erzeugt, die registriert werden kann. Dafür wurde ein modulares System entwickelt, das für einen IR-Strahler in einem TO-39-Gehäuse ausgelegt ist. Strahler und Empfänger sind an den gegenüberliegenden Enden angebracht.

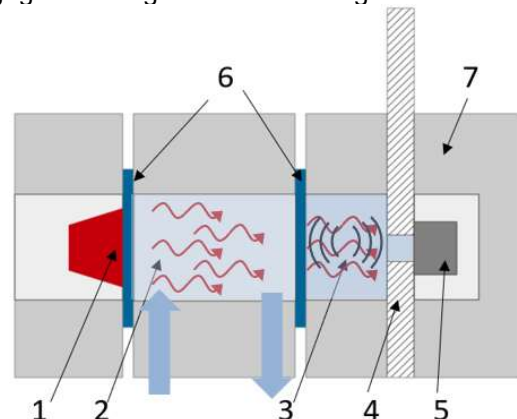


Abb. 1: Schematischer Aufbau: 1 Strahler, 2 Probenraum, 3 Detektionskammer, 4 PCB, 5 MEMS-Mikrofon, 6 Optische Fenster, 7 Abstandshalter und Rückplatte

Zur Konzentrationsbestimmung wird das Gas in den Probenraum geleitet und gepulst bestrahlt (Abb. 1). Da die stärksten Absorptionen im Bereich zwischen ca. 8 und 13 μm liegen, wird ein breitbandiger Infrarotstrahler genutzt. Das Licht wird entsprechend des Lambert-Beer-Gesetzes geschwächt. Das transmittierte Licht gelangt schließlich in die Detektionskammer und dient dort zur Schallerzeugung. Für eine selektive und empfindliche Erfassung wird die Kammer gewöhnlich mit dem hoch konzentrierten Zielgas gefüllt. Das durch die Lichtabsorption erzeugte akustische Signal wird von einem MEMS-Mikrofon erfasst. Die Frequenz des erzeugten Schalldrucks entspricht der des jeweils stimulierenden Lichts. Durch die niedrigen Modulationsfrequenzen der thermischen Strahler resultiert bei der hier verfolgten Lösung eine enorme Wellenlänge. Die Detektion findet daher im sog. Nahfeld statt, das stark durch die Zellgeometrie beeinflusst wird. Dadurch war es zunächst notwendig, den Einfluss des Designs auf das akustische Signal zu betrachten.

Setup

Entsprechend der Zielstellung dieser Arbeit wurden nach Möglichkeit Komponenten genutzt, die entweder kommerziell erhältlich sind oder die durch Skalierung preisgünstig in größeren Stückzahlen hergestellt werden können. Der Aufbau wurde daher mit thermischen Strahlern realisiert, die gewöhnlich mit einer wenig aufwendigen Elektronik bei der Steuerung und Detektion auskommen.

Das Setup basiert auf einem kommerziellen bottom-port MEMS-Mikrofon ICS-40730 (TDK). Es zeichnet sich durch eine sehr hohe Empfindlichkeit auch bei sehr tiefen Frequenzen aus. Es wurde auf einer PCB aus FR4 aufgebaut (Abb.2).

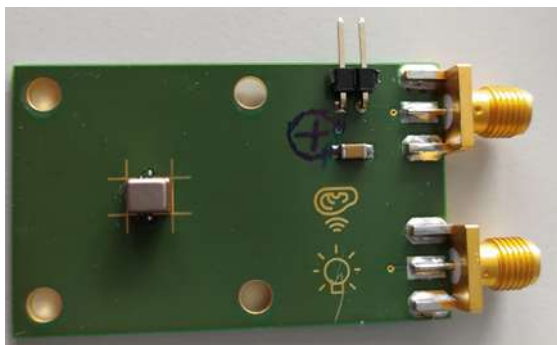


Abb.2: PCB mit MEMS-Mikrofon

Der Strahler wurde als Array aus 16 gleichartigen Einzelemitter aufgebaut, sodass sie mit vergleichsweise hoher Frequenz getaktet werden können [3]. Die einzelnen IR-Strahler wurden direkt auf einen TO-39 Sockel geklebt und durch Drahtbonden elektrisch verbunden. Sie sind in Zweiergruppen parallel verschaltet und haben eine nominale Leistungsaufnahme von insgesamt 1600 mW. Um den Anteil kurzwelligen Lichts im Vergleich zum Nutzlicht ($> 8 \mu\text{m}$) zu verringern und damit Falschlicht zu reduzieren, wurde das Array bei einer deutlich geringeren Leistung betrieben. Der Pulsbetrieb der Baugruppe wird mit einer Transistorschaltung realisiert, die einen Spannungsoffset und eine -amplitude generiert. Die Versorgung erfolgt jeweils spannungsbegrenzt. Der IR-Strahler und der Empfänger werden über einen Lock-In Verstärker (HF2LI, 50 MHz, 210 MSa/s, Fa. Zürich Instruments) synchronisiert. Dabei wird mit einer sinusförmigen Anregung gearbeitet.

Für die Versuche wurde die Pulsfrequenz des Emitters im Sweep-Modus des Lock-in-Verstärkers zwischen 1 und 100 Hz variiert. Bei geschlossener Kammer konnte durch Anwendung des Low-Pass-Filters mit einem Abklingverhalten von mindestens 24 dB/oct eine weitest gehende Ausblendung der Umgebungsgeräusche erreicht werden, sodass keine zusätzliche Abschirmung erforderlich war.

Design der Detektionskammer

Das Detektionssystem wurde optimiert, indem die Abhängigkeit des akustischen Signals mit unterschiedlichen Gasfüllungen vom Design der Detektionskammer ermittelt wurde. Dazu wurde der Aufbau ohne Probenkammer verwendet, sodass die Detektionskammer direkt bestrahlt wird (Abb. 3). Es wurden mehrere systematisch variierte Kammerkonstruktionen mit zylindrischem bzw. konischem Innenraum untersucht, die sich auch in der Länge des optischen Pfades unterscheiden (Tab. 1). Für eine gute Lichtausbeute sind alle beleuchteten Bereiche poliert ausgeführt. Die Schallgeneration wurde bei Füllung mit Luft und reinem Stickstoff betrachtet. Das Nutzsignal stellt das Differenzsignal aus beiden Füllungen dar (Abb. 4). Es zeigt sich bei allen bisher durchgeführten Versuchen, dass die intensivsten akustischen Signale bei einer sehr niedrigen Stimulationsfrequenz des thermischen Emitters erreicht wird ($< 10 \text{ Hz}$).

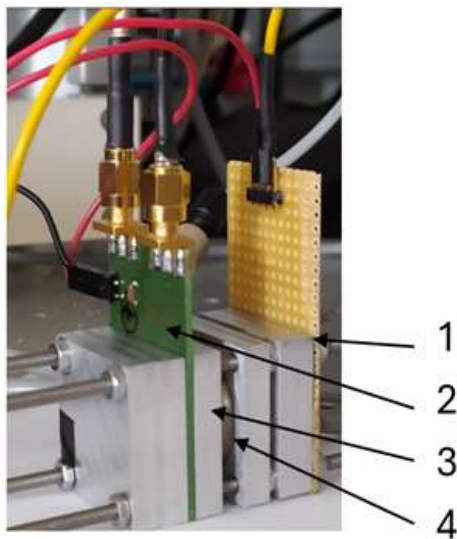


Abb.3: Detektionseinheit des photoakustischen Sensors, 1 Halterung Strahler-Array, 2 MEMS-Mikrofon, 3 Detektionskammer, 4 Ge-Fenster, AR-coated 7 – 12 μ m

Tab. 1: Untersuchte Kammerdesigns, ¹⁾ die Konusöffnung ist zum Mikrofon gerichtet

	Form	Länge in mm	Ø in mm
Z8	zylindrisch	8	8,5
Z12	zylindrisch	12	8,5
Z50	zylindrisch	50	8,5
K8	konisch ¹⁾	8	8,5 / 12,5

Im Falle der luftgefüllten Kammer absorbieren das enthaltene CO₂ sowie der Wasserdampf die IR-Strahlung, sodass ein akustisches Signal resultiert. Obwohl reiner Stickstoff IR-Licht nicht absorbiert, wurde auch bei dieser Füllung ein intensives Signal erhalten (Tab. 2, Abb. 4), was auf thermomechanische Wechselwirkungen des IR-Lichts mit den Gehäusekomponenten, die sich im Strahlengang befinden (Fenster, Seitenwände der Kammer, Leiterplatte usw.) zurückzuführen ist und ein Hintergrundsignal bildet.

Bei den Kammern mit den kurzen optischen Wegen (Z8, Z12 und K8) wird hinsichtlich des akustischen Signals ein Maximum bei einer Stimulationsfrequenz von ca. 5 Hz erhalten (Abb. 4, oben). In der längeren Kammer Z50 wird ein deutlich verändertes akustische Verhalten der Gase beobachtet. (Abb. 4, unten). Bei beiden Gasen fällt das Signal nahezu im gesamten betrachteten Bereich mit

zunehmender Frequenz monoton ab. Ein lokales Maximum ergibt sich nur im Differenzsignal. Letzteres liegt bei deutlich höheren Frequenzen als bei den übrigen Kammern (ca. 17 Hz).

Um den Einfluss der Seitenwände auf das Hintergrundsignal zu reduzieren, wurde die konische Zelle K8 entworfen. Sie weist dieselbe Länge auf wie die zylindrische Zelle Z8, besitzt allerdings ein deutlich größeres Innenvolumen. Diese wurde so eingesetzt, dass die größere Öffnung zum Mikrofon zeigt. Auf diese Weise werden die Kammerwände abgeschattet.

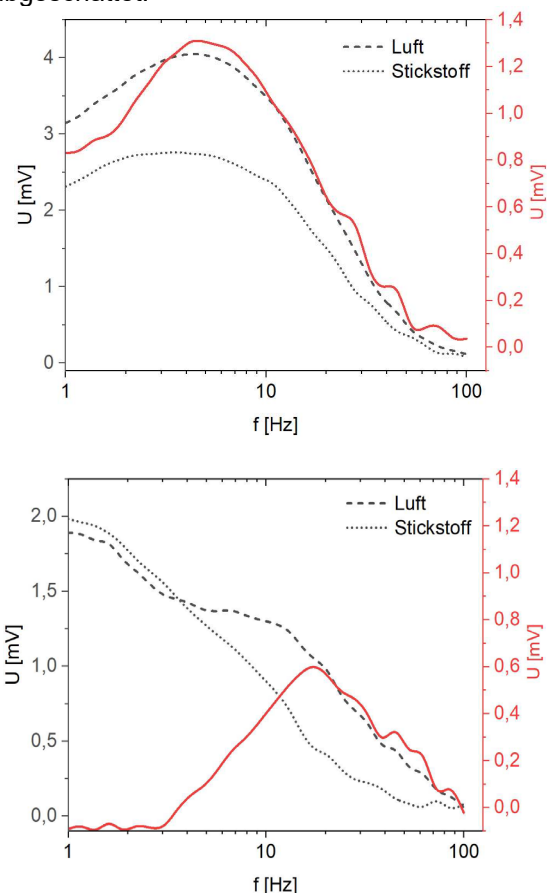


Abb. 4: Frequenzabhängigkeit des akustischen Signals, erhalten mit dem Emitterarray bei 330 mW; Z8 (oben) und Z50 (unten), rot: Differenzsignal

Es zeigt sich jedoch, dass sich bei der konischen Zelle zwar das Hintergrundsignal verringert, aber der in Luft gemessene Wert nimmt noch deutlich stärker ab, sodass sich letztlich das Nutzsignal reduziert. Die Wechselwirkungen mit den Seitenwänden sind offenbar weniger bedeutsam für die Qualität der Gasmessung. Wie in Tab. 2 zu erkennen ist, wird von allen untersuchten Varianten das größte Nutzsignal mit Z8 erhalten.

Tab. 2: Maximale akustische Signale, erhalten bei 330 mW, ^{a)} bei 17 Hz

	Luft $U_{\max}$ in mV	Stickstoff $U_{\max}$ in mV	Nutzsignal ΔU in mV
Z8	4,06	2,74	1,32
Z12	3,20	2,18	1,02
Z50	1,04 ^{a)}	0,43 ^{a)}	0,61
K8	2,73	2,20	0,53

Dieses Ergebnis steht im Einklang mit früheren Arbeiten [4-7].

Zweikammeraufbau

Für die angestrebten Messungen im Durchfluss wurden erste Versuche mit einem 5 cm langen Probenraum und der luftgefüllten Detektionskammer Z8 durchgeführt. Dabei zeigte sich allerdings, dass die Signalhöhe auch von äußeren Bedingungen des Probenraums und den Gaszuleitungen (wie Krümmungen oder geöffnete bzw. geschlossene Hähne) abhängt. Vermutlich ist dies auf mechanische Kopplungen zwischen beiden Kammern zurückzuführen, sodass das in der Detektionskammer erzeugte akustische Signal modifiziert wird. Weitere Arbeiten zur Entkopplung der Komponenten sind daher notwendig.

Elektronische Störungen

Um elektronische Einflüsse auf das akustische Signal zu bewerten und das System zu optimieren, wurde das Übersprechen zwischen den elektronischen Komponenten untersucht. Zu diesem Zweck wurde der Emitter im Messsystem betrieben, indem die Rückseite der Strahlerhalterung der Messzelle zugewandt wurde, sodass weder Licht- noch Wärmeeinflüsse in die Kammer gelangen konnten. Es zeigte sich, dass bei Betrieb des Emitters ein Mikrofonsignal von etwa 80 μV auftritt, unabhängig von der Strahlungsleistung und der Kammerfüllung. Der Einfluss der Kabelanordnung war noch geringer.

Schlussfolgerungen

Zur Bestimmung des Ammoniakgehalts am Ausgang eines Crackers soll ein photoakustisches System entwickelt werden, das mit hoher Empfindlichkeit und kurzer Ansprechzeit kontinuierlich oder in kurzen Zeitintervallen messen kann.

Ein erster Aufbau wurde realisiert und zunächst mit Labormesstechnik betrieben. In Übereinstimmung mit Ergebnissen anderer Gruppen wurde gefunden, dass die Kammer

mit dem kleinsten Innenvolumen (Z8) das stärkste Nutzsignal liefert. Die Versuche zeigen auch, dass das starke Hintergrundsignal überwiegend auf die thermomechanischen Wechselwirkungen des IR-Lichts mit den Komponenten des Messsystems zurückzuführen ist, wobei den Seitenwänden eine geringere Bedeutung zukommt.

Eine deutliche Effektivierung der mechanischen Entkopplung der Komponenten des Messsystems muss zudem erreicht werden, um reproduzierbare Messergebnisse bei Integration in ein Rohrleitungssystem zu erhalten. Weitere Entwicklungen dazu sind im Gange.

Danksagung

Wir danken dem Bundesministerium für Wirtschaft und Energie (BMWE) für die finanzielle Unterstützung des Projekts NH3-BZ (49MF220229).

Gefördert durch:



Bundesministerium
für Wirtschaft
und Energie

aufgrund eines Beschlusses
des Deutschen Bundestages

Literaturnachweis

- [1] O. E. Bonilla-Manrique, J.E. Posada-Roman, J. A. Garcia-Souto, M. Ruiz-Llata, *Sensors* 19, 2890 (2029); doi:10.3390/s19132890
- [2] S. Palzer, *Sensors* 20, 2745 (2020); doi:10.3390/s20092745
- [3] T. Schildhauer, A. T. Winzer, J. Baldauf, Fast IR-MEMS-emitters for NDIR applications Terahertz, RF, Millimeter, and Submillimeter-Wave Technology and Applications XVII. SPIE, 2024.
- [4] M. Dreßler, Entwicklung eines auf dem Prinzip der Photoakustik beruhenden Detektors für Spuren wasserfreien, gasförmigen Fluorwasserstoffs, Dr. Ing. thesis, Universität der Bundeswehr Hamburg, Hamburg 2010
- [5] Y.-H. Pao, *Optoacoustic Spectroscopy and Detection*, pp. 194, Academic Press, New York, (1977)
- [6] A. Rosencwaig, A.: *Photoacoustics and Photoacoustics Spectroscopy*, John Wiley & Sons, (1980)
- [7] F.J.M. Harren, S.M. Cristescu, Part of book (Meyers, R.A. (ed.), *Encyclopedia of Analytical Chemistry Photoacoustic Spectroscopy in Trace Gas Monitoring* (2019); <https://doi.org/10.1002/9780470027318.a0718.pub3>

Assessing the Controllability of Modular UAV Swarms over 5G Networks via ROS 2

Kevin Tebbe¹, Robert Pommeranz¹, Gaurav Kapoor¹, Ralf Heynicke¹, Gerd Scholl¹

¹Chair of Electrical Measurement Engineering, Helmut-Schmidt-University, University of German Armed Forces, Hamburg

Abstract

Multi-UAV operations are constrained by legacy point-to-point radio links that scale poorly in range, bandwidth management, and quality of service. In this feasibility study, the controllability and reliability of a highly modular Robot Operating System 2 based UAV swarm interconnected via private 5G cellular networks are evaluated. After integrating multiple aerial vehicles into a 5G architecture, the communication performance is compared against standard WLAN by profiling end-to-end latency and throughput under representative mission and sensor traffic. The findings provide practical design guidelines for deploying UAV swarms over cellular infrastructure, particularly in Beyond-Line-of-Sight scenarios, where scalability and coverage outweigh raw latency performance.

Keywords: C-UAS, 5G, ROS 2, Swarm, PX4-Autopilot

Introduction

Unmanned Aerial Vehicles (UAVs) are increasingly deployed in complex, distributed scenarios such as disaster response, infrastructure monitoring, and cooperative multi-agent missions. Traditional UAV communication relies predominantly on point-to-point radio links, such as ExpressLRS, FrSky R9 or proprietary Links from DJI or Walksnail. While these systems provide low-latency links for individual platforms, they do not scale efficiently with increasing swarm size, nor do they support flexible multi-node communication patterns required for collaborative autonomy. As swarm sizes grow, the communication paradigm must evolve toward point-to-multipoint (P2X) or fully networked architectures. Cellular communication systems, particularly Fifth Generation (5G) Standalone (SA) networks, offer a promising alternative due to their inherent support for wide-area coverage, mobility management, and quality-of-service (QoS) differentiation. However, unlike local wireless technologies such as Wireless Local Area Networks (WLAN), 5G networks introduce additional architectural complexity, including radio scheduling, core network traversal, and transport-layer buffering. As a result, the communication network becomes an active component of the control loop rather than a transparent channel. This shift raises a fundamental question: to what extent do the latency and reliability characteristics of cellular networks constrain the controllability of UAV

swarms? To address this question, this paper presents an experimental evaluation of a modular UAV swarm architecture based on PX4-Autopilot and Robot Operating System 2 (ROS2). Unlike prior work focusing primarily on network-level performance metrics, this study explicitly links measured communication characteristics to control-relevant constraints such as achievable control loop rates, watchdog design, and synchronization accuracy. In doing so, the paper frames the communication network as a dynamic system component that directly influences swarm behavior and operational safety. The results provide a quantitative comparison of communication technologies and practical design guidelines for deploying UAV swarms over wide-area cellular infrastructure, particularly in Beyond-Line-of-Sight (BLoS) scenarios.

Related Work

Recent research on multi-UAV systems has increasingly focused on modular software and hardware architectures that support scalable swarm operation. In particular, middleware-based approaches built on PX4-Autopilot and ROS2 have gained attention because they enable flexible integration of heterogeneous vehicles, companion computers, and distributed mission logic. A modular and scalable ROS2/PX4-based architecture for heterogeneous UAV swarms was presented in [1], providing the system-level foundation for the platform used in this paper. In parallel, the role

of 5G communication for cyber-physical and industrial real-time systems has been investigated in a growing body of literature. Prior work has shown that private 5G SA networks can provide predictable connectivity and offer advantages such as managed spectrum access, mobility support, and QoS differentiation. The private 5G SA campus network used in this work, including its baseline performance characteristics, is described in [2]. However, existing studies often focus either on communication infrastructure metrics without explicitly linking network behavior to control-relevant constraints in real multi-UAV deployments. In particular, there is still limited experimental work on how end-to-end latency, jitter, and packet loss affect the practical controllability of ROS2-based UAV swarms operating over cellular 5G networks. This paper addresses the gap by experimentally evaluating a modular PX4/ROS2 UAV swarm by interpreting the measured application-level latency and reliability in terms of operational swarm control constraints.

System Architecture and Experimental Setup

The UAVs in this experiment are configured as modular research platforms. They consist of a Pixhawk 6X Flight Control Unit (FCU) running PX4-Autopilot version 1.15, an Nvidia Jetson Orin NX as a companion computer for high-level tasks, and a 5G modem (Quectel RMU500EK or Waveshare Dongle with RM520N-GL). Figure 1 illustrates the hardware used in this experimental setup, whereas Figure 2 depicts the logical network topology, including the protocols and communication pathways between the UAVs and the ground station. The logical connections within the 5G network are omitted from the diagram for clarity. The experimental setup involves interconnecting the Ground Control Station (GCS), designated as the benchmark node, and the UAVs via three different physical layers for comparison: LAN, WLAN, and a private 5G campus network at Helmut-Schmidt-University (HSU). Figure 3 illustrates the communication pathways and protocols between the vehicles and the GCS. This modular architecture enables a clear separation between flight control, onboard processing, and communication layers, allowing the isolated evaluation of network-induced effects on swarm operation.

Measurement Scenario

The benchmark system implements a distributed ROS 2 measurement campaign for one-way message latency measurements in multi-UAV scenarios. The system is structured

as a client-server topology, where a GCS serves as the time synchronization master and the sender nodes (UAVs) that are time-synchronized to the GCS via Network Time Protocol (NTP).



Fig. 1 UAV experimental setup showing Nvidia Jetson, Pixhawk 6X, and Quectel RMU500EK 5G modem mounted on a Holybro X650 Frame.

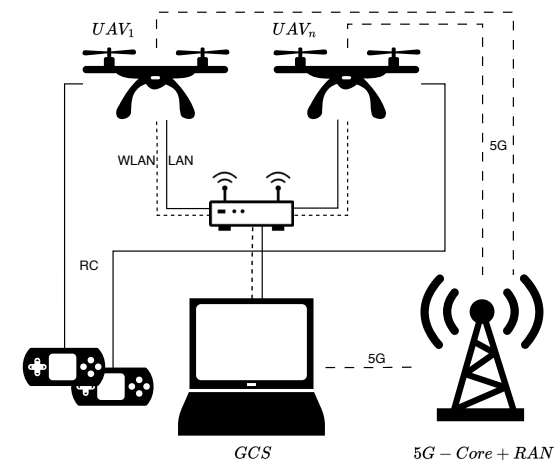


Fig.2 Simplified representation of the benchmark measurement setup for KPI measurement using LAN, WLAN and 5G cellular network.

In this work, latency in this publication is defined as the time difference between message generation at the UAV and reception at the GCS:

$$L_{\text{end-to-end}} = t_{\text{recv}} - t_{\text{send}}, \quad (1)$$

whereas t_{recv} is the timestamp of message reception at the GCS (ROS 2 system clock) and t_{send} is the timestamp of message publication at the UAV (ROS 2 system clock). Clock synchronization across the UAV and GCS is maintained via NTP. Although Precision Time Protocol (PTP) is generally more accurate than NTP, the achieved NTP synchronization accuracy of 1 ms – 2 ms is sufficient for the

latency measurements in this publication, which are in the order of tens to hundreds of ms [3], [4].

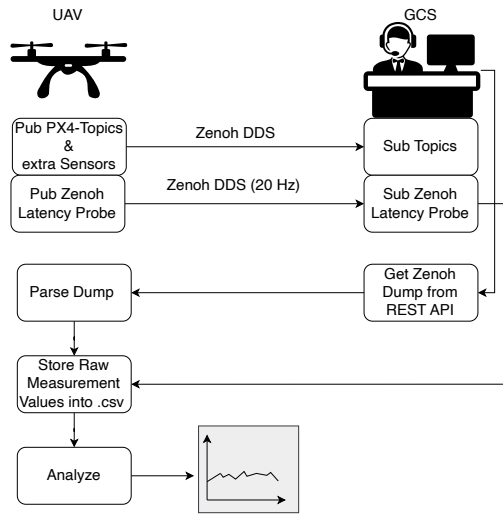


Fig. 3 Benchmark setup for ROS 2 based latency measurements in multi-UAV scenarios. The figure illustrates the generic data flow from a UAV to GCS.

The UAV messages from the PX4 firmware are not using sequence numbers, which makes it difficult to determine the number of lost packets. Therefore, packet loss is determined by analyzing the dedicated latency probe stream on a dedicated latency probe topic using Zenoh (a lightweight communication middleware, detailed in Sec. Network Benchmarking), which contains a sequence number in each JSON payload. By comparing consecutive sequence numbers in the probe, missing probes can be identified and packet loss calculated from the gap count.

Message-Callback Architecture:

The measurement framework follows a dual-path design to capture both end-to-end message delay and control-plane dynamics of the Zenoh edge-communication protocol [5]. The execution flow of each node is determined by command-line initialization arguments passed to a unified measurement script. When launched in GCS mode, the orchestration layer starts a periodic Zenoh dump process and latency subscribers in parallel. When launched in UAV mode, only the probe publisher is active. **Active Probe Path:** A ROS 2 probe publisher transmits JSON payloads containing sequence number and sender timestamp. The subscriber callback extracts the payload, computes the one-way latency, and appends samples to database. For sample i , latency is computed as defined in Eq. 1. with acceptance constraint.

$$0 \leq L_i^{\text{ms}} \leq 6 \times 10^5 [\text{ms}]. \quad (3)$$

Let L denote the sorted latency sample set and let $a \in (0,1)$ represent the target probability level for the desired quantile (e.g., $a = 0.95$ for the 95th percentile). The reported quantiles are then defined as:

$$Q_q = a_{[q(N-1)]}, \quad q \in \{0.50, 0.95, 0.99\}. \quad (4)$$

Snapshot Key Performance Indicator (KPI) Path:

While the primary objective of this study is the evaluation of the end-to-end latency to assess swarm controllability, network throughput is monitored as a secondary KPI. Throughput is recorded not to evaluate maximum capacity, but to verify that the network bandwidth is not exhausted, ensuring that the observed latencies are not artifacts of network saturation. To contextualize these primary physical-layer metrics, the system simultaneously monitors the internal state of the middleware. At fixed interval Δt , the system queries the Zenoh administrative REST API to store a JSON snapshot of the middleware's internal metrics. Runtime topics and endpoint metadata are parsed from each snapshot to compute per-interval KPIs. The evaluated metrics encompass total records, topic cardinality at runtime, the number of unique routers and participants, writers/reader counts, QoS class counts, and topic churn relative to previous interval. All measurements are performed under identical software and configuration settings to ensure comparability across different communication links. The use of an application-layer latency probe ensures that the measured values reflect end-to-end communication performance as experienced by the control application.

Throughput Estimation:

Two throughput estimates are logged per interval. Admin-snapshot throughput:

$$B_{\text{admin}} = \frac{\text{bytes}_{\text{dump}}}{\max(\Delta t, 10^{-6})}. \quad (7)$$

All metrics are appended as timestamped CSV rows, yielding a synchronized dataset for latency, topology, QoS, and bandwidth analysis.

UAV Network Performance Characterization

Robust and reliable communication link is essential for controlling UAVs, whether individually or in a swarm. The performance requirements vary depending on the level of autonomy [6]:

- **Manual Control:** The operator sends direct commands (stick input). This requires low latency to ensure responsive handling.
- **Automatic Control:** The UAV follows a pre-programmed flight plan or mission. Real-time operator input is minimal, but telemetry monitoring remains critical.
- **Autonomous Control:** The vehicle makes decisions based on onboard sensors and algorithms. Communication is primarily used for high-level coordination and state updates, but low latency is still important for timely decision making and responsiveness.

Latency requirements depend on the environment and the flight type (Line-of-Sight (LoS) vs. Beyond-Line-of-Sight (BLoS)). For manual control at speeds up to 60 km/h, a maximum latency of 40 ms is recommended by 3GPP [7] as the upper limit for safe operation. Studies indicate that while latencies below 170 ms are manageable, delays between 300 ms-500 ms can significantly degrade operator performance and necessitate active obstacle avoidance systems [8]. For video streaming, latency requirements range from 20 ms-200 ms depending on the application and video quality [7].

Network Benchmarking

This section details the measurement campaign used to evaluate the controllability of modular UAV swarms over various physical networks (Ethernet, WLAN, and 5G). This triad was chosen to systematically isolate the latency bottleneck of the system: Ethernet establishes the hardware and middleware baseline devoid of wireless interference, WLAN serves as a high performance local wireless reference using Layer-2 switching, and private 5G reflects the scalable wide-area cellular communication required for operational BLoS deployments. It should be noted that this comparison highlights architectural differences between local and wide-area communication systems rather than purely physical-layer performance. The measurements were conducted using a Private 5G SA Campus Network at HSU. The network operates in the 3.7 GHz band (n78) with a 100 MHz bandwidth, utilizing Ericsson 5G Core and Radio Access Network (RAN) components. For this study, an enhanced Mobile Broadband (eMBB) slice was used. The testbed follows a client-server topology where the GCS acts as the time synchronization master. Zenoh Router Role: In this benchmark setup, the GCS is configured as the central Zenoh router. It is important to note that for a production-ready, real-world deployment configuring Zenoh for

P2P communication would likely be the preferred approach to further minimize end-to-end latencies. However, for the purpose of this comparative study, a routed topology was chosen to ensure a consistent and reproducible baseline across all network types. While a P2P configuration could reduce absolute latency values, the routed setup was intentionally chosen to ensure reproducibility and comparability across all communication technologies.

- **Time Sync:** Synchronization between UAVs and the GCS is maintained via NTP (chrony [9]), achieving and accuracy of approximately 0.1 ms–2 ms, which is sufficient for latency measurement in the range of tens to hundreds of ms.
- **Data Flow and Middleware:** ROS2 communication is bridged to Zenoh using the *zenoh-bridge-ros2dds*, enabling efficient data transport over UDP-based communication links. It serves as an optimized alternative to the standard Data Distribution Service (DDS) in ROS 2, specially targeting performance improvements over constrained or wireless networks [5].
- **KPI Collection:** A dedicated latency probe publishes JavaScript Object Notation (JSON) payloads (sequence numbers and timestamps) at 20 Hz. One-way latency is calculated as shown in Eq. (2) based on timestamp differences between sender and receiver.

The 5G network is configured to provide different slices with different 5G QoS Indicator (5QI) values, which allows the evaluation of the impact of different QoS settings on the network performance and swarm control metrics. In this study, only eMBB slice was available, which is not optimized for ultra-low latency communication. The KPIs used for the evaluation in this publication were taken from [2] and shown in Tab. 1:

Tab. 1: Private 5G Network Performance (HSU) adapted from [2]

KPIs	Average
Round-Trip Time^a	20.4 ms ^b
Throughput (DL)^c	912 Mbps
Throughput (UL)^c	110 Mbps

^a IRTT [10]

^b with 2 · 1.2 ms latency added from Ethernet

^c iPerf3

Measurement methodology and system architecture:

The testbed follows a two-role topology. On the aerial side, each node consists of a Pixhawk 6X and an NVIDIA Jetson Orin NX. On the ground

side, a dedicated ground station performs KPI collection. Communication is tunneled from ROS 2 through *zenoh-bridge-ros2dds* over Zenoh (UDP) across Ethernet, WLAN, and private 5G links. Before each run, clock synchronization is established using NTP (chrony). In GCS mode, the script acquires Zenoh middleware KPIs via the official Zenoh REST API and stores all outputs in a common measurement directory.

End-to-end latency:

Each UAV publisher emits *std_msgs/String* probe messages at rate f (default 20 Hz) on

$$\text{topic} = \text{TOPIC_BASE}/\text{uav_}\langle id \rangle. \quad (9)$$

The publisher period is defined as

$$T_{\text{pub}} = \frac{1}{\max(hz, 0.1)}, \quad (10)$$

where hz denotes the requested update frequency. Each payload contains a sequence number and sender timestamp t_s . To convert timestamp tuples (sec, nsec) to floating point seconds. A dedicated ROS 2 latency probe publishes probe messages on topic */zenoh_latency_probe*. Each probe payload is a JSON object containing a sequence number and sender timestamp:

$$\text{payload} = \{\text{seq}":n,\text{send_ns}":t_{\text{send}}\}. \quad (12)$$

Samples with $\text{latency}_{\text{ms}} < 0$ or $\text{latency}_{\text{ms}} > 600,000$ are discarded as anomalies (clock violations or extreme outliers). Every N_{report} samples (default $N_{\text{report}} = 100$), percentile KPIs are computed from sorted latency samples a :

$$p_q = a_{\lfloor q \cdot (n-1) \rfloor}, \quad q \in \{0.50, 0.95, 0.99\}, \quad (14)$$

corresponding to median (p_{50}), tail latency (p_{95}), and extreme tail latency (p_{99}), and logged to console for real-time monitoring.

Network topology snapshot KPIs:

Every T seconds (default $T = 5$ s), queries the Zenoh admin space through the official Zenoh REST API and parses the returned JSON records.

Network throughput:

Two complementary throughput estimators are logged per interval. Application-layer proxy from admin dump size:

$$\text{admin_bps} = \frac{B_{\text{dump}}}{\max(\Delta t, 10^{-6})}. \quad (17)$$

Wire-level proxy from UDP capture:

$$\text{udp_bps} = \frac{B_{\text{udp}}}{\max(T_{\text{sample}}, 10^{-6})}. \quad (18)$$

Both throughput metrics, alongside the raw latency timestamps and sequence numbers, are logged locally as comma-separated values

to form the experimental dataset. The offline post-analysis subsequently parses these raw logs to compute the statistical network KPIs. All measurements were conducted under comparable system configurations and identical software setups to ensure a fair comparison between communication technologies. The three latency series are summarized in Fig. 3 to 5. Ethernet is included as a cable-bound reference only and is not considered a deployable link for UAV swarm operation.

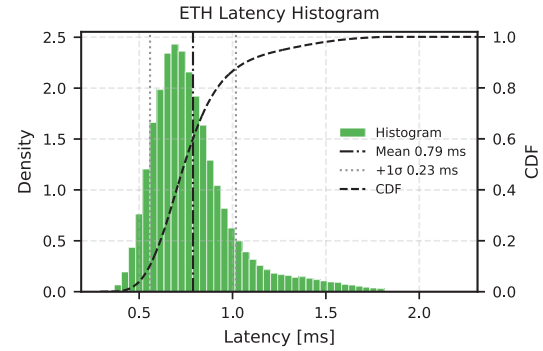


Fig. 4 Normalized Ethernet Latency histogram

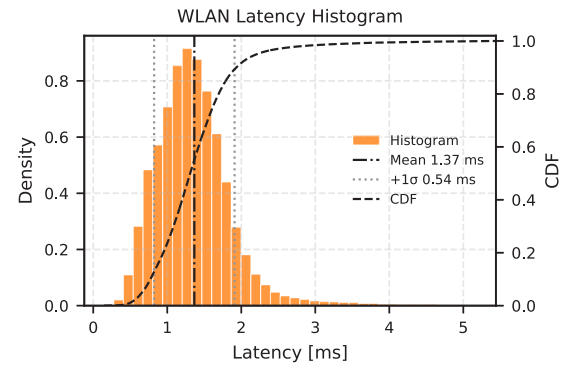


Fig. 5 Normalized WLAN Latency histogram

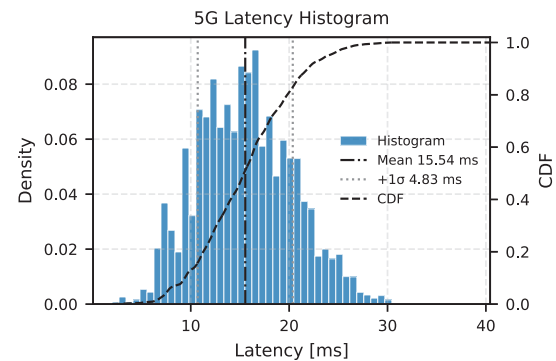


Fig. 6 Normalized 5G Latency histogram

Its measured latency level is lowest (*median* ≈ 0.74 ms, $p_{95} \approx 1.34$ ms), which defines the benchmark floor for the setup. For the relevant wireless comparison, WiFi and 5G show clearly

separated latency ranges. WLAN is centered at lower one-way delay (*median* ≈ 1.31 ms, $p_{95} \approx 2.31$ ms), while 5G exhibits higher central tendency and broader spread (*median* ≈ 15.46 ms, $p_{95} \approx 24.58$ ms). The reported means and standard deviations are shown directly in each histogram. At this stage, this section remains descriptive and does not derive control performance implications. The benchmark and wireless runs had durations of approximately 60 min yielding 10^5 one-way latency samples, respectively. Sequence-gap-based packet-loss estimates over all UAV streams were 0.000% (Ethernet), 0.000% (WLAN), and 0.035% (5G). Notably, this minimal packet loss on the 5G link occurred despite the static setup and optimal Received Signal Strength Indicator (RSSI) / Reference Signal Received Quality (RSRQ) conditions. The minimal packet loss observed exclusively on the 5G link can be attributed to the inherent characteristics of the cellular radio environment combined with the UDP-based transport layer. Unlike the controlled Ethernet baseline and the short-range WLAN environment, 5G signals are subject to transient channel fading. Because the evaluated network slice was configured for eMBB rather than URLLC, the MAC-layer Hybrid Automatic Repeat Request (HARQ) mechanisms possess strict retransmission limits. Once these lower-layer limits are exceeded during brief signal degradation, the packet is discarded. Since Zenoh relies (in this setup) on UDP without transport-layer acknowledgements in this configuration, such discarded frames translate directly into application-level packet loss. Consequently, in dynamic flight scenarios involving handovers and BLoS conditions, this packet loss rate is expected to increase, further emphasizing the need for robust onboard fail-safe mechanisms. All measurements used synchronized one-way probe logging and identical GCS-side KPI extraction; thus, packet-loss values are reported as application-level estimates derived from missing sequence numbers.

Analysis

The latency histograms in Fig. 5 and Fig. 6 reveal a clear separation between the two wireless communication technologies. The 5G latency distribution is centered around a higher median and has a broader spread compared to WLAN. The median latency for 5G is approximately 15.46 ms, while for WLAN it is around 1.31 ms. This corresponds to an approximately 12-fold increase in mean latency for 5G compared to WLAN. This significant difference in latency can be attributed to the architectural characteristics of the respective

communication systems. The architectural divergence is visually summarized in Fig. 6, which contrasts the direct localized data pathways of a WLAN Access Point (AP) against the multi-stage transversal required within the 5G RAN and CN. In contrast to LAN or WLAN, 5G networks behave more like a Wide Area Network (WAN) with higher latency and jitter due to the additional processing and network traversal involved [11]. Furthermore, Time Division Duplex (TDD) patterns, especially using eMBB slices, increase the latency. This pattern (typically 4:1) means that the network is only available for uplink transmission of 20% of the time, which can significantly increase the latency for uplink data. In contrast WLAN uses a completely different Medium Access Control (MAC) and layer-2 switching mechanism, which can provide lower latency under low-load scenarios. Even though the measured latency for 5G is significantly higher than WLAN, it is still within the defined latency requirement for controlling a UAV with direct stick input at a speed of 16.7 ms (60 km/h), which was defined as 40 ms in [5]. This highlights that network-induced delay must be explicitly considered as part of the control loop design rather than treated as a negligible transport effect. Operating an autonomous UAV swarm communicating over wireless networks is limited by the worst-case latency due to the need for low latency and high reliability for control and coordination messages. The higher latency and jitter must be taken into account when designing the control algorithms and operations a UAV swarm is capable of. For example, a swarm that relies on real-time communication for collision avoidance or formation control must be constrained in terms of maximum speed and maneuverability to ensure safe operation. This becomes particularly critical for tightly coupled swarm behaviors such as formation control or synchronized task execution. Algorithms requiring inter-UAV temporal alignment should implement a jitter buffer or use GNSS-time-stamped command activation (execute at $t_s + \Delta$) rather than arrival-triggered execution [12]. Evaluating dedicated low-latency 5G slices (Ultra-Reliable Low-Latency Communication (URLLC)) rather than the eMBB slice used in these campaigns represents a direct next step for reducing the latency and tail variability [13]. Special NR slices with low-latency configurations could potentially reduce the latency, but in this measurement campaign, the only supported eMBB slice was used, which is not optimized for low latency.

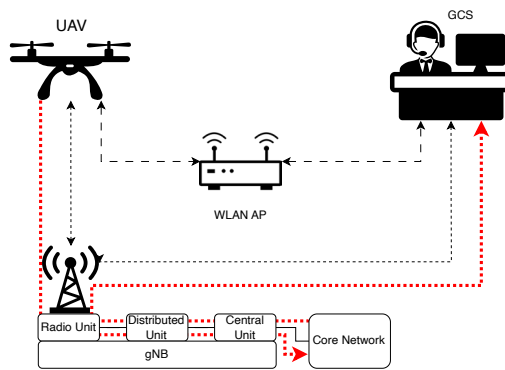


Fig. 7 Simplified comparison of data streams over cellular 5G and WLAN

Tail Latency and Watchdog Design:

Beyond median values, the tail of the latency distribution is critical for safety-critical operation. In the 5G trials, while the median was 15 ms, late arrivals reached over 400 ms. Watchdog mechanisms must reliably distinguish between such delays and actual packet loss. Consequently, system design must be driven by worst-case latency rather than average performance. It is recommended to derive fail-safe thresholds from the 99th percentile ($p_{99} \approx 30\text{ms}$) with an additional safety margin. Crucially, critical failsafe behaviors (e.g., Return-to-Home) must reside onboard the UAV to avoid dependence on delayed GCS commands.

Control Loop Rate Constraints:

A median one-way delay of 15.46 ms implies a minimum round-trip time of 31 ms. This limits the effective GCS-side control loop to approximately 32 Hz. For predictable operation considering the p_{95} values, a schedulable rate of 20 Hz is more realistic. High-frequency safety-critical reactions must therefore be handled locally on the UAV.

Jitter-Aware Command Timing:

The 5G standard deviation of 5.24 ms introduces timing jitter that is non-negligible relative to the median latency. For time sensitive swarm maneuvers — synchronized arming sequences, coordinated trajectory transitions, or timed payload releases — this jitter introduces an inter-UAV temporal uncertainty of $\pm 5\text{ ms}$ per node.

Middleware QoS and Message Scheduling:

The measured 5G application-layer throughput is stable at 0.140 MB/s with a narrow distribution, indicating that bandwidth saturation is not the primary constraint in the observed scenario. Instead, middleware-layer scheduling becomes relevant. ROS 2 via Zenoh bridges provides configurable QoS profiles: reliable (with retransmit acknowledgment) and best-

effort. For safety-critical topics — emergency stop, mode transitions, collision-avoidance triggers — a reliable QoS profile with a retransmit timeout aligned to p_{99} is advisable to avoid silent message loss. Non-critical telemetry, such as periodic GPS state logging, should use best-effort QoS profiles to prevent head-of-line blocking effects that would inflate the latency of critical messages. This separation prevents head-of-line blocking and preserves latency for critical control messages. Furthermore, the Zenoh session and transport layer parameters (scouting intervals, transport unicast/multicast configuration, and ROS 2 domain ID isolation) should be tuned per deployment link: configurations optimized for low-latency WLAN may not perform equivalently over a cellular 5G link with different PDCP and RLC retransmit behavior. Evaluating dedicated low latency 5G slices (URLLC) rather than the eMBB slice used in this campaign represents a direct next step for reducing the observed median latency and tail variability.

Latency Optimization and Deployment Considerations:

The analysis of 5G performance (median latency $\approx 15.46\text{ ms}$) vs. WLAN ($\approx 1.31\text{ ms}$) shows a clear 12-fold difference. As discussed in the methodology, the use of a central router at the GCS adds a specific processing overhead. While moving to a Zenoh P2P configuration could reduce the absolute latency values, the fundamental bottlenecks of the 5G eMBB slice — specifically the TDD pattern would remain the dominant factor for the 5G lag. Therefore, the relative performance gap between WLAN (Layer-2 switching) and 5G (WAN-like traversal) is expected to be robust against middleware-level topology changes.

Conclusion on Network Selection:

Although this campaign reports lower latency for WLAN under local test conditions, private 5G remains necessary for operational UAV swarms once deployment extends beyond confined, interference-controlled environments. Compared to WLAN, private 5G provides managed spectrum access, mobility support with handover, network slicing, and enforceable QoS policies over larger service areas, which are essential for predictable command-and-control performance under multi-user load. Accordingly, WLAN should be treated as a high-performance short-range link, while private 5G represents the scalable wide-area control infrastructure. In addition, when swarm operation can leverage public cellular infrastructure, the mission-level communication range is theoretically unbounded (i.e., constrained by network coverage rather than hotspot footprint), which is fundamentally not

achievable with standalone WLAN architectures. Therefore, the use of private 5G is justified not because it always minimizes latency, but because it enables controlled, secure, and geographically scalable swarm operations with a practical pathway toward beyond-visual-line-of-sight deployment.

Conclusion

Although WLAN provided lower latency in local tests ($p_{95} \approx 2.31$ ms), Private 5G ($p_{95} \approx 24.58$ ms) is the superior choice for operational swarms due to its managed spectrum, mobility support (handover), and scalability over large areas. While WLAN serves as a high-performance short-range link, 5G enables geographically unbounded operations. The use of private 5G is justified not by minimizing latency in all cases, but by providing a secure, scalable, and professional infrastructure BLoS deployments.

Acknowledgement

This work is a part of the “Digital Sensor-2-Cloud Campus Platform” (DS2CCP) project, which is funded by the Federal Ministry of Defense under the dtcc.bw program. Project website:

<https://dtccbw.de/home/forschung/hsu/projekt-ds2ccp/projekt-ds2ccp>

References

- [1] R. Pommeranz, K. Tebbe, R. Heynicke, and G. Scholl, “A Modular and Scalable System Architecture for Heterogeneous UAV Swarms Using ROS 2 and PX4-Autopilot,” Oct. 31, 2025. doi: 10.48550/arXiv.2510.27327.
- [2] H. Beuster, K. Tebbe, T. R. Doebbert, and G. Scholl, “Measurements of the Safety Function Response Time on a Private 5G and IO-Link Wireless Testbed,” in *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, Sep. 2024, pp. 1–4. doi: 10.1109/ETFA61755.2024.10710762.
- [3] D. L. Mills, “Internet time synchronization: the network time protocol,” *IEEE Trans. Commun.*, vol. 39, no. 10, pp. 1482–1493, Oct. 1991, doi: 10.1109/26.103043.
- [4] *IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems*. doi: 10.1109/IEEESTD.2008.4579760.
- [5] *eclipse-zenoh/zenoh*. (Apr. 15, 2026). Rust. Eclipse zenoh. Accessed: Apr. 15, 2026. [Online]. Available: <https://github.com/eclipse-zenoh/zenoh>
- [6] Luftfahrt Bundesamt (LBA), “Operations Manual / OM Betriebshandbuch für den Betrieb von unbemannten Flugsystemen (UAS).” Jan. 28, 2026.
- [7] ETSI, “Unmanned Aerial System (UAS) support in 3GPP (3GPP TS 22.125 version 17.6.0 Release 17).” Apr. 2022. [Online]. Available: https://www.etsi.org/deliver/etsi_ts/122100_122199/122125/17.06.00_60/ts_122125v170600p.pdf
- [8] T. Zhang, “Toward Automated Vehicle Teleoperation: Vision, Opportunities, and Challenges,” *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11347–11354, Dec. 2020, doi: 10.1109/JIOT.2020.3028766.
- [9] “chrony – Introduction.” Accessed: Apr. 15, 2026. [Online]. Available: <https://chrony-project.org/>
- [10] P. Heist, *IRTT*. (May 15, 2021). [Online]. Available: <https://github.com/heistp/irtt>
- [11] M. Seliem, D. Pesch, U. Roedig, and C. J. Sreenan, “Comparative Analysis of 5G and Wi-Fi Integration with TSN for Industrial Applications,” in *2025 IEEE Wireless Communications and Networking Conference (WCNC)*, Milan, Italy: IEEE, Mar. 2025, pp. 1–6. doi: 10.1109/WCNC61545.2025.10978747.
- [12] R. Imai and R. Kubo, “Introducing jitter buffers in networked control systems with communication disturbance observer under time-varying communication delays,” in *IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society*, Yokohama: IEEE, Nov. 2015, pp. 002956–002961. doi: 10.1109/IECON.2015.7392552.
- [13] L. Guijarro, J.-R. Vidal, and V. Pla, “Competitive advantage of URLLC vs. eMBB for supporting timeliness-relevant services,” Apr. 26, 2024, *arXiv: arXiv:2402.03922*. doi: 10.48550/arXiv.2402.03922.

Neuartiger Sensor zur Messung der Schmierfilm-Eigenschaften im Zahnradkontakt

Christoph Schneidhofer¹, Michele Schirru¹, Nico Bernegger¹

¹AC2T research GmbH, Viktor Kaplan-Straße 2/c, 2700 Wiener Neustadt, Österreich

Zusammenfassung

Die Kenntnis des Schmierungs Zustands eines Zahnradkontaktes ist entscheidend für die Auslegung von Getrieben. In dieser Studie wird ein neuartiger Sensor mit zugehöriger Steuerelektronik vorgestellt, der die in-situ- und Echtzeitmessung der Schmierfilmdicke in rotierenden Kontaktsituationen ermöglicht. Hier ist in Getriebenanwendungen der Zweischeibenversuch eine Modellmethode zur Charakterisierung von Schmierstoffen, da er entscheidende Einblicke in Reibung, Verschleiß und den Schmierzustand liefert. Um dies zu beobachten wurde das neuartige Sensorsystem direkt in eine der Scheiben integriert, sodass bei jeder Umdrehung Informationen über den Schmierungs Zustand im Kontaktpunkt gewonnen werden können. Dieser Durchbruch in der Messtechnik ermöglicht eine bisher unerreichte Auflösung bei der dynamischen Erfassung der Schmierungs Zustände unter realistischen Belastungs- und Drehzahlbedingungen, wie sie in Getriebekontakten auftreten. Durch die direkte Verknüpfung tribologischer Leistungsmerkmale mit der momentanen Filmdicke bietet diese Technologie ein leistungsstarkes Werkzeug zur Optimierung von Schmierstoffformulierungen und zur Getriebeauslegung und überbrückt die Lücke zwischen Laborversuchen und realem Getriebebetrieb.

Diese Arbeit beschreibt die neuartige Messelektronik, die mittels der aktiven ultraschallbasierten Reflexionsmethode die Filmdicke im Kontaktpunkt des Sensors zugänglich macht. Darüber hinaus nutzt die Elektronik Energy-Harvesting, um sich selbst mit Energie zu versorgen, sowie eine kabellose Datenübertragung über WLAN. Die gesamte Messkette wird anhand von Messungen am Zweischeiben-tribometer demonstriert.

Keywords: Filmdickenmessung, Ultraschall- Reflexionsmethode, Tribologie, Energy Harvesting, Getriebeanwendung

Einleitung

Die Schmierung von Zahnradkontakten stellt einen entscheidenden Einflussfaktor für die Effizienz, Lebensdauer und Zuverlässigkeit von Getrieben dar. Insbesondere die Schmierfilmdicke bestimmt maßgeblich den Betriebszustand im Kontakt und erlaubt die Unterscheidung zwischen Grenz-, Misch- und Vollschmierung. Eine präzise Kenntnis dieser Größe ist daher von zentraler Bedeutung für die Auslegung von Getrieben sowie für die Entwicklung und Bewertung von Schmierstoffen. [1], [2], [3]

Zur experimentellen Untersuchung tribologischer Systeme mit hochbelasteten Kontakten hat sich der Zweischeibenversuch als etablierte Modellmethode bewährt. Er ermöglicht die gezielte Analyse von Reibung, Verschleiß und Schmierungs Zuständen unter definierten Belastungs- und Geschwindigkeitsbedingungen. [4] Die Bestimmung der aktuellen Schmierfilmdicke ist in der Realität sehr schwierig und kann mittels optischer,

elektrischer oder akustischer Messverfahren erfolgen. [2], [5], [6] Diese Ansätze sind jedoch häufig auf stationäre oder vereinfachte Randbedingungen beschränkt und erlauben nur eine begrenzte zeitliche und örtliche Auflösung. Insbesondere die direkte, *in-situ* Messung der Filmdicke in einem realitätsnah rotierenden Kontakt stellt weiterhin eine messtechnische Herausforderung dar. [3]

Vor diesem Hintergrund wird in der vorliegenden Arbeit ein neuartiges Sensorsystem vorgestellt, das die Echtzeitmessung der Schmierfilmdicke direkt im Kontaktpunkt eines Zweischeibentribometers ermöglicht. Hierzu wird ein Sensor in eine rotierende Scheibe integriert, wodurch bei jeder Umdrehung Messdaten aus dem Kontakt gewonnen werden können. Die Bestimmung der Filmdicke erfolgt mittels einer aktiven ultraschallbasierten Methode [1], [7], [8]. Ergänzend ist das System mit einer energieautarken Versorgung durch Energy Harvesting sowie einer kabellosen Datenübertragung ausgestattet.

Ziel der Arbeit ist die Entwicklung und experimentelle Validierung dieser Messkette unter realistischen Betriebsbedingungen. Durch die Kombination aus integrierter Sensorik, energieautarker Elektronik und drahtloser Kommunikation wird ein neuer Ansatz zur hochauflösenden Erfassung dynamischer Schmierzustände geschaffen, der die Lücke zwischen Laboruntersuchungen und realem Getriebebetrieb schließt.

Zunächst wird das Konzept des Energy Harvestings vorgestellt, gefolgt von der Beschreibung der entwickelten Messelektronik. Anschließend werden der Aufbau und die Versuchsanordnung des Zweischeibentribometers erläutert sowie die experimentellen Ergebnisse der Filmdickenmessung diskutiert.

Energy-Harvesting

Bei der vorliegenden Fragestellung zur Messung der Filmdicke beim Zweischeibentribometer befinden sich beide Kontaktpartner des Tribokontaktes in Bewegung, wodurch keine direkte elektrische Verbindung des Ultraschallsensors zu einer stationären Messelektronik möglich ist. Demnach ist die Überführung der Messelektronik auf einen rotierenden Teil zielführend.

In den vergangenen Jahren haben sich Energy-Harvesting Technologien, insbesondere im Bereich der Mechatronik und Sensorik, schnell weiterentwickelt [9]. Dabei rückt die Fähigkeit, Energie direkt aus der Umgebung zu gewinnen und für den Betrieb autonomer Systeme nutzbar zu machen, zunehmend in den Fokus, vor allem in Szenarien, in denen konventionelle Energiequellen schwer zugänglich oder unpraktisch sind. In Abb. 1 werden verschiedene Technologien und ihre Anwendungen verglichen.

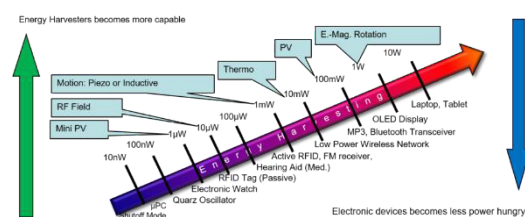


Abb. 1: Vergleich von Energy-Harvesting Technologien und Anwendungen [10]

Für die vorliegende Messanordnung wurde für das Energy-Harvesting ein elektromagnetisches Prinzip gewählt. Elektromagnetisches

Energy-Harvesting basiert auf dem Prinzip der Induktion. Das Generatorprinzip basiert auf der Bewegung von Magnetfeldern relativ zu einer Spule. Bewegt sich eine Spule durch ein Magnetfeld, oder auch umgekehrt, wird in der Spule eine elektrische Spannung induziert. Ein praktisches Beispiel ist ein rotierender Sensor, neben dem ein Magnet angebracht ist. Durch die Drehbewegung wird in der Spule auf dem Sensor elektrische Energie erzeugt. Dieser Ansatz bietet sich beispielsweise in Maschinen oder Fahrzeugen an, in denen mechanische Bewegung ohnehin vorhanden ist, wie es in der vorliegenden Aufgabenstellung beim Zweischeibentribometer der Fall ist. [11], [12]

Bei dem hier angewendeten elektromagnetischen Energy-Harvesting-System wird aus der Rotationsbewegung der Maschine die elektrische Energie für die Spannungsversorgung zuverlässig sichergestellt. Dabei sind eine definierte Anzahl an Spulen ringförmig auf der Messplatine angeordnet, die sich bei der Rotationsbewegung an feststehenden Magneten vorbeibewegen und somit durch Induktion eine Spannung induziert wird, die nachfolgend für die Power Supply aufbereitet wird.

Messelektronik

Abb. 2 zeigt das Blockschaftbild der entwickelten Messelektronik zur Bestimmung der Schmierfilmdicke bei der rotierenden Kontaktstelle im Zweischeibentribometer. Die Energieversorgung der Messelektronik wird mit dem zuvor beschriebenen Energy-Harvester auf elektromagnetischem Prinzip bewerkstelligt.

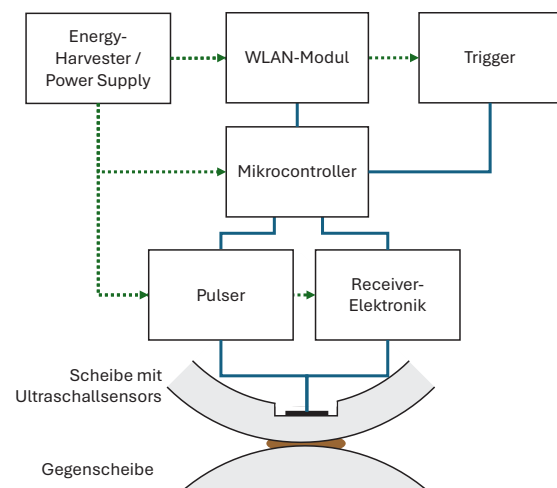


Abb. 2: Blockschaftbild der Messelektronik

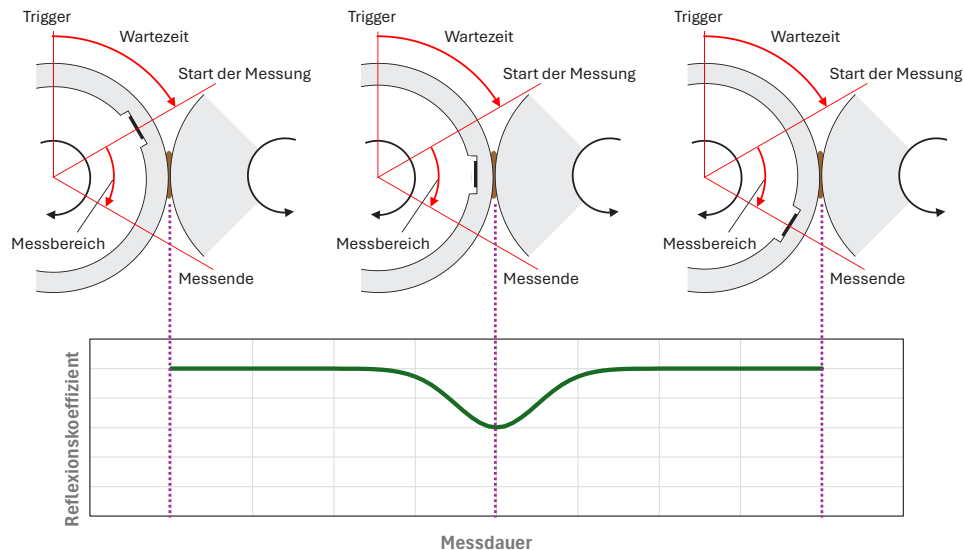


Abb. 3: Schematische Darstellung des Messablaufs und des resultierenden Messsignals anhand des Reflexionskoeffizienten

Für die Ablaufsteuerung wird ein Mikrocontroller verwendet, der die Anregung des Ultraschallsensors durch den Sender ansteuert. Die Reflexion des Ultraschallpulses, welche die Information der Schmierfilmdicke enthält, wird von der Empfänger-Elektronik verarbeitet und die nötige Amplituden-Information der Ultraschallreflexion extrahiert und an den Mikrocontroller zurückgesendet.

Bei jeder Umdrehung wird ein Triggersignal ausgelöst, um die Position der Scheibe in Bezug zur Kontaktstelle zu erkennen. Basierend auf dem Triggersignal wird der zeitliche Ablauf der Messung gesteuert. In Abb. 3 ist der Messablauf und das dazu resultierende Messsignal anhand des Reflexionskoeffizienten dargestellt. Befindet sich der Sensor vor der Kontaktstelle, wie in Abb. 3 links dargestellt, so wird sämtliche Energie des Ultraschallpulses an der Scheibenoberfläche reflektiert, wodurch der Reflexionskoeffizient 1 als Referenz gemessen wird. Bewegt sich der Ultraschallsensor in die Kontaktzone (Abb. 3 Mitte), wird die Ultraschallreflexion entsprechend der Kontaktsteifigkeit, welche unter anderem die Information der Schmierfilmdicke beinhaltet, abgeschwächt. Verlässt der Ultraschallsensor wieder die Kontaktzone (Abb. 3 rechts), geht der gemessene Reflexionskoeffizient zum Referenzwert 1 zurück.

Die Amplitudeninformation zur Bestimmung des Reflexionskoeffizienten wird während des gesamten Messablaufs vom Mikrocontroller aufgezeichnet. Nach abgeschlossener Messung werden die Messdaten über das WLAN-Modul zu einem PC zur Messdatenspeicherung und nachfolgenden Auswertung übermittelt.

Zweischeibentribometer – Aufbau und Versuchsanordnung

Ein Zweischeibentribometer ist eine spezielle Bauform eines Tribometers zur Untersuchung von Reibung und Verschleiß zwischen zwei rotierenden Scheiben. Diese Konfiguration ermöglicht die Nachbildung von hochbelasteten Roll- und Gleitkontakten unter realistischen Bedingungen, wie sie beispielsweise bei Lagern und Getrieben vorkommen. Durch die Relativbewegung und (leicht) unterschiedliche Drehzahlen entstehen definierte Schlupfbedingungen in der Kontaktzone. Der Anpressdruck, die Relativgeschwindigkeit und die Schmierbedingungen können dabei eingestellt und variiert werden. Mit einem Zweischeibentribometer lassen sich Verschleißraten, Reibkoeffizienten sowie Materialübertragungen unter unterschiedlichen Betriebsbedingungen experimentell ermitteln. Je nach Scheibengeometrie und Anordnung lassen sich Punkt- bzw. Linienkontakte nachstellen. [13], [14], [15]

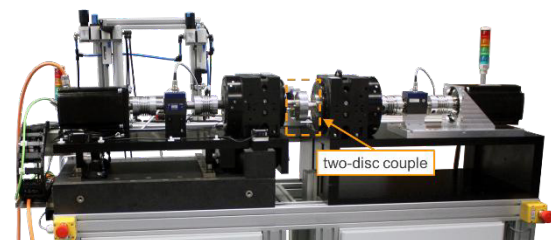


Abb. 4: Zweischeibentribometer bei AC2T research GmbH

Abb. 4 zeigt den Zweiseibentribometer, der für diese Arbeit herangezogen wurde. Er besteht aus zwei unabhängig voneinander angetriebenen Achsen 1 und 2, die jeweils Scheiben mit unterschiedlichen Durchmessern aufnehmen können. In der vorliegenden Arbeit wurde beide Scheiben aus Edelstahl 1.4301 gefertigt. Die Scheibe 1 mit einem Durchmesser von 140 mm und eine Breite von 35 mm hat im Inneren eine Ausnehmung zur Aufnahme der Messelektronik. Abb. 5 zeigt die Scheibe mit montierter Messelektronik. Die Scheibe hat 4 innenliegende Nuten auf dem Umfang verteilt zur Montage der Ultraschallsensoren, wobei für diese Arbeit nur ein Sensor montiert und verwendet wurde. Die Gegenscheibe (Scheibe 2) (nicht dargestellt) hat einen Durchmesser von 100 mm und eine Breite von 4 mm an der Kontaktstelle, wobei diese mit einem Radius von 10 mm versehen ist, um eine punktförmige Kontaktsituation zu erreichen.



Abb. 5: Prototyp der Messelektronik montiert auf Scheibe 1 für die Versuche im Zweiseibentribometer

Für die Evaluierung der Filmdickenmessung wurde die Scheibe 1 mit einer Drehzahl von 700 U/min angetrieben und mit einer Kraft von 700 N gegen die Scheibe 2 gepresst. Dadurch entsteht eine Hertzsche Kontaktpressung von 1,8 GPa. Als Schmierung wurde ein Schmierfett der Konsistenzklasse NLGI 2 mit Lithiumverdicker und mineralischem Basisöl mit einer Viskosität von 100 mm²/s bei 40°C verwendet. Eine definierte Menge des Schmierfettes wurde initial aufgebracht und während dem Versuch wurde nicht nachgeschmiert.

In dieser Konfiguration wurde das Experiment für eine Dauer von 6 Minuten durchgeführt, wobei jede Minute eine Messung der Schmierfilmdicke für ein Dauer von 4 Umdrehungen durchgeführt wurde.

Ergebnisse

Abb. 6 zeigt exemplarisch den gemessenen Trend des Reflexionskoeffizient während einer kompletten Messung bestehend aus dem Startimpuls gefolgt von 200 Messwerten des Referenzwertes. An diesen angeschlossen ist die Messung des Reflexionskoeffizienten während vier aufeinanderfolgenden Kontaktzyklen wie in Abb. 6 ersichtlich ist. Jeder Kontaktzyklus besteht aus 400 Messwerten. In der Mitte jedes Zyklus ist der charakteristische Einbruch im Reflexionskoeffizient zu erkennen da sich hier der Sensor genau an der Kontaktzone befindet.

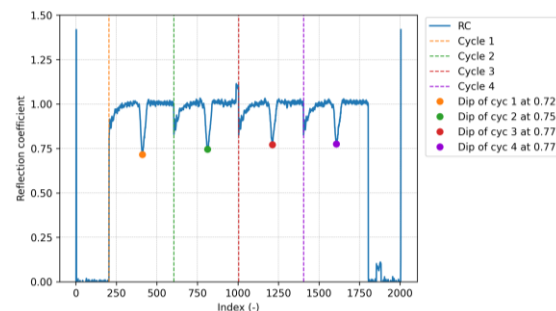


Abb. 6: Entwicklung des Reflexionskoeffizienten während einer Messung mit 4 Zyklen

Der Reflexionskoeffizient in der Kontaktzone liefert die Information der Steifigkeit des Kontaktes, woraus die Schmierfilmdicke errechnet werden kann. Anschließend an die vier Messzyklen wird erneut der Referenzwert bestehend aus 200 Messwerten gemessen und mit dem Stoppimpuls die Messung beendet. Wie der Abbildung zu entnehmen ist, zeigen die vier aufeinanderfolgenden Zyklen lediglich eine geringe Variation im gemessenen Reflexionskoeffizient.

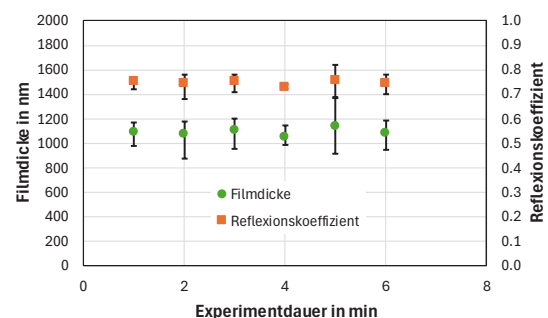


Abb. 7: Trend der berechneten Schmierfilmdicke sowie des gemessenen Reflexionskoeffizient während des Experiments

Abb. 7 zeigt den Trend der berechneten Schmierfilmdicke sowie des gemessenen Reflexionskoeffizient während des Experiments. Die Messung zeigt, dass über den Zeitraum des Experimentes eine stabile Messung des Reflexionskoeffizienten durchgeführt werden kann und hieraus ein stabiler Wert für die Schmierfilmdicke berechnet werden kann.

Zusammenfassung und Schlussfolgerung

Ein neuartiger Sensor mit zugehöriger Steuer-elektronik zur *in-situ* und Echtzeitmessung der Schmierfilmdicke in rotierenden Kontakt-situationen wurde entwickelt und aufgebaut. Hierbei wird die Reflexionsmethode eines Ultraschallimpulses genutzt, um die Filmdicke im Kontaktpunkt des Sensors zugänglich zu machen. Ein Energy-Harvester zur Spannungsversorgung sowie eine Datenübertragung über WLAN ermöglichen die autarke Messung der Schmierfilmdicke bei rotierenden Probenkörpern. Die gesamte Messkette wurde anhand von Messungen am Zweischeiben-tribometer demonstriert. Hierbei wurden über die Experimentdauer stabile Signale beobachtet, die eine Bestimmung der Filmdicke in der Kontaktzone ermöglichen.

Die durchgeführten Experimente zeigen detaillierte Einblicke in den Schmierungs-zustand im Kontaktpunkt. Dadurch ist die dynamische *in-situ* Erfassung des Schmierungs-zustandes unter realistischen Belastungs- und Drehzahlbedingungen, wie sie in Getriebekontakten auftreten, möglich.

Danksagung

Die präsentierten Forschungsergebnisse wurden im Rahmen des COMET-Zentrums InTribology (FFG-Nr. 906860), einem Projekt des „Kompetenzzentrums für Tribologie“ (AC2T research GmbH), erarbeitet. InTribology wird durch das COMET – Competence Centers for Excellent Technologies Programm von den Bundesministerien BMIMI und BMWET als auch den Bundesländern Niederösterreich und Vorarlberg auf Basis der finanziellen Unterstützung durch die beteiligten Projektpartner gefördert. **COMET wird durch die Österreichische Forschungsförderungsgesellschaft (FFG) abgewickelt.** Die Entwicklung der Messmethode wurde durch das Interreg-Projekt IPMAI - Interpretierbare präskriptive Instandhaltung mit künstlicher Intelligenz (ATCZ00060) unterstützt.

Literaturnachweis

- [1] M. Schirru und M. Varga, „A Review of Ultrasonic Reflectometry for the Physical Characterization of Lubricated Tribological Contacts: History, Methods, Devices, and Technological Trends“, 1. Dezember 2022, *Springer*. doi: 10.1007/s11249-022-01670-8.
- [2] P. Dou u. a., „A finite-element-aided ultrasonic method for measuring central oil-film thickness in a roller-raceway tribo-pair“, *Friction*, Bd. 10, Nr. 6, S. 944–962, Juni 2022, doi: 10.1007/s40544-021-0544-y.
- [3] P. Dou u. a., „Review of ultrasonic-based technology for oil film thickness measurement in lubrication“, *Tribol. Int.*, Bd. 165, S. 107290, Jän. 2022, doi: 10.1016/J.TRIBOINT.2021.107290.
- [4] G. Amine, N. Fillot, D. Philippon, N. Devaux, J. Dufils, und E. Macron, „Dual experimental-numerical study of oil film thickness and friction in a wide elliptical TEHL contact: From pure rolling to opposite sliding“, *Tribol. Int.*, Bd. 184, S. 108466, Juni 2023, doi: 10.1016/J.TRIBOINT.2023.108466.
- [5] O. Marklund und L. Gustafsson, „Interferometry-based measurements of oil-film thickness“, *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology*, Bd. 215, Nr. 3, S. 243–259, März 2001, doi: 10.1243/1350650011543510.
- [6] N. Marx, J. Guegan, und H. A. Spikes, „Elastohydrodynamic film thickness of soft EHL contacts using optical interferometry“, *Tribol. Int.*, Bd. 99, S. 267–277, Juli 2016, doi: 10.1016/J.TRIBOINT.2016.03.020.
- [7] W. A. Gray und R. S. Dwyer-Joyce, „Ultrasonic resonance approach to measure rolling element bearing raceway films in-situ: A practical approach“, *Tribol. Int.*, Bd. 214, S. 111141, Feb. 2026, doi: 10.1016/J.TRIBOINT.2025.111141.
- [8] R. S. Dwyer-Joyce, T. Reddyhoff, und J. Zhu, „Ultrasonic Measurement for Film Thickness and Solid Contact in Elastohydrodynamic Lubrication“, *J. Tribol.*, Bd. 133, Nr. 3, Juni 2011, doi: 10.1115/1.4004105.
- [9] M. Chetto und A. Queudet, „3 - Harnessing Ambient Energy for Embedded Systems“, in *Energy Autonomy of Real-Time Systems*, M. Chetto und A. Queudet, Hrsg., Elsevier, 2016, S. 57–83. doi: <https://doi.org/10.1016/B978-1-78548-125-3.50003-8>.
- [10] M. Hayes, „Synergies between Energy Harvesting and Power Electronics“, in *IEEE PELS Young Professionals Webinar*, IEEE Power Electronics Society, Okt. 2016, S. 33.
- [11] H. Fu u. a., „Rotational energy harvesting for self-powered sensing“, *Joule*, Bd. 5, Nr. 5, S. 1074–1118, Mai 2021, doi: 10.1016/J.JOULE.2021.03.006.

- [12] D. Dinulovic, M. Brooks, M. Haug, und T. Petrovic, „Rotational Electromagnetic Energy Harvesting System“, *Phys. Procedia*, Bd. 75, S. 1244–1251, Jän. 2015, doi: 10.1016/J.PHPRO.2015.12.137.
- [13] S. K. Peesapati, J. Prost, G. Vorlaufer, M. Varga, und C. Gachot, „Chatter Detection and Identification Based on Mode Decompositions“, *Adv. Eng. Mater.*, Bd. 27, Nr. 23, Dez. 2025, doi: 10.1002/adem.202500382.
- [14] A. M. Puhwein, A. Biederbeck, B. Jakab, und M. Varga, „Experimental Modal Analysis of a Twin-disc Tribometer using a 3D-scanning Laser Doppler Vibrometer“, in *Computer Vision & Laser Vibrometry, Vol. 6*, Denmark: River Publishers, 2025, S. 103–110. doi: 10.13052/97887-438-0151-1_12.
- [15] G. Brunbauer u. a., „Finite element modal analysis of a twin-disc tribometer: Sensitivity analysis and experimental validation using vibrometry“, *Tribol. Int.*, Bd. 219, S. 111778, Juli 2026, doi: 10.1016/J.TRIBOINT.2026.111778.

Ein neuartiger Sensor für Öl-Condition Monitoring in extremen Umgebungsbedingungen

Michele Schirru¹

¹AC2T research GmbH, Viktor-Kaplan-Straße 2/C 2700 Wiener Neustadt, Österreich
Michele.schirru@ac2t.at

Zusammenfassung

Die Zustandsüberwachung von Schmierstoffen in kritischen mechanischen Systemen ist von entscheidender Bedeutung für die Gewährleistung von Zuverlässigkeit, Effizienz und Sicherheit. Unter den Schmierstoffeigenschaften stellt die Viskosität einen zentralen Indikator für Alterungsprozesse wie Oxidation, Kontamination und Gasabsorption dar. In dieser Arbeit wird ein neuartiges Ultraschallviskosimeter vorgestellt, das Echtzeit- und *In-situ*-Viskositätsmessungen unter extremen Bedingungen, einschließlich hoher Drücke und Temperaturen, ermöglicht. Der Sensor nutzt hochfrequente transversale Ultraschallwellen (Scherwellen) sowie eine innovative polymere Anpassungsschicht zur Erhöhung der Sensitivität durch Reduktion von akustischen Impedanzunterschieden. Der vorgestellte Ansatz ermöglicht eine präzise Überwachung über einen breiten Viskositätsbereich und zeigt eine hohe Robustheit in anspruchsvollen Umgebungen, was durch Messungen von Gasabsorptionsprozessen in unter Druck stehenden ionischen Flüssigkeiten validiert wurde.

Keywords: Viskosität, Rheologie, Hochdruck, Zustandsüberwachung

Einleitung

Die Schmierung spielt eine fundamentale Rolle im Betrieb industrieller Systeme wie Getriebe, Verdichter, Turbinen und hydraulische Anlagen. Die Leistungsfähigkeit und Lebensdauer dieser Systeme hängen maßgeblich vom Zustand des Schmierstoffs ab, der verschiedenen Degradationsmechanismen wie Oxidation, thermischer Zersetzung, Kontamination und Gasabsorption unterliegt.

Unter den physikochemischen Eigenschaften von Schmierstoffen gilt die Viskosität als einer der wichtigsten Indikatoren für den Schmierstoffzustand, da sie direkt die Schmierfilmbildung, Reibung und Verschleiß beeinflusst.

Änderungen der Viskosität können als Frühindikatoren für Systemstörungen dienen. Beispielsweise führt Oxidation typischerweise zu einer Erhöhung der Viskosität, während Kontamination oder Kraftstoffverdünnung eine Verringerung bewirken können. Auch die Gasabsorption, insbesondere unter Hochdruckbedingungen, kann die Viskosität erheblich beeinflussen und somit die Schmierleistung in Anwendungen wie Verdichtern oder Druckreaktoren verändern.

Trotz ihrer großen Bedeutung stellt die zuverlässige *In-situ*-Messung der Viskosität unter anspruchsvollen Umgebungsbedingungen mit hohen Drücken und Temperaturen eine erhebliche Herausforderung dar. Konventionelle Laborviskosimeter liefern zwar präzise Ergebnisse, sind jedoch für die kontinuierliche Überwachung ungeeignet. Bestehende Inline-Sensortechnologien, wie kapazitive oder induktive Sensoren, weisen häufig eine begrenzte Empfindlichkeit auf, insbesondere bei Medien mit niedriger Permittivität oder bei geringen Änderungen der Schmierstoffzusammensetzung. Zudem sind diese Sensoren in der Regel nicht für den Betrieb unter extremen Bedingungen wie hohem Druck, hohen Temperaturen oder chemisch aggressiven Medien ausgelegt [1].

Ultraschallbasierte Verfahren haben sich als vielversprechende Alternativen zur Charakterisierung von Fluideigenschaften etabliert [2]. Allerdings sind klassische Ultraschall-Reflexionsmethoden häufig durch starke akustische Impedanzunterschiede zwischen festen Sensormaterialien (z. B. Metallen) und Flüssigkeiten limitiert [3]. Diese führt zu schwachen Transmissions- und Reflexionssignalen und somit zu einer reduzierten Messgenauigkeit und Sensitivität.

Zur Überwindung dieser Herausforderungen wird in dieser Arbeit ein neuartiges, auf Ultraschallreflexion basierendes Viskosimeter vorgestellt, das Scherwellen sowie eine gezielt ausgelegte Anpassungsschicht nutzt, um die Messsensitivität und Robustheit signifikant zu verbessern [4]. Der vorgeschlagene Sensor ist speziell für den Einsatz in anspruchsvollen Umgebungen konzipiert und ermöglicht eine präzise Echtzeitüberwachung der Viskosität in Anwendungen, in denen konventionelle Sensoren versagen.

Materialien

Der entwickelte Sensor basiert auf einem aktiven Ultraschall-Reflexionsmessprinzip und ist speziell für den Einsatz in anspruchsvollen industriellen Umgebungen ausgelegt. Die Hauptkomponenten und die Funktionsweise sind schematisch in Abbildung 1 dargestellt.

Die Sensorarchitektur umfasst folgende wesentliche Komponenten: einen piezoelektrischen Scherwandler, eine Verzögerungsstrecke (Delay Line) sowie eine maßgeschneiderte polymere Anpassungsschicht.

Der piezoelektrische Wandler arbeitet im Schermodus in einem Frequenzbereich von etwa 1 bis 20 MHz. Scherwellen sind besonders geeignet für Viskositätsmessungen, da ihre Wechselwirkung mit Flüssigkeiten direkt mit den viskosen Eigenschaften des Fluids verknüpft ist. Der Wandler erzeugt hochfrequente Ultraschallwellen, die sich durch die Sensorstruktur bis zur Grenzfläche mit der Prüfflüssigkeit ausbreiten.

Zum Schutz des Wandlers vor direkter Exposition gegenüber der anspruchsvollen Betriebsumgebung ist dieser mit einer Verzögerungsstrecke verbunden. Diese erfüllt mehrere Funktionen: mechanischer Schutz, kontrollierte Wellenausbreitung sowie die Definition einer klar abgegrenzten Wechselwirkungszone zwischen Ultraschallsignal und Messmedium. Darüber hinaus ist die Verzögerungsstrecke typischerweise mit einem Gewinde versehen, um eine direkte Integration in mechanische Systeme wie Ölsümpfe oder Druckbehälter zu ermöglichen.

Eine zentrale Innovation des Sensors ist die Einführung einer polymerbasierten Anpassungsschicht zwischen Verzögerungsstrecke und Prüffluid [5]. Diese Schicht reduziert gezielt die akustische Impedanzfehlانpassung zwischen der festen

Sensorstruktur und der Flüssigkeit. Durch eine geeignete Wahl der Materialeigenschaften und eine Schichtdicke von einem Viertel der Ultraschallwellenlänge wird eine Resonanzbedingung zwischen einfallenden und reflektierten Scherwellen erzeugt [6].

Diese Resonanz führt zu einer signifikanten Erhöhung der Sensitivität gegenüber Änderungen der mechanischen Eigenschaften des Fluids, insbesondere der Viskosität. Im Gegensatz zu konventionellen Ultraschallsensoren ermöglicht dieses Design eine starke Kopplung zwischen Ultraschallwelle und Flüssigkeit, selbst bei niedrigviskosen oder niedrigpermittiven Medien.

Die polymere Anpassungsschicht ist zudem für den Einsatz unter extremen Bedingungen wie erhöhten Temperaturen und hohen Drücken ausgelegt und gewährleistet eine stabile Langzeitperformance.

Der Sensor ist in einem industrietauglichen Gehäuse mit entsprechendem Schutzgrad integriert und kann in realen Anwendungen wie Getrieben, Verdichtern und Druckreaktoren eingesetzt werden. Eine elektronische Steuereinheit übernimmt die Ansteuerung des Wandlers, die Signalerfassung sowie die Echtzeit-Datenverarbeitung.

Datenanalyse und Sensorkalibrierung

Das Messprinzip des vorgeschlagenen Sensors basiert auf der Analyse des Ultraschall-Reflexionskoeffizienten an der Grenzfläche zwischen Anpassungsschicht und Prüffluid. Trifft eine transversale Ultraschallwelle auf diese Grenzfläche, wird ein Teil der Welle reflektiert, während ein anderer Teil im Fluid gedämpft wird.

Die Elektronik wandelt das vom piezoelektrischen Wandler empfangene Zeitsignal in den Frequenzbereich um. Anschließend wird die Amplitude bei der Resonanzfrequenz der resonanten Anpassungsschicht zur Berechnung des Reflexionskoeffizienten herangezogen [5].

Der Reflexionskoeffizient ist die zentrale Messgröße des Ultraschallsensors und hängt von dem akustischen Impedanzunterschied zwischen Anpassungsschicht und Fluid ab. Da die akustische Impedanz einer Flüssigkeit durch deren Dichte und Viskosität bestimmt wird, führen Änderungen der Viskosität zu messbaren Variationen des reflektierten Signals.

Der Reflexionskoeffizient lässt sich experimentell wie folgt bestimmen:

$$R = \frac{A_m}{A_r} \quad (1)$$

wobei R den Reflexionskoeffizienten, A_r die Referenzamplitude und A_m die im Betrieb gemessene Amplitude des reflektierten Signals bezeichnet.

Der Referenzwert A_r wird im Rahmen eines Kalibrierverfahrens bestimmt, bei dem die Signalamplitude über den relevanten Temperaturbereich hinweg gegen Luft gemessen wird (vgl. Abbildung 2) [7].

Da ein Betrieb unter thermischen Lastwechseln zu erwarten ist, stellt die Auswahl eines geeigneten Wandlermaterials einen kritischen Aspekt des Sensordesigns dar. Insbesondere muss ein piezoelektrisches Material mit geringer temperaturabhängiger Drift der dielektrischen Eigenschaften gewählt werden, um die Langzeitstabilität der Referenzfunktion auch unter thermischer Zyklierung sicherzustellen.

Nach Etablierung der Referenzfunktion kann der Reflexionskoeffizient *in-situ* bestimmt und mittels eines geeigneten rheologischen Modells, beispielsweise eines Newton- oder Maxwell-Modells, in eine Viskosität überführt werden [5,8].

Im Rahmen dieser Arbeit wird jedoch ein praxisorientierter Ansatz verfolgt: Es konnte gezeigt werden, dass die Beziehung zwischen Reflexionskoeffizient und Viskosität für Kalibrieröle sowohl durch ein Maxwell-Modell als auch durch eine Exponentialfunktion beschrieben werden kann [2]. Dadurch ist es möglich, den Reflexionskoeffizienten direkt mit der unbekannten Viskosität zu korrelieren, indem der Sensor mit Referenzölen bekannter Viskosität kalibriert wird.

Ein entsprechendes Beispiel ist in Abbildung 3 dargestellt, in der der Reflexionskoeffizient in Abhängigkeit von der Viskosität bekannter Proben aufgetragen ist. Hierzu werden standardisierte Kalibrieröle (z. B. Canon-Normöle für Viskosimeterkalibrierung) verwendet. Es zeigt sich, dass die Messpunkte verschiedener Öle auf derselben Exponentialfunktion liegen.

Das Ergebnis dieser Kalibrierung ist somit eine direkte Zuordnung zwischen Reflexionskoeffizient und Viskosität, ohne dass

zusätzliche physikalische Eigenschaften der untersuchten Probe bekannt sein müssen.

Nach erfolgter Kalibrierung kann der Sensor Echtzeitmessungen des Reflexionskoeffizienten direkt in Viskositätswerte umrechnen. Der Einsatz hochfrequenter Scherwellen gewährleistet dabei eine hohe Sensitivität gegenüber den Basisfluideigenschaften und ermöglicht die Detektion subtiler Veränderungen, etwa durch Oxidation oder Gasabsorption.

Zur Verbesserung der Messgenauigkeit und Robustheit werden Signalverarbeitungsverfahren wie Rauschunterdrückung, Frequenzbereichsanalyse und Temperaturkompensation eingesetzt. Dadurch wird eine zuverlässige und reproduzierbare Viskositätsmessung unter variierenden Betriebsbedingungen ermöglicht.

Experimentelle Ergebnisse: Messung der Gasabsorption.

Zur Demonstration der Leistungsfähigkeit des vorgeschlagenen Sensors unter extremen Bedingungen wurden Experimente zur Erfassung von Viskositätsänderungen infolge von Gasabsorption in einer ionischen Flüssigkeit unter Hochdruckbedingungen durchgeführt (vgl. Abbildung 4).

Der Versuchsaufbau besteht aus einem Druckreaktor, in den der Sensor direkt integriert ist. Der Reaktor arbeitet bei Drücken von bis zu 80 bar und Temperaturen um 50 °C, was typischen Bedingungen industrieller Anwendungen mit Gas-Flüssig-Wechselwirkungen entspricht. Die Messergebnisse sind in Abbildung 5 dargestellt.

Die in den Experimenten verwendete ionische Flüssigkeit ist speziell so ausgelegt, dass sie eine minimale Wasserstoffabsorption aufweist. Dies wird durch Messungen bestätigt, die bei Wasserstoffbeaufschlagung konstante Viskositätswerte über die Zeit zeigen. Der Sensor detektiert zuverlässig das Ausbleiben signifikanter Viskositätsänderungen und demonstriert damit seine Stabilität unter konstanten Bedingungen.

Im Gegensatz dazu führt die Einleitung von Stickstoff in das System zu einer Absorption des Gases durch die ionische Flüssigkeit, was zu messbaren Änderungen der Viskosität führt. Der Sensor erfasst diese Veränderungen in Echtzeit und zeigt eine klare Variation des

reflektierten Ultraschallsignals, die dem Absorptionsprozess entspricht.

Die Ergebnisse verdeutlichen die hohe Sensitivität des Sensors gegenüber gasinduzierten Viskositätsänderungen, selbst unter Hochdruckbedingungen. Diese Eigenschaft ist insbesondere für Anwendungen in Verdichtern, chemischen Reaktoren und Energiesystemen von großer Relevanz, bei denen Gasverunreinigungen oder -absorption die Schmierstoffleistung erheblich beeinflussen können.

Die Möglichkeit, diese Messungen *in-situ* und ohne Probenentnahme oder externe Analytik durchzuführen, stellt einen wesentlichen Vorteil gegenüber konventionellen Methoden dar. Darüber hinaus gewährleistet die robuste Auslegung des Sensors einen zuverlässigen Betrieb in Umgebungen, in denen herkömmliche Sensoren versagen würden.

Schlussfolgerungen

In dieser Arbeit wurde ein neuartiges Ultraschallviskosimeter zur *In-situ*-Zustandsüberwachung von Schmierstoffen unter extremen Umgebungsbedingungen vorgestellt. Der Sensor kombiniert hochfrequente Scherwellenmessungen mit einer innovativen polymeren Anpassungsschicht, um die Einschränkungen konventioneller Ultraschall- und elektrischer Messverfahren zu überwinden.

Ein wesentlicher Vorteil des Sensors besteht darin, dass nach geeigneter Kalibrierung eine quantitative Bestimmung der Viskosität unabhängig von detaillierten physikalischen Eigenschaften des Prüffluids möglich ist.

Die zentrale Innovation liegt in der Verwendung der Anpassungsschicht zur Verbesserung der akustischen Kopplung und zur Erzeugung eines Resonanzeffekts, wodurch die Sensitivität gegenüber Viskositätsänderungen signifikant erhöht wird. Dieser Ansatz ermöglicht präzise Messungen über einen breiten Bereich von Fluidtypen, einschließlich niedrigviskoser Öle und komplexer Schmierstoffe.

Der Sensor zeigt eine robuste Leistungsfähigkeit unter rauen Bedingungen, insbesondere bei hohen Drücken und Temperaturen, und ist somit für anspruchsvolle industrielle Anwendungen geeignet. Die experimentellen Ergebnisse bestätigen die Fähigkeit des Sensors, Viskositätsänderungen infolge von Gasabsorption in Echtzeit zu

detektieren und unterstreichen sein Potenzial für fortschrittliche Zustandsüberwachungssysteme.

Insgesamt stellt der vorgestellte Sensor einen bedeutenden Fortschritt im Bereich der Schmierstoffüberwachung dar und bietet eine vielseitige sowie zuverlässige Lösung für Anwendungen, in denen konventionelle Messverfahren nicht einsetzbar sind.

Abbildungen

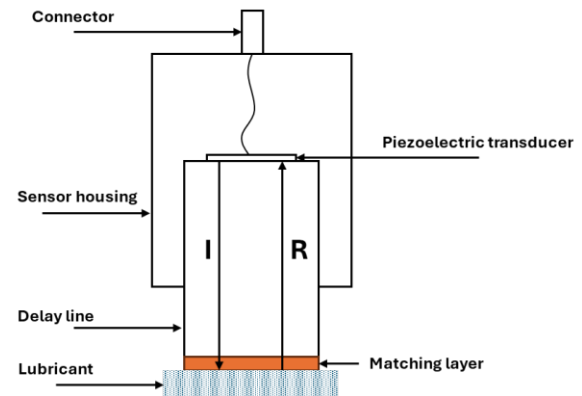


Abbildung 1: Schematische Darstellung der Hauptkomponenten des Sensors und seiner Funktionsweise

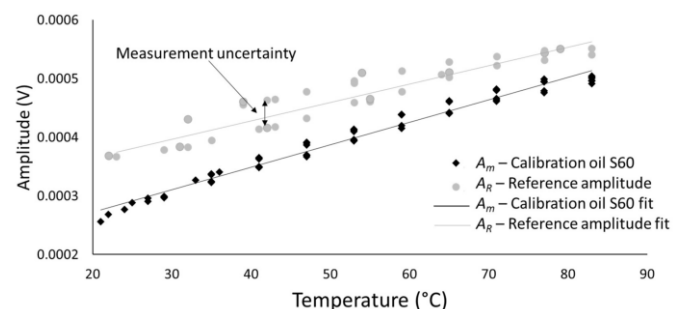


Abbildung 2: Referenzamplitude, gemessen über einen Temperaturbereich, und Messamplitude, ermittelt anhand eines Kalibrierungsstandards (Viskositätsmineralöl) [7]

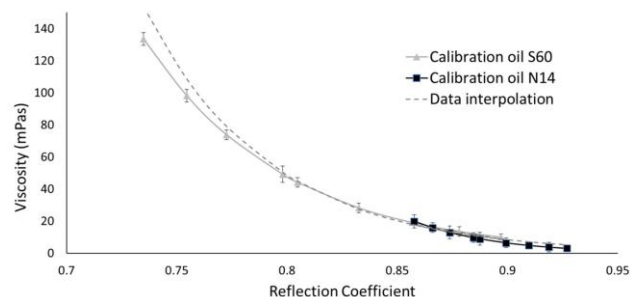


Abbildung 3: Reflexionskoeffizienten-Viskositätskurve für zwei Mineralöle mit Kalibrierviskosität

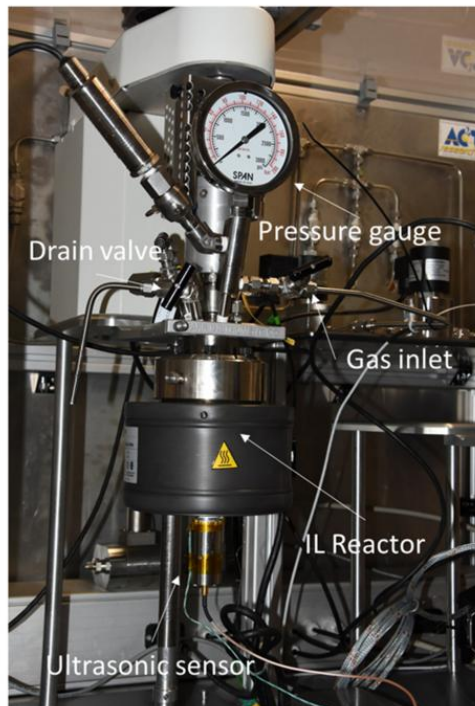


Abbildung 4: Versuchsaufbau zur Messung der Viskosität von Ölen unter Druck und bei Gassättigung [7].

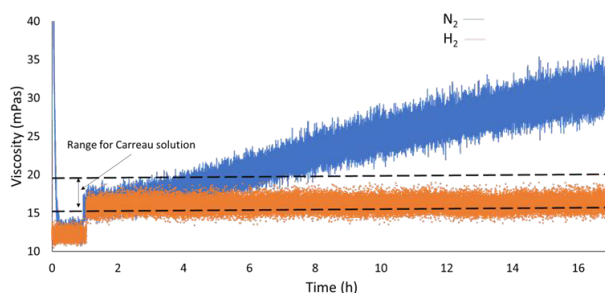


Abbildung 5: Viskositätsänderung aufgrund der Gasaufnahme von Stickstoff und Wasserstoff in einer ionischen Flüssigkeit in einem Reaktor bei 50 °C und 80 bar [7].

Aknowledgment

Diese Arbeit wurde im Rahmen des COMET Zentrums InTribology (FFG-Nr. 906860), einem Projekt des „Kompetenzzentrums für Tribologie“ (AC2T research GmbH), durchgeführt. InTribology wird durch das COMET – Competence Centers for Excellent Technologies Programm von den Bundesministerien BMIMI und BMWET als auch den Bundesländern Niederösterreich und Vorarlberg auf Basis der finanziellen Unterstützung durch die beteiligten Projektpartner gefördert. COMET wird durch die Österreichische Forschungsförderungsgesellschaft (FFG) abgewickelt.

Literaturnachweis

- [1] Shah, Raj, Rui Chen, and Vasu Bala. "Viscosity measurement at High Temperature and Varying Shear Rates for Engine Lubricants." *Petro Industry News* (2020).
- [2] Schirru, M., and M. Varga. "A review of ultrasonic reflectometry for the physical characterization of lubricated tribological contacts: history, methods, devices, and technological trends." *Tribology Letters* 70.4 (2022): 129.
- [3] Franco, Ediguer E., and Flávio Buiocchi. "Ultrasonic measurement of viscosity: Signal processing methodologies." *Ultrasonics* 91 (2019): 213-219.
- [4] Schirru, M., Li, X., Cadeddu, M., & Dwyer-Joyce, R. S. (2019). Development of a shear ultrasonic spectroscopy technique for the evaluation of viscoelastic fluid properties: Theory and experimental validation. *Ultrasonics*, 94, 364-375.
- [5] Schirru, M., Mills, R., Dwyer-Joyce, R., Smith, O., & Sutton, M. (2015). Viscosity measurement in a lubricant film using an ultrasonically resonating matching layer. *Tribology Letters*, 60(3), 42.
- [6] Fei, C., Ma, J., Chiu, C. T., Williams, J. A., Fong, W., Chen, Z., ... & Zhou, Q. (2015). Design of matching layers for high-frequency ultrasonic transducers. *Applied physics letters*, 107(12).
- [7] Schirru, Michele, and Michael Adler. "An ultrasonic rheometer to measure gas absorption in ionic liquids: design, calibration and testing." *Sensors* 20.12 (2020): 3544.
- [8] Franco, Ediguer E., et al. "Viscosity measurement of Newtonian liquids using the complex reflection coefficient." *IEEE transactions on ultrasonics, ferroelectrics, and frequency control* 55.10 (2008): 2247-2253.

Benutzungsgerechte Untersuchung der Wirkung von Gleitsichtgläsern

Friedrich Fleischmann¹, David Hilbig¹

¹Hochschule Bremen, Neustadtswall 30, 28199 Bremen, Deutschland

+49 421 5905 3453, (friedrich.fleischmann), (david.hilbig)@hs-bremen.de

Zusammenfassung

Moderne Gleitsichtgläser bestehen aus komplexen Freiformflächen. Sie werden zunehmend für die individuellen Bedürfnisse der Benutzer optimiert, was die Komplexität ihrer Charakterisierung erhöht. Gerade die für Brillenträger relevante und preislich sensitive Charakteristika wie Breite und Form des sogenannten Progressionskanals werden von aktuellen Normen nicht überprüft. Ein weiterer Anspruch an moderne Gläser ist zudem die korrekte Auslegung des Glases an die Gebrauchssituation, und damit die Blickrichtung während des Gebrauchs, durch die verschiedenen Bereiche des Glases unter anderen Strahlwinkeln wirksam werden. Die Individualisierung der Fertigung bedingt jedoch auch Messverfahren, die in der Lage sind, die Wirkung der Sehhilfe auf das Individuum in Gebrauchssituationen zu bestimmen. Der Beitrag gibt einen Überblick über Realisierungen des „Experimentellen Raytracing (ERT)“, mit denen eine Reihe funktionsrelevanter Messgrößen in Gebrauchssituationen bestimmt werden können und stellt eine Realisierung einer Anordnung dafür vor. Erste Ergebnisse am Beispiel einer Festbrennweiten-Brille werden vorgestellt.

Keywords: Optisches Messverfahren, experimental raytracing, Freiformen, Brillenmessung

Einführung

Mit zunehmenden Alter verliert das Auge seine Akkomodationsfähigkeit, wodurch für scharfes Sehen im Nah-, Mittel- und Fernbereich Sehhilfen mit unterschiedlicher Brechkraft notwendig werden. Gleitsichtbrillen („Progressive Addition Lenses“, PAL) bieten den Vorteil kontinuierlicher Übergänge zwischen diesen Sehbereichen. Dies wird durch eine allmähliche Zunahme der optischen Stärke entlang der Meridianachse erreicht. Allerdings lässt sich dieser Gradient nicht ohne Astigmatismus oder Zylinder in den Bereichen neben der Hauptmeridianachse erzielen. In diesen Übergangsbereichen kommt es zu Verzerrungen, die das scharfe Sichtfeld entlang der horizontalen Achse stark einschränken, insbesondere beim Sehen im Zwischen- und Nahbereich.

Moderne Gleitsichtgläser bestehen aus komplexen Freiformflächen. Fortschritte der Fertigungstechnik ermöglichen inzwischen eine bezahlbare Herstellung solcher optisch wirksamen Freiformen, weshalb die Gläser zunehmend für die individuellen Bedürfnisse der Benutzer optimiert werden. Zur Qualitätsbeurteilung beim Hersteller der Gläser werden Verfahren zur Topografiemessung wie Profilometrie oder Interferometrie eingesetzt, die optische Funktion kann mit Moiré-Deflektometrie [1] oder Shack-Hartmann-Sensoren [2, 3] un-

tersucht werden. Die Vermessung der gefertigten Brillengläser erfolgt zumeist durch Focimeter und beschränkt sich auf die Bestimmung der Brechkraft in einigen ausgewählten Messpunkten [4].

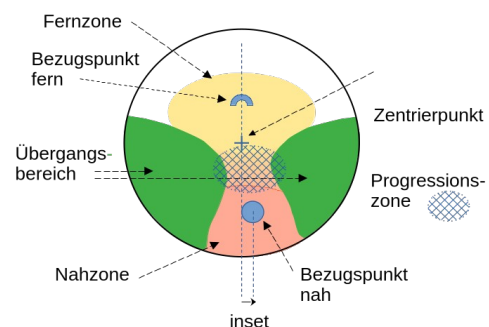


Abb. 1: Veranschaulichung eines Gleitsichtglases und relevanter Bereiche.

Gleitsichtbrillen sollen scharfes Sehen gerade in einer Vielzahl von unterschiedlichen Gebrauchsfällen ermöglichen, beispielsweise im Wechsel von Fernsicht zu Nahsicht während des Autofahrens oder bei der Büroarbeit. Herausfordernd ist dabei die Breite des Bereichs, in dem eine scharfe Sicht möglich ist, was direkt auf den Komfort im Gebrauch wirkt. In Abb. 1 sind charakteristische Bereiche eines (linken) Gleitsichtglases dargestellt. In den

Bezugspunkten nah bzw. fern wird üblicher Weise die Brechkraft überprüft. Das Inset gibt den Versatz zwischen Fern- und Nahsicht an, um die Konvergenz auszugleichen. In den Übergangsbereichen ist keine scharfe Wahrnehmung möglich.

Wie Minkwitz herausgearbeitet hat, bedingen sich die Addition längs einer Scheitellinie und die Zunahme des Astigmatismus senkrecht zu dieser gegenseitig [5], was im praktischen Fall dazu führt, dass bei stärkeren Brechkraftunterschieden zwischen den zu realisierenden Sehbereichen das brauchbare Sichtfeld seitlich stärker eingeengt wird. Dies drückt sich in der Breite der Progressionszone aus. Durch geschickte Verteilung der unvermeidlichen Aberrationszonen kann die individuelle Wirkung auf den Anwender und die Gestaltung des Übergangsbereichs optimiert werden.

Eine Beurteilung vieler qualitätsbestimmender Eigenschaften ist gerade bei individualisierten Gläsern mit den genannten Verfahren nicht möglich. Zudem ermöglichen moderne additive Fertigungsverfahren nicht nur die Herstellung individueller Gläser, sondern eröffnen auch die Möglichkeit einer lokalen Variation des Brechungsindex und damit auch eine neue Form, die bisherige prinzipielle Einschränkungen zu überwinden. Die Individualisierung der Fertigung bedingt jedoch auch Messverfahren, die in der Lage sind, die Wirkung der gefertigten Sehhilfe auf das Individuum in Gebrauchssituationen zu bestimmen. [6,7] In diesem Beitrag werden verschiedene Anwendungen des „Experimentellen Raytracing (ERT)“ zur Bestimmung relevanter Parameter mit hoher lokaler Auflösung und unter Simulation des Gebrauchs durch den Nutzer vorgestellt.

Methode

Das Verfahren des ERT ist in der optischen Messtechnik lange bekannt [8-10]. Beim ERT wird das zu prüfende Gerät (DUT) mit einem schmalen Lichtstrahl abgerastert, der in Abb.2

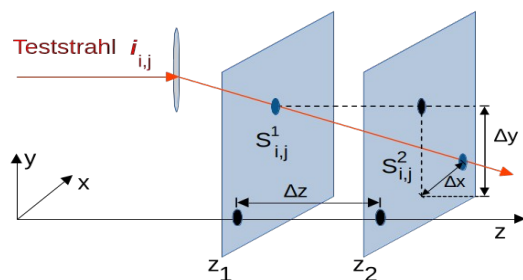


Abb. 2: Prinzip des ERT. Für jeden Messpunkt $i_{i,j}(x_{i,j}, y_{i,j})$ wird der Strahlschwerpunkt $S_{i,j}$ in 2 Ebenen gemessen.

als Teststrahl i bezeichnet wird. Die Richtung des gebrochenen Strahls wird anhand seines Schnittpunkts mit zwei parallelen Bildebenen ermittelt, wodurch ein Strahlgradient bestimmt werden kann. In beiden Ebenen z_1 und z_2 im Abstand Δz werden die Schwerpunktpositionen $S^1_{i,j}$ und $S^2_{i,j}$ aus den von einem CMOS-Sensor erfassten Intensitätsverteilungen berechnet. Der Gradient wird durch die Steigungen T_x und T_y des Strahls in x - und y -Richtung bestimmt und mit Gl. (1) berechnet, wobei die Position des Strahlschwerpunkts $S_{i,j}$ für jede Teststrahlposition $i_{i,j}$ als $z_1(x_1, y_1)$ bzw. als $z_2(x_2, y_2)$ benannt wird.

$$T_x = \frac{x_1 - x_2}{z_1 - z_2} = \frac{\Delta x}{\Delta z}; T_y = \frac{y_1 - y_2}{z_1 - z_2} = \frac{\Delta y}{\Delta z} \quad (1)$$

Über den Zusammenhang der partiellen Ableitungen der Wellenfront mit den gemessenen Strahlgradienten kann mit Gl. (2) durch modale Integration die Wellenfront rekonstruiert werden [10-13] und hieraus Parameter wie Brechkraft, Punktverbreiterungsfunktion (PSF), oder Modulationstransferfunktion (MTF) berechnet werden.

$$\nabla W = \begin{pmatrix} \frac{\partial W(x, y)}{\partial x} \\ \frac{\partial W(x, y)}{\partial y} \end{pmatrix} = \begin{pmatrix} T_x \\ T_y \end{pmatrix}. \quad (2)$$

Die PSF stellt die Punktantwort des optischen Systems, hier des Gleitsichtglases, dar.

Die Modulationsübertragungsfunktion (MTF) ist ein bekanntes Qualitätsmaß und beschreibt das Auflösungsvermögen im Frequenzbereich. Indem der Kontrastverlust vom Objekt zum Bild über der Raumfrequenz aufgetragen wird. Mathematisch ergibt sie sich aus dem Betrag der optischen Transferfunktion (OTF), der Fouriertransformierten der PSF.

Gerade bei Gleitsichtbrillen ist eine ortsauflöste Bestimmung der Qualitätsparameter erforderlich, da die optische Wirkung ja gezielt über das Glas variieren soll.

Wie Abb. 3 illustriert, wird in realen Anwendungsfällen nur ein kleiner Teil des Glases benutzt. Nur ein enger Bereich des Lichtes dringt nach Durchgang durch das Glas in die Pupille, wobei die Größe des Bereichs von der Brechkraft des Glases, dem Hornhautscheitelabstand und dem aktuellen, lichtabhängigen, Pupillendurchmesser abhängt. Die Wirkung des Glases in diesem Bereich, der sog. projizierten Pupille, bestimmt die Wahrnehmung des Tragenden. Erfolgt die ERT-Messung be-

schränkt auf diesen Bereich, ergibt sich die Charakterisierung des gewählten Parameters im Gebrauchsbereich, und bei Auswertung einer Vielzahl dieser Bereiche die orts aufgelöste Charakterisierung des gesamten Glases.

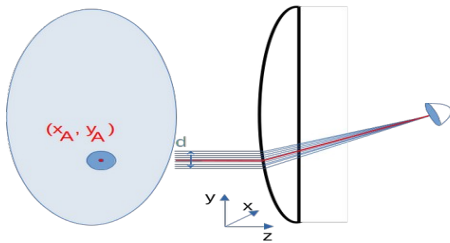


Abb. 3: Ortsaufgelöste Messung in einer Subapertur. Um jeden Messpunkt (x_A, y_A) , wird im Bereich der projizierten Pupille mit Durchmesser d die Wellenfront rekonstruiert.

Ein Beispiel hierfür ist in Abb. 4 für den Fall der Messung an 5 Positionen entlang der Meridionalebene einer um die typische 18° -Verkipfung des Brillenglases in der Fassung dargestellt. Die rot markierten Punkte symbolisieren die Messbereiche der untersuchten projizierten Pupillen.

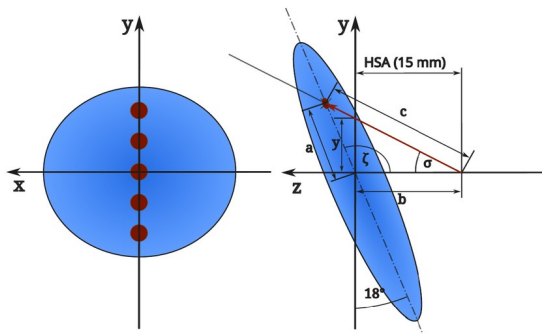


Abb. 4: Verkipfung des Glases zur Simulation der Gebrauchslage. Die Messung erfolgt hier an 5 Messpositionen der projizierten Pupille (rot markiert).

Aufbau

Zur Realisierung des Verfahrens wurde eine Kombination von 2 Translations- und Rotationsachsen zur Halterung des DUT und dessen Ausrichtung im Raum verwendet.

Als Teststrahlquelle diente eine Laserdiode Roithner SPL520-1-3-PD, betrieben unterhalb der Laserschwelle bei 27 mA, als Kamera eine boardlevel Photonfocus BL1-D2080-160-G2, mit Pixelgröße $8 \mu\text{m} \times 8 \mu\text{m}$.

Die prinzipielle Anordnung ist in Abb. 5 skizziert.

Anstelle der zur Realisierung des ERT erforderlichen z-Achse zum Anfahren der Messpositionen des Teststrahls wurde ein neuartiger integrierter Gradientensensor eingesetzt. In einem auf den Bildaufnehmer aufgesetztem Etalon erzeugt der eintreffende Teststrahl eine Reihe von Mehrfachreflektionen, aus deren Intensitätsverteilung dann der Eintrittswinkel berechnet wird. Um eine sichere Bestimmung der Strahlwinkel zu gewährleisten, wurde der Sensor um 5° aus der Normalen verkippt [15].

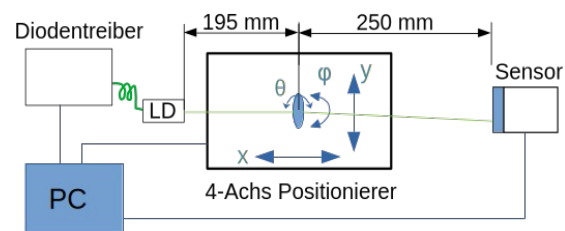


Abb. 5: Schematische Darstellung des Messsystems.

Die sich daraus ergebende Realisierung des 4-Achs-Positionierers und des Messsystems ist in Abb. 6 dargestellt.

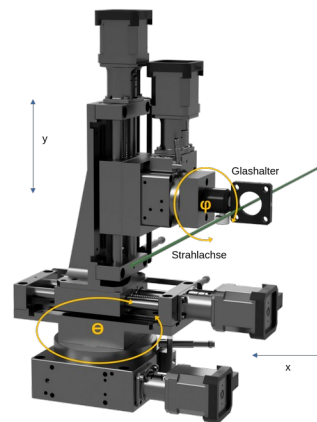


Abb. 6: Darstellung der Bewegungseinrichtung zur Halterung und Verkipfung eines Brillenglases zur Simulation der Gebrauchslage.

Verfahren

Das Brillenglas wird in den Glashalter der Bewegungseinrichtung montiert und senkrecht zum Strahlengang des Teststrahls ausgerichtet. Die PC-Ansteuerung des Positionierers richtet das Glas entsprechend der Neigung der Blickrichtung in der Vertikalen um Winkel ϕ und zur Horizontalen um θ aus, die sich durch die Drehungen ergebenden Verschiebungen des Durchtrittspunktes des Teststrahls werden durch Verfahren der Translationsachsen kompensiert. Um den Durchtrittspunkt herum wird ohne weitere Korrekturen der Bereich der

projizierten Pupille abgescannt. Aus der in dem Bereich bestimmten Wellefront wird dann die lokale Brechkraft berechnet.

Ergebnisse

Erste Ergebnisse zeigen die prinzipielle Funktionsfähigkeit der Anordnung. Abb. 5 zeigt exemplarisch die Ergebnisse einer Messung eines Festbrennweitenglasses bei Verkippung entlang der Drehachse ϕ . Die aus den gemessenen Verteilungen ermittelten Brechwerte enthalten neben der Brechkraft noch weitere Anteile wie Astigmatismus. Wiederholungsmessungen zeigen eine gute Wiederholbarkeit der Ergebnisse an den einzelnen Positionen.

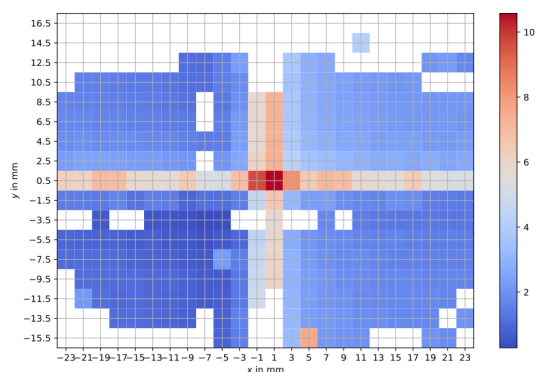


Abb. 7: Aus der Messung bestimmte Brechkraft des kompletten Brillenglasses bei 18° geneigtem Glas. Darstellung inklusive höherer Terme.

Diskussion

Die bislang gewonnenen Ergebnisse decken sich nicht mit den Erwartungen. Wie in Abb. 7 ersichtlich, wird gerade im mittleren Bereich eine sehr hohe Brechkraft, die rapide abnimmt, ermittelt, was beim Messobjekt mit fest 1,5 Dpt nicht plausibel ist und auf eine fehlerhafte Berechnung der Brechwerte aus den Strahlgradienten zurückzuführen sein dürfte. Hierfür kommen primär zwei Fehlerquellen in Frage. Eine erste Analyse zeigt eine unerwartet hohe Fehlerrate der automatisierten Schwerpunktsbestimmung der Mehrfachreflektionen im Gradientensensor, wodurch die Einfallswinkel sprunghaft falsch ermittelt werden. Dies erklärt jedoch nicht die systematisch falschen Werte im Bereich der Meridonal- und Sagittalebene. Ein Vergleich mit ZEMAX-simulierten Messungen deutet auf eine fehlerhafte Bestimmung der Wellenfronten in den projizierten Pupillen hin.

Literaturnachweis

- [1] E. De Lestrangé-Angineur, C. Kee, „Optical performance of progressive addition lenses (PALs) with astigmatic prescription. Sci Rep 2021 (11), 2984 (2021); doi: 10.1038/s41598-021-82697-0
- [2] E. Villegas, P. Artal, „Visual acuity and optical parameters in progressive-power lenses“ Optom. Vis. Sci. 83, 672–681 (2006); doi: 10.1097/01.opx.0000232818.37832.ee
- [3] D. R. Neal, R. Copland J., D. A. Neal, D. Topa, P. Riera, „Measurement of lens focal length using multicurvature analysis of Shack–Hartmann wavefront data, Proc SPIE, 5523 (2004), 243; doi: 10.1117/12.561772
- [4] E.A. Barbosa, D.M. Silva, C.E. Nascimento, F.L. Galvão, J.C.R. Mittani, „Progressive power lens measurement by low coherence speckle interferometry,“ Optics and Lasers in Engineering, Volume 51, Issue 7, 898-906 (2013); doi: 10.1016/j.optlaseng.2013.02.007
- [5] G. Minkwitz, „Über den Flächenastigmatismus Bei Gewissen Symmetrischen Asphären,“ Optica Acta: International Journal of Optics, 10(3), 223–227 (1963); doi: 10.1080/713817794
- [6] E. Pascual, JA Gómez-Pedrero, J. Alonso „Theoretical performance of progressive addition lenses with poorly measured individual parameters.“ Ophthalmic Physiol Opt.2023 Mar;43(2) 244-253 (2023); doi: 10.1111/opo.13088
- [7] E. Villegas, P. Artal, „Spatially resolved wavefront aberrations of ophthalmic progressive-power lenses in normal viewing conditions,“ Optom. Vis. Sci 2003; 80:106–114 (2003); doi: 10.1002/j.1538-9235.2003.tb01601.x
- [8] G. Häusler, G. Schneider, „Testing optics by experimental raytracing with a lateral effect photodiode,“ Appl. Opt 1988;27:5160–5164 (1988); doi: 10.1364/AO.27.005160
- [9] G. Gutierrez, D. Hilbig, F. Fleischmann, T. Henning, „Locally resolved characterization of progressive addition lenses by calculation of the modulation transfer function using experimental ray tracing.“ Proc. of SPIE 2017; Photonic Instrumentation Engineering IV, 101100E; doi: 10.1117/12.2251992
- [10] F. Fleischmann, D. Hilbig, "Use of experimental ray tracing for locally resolved characterization of progressive addition and contact lenses" Current Directions in Biomedical Engineering, vol. 11, no. 1, 2025, pp. 17-19 (2025); doi: 10.1515/cdbme-2025-0105
- [11] B. Nijboer, "The diffraction theory of optical aberrations: Part II: Diffraction pattern in the presence of small aberrations," Physica 13(10), 605–620 (1947); doi: 10.1016/0031-8914(47)90052-9
- [12] E. Wolf, "On a New Aberration Function of Optical Instruments," J. Opt. Soc. Am. 42(8), 547–552 (1952); doi: 10.1364/JOSA.42.000547
- [13] J. Rayces, "Exact relation between wave aberration and ray aberration," Optica Acta:

- International Journal of Optics 11, 85–88 (1964); doi: 10.1080/713817854
- [14] U. Ceyhan et al., "Inspection of aspherical lenses by wavefront analysis," Proc. SPIE 7389, 73893L–73893L–11 (2009); doi: 10.1117/12.834224
- [15] N. Mueller, R. Luening and F. Fleischmann, "Optical Beam Gradient Sensor," *Sensors and Measuring Systems; 21th ITG/GMA-Symposium*, Nuremberg, Germany, 1-7 (2022); <https://ieeexplore.ieee.org/servlet/opac?punumber=9861877>

Prototyp eines piezoelektrischen Energy Harvesters zur Umwandlung menschlicher Bewegungsenergie für die Versorgung von Wearables

B. Keune¹, M. Delle¹, C. Dirks¹, N. Boushaba¹, D. Brakensiek¹, A. Möller¹

¹THGA, Herner Str. 45, 44787 Bochum, Deutschland

Zusammenfassung

Diese Arbeit stellt den technischen Entwicklungsprozess und -fortschritt eines eigenen Prototyps zum Energy Harvesting menschlicher Bewegungsenergie für Lehr- und Forschungszwecke an der Technischen Hochschule Georg Agricola detailliert dar. Der Prototyp besteht aus piezoelektrischen Biegewandlern als elektrische Energiequelle, einer Power-Management-Platine zur Energieformung und Steuerung der Energieflüsse sowie einem Lithium-Ionen-Akkumulator als elektrochemischen Energiespeicher. Die Power-Management-Platine ist als Aufsteckplatine für einen Mikrocontroller vom Typ ESP32-C3-Mini-1 gestaltet. Die inneren Zustandsgrößen des Harvesters werden über den Mikrocontroller drahtlos per Bluetooth-Low-Energy versendet und an zentraler Stelle ausgewertet. Die funktionierende Wirkungskette des piezoelektrischen Energy Harvesters kann auf diese Weise experimentell nachgewiesen werden. Dabei zeichnet sich insbesondere die eigens entwickelte Power-Management-Platine durch ihre Effizienz aus. Schwächen des Prototyps zeigen sich noch in dem geringen Energieertrag der verwendeten piezoelektrischen Wandler sowie in dem relativ hohen Energiebedarf des Mikrocontrollers zum Versenden von Status-Nachrichten per Bluetooth.

Keywords: Bewegungsenergie, Energy Harvesting, piezoelektrischer Wandler, Wearables

Motivation – Was sind Wearables?

Als Alltagsgegenstände genutzte, tragbare elektronische Geräte wie Smartwatches und Fitnesstracker werden als sogenannte Wearables bezeichnet, vgl. Abb. 1.



Abb. 1: Beispiel-Produktabbildungen von Wearables (Bildquelle: Onlineversandhändler)

Moderne Wearables können über eine Reihe technischer Funktionen in den übergeordneten Kernfunktionen Konnektivität, Sicherheit, Gesundheits- und Fitnesstracking sowie Entertainment verfügen. Dazu zählen die Anzeige von Benachrichtigungen, WLAN für Fernupdates, GPS zur Navigation, NFC für kontaktloses Bezahlen, Sturzerkennung mittels Beschleunigungsmesser und Gyroskop, NotrufFunktionen, das Zählen von Schritten, Herzfrequenzmessung, Schlafanalyse und das Erstellen von EKGs. Immer neue technische Funktionen wie Sprachassistenten für eine komfortablere Bedienung und die Integration in alltägliche Geschäftsprozesse führen zu einer

steigenden Durchdringung von Wearables innerhalb der Gesellschaft. Ebenso wächst mit der zunehmenden Verbreitung solcher Geräte ihre Bedeutung in der individuellen Leistungsoptimierung und Gesundheitsüberwachung [1]. Wearables entwickeln sich somit zum unverzichtbaren Werkzeug im Alltag und zu einer treibenden Kraft technologischer Entwicklung.

Ihren Erfolg verdanken Wearables vor allem ihrem praktischen Nutzen und weitestgehend ihrer Gestaltung als mobile Endgeräte mit hohem Tragekomfort am menschlichen Körper, sei es als Armband oder Sonnenbrille. Prognosen erwarten, dass dieser Trend auch in Zukunft anhält. Denn der Mensch als mobiles Wesen verlangt nach effektiven Werkzeugen, um die technischen Möglichkeiten der Mobilität auszuschöpfen. Aus sozioökonomischer Perspektive umfasst der Begriff Mobilität sowohl die physische Bewegung im Raum als auch den sozialen Auf- oder Abstieg in der Gesellschaft. So wird die räumliche Erreichbarkeit von Bildung, Arbeit, Gesundheitswesen und finanziellen Mitteln als Schlüssel zur gesellschaftlichen Teilhabe und ein wesentliches volkswirtschaftliches Potential zur Produktivitätssteigerung betrachtet.

Ziel – Energy Harvesting für Wearables

Neben den bereits im Alltag etablierten Wearables erreichen daher immer neue Ideen aus Forschung und Entwicklung die Schwelle zur Praxis- und Alltagstauglichkeit, vgl. Abb. 2.

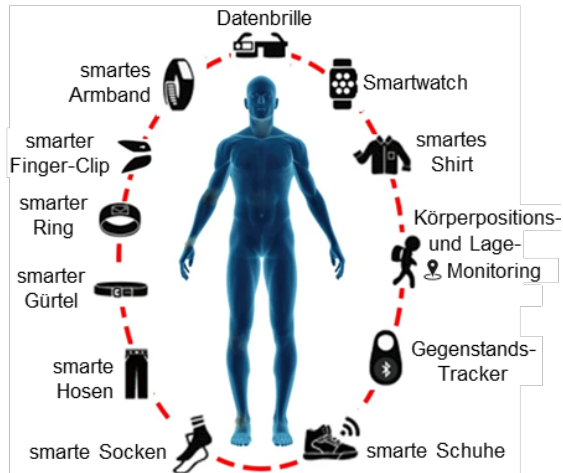


Abb. 2: Illustrative Darstellung von Wearable-Technologien [2] (Text vom Englischen ins Deutsche übersetzt)

Als elektronische Geräte hängt das Funktionieren von Wearables insbesondere von einer zuverlässigen und stabilen Energieversorgung ab. Dazu verfügen sie über elektrochemische Energiespeicher, vor allem Lithium-Ionen-Akkus mit typischen Kapazitäten zwischen 150mAh und 475mAh. Je nach Anwendung bedeutet das eine Laufzeit von 6 bis 30 Stunden. Durch Alterungsprozesse und abhängig vom individuellen Nutzungsverhalten kann diese durchschnittliche Gerätelauzeit deutlich sinken. Über geeignete Schnittstellen lassen sich Wearables bzw. ihre Energiespeicher bequem aufladen. Dies kann sowohl drahtgebunden passieren, z.B. über USB-C, oder drahtlos über den Qi-Standard. Wohingegen Wearables als mobile Geräte ausgestaltet sind, stammt die elektrische Energie dafür jedoch in der Regel aus einem stationären elektrischen Stromverteilnetz. In dicht besiedelten, urbanen Gebieten verfügt dieses Stromnetz flächendeckend über Netzanchluss- bzw. Ladepunkte. Immer häufiger finden sich auch Lademöglichkeiten an öffentlichen Plätzen in Form von Solarbänken und Straßenlaternen. Ziel solcher kommunaler Initiativen besteht darin, die Aufenthaltsqualität in öffentlichen Bereichen zu verbessern und gleichzeitig eine moderne Infrastruktur zu schaffen, die es ermöglicht, mobile Endgeräte möglichst bequem vor Ort aufzuladen. In dünn besiedelten, ländlichen Gebieten mit einer weniger gut ausgebauten

Infrastruktur, lassen sich Wearables durch tragbare Solarpaneele, Kurbelgeneratoren oder auch per Powerbank aufladen.

Diese praktischen Lösungen stehen jedoch den wesentlichen Vorteilen von Wearables als leichtgewichtige, mobile Endgeräte entgegen. Denn entweder muss zum Aufladen des Wearables eine verfügbarer Ladepunkt in der Umgebung verfügbar sein oder es müssen Hilfsmittel transportiert werden, die keinen vergleichbaren Tragekomfort bieten, deutlich schwerer wiegen oder selbst nur eine begrenzte Energiemenge bieten. Stattdessen werden unmittelbar am Träger verfügbare, leichtgewichtige Lösungen mit kontinuierlicher Ladefähigkeit benötigt, um längere Gerätelauzeiten von Wearables bei gleichzeitig hohem Trägerkomfort zu erreichen. Mit dieser Zielsetzung werden die technischen Einsatzmöglichkeiten von Energy Harvesting erforscht und stetig weiterentwickelt [3]. Energy Harvesting bezeichnet die Gewinnung typischerweise eher kleiner, in unmittelbarer Umgebung zum Endverbraucher verfügbarer Energiemengen, zum Beispiel aus Strahlung, Wärme oder Vibration. Quellen für Energy Harvesting sind thermische, kinetische und triboelektrische Energie, aber auch Solarenergie sowie elektromagnetische Energie aus hochfrequenten Feldern [4]. Theoretische und praktische Untersuchungen zeigen, dass der Mensch selbst als mögliche Energiequelle für Wearables ein beachtliches Ertragspotential bietet [5], vgl. Abb. 3.

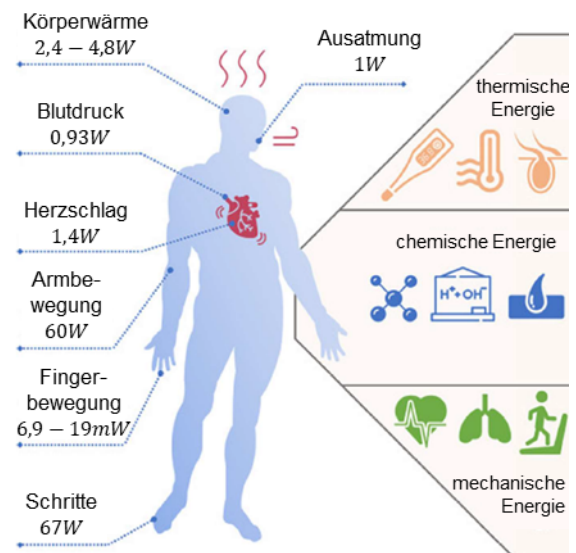


Abb. 3: Energieertragspotential Mensch [5] (Text vom Englischen ins Deutsche übersetzt)

Konzept – Schrittennergie nutzen

Insbesondere menschliche Bewegungsenergie bietet laut wissenschaftlicher Studien ein

ausreichendes Ertragspotential zur Energieversorgung von Wearables mittels Energy Harvesting [5]. Diese Energie kann über geeignete Energiewandler während des Gehens gewonnen, gespeichert und genutzt werden. Zur Energiewandlung können verschiedene Prinzipien Anwendung finden. Beim Induktionsprinzip wird eine elektrische Spannung aufgrund der zeitlichen Änderung des magnetischen Flusses durch eine Leiterschleife induziert. Bei der Energiewandlung mittels des piezoelektrischen Effekts wird hingegen die Eigenschaft bestimmter Werkstoffe ausgenutzt, bei ihrer Verformung durch äußere Krafteinwirkung einen elektrischen Strom treiben zu können. Vorteilhaft ist, dass sich die Bauformen piezoelektrischer Wandler relativ gut in Schuhwerk integrieren lassen, um dadurch die Bodenreaktionskräfte menschlicher Schrittenergie zu nutzen.

Bodenreaktionskräfte sind jene Kräfte, die der Boden gemäß dem dritten Newtonschen Axiom als Reaktion auf eine vom Menschen beim Gehen oder Laufen ausgeübte Kraft entgegengesetzt aufbringt. In der Biomechanik und Ganganalyse werden Bodenreaktionskräfte in drei Dimensionen gemessen: vertikal, vor/zurück und lateral. Beim Gehen entsteht für die vertikale Bodenreaktionskraft ein charakteristischer zweigipfliger Kurvenverlauf, vgl. Abb. 4.

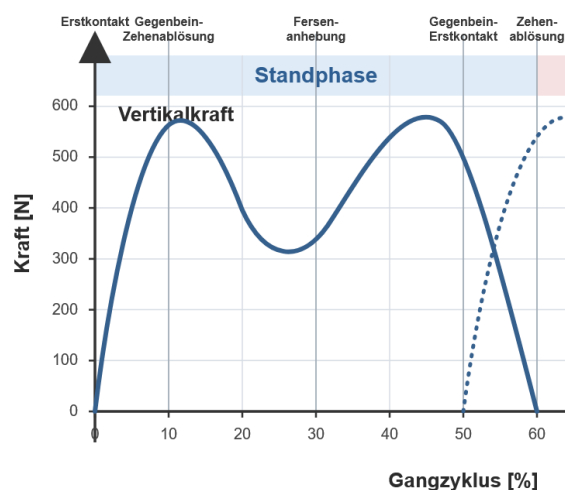


Abb. 4: Vertikale Bodenreaktionskraft über den Gangzyklus für ein Bein basierend auf [6]

Der erste Spitzenwert eines Gangzyklus tritt kurz nach dem Fersenkontakt auf und liegt oft über dem eigentlichen Körpergewicht. Die zweite Spitze entsteht am Ende der Standphase, wenn sich mit dem Vorfuß vom Boden abgestoßen wird. Zwischen diesen Spitzen befindet sich eine ausgeprägte Standphase. Dabei ist die Kraft geringer als

das Körpergewicht, da sich der Körperschwerpunkt über dem Fuß und dieser sich in der Abrollbewegung befindet. Ein typischer Gangzyklus beim Menschen wiederholt sich mit einer Frequenz von ca. zwei Hertz.

Das Butterworth-Van-Dyke-Ersatzschaltbild modelliert das elektrische Verhalten von piezoelektrischen Wandlern nahe ihrer Resonanzfrequenz, indem es das mechanische Schwingungsverhalten in rein elektrische Komponenten übersetzt. Für Schaltungssimulationen weit unterhalb der Resonanzfrequenz kann ein piezoelektrischer Wandler vereinfacht als Parallelschaltung aus Wechselstromquelle, Wirkwiderstand für dielektrische Verluste und statische Kapazität modelliert werden. Wird ein piezoelektrischer Biegewandler verformt, so entsteht durch die Verschiebung der Ladungsschwerpunkte im Kristallgitter eine elektrische Spannung. Bei einer periodischen Verformung biegt sich der Wandler abwechselnd nach oben und nach unten bzw. erfährt eine positive und negative Auslenkung. Dadurch wechselt das Vorzeichen der Ladungstrennung mit jedem Richtungswechsel und es entsteht eine Wechselspannung, die der Schwingungsfrequenz folgt. Dabei wechselt die Polarität in der Ausgangsspannung des Wandlers. Dieses Verhalten wird genutzt, um Vibrationen in elektrische Signale umzuwandeln. Je signifikanter die Verformung des piezoelektrischen Wandlers, desto größer seine erzeugte elektrische Spannung. Um aus einer solch fluktuierenden Energiequelle einen maximalen Leistungsertrag zu erzielen, wird ein effektives Power-Management benötigt. Dabei wird eine Optimierung der Leistungsabgabe durch Gleichrichtung der Quellspannung, Anpassung des Spannungslevels mittels Gleichstromsteller sowie durch ein effektives Maximum Power Point Tracking (MPPT) erreicht.

Die erreichbare Leistung aus solchen piezoelektrischen Energy Harvestern ist i.d.R. zu klein für einen kontinuierlichen Betrieb von typischen Wearables, sodass diese vielmehr aus einem Energiespeicher gespeist werden müssen. Um elektrische Energie zu speichern, eignen sich für mobile Anwendungen insbesondere Lithium-Ionen-Akkus aufgrund ihrer sehr hohen Energiedichte bei geringem Gewicht. Sie sind schnellladefähig, haben keinen Memory-Effekt und zeichnen sich durch eine geringe Selbstentladung sowie einer langen Lebensdauer mit vielen Ladezyklen aus. Dafür erfordern sie ein sicheres Batterielademanagement, um Überladung, Tiefenentladung oder Überhitzung zu

verhindern. Neben der zentralen MPPT-Funktionalität und dem Schutz des Energiespeichers, bestimmt das Power-Management die Leistungsflüsse zwischen Energiespeicher und dem elektrischen Verbraucher. Somit ergibt sich als Systemkonzept für einen Energy Harvester zur Umwandlung menschlicher Bewegungsenergie ein allgemeiner Aufbau bestehend aus drei zentralen Komponenten:

- (1) einer Zusammenschaltung piezoelektrischer Elemente als Energiewandler,
- (2) einer Power-Management-Platine zur effizienten Energieformung und Steuerung der Energieflüsse und
- (3) einem Akkumulator als elektrochemischer Energiespeicher.

Machbarkeitsnachweis – Prototyp

Ein Prototyp dient dem eigenen technischen Machbarkeitsnachweis eines piezoelektrischen Energy Harvesters zur Schrittennergieumwandlung. Für Menschen typische Gangzyklen haben eine Schrittfrequenz von nur ein bis drei Hertz. Die Resonanzfrequenz des piezoelektrischen Wandlers sollte also möglichst niedrig sein. Gleichzeitig muss er der mechanischen Belastung durch die wirkenden Gewichts- bzw. Bodenreaktionskräfte standhalten. Als Energiequelle wurde daher der piezoelektrische Biegewandler S118-J1SS-1808YB verwendet. Wie üblich für Piezoelemente, besteht er aus einem Blei-Zirkonat-Titanat-Verbund. Tab. 1 zeigt seine wesentlichen Spezifikationen.

Tab. 1: Spezifikationen S118-J1SS-1808YB

Parameter	Wert
Resonanzfrequenz	130 Hz
Leerlaufspannung	± 6 V
Lastwiderstand	15,7 kOhm
Nennleistung	0,7 mW
Nennbeschleunigung	± 8 G's
Abmessungen (mm)	55,4 x 23,4 x 0,46
Gewicht	9,537 g

Da die Resonanzfrequenz des piezoelektrischen Wandlers immer noch deutlich über der Schrittfrequenz liegt, werden drei Biegewandler vom gleichen Typ miteinander parallel verschaltet, um eine ausreichend hohe, messbare Stromstärke zu erreichen. Ein per 3D-Drucker gefertigtes Gehäuse zur provisorischen Befestigung an einer Schuhspitze und gleichartiger Anregung der Biegewandler ist in Abb. 5 dargestellt.

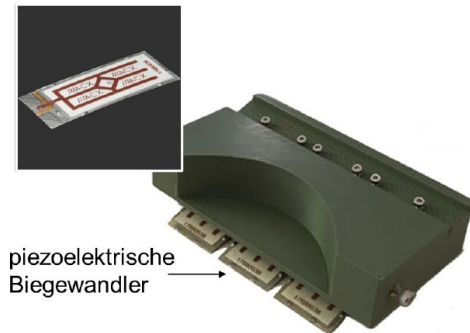


Abb. 5: Foto von S118-J1SS-1808YB

Für die Entwicklung der Power-Management-Platine wurde der PMIC AEM00330 von e-Peas ausgewählt. Der AEM00330 integriert effiziente Buck/Boost-Wandler anstelle verlustbehafteter LDOs (Low Dropout Regulatoren) und hat einen integrierten Shipping Mode, der es ermöglicht, den Ruhestrom des Systems im inaktiven Zustand auf ein Minimum zu reduzieren [7]. Buck/Boost-Wandler verändern die Spannung durch schnelles Schalten, wohingegen LDOs die überschüssige Energie an Widerständen in Wärme umwandeln. Somit eignet sich der AEM00330 besonders für kompakte und niedrigleistungsfähige Systeme wie drahtlose Sensorknoten. Neben der Lastversorgung übernimmt der IC auch wichtige Schutz- und Steuerfunktionen, um die Lebensdauer des Energiespeichers zu verlängern und den zuverlässigen Betrieb zu gewährleisten. Darüber hinaus ist der IC für alle gängigen Typen elektrochemischer Energiespeicher geeignet und bietet eine passende Ladespannung. Der Chip integriert somit als zentraler Baustein wesentliche Funktionen der Platine. Dazu zählen

- ✓ Gleichspannungskonverter mit hohem Wirkungsgrad,
- ✓ Energiespeicherverwaltung,
- ✓ Maximum Power Point Tracking (MPPT),
- ✓ Ladeschutzmechanismen,
- ✓ Systemstart aus energielosem Zustand (Kaltstart).

Durch seine externe Beschaltung, kann der AEM00330 für verschiedene Anwendungsfälle, z.B. unterschiedliche Quellen und Lasten, konfiguriert werden. Um den Zustand des Energy Harvesters und insbesondere den Ladezustand (engl. SoC) des Energiespeichers zu monitorieren, wurde der AEM00330 als PMIC auf einer eigens entwickelten Platine inklusive Dioden-Gleichrichterschaltung eingesetzt. Diese ist als HAT, also als Aufstreckplatine, für den Mikrocontroller ESP32-C3-Mini gestaltet.

Dieser Mikrocontroller hat einen geringeren Energieverbrauch als z.B. die verwandte Entwicklungsplatine ESP32-Wroom und auch kleinere Abmessungen, verfügt jedoch über die gleichen Konnektivitätsmöglichkeiten per Bluetooth-Low-Energy (BLE). Der ESP32-C3-Mini basiert auf einer energieeffizienten 32-Bit-RISC-V-CPU und unterstützt verschiedener Low-Power-Betriebsmodi (Sleep, Deep-Sleep, Modem-Sleep). Dadurch kann die Stromaufnahme im Ruhezustand auf unter $5\mu\text{A}$ sinken. Dies ist entscheidend, um die durch das Piezoelement gewonnene, geringe Energiemenge möglichst effizient zu nutzen. Somit eignet sich der ausgesuchte Mikrocontroller im direkten Vergleich besser für die Anwendung im Energy Harvesting. Abb. 6 zeigt die fertige, mit KiCad designte Power-Management-Platine ohne und mit aufgestecktem Mikrocontroller [8].

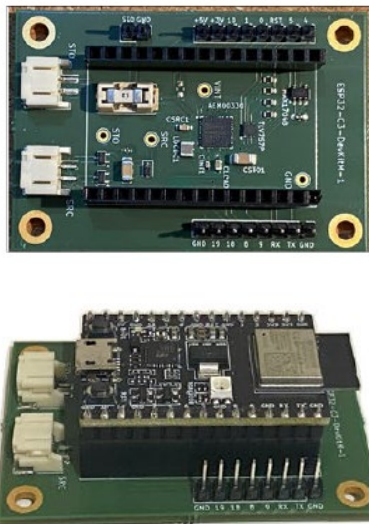


Abb. 6: Foto von Power-Management-Platine ohne und mit aufgestecktem ESP32-C3-MINI-1

Der AEM00330 startet im sogenannten Reset State, in dem alle internen Zustände zurückgesetzt sind und keine Spannungen anliegen. Sobald am Eingang der Power-Management-Platine durch den piezoelektrischen Biegewandler eine Eingangsspannung $V_{SRC} > 275\text{mV}$ sowie eine Eingangsleistung $P_{SRC} > 3\mu\text{W}$ erreicht werden, wechselt der AEM00330 in den Wake-Up State. Erreicht anschließend die interne Spannung VINT den Wert von $2,2\text{V}$, ist der Cold Start abgeschlossen und der Chip befindet sich im Start State. Abhängig von der gesetzten Priorisierung (über den Pin STO PRIO) beginnt der AEM00330 nun, entweder primär den Speicher (STO) zu laden oder direkt die Last (LOAD) zu versorgen. Sobald beide Ausgänge die jeweils konfigurierten

Schwellenwerte erreicht haben, wechselt das System in den Supply State. Im Supply State steht der volle Funktionsumfang des Power-Managements zur Verfügung. Der AEM00330 kann nun dynamisch Energie zwischen Quelle, Speicher und Last verteilen, je nachdem, wo Energie verfügbar ist oder benötigt wird. Idealerweise befindet sich der Chip dauerhaft in dieser Phase, da sie einen stabilen und effizienten Betrieb des Gesamtsystems ermöglicht.

Zur Überwachung des Ladezustands und der Spannung des eingesetzten Lithium-Ionen-Akkus wird der Fuel Gauge MAX17048 von Analog Devices verwendet. Das Bauteil bietet eine Lösung zur Ladezustandserfassung mit minimalem Energieverbrauch und ist damit gut geeignet für energieeffiziente Systeme. Der MAX17048 übernimmt die Aufgabe, den SoC des Akkus zu berechnen und dessen Spannung kontinuierlich zu überwachen. Diese Informationen werden über seine I2C-Schnittstelle an den Mikrocontroller ESP32-C3-Mini übertragen. Die Verbindung erfolgt über das Aufstecken auf einer Stiftleiste. Als Energiespeicher wurde der Lithium-Ionen-Akku Samsung INR18650-25R mit einer Nennspannung von $3,6\text{V}$, einer Kapazität von 2500mAh , einer max. Ladespannung von $4,2\text{V}$, einem max. Ladestrom von 4A und einer Entladeschlussspannung von $2,5\text{V}$ verwendet. Verglichen mit dem Energieertrag des piezoelektrischen Wandlers ist der Energiespeicher überdimensioniert, erfüllt jedoch für erste Untersuchungen am Prototyp seinen Zweck.

Ergebnisse & Ausblick

Abb. 7 zeigt den Aufbau des Demonstrators bestehend aus parallel geschalteten, piezoelektrischen Biegewandlern, Power-Management-Platine mit aufgestecktem Mikrocontroller, angeschlossenen Batteriespeicher und einem weiteren Mikrocontroller vom Typ ESP32-Wroom mit OLED-Display als Datensinke und zur Anzeige per BLE empfangener Zustandsdaten.

Ein erster technischer Machbarkeitsnachweis der Wirkungskette des piezoelektrischen Energy Harvesters verlief erfolgreich. Durch wiederholtes Drücken und damit Anregung der Biegewandler, konnte eine ausreichend hohe Leistung für den Kaltstart der Power-Management-Platine erreicht werden. Ein Ladevorgang des Lithium-Ionen-Akkus konnte messtechnisch nachgewiesen werden.

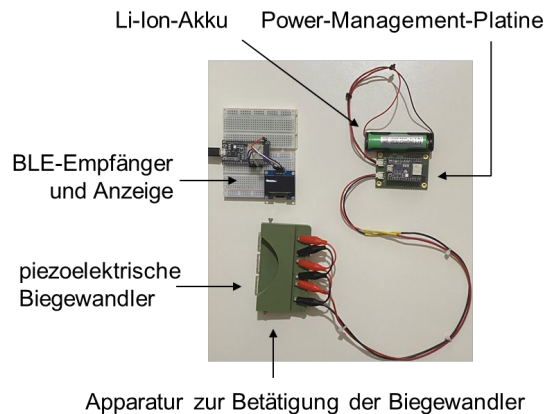


Abb. 7: piezoelektrischer Energy Harvester [8]

Die inneren Zustände, wie bspw. der SoC, wurden zyklisch auf dem OLED-Display aktualisiert. Der gewonnene Leistungsertrag konnte über weite Bereiche wiederholter Versuchsdurchführungen aufrecht erhalten bleiben. Jedoch war die Ladeleistung aus Ladespannung von 3,6V und einem Ladestrom am Akku mit 0,1mA noch sehr niedrig. Tatsächlich reichte diese nicht aus für eine nennenswerte Ladung des 2500mAh-Akkus. Darüber hinaus wurden bei der Versorgung der Last, also des Mikrocontrollers ESP32-C3-Mini, Stromspitzen von 60mA beim Versenden von BLE-Statusnachrichten registriert. Aus diesem Grund wurde der Akku vor Versuchsstart über ein Labornetzgerät mit Strombegrenzung aufgeladen. Mit dem erfolgreichen Kaltstart des Energy Harvesters bei Anregung der piezoelektrischen Energiewandler, konnte dieser die Last aus dem bereits aufgeladenen Akku speisen. Insgesamt ist der entwickelte Prototyp und Demonstrator-Aufbau unter der Maßgabe eines Rapid-Prototyping-Prozesses mit Nachweis der technischen Kernfunktionalität ein Erfolg. Jedoch gilt es für zukünftige Weiterentwicklungen dem noch relativ hohen Energiebedarf des Mikrocontrollers zur reinen Messdatenversendung und -protokollierung zu begegnen. Insbesondere bei der BLE-Kommunikation kann ein dedizierter BLE-Chip auf der Power-Management-Platine eine signifikante Energieeinsparung gegenüber einem HAT-Design mit ESP32-Mikrocontroller bedeuten. Ultra-Low-Power BLE-SoCs benötigen im Tiefschlaf oft nur wenige μA , ihre Spitzenströme beim Senden und Empfangen liegen typischerweise nur bei wenigen mA und damit deutlich unter den im Versuchsaufbau registrierten 60mA vom ESP32. Darüber hinaus ist der verwendete Lithium-Ionen-Akku im Vergleich mit typischen Kapazitäten von Wearables mit 2500mAh überdimensioniert, sodass sich gespeicherte Ladungen nicht

unmittelbar in einem höheren SoC widerspiegeln. Eine wesentliche Schwäche zeigt sich auch in dem Energieertrag, der mit einer erreichten Leistung im μW -Bereich zu niedrig für ein ausreichend schnelles Wiederaufladen des Akkus oder den Betrieb eines Wearable-Gerätes ist. Mit dem neu verfügbaren ePeas-PMIC AEM13921 lässt sich Energie aus zwei separaten Energiequellen gewinnen, sodass ein verbesserter piezoelektrischer Wandler kombiniert mit einem Solarmodul ein theoretisch ausreichendes Ertragspotential bieten kann. Und schließlich ist der Prototyp mit Gesamtkosten von ca. 100€ noch weit entfernt von einem wirtschaftlichen Use-Case.

Literatur

- [1] Liu, F., Cui, F., Tong, Y. et al. Research Progress on Human Body Energy Harvesting and Storage Systems for Wearable Electronic Devices. *J. Electron. Mater.* 54, 8103–8143 (2025). <https://doi.org/10.1007/s11664-025-12232-8>
- [2] Rodrigues, J., Segundo, D., Junqueira, H., Sabino, M., Prince, R., Al-Muhtadi, J., de Albuquerque, V. Enabling Technologies for the Internet of Health Things. *IEEE Access* (2018) DOI 10.1109/ACCESS.2017.2789329.
- [3] Khalid, S., Raouf, I., Khan, A. et al. A Review of Human-Powered Energy Harvesting for Smart Electronics: Recent Progress and Challenges. *Int. J. of Precis. Eng. and Manuf.-Green Tech.* 6, 821–851 (2019). <https://doi.org/10.1007/s40684-019-00144-y>
- [4] Spies, P., Pollak, M., Mateu, L. *Handbook of Energy Harvesting Power Supplies and Applications*. Verlag Pan Stanford Publishing Pte Ltd, 2015, ISBN-13: 978-9814241861.
- [5] Zou, Y., Bo, L., Li, Z. Recent progress in human body energy harvesting for smart bioelectronic system. *Fundamental Research*, Volume 1, Issue 3, 2021, Pages 364-382, ISSN 2667-3258, <https://doi.org/10.1016/j.fmre.2021.05.002>.
- [6] Boushaba, N. (2025). Analyse und prototypische Entwicklung eines piezoelektrischen Energy Harvesters zur Schrittenenergieumwandlung. (unveröffentlichte Bachelorarbeit). THGA.
- [7] E-peas S.A. AEM00330: Photovoltaischer Konstantspannungs-Buck-Boost-Wandler mit DCDC-Ausgang. 2025. url: <https://e-peas.com/documents/AEM00330/DSAEM00330.pdf> (abgerufen am 03. 08. 2025).
- [8] Delle, M. (2025). Piezoelektrisches Energy Harvesting zur Versorgung eines leistungsarmen Mikrocontrollers - Design einer Leiterplatte mit integrierter Stromversorgungsschaltung zur Umwandlung von Bewegungsenergie in elektrische Energie (unveröffentlichte Bachelorarbeit). THGA.

Von der Elektrode zur Cloud: Leckagemonitoring mit kapazitiver Sensorik und mioty®

Johannes Ehrlich¹, Bernhard Brunner¹, Thomas Nierla¹, Matteo Benkert¹, Gerhard Domann¹, Ferdinand Kemeth², Alexej Jarresch², Anja Meldau²

¹ Fraunhofer-Institut für Silicatforschung ISC, Neunerplatz 2, 97082 Würzburg
+49 931 4100-416, bernhard.brunner@isc.fraunhofer.de

² Fraunhofer-Institut für Integrierte Schaltungen IIS, Nordostpark 84, 90411 Nürnberg
+49 911 58061-3330, ferdinand.kemeth@iis.fraunhofer.de

Die Detektion von Flüssigkeitsleckagen ist sowohl im privaten als auch im industriellen Bereich wichtig, um Sach- und Personenschäden zu verhindern. Herkömmliche Leckageerkennungen nutzen meist ein resistives Messverfahren, das nur funktioniert, wenn die Flüssigkeit die Elektroden benetzt und eine ausreichende Leitfähigkeit aufweist. Wasser ist hierbei oft kein Problem, Lösemittel, Öle, Treibstoffe oder Chemikalien sind jedoch schwierig oder gar nicht zu erkennen. Die Autoren dieser Arbeit setzen auf ein kapazitives Messprinzip, das flexible und auch elastische Materialien für besonders herausfordernde Anwendungsfälle der Sensorik nutzt. Der Vorteil liegt darin, daß bereits der erste Tropfen, also ca. 100 µl, unabhängig von dessen Auftreffstelle auf der kapazitiven Detektionsfläche nachgewiesen werden kann. Die Sensorstruktur besteht aus einer gedruckten Interdigitalelektrode auf einer flexiblen oder elastischen Trägerfolie, die in unterschiedlichen Größen kostengünstig hergestellt werden kann. Der Sensoraufbau ist durch ein elastisches Dielektrikum gekapselt, damit bleibt das Elektrodenmaterial vor Umwelteinflüssen und aggressiven Medien geschützt. Durch die Nutzung von hochelastischem Silikon wird es möglich, Sensoren auch auf sehr stark gekrümmten Oberflächen z. B. Rohrverbindungen, anzubringen. Tests belegten den Nachweis von Wasser, Isopropanol, Aceton, Benzin, Säuren und Laugen ab einer Tropfengröße von 100 µl. Eine eigens entwickelte Elektronik mit Langzeitbatteriebetrieb (10 Jahre) wertet das kapazitive Signal energieeffizient aus und setzt auf ein fortschrittliches Datenübertragungssystem (mioty® LPWAN) mit sicherem Protokoll sowie Heartbeat- und Fernwartungsoptionen.

Keywords: Flüssigkeitssensorik, kapazitives Messsystem, energieeffizient, Datenübertragung mit hoher Reichweite

Einleitung

Allein in Deutschland ereignen sich laut dem Gesamtverband der Versicherer ca. 1,1 Mio. Leitungswasserschäden pro Jahr in Privathaushalten. Durchschnittlich müssen die Versicherungen ca. 4.500 € pro Schaden über die Wohngebäudeversicherung regulieren. Zusätzlich kommen jährlich etwa 150.000 Schadensfälle in der Hausratversicherung mit einer Regulierungssumme von 350 Mio. € hinzu [1]. Im industriellen Umfeld sind insbesondere Chemieanlagen, Raffinerien sowie Rohrleitungen, Tanks und Gefahrstofflager mit höchster Priorität zu überwachen, um Risiken für das Personal und Schäden für die Umwelt möglichst gering zu halten.

Eine Sensorik, die sich einfach auch an unzugänglichen oder gekrümmten Flächen nachrüsten lässt, kabellos über einen langen Zeitraum Daten an eine Leitstelle übermittelt und zuverlässig kleinste Mengen ausgetretener

Flüssigkeit detektiert, kann dabei helfen, Risiken und wirtschaftliche Schäden zu verringern.

Funktionsweise der Leckagesensorik ab dem ersten Tropfen

Die Funktionsweise der Leckagesensorik beruht auf einem kapazitiven Messprinzip, wie in Abb. 1 dargestellt. Die Interdigitalstrukturen bilden dabei elektrisch gesehen einen Kondensator, dessen Feldverlauf durch die geometrischen Abmessungen und Materialparameter bestimmt ist (siehe Abb. 1, links unten). Wichtig hierbei ist die Permittivität der eingesetzten Materialien; in dem erläuterten Fall wird vorwiegend mit Polymeren wie Silikon ($\epsilon_r = 2,8$) oder Polyimid ($\epsilon_r = 3,4$) gearbeitet.

Wird durch eine Leckage, also einem Tropfen eines Mediums, die Zusammensetzung der Permittivität des Sensoraufbaus und damit der Feldverlauf des Kondensators verändert, wird dies durch eine Veränderung der Kapazität

messbar (siehe Abb. 1, rechts unten). Besonders gut messbar sind dabei Flüssigkeiten, die im Vergleich zu den verwendeten Grundmaterialien eine höhere Permittivität aufweisen. Beispielhaft ist hierfür Wasser mit einer Permittivität von ca. 80.

Dieser Vorteil kann aber gleichzeitig ein Nachteil sein: Fällt der Unterschied zu gering aus, ist eine sichere Detektion gegebenenfalls nicht mehr möglich, oder es werden größere Flüssigkeitsmengen benötigt, um einen messbaren Unterschied hervorzurufen. Ein großer Vorteil des kapazitiven Messprinzips ist, dass es keine offenen Elektrodenbereiche gibt. Dies verringert den generellen Einfluss von Umweltfaktoren, die z. B. zu einer Korrosion der Elektroden führen und den eingesetzten Sensor frühzeitig zerstören können.

Ein weiterer Vorteil ist, dass das zu detektierende Medium keine erhöhte Leitfähigkeit aufweisen muss. Dies erweitert den Einsatzbereich auch auf potentiell isolierende Flüssigkeiten wie Lösemittel und weitere technische Medien.

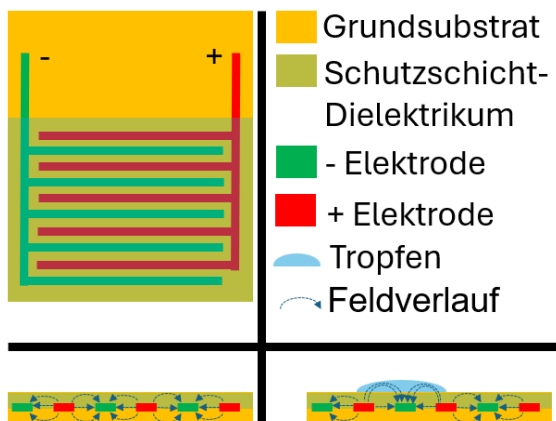


Abb. 1 Funktionsweise der kapazitiv arbeitenden Leckagesensorik

Materialien und Methoden

Die Herstellung der Leckagesensoren erfolgt zunächst als Einzelsensor auf einem elastischen Silikonsubstrat. Dazu werden die einzelnen Schichten, wie in Abb. 2 dargestellt, schrittweise nacheinander aufgebaut. Das kommerziell erhältliche 2-Komponenten-Silikon RT601 von Wacker wird nach dem vom Hersteller vorgegebenen Mischverhältnis in einem Speedmixer im flüssigen Zustand für ca. 3 Minuten gemischt und anschließend in einer Dicke von ca. 200 µm auf eine Glasplatte mit einem Rakel aufgerakelt. Der so entstandene Folienfilm wird anschließend thermisch mittels IR-Strahlung für 10 Minuten vernetzt. Direkt im Anschluss wird mit einem Rakel eine Elektrodenschicht, bestehend aus Carbon

Black mit Silikonbinder, in einer Schichtdicke von ca. 30 µm vollflächig aufgerakelt und ebenfalls anschließend thermisch vernetzt. Somit sind sowohl die Träger-, Isolier- und auch Elektrodenschichten sehr elastisch. Nach diesem Schritt erfolgt die Strukturierung der Interdigitalstruktur mittels Laserabtrag mit einem Gravurlaser. Dabei werden Interdigitalstrukturen in Negativform bis auf das elastische Silikonsubstrat abgetragen.

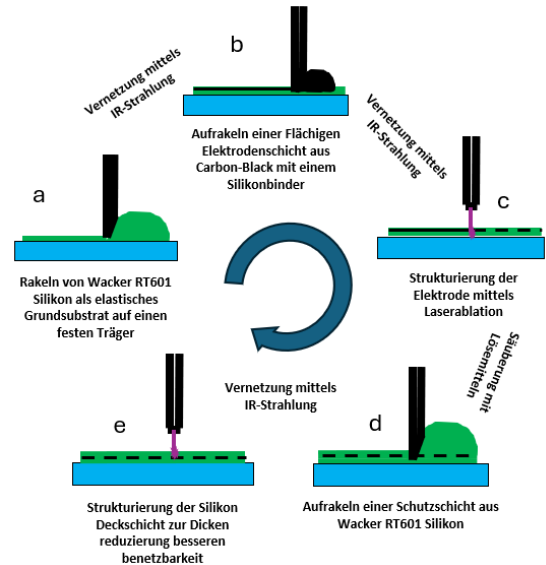


Abb. 2 Schematische Herstellung der Interdigitalstrukturen durch Laserabtrag.

Durch Vorversuche wurde ermittelt, dass eine Leiterbahnbreite von 0,5 mm und ein Leiterbahnabstand von 0,5 mm stabil und ohne Kurzschlüsse hergestellt werden können. Diese Parameter werden damit für alle folgenden Proben und Versuche festgelegt. Die aktive Fläche der angefertigten Interdigitalstruktur eines Einzelsensors beträgt $65 \times 50 \text{ mm}^2$. Nach dem Laserabtrag wird die Interdigitalstruktur mit einem Dielektrikum als oberste Schutzschicht mit einer Dicke von ca. 60–100 µm aufgerakelt. Die hergestellte Interdigitalstruktur auf elastischem Silikon ist in Abb. 3a dargestellt. Neben den Interdigitalstrukturen aus elastischem Silikon wurden in nahezu demselben Herstellungsprozess wie in Abb. 3 Interdigitalstrukturen auf flexiblem Polyimidsubstrat hergestellt (Abb. 3b). Dabei ersetzt das flexible Polyimid das elastische Silikon als Grundsubstrat. Bei diesem Sensortyp wird keine zusätzliche Schicht für die Elektroden benötigt; durch den Lasergravurprozess wird die Oberfläche des Polyimids verbrannt, und es entsteht eine leitfähige Elektrodenschicht, die direkt genutzt werden kann. Dies reduziert den Aufwand für einen späteren Hoch-

skalierungsprozess, da der Schritt der Elektro-
drierung und Vernetzung entfällt.
Die flexiblen Polyimidsubstrate werden nach
dem Reinigen mit Lösemittel ebenfalls mit einer
Deckschicht als isolierendem Dielektrikum und
Schutzschicht aus Silikon mit einer Dicke von
60–100 μm überzogen. Im weiteren Verlauf
wurden nur noch Interdigitalstrukturen aus
flexiblem Silikonsubstrat betrachtet.

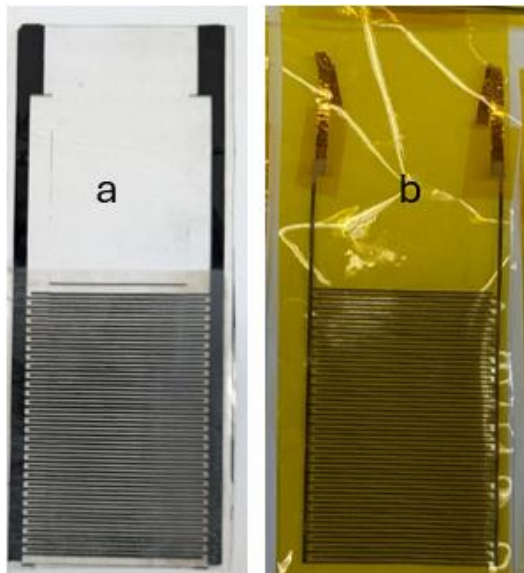


Abb. 3 Hergestellte Interdigitalstrukturen auf
a.) elastischem Silikonsubstrat
b.) flexiblem Polyimid Substrat

Als Weiterentwicklung des Einzelsensors wurde
die Geometrie der Einzelsensoren in ein
Sensorarray bestehend aus 16 Einzelsensoren
mittels CAD-Layout überführt und an-
schließend durch Siebdruck hergestellt. Dazu
wurde eine PET-Folie von DuPont als
Basissubstrat verwendet und das Elektroden-
muster mit einem Siebdrucker aufgedruckt. Als
leitfähige Tinte zur Herstellung der Elektroden
kam diesmal EDAG 479SS von LOCTITE zum
Einsatz (Abb. 4). Die Tinte unterscheidet sich im
Wesentlichen durch die Verwendung von Silber
anstatt Carbon Black als leitfähige Partikel. Das
Elektrodenmuster wurde abschließend mit
einer Silikonschicht aus RT601 mit einer Dicke
von 60–100 μm als Schutz und Dielektrikum
durch Rakeln überschichtet.
Das so entstandene Leckagesensor-Array
weist damit nicht nur eine deutlich vergrößerte
Detektionsfläche (ca. DIN A4) auf, sondern es
ist nun auch möglich, die Leckage orts aufgelöst
zu detektieren. Als Anbindung an die
Auswerteelektronik kommt ein Steckver-
binder mit Crimpkontakten der Firma Amphenol (Typ
FCI Clincher 65801-009LF) zum Einsatz.

Gerade die Ortsauflösung kann bei komplexen
Überwachungsaufgaben wie z. B. Verteilern
oder engen Rohrleitungssystemen von Vorteil
sein. Prinzipiell ist die Ortsauflösung auch
skalierbar.

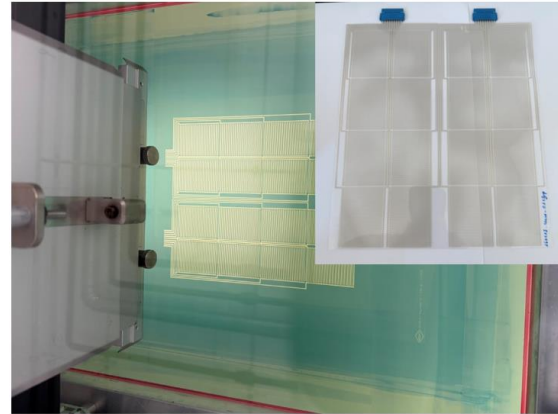


Abb. 4 Siebdruck des Leckagesensor Array

Test der Leckagesensorik

Ziel der Leckagesensorik ist der Nachweis
verschiedener Flüssigkeiten ab einer
Tropfengröße von 100 μl . Um diese Menge
gezielt aufzubringen, wurde eine Eppendorf-
Pipette mit Einwegspitzen verwendet, dadurch
ist eine Flüssigkeitsabgabe von definiert 100 μl
gewährleistet. Exemplarisch ist in Abb. 5 ein
100 μl großer Wassertropfen auf einem
Leckagesensor aus elastischem Silikon
dargestellt.

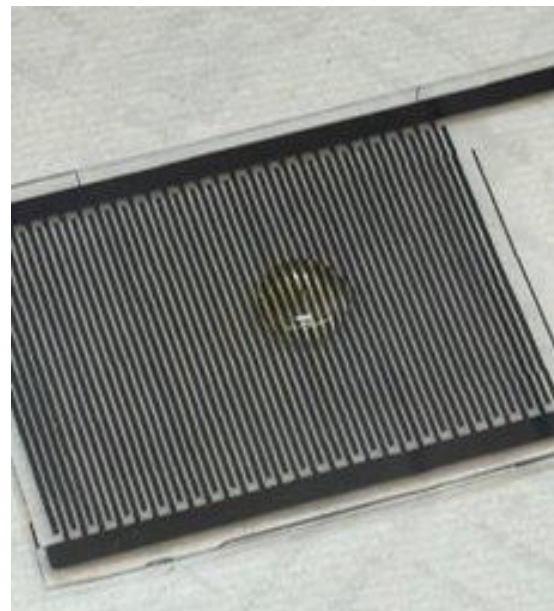


Abb. 5 Elastischer Silikon Leckagesensor mit
100 μl Wassertropfen.

Als Messelektronik kam die weiter unten
beschriebene Testelektronik zum Einsatz. Die

Daten wurden über eine serielle Schnittstelle kontinuierlich während des Auftropfens ausgelesen, abgespeichert und anschließend in Origin 2024 ausgewertet. Auf eine Umrechnung des kapazitiven Messsignals wurde verzichtet, da im Wesentlichen die Signaländerung während einer Leckage von Interesse ist, daher sind die Messdaten als Rohdaten-Ausgangssignal dargestellt.

Die kapazitiv arbeitende Struktur des Leckagesensors wird elektronisch betrachtet in einen RC-Schwingkreis geschaltet, sodass ihre Kapazität über die Frequenz des Schwingkreises indirekt ausgewertet werden kann. Dabei zählt der Mikrocontroller die Anzahl der Be- und Entladezyklen innerhalb einer bestimmten Zeit. Zur besseren Vergleichbarkeit wird das Ausgangssignal jedes Sensors auf 0 genormt. In den Abbildungen ist jeweils der Rohdatenwert minus der Ausgangswert dargestellt, d. h. erhöht sich die Kapazität der Sensorstruktur durch das Auftropfen der Flüssigkeit, so verringert sich die gemessene Anzahl an Ladezyklen und der dargestellte Wert wird negativ.

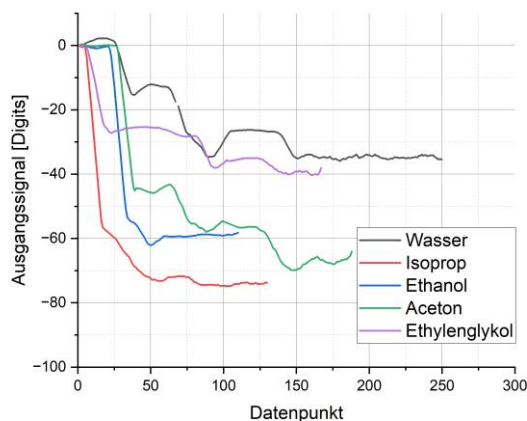


Abb. 6 Messsignal auf dem Leckagesensor von 100µl Wasser und gängige Lösemittel

In Abb. 6 ist die Messung für Wasser und gängige Lösemittel dargestellt. Das Messsignal wurde auf 0 normiert und direkt während der ersten Sekunden des Auftropfens aufgezeichnet. Da sowohl Wasser als auch Lösemittel eine Permittivität > 10 besitzen, ändert sich – wie zu erwarten – das kapazitive Messsignal. Allerdings spielt hier auch die Benetzungsfähigkeit eine Rolle: Wasser bildet einen Tropfen wie in Abb. 4, während die Lösemittel stärker zerfließen und dadurch eine größere Fläche der Interdigitalstruktur überdecken. Dies erklärt, warum die Signal-

änderungen der Lösemittel größer ausfallen als die des reinen Wassers.

Eine Änderung des Messsignals um mindestens 10 Digits während des Auftropfens kann hier festgestellt werden und als Kriterium für eine sichere Detektion herangezogen werden.

Tab. 1 Permittivität der getesteten Medien – Alle Werte beziehen sich auf Standardangaben aus der Literatur [2][3].

Medium	Permittivität
Wasser	Ca. 80
Isopropanol	Ca. 22
Ethanol	Ca. 20
Aceton	Ca. 26
Ethylenglykol	Ca. 39
Essigsäure	Ca. 8
Ammoniak	Ca. 28
Natronlauge	Ca. 80
Benzin	Ca. 2-3
Diesel	Ca. 2-3
Motoröl / Altöl	Ca. 2.3

In Abb. 7 ist das Messsignal für Essigsäure, Natronlauge und Ammoniak dargestellt. Hier hängt die Permittivität ebenfalls stark von der Konzentration ab, wobei die Permittivität der Natronlauge im Bereich des Wassers liegt, Essigsäure bei ca. 8 bzw. Ammoniak in verdünnter Konzentration bei ca. 28. Auch hier spielt die Benetzung eine Rolle; die Signaländerung von > 10 Digits lässt auch hier eine sichere Detektion zu.

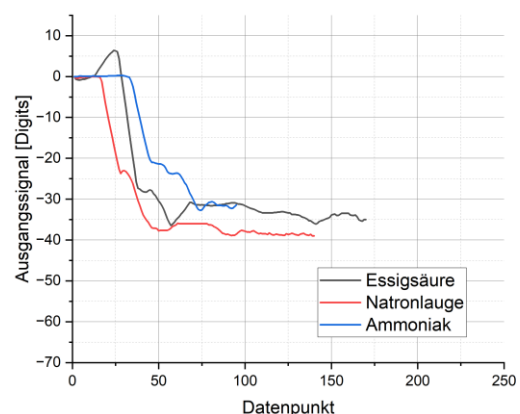


Abb. 7 Messsignal des Leckagesensors für 100µl Säuren und Laugen

Als weitere Substanzklasse wurden in Abb. 8 Kraftstoffe wie Benzin und Diesel sowie konventionelles Motoröl bzw. Altöl getestet. Der deutlichste Unterschied bei den Ergebnissen der Kraftstoffe und Öle ist, dass die

Signaländerungen beim Auftropfen der Medien deutlich geringer sind und mit < 10 Digits nicht sicher detektiert werden können. Dies ist auf die geringeren Permittivitäten der Medien zurückzuführen, die im Bereich zwischen 2 und 3 liegen. Hier würden gegebenenfalls Anpassungen der Interdigitalstruktur, etwa feinere Abstände oder dünnere Dielektrika, helfen. Ebenfalls möglich wäre der Einsatz anderer Materialien als Silikon oder Polyimid für das Grundsubstrat, da deren Permittivität ähnlich der des getesteten Mediums ist und somit nur eine geringe Verschiebung des elektrischen Feldes für die kapazitive Messung ermöglicht. Neben diesen Änderungen am Sensor selbst lässt sich aber sagen, daß bei einer höheren Überdeckungsfläche auch die Änderung der Kapazität deutlicher ausfällt.

Auffällig ist der Signalverlauf von Benzin: Dieser steigt nach kurzzeitigem Abfall wieder in den positiven Bereich an. Benzin durchdringt aufgrund seiner chemischen Struktur die Silikonschutzschicht und wirkt direkt auf die Interdigitalelektroden ein. Dabei wird die kapazitive Messung vermutlich dahingehend gestört, dass kein ausreichend großer Isolationswiderstand mehr vorhanden ist und vermutlich Kurzschlüsse zwischen den Elektroden zu diesem Verhalten führen.

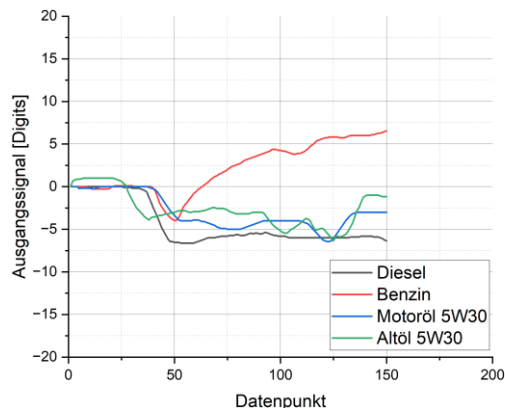


Abb. 8 Messsignal des Leckagesensors für 100µl Kraftstoffe und Öle

Auswerteelektronik

Der Sensor selbst wird durch eine energieeffiziente Elektronik komplettiert (Abb.9), welche das kapazitive Messsignal auswertet. Als moderne energieeffiziente System-on-Chip-Plattform (Mikrocontroller inkl. Sub-GHz-Transceiver) kommt EFR32FG23 von Silicon Labs zum Einsatz. Die Detektion arbeitet ereignis-gesteuert: Der erste auftreffende Tropfen weckt dabei die Elektronik auf und aktiviert die Auswertung und Datenübertragung. Das Betriebskonzept sieht dabei konkret auch

die Implementierung eines Heart-beats, die Übertragung des Batteriestatus und die Möglichkeit einer Ferndiagnose vor.

Das Projektteam setzte sich zum Ziel, den Betrieb der Elektronik mit Batterien über ca. 10 Jahre zu gewährleisten. Damit ist man unabhängig von einer vorhandenen Energieinfrastruktur und auch die Installation an schwer zugängliche Orte ist problemlos möglich („install and forget“). Das am Fraunhofer IIS entwickelte Funkprotokoll mioty® sorgt für eine energieeffiziente und zuverlässige Datenübertragung und integriert sich problemlos in vorhandene Überwachungsinfrastrukturen.



Abb. 9 Sensorelektronik mit Stromversorgung

mioty® ist ein LPWAN-Funkstandard, der mit Telegram-Splitting (TSMA) arbeitet: Jedes Telegramm wird in viele kurze Subpakete zerlegt und zeitlich/frequenzmäßig verteilt. Dadurch steigt die Robustheit gegenüber Störungen und Kollisionen, bei zugleich sehr geringer On-Airtime. mioty® nutzt lizenzfreie Sub-GHz-Bänder (z. B. EU 868 MHz SRD) und erfüllt die jeweiligen SRD-Vorgaben inklusive Duty-Cycle-Beschränkungen; die kurze Sendezeit unterstützt eine hohe Netzdichte und Skalierbarkeit.

Für batteriebetriebene Geräte ist mioty® besonders energieeffizient: Der Betrieb ist uplink-zentriert, die Sendephasen sind extrem kurz und lange Schlafzeiten dominieren, was typischerweise Batterielaufzeiten von 10+ Jahren bei wenigen Meldungen pro Tag ermöglicht. In der hier beschriebenen Lösung konnte ein Energiebedarf pro Messvorgang und Übertragung von etwa 16 µWh erreicht werden bei einem Standby-Stromverbrauch von ca. 34 µA (1 Sensor) bzw. ca. 46 µA (Sensor-Array mit 16 Sensoren).

Die Reichweite beträgt je nach Umgebung mehrere Kilometer bei tiefer Gebäudedurchdringung. Die Netztopologie ist sternförmig mit zentralen Basisstationen die eine Vielzahl von batteriebetriebenen Sensorknoten bedienen können. Die Anbindung ans Backend erfolgt IP-seitig (z. B. MQTT/REST), wodurch sich mioty® nahtlos in Leit-, Cloud- und Monitoring-Systeme integrieren lässt. Downlinks sind verfügbar, werden aber bewusst sparsam

genutzt (Konfiguration). Sicherheitsfunktionen wie Verschlüsselung und Integritätsschutz (z. B. AES-128) sowie Geräteauthentisierung werden unterstützt.

Für den Leckagesensor empfiehlt sich ein ereignisbasierter Uplink (Alarm sofort senden, optional kurzzeitige Wiederholung zur Erhöhung der Zustellsicherheit), ergänzt um seltene Heartbeats (z. B. täglich) mit Batteriestatus und Selbsttest-Ergebnissen. Konfigurierbare Schwellwerte und Haltezeiten helfen, Fehlalarme durch Kondensat oder Spritzwasser zu vermeiden. Diese Betriebsweise vereint schnelle Alarmierung, hohe Zuverlässigkeit und sehr niedrigen Energieverbrauch.



Abb. 10 Sensorarray mit angebundener Elektronik während Tests

Schlussfolgerungen und Ausblick

In dieser Arbeit wurde eine auf einem kapazitiv arbeitenden Messprinzip basierende Leckagesensorik in unterschiedlichen Ausbaustufen erfolgreich getestet. Beginnend mit Einzelsensoren aus vollständig elastischem Silikonelastomer für eine flexible Anpassung an die Anwendung, über hybride Aufbauten mittels Polyimid-Grundsustrat, per Lasergravur hergestellten Elektroden und elastischer Übersichtung, bis hin zu flexiblen Sensorarrays in der Größe eines DIN-A4-Blattes (Abb. 10), die mittels Siebdruck hergestellt wurden.

Dabei zeigte sich gerade für das Siebdruckverfahren und die Elektroderung mittels Lasergravur eine potentiell kostengünstige Skalierung in größere Stückzahlen, da diese Prozesse industriell etabliert sind.

Die Leckagesensorik wurde erfolgreich mit Wasser, Lösemitteln, Säuren und Laugen getestet, und der Nachweis einer Detektion ab dem ersten Tropfen (100 µl) erbracht. Die Detektion von Medien mit einer dem Sensor ähnlichen Permittivität ist zwar möglich, aber eine sichere Detektion ist erst ab größeren Mengen oder mit geänderten Sensorparametern möglich. Die dem Sensor zugehörige Elektronik wurde erfolgreich mittels einer kapazitiven Engine umgesetzt und die Übertragung der Tropfenevents mittels des mioty®-Standards erfolgreich realisiert.

In folgenden Arbeiten sollen die Einsatzbereiche auch unter erschwerten Bedingungen über längere Zeiträume validiert werden. Durch die flexible Ausgestaltung der Sensorik sollen Anwendungen in der Prozessmesstechnik, beim Überwachen schwer zugänglicher Baustrukturen im privaten oder gewerblichen Bereich und gegebenenfalls Anwendungen im Agrarbereich evaluiert werden.

Quellen

- [1] <https://www.gdv.de/gdv/statistik/statistiken-zur-deutschen-versicherungswirtschaft-uebersicht/schaden-und-unfallversicherung#sachversicherung>
- [2] Rumble, J. R. (Ed.). (2023). *CRC Handbook of Chemistry and Physics* (104th ed.). CRC Press. Abschnitt: Dielectric Constants of Pure Liquids.
- [3] Perry, R. H., Green, D. W., & Maloney, J. O. (Eds.). (2019). *Perry's Chemical Engineers' Handbook* (9th ed.). McGraw-Hill.

Latenz- und Energievergleich drahtloser Protokolle für haptische VR-Systeme

Arne Kohlhaupt¹, Ulrike Steinmann¹

¹ IFAT, Universitätspl. 2, 39106 Magdeburg, Germany

Zusammenfassung

Haptisches Feedback in Virtual-Reality-Systemen erfordert eine geringe Übertragungslatenz, um eine hohe Immersivität zu gewährleisten und Cybersickness zu vermeiden. Je geringer die Übertragungslatenz, desto größer ist die verfügbare Zeitreserve für die Ansteuerung haptischer Aktuatoren und den Einsatz komplexerer Feedbacksysteme. Die Übertragungslatenz bildet dabei einen entscheidenden Teil der Gesamtsystemlatenz, da die entsprechende Steuernachricht vollständig übertragen worden sein muss, bevor ein Aktuator ausgelöst werden kann. In der vorliegenden Arbeit werden die drahtlosen Kommunikationsprotokolle ESP-NOW, WLAN-UDP und Bluetooth Low Energy 5.3 hinsichtlich ihrer Round-Trip-Zeit sowie ihres Stromverbrauchs auf dem Mikrocontroller ESP32-C6 messtechnisch verglichen. Der ESP32-C6 wurde als Hardwareplattform gewählt, da er alle drei Protokolle nativ unterstützt und damit einen direkten Vergleich unter identischen Hardwarebedingungen ermöglicht. Je Protokoll wurden 1000 Messungen mit einer Sendepause von 100 ms durchgeführt. Der Stromverlauf wurde auf der Client-Seite mit dem Power Profiler Kit 2 von Nordic Semiconductor bei einer Versorgungsspannung von 3,3 V aufgezeichnet. ESP-NOW erzielte mit einer mittleren Round-Trip-Zeit von 2,626 ms und einer Standardabweichung von 0,607 ms die geringste und reproduzierbarste Latenz. WLAN-UDP erreichte einen ähnlichen Mittelwert von 3,941 ms, wies jedoch mit einer Standardabweichung von 2,443 ms und einem 95. Perzentil von 9,311 ms eine für Echtzeitsysteme problematische Streuung auf. Bluetooth Low Energy 5.3 zeigte mit 13,007 ms die höchste Latenz, überzeugte jedoch durch die geringste Streuung aller Protokolle sowie einen um rund 41 Prozent reduzierten mittleren Stromverbrauch von 50,430 mA gegenüber ESP-NOW mit 85,100 mA und WLAN-UDP mit 84,145 mA. Alle drei Protokolle unterschreiten die für haptisches Feedback kritische Wahrnehmungsschwelle von 20 ms als One-Way-Latenz deutlich, womit die grundlegende Eignung aller Protokolle für den betrachteten Anwendungsfall bestätigt wird.

Keywords: Haptisches Feedback, Übertragungslatenz, ESP32-C6, drahtlose Kommunikation, Virtual Reality

Einleitung und Motivation

Haptisches Feedback in Virtual-Reality-Systemen stellt hohe Anforderungen an die Übertragungslatenz drahtloser Kommunikationsprotokolle. Je geringer die Übertragungslatenz, desto größer ist die verfügbare Zeitreserve für die Ansteuerung haptischer Aktuatoren. Dies ermöglicht den Einsatz komplexerer Feedbacksysteme und erweitert das Spektrum nutzbarer haptischer Technologien. Eine hohe Übertragungslatenz führt hingegen zu einer verminderten Immersivität und kann Cybersickness begünstigen. Bevor ein haptischer Aktuator ausgelöst werden kann, muss die entsprechende Steuernachricht vollständig übertragen worden sein. Die Übertragungslatenz bildet damit einen entscheidenden Teil der Gesamtsystemlatenz.

Im Rahmen des VRHap-Projekts wird ein eigenes Aktuatoriksystem auf Basis des ESP32-C6 entwickelt, das Anwendungen in Training, Therapie und Diagnostik adressiert. Für den Aufbau eines Host-Client-Systems mit bis zu zehn gleichzeitigen Clients ist die Kenntnis der protokollspezifischen Latenzen und des Energieverbrauchs eine wesentliche Voraussetzung für die Systemauslegung, da es sich um eine mobile, akkubetriebene Anwendung handelt. In der vorliegenden Arbeit werden daher die Protokolle ESP-NOW, WLAN-UDP und BLE 5.3 hinsichtlich ihrer Round-Trip-Time (RTT) sowie ihres Energieverbrauchs auf dem ESP32-C6 messtechnisch verglichen.

Technische Grundlagen

Die Latenz eines Netzwerks beschreibt die Zeit, die eine Nachricht vom Sender zum Empfänger benötigt, und setzt sich aus Ausbreitungs-, Übertragungs- und Wartezeit zusammen [1]. Bei kleinen Nachrichten dominieren die protokollspezifischen Verzögerungsanteile, da die Übertragungsverzögerung vernachlässigbar klein wird [1]. Da die Übertragungsdistanz zwischen Host- und Client-ESP32 typischerweise unter einem Meter beträgt, ist die Ausbreitungsverzögerung elektromagnetischer Wellen vernachlässigbar klein gegenüber den protokollspezifischen Verzögerungsanteilen [1]. Hohe Latenz in VR-Systemen kann Cybersickness begünstigen [2], während für haptisches Feedback Wahrnehmungsschwellen von 20 bis 200 ms beschrieben werden [3].

Als Hardwareplattform wurde der ESP32-C6 gewählt, da er WiFi 6, BLE 5.3 und ESP-NOW nativ unterstützt und damit einen direkten Protokollvergleich unter identischen Hardwarebedingungen ermöglicht [4]. ESP-NOW ist ein von Espressif entwickeltes verbindungsloses Protokoll, das Anwendungsdaten direkt in herstellerspezifischen Action Frames des IEEE-802.11-Standards überträgt, ohne dass ein Verbindungsaufbau oder ein Access Point erforderlich ist [5]. Die Standardbitrate beträgt 1 Mbps bei einer maximalen Nutzlast von 250 Bytes pro Paket [5].

WLAN-UDP nutzt das verbindungslose User Datagram Protocol über einen IEEE-802.11-Access-Point. Durch den Verzicht auf Verbindungsaufbau und Bestätigungsmechanismen ist UDP gegenüber TCP latenzärmer [1]. Gemäß IEEE-802.11-Standard sendet ein Access Point kontinuierlich Beacon Frames in einem Standardintervall von 102,4 ms [6], was einen erhöhten Grundenergieverbrauch auch ohne aktive Datenübertragung verursacht.

BLE 5.3 kommuniziert in festen Verbindungsintervallen, in denen Sender und Empfänger Daten austauschen und anschließend in einen inaktiven Zustand wechseln können [7]. Dieses Prinzip reduziert den Energieverbrauch erheblich, begrenzt jedoch die minimale erreichbare Latenz strukturell nach unten, da Nachrichten erst im nächsten Verbindungsintervall übertragen werden können.

Methodik und Versuchsaufbau

Für die RTT-Messung wurden zwei ESP32-C6 Dev-Boards eingesetzt: ein Host-Board (H) und ein Client-Board (C), wie in Abbildung 1 dargestellt.

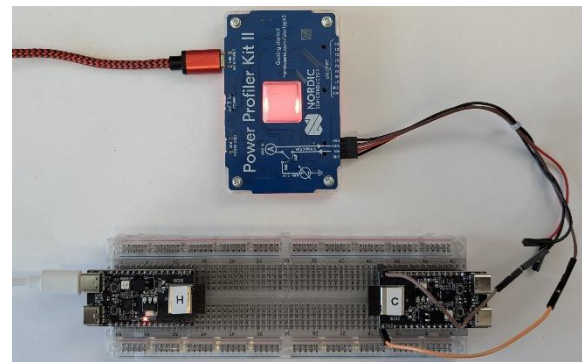


Abb. 1: Versuchsaufbau mit ESP32-C6 Host (H), ESP32-C6 Client (C) und Power Profiler Kit 2 (Nordic Semiconductor)

Der Host sendet eine 18-Byte-Nachricht („255255255255255255“), die einem realistischen Motorbefehl für das haptische Feedbacksystem entspricht. Die Nachricht ermöglicht die individuelle Ansteuerung von sechs unabhängigen Aktuatoren und wird vom Client als Echo zurückgesendet, woraus der Host die RTT berechnet. Je Protokoll wurden 1000 Messungen mit einer Sendepause von 100 ms zwischen den Messungen durchgeführt. Die Strommessung erfolgte auf der Client-Seite mit dem Power Profiler Kit 2 (Nordic Semiconductor) bei einer eingestellten Versorgungsspannung von 3,3 VDC. Zur Auswertung wurde jeweils ein stabiles Messfenster nach dem Einlaufen der Verbindung gewählt, um Ein- und Auslaufeffekte auszuschließen.

Round-Trip-Time

Für den Vergleich der Übertragungsprotokolle wurden je 1000 RTT-Messungen mit einer Sendepause von 100 ms zwischen den Messungen aufgezeichnet.

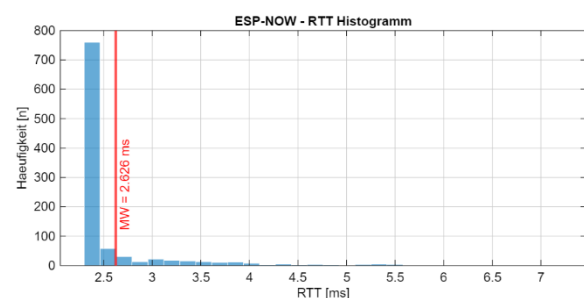


Abb. 2: RTT von ESP-NOW (n=1000)

Abbildung 2 zeigt exemplarisch die RTT-Verteilung von ESP-NOW. Die Verteilung weist ein stabiles Grundniveau bei ~2,39 ms auf, mit einer rechtsschiefen Charakteristik durch sporadisch auftretende erhöhte Latenzwerte. Dieser Verlauf ist charakteristisch für verbindungslose Protokolle, bei denen gelegentliche Kanalzugriffskonflikte zu erhöhten Latenzwerten führen. Mit einem Mittelwert von 2,626 ms, einer Standardabweichung von 0,607 ms und einem 95. Perzentil von 3,866 ms zeigt ESP-NOW eine für Echtzeitsysteme gut geeignete Latenzcharakteristik.

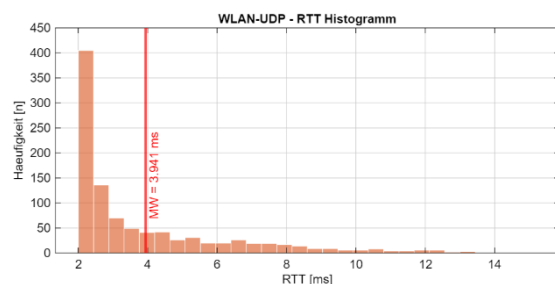


Abb. 3: RTT von WLAN-UDP (n=1000)

Abbildung 3 zeigt die RTT-Verteilung von WLAN-UDP. Im direkten Vergleich zu ESP-NOW ist die deutlich ausgeprägtere Streuung erkennbar: Während ESP-NOW den Großteil der Messungen in einem engen Bereich um ~2,39 ms konzentriert, erstreckt sich die WLAN-UDP-Verteilung mit einem langen Schwanz bis über 15 ms. Dies spiegelt sich in der Standardabweichung von 2,443 ms wider, die mehr als viermal so hoch ist wie bei ESP-NOW (0,607 ms). Das 95. Perzentil liegt bei 9,311 ms, was bedeutet, dass 5 % aller Pakete eine Latenz von über 9 ms aufweisen. Diese Variabilität ist auf die Zugriffssteuerung des vollständigen IEEE-802.11-Stacks zurückzuführen, bei dem konkurrierende Prozesse wie Beacon-Frame-Verwaltung zu unvorhersehbaren Verzögerungsspitzen führen können.

Abbildung 4 stellt die mittlere RTT aller drei Protokolle gegenüber.

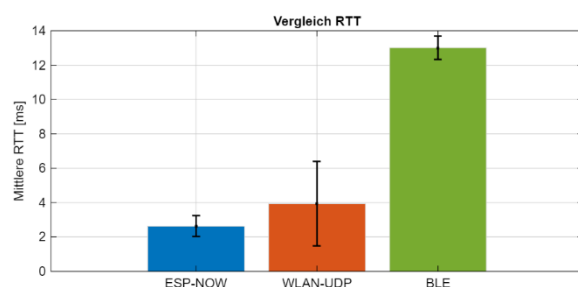


Abb. 4: Vergleich RTT alle drei Protokolle

ESP-NOW erzielt mit 2,626 ms den geringsten Mittelwert, gefolgt von WLAN-UDP mit 3,941 ms. BLE 5.3 weist mit 13,007 ms die höchste Latenz auf, zeigt jedoch mit einer Standardabweichung von 0,663 ms die geringste Streuung aller Protokolle. WLAN-UDP hingegen weist trotz ähnlichem Mittelwert wie ESP-NOW eine deutlich höhere Standardabweichung von 2,443 ms und ein 95. Perzentil von 9,311 ms auf, was für Echtzeitsysteme problematisch sein kann.

Tab. 1: Statistische Kenngrößen der RTT-Messungen (n = 1000)

Kenngröße	ESP-NOW [ms]	WLAN-UDP [ms]	BLE 5.3 [ms]
Mittelwert	2,626	3,941	13,007
Standardabweichung	0,607	2,443	0,663
Median	2,398	2,667	12,845
Modalwert	2,390	2,170	12,370
Minimum	2,385	2,117	12,310
Maximum	7,195	15,113	15,843
95. Perzentil	3,866	9,311	14,944
99. Perzentil	5,460	12,521	15,749

Tabelle 1 fasst alle statistischen Kenngrößen der RTT-Messungen zusammen. Alle drei Protokolle unterschreiten als One-Way-Latenz (RTT/2) die für haptisches Feedback kritische Wahrnehmungsschwelle von 20 ms deutlich.

Energieverbrauch

Der Energieverbrauch wurde auf der Client-Seite des ESP32-C6 mit dem Power Profiler Kit 2 bei einer Versorgungsspannung von 3,3 V aufgezeichnet. Zur Auswertung wurde jeweils ein stabiles Messfenster nach dem Einlaufen der Verbindung gewählt: 98 s für ESP-NOW, 101 s für WLAN-UDP und 112 s für BLE 5.3.

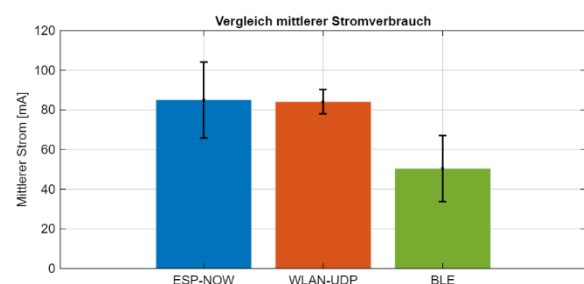


Abb. 5: Vergleich mittlerer Stromverbrauch

Abbildung 5 stellt den mittleren Stromverbrauch aller drei Protokolle gegenüber. ESP-NOW und WLAN-UDP weisen mit 85,100 mA bzw. 84,145 mA nahezu identische Werte auf. BLE 5.3 zeigt mit 50,430 mA einen um rund 41 % reduzierten Stromverbrauch gegenüber den beiden WLAN-basierten Protokollen. Bemerkenswert ist die deutlich geringere Standardabweichung von WLAN-UDP (6,108 mA) gegenüber ESP-NOW (19,155 mA), da der ESP32 im softAP-Modus durch den kontinuierlichen Beacon-Frame-Betrieb keinen ausgeprägten Ruhezustand zwischen den Paketen erreicht. Die erhöhte Standardabweichung bei BLE 5.3 (16,795 mA) ist auf die periodischen Aktivitätspulse im Verbindungsintervall zurückzuführen.

Tab. 2: *Kenngößen des Energieverbrauchs (Client-Seite, 3,3 V)*

<i>Kenngöße</i>	<i>ESP-NOW</i>	<i>WLAN-UDP</i>	<i>BLE 5.3</i>
<i>Mittlerer Strom</i>	85,1 mA	84,145 mA	50,430 mA
<i>Mittlere Leistung</i>	280,8 mW	277,7 mW	166,4 mW
<i>Energie pro Paket</i>	28,848 µJ	28,883 µJ	18,808 µJ
<i>Ladung pro Paket</i>	2428,3 µAh	2431,3 µAh	1583,2 µAh
<i>Ladung gesamt</i>	2,317 mAh	2,361 mAh	1,569 mAh

Tabelle 2 fasst alle Kenngößen des Energieverbrauchs zusammen. Für akkubetriebene mobile Anwendungen stellt der deutlich geringere Stromverbrauch von BLE 5.3 einen relevanten Vorteil dar, sofern die höhere Latenz von ~13 ms im jeweiligen Anwendungsfall tolerierbar ist.

Diskussion

ESP-NOW erzielt mit einem RTT-Mittelwert von 2,626 ms die geringste Übertragungsverzögerung und ist damit für latenzkritische Anwendungen am besten geeignet. Die Standardabweichung von 0,607 ms ist gering, jedoch treten vereinzelt Ausreißer bis 7,195 ms auf. WLAN-UDP weist mit einem Mittelwert von 3,941 ms eine ähnliche Grundlatenz auf, jedoch ist die Streuung mit einer Standardabweichung von 2,443 ms und einem 95. Perzentil von 9,311 ms für Echtzeitsysteme problematisch. BLE 5.3 zeigt mit 13,007 ms den höchsten Mittelwert,

gleichzeitig aber die geringste Streuung aller drei Protokolle.

Für das haptische VR-System ist nicht nur der Mittelwert entscheidend, sondern auch die Vorhersagbarkeit der Latenz. Unerwartete Verzögerungsspitzen können das haptische Feedback inkonsistent machen und die Immersion beeinträchtigen. Unter diesem Gesichtspunkt zeigt ESP-NOW das beste Gesamtbild aus niedrigem Mittelwert und vertretbarer Streuung.

Beim Energieverbrauch ergibt sich ein klares Bild: BLE 5.3 verbraucht mit 50,430 mA mittlerem Strom rund 41 % weniger als ESP-NOW (85,100 mA) und WLAN-UDP (84,145 mA), die nahezu identische Werte aufweisen. Bemerkenswert ist die nahezu identische Energieaufnahme von ESP-NOW und WLAN-UDP trotz des deutlich komplexeren IEEE-802.11-Stacks bei WLAN-UDP. Der kontinuierliche Beacon-Frame-Betrieb im softAP-Modus gleicht den potenziellen Energievorteil durch seltenere Nutzdatenübertragungen aus.

Alle drei Protokolle unterschreiten als One-Way-Latenz (RTT/2) die für haptisches Feedback als kritisch geltende Wahrnehmungsschwelle von 20 ms deutlich. Für den geplanten Einsatz im VRHap-System empfiehlt sich ESP-NOW als primäres Protokoll, da es die geringste Latenz bei vertretbarer Streuung bietet. Für Szenarien mit strikten Energiebeschränkungen stellt BLE 5.3 eine valide Alternative dar, sofern die höhere Latenz von ~13 ms toleriert werden kann. WLAN-UDP bietet gegenüber den beiden anderen Protokollen weder einen Latenz- noch einen Energievorteil.

Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurden die Kommunikationsprotokolle ESP-NOW, WLAN-UDP und BLE 5.3 hinsichtlich ihrer Eignung für den Einsatz in latenzoptimierten haptischen VR-Systemen messtechnisch untersucht und verglichen. Als Hardwareplattform wurde der ESP32-C6 von Espressif Systems gewählt, da er alle drei Protokolle nativ unterstützt und damit einen direkten Vergleich unter identischen Hardwarebedingungen ermöglicht. Die Bewertung erfolgte anhand zweier zentraler Kriterien: der Round-Trip-Zeit als Maß für die Übertragungslatenz sowie dem mittleren Stromverbrauch als Maß für die Energieeffizienz im mobilen Betrieb.

Die RTT-Messungen zeigen, dass ESP-NOW mit einem Mittelwert von 2,626 ms und einer Standardabweichung von 0,607 ms die

geringste und zugleich gut reproduzierbare Latenz erzielt. WLAN-UDP erreicht zwar einen ähnlichen Modalwert, weist jedoch mit einer Standardabweichung von 2,443 ms und einem 95. Perzentil von 9,311 ms eine für Echtzeitsysteme problematische Streuung auf. BLE 5.3 zeigt mit 13,007 ms die höchste mittlere Latenz, zeichnet sich jedoch durch die geringste Streuung aller Protokolle aus, da das feste Verbindungsintervall sowohl die minimale als auch die maximale Übertragungszeit begrenzt. Hinsichtlich des Energieverbrauchs zeigt BLE 5.3 mit einem mittleren Stromverbrauch von 50,430 mA einen um rund 41 % reduzierten Verbrauch gegenüber ESP-NOW (85,100 mA) und WLAN-UDP (84,145 mA), die nahezu identische Werte aufweisen. Alle drei Protokolle unterschreiten die für haptisches Feedback kritische One-Way-Latenz von 20 ms deutlich, womit die grundlegende Eignung aller Protokolle für den betrachteten Anwendungsfall bestätigt wird. Für das VRHap-System empfiehlt sich ESP-NOW als primäres Protokoll für die drahtlose Kommunikation zwischen Host und Client, da es die beste Kombination aus geringer Latenz, niedrigem Jitter und einfacher Implementierung bietet. WLAN-UDP bietet gegenüber den beiden anderen Protokollen weder einen Latenz- noch einen Energievorteil und eignet sich vor allem dann, wenn eine direkte Integration in bestehende WLAN-Infrastrukturen erforderlich ist oder größere Datenmengen übertragen werden sollen. BLE 5.3 stellt aufgrund seines deutlich geringeren Stromverbrauchs eine energieeffiziente Alternative für akkubetriebene Anwendungen dar, sofern die höhere und durch das Verbindungsintervall bedingte Latenz von ~13 ms im jeweiligen Anwendungsfall tolerierbar ist. Die vorliegende Untersuchung beschränkt sich auf eine Point-to-Point-Verbindung zwischen einem Host- und einem Client-ESP32. In einem nächsten Schritt soll die End-to-End-Latenz zwischen der Meta Quest 3 und dem ESP32-Host über UDP gemessen werden, um die Gesamtverzögerung des VRHap-Systems vom Rendering-Zeitpunkt bis zur haptischen Ausgabe vollständig zu charakterisieren. Erste Vorversuche deuten auf eine One-Way-Latenz von ~14,7 ms für diesen Pfad hin, sodass die Gesamtlatenz Quest 3 → Host → Client bei Verwendung von ESP-NOW auf der letzten Meile mit ~16 ms unterhalb der für haptisches Feedback kritischen Wahrnehmungsschwelle von 20 ms verbleibt. Darüber hinaus soll die Skalierbarkeit der Protokolle auf mehrere gleichzeitige Clients evaluiert werden, da im realen Betrieb beide Hände unabhängig voneinander mit haptischem Feedback versorgt

werden müssen. ESP-NOW bietet hier durch seine Mesh-Netzwerk-Fähigkeit und die Unterstützung von bis zu 20 Peer-Geräten [8] besondere Vorteile und ermöglicht den Aufbau komplexerer Aktuatoriksysteme ohne zentrale WLAN-Infrastruktur. Abschließend ist die Entwicklung eines vollständigen haptischen Aktuatoriksystems auf Basis des ESP32-C6 geplant, das die in dieser Arbeit gewonnenen Erkenntnisse direkt umsetzt und eine vollständige Integration mit der Meta Quest 3 für Anwendungen in Training, Therapie und Diagnostik ermöglicht.

References

- [1] C. Baun, *Computernetze kompakt*. Springer Berlin Heidelberg, Berlin, Heidelberg; 2022.
- [2] J.-P. Stauffert, F. Niebling, M. E. Latoschik, Latency and Cybersickness: Impact, Causes, and Measures. A Review, *Frontiers in Virtual Reality* 1, 582204 (2020); doi: 10.3389/frvir.2020.582204.
- [3] H. Alsuradi, M. Eid, An ensemble deep learning approach to evaluate haptic delay from a single trial EEG data, *Frontiers in Robotics and AI* 9, 1013043 (2022); doi: 10.3389/frobt.2022.1013043.
- [4] esp32-c6_datasheet_en, https://documentation.espressif.com/esp32-c6_datasheet_en.pdf. Accessed May 18, 2026.
- [5] ESP-NOW - ESP32 - — ESP-IDF Programming Guide latest documentation, https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/network/esp_now.html. Accessed May 18, 2026.
- [6] IEEE Standard for Information Technology--Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks--Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications.
- [7] P. Bulić, G. Kojek, A. Biasizzo, Data Transmission Efficiency in Bluetooth Low Energy Versions, *Sensors (Basel, Switzerland)* 19 (2019); doi: 10.3390/s19173746.

- [8] J. A. Cujilema Paguay, G. A. Hidalgo Brito, D. L. Hernandez Rojas, J. J. Cartuche Calva, Secure home automation system based on ESP-NOW mesh network, MQTT and Home Assistant platform, *IEEE Latin America Transactions* 21, 829–838 (2023); doi: 10.1109/TLA.2023.10244182.

Titel: ThinKIsense – KI in batteriebetriebenen Sensorknoten für die Prozessautomatisierung

Schwerpunkt: B3

Autor(en):

Dr. Stefan von Dosky, Stefan Klehr, Dr. Gilbert Erdler

Anschrift(en):

stefan.von_dosky@siemens.com

0172 1556960

Siemens AG, DI PA MI ST IP

Östliche Rheinbrückenstr. 50, 76187 Karlsruhe

In Kooperation mit NXP Semiconductors Germany GmbH, Netzsch Pumpen & Systeme GmbH, Technische Universität Darmstadt, FZI Forschungszentrum Informatik

Gefördert durch VDI/VDE Innovation + Technik GmbH, Projektträger für das Bundesministerium für Forschung, Technologie und Raumfahrt
Mikroelektronik, Supercomputing, Robotik

1. Motivation und Hintergrund

Die Industrie 4.0 erfordert intelligente, energieeffiziente und resiliente Systeme. Künstliche Intelligenz (KI) ermöglicht prädiktive Analysen und Echtzeitentscheidungen, insbesondere für Anwendungen wie **Predictive Maintenance**. Der aktuelle Stand der Technik nutzt batteriebetriebene Sensoren, die Rohdaten in die Cloud übertragen – dies verursacht hohen Energieverbrauch und begrenzt die Batterielaufzeit.

Projektziel: Entwicklung des **ThinKIsense Sensors**, der KI-Inferenz direkt am Sensorknoten durchführt. Nur verdichtete Ergebnisse (KPIs) werden übertragen. Ziel ist eine Batterielaufzeit von **8 Jahren**.

2. Technische Herausforderungen

- **Energieeffizienz:** Minimierung des Energieverbrauchs für Datenübertragung und Inferenz.
 - **Hardware-Ansätze:** Vergleich von drei Beschleuniger-Architekturen:
 - Klassische neuronale Netze auf ARM-Mikrocontroller (MCX-N947)
 - RISC-V mit ISA-Erweiterungen für KI
 - Neuromorphe Beschleuniger (Spiking Neural Networks)
 - **Sicherheit:** Sichere Übertragung und Steuerung intelligenter Sensoren.
 - **Demonstrator:** Funktionsfähige Förderpumpe mit Echtzeit-Datenvisualisierung.
-

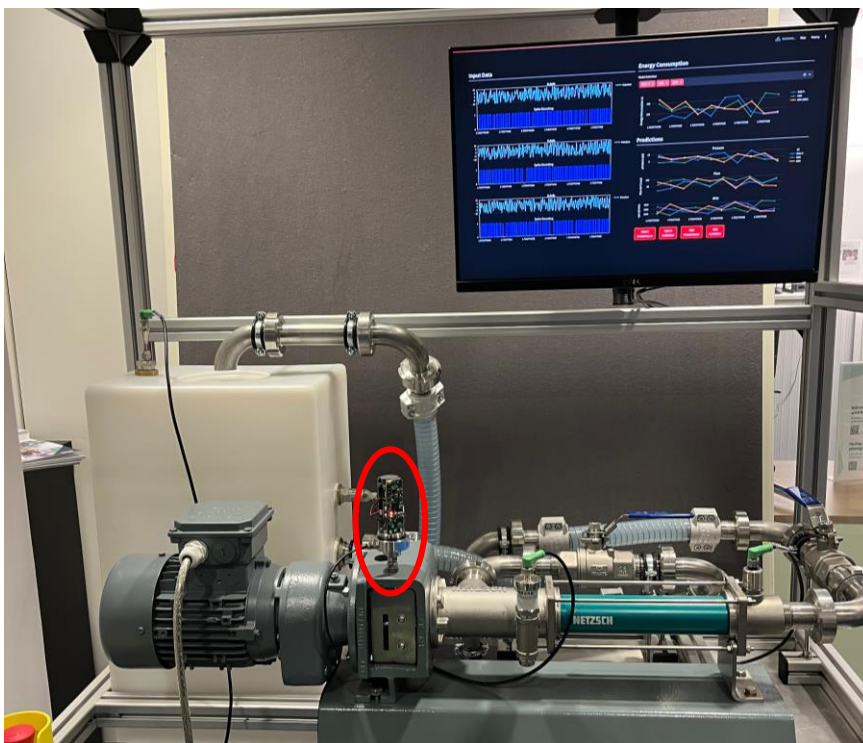
3. Ergebnisse und Erkenntnisse

- **Energieprofil:** Datenübertragung ist der größte Verbraucher. KI-Inferenz on the edge reduziert den Energiebedarf drastisch.
- **Benchmark:** MLPerf-Tests für ARM und RISC-V; neuromorphe Ansätze zeigen Potenzial für signifikante Energieeinsparungen.
- **Hybrider Ansatz:** Kombination aus Edge-KI für Echtzeit und zentralisierter KI für langfristige Analysen.
- **Chancen:** Verlängerte Wartungsintervalle, mögliche Integration von Energy Harvesting für quasi unbegrenzte Laufzeit.

4. Relevanz für die Industrie

- Beitrag zur **Nachhaltigkeit** und **Resilienz** in vernetzten Produktionsumgebungen.
- Reduktion von Betriebskosten durch längere Batterielaufzeiten.
- Skalierbarkeit für verschiedene industrielle Anwendungen.

Im Vortrag wird das abgeschlossene Projekt zusammengefasst und dessen technische Details vertieft.



Projektdemonstrator mit ThinkKIsense Sensor (rot)

Optimierung der Sensorleistung: Einfluss von Schutzmembranen auf die Ansprechzeit

*Michael Schindler, M.Sc., Mihir Khadilkar, Ph.D., Benedikt Hirler, B.Sc., Ertan Anar, Dr. Peter Hils
W.L. Gore & Associates GmbH, Hermann-Oberth-Str. 26, 85640 Putzbrunn, Deutschland*

Zusammenfassung

Die vorliegende Studie untersucht den Einfluss von Schutzmembranen der Firma W.L. Gore auf die Ansprechzeit T₆₃ eines NDIR CO₂-Sensors und analysiert den Zusammenhang zwischen Membraneigenschaften und Sensorperformance mittels eines statistischen Modells. Je nach Membrantyp, Gasaustauschfläche und Zielgaskonzentration erhöhen die Membranen die Ansprechzeit des Sensors von 4 – 5 s, bzw. 13 – 15 s um 0,5 – 4,5 s. Bei großer Gasaustauschfläche dominiert dabei die Grenzflächenströmung den Gastransport, wodurch alle Membranen ein ähnliches Ergebnis zeigen, während bei reduzierter Gasaustauschfläche die Membrandiffusion limitierend wirkt und deutliche Unterschiede zwischen den Membranen aufzeigt. Statistisch zeigen insbesondere Porengröße sowie die Wechselwirkung von Membrandicke und Porosität einen wesentlichen Einfluss. Dünne oder hochporöse Membranen verbessern die Ansprechzeit ebenso wie größere Porendurchmesser.

Keywords: Gassensorik

Einleitung

Je nach Messverfahren benötigen manche Sensoren eine offene Schnittstelle zu der von ihnen zu charakterisierenden Umgebung. Dies trifft insbesondere auf Gassensoren zu. Das Zielgas soll möglichst ungehindert durch eine Schnittstelle an das Sensorelement strömen, damit es dort schnell und vollständig detektiert werden kann. Wenn die Schnittstelle ungeschützt ist, können jedoch auch Kontaminationen wie Staub oder Flüssigkeiten auf das Sensorelement treffen und die Langlebigkeit oder Signalintegrität negativ beeinflussen. Schutzmembranen aus ePTFE können in anspruchsvollen Umgebungen helfen die Sensorgenauigkeit zu verbessern und den Datenverlust zu minimieren [1]. Da die Membran einen Diffusionswiderstand darstellt, kann sich die Ansprechzeit des Sensors durch dieses Schutzelement verlängern. Im Folgenden wird diese Verzögerung für unterschiedliche Membrantypen in einem ausgewählten Messsystem untersucht.

Theoretischer Hintergrund

Der Gastransport an und durch die Membran kann in drei Schritte unterteilt werden:

- 1) Konvektiver Gastransport von der Umgebung an die äußere Grenzschicht heran
- 2) Diffusiver Gastransport durch die Membran hindurch

- 3) Konvektiver Gastransport von der inneren Grenzschicht in die Sensorkavität

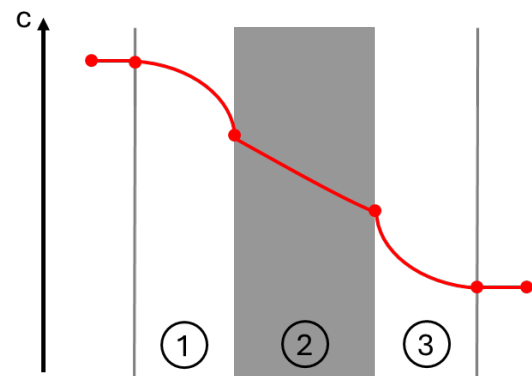


Abb. 1: schematische Darstellung der Gaskonzentration c während des Gastransports von der Umgebung in die Sensorkavität.

Die Vorgänge in der Grenzfläche sind unabhängig von den Membraneigenschaften.

Der Diffusionsstrom J durch die Membran mit der Fläche A ist nach dem 1. Fick'schen Gesetz Gl. (1) proportional zur Diffusionskonstante D und dem Konzentrationsgradienten dc/dx :

$$J = -D \cdot A \cdot dc/dx \quad (1)$$

Hier können bestimmte Membraneigenschaften wie Dicke, Porosität und Porengrößenverteilung

eine Rolle spielen, da sie die effektive Fläche A sowie den Abstand dx beeinflussen. Der Einfluss der Fläche A wird durch Graunke et al. für nicht-poröse Fluoropolymerfilme bestätigt [2].

Die mittlere Porengröße der in dieser Arbeit untersuchten Membranen liegt bei ca. $0,5 - 5 \mu\text{m}$. In diesem Bereich ist die mittlere freie Weglänge der Moleküle deutlich kleiner als der Porendurchmesser. Somit limitiert die Knudsen Diffusion den Diffusionsstrom nicht mehr. Es überwiegt die allgemeine Molekulardiffusion [3].

Die Wechselwirkung des verwendeten Membranmaterials PTFE mit unterschiedlichen Zielgasen wurde u.a. von Graunke et al. untersucht. Je nach Polarität des Zielgases können sich unterschiedlich starke Wechselwirkungen ergeben, die die Ansprechzeit des Sensors zusätzlich verlängern können. Im Folgenden soll nur CO_2 als Zielgas untersucht werden, um den Einfluss unterschiedlicher Wechselwirkungen mit verschiedenen Zielgasen auszuschließen.

Versuchsaufbau

Für die Messung der Ansprechzeit wurde ein spezieller Versuchsaufbau konstruiert, der in Abb. 2 zu sehen ist.

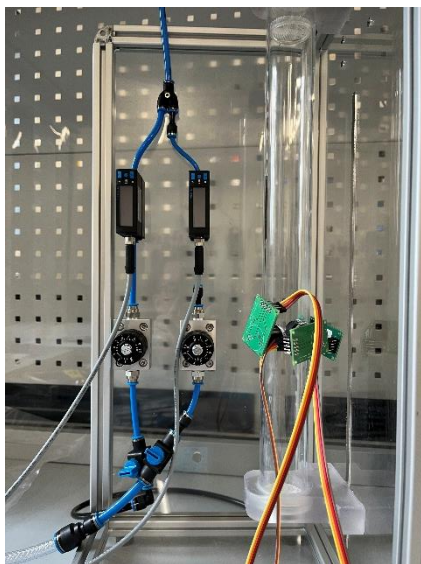


Abb. 2: Versuchsaufbau zur Messung der Ansprechzeit mit Gasregler auf der linken Seite und senkrechten Strömungsrohr mit Sensoren rechts.

Abb. 3 zeigt den schematischen Aufbau. Im vorderen Bereich können Druckluft und Zielgas geregelt und per Ventil ab- und zugeschaltet werden.

Die Gasmischung strömt am oberen Ende des senkrechten Strömungsrohres durch einen Diffusor und wird so gleichmäßig vermischt. In einer laminaren Strömung strömt das Gasgemisch an den Sensoren vorbei, die sich radial angeordnet alle im gleichen Abstand zum Diffusor befinden.

Am unteren Ende des Rohres befindet sich ein weiterer Diffusor. Dieser erzeugt einen gleichmäßigen Staudruck und ermöglicht so eine reproduzierbare Messung.

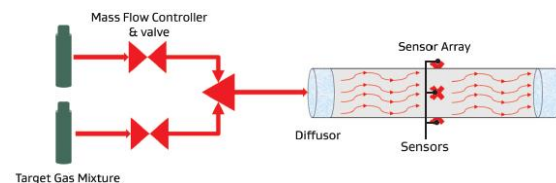


Abb. 3: Schematische Darstellung des Versuchsaufbaus zur Messung der Ansprechzeit.

Als Sensor wird der CO_2 NDIR Sensor SprintIR-WX-20 von Gas Sensing Solutions verwendet. Dieser wurde ausgewählt, da er über eine sehr schnelle Messrate von 20 Messungen pro Sekunde verfügt und die Änderungsgeschwindigkeit des CO_2 -Signals somit nicht vom Sensor selbst, sondern von der applizierten Membran limitiert wird.

Es befinden sich parallel vier Sensoren im Versuchsaufbau. Ein Sensor dient dabei immer als Referenz und hat keine Schutzmembran. Die anderen drei Sensoren werden durch unterschiedliche mikroporöse Membranen des Herstellers W.L. Gore & Associates geschützt:



Abb. 4: Referenzsensor ohne Membran und Sensor mit Schutzmembran im Strömungsrohr angeordnet.

Durchführung

Mit dem gezeigten Aufbau werden Versuche durchgeführt, bei denen die folgenden Parameter variiert werden:

- Membrantyp
- Aktive Fläche der Membran
- CO₂-Konzentration

Folgende Membrantypen der Firma Gore werden charakterisiert:

Tab. 1: Verwendete Membrantypen

Membran Bezeichnung	Status
MMS311	kommerziell
MMS321	kommerziell
MMS331	kommerziell
MMS341	kommerziell
RD1	Nicht kommerziell
RD2	Nicht kommerziell

Zur Entwicklung eines statistischen Modells werden auch nicht-kommerzielle Membranen getestet, die RD1 und RD2 bezeichnet werden.

Es ist zu beachten, dass die unterschiedlichen Membranen wie in Abb. 5 beispielhaft zu sehen unterschiedliche Porengeometrien aufweisen. Bei der statistischen Auswertung wird deshalb für alle Typen ein mathematisch berechneter Äquivalenzdurchmesser für die Porengröße angegeben.

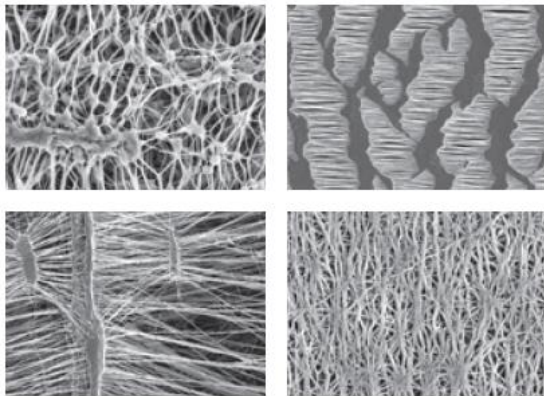


Abb. 5: Beispiele für unterschiedliche Porenstrukturen der verwendeten Membranen.

Die ursprüngliche aktive Fläche A des Sensors hat einen Durchmesser von 14 mm. Auf diese kann ein gasundurchlässiger Klebering aufgebracht werden, der die aktive Fläche verkleinert, wie es bei dem Referenzsensor in Abb. 4 zu sehen ist. Somit gibt es zwei untersuchte aktive Flächen:

- Ø 14 mm (original)
- Ø 7 mm (reduziert durch Klebering)

Der verwendete Sensor hat einen Messbereich von 0 - 20 %CO₂. Aus diesem Bereich werden 0,5 %CO₂ und 15 %CO₂ für die Studie gewählt.

Zur statistischen Auswertung gibt es jeweils fünf Wiederholungsversuche pro Parameterkombination.

Auswertung

Die Auswertung der Ansprechzeit erfolgt anhand von T63. Diese entspricht, wie in Abb. 6 beispielhaft dargestellt, der Zeit bis 63 % der Konzentrationsänderung abgelaufen ist.

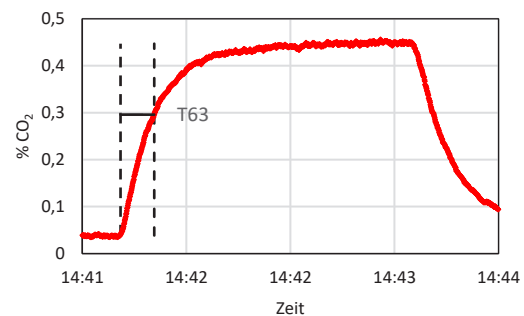


Abb. 6: Visualisierung der Ansprechzeit T63.

Für die Bestimmung des Einflusses der Membran wird für jeden Lauf die Ansprechzeit des Referenzsensors von der Ansprechzeit des Sensors mit Membran abgezogen:

$$t_{\text{Membran}} = t_{\text{Gesamt}} - t_{\text{Referenz}} \quad (\text{Gl.2})$$

Der Zusammenhang zwischen den Membraneigenschaften und T63 wird mit der Methode der kleinsten Quadrate ermittelt. Das Model stellt eine Korrelation der Ansprechzeit mit den Eigenschaften Dicke, Porengröße und Porosität auf.

Ergebnisse

Ansprechzeit des Referenzsensors ohne Membran

Um die Fähigkeit des Messsystems zu evaluieren werden zunächst die Ergebnisse des Referenzsensors betrachtet. Hier gilt nach Gl. (2):

$$t_{\text{Gesamt}} = t_{\text{Referenz}}$$

Abb. 7 zeigt eine Übersicht über die Ansprechzeiten des Referenzsensors für alle Kombinationen von aktiver Fläche und CO₂-Konzentration. Pro Kombination gibt es 10 Einzelwerte.

Es zeigt sich, dass der Mittelwert der Ansprechzeit des Sensors für beide Gaskonzentrationen sehr ähnlich ist. Für den

originalen Durchmesser von 14 mm liegt sie bei 4 – 5 s. Die Reduzierung des Durchmessers auf 7 mm führt zu einer Verlängerung der Ansprechzeit auf 13 – 15 s.

Bei der geringeren Gaskonzentration von 0,5 % ist die Standardabweichung des Referenzsensors leicht erhöht gegenüber der CO₂-Konzentration von 15 %.

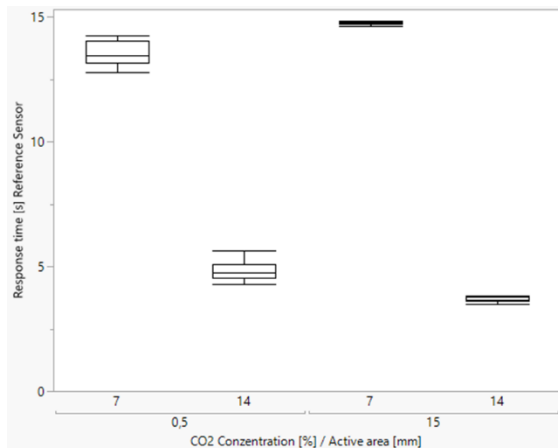


Abb. 7: T63 des Referenzsensors für die beiden aktiven Flächen und Gaskonzentrationen.

Im nächsten Schritt wird betrachtet welchen Einfluss die Schutzmembranen auf T63 haben.

Anteil der Membranen an der Ansprechzeit

Für den Vergleich der Schutzmembranen werden die vier kommerziellen Membranen aus Tab. 1 genutzt.

Abb. 8 zeigt den Anteil der Membran t_{Membran} an der gesamten Ansprechzeit des geschützten Sensors t_{Gesamt} für die CO₂-Konzentration 0,5% und Abb. 9 für 15%.

Aus dem Vergleich der beiden Abbildungen sind folgende Erkenntnisse ersichtlich:

Bei der vollen aktiven Fläche mit Durchmesser 14 mm verhalten sich alle vier untersuchten Membrantypen ähnlich. Dies gilt für beide CO₂-Konzentrationen. Bei 0,5 % beträgt t_{Membran} ca. 1,5 s, bei 15 % liegt der Wert bei ca. 3,0 s.

Beim verkleinerten Durchmesser von 7 mm zeigt sich ein deutlicher Unterschied zwischen den Membranen, der für beide CO₂-Konzentrationen von der Tendenz ähnlich ist. MMS311 hat die höchste Ansprechzeit mit 2,0 s (0,5 %) bzw. 4,5 s (15 %) und MMS341 den niedrigsten Wert mit 0,5 s (0,5 %) bzw. 1,5 s (15 %).

Dies deutet darauf hin, dass bei der großen Fläche der Diffusionsstrom nicht durch die Membran, sondern durch die Grenzflächenströmung limitiert ist. Bei der kleineren Fläche ist der Diffusionsstrom durch die Membran gemäß dem Fick'schen Gesetz Gl. (1) geringer, wodurch der Membraneinfluss zum Tragen kommt.

In diesem Fall ist die Auswahl der optimalen Membran wichtiger als bei großen Sensorgeometrien.

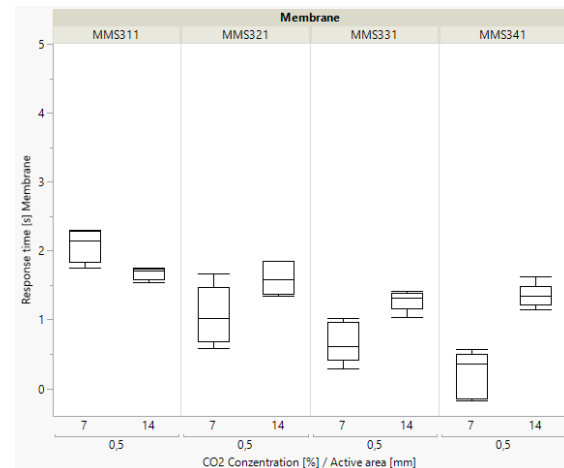


Abb. 8: Anteil der Membran t_{Membran} an T63 für die CO₂-Konzentration 0,5 %.

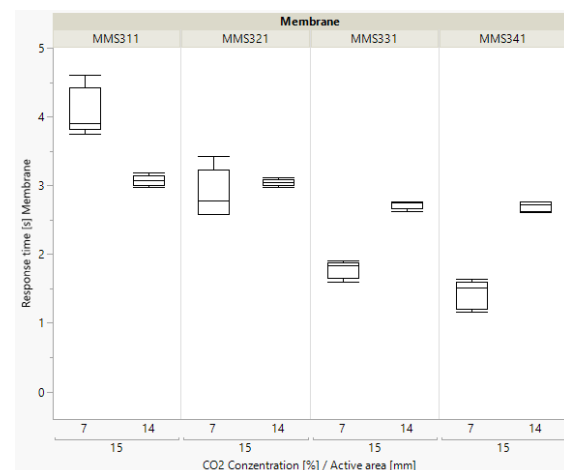


Abb. 9: Anteil der Membran t_{Membran} an T63 für die CO₂-Konzentration 15 %.

Weiterhin ist bei der niedrigeren CO₂-Konzentration der absolute Anteil der Membran t_{Membran} mit 0,5 – 2,3 s geringer als bei der höheren CO₂-Konzentration mit 1,5 – 4,5 s, obwohl der Referenzsensor in Abb. 7 keine signifikanten Unterschiede zeigt.

Dies kann darauf zurückzuführen sein, dass bei der geringeren Konzentration weniger CO₂-

Moleküle die Membran passieren müssen bis der Sensor den korrekten Wert anzeigt. Weiterhin stellen sich bei beiden Konzentrationen unterschiedliche Volumenströme im Strömungsrohr ein, da der Massedurchfluss des Trägergases Luft konstant ist. Dies kann zu unterschiedlichen Grenzflächen- oder Strömungseffekten führen, was wiederum einen Einfluss auf die Ansprechzeit der Sensoren haben kann.

Anhand dieser Erkenntnisse wird entschieden das statistische Modell für den Membraneinfluss nur auf Basis der Ergebnisse mit Durchmesser 7 mm und der CO₂-Konzentration 15 % aufzustellen, da hier der Einfluss der Membranen am deutlichsten ersichtlich ist.

Einfluss der Membraneigenschaften auf die Ansprechzeit

Zur Untersuchung welche der Membraneigenschaften Dicke, Porengröße und Porosität welchen Einfluss auf die Ansprechzeit hat, wird die Methode der kleinsten Quadrate angewendet.

Da die vier oben aufgeführten, kommerziellen Schutzmembranen teilweise ähnliche Eigenschaften aufweisen, werden für die Generierung des statistischen Modells noch zwei weitere, nicht-kommerzielle PTFE-Membranen charakterisiert.

Auf Basis dieser sechs Membranen lässt sich, wie in Abb. 10 zu sehen, ein Modell mit Bestimmtheitsmaß $R^2 = 0,91$ erstellen. Die Wurzel des mittleren quadratischen Fehlers (RMSE) liegt dabei bei 0,2109, was in etwa der in den Versuchen beobachteten Standardabweichung entspricht. Das Modell ist mit einem p-Wert $\leq 0,0001$ statistisch signifikant.

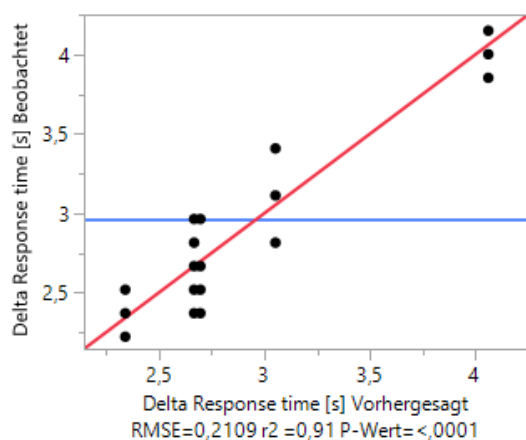


Abb. 10: Beobachtete Werte über Vorhersage für das erzeugte Modell.

Das Modell zeigt, dass die Porengröße den größten Einfluss einer einzelnen Membraneigenschaft hat. Je größer die Pore, desto schneller die Ansprechzeit.

Weiterhin hat die Dicke in Wechselwirkung mit der Porosität einen großen Einfluss wie in Abb. 11 zu sehen ist.

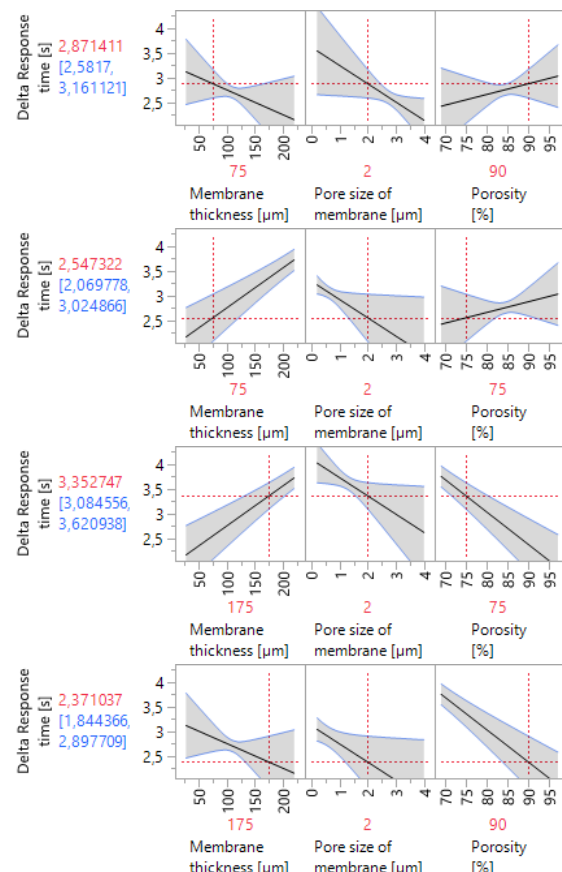


Abb. 11: Vorhersageanalyse für den Einfluss der Membraneigenschaften auf t_{Membran} .

Der Einfluss der Porosität ist bei dickeren Membranen anders als bei dünneren Membranen und umgekehrt.

- Bei dünnen Membranen steigt die Ansprechzeit mit zunehmender Porosität.
- Bei dicken Membranen fällt die Ansprechzeit mit zunehmender Porosität.
- Bei hohen Porositäten sinkt die Ansprechzeit mit zunehmender Dicke.
- Bei niedrigen Porositäten steigt die Ansprechzeit mit zunehmender Dicke.

Dabei ist zu beachten, dass die Membranen des Herstellers Gore im Vergleich zu anderen kommerziell verfügbaren Membranen zwar eine sehr gleichmäßige Struktur aufweisen, die Porenstrukturen der unterschiedlichen Membranen allerdings nicht einheitlich sind und

keine der Membranen ideal runde und ideal gleichmäßig verteilte Poren hat (s. Abb. 5). Dadurch kommt es zu einer gewissen Unsicherheit bei dem Modell und die Vorhersage ist nicht 1:1 auf weitere Membrantypen anwendbar.

Ebenso kann der verwendete Sensor und das Zielgas einen Einfluss auf die Übertragbarkeit des Modells auf andere Anwendungen haben.

Zusammenfassung

In dieser Studie wird der Einfluss von Schutzmembranen der Firma W.L. Gore auf die Ansprechzeit T₆₃ eines NDIR CO₂-Sensors untersucht. Weiterhin wird ein statistisches Modell aufgestellt, das den Zusammenhang zwischen den Membraneigenschaften Dicke, Porengröße und Porosität mit der Ansprechzeit zeigt.

Ohne Schutzmembran haben die eingesetzten Sensoren eine Ansprechzeit von 4-5 s bei Nutzung der vollen Gasaustauschfläche. Bei Halbierung des aktiven Durchmessers von 14 mm auf 7 mm erhöht sich dieser Wert auf 13 - 15 s unabhängig von der Gaskonzentration. Die untersuchten Membranen addieren je nach Membrantyp, Austauschfläche und Gaskonzentration 0,5 – 4,5 s zur Ansprechzeit des Sensors. Dabei dominiert bei voller Fläche die Grenzflächenströmung den Gastransport, wodurch T₆₃ für alle Membranen vergleichbar ist. Bei Verringerung der Fläche limitiert die Membrandiffusion den Gastransport, wobei sich ein signifikanter Unterschied zwischen den getesteten Membranen zeigt.

Die CO₂-Konzentration hat keinen Einfluss darauf, welche Membran die schnellste Ansprechzeit des Sensors zur Folge hat.

Die statistische Auswertung der Ansprechzeiten der unterschiedlichen Membranen kann einen Zusammenhang mit den Membraneigenschaften Dicke, Porengröße und Porosität zeigen. Den größten Einfluss hat dabei die Wechselwirkung aus Dicke und Porosität.

Die Porengröße hat den größten Einfluss einer einzelnen Membraneigenschaft. Je größer die Poren sind, desto schneller kann das Gas hindurch diffundieren. Dabei ist zu berücksichtigen, dass die Schutzwirkung gegen das Eindringen von Flüssigkeiten und Partikeln mit zunehmender Porengröße abnimmt.

Die Studie belegt somit, dass insbesondere bei kleineren Sensoren die Auswahl einer Membran mit geeigneten Eigenschaften einen

signifikanten Einfluss auf die Ansprechzeit eines Gassensors haben kann.

Literaturnachweis

- [1] R. Parsiegel, M. Budag Becker, P. Try, M. Gebhard, *Virtual MOS Sensor Array Design for Ammonia Monitoring in Pig Barns*, *Sensors* **2025**, 25, 2617; doi: 10.3390/s25082617
- [2] T. Graunke, K. Schmitt, S. Raible, J. Wöllenstein, *Towards Enhanced Gas Sensor Performance with Fluoropolymer Membranes*, *Sensors* **2016**, 16(10), 1605; doi: 10.3390/s16101605
- [3] D.M. Tartakovsky, M. Dentz, *Transport in Porous Media* 130, 105–127 (2019); doi: 10.1007/s11242-019-01262-6

Körperschallanalyse zur Fehlererkennung bei Fahrzeugkomponenten

Martin Bussas¹

¹Trout GmbH, Parkstraße 28, 34119 Kassel, Germany

Zusammenfassung

Rotierende und bewegliche Fahrzeugkomponenten wie Radlager oder Antriebswellen unterliegen im Betrieb hohen mechanischen Belastungen. Mit zunehmender Laufleistung können dadurch Verschleiß, Materialermüdung oder Mikrorisse entstehen, die – wenn sie unentdeckt bleiben – zu Funktionsstörungen oder gar Ausfällen führen können. Konventionelle Prüfmethode erkennen Schäden häufig erst in einem fortgeschrittenen Stadium oder erfordern aufwändige Demontagearbeiten. Eine frühe und nicht-invasive Detektion solcher Defekte stellt daher eine wesentliche Herausforderung im Bereich der Fahrzeugdiagnose und Zustandsüberwachung dar. Die Erfassung des strukturgetragenen Schalls (Körperschall) bietet hier einen vielversprechenden Ansatz: Durch die Analyse des Schallspektrums lassen sich charakteristische Anomalien oder Muster identifizieren, die auf beginnende Schäden hinweisen.

Keywords: Machine-Learning, Fahrzeugdiagnose, Frequenzspektrum, Körperschall, Instandhaltung

Körperschall

Körperschallsensoren, auch als Strukturschallsensoren bekannt, messen die durch mechanische Quellen in einem Material erzeugten Schwingungen. Diese Sensoren sind besonders nützlich für die Zustandsüberwachung und die vorbeugende Wartung von Maschinen und Fahrzeugen, einschließlich militärischer Radfahrzeuge. Sie werden hier verwendet, um Defekte an bewegten oder rotierenden Teilen zu detektieren, indem Veränderungen in den Schwingungsmustern erkannt werden, die auf Probleme wie Unwucht, Verschleiß, Risse oder andere Schäden hinweisen.



Abb. 1: Sensieren des Körperschalls mit zwei MEMS Mikrofonen.

Zwei Mikrofone erfassen kontinuierlich die durch die Fahrzeugkomponenten erzeugten und transmittierten Schwingungen, sowohl im Ruhezustand (Transport) als auch während der Fahrt.

Die Schallerfassung wird nach Geschwindigkeitsintervallen differenziert, und es wird zwischen Gelände und Straße unterschieden. Die aufgenommenen Daten werden einmal in Echtzeit analysiert und auch im Nachgang, um Muster oder Anomalien zu erkennen. Dies erfolgt durch fortgeschrittene Algorithmen aus dem Umfeld des maschinellen Lernens, um auch geringfügige Veränderungen zu identifizieren, die auf potenzielle Defekte hinweisen. Weitere Parameter aus Kinematik, Umwelt und Fahrzeugdaten aus CAN-Bus-Kommunikation werden parallel gespeichert und finden in die Auswertung Eingang.

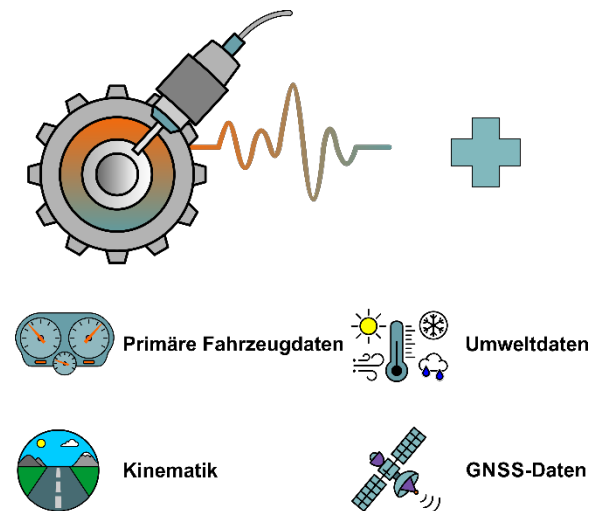


Abb. 2: Datenerfassung Schall mit Fahrzeugdaten, Umwelt, Kinematik und ggf. Position

Bei der Erkennung von Anomalien wird eine Benachrichtigungskette initiiert, um die spezifische Ursache zu identifizieren und vorbeugende Wartungsmaßnahmen durchzuführen, bevor ein schwerwiegender Ausfall auftritt.

Im Gegensatz zu anderen Diagnosemethoden, die direkten Zugang zu den zu überwachenden Teilen erfordern, können Körperschallsensoren Messungen durchführen, ohne in den Betrieb des Fahrzeugs einzugreifen.

Zustandsdiagnose

Fahrzeuge erzeugen während ihres gesamten Betriebs kontinuierlich mechanische, akustische und dynamische Signale. Vibrationen, Körperschall und Geräusche spiegeln dabei den aktuellen Zustand einzelner Komponenten und Systeme wider. Die zentrale Herausforderung besteht jedoch darin, diese komplexen Signale nicht nur zu erfassen, sondern sie auch zuverlässig zu interpretieren und in belastbare diagnostische Erkenntnisse zu überführen.

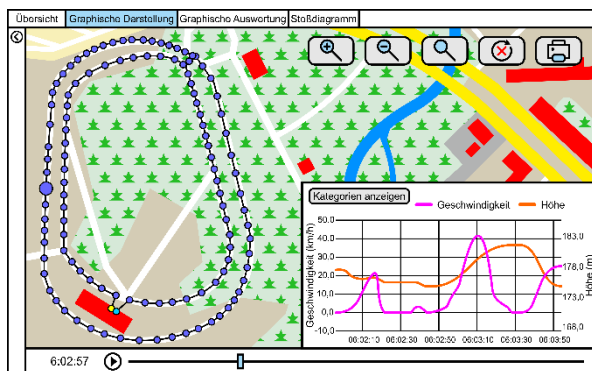


Abb. 3: Fahrzeugdatengewinnung auf der Teststrecke

Mit dem Vehicle-Data-Analysis-System VeDAS wurde ein kompaktes, autarkes Diagnosesystem entwickelt, das speziell für den Einsatz auf gepanzerten Radfahrzeugen konzipiert ist. Das System ermöglicht die kontinuierliche Erfassung und Speicherung hochauflösender Betriebsdaten unter realen Einsatzbedingungen. Ein besonderer Vorteil des Systems besteht darin, dass dank eines leistungsstarken Akkus die Datenerfassung unabhängig vom Betriebszustand des Fahrzeugs erfolgen kann. Dadurch lassen sich beispielsweise während Transportphasen relevante Zustandsinformationen gewinnen. Diese kontinuierliche und betriebsnahe Datenerhebung schafft die Grundlage für ein unverfälschtes Bild des tatsächlichen Fahrzeugverhaltens unter realen Belastungs- und Umweltbedingungen. Im Gegensatz zu

Labor- oder Prüfstandsdaten spiegeln die erfassten Informationen die tatsächlichen Einsatzszenarien wider und ermöglichen dadurch eine deutlich realitätsnähere Zustandsbewertung.



Abb. 4: Aufnahmebox für Schall, Kinematik, Umwelt- und Fahrzeugdaten

Körperschall entsteht durch mechanische Wechselwirkungen innerhalb des Fahrzeugs und breitet sich über die Fahrzeugstruktur aus. Verschleißerscheinungen, Materialermüdung, Lagerdefekte oder strukturelle Schäden erzeugen charakteristische Frequenzmuster, die sich häufig bereits in frühen Entstehungsphasen nachweisen lassen. Das System „hört“ somit direkt in die Fahrzeugstruktur hinein und erfasst Veränderungen, bevor diese zu sichtbaren Fehlfunktionen oder kritischen Ausfällen führen. Zur Auswertung dieser Signale werden akustische und schwingungsbasierte Rohdaten zunächst in spektrale und frequenzbasierte Merkmalsräume transformiert.

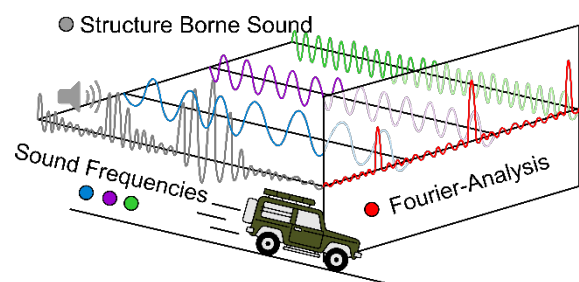


Abb. 5: Vom Geräusch zum Frequenzspektrum mit Fast-Fourier-Transformation

Anschließend kommen Verfahren des maschinellen Lernens zum Einsatz, um charakteristische Signaturen für Verschleiß, Schwachstellen und Defekte zu identifizieren. Ziel ist die Entwicklung robuster Modelle, die in der Lage sind, wiederkehrende Muster zuverlässig von betriebsbedingten Variationen und Umwelteinflüssen zu unterscheiden. Die Herausforderung liegt dabei in der hohen

Komplexität realer Einsatzdaten. Im praktischen Betrieb sind die Sensorsignale zahlreichen Störeinflüssen ausgesetzt. Unterschiedliche Fahrbahn- und Geländebeschaffenheiten oder Umweltbedingungen führen zu stark variierenden Signalcharakteristika. Zusätzlich steht die Entwicklung datengetriebener Diagnosemodelle vor dem Problem begrenzter gelabelter Datensätze. Kritische Fehlerzustände treten vergleichsweise selten auf, wodurch nur wenige eindeutig annotierte Beispiele verfügbar sind. Um trotzdem aus Schwingungen und Körperschall verwertbares Wissen zu generieren und damit Schall in belastbare Erkenntnisse zu transformieren, müssen folgende Hürden genommen werden.

- Entwicklung robuster Lernverfahren zur Mustererkennung unter eingeschränkter Datenverfügbarkeit
- Multimodale Sensorfusion, Kombination und Synchronisierung heterogener Datenquellen wie Kinematik, Körperschall, Umgebungsdaten und CAN-Bus-Informationen.

Datenreduktion

Zur effizienten und zielgerichteten Erfassung des Strukturschalls wird ein segmentiertes Messverfahren angewendet, das auf die charakteristischen Frequenzbereiche der verschiedenen Schadensarten abgestimmt ist. Die aufgenommenen Schallsignale werden einer Fast-Fourier-Transformation unterzogen, auf einer internen SD-Karte gespeichert und anschließend über ein Machine-Learning-Verfahren miteinander verglichen: Ein- und Ausgangsfahrt, verschiedene Drehzahlen, unterschiedliche Lastzustände.

Die kontinuierliche und vollständige Erfassung des gesamten Frequenzbereichs würde zu sehr großen Datenmengen und damit zu hohen Anforderungen an Speicherung und Auswertung führen. Aus diesem Grund werden die Messungen in vier Teilbereiche, die für eine individuelle Abstimmung auf einen bestimmten Fahrzeugtyp weiter differenziert werden können, unterteilt:

- Niedrige Frequenzen (6 Hz – 512 Hz): Lange Messzeiten erforderlich, Erkennung von Unwuchten oder großflächigen Strukturveränderungen.
- Mittlere Frequenzen (512 Hz – 1536 Hz): Typische Anzeichen für beginnende Lager- oder Getriebebeschäden.
- Hohe Frequenzen (1536 Hz – 4608 Hz): Empfindlich gegenüber Materialermüdung und beginnendem Versagen. Kürzere Messzeiten ausreichend.

- Höchste Frequenzen (4608 Hz – 13824 Hz): Empfindlich gegenüber Mikrorissen. Kürzeste Messzeiten ausreichend.

Diese Aufteilung reduziert das Datenvolumen erheblich und ermöglicht gleichzeitig eine detaillierte Analyse in allen relevanten Frequenzbereichen.

Signalverarbeitung

Die über die Zeit erfassten akustischen Signale werden mittels Fast-Fourier-Transformation (FFT) in den Frequenzbereich überführt. So lassen sich die spektralen Komponenten des Körperschalls präzise darstellen und charakteristische Frequenzen von Defekten (z.B. harmonische Schwingungen beschädigter Lager) identifizieren.

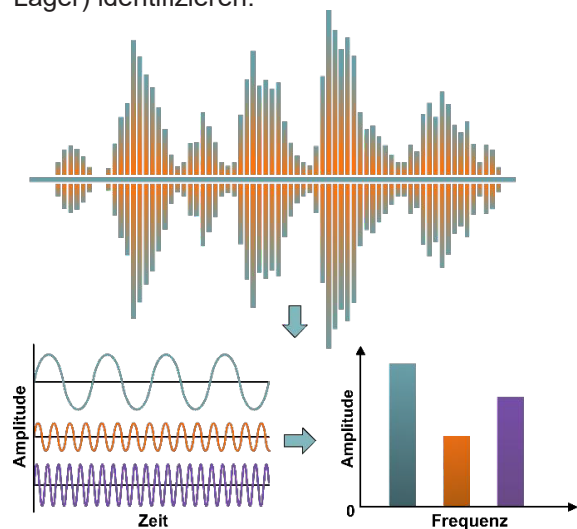


Abb. 5: Überführung akustischer Signale in ein Frequenzspektrum

Defekterkennung durch Musteranalyse

Ein zentraler Bestandteil der datenbasierten Zustandsüberwachung ist die automatisierte Defekterkennung durch Musteranalyse. Hierbei werden aktuelle Messdaten kontinuierlich mit zuvor definierten Referenzzuständen verglichen, die das charakteristische Schwingungs- und Körperschallverhalten eines technisch intakten Fahrzeugs beschreiben.

Mechanische Defekte, Verschleißerscheinungen oder strukturelle Veränderungen führen häufig zu charakteristischen Veränderungen innerhalb dieser Spektren. Dazu zählen unter anderem erhöhte Amplituden einzelner Frequenzbereiche, neu auftretende Resonanzen, harmonische Überlagerungen oder Verschiebungen dominanter Frequenzanteile. Solche Veränderungen entstehen beispielsweise durch Unwuchten, Lagerschäden, Materialermüdung oder veränderte Kopplungsverhältnisse innerhalb der Fahrzeugstruktur.

Im Rahmen der Musteranalyse werden diese spektralen Veränderungen automatisiert identifiziert und hinsichtlich ihrer Relevanz bewertet. Verfahren des maschinellen Lernens ermöglichen dabei die Erkennung komplexer Zusammenhänge und wiederkehrender Signurmuster, die sich einer rein manuellen Analyse häufig entziehen. Die systematische Analyse und der Vergleich dieser Datenmengen übersteigen die Möglichkeiten einer rein menschlichen Auswertung deutlich und erfordern automatisierte, datengetriebene Analyseverfahren.

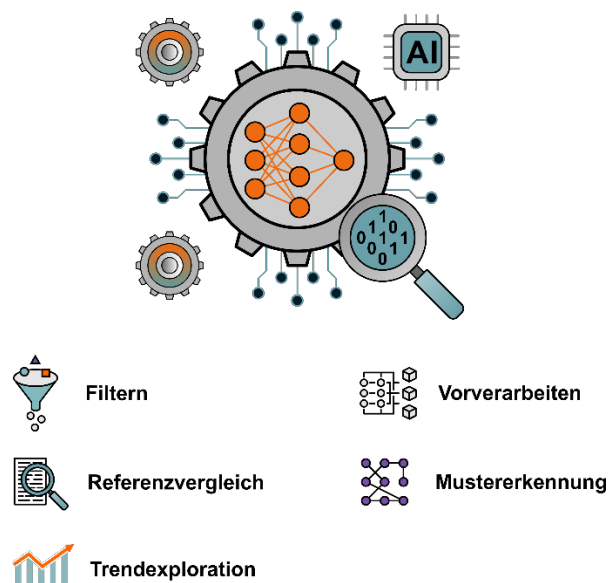


Abb. 6: Defekterkennung durch Musteranalyse

Vorausschauende Wartung

Die nachfolgend dargestellten Spektren zeigen exemplarisch den Vergleich zwischen einer Eingangsfahrt eines defekten Fahrzeugs und einer Ausgangsfahrt nach erfolgter Reparatur. Im Frequenzspektrum der Eingangsfahrt sind deutliche Auffälligkeiten sichtbar. Mehrere Frequenzbereiche weisen stark erhöhte Amplituden auf, die auf zusätzliche mechanische Schwingungsanteile und strukturelle Anregungen hindeuten. Darüber hinaus zeigen sich Verschiebungen charakteristischer Frequenzanteile gegenüber dem Referenzzustand. Sie wurden rot umringt. Im dargestellten Beispiel sind die Unterschiede zwischen beiden Spektren bereits visuell gut erkennbar. In realen Einsatzszenarien fallen solche Abweichungen jedoch häufig deutlich geringer aus und verteilen sich über zahlreiche Frequenzbereiche sowie unterschiedliche Betriebszustände.

Darüber hinaus ist die Anzahl der Spektren aus verschiedenen Fahrten, die miteinander verglichen werden, erheblich. Eine zuverlässige Bewertung dieser Datenmengen ist daher ohne automatisierte Analyseverfahren kaum realisierbar. Genau an dieser Stelle entfalten datengetriebene Methoden und maschinelles Lernen ihren entscheidenden Vorteil. Sie ermöglichen die systematische Erkennung selbst kleinster Musterabweichungen und schaffen damit die Grundlage für eine skalierbare, reproduzierbare und belastbare Zustandsdiagnose.

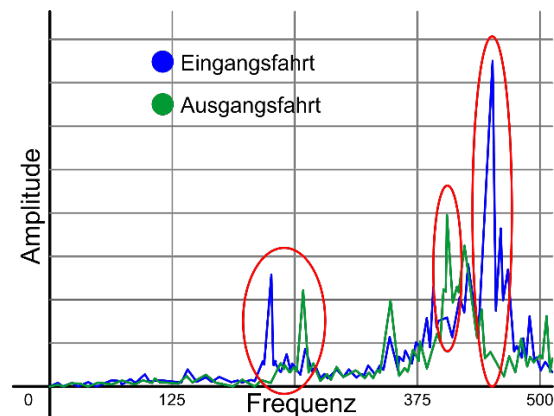


Abb. 7: Identifizierte Frequenzanomalien zwischen Ein- und Ausgangsfahrt

Fazit

Die Analyse des strukturgetragenen Schalls stellt ein leistungsfähiges Werkzeug zur frühen Erkennung von Schäden an rotierenden Komponenten dar. Durch die Kombination aus segmentierter Messung, FFT-Auswertung und vergleichender Musteranalyse entsteht ein Verfahren, das sowohl technisch effizient als auch wirtschaftlich sinnvoll ist. Es ermöglicht nicht nur rechtzeitige Reparaturen, sondern auch die Implementierung moderner, zustandsbasierter Wartungskonzepte im Fahrzeugbereich.

Literaturnachweis

- [1] Vehicle-Data-Analysis-System (VeDAS) ein erster Schritt hin zur Planung und Vorbereitung bedarfsorientierter Wartung, CPM Forum für Rüstung und Sicherheit Heft 4 | 2023 Seite 66ff
- [2] Michael Möser, Wolfgang Kropp, Körperschall Physikalische Grundlagen und technische Anwendungen, 2010 ISBN: 978-3-540-49048-7
- [3] Michael Möser, Körperschall-Messtechnik, 2018 ISBN: 978-3-662-56621-3
- [4] Bernd Orend & Ingo Meyer Schadensfrüherkennung mittels Körperschall, 2009 MTZ – Motortechnische Zeitschrift Volume 70, Seiten 386 - 391

Development and validation of a multimodal sonar measurement system for object detection in water bodies

*Sabrina Lennartz, Michael Schneider, Hubert Mantz, Thomas Walter,
University of Applied Sciences Ulm, Prittwitzstraße 10, 89075 Ulm, Germany*

Summary

This paper presents an autonomous, multimodal measurement system designed to detect man-made objects in bodies of water. The aim is the implementation of a system, based on an unmanned surveillance vehicle (USV) equipped with acoustic sensors for detecting underwater objects from various angles. The collected data is then being georeferenced and combined using imaging techniques to create high-resolution 2D and 3D representations. To develop and validate data processing, a rail-based measurement system was set up, enabling precisely georeferenced measurements in a water basin containing defined targets. In both cases, five single-beam sonars arranged in an array, as well as a side-scan sonar, are used. The array provides broad cross-range coverage while maintaining good resolution. Conversely, the side-scan sonar has very high along-track resolution and low angular resolution. Improved data processing is based on three principles: the use of multiple single-beam sonars, the combination of single-beam and side scan sonars and the fusion of data from multiple angles. In test scenarios involving known targets, the system demonstrated significantly improved image quality. Based on these findings, data processing using data collected by the USV can now be tested in a field trial.

Keywords: 3D reconstruction, multimodal sensor technology, multilook, underwater object detection

Introduction

Over 70 % of the Earth's surface is covered by water [1]. Due to its vast extent and depth, as well as factors such as limited visibility caused by light absorption and turbidity, dynamic environmental factors like currents and changing water properties, and general environmental conditions, researching these ecosystems poses a major challenge [2]. A prime example is ALDFG (abandoned, lost, or otherwise discarded fishing gear), which accounts for around 10 % of marine debris [3]. Fish, marine mammals, and birds can become entangled in the nets, and their decomposition leads to the creation of microplastics [4].

Complete, accurate, and comprehensive data are required for the detection and classification of underwater targets. The strengths of both visual and acoustic sensors can be combined using a multimodal sensor fusion system [2]. Better image quality enables users with limited resources or without a scientific background to obtain meaningful results.

This article presents three approaches to improve data processing without requiring expensive hardware. These approaches were tested in a defined setup. For this purpose, targets were placed in a pool, and the sonars

were moved using a rail-based system. This revealed the positive impact of using different sonars and fusing data from different poses. These findings and insights can now be adapted to a USV (unmanned surveillance vehicle)-based system, which introduces additional challenges, such as wave motion, but also offers more possibilities, such as the more flexible incorporation of different viewpoints of the target object. Thus, this work lays the foundation for a framework capable of combining data from various acoustic and non-acoustic sensors, thereby improving underwater imaging.

State of the art

The main difficulty in detecting ALDFGs is that they become entangled in structures such as rocks and wrecks or are pressed to the seabed by their own weight. Distinguishing the nets from the various, often complex, inhomogeneous surfaces is challenging [5]. Current approaches largely rely on artificial intelligence. Regardless of whether classification is performed by humans or artificial intelligence, there is a high false-positive rate. Similar objects, such as sediment waves or cables are often mistaken for net ropes [6].

Higher image quality would lead to better classification. One effective way to achieve this is to combine complementary perspectives and sensor modalities, as demonstrated for the fusion of multiple acoustic systems and for acoustic–optical combinations [2]. How such data is fused best remains an open question. Feature-based techniques, which dominate the field, tend to struggle underwater due to low contrast, turbidity, and the lack of stable visual features. More recent approaches instead exploit navigation information, which requires accurate knowledge of the sensor platform's position and orientation [2, 6]. This work follows the latter line, fusing data from multiple acoustic sensors through precise georeferencing.

Proposed approaches

As outlined above, the main problem lies in the poor quality of much of the collected sonar data. The three approaches presented here do not aim to improve the collection of data but its processing. First, five single-beam sonars are arranged side by side in an array to cover a larger area without sacrificing cross-range resolution. Second, data from a side-scan sonar are added, allowing the complementary beam characteristics of the two sonar types to be fused. Third, targets are observed from multiple angles and combined through suitable algorithms to achieve greater clarity. Table 1 lists the key specifications of the sonars used, the Sounder S500 single-beam sonar by and the Omniscan 450 Side-Scan sonar both from Cerulean Sonar [7].

Table 1: Specification of the sonars used [7].

Parameter	Single-Beam	Side-Scan
Frequency [kHz]	500	450
Beam Width [°]	5	0.5
Beam Height [°]	-	50
Resolution [mm]	3	< 1/1200 th of range
Max. Range [m]	100	150

Measurement system USV

The goal of this project is the improved processing of data recorded by an autonomous USV. This is based on a small catamaran to the sensors can be attached. The single-beam sonars transmit sequentially to avoid interference between them. They look downward, while the side-scan sonar is mounted on the hull and pointed slightly downward in the cross-range direction. Figure 1 shows the USV. On either side are two gray GPS antennas. Between them, at regular

intervals, are the single-beam sonars. The one at the very front can be seen in red on the side of the hull. Behind it, underwater, is the side-scan sonar mounted on the hull of the boat. The light gray box in the center of the USV contains the necessary electronics for the sensors.



Figure 1: Image of the USV, equipped with sonar and GPS antennas

Measurement system pool

To enable experiments in a controlled environment, an additional measurement system was set up. A carriage is mounted on a linear axis, to which the various sonars can be attached, as shown in Figure 2.

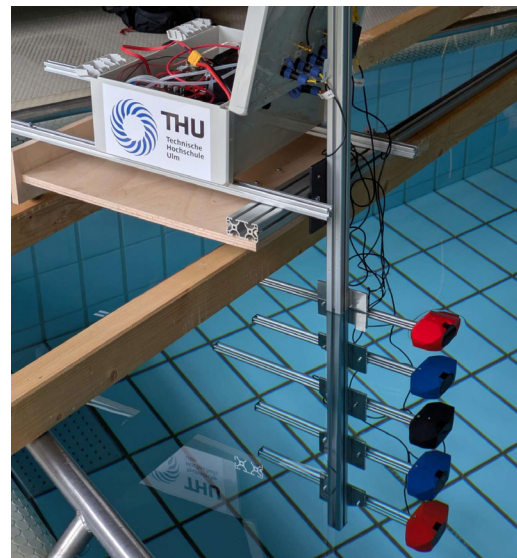


Figure 2: Picture of the measurement system

The pool measured 6.5 by 10 m, with a constant depth of 1.2 m.

Two vertical poles with a diameter of 6mm each (red circle) and two plastic boxes, one of which was covered with a net to simulate the real-life situation of a snagged net, were used as targets.

The experimental setup can be seen in Figure 3. The measuring system was positioned successively in the two corners to the left and right of the capture point.

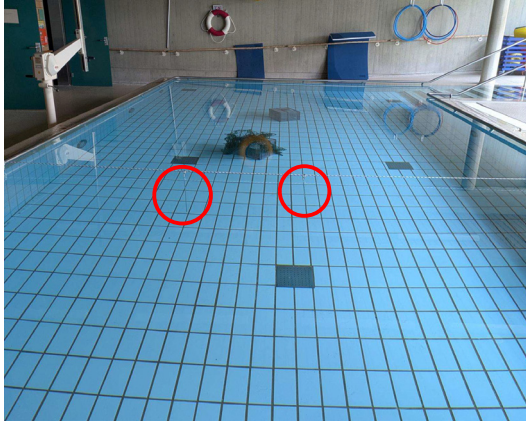


Figure 3: Image of the experimental setup. Front: Leash with two vertical metal rods, Middle: Box with net, Back: Box without net

Before presenting the scans from this experiment, the two coordinate systems used for the axis labels are briefly introduced. The parent coordinate system consists of the X and Y axes, which represent the sides of the pool, and the Z axis, which extends into the pool. Since the rail was positioned at an angle in a corner, a second coordinate system rotated relative to the main coordinate system was established for it. Here, the sonars are moved in the x_s direction, lie underneath each other on the z_s axis, and face in the y_s direction. The coordinate system definitions are illustrated in Figure 4.

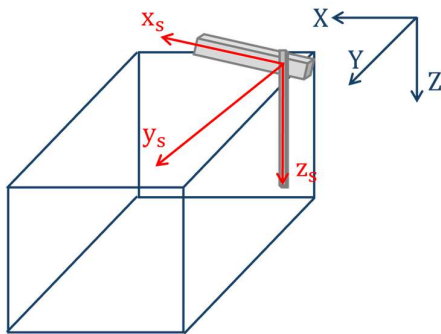


Figure 4: Schematic diagram of the parent coordinates (blue) and the measurement coordinates (red)

First Approach: Sonar Array

To use the array of single-beam sonar correctly, certain constraints must be met, which stem from the fundamental characteristics of sonars. The basic principle of a sonar is to measure the time t between transmitting and receiving acoustic signals. Based on the speed of sound c , which depends on temperature, pressure and salinity, Eq. (1) calculates the range l [1].

$$l = \frac{c \cdot t}{2} \quad (1)$$

The diameter d_v of the circular field of view changes depending on the distance from the sonar. To provide seamless data the fields of view of the sonar in the array must overlap. The diameter d_v of these circular fields of view can be calculated at a distance l by Eq. (2)

$$d_v = 2 \cdot l \cdot \tan\left(\frac{\alpha}{2}\right) \quad (2)$$

It becomes clear that a certain distance from the target is necessary or, conversely, a certain distance d_s between the sonars must not be exceeded, to achieve overlap, as in Figure 5.

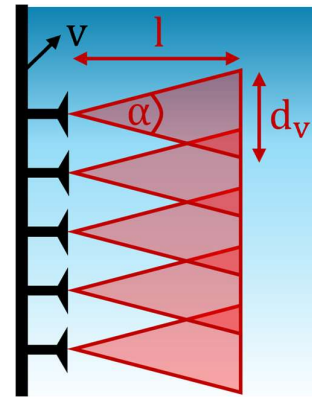


Figure 5: Drawing illustrating how the distance between the sonars, their opening angle, their distance to the target and the overlapping areas are related

Besides the overlap between adjacent sonars, consecutive pings of the same sonar must also overlap. This imposes an upper bound on the travel speed: within the ping interval t , the platform moves a distance $s = v \cdot t$, which must not exceed the footprint diameter d_v . The resulting speed limit is given by Eq. (3). Since d_v grows with the distance l to the target, the admissible travel speed increases accordingly.

$$v < \frac{2 \cdot l \cdot \tan\left(\frac{\alpha}{2}\right)}{t} \quad (3)$$

Under this condition, consecutive scans overlap as shown in Figure 6.

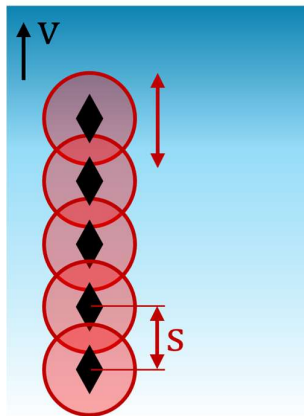


Figure 6: Drawing to illustrate how the measurement speed and the opening angle are related.

This knowledge must be applied to measurement systems. The single-beam sonars have a beam width α of 5° . In the case of the pool measurement, the nearest box - which was the focus of this study - was in a distance of 4.5 m, which would require a sonar distance of at least 39 cm. To have appropriate coverage, 18 cm was chosen.

In the case of the USV measurement vessel (as shown in Figure 1), the boat imposes certain constraints due to its twin hulls. With an even spacing between the five sonars, they can be mounted 22 cm apart from each other. This results in a minimum scan range of 5 m.

The speed of the linear axis can be specified individually. For a target at a distance of 4 m and a ping interval of 0.5 s for each sonar, this results in a maximum measurement speed of 0.78 m/s. In this study, a measurement speed of 0.7 m/s was used.

The USV has a minimum speed of 1 m/s. This results in a ping interval of 0.5 s. Consequently, the minimum distance to the targets is 5.7 m.

Second Approach: Sonar Combination

Effective fusion of different sonars requires their individual characteristics to be considered, most importantly the radiation pattern and the operating frequency.

Single-beam sonars, usually referred to as down-looking sonars due to their orientation, have a highly focused beam pattern. This results in a high degree of data clarity. However, a drawback is their limited coverage area. Side-scan sonars, on the other hand, have a very wide beam angle in one direction - usually perpendicular to the direction of travel - and a

very narrow beam angle in the direction of travel. This results in very high resolution in the direction of motion, while losing angular information. At the same time, their lateral field of view makes them particularly well-suited for scanning large areas. Combining the two sonar types yields both high depth resolution and precise lateral coverage [1,8].

In addition to correctly aligning the sonars and interpreting their fields of view, an understanding of the measured intensities is important for correctly combining the data.

The sonar equation determines the echo strength (EL) using the source level (SL), the transmission loss (TL), and the target strength (TS), as shown in the Eq. (4).

$$EL = SL - 2 * TL + TS \quad (4)$$

The strength of the echo is thus significantly influenced by the transmitters' power, the frequency (and thus the wavelength), the resulting absorption in the water, and the characteristics of the target [1]. Normalizing the data is therefore necessary to achieve a uniform greyscale distribution when superimposing scans.

The effect of proper normalization was investigated in the pool. The unique feature of the sonars used is that their data is scaled to the same maximum value. However, this results in scans with backgrounds of varying intensity depending on the target intensity for each individual scan. Each sonar must therefore be normalized individually and all sonars jointly against each other. To address this, the median of the measured values was calculated for each individual sonar measurement. This median corresponds to the value of the background, which can be used as a reference value. A scaling factor was then calculated to bring the individual scans onto a uniform scale. Figure 7 shows the improvement achieved. In the figure on the right, the data have been normalized as explained. The improvement in background uniformity across the scan is evident.

The more uniform background is clearly visible. This makes the targets stand out more clearly. At $R_1 = (0.0, 2.7)$ and $R_2 = (1.0, 2.6)$ thin metal rods are located (red circles), at $B_1 = (1.0, 4.5)$ there is a box surrounded by a net (blue circle); and at $B_2 = (1.5, 6.2)$ there is a box without a net (green circle).

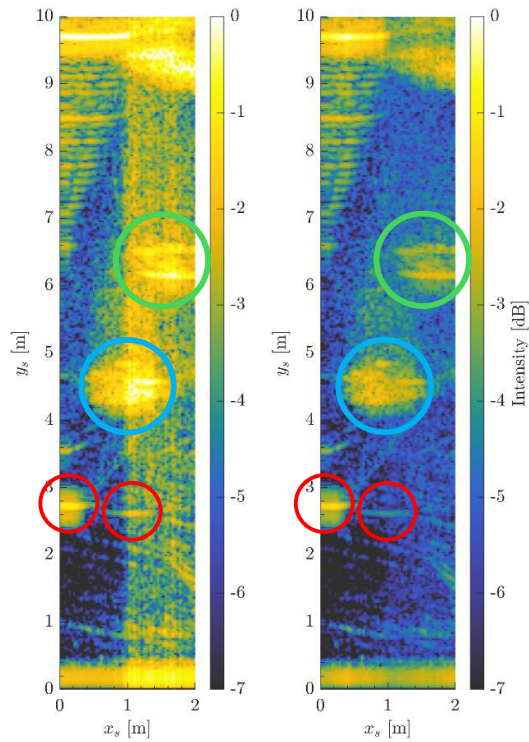


Figure 7: Comparison of the same data. The left one is not further processed while the right image has undergone normalization

In addition to a visual assessment of the improvement in scan quality, background uniformity, the CNR (Contrast to noise ratio) [9] and the SNR (Signal to noise ratio) [10] were calculated as quantitative evaluation metrics. When it comes to background uniformity, a low value is desirable, whereas a high value is advantageous for CNR and SNR [8,10].

For this purpose, the ranges $0.5 < y_s < 2.0$, $3.0 < y_s < 4.0$, and $5.0 < y_s < 6.0$ were defined as the background, and the position of the box with the net (blue circle) and the rod on the right were defined as targets. Background uniformity is calculated as the ratio of the standard deviation to the mean. The CNR is calculated as the ratio of the difference between the means of the target and the background to the standard deviation of the background [9]. The SNR, on the other hand, is calculated as the ratio of the signal mean to the standard deviation of the background [1].

The significant improvement in both metrics resulting from normalization is shown in Table 2. This is an example of the sonar shown in Figure 7, the other sonars had the same output.

Table 2: Comparison of the background uniformity, the CNR and the SNR of two targets between the original data and the normalized data of the sonar shown in Figure 7.

	Original data	Normalized data
Uniformity	0.35	0.27
CNR _{Rod}	1.01	2.55
CNR _{Box}	2.42	3.24
SNR _{Rod}	3.89	6.98
SNR _{Box}	5.31	6.29

However, when combining different sonar systems, the impact of normalization on the comparability and superimposition of two sonar scans is crucial. Figure 8 shows the superimposition of a single-beam sonar with a side-scan sonar. In the left case, the data is not normalized; in the right case, it is.

It becomes clear that adjusting the pixel values darkens the left side of the scan. This eliminates the false impression of the target on the left, stretching from $Y = 4.5$ until 8.

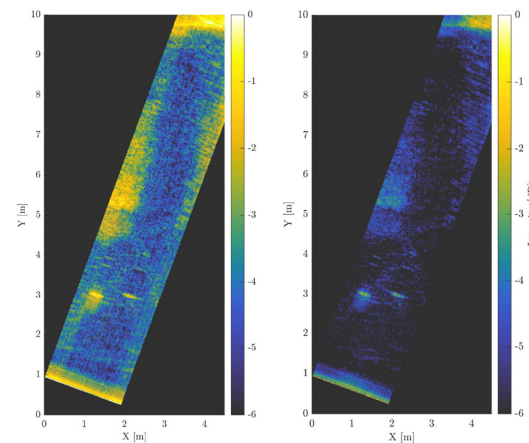


Figure 8: Comparison of the same fused data of one single beam and one side scan sonar. In the left the data is only added to each other, while in the right image had undergone normalization

Third Approach: Multilook

To achieve the third approach - the multilook method - precise georeferencing is necessary. More specifically, at every point in time when a measurement is taken, it must be known where the sonar is located and in which direction it is pointing.

With increasing distance to the target, it is also essential to consider the varying speed of sound in water. While an incorrect assumption of the speed in a single measurement merely results in a shift in the intensities, in a multilook measurement it prevents the correct superposition of the various measurements.

Several models exist for computing the speed of sound in water. Besides simplified approximations, the UNESCO, Del Grosso, and Wilson formulas are established as accurate references [1, 6].

To investigate the effect of using multiple viewing angles, data was collected from two corners of the pool. For correct overlay, each dataset was georeferenced individually and then fused pixelwise, either by multiplying or by adding the corresponding pixel values. In both cases they must be normalized afterwards to adjust the brightness, to prevent overestimation caused by multiple sightings. Figure 9 shows the results of the multiplication on the left and the addition on the right.

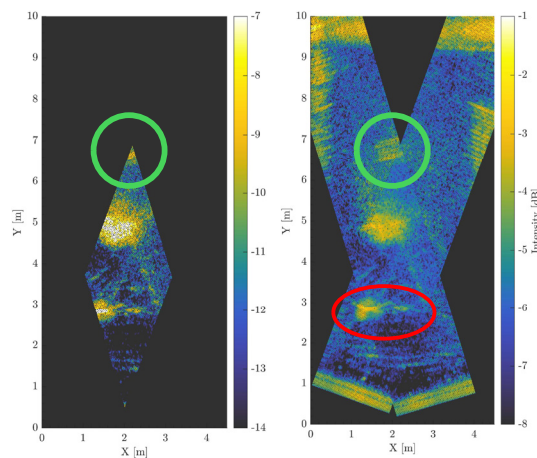


Figure 9: Comparison of multiplication (left) and addition (right) of the pixel values of two poses, measured by one single-beam sonar.

Returning to Figure 7, the metal rods appear as hyperbolic shapes typical of point-like scatterers in sonar data. The same rods can be seen in the right picture at $R_1 = (1.5, 2.8)$ and $R_2 = (2.2, 2.8)$, marked with a red circle. For R_1

one can see how the two hyperbolic shapes overlay to a cross, marking the real position of the rod and proving the correct fusion. Both ways of fusion do have advantages and disadvantages. Multiplication yields sharper targets, but any target visible from only one pose is lost — such as the upper box at $B_2 = (2, 6.8)$, marked in green.

This time, the improvement in data quality achieved by the multilook method was quantitatively verified in addition to the visual assessment. For this purpose, the two metal rods were defined as targets, and the CNR and SNR were determined again relative to them. This was performed for the individual sonars as well as for the combination of two directions. The results are shown in Table 3, again representative of the sonar depicted in Figure 8.

Table 3: Comparison of the SNR and CNR of the two rods between the data from a single sonar and the combination of them

Data set	SNR	CNR
Right corner single-beam	2.43	0.45
Left corner single-beam	3.60	1.41
Side-Scan	4.50	1.15
Both corners single-beam	6.52	2.12
All data	9.71	2.73

To combine these three approaches, array of multiple single beam sonars, combination of single-beam sonar and a side-scan sonar and the view out of multiple poses, the total data was combined. After georeferencing and normalization the 3D plot in Figure 10 was created. To make structures recognizable, a threshold was introduced to define which values should be rendered transparent. This makes the targets visible without completely losing the sense of space. Clearly visible are the two rods at $R_1 = (1.5, 2.8)$ and $R_2 = (2.2, 2.8)$ and the two boxes at $B_1 = (2.0, 5.0)$ and $B_2 = (2.0, 6.8)$. Thanks to the additional information about the length of the Z-axis, objects can also be distinguished from artifacts more effectively than is the case in a 2D representation.

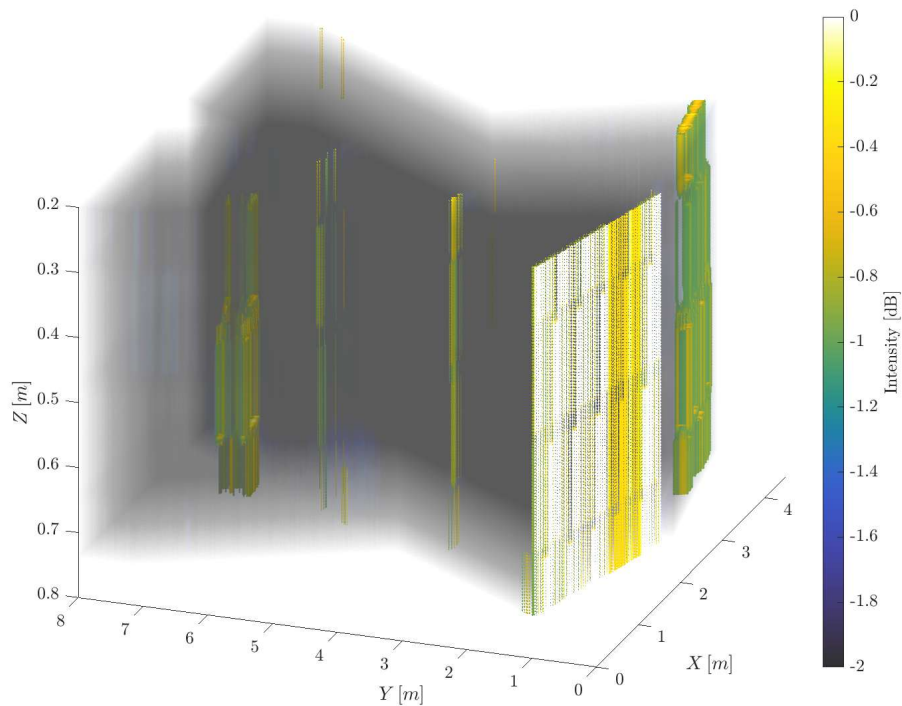


Figure 10: A total 3D view of the pool including the data from the different single beam sonar as well as from the side scan sonar from two different views.

Conclusion

This paper presented a three-stage processing pipeline for improving target clarity in sonar scans, validated in a pool experiment with a rail-based measurement system. First, several single-beam sonars were mounted one below the other, orthogonally to the direction of motion. This made it possible to cover a larger area without losing resolution. In the second stage, data from a side-scan sonar was additionally utilized. To ensure consistent use of the various sonars and account for varying conditions within a single scan, the data was normalized using the median of each scan. This resulted in more accurate target detection, which can be measured by improved background uniformity, CNR and SNR. The different fusion techniques based on multiplication and addition showed different positive features. While multiplication led to greater focus, addition provided a better overall picture. The third stage, data from two viewing poses were fused, which allowed targets to be localized more precisely and reduced artifacts even with this minimal number of perspectives. The CNR and SNR metrics improved significantly again. Combining all three stages

produced a consistent 3D representation of the pool in which every target could be identified.

Outlook

As outlined in the introduction, the long-term goal is an autonomous multimodal system capable of collecting data from real-world environments. On a USV, sonar positioning is affected by wave motion and GPS inaccuracies, in contrast to the well-defined geometry of the rail system. The use of GPS for coordinating acquisition and the presence of wave motion present additional challenges. Future work will focus on applying the findings to this system.

A second direction concerns the experimental setting itself. The pool environment does not reproduce the conditions encountered in the field such as varying bathymetry, sediment backscatter, currents, and biological clutter. Future experiments in natural water bodies will therefore be carried out to assess the robustness of the approach and to gain deeper insights into the detection and classification of ALDFGs.

Acknowledgment

The authors would like to thank the Friedrich-von-Bodelschwingh school for providing their pool for the measurements which were used in this paper.

References

- [1] X. Lurton, *An Introduction to Underwater Acoustics, Principles and Applications*; ISBN 3-540-42967-0, Springer Verlag, 2002
- [2] F.F.R. Merveille, B. Jia, Z. Xu, B. Fred, *Advancements in Sensor Fusion for Underwater SLAM: A Review on Enhanced Navigation and Environmental Perception*, *Sensors* 2024, 24; doi: <https://doi.org/10.3390/s24237490>
- [3] M. Morfin, S. Méhault, L. Miquerol, Q. C. Dufau, D. Kopp, Using citizen science to inventory and map abandoned, lost, or otherwise discarded fishing gear (ALDFG), *Marine Pollution Bulletin*, Volume 220, 2025; <https://doi.org/10.1016/j.marpolbul.2025.118418>
- [4] C. Wilcox, G. Heathcote, J. Goldberg, R. Gunn, D. Peel, B. D. Hardesty, *Understanding the sources and effects of abandoned, lost, and discarded fishing gear on marine turtles in northern Australia*, *Conservation Biology* Volume 29, Issue 1 pp. 198-206; doi: <https://doi.org/10.1111/cobi.12355>
- [5] M. Schumacher, M. Götz, G. Dederer, *Nutzung künstlicher Intelligenz zur Aufspürung von Geisternetzen in Küstengewässern*, *Journal of Applied Hydrography*, 03/2024; doi: 10.23784/HN127-04
- [6] F. J. Millero, T. Kubinski, *Speed of sound in sea water as a function of temperature and salinity at 1 atm*, *Journal of the Acoustical Society of America* 57, 312-319 (1975); doi: <https://doi.org/10.1121/1.380462>
- [7] <https://ceruleansonar.com/> (last visited: 15.04.2026)
- [8] H. Liu, X. Ye, H. Zhou, H. Huang, *A Mapping Method Fusing Forward-Looking Sonar and Side-Scan Sonar*, *Journal of Marine Science and Engineering* 13, no. 1: 166; <https://doi.org/10.3390/jmse13010166>
- [9] K. Hong-Gi, S. Jungmin, K. Soo Mee, *Underwater Optical-Sonar Image Fusion Systems*, *Sensors* 2022, 2, 8445; doi: <https://doi.org/10.3390/s22218445>
- [10] N. Z. Kolev, *Sonar Systems*; ISBN 978-953-307-345-3

Comparative Study of 10-mm Length Commercial Erbium Doped Optical Fibers for Temperature Sensing with a Fluorescence Decay Method

*Roman Kruglov, Katharina Rostan, Rainer Engelbrecht
Polymer Optical Fiber Application Center (POF-AC),*

Technische Hochschule Nürnberg Georg Simon Ohm, Wassertorstr. 10, 90489 Nuremberg, Germany

Abstract

This paper shows the experimental results on the temperature sensitivity of the fluorescence decay time in commercial rare-earth-doped silica fibers within the ambient temperature range from 0°C to 60°C by pumping with a laser diode at 976 nm. The average temperature sensitivity of the 10-mm fiber length sensors defined as a slope of a linear regression over the measured temperature range has achieved 2.51 $\mu\text{s}/^\circ\text{C}$ and 1.54 $\mu\text{s}/^\circ\text{C}$ for Er^{3+} -doped fibers with low and high ions concentration respectively. The average sensitivity of 1.99 $\mu\text{s}/^\circ\text{C}$ has been achieved with $\text{Yb}^{3+}/\text{Er}^{3+}$ -codoped fiber. A short-term precision calculated as an overall standard deviation of the measured fluorescence decay time within the ambient temperature range was 0.88 μs and 0.64 μs for Er^{3+} -doped fibers with low and high ions concentration; and 4.04 μs for $\text{Yb}^{3+}/\text{Er}^{3+}$ -codoped fiber respectively.

Keywords: Fiber-optic sensor, temperature sensing, rare-earth-doped fiber, fluorescence.

Introduction

Employing fiber-optic sensors for temperature monitoring ensures a precise measurement while withstanding harsh environments. Due to its small thermal mass, mechanical flexibility, electrical insulation and low loss, an optical fiber can be advantageous for temperature sensing in remote, hard accessible physical spaces. Well-established approaches for distributed and point sensing are mainly based on scattering methods, fiber Bragg gratings and fluorescence techniques [1]. The various rare-earth-doped waveguides and fibers are often applied as an intrinsic probe for temperature sensing, where the fluorescence intensity ratio (FIR) of two radiative transitions is analyzed or the fluorescence decay time is measured [2, 3]. Typical length of the sensing rare-earth-doped fibers may vary from a few centimeters to several tens of centimeters to provide appropriate fluorescence power as well as adequate pump absorption [2-4].

In this work we analyze and compare the temperature sensitivity of the fluorescence decay time in commercial rare-earth-doped silica fibers (LIEKKI® Er30-4/125, LIEKKI® Er110-4/125, Fibercore® DF1500Y) within the ambient temperature range from 0°C to 60°C by pumping with a laser diode at 976 nm. The length of the sensing fibers has been chosen as short as 10 mm allowing a single-point

probe implementation. These results aim to determine the potential of commercial fibers that are commonly used in telecommunication infrastructure for fluorescence-based thermometry at ambient temperatures, requiring higher sensitivity and precision as for reported high-temperature applications [4, 5].

Measurement Setup

The schematic diagram of the setup to measure the fluorescence decay time is shown in Fig. 1.

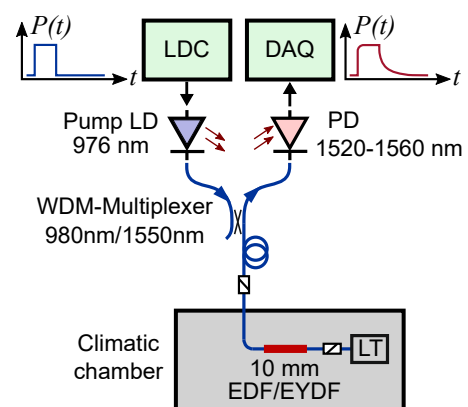


Fig.1: Schematic diagram of the experimental setup to measure a temperature dependent fluorescence decay time.

Tab. 1: Parameters of the rare-earth-doped silica fibers under test.

Fiber	LIEKKI® Er30-4/125 (Er ³⁺ -doped)	LIEKKI® Er110-4/125 (Er ³⁺ -doped)	Fibercore® DF1500Y (Yb ³⁺ /Er ³⁺ -codoped)
Parameter			
Mode Field Diameter @ 1550 nm, [μm]	6.5 ± 0.5	6.5 ± 0.5	5.3 - 6.8
Numerical Aperture (nominal)	0.2	0.2	0.2 - 0.24
Peak Core Absorption @ 1530 nm, [dB/m]	30 ± 3	110 ± 10	20
Concentration Er ³⁺ , [ions/m ³]	1.94 · 10 ²⁵	6.6 · 10 ²⁵	-
Concentration Er ³⁺ , [ppm wt]	2450	8320	-

Three types of single mode silica fibers doped with Er³⁺ ions (EDF) and Yb³⁺/Er³⁺ ions (EYDF) were tested in a climatic chamber (see Table 1).

An FBG-stabilized pump laser diode (LD) emitting at 976 nm with a maximum power of

300 mW was modulated directly through an external analog modulation input of the laser diode controller (LDC). The duration and period of the rectangular modulation pulses were set to 20 ms and 100 ms respectively. A 10-mm length rare-earth-doped silica fiber was spliced to the lead-in HI1060 fiber, while the far end of the sample was spliced to an FC/APC fiber pigtail and secured with a fiber optic light trap (LT). The length of the fiber under test is reasonable for the point sensor implementation, and still provides a sufficient fluorescence emission. A fused fiber optic wavelength division (WDM) multiplexer 980 nm / 1550 nm was required to launch a pump light into the rare-earth-doped fiber and split a backward fluorescence decay signal to an InGaAs photodetector (PD) operating in a photoconductive mode. The output signal, amplified by a transimpedance amplifier, was then acquired by an oscilloscope with an averaging over 10 waveforms.

Calibration of probes

First, the spectral fluorescence intensity profiles have been measured with an optical spectrum analyzer (OSA) Yokogawa AQ6370D in a backward propagation direction with respect to temperature and pump power P_p (see Fig. 2, a-c).

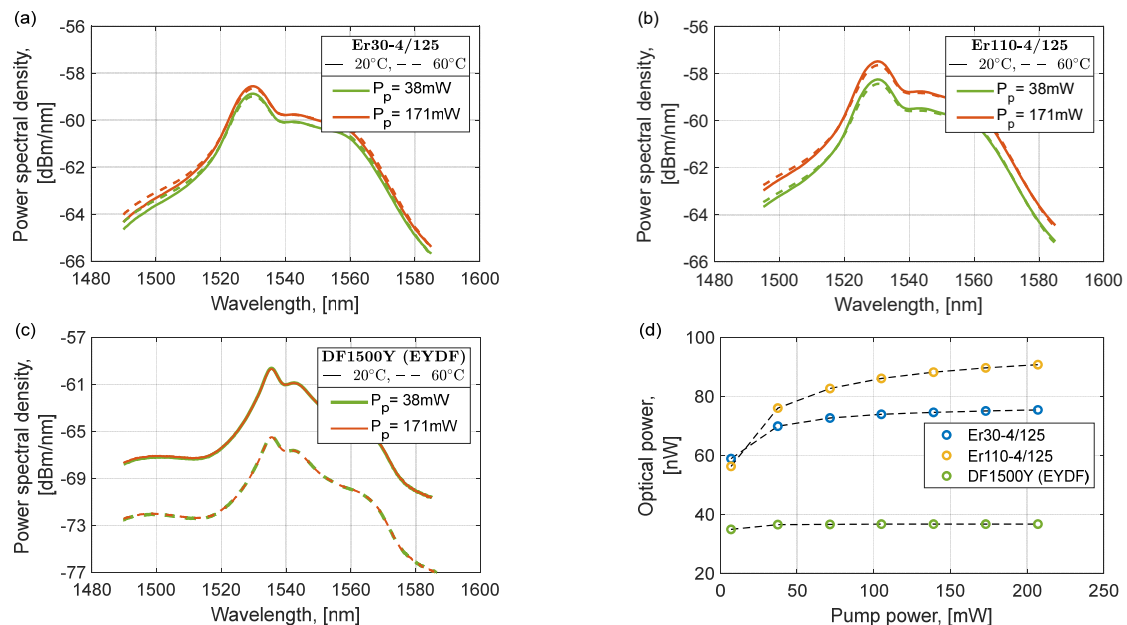


Fig. 2: Spectral fluorescence intensity profiles of the rare-earth-doped fibers by pumping at 976 nm (a-c); Fluorescence intensity near 1550 nm as a function of the pump pulse power measured at a temperature of 20 °C (d).

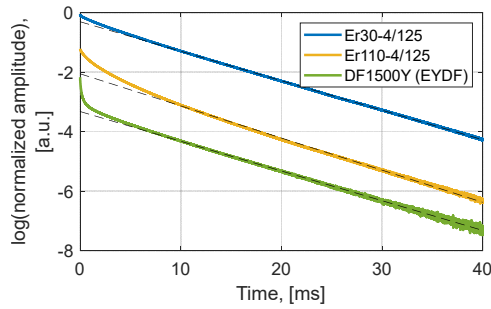


Fig. 3: Normalized fluorescence decay signals measured at a temperature of 20 °C and pump pulse power of 58 mW.

Both EDF demonstrated a similar behavior. The spectral change with temperature is related to the temperature dependent nonradiative relaxation of erbium ions [2] as well as the thermal distribution of populations between the components of the $^4I_{15/2}$ and $^4I_{13/2}$ levels [6]. The overall fluorescence intensity in EDF rises with a pump power indicating that the fibers are not saturated yet (see Fig. 2, d). The EYDF spectra demonstrated a stronger temperature dependence of the spectral fluorescence profile, and at the same time, the fiber was saturated by the pump power above 40 mW. The $\text{Yb}^{3+}/\text{Er}^{3+}$ -dopant system results in a dual-band fluorescence (Fig. 2, c shows the spectral fluorescence profile of erbium ions only).

The ratio between emission intensities in two bands varies with temperature, since the energy transfer between erbium and ytterbium ions is thermally dependent [2]. Thus, the EYDF can be potentially used for FIR thermometry.

Fig. 3 shows an example of the fluorescence decay signals for three fiber samples measured at a temperature of 20 °C after the pumping pulse has been turned off. The logarithmic scaling of the amplitude axis makes it clear that the decay signals cannot be described with a single exponential function (shown by the dashed lines in Fig. 3) since an additional faster fluorescence process is noticeable within the first 2-4 ms. This rapid initial decay may be attributed to stimulated emission. Additionally, a wide emission spectrum of ytterbium ions in EYDF may affect the decay signal by leakage through a WDM multiplexer.

The time constant related to the radiative transition from level $^4I_{15/2}$ to ground level $^4I_{13/2}$ has been estimated with the nonlinear least squares method by fitting a single exponential model to the decay curve captured by an oscilloscope and averaged over 10 consecutive waveforms with a period of 100 ms. The application of the fitting interval starting 2 ms after the pumping pulse had been turned off allows to exclude the rapid initial fluorescence decay from the evaluation, and thus ensures the simplicity of the model.

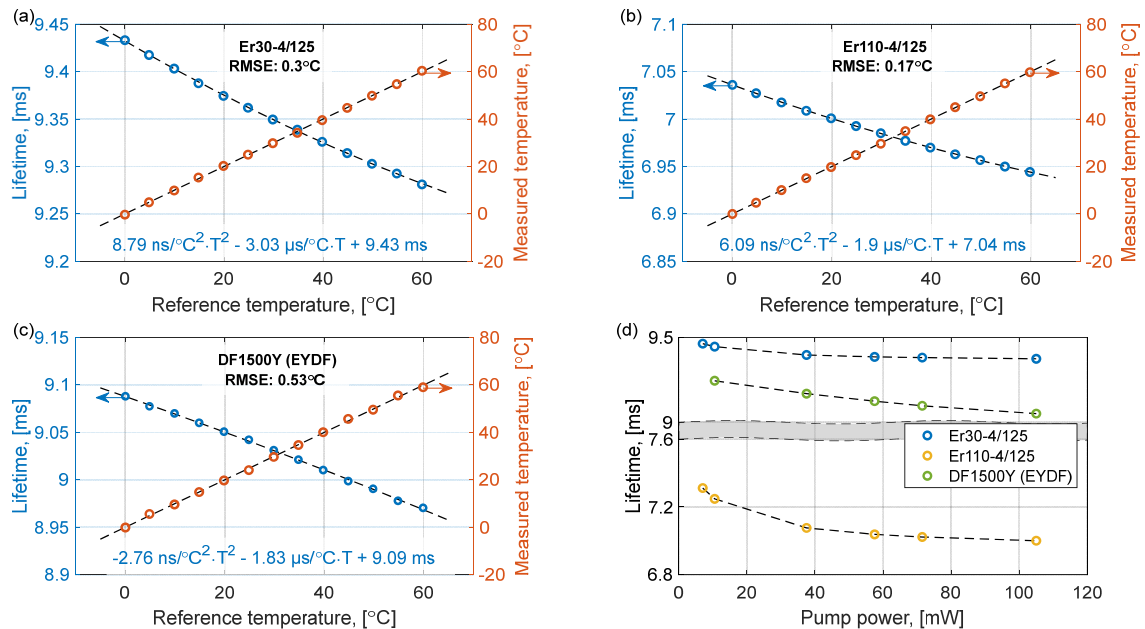


Fig. 4: Calibration curves of the rare-earth-doped fibers as a temperature sensor by the pump pulse power of 105 mW within the ambient temperature range from 0 °C to 60 °C (a-c); Fluorescence decay lifetime as a function of the pump pulse power measured at temperature of 20 °C (d).

The calibration of the fibers as a temperature sensor shown in Fig. 4 has been performed by measuring the fluorescence lifetime as a function of temperature within the range from 0°C to 60°C by the pump pulse power of 105 mW. The lifetime measurements at every temperature level have been repeated 10 times with a one minute interval allowing an estimation of a short-term precision. The curves have then been fitted with a 2nd degree polynomial whose coefficients are also shown in the figure. The root mean square error (RMSE) for 2nd degree calibration polynomials was as low as 0.3°C, 0.17°C and 0.53°C for EDF Er30-4/125, Er110-4/125 and EYDF DF1500Y respectively. The average temperature sensitivity of the probes (S_T) calculated as a slope of a linear regression over the range from 0°C to 60°C achieved 2.51 $\mu\text{s}/^\circ\text{C}$, 1.54 $\mu\text{s}/^\circ\text{C}$ and 1.99 $\mu\text{s}/^\circ\text{C}$ respectively. A short-term precision defined as an overall standard deviation of the measured fluorescence decay time and computed over the total number of trials and temperature levels was 0.88 μs , 0.64 μs and 4.04 μs respectively.

By repeating the calibration measurements with a pump pulse power ranging from 10 mW to 105 mW the variation of the average temperature sensitivity for EDF probes remained below 1 %; and for EYDF probe it was 5 %. All fibers under test have demonstrated a similar behavior of the fluorescence lifetime depending on a pump power that is approaching an asymptotic value (see Fig. 4, d). The minimal lifetime variation with a pump power was observed for EDF Er30-4/125.

Conclusion

Characterization and comparison of the temperature behavior of commercial Er^{3+} -doped and $\text{Yb}^{3+}/\text{Er}^{3+}$ -codoped fibers within the ambient range from 0°C to 60°C has been reported in the paper. The measurements have been carried out with 10-mm fiber length samples enabling the design of the compact, intrinsic probes for point temperature sensing. It has been shown that the fluorescence lifetime as well as the fluorescence intensity depend on temperature and pumping power. At the same time, the average temperature sensitivity of the probes has been remained nearly constant within the wide range of the pump power. The fiber Er30-4/125 with a lower Er^{3+} -concentration has demonstrated higher average temperature sensitivity of 2.51 $\mu\text{s}/^\circ\text{C}$ with an overall standard deviation of the measured fluorescence decay time of 0.88 μs and the considerably weaker dependency of

the fluorescence lifetime on the pump power. The fiber calibration curves have been fitted with 2nd degree polynomials with RMSE of 0.3°C, 0.17°C and 0.53°C for EDF Er30-4/125, Er110-4/125 and EYDF DF1500Y respectively.

Acknowledgements

This research was funded in part by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) as Forschungsimpuls FIP 8/1 – 2024, Project number 528480942 and the European Union by means of the European Regional Development Fund (EFRE), Project “BaLista”.

References

- [1] P. Lu, N. Lalam, M. Badar, B. Liu, B. T. Chorpining, M. P. Buric, P. R. Ohodnicki, “Distributed optical fiber sensing: Review and perspective,” *Appl. Phys. Rev.*, 6(4), 041302 (2019); doi: 10.1063/1.5113955
- [2] S. Baek, Y. Jeong, J. Nilsson, J.K. Sahu, B. Lee, “Temperature-dependent fluorescence characteristics of an ytterbium-sensitized erbium-doped silica fiber for sensor applications,” *Optical Fiber Technology*, 12(1), 10-19 (2006); doi: 10.1016/j.yofte.2005.04.002
- [3] G. Paez, V. Lopez, M. Strojnik, “Variable time constant of erbium-doped temperature sensor,” *Proc. SPIE 5152, Infrared Spaceborne Remote Sensing XI*, (2003); doi: 10.1117/12.506862
- [4] D. I. Forsyth, T. Sun, K. T. V. Grattan, S. A. Wade, S. F. Collins, “Characteristics of doped optical fiber for fluorescence-based fiber optic temperature systems,” *Rev. Sci. Instrum.*, 74(12), 5215-5218 (2003); doi: 10.1063/1.1623624
- [5] Z. Y. Zhang, K. T. V. Grattan, A. W. Palmer, B. T. Meggitt, T. Sun, “Fluorescence decay-time characteristics of erbium-doped optical fiber at elevated temperatures,” *Rev. Sci. Instrum.*, 68(7), 2764-2766 (1997); doi: 10.1063/1.1148192
- [6] N. Kagi, A. Oyobe, K. Nakamura, “Temperature dependence of the gain in erbium-doped fibers,” *J. of Lightw. Technol.*, 9(2), 261-265 (2002); doi: 10.1109/50.65884

Tool-Center-Point-Kalibrierung eines Werkzeuges für Kaltschmiedeprozesse

Benedikt Loris¹, Markus Gapp¹, Stefan J. Rupitsch², Tobias Weiser¹

¹Hochschule Kempten, Institut für angewandte KI und Robotik, Schwabenstraße 102, 87616 Marktoberdorf

²Universität Freiburg, Institut für Mikrosystemtechnik – IMTEK, Georges-Köhler-Allee 201 D-79110 Freiburg

Zusammenfassung

Kaltschmiedeprozesse finden vorrangig Anwendung bei der Bearbeitung von Oberflächen aus Metall. Bei der Verwendung des Werkzeugs an einem Roboter ist der Tool-Center-Point (TCP) des Roboters für eine akkurate Bahnbewegung des Werkzeugs erforderlich. Da der Prozess mit einer Halbkugel realisiert wird, sind die in der Robotersteuerung verfügbaren Möglichkeiten zur TCP-Ermittlung nur bedingt geeignet. Mittels einer Low-Cost-Kamera bei gleichzeitig hoher Auflösung wird in diesem Beitrag der TCP durch ein externes System hochgenau ermittelt.

Keywords: Tool-Center-Point-Kalibrierung, Kaltschmiedeprozess, CMOS-Low-Cost-Kamera

Einführung

Kaltschmiedeverfahren werden verwendet, um Oberflächen zu vergüten. In diesem Kontext sind die Aspekte einer automatischen Bearbeitung von Interesse, beispielsweise unter Einsatz einer CNC-Maschine oder eines Roboters. Dabei ist der reduzierte Energieaufwand für den Prozess von Relevanz, bei gleichzeitig kleinflächiger Bearbeitung prozessrelevanter Oberflächenbereiche.

Die Ausführung von Kaltschmiedeprozessen erfolgt in Abhängigkeit des Materials mit unterschiedlichen Kugeldurchmessern [1]. Es konnte festgestellt werden, dass die Position der Halbkugel einen signifikanten Einfluss auf die Qualität des Prozesses hat. Im einfachsten Fall werden die verfügbaren Werkzeuge von einem Roboter geführt, dessen "Tool-Center-Point" (TCP) mit möglichst hoher Präzision bestimmt werden muss (vgl. [2]). Im Gegensatz zu [3], in dem kein physikalischer TCP zum Einsatz kommt, beschreibt dieser Beitrag eine Methode, mit der sich die Kugel effizient und mit einer kostengünstigen Kamera berührungslos [4] einmessen lässt.

Hardware- und Netzwerkarchitektur

Zu diesem Zweck wurde das Kaltschmiedewerkzeug "Forgefix" [5] der Firma Pokolm an einem KUKA KR3 R540 adaptiert.

Der Kameraschlitten zur TCP-Vermessung ist seitlich innerhalb der Zelle positioniert. Die

Kamera ist auf einer 3D-gedruckten Rundschiene montiert, die an ihrer Oberseite einen Zahnkranz aufweist. Die Bewegung der Kamera erfolgt auf einem Radius von ca. 400 mm über einen Winkel von 35°. Der Prozess wird durch einen Schrittmotor, zwei Zahnriemenscheiben, einen Zahnriemen und ein gedrucktes Zahnrad auf einer kugelgelagerten Welle realisiert. Der Verfahrweg der Kamera wird durch zwei Endschräler begrenzt, die darüber hinaus zur wiederholten Positionierung der Kamera für die Stereo-Bildaufnahme dienen.

Die Initiierung der Verfahrbewegung erfolgt durch einen Raspberry Pi, welcher mit einer Codesys-Applikation ausgestattet ist und eine Kommunikation mit einem Arduino ermöglicht. Gleichzeitig fungiert er als Kommunikationsschnittstelle zur Robotersteuerung, um Variablen per PROFINET auszutauschen, sowie zur Interaktion mit der installierten Kamera-Visualisierung. An den Arduino sind ein Schrittmotortreiber und die beiden Endschräler angeschlossen, um die Bewegung zu realisieren. Die vorliegende Kamera ist mit einem USB-Anschluss ausgestattet und ermöglicht die Aufnahme von Bildern mit einer maximalen Auflösung von 3264 x 2448 Pixeln bei einer Bildrate von 15 Bildern pro Sekunde im MJPEG- oder YUY2-Format.

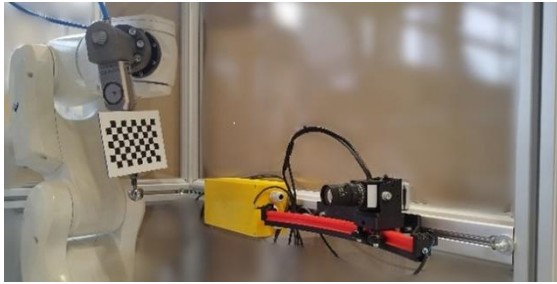


Abb. 1: Roboterzelle mit Kameraschlitten, Werkzeug und Kalibriermuster.

Aufbau der Software-Architektur

Das entwickelte Programm umfasst die Funktionen, Bildaufnahme und -verarbeitung, Kamerakalibrierung sowie die Kommunikation mit einer externen Steuerung über OPC UA. Die Struktur ist objektorientiert aufgebaut und nutzt eine grafische Benutzeroberfläche, die auf PySide6 basiert.

Zunächst wird die Kamera automatisch erkannt und initialisiert. Im Anschluss erfolgt eine kontinuierliche Aufnahme von Bilddaten mit definierter Auflösung und Bildrate. Die aufgenommenen Bilder werden gedreht und über ein Signal an andere Programmteile weitergeleitet. Darüber hinaus besteht die Option, Einzelbilder gezielt abzuspeichern, wobei zwischen Kalibrierungsbildern und sogenannten TCP-Bildern differenziert wird. Die Speicherung erfolgt in einer strukturierten Ordnerhierarchie, die durch eine fortlaufende Nummerierung gekennzeichnet ist.

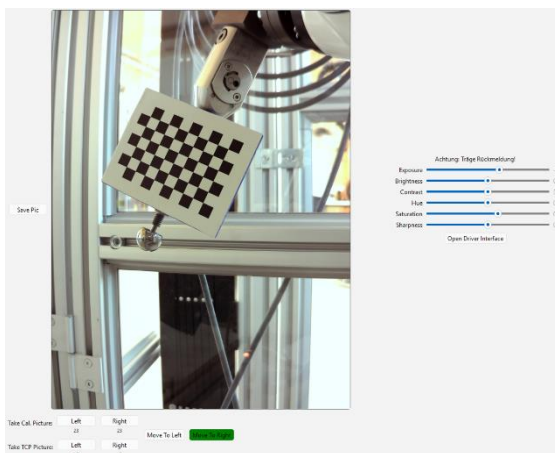


Abb. 2: „Image Capture“ der Visualisierung: Einstellen der Kameraparameter, Speicherung der Bilder für die Kalibrierung und Trigger des Kameraschlittens.

Parallel dazu wird eine Bildverarbeitung durchgeführt. Innerhalb eines auswählbaren

Bildausschnitts (Region of Interest, ROI) wird eine mehrstufige Verarbeitung durchgeführt. Diese umfasst die Grauwertumwandlung, die Kontrastverbesserung mittels CLAHE, die Glättung, die Kantendetektion über Sobel-Operatoren sowie die binäre Segmentierung durch Schwellwertverfahren. Im Anschluss sind morphologische Operationen wie Lochfüllung und Rauschunterdrückung vorgesehen.

Die Bestimmung geometrischer Eigenschaften erfolgt auf der Basis des resultierenden Binärbildes (Abb. 3). Dabei werden der Schwerpunkt sowie die Hauptachsen unter Verwendung der Kovarianzanalyse berechnet. Im Anschluss erfolgt eine Rotation des Bildes, wobei ein Referenzbild zur Bestimmung der optimalen Orientierung herangezogen wird. Nach der Ausrichtung werden Kanten extrahiert und relevante Punkte selektiert, um schließlich eine Kreisapproximation mittels Ausgleichsrechnung durchzuführen. Aus diesen Werten werden anschließend der Mittelpunkt und der Radius berechnet und sowohl lokal als auch global im Bildkoordinatensystem angegeben.

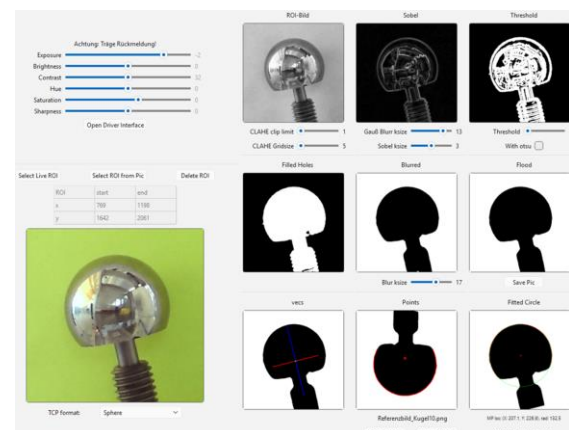


Abb. 3: „Image Processing“ innerhalb der Visualisierung: Extrahieren der Kugelgeometrie im ROI-Bild durch anpassbare Parameter und Anzeige der Pixel- und Bildkoordinaten.

Ein weiterer essenzieller Bestandteil ist die Kamerakalibrierung (Abb. 4). In diesem Verfahren werden anhand von Schachbrettmustern sowohl intrinsische Parameter, wie etwa die Verzerrungskoeffizienten und die Kameramatrix, als auch extrinsische Parameter für ein Stereokamerasystem bestimmt. Darüber hinaus besteht die Möglichkeit, die Kalibrierdaten zu speichern und zu laden. Auf der Basis der Kalibrierung ist auch eine 3D-Rekonstruktion durch Triangulation möglich.

Image Capture		Image Processing		Calibration			
Mono Calibration	Error: Reprojektionserror: 0.4528	Camera Matrix:	4553.887	0.000	1522.766	Save to File	
			0.000	4555.775	1558.091		
			0.000	0.000	1.000		
			Load from File				
Stereo Calibration	Error: Reprojektionserror: 1.0125	R:	0.810	-0.006	0.587	T: -0.225	Save to File
			0.004	1.000	0.004		
			-0.587	-0.001	0.001		
			0.068				
			Load from File				

Abb. 4: „Calibration“ der Visualisierung: Durchführen der Kamerakalibrierung (Mono und Stereo) mit Fehlerbetrachtung, Anzeige der Kalibrierwerte und Speichern/Laden der Kalibrierdaten.

Die Anwendung verwendet zudem eine Schnittstelle zu einem OPC-UA-Server. Es besteht die Möglichkeit zur Steuerung des Kameraschlittens (links/rechts).

Die grafische Benutzeroberfläche dient der Anzeige des Livebildes, der Zwischenschritte der Bildverarbeitung sowie der Parametrierung. Mittels diverser Bedienelemente können Kameraeinstellungen angepasst, Bildverarbeitungsparameter modifiziert, Kalibrierungen initialisiert und Bilder aufgenommen werden. Darüber hinaus erlaubt die Oberfläche die Selektion von Bildausschnitten sowie die Anzeige berechneter Ergebnisse, wie etwa des ermittelten Mittelpunkts.

Methode und Datenerfassung

Zunächst erfolgt die Bestimmung der Kameramatrix über ein Schachbrett-Kalibriermuster am Werkzeugschaft. Zu diesem Zweck wird der Roboter auf verschiedene Positionen und Orientierungen verfahren und anschließend jeweils ein Bild in beiden Kamerapositionen aufgenommen. Die Bestimmung der Kameraparameter erfolgt durch einen Algorithmus von OpenCV [6]. Die zugrunde liegende Kalibriermethode nach Zhang [7], Gl. (1), bildet die Basis dieses Algorithmus.

$$s\tilde{m} = A \cdot [R \quad t] \cdot \tilde{M}$$

mit s - Skalierungsfaktor,
 $\tilde{m}$ - Bildpunkt in homogenen Koordinaten,
 A - intrinsische Kameramatrix,
 R, t - Rotation und Translation,
 $\tilde{M}$ - 3D-Weltpunkt

Die intrinsische Matrix lautet (Gl. 2):

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

mit α, β - Brennweite in Pixeln,
 γ - Scherparameter,
 u_0, v_0 - Hauptpunkt

In der Folge wird der Prozess in einer iterativen Abfolge wiederholt, wobei die Bildaufnahmen über definierte Kamerapositionen (Endschalter am Aufbau) erfolgen. Nach Abschluss des Speichervorgangs der Bilddateien erfolgt die Stereokalibrierung. Zu diesem Zweck werden mit dem Roboter verschiedene Posen angefahren und jeweils ein Bild in der linken und rechten Endlage mit der Kamera erfasst. Die konstruktive Gestaltung des Werkzeugs ermöglicht die Erfassung von Tiefeninformationen aus den Aufnahmen. Im Rahmen einer 4-Punkt-Kalibrierung des TCP kann nach jeder Umorientierung des Roboters die Position der Mitte der Werkzeug-Halbkugel bestimmt werden. Gemäß der vorliegenden Evidenz wird der Roboter nach einer erfolgten Umorientierung mittels eines Programmaufrufs und übergebenen Variablen wieder auf den ursprünglichen Mittelpunkt bewegt. Die Ergebnisse der TCP-Kalibrierung werden mit einem Lasertracker der Fa. API verifiziert.

Ergebnisse

Die Abweichung der TCP-Vermessung von 0,17 mm beim Abgleich mit der Kamera wird in Tab. 1 dargestellt. Dieser Wert ist um mindestens den Faktor zwei kleiner als derjenige, der mit dem Roboter selbst ermittelt werden kann.

Tab. 1: Ergebnis der KUKA 4-Punkt-Kalibrierung mit Transformation, die Abweichung liegt bei 0,17 mm.

mm	x	y	z	Fehler
TCP	198,42	0,3	61,67	0,17

Für die Messung mit dem Lasertracker wurde ein Referenzbild der Werkzeugkugel aufgenommen und dessen Mittelpunkt bestimmt. Nach der manuellen Umorientierung des Werkzeugs erfolgte die Vermessung des Kugelschwerpunkts mit der Kamera. Die Position wurde so lange korrigiert, bis die Abweichung zwischen Referenzmittelpunkt und neuem Mittelpunkt entsprechend der Kamera in allen drei Achsen einen Wert von kleiner als 0,02 mm hat. Im Rahmen der Untersuchung wurde ein Vergleich der beiden Positionen der Kugel unter Zuhilfenahme des Lasertrackers durchgeführt. Dabei konnte eine Abweichung von 0,1 mm (Abb. 5) festgestellt werden.

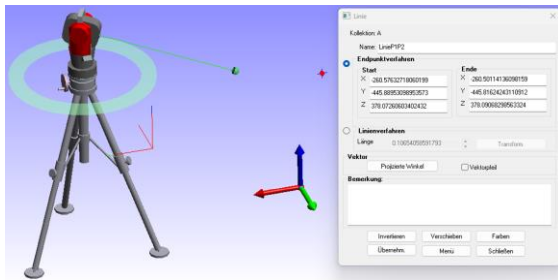


Abb. 5: Vergleichsmessung mit dem Lasertracker, die Abweichung liegt bei 0,1 mm.

Diskussion

Die vorliegende TCP-Vermessung ermöglicht eine mindestens zweifache Verbesserung der Ermittlung des TCPs. Im Vergleich zu Koordinatenmessmaschinen, Messarmen oder optischen Systemen ist der verwendete Aufbau erheblich kostengünstiger. Es besteht die Möglichkeit, dass dieser Aufbau bei Roboterzellen mit einer regelmäßigen TCP-Prüfung dauerhaft verbaut ist.

Die Vergleichsmessung mit einem Lasertracker zeigt eine Abweichung in ähnlicher Größenordnung und bestätigt daher den kamerabasierten Ansatz.

Die Anforderungen an die Genauigkeit des Messsystems hängen von den geforderten Prozessparametern ab. In der Hubrichtung des Werkzeuges ist eine Abweichung weniger kritisch, da das Werkzeug selbst einen Hub von ca. 2 mm zurücklegen kann. Wichtiger ist die Genauigkeit in Bahnrichtung, da der Zeilenabstand des Werkzeuges hier einen Wert von 0,1 mm erreichen kann.

Fazit und Ausblick

Die vorgestellte Methodik zur kamerabasierten Tool-Center-Point-Kalibrierung eines Kaltschmiedewerkzeugs zeigt, dass durch den Einsatz einer kostengünstigen CMOS-Kamera eine hochgenaue und berührungslose Vermessung des TCPs realisiert werden kann. Im Vergleich zur manuellen 4-Punkt-Kalibrierung kann die Genauigkeit deutlich gesteigert werden.

Die Verifikation mittels Lasertracker bestätigte die erreichte Messgenauigkeit sowie die Eignung des entwickelten Ansatzes für Kaltschmiedeprozesse. Darüber hinaus konnte gezeigt werden, dass sich der entwickelte Aufbau in eine bestehende Roboterzelle integrieren lässt und eine automatisierte Kalibrierung ermöglicht. Besonders die

Kombination aus stereoskopischer Bildaufnahme, Bildverarbeitung und Triangulation erwies sich als geeignet, um die Position der Werkzeugkugel reproduzierbar zu bestimmen.

Im Rahmen zukünftiger Arbeiten könnten weitere Optimierungen des Systems untersucht werden. Dabei ist zu evaluieren, inwieweit eine Reduzierung der Bildauflösung bei vergleichbarer Genauigkeit zu einer geringeren Rechenzeiten beitragen kann. Zusätzlich soll geprüft werden, ob durch eine Mittelwertbildung über mehrere Livebilder eine weitere Stabilisierung und Verbesserung der Messergebnisse erreicht werden kann. Weiterhin ist der Einfluss der Größe des verwendeten Schachbrett-Kalibriermusters auf die Genauigkeit der intrinsischen und extrinsischen Kamerakalibrierung zu analysieren.

Literaturnachweis

- [1] Wai Luen Chan, Henry Kuo Feng Cheng, Hammer peening technology—the past, present, and future; Springer, The International Journal of Advanced Manufacturing Technology (2022), doi: 10.1007/s00170-021-07993-5
- [2] Albert Nubiola, IlianA.Bonev, Absolute calibration of an ABB IRB 1600 robot using a lasertracker, Robotics and Computer-Integrated Manufacturing 29 (2013) 236–245, doi: 10.1016/j.rcim.2012.06.004
- [3] Cheng-Kai Huang, Hao-Yun Zhan, A novel calibration method for robotic arm tool center point without a physical tool apex, The International Journal of Advanced Manufacturing Technology (2025), doi: 10.1007/s00170-024-14906-9
- [4] Fakherddine Fares, Haïfa Souifi, Yassine Bouslimani, Mohsen Ghribi, Tool Center Point Calibration Method for an Industrial Robots based on Spheres Fitting Method, IEEE International Symposium on Robotic and Sensors Environments (2021), doi: 10.1109/ROSE52750.2021.9611759
- [5] Forgefix, Pokolm, Produkte, <https://www.pokolm.com/produkte/forgefix> [abgerufen am 29.04.2026]
- [6] OpenCV, Camera Calibration, https://docs.opencv.org/4.x/dc/dbb/tutorial_py_calibration.html [abgerufen am 29.04.2026]
- [7] Z. Zhang, A Flexible New Technique for Camera Calibration, IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 22, no. 11 (2000) 1330-1334, doi: 10.1109/34.888718

Development of an Interactive Search System for Standardised Documents in Legal Metrology and Comparison with a Retrieval-Augmented Generation Approach

Zihao Wang¹, Marko Esche¹, Martin Nischwitz¹, Hans Grönniger²

¹Physikalisch-Technische Bundesanstalt, Department 8.5 „Metrological IT“, Abbestr. 2-10, 10587 Berlin, Germany, Tel.: 030 3481 7975, E-Mail: marko.esche@ptb.de

²Ostfalia University of Applied Sciences, Faculty of Computer Science/IT, Am Exer 2, 38302 Wolfenbüttel, Germany, Tel.: 05331 939 31190, E-Mail: h.groenniger@ostfalia.de

Abstract

Modern measurement systems are becoming increasingly diverse and complex, with software playing an ever more important role. Ensuring the compliance of software protection features with legal requirements has thus become a core task in legal metrology. Focusing on technical solutions harmonised in WELMEC Guide 7.3, this paper develops and evaluates two types of document content retrieval systems: a custom interactive search system based on automated scripts, database, and a search engine, which supports structured and full-text search, fuzzy matching, and relevance ranking; and a Retrieval-Augmented Generation (RAG) approach that integrates dense vector retrieval with large language models, enabling natural language queries. Experimental results demonstrate that the interactive search system achieves stable and accurate performance in locating target entries and handling complex queries, whereas the RAG system, which relies heavily on model quality and prompt design, often struggles to maintain accuracy and stability. This paper confirms the practical value of the interactive search system in supporting software examination and provides a reference for the applicability of the RAG approach.

Keywords: Document Content Retrieval, Interactive Search System, Retrieval-Augmented Generation (RAG), Large Language Model (LLM), Legal Metrology

Introduction

The Physikalisch-Technische Bundesanstalt (PTB) is the National Metrology Institute of Germany and, among other tasks, is responsible for the evaluation of software in electronic measuring instruments, e.g., taximeters, electricity meters, heat meters, based on the guides (hereafter referred to as “Guide”) published by the European Cooperation in Legal Metrology (WELMEC). Among these, the Software Guide 7.2:2025 [1] serves as a key reference both for software examiners and for manufacturers of measuring instruments when preparing their documentation for conformity assessment.

Following the current revision of Guide 7.3, the new version of Guide 7.3 primarily covers the acceptable solutions transferred and adapted from the central requirement document Guide 7.2, with each solution having four keyword attributes. A typical use case involves developers and examiners locating and consulting all relevant acceptable solutions based on specific keywords. However, the guide is avail-

able only in paper format or as a PDF file, contributing to inefficient information retrieval.

To further advance the digital transformation of standards, the International Electrotechnical Commission (IEC) and the International Organization for Standardization (ISO) jointly introduced SMART Standards, which stands for Standards that are Machine Applicable, Readable, and Transferable. This concept has been widely adopted by multiple organisations and initiatives. Its most fundamental requirement is to use machine-readable formats, e.g., XML or JSON, to store document content, enabling automated processing and facilitating subsequent applications [2].

Among the vast collection of standardised documents, a significant portion serves only for content retrieval, without requiring integration or interaction with machine systems. How to efficiently retrieve the content of such documents has become an important research topic. Although custom-built database search systems can satisfy the retrieval needs of different standardised documents, they suffer from high development and maintenance costs

and limited flexibility when adapting to varying content structure or query requirements. To reduce the cost of manual processing, there is a growing need to leverage AI technology to enable more intelligent document content retrieval.

Dense retrieval methods that combine text embeddings and vector-based queries have been widely applied in information retrieval tasks [3, 4]. However, relying solely on semantic matching makes it difficult to model logical relationships between contexts, which imposes limitations when querying structured information. Recent advances in large language models (LLMs) for natural language understanding, reasoning, and generation [5] have spurred the development of Retrieval-Augmented Generation (RAG) methods [6, 7]. These methods provide the retrieved relevant passages as context to LLMs, enabling them to answer natural language queries about documents. This not only improves the accuracy of results [8, 9] but also reduces the risk of producing outputs that appear plausible but do not align with the provided source inputs or factual information, i.e., hallucination [10].

This paper focuses on transforming key tabular content from the new Guide 7.3 into structured, machine-readable data through automated scripts, and developing an interactive web application to support efficient search and content access.

It also aims to evaluate the performance of RAG systems when applied to this document under different configurations of the retrieval and generation components. Experiments will be conducted using a set of quality metrics to assess the reliability of the interactive search system and the RAG systems. The goal is to summarise the applicability of the two systems in this specific use case.

The remainder of this paper is structured as follows. Section 2 introduces the components of the interactive search system. Section 3 describes the construction of the RAG system. Section 4 covers the experimental setup, evaluation metrics and the experimental results in tabular form. Section 5 visualises and analyses the results, and discusses the limitations of the current experiments. Finally, Section 6 concludes the paper and outlines possible directions for future work.

Interactive Search System

The developed system employs a frontend and backend separated architecture and is primarily composed of four modules: frontend, backend, database, and search engine.

The frontend is built using React to construct the web user interface, responsible for receiv-

ing user queries, performing format validation, and displaying search results. The backend, based on the Flask framework, provides a RESTful endpoint and parses query parameters to construct the corresponding database queries. MongoDB is used to store all acceptable solutions from Guide 7.3 and exchanges data with the backend, while Typesense, as a search engine, supports full-text search functionalities and directly responds to requests from the frontend.

All modules are run as Docker containers and are managed collectively via Docker Compose. Figure 1 illustrates the data communication among the containers. The init-typesense container, which is responsible for initialising the Typesense data, is executed only once.

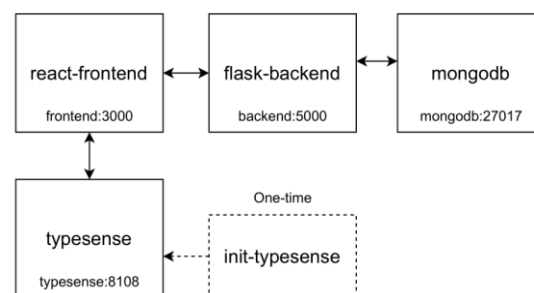


Fig. 1: Communication among docker containers

As shown in the table in Figure 2, each acceptable solution contains seven sections: "Name", "Description", "Variants", "Conditions", "Keywords", "Required Documentation", and "Functional Checks". There are four types of keywords: "Assets", "Security properties", "Attack vectors" and "Status". Each type can contain one or more specific values, with "all" representing the complete set of values for that category.

WELMEC Guide 7.3

Acceptable Solutions

Assets	Security properties	Attack vectors	Status
Full-text Search bootloader			

Found 3 results.

Solution: bootloader secured by trusted environment

Assets: software

Security properties: integrity, availability, authenticity

Attack vectors: ID1

Status: active

Solution: bootloader secured with a password

Assets: software

Security properties: integrity, availability, authenticity

Attack vectors: ID1

Status: active

Solution: secure boot

Assets: software

Security properties: integrity, availability, authenticity

Attack vectors: ID1

Status: active

Solution: bootloader secured with a password

Risk Class B	Risk Class C	Risk Class D
General - The boot configuration is secured by a password, see 4.4.2.1. - The integrity of the legally relevant parts of the operating system is checked by the bootloader by means of a reference value, see 4.1.1. - The integrity of the bootloader is physically protected by means of a hardware seal, see 4.4.3.1.		
Variants		
Conditions		
Keywords - Assets: software - Security properties: integrity, availability, authenticity - Attack vectors: ID1 - Status: active		
Required Documentation Description of the boot order and the interactions between boot configuration, bootloader and operating system.		
Functional Checks - Check that modification of the boot configuration is not possible without entering a password. - Check that booting of a modified boot configuration is not possible.		

Fig. 2: User interface of the interactive search system

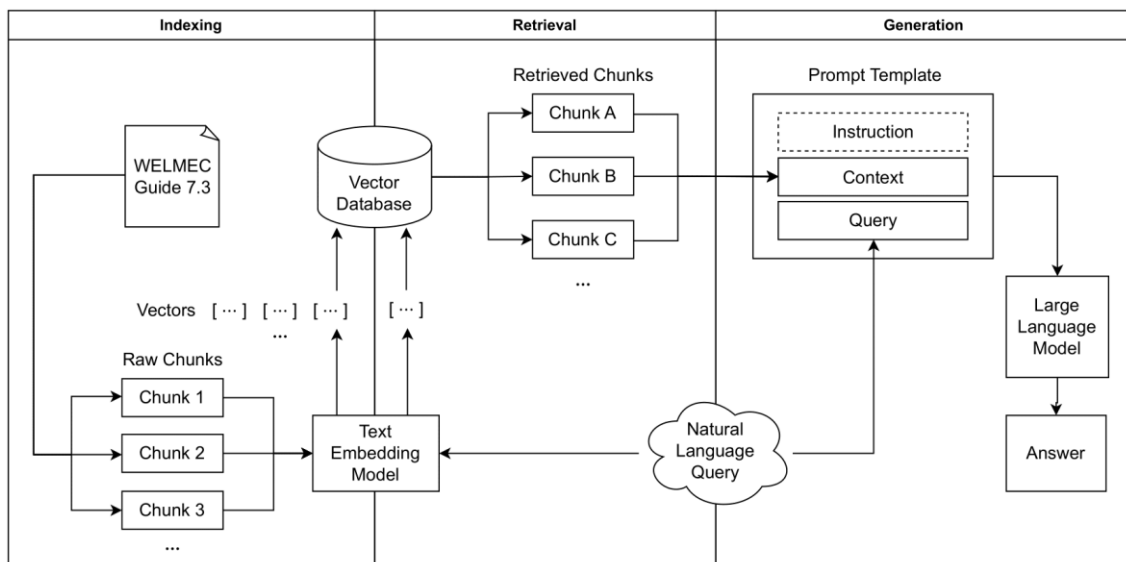


Fig. 3: Workflow of the RAG system

Four keyword input fields are provided at the top, corresponding to the four types of keywords in the “Keywords” section of an acceptable solution, along with a full-text search input field. The lower part is divided into two columns: the left side displays all retrieved results in the form of clickable cards, while the right side renders the title and table content of the selected acceptable solution in its original format, thereby ensuring that the presentation remains consistent with the official document.

Retrieval-Augmented Generation System

Based on the LlamaIndex framework, a basic RAG pipeline is implemented using exclusively a dense retrieval strategy. The overall workflow of the system is illustrated in Figure 3.

It begins with data ingestion, where the input DOCX document is parsed according to headings and tables, enabling semantic chunking of the Guide 7.3. All chunks are converted into dense vectors using a text embedding model and stored in the Chroma vector database. These stored vectors can be reused without being regenerated each time.

To optimise retrieval results, two parameters are configured: `similarity_top_k`, which limits the number of chunks returned from the matching results; and `similarity_cutoff`, which filters out results with low similarity scores.

For the LLM, the temperature parameter controls the randomness of the generated responses, with lower values leading to more stable outputs. Additionally, a prompt template can be provided to improve the quality of the generated content. The basic version, represented by `basic_prompt`, includes the retrieved content along with the user query.

Comparative experiments with various parameters and configurations will be conducted in the next section to evaluate the performance of different RAG systems.

Experiments and Results

In order to better analyse the performance of dense retrieval in identifying relevant texts and the further optimisation of results by LLMs, the retrieval and generation components of the RAG system are evaluated separately. OpenLLMetry, an open-source project for tracking and monitoring LLM applications, is employed here to observe the text chunks obtained during the retrieval stage.

The test cases are divided into two sets: Test Set 1 and Test Set 2. Test Set 1 contains 20 queries of varying complexity, using keyword values that actually exist in acceptable solutions. Test Set 2 contains 10 queries with non-existent values to test the system’s reliability when queries do not match any acceptable solutions. The following Experiments 1–4 are conducted on Test Set 1, while Experiment 5 is conducted on Test Set 2.

Experiment Set 1 (E1) provides a comprehensive performance comparison between the interactive search system and the RAG systems, using Test Set 1. In this experiment, OpenAI’s text embedding models, `text-embedding-3-small` and `text-embedding-3-large`, are used to generate dense vectors of different dimensions. To ensure retrieval completeness even in extreme cases, the `similarity_top_k` parameter is set higher than the total number of acceptable solutions (here set to 100), while `similarity_cutoff` is set to 0.2 to prevent relevant results from being prematurely excluded. The generation component uses

two general-purpose models, gpt-4omini-2024-07-18 and gpt-4.1-2025-04-14, as well as a reasoning model, o4-mini-2025-04-16, with the temperature parameter set to 0 to ensure stable responses.

To further enhance the quality of the generated outputs, a custom prompt template is designed. This template includes task analysis, definitions of the acceptable solution, and instructions on how to handle situations where no relevant information is found, with the aim of mitigating model hallucinations. It is noteworthy that the keywords fields may contain the value "all". To ensure the LLM accurately matches these fields when processing specific queries, the prompt explicitly specifies all corresponding values.

An excessively low similarity_cutoff value may place a high burden on the LLM and may include irrelevant results, while a value that is too high may unnecessarily exclude relevant results. To illustrate the effect of this parameter, different similarity_cutoff values (0.25, 0.30, and 0.35) are used in Experiment Set 2 (E2), with text-embedding-3-small employed for text embeddings.

If the temperature is not constrained, the LLM generates more diverse outputs, which are potentially less reliable or may even contain fabricated information. Experiment Set 3 (E3) quantifies this phenomenon by testing temperature values of 0.3, 0.6, 0.9 and 1.2, while the remaining configurations are set as follows: similarity_cutoff of 0.2, custom prompt, text-embedding-3-small and o4-mini-2025-04-16.

To demonstrate the impact of prompt quality on three different LLMs, Experiment Set 4 (E4) compares results using the custom prompt versus using only the retrieved text chunks and the user query as the prompt. Other configurations are: similarity_cutoff of 0.2, temperature of 0, and textembedding-3-small.

Finally, Experiment Set 5 (E5) examines whether the interactive search system and the RAG systems using three different LLMs return no results when queried for non-existent values. The settings are: similarity_cutoff of 0.2, temperature of 0, custom prompt, and text-embedding-3-small.

The evaluation is conducted using five traditional quality metrics, namely Precision (P), Recall (R), F1 Score (F1), Fallout (F), and Miss Rate (M). In addition, the Hallucination Rate (H) is calculated to evaluate the generation component of the RAG systems. Hallucination is calculated as the ratio of fabricated items generated by the LLM to the total number of generated items, representing the probability that the LLM produces incorrect information.

Table 1 presents the results of the RAG system using two different text embedding models and three different LLMs in E1. The results of the interactive search system are also listed in the table for comparison.

Tables 2 shows how the retrieval performance varies with similarity_cutoff values of 0.25, 0.30, and 0.35 in E2.

Table 3 presents the results of E3, illustrating how different temperature settings affect the generation performance of the RAG system.

Tab. 1: Results of different RAG systems and the interactive search system (E1)

Model		Retrieval					Generation					
		P (%)	R (%)	F1 (%)	F (%)	M (%)	P (%)	R (%)	F1 (%)	F (%)	M (%)	H (%)
embedding-small	4o-mini	43.33	100.00	55.91	100.00	0.00	54.54	21.03	24.33	8.73	78.97	0.00
	4.1						78.28	51.69	57.02	4.90	43.31	0.00
	o4-mini						100.00	94.43	96.02	0.00	5.57	0.00
embedding-large	4o-mini	43.33	100.00	55.91	100.00	0.00	68.06	25.30	33.90	3.56	74.70	0.00
	4.1						86.70	45.44	54.04	5.63	54.56	0.00
	o4-mini						98.83	95.82	96.12	2.69	4.18	0.00
Interactive Search System		95.00	95.00	95.00	0.00	5.00	-	-	-	-	-	-

Tab. 2: Results of different similarity_cutoff thresholds (E2)

similarity_cutoff	Retrieval				
	P (%)	R (%)	F1 (%)	F (%)	M (%)
0.25	43.46	93.71	55.28	96.00	6.29
0.30	45.25	58.39	43.59	61.88	41.61
0.35	54.82	22.87	25.44	12.44	77.13

Tab. 3: Results of different temperature values (E3)

temperature	Generation					
	P (%)	R (%)	F1 (%)	F (%)	M (%)	H (%)
0.3	94.81	85.21	86.57	3.19	14.79	0.00
0.6	97.62	94.17	93.90	6.50	5.83	0.00
0.9	100.00	88.04	90.27	0.00	11.96	0.00
1.2	100.00	95.19	95.84	0.00	4.81	0.00

Tab. 4: Results of different LLMs using basic prompt (E4)

LLM	Generation					
	P (%)	R (%)	F1 (%)	F (%)	M (%)	H (%)
4o-mini	48.09	22.30	26.03	13.32	77.70	0.00
4.1	74.50	30.76	39.79	0.22	69.24	0.00
o4-mini	70.00	24.79	33.40	0.00	75.21	0.00

Table 4 shows the generation performance of three different LLMs with only the retrieved text chunks and the user query as the prompt in E4. Table 5 records the Fallout and Hallucination Rate of RAG systems using three different LLMs when handling queries with non-existent values in E5, while for the interactive search system, only the Fallout is recorded.

Discussion

Figure 4 presents a comparison between the output quality of the different RAG systems and the performance of the interactive search system. Among them, RAG systems using the reasoning model o4-mini could better understand the task instructions in the custom prompt and performed well, yet they still failed to return all relevant acceptable solutions exclusively.

It is noteworthy that the interactive search system did not achieve a 100% F1 score and a 0% Miss Rate. This resulted from its reliance on boolean matching, which leads to no results being returned for queries containing spelling errors. In contrast, high-quality LLMs can comprehend instructions regarding typo handling and therefore demonstrate an advantage in responding to such queries.

With the similarity_cutoff set to 0.2, the size of the text embedding model did not affect the

Tab. 5: Results of queries with non-existent values (E5)

LLM	F (%)	H (%)
4o-mini	3.61	0.00
4.1	3.33	0.00
o4-mini	3.06	0.00
Interactive Search System	0.00	-

results, and the retrievers in each RAG system returned all acceptable solutions. The only difference was that the similarity scores obtained for the same query varied across vector databases, causing the acceptable solutions within the LLMs' context to be ranked differently. However, this mainly affected the outputs of non-reasoning models, while the reasoning model was hardly influenced, and the similarity score remains an uncontrollable variable with no research significance.

Table 1 shows that, under this configuration, the retrieval component of the RAG system obtained identical results across all five metrics. Due to an excessively low cutoff value, the retriever passed all acceptable solutions to the LLM, which achieved the highest recall but led to very poor precision, resulting in an F1 score of only around 50%. Figure 5 illustrates the changes after increasing the similarity_cutoff value: the accuracy of the results does not improve significantly, while recall drops sharply, causing more relevant acceptable solutions to be missed.

Through analysis of the test results, it was found that the keywords specified in user queries did not effectively constrain the retrieval stage. For example, for the query "List the names of all acceptable solutions where the 'Assets' field contains 'software'.", acceptable solutions whose assets field contains the

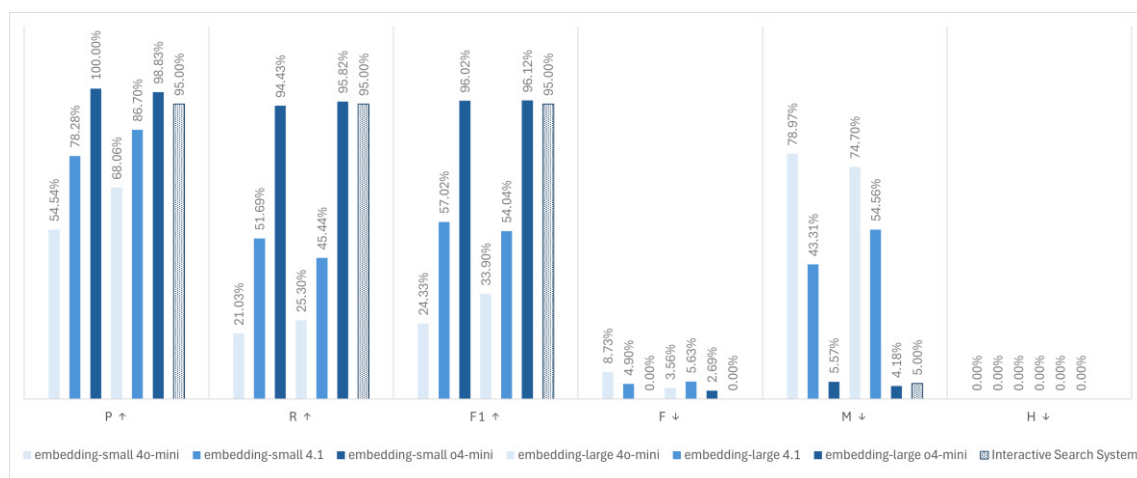


Fig. 4: A comprehensive performance comparison between RAG systems using different models and the interactive search system (E1)

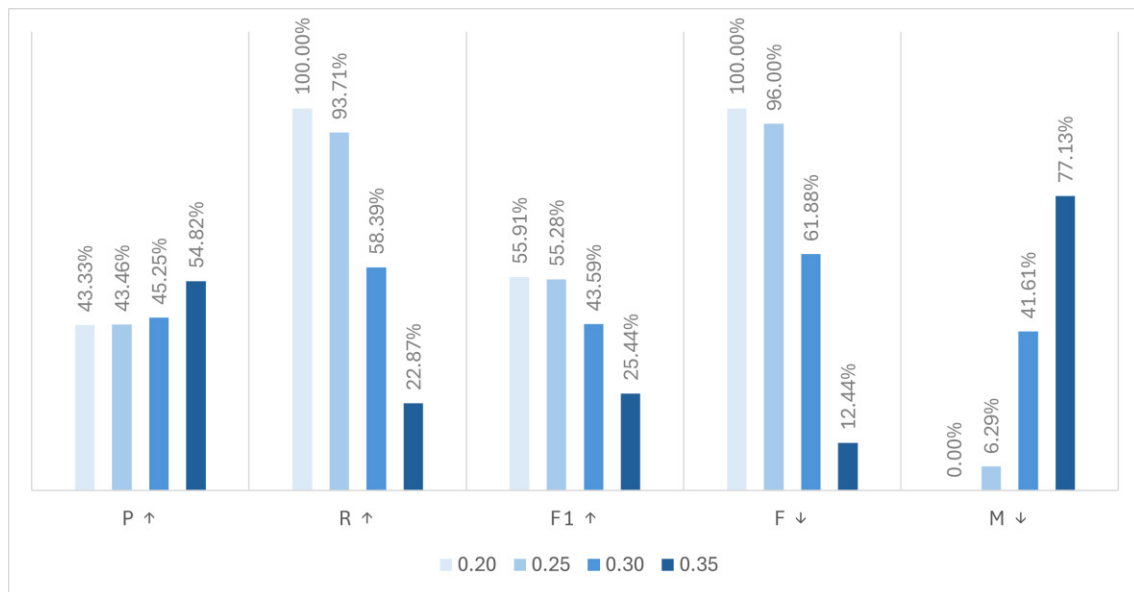


Fig. 5: Comparison of the retrieval performance of the RAG system with varying similarity_cutoff thresholds (E2)

term “software” do not necessarily obtain higher scores in similarity comparisons. The reason is that, apart from the “Keywords” section, semantic information in other sections negatively affects vector index construction and matching. This implies that increasing the similarity_cutoff value may lead to the loss of relevant documents during retrieval, thereby reducing overall usability.

Figure 6 shows the impact of the temperature parameter on the performance of the LLM. As the temperature value increases, the F1 score fluctuates without showing a clear trend, and this also applies to the four other metrics. Notably, the Hallucination Rate remains at 0%.

the LLMs successfully followed the instructions in the custom prompt and completed tasks without producing fabricated answers. In general, when using high-quality LLMs with a well-designed prompt, the setting of the temperature has a negligible effect on the performance. Figure 7 illustrates the impact of prompt quality on the performance of different LLMs. When no explicit instructions or special-case handling were provided, all three LLMs performed poorly. Unexpectedly, even without explicitly instructing the model not to fabricate information, no hallucinations occurred in the task of returning title texts.

When more detailed descriptions were added, the RAG system using low-quality model 4o-

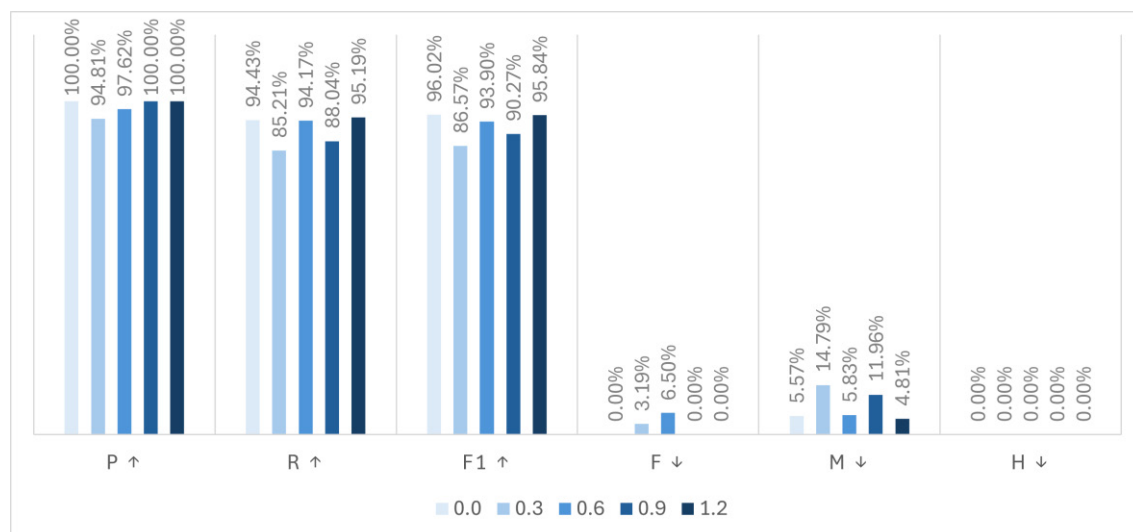


Fig. 6: Comparison of the generation performance of the RAG system with varying temperature values (E3)

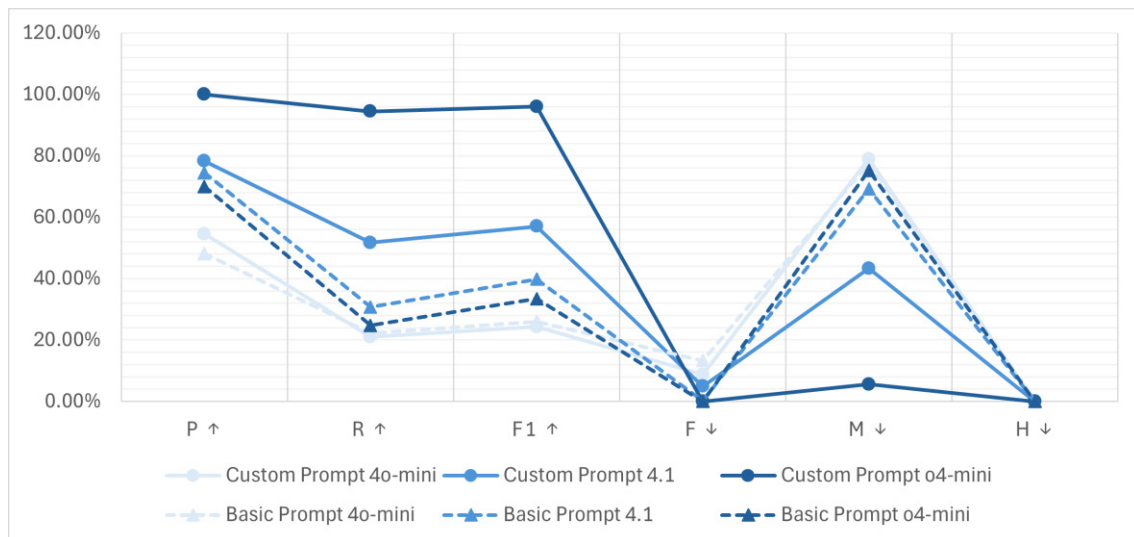


Fig. 7: Comparison of the generation performance of RAG systems using different LLMs under basic prompt (E4)

mini maintained almost the same level of performance. In contrast, the RAG system built on the reasoning model o4-mini nearly tripled its F1 score, with a significant drop in the Miss Rate. This improvement is mainly attributed to translating statements containing “all” into their corresponding values in the prompt, allowing a large number of relevant acceptable solutions to be retrieved. Despite some improvement observed with the 4.1 model, its overall performance remains insufficient for practical use. As shown in Figure 8, in the final set of experiments, when faced with keywords that do not exist in the documents, all RAG systems were able to maintain a Hallucination Rate of 0%, thereby ensuring a certain level of reliability. However, even the best among the three tested LLMs could not definitively determine that there are no matching documents in every test

case. Therefore, in such a scenario, RAG systems cannot guarantee complete correctness. The main limitation of this test is that there are only 36 acceptable solutions in Guide 7.3. Even if all of them are passed to the LLM, it does not exceed the context window limit. When the number increases, using a relatively low similarity_cutoff value may result in LLM failure or forced truncation of excess text. Future work should test documents containing more of these acceptable solutions to further evaluate the system’s usability. Moreover, each test case under each experimental configuration was executed a single time through the RAG pipeline and the results were recorded. Especially for RAG systems with higher temperature, the results may show certain variability. Future work could conduct multiple runs and average the results to obtain more reliable conclusions.

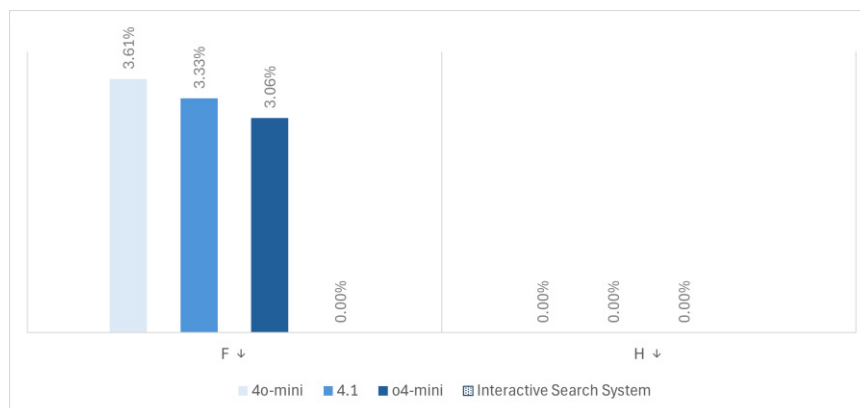


Fig. 8: Comparison of the performance of RAG systems and interactive search system on queries with non-existent values (E5)

Conclusion and Outlook

In this paper, an interactive search system was developed, which supports both structured and full-text searches of acceptable solutions contained in WELMEC Guide 7.3, with results displayed on the web interface in their original format. The system automates the data processing workflow through multiple scripts for data extraction, database storage, and search engine import, thereby minimising manual work and associated costs. In addition, it supports complex queries, fuzzy matching, result highlighting, and relevance ranking, effectively improving document retrieval efficiency and user query experience.

This paper also implemented a RAG approach that combines dense vector retrieval with LLMs. This method does not require a data model and allows generating answers directly from natural language queries on the document content. It reduces development cost and can be easily applied to a wide range of documents.

Experimental results show that the interactive search system performs excellently in all tests, except in cases involving spelling errors. It accurately locates target acceptable solutions and maintains high precision even in complex query scenarios, demonstrating its applicability for efficient content retrieval of standardised documents. In contrast, the RAG systems using general-purpose LLMs failed to meet practical requirements across various parameters and configurations; even with a reasoning LLM, a well-designed prompt is necessary to achieve basic usability, yet complete accuracy and stability could not be guaranteed, which presents difficulties for practical applications.

Some limitations still exist. Firstly, the retrieval component currently relies solely on dense retrieval, and the inclusion of irrelevant text in vector construction has reduced retrieval performance. Future work could consider constructing vectors only from the “Keywords” section of acceptable solutions, adopting sparse retrieval methods (e.g., BM25), or employing a hybrid retrieval strategy. Secondly, in the current prompt template, retrieved acceptable solutions are ranked in descending order of similarity score, meaning those with lower similarity scores are positioned closer to the user query. According to [11], placing the target document near the query improves accuracy, and the study also suggests that adding irrelevant documents may further enhance system performance. These findings provide valuable insights for future work.

References

- [1] “WELMEC 7.2 Software Guide,” European Cooperation in Legal Metrology, WELMEC Secretariat, Braunschweig, Standard, Sep. 2025
- [2] “Scenarios for Digitizing Standardization and Standards,” DIN e. V. and DKE Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE, Berlin and Frankfurt am Main, White Paper, June 2021
- [3] J. Herzig, T. Müller, S. Krichene, and J. Eisen-schlos, “Open domain question answering over tables via dense retrieval,” in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2021, pp. 512–519, doi: 10.18653/v1/2021.naacl-main.43
- [4] A. Asai, X. Yu, J. Kasai, and H. Hajishirzi, “One question answering model for many languages with cross-lingual dense passage retrieval,” *Advances in Neural Information Processing Systems*, vol. 34, pp. 7547–7560, 2021
- [5] W. X. Zhao, K. Zhou, J. Li, T. Tang, X. Wang, Y. Hou, Y. Min, B. Zhang, J. Zhang, Z. Dong, et al., “A survey of large language models,” *arXiv preprint arXiv:2303.18223*, 2023, doi: 10.48550/arXiv.2303.18223
- [6] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel et al., “Retrieval-augmented generation for knowledge-intensive nlp tasks,” *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020
- [7] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, H. Wang, and H. Wang, “Retrieval-augmented generation for large language models: A survey,” *arXiv preprint arXiv:2312.10997v5*, 2024, doi: 10.48550/arXiv.2312.10997
- [8] C. Melton, A. Sorokine, and S. Peterson, “Evaluating Retrieval Augmented Generative Models for Document Queries in Transportation Safety,” *arXiv preprint arXiv:2504.07022*, 2025, doi: 10.48550/arXiv.2504.07022
- [9] Y. H. Ke, L. Jin, K. Elangovan, H. R. Abdullah, N. Liu, A. T. H. Sia, C. R. Soh, J. Y. M. Tung, J. C. L. Ong, C.-F. Kuo et al., “Retrieval augmented generation for 10 large language models and its generalizability in assessing medical fitness,” *npj Digital Medicine*, vol. 8, no. 1, p. 187, 2025, doi: 10.1038/s41746-025-01519-z
- [10] H. Ye, T. Liu, A. Zhang, W. Hua, and W. Jia, “Cognitive mirage: A review of hallucinations in large language models,” *arXiv preprint arXiv:2309.06794*, 2023, doi: 10.48550/arXiv.2309.06794
- [11] F. Cuconasu, G. Trappolini, F. Siciliano, S. Filice, C. Campagnano, Y. Maarek, N. Tonello, and F. Silvestri, “The power of noise: Redefining retrieval for RAG systems,” in *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2024, pp. 719–729, doi: 10.1145/3626772.3657834

Optimierung der Lage der Roboterbasis durch Kombination verschiedener Messverfahren

Markus Gapp¹, Tobias Weiser¹

*¹Hochschule Kempten, Institut für angewandte KI und Robotik, Schwabenstraße 102, 87616 Marktoberdorf
markus.gapp@hs-kempten.de*

Zusammenfassung

Die Steigerung der Absolutgenauigkeit von Robotern stellt seit vielen Jahren einen signifikanten Forschungszweig dar. Ein wesentlicher Aspekt ist hierbei die Simulation von Roboterzellen sowie die darauffolgende Offline-Generierung von Roboterprogrammen. Die Übereinstimmung der Koordinaten des Programms mit der realen Zelle muss dabei möglichst akkurat gewährleistet sein. Obwohl sich anfangs nur wenige Roboterhersteller mit dieser Thematik befassten, bieten inzwischen die meisten Industrieroboterhersteller und auch diverse Dienstleister eine Verbesserung der Roboter Genauigkeit an. Vor der Kalibrierung sind jedoch zwei wesentliche Vorarbeiten zu leisten: Die Ermittlung des Tool-Center-Points (TCP) und die Transformation zwischen dem Messsystem und der Roboterbasis. Der Fokus dieses Beitrags liegt auf der Basis des Roboters.

Keywords: Robotik, Genauigkeit, Roboterbasis, optische Messverfahren

Einführung und Stand der Technik

Für einen Roboter mit hoher Absolutgenauigkeit [1] ist ein möglichst genauer TCP [2] mit Bezug auf den Roboterflansch und die Position der Roboterbasis von grundlegender Bedeutung. In der vorliegenden Abhandlung wird die Präzision der Basis des Roboters erörtert.

Die Ermittlung der Roboterbasis (RBF – Robot Base Frame) kann grundsätzlich mit den beiden Methoden "direkt" oder "indirekt" erfolgen.

Bei den indirekten Methoden erfolgt die Ermittlung der Transformation zwischen Messsystem und Roboterfuß/-basis mittels eines Optimierungsverfahrens, z. B. der Singulärwertzerlegung (SVD – Singular Value Decomposition [3]). Eine alternative Variante zur Bestimmung des Koordinatensystems besteht in der Messung der bewegten Achsen 1 und 2 über ein Kreis-Fitting [1, 4].

Die L-Methode sowie weitere indirekte Methoden wurden in [5] detailliert beschrieben.

Die direkten Methoden bedienen sich geometrisch bekannter Elemente auf der Aufstellplatte, die einen Bezug zur

Roboterbasis aufweisen. Abschließend ist es von essenzieller Bedeutung, dass der aufzustellende Roboter per Aufnahmestifte mit Passung auf die Platte gestellt wird. Andernfalls können erste Abweichungen auftreten. Nach Festlegung der geometrischen Elemente kann die Roboterbasis mithilfe der Vektorrechnung ermittelt werden.

Hinsichtlich der verwendeten Messsysteme in der Robotik wurde anfangs mit Theodoliten gearbeitet. Darauf aufbauend wurden kamerabasierte Systeme entwickelt [6, 7], die aufgrund der kontinuierlichen Steigerung von Auflösung und Genauigkeit auch heute [8] noch im Einsatz sind. Einen signifikanten Beitrag zur Präzision bei Robotikanwendungen leisten seit 1988 Lasertracker der Fa. API (Automated Precision Inc.), die entweder mit einer Laufzeitmessung arbeiten oder basierend auf der Interferometrie Genauigkeiten von unter 10 µm erzielen. Ein weiterer Vorteil des Lasertrackers ist die Erfassung eines Arbeitsraums von bis zu 80 m im Radius.

Messtechnischer Aufbau

In vorliegendem Beitrag findet eine CNC-bearbeitete Montageplatte als Roboter aufstellplatte Verwendung (vgl. Abb. 1). Sie eignet sich für den Einsatz mit verschiedenen Robotertypen und bietet zusätzliche

geometrische Elemente, die zur Ermittlung des Basiskoordinatensystems genutzt werden können.

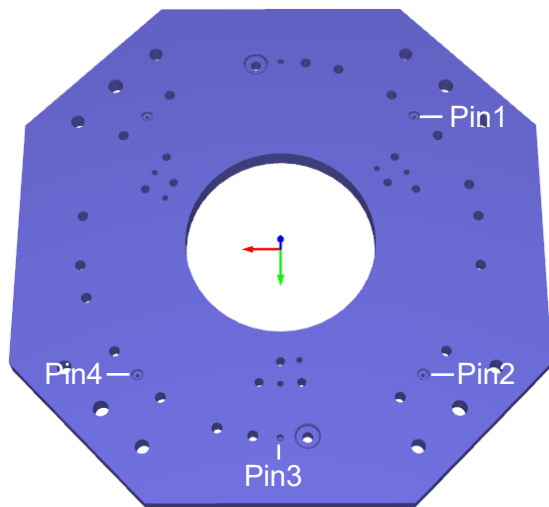


Abb. 1: CNC-bearbeitete Roboterplatte zur Montage unterschiedlicher Roboter-typen über Zentrierstifte. Die gemessenen Bohrungen zur Aufnahme der Zentrierstifte, sind gekennzeichnet.

Als Roboter wurde ein KR30HA_3 verwendet, welcher eine maximale Reichweite von 2033 mm aufweist und somit der Kategorie des sogenannten "Medium Range" zugeordnet werden kann. Die Option "positioniergenau" war deaktiviert. Mit dieser Option bietet KUKA eine bessere, absolute Positionierung an, die in Anlehnung an [7] modellbasiert zum Einsatz kommt.

Das Messsystem ist ein Lasertracker der Firma API vom Typ RADIANT PLUS, der mit Laufzeitmessung arbeitet, jedoch auch 6D-Messungen durchführen könnte. Gemäß den Herstellerangaben beträgt die Grundgenauigkeit $15\text{ }\mu\text{m} + 5\text{ }\mu\text{m/m}$. Die vorliegenden Versuche im temperaturstabilen Labor weisen eine Messgenauigkeit von ca. $10\text{ }\mu\text{m}$ im verwendeten Messbereich bis ca. 3 m auf.

Für die Messungen wurde ein Standard 1,5" Reflektor (SMR-Spherically Mounted Retroreflector) verwendet. Alternativ fand das von API verfügbare "Active Target" Anwendung, welches mit zwei Motoren ausgestattet ist und eine Nachführung des Laserstrahls ermöglicht. Dadurch ist die Erfassung großer Bewegungsbereiche möglich. Dieser Vorteil wird insbesondere bei der Messung der Achse-1-Bewegung genutzt.

Methode und Datenerfassung

Die Ermittlung des Basiskoordinatensystems eines Roboters erfolgt in der Regel auf Basis geometrischer Elemente. Im darauffolgenden Schritt wird das Roboter-Modell mittels sogenannter Zentrierstifte auf der Montageplatte positioniert. Dies kann in Abweichungen bei der Transformation zwischen dem Messsystem und der realen Basis des Roboters resultieren und hängt von der Position des Roboters ab. Beim verwendeten Typ wurde die max. Abweichung im mittleren Arbeitsbereich entsprechend Gl. (1) zu

$$AP \leq 0,06\text{ mm}$$

berechnet. Dieser Wert wurde unter Verwendung der AP-Accuracy of Pose (ISO 9283 [9]) ermittelt, der Strahlensatz kam zur Anwendung. Der mittlere Arbeitsbereich ist zudem definiert als der in der ISO 9283 festgelegte ISO-Würfel. Für die Berechnung wurde die Position P1, der Würfelmittelpunkt, genutzt, siehe Abb. 2.

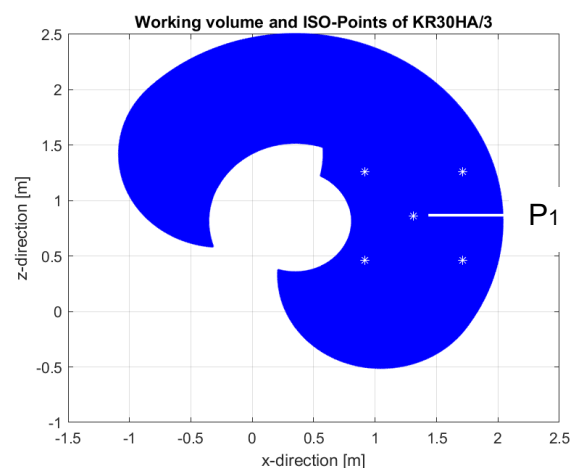


Abb. 2: Seitenansicht des Roboterarbeitsbereiches, der Roboterfuß steht im Ursprung, P_1 entspricht der Mitte des Würfels.

Abb. 3 zeigt die durchgeführten Messungen. Die Aufzeichnung der Achsbewegungen ist besonders gut sichtbar. Die Koordinatensysteme der Roboterbasis sind unterhalb der Achse-1-Aufzeichnung lokalisiert. Darüber hinaus wird das Koordinatensystem dargestellt, welches aus den Achsbewegungen und der Ebene (vgl. Abb. 4) konstruiert wurde. Aufgrund der präferierten Übersichtlichkeit werden die Messungen der Ebene jedoch nicht dargestellt. Die Pin-Aufnahmen auf der Aufstellplatte sind lediglich als ein Punkt zu sehen, dies führt bei

dieser Auflösung zu einer kaum wahrnehmbaren Erkennbarkeit. Auf der rechten Seite des Bildes befindet sich der Lasertracker des Herstellers API.

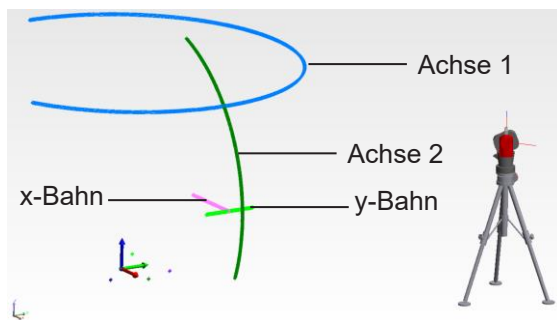


Abb. 3: Durchgeführte Messungen an der Aufstellplatte und die aufgezeichneten Bewegungen des Roboters (Achse 1 und 2, x- und y-Bahn [5]) ohne die Messungen auf der Ebene der Platte (siehe Abb. 4).

Im Rahmen der indirekten Methoden wurden folgende Varianten untersucht:

- Ermittlung über die Bewegungen der Achsen 1 und 2 des Roboters (1. Methode, M1), Prozedur aus RoboDK, Modul Roboterkalibrierung, [10]).
- Ermittlung mit Hilfe von Soll- und Ist-Positionen mittels der Singulärwertzerlegung (2. Methode, M2).

Für die direkten Methoden wurden die Zentrierbohrungen auf der Aufstellplatte Abb. 1 verwendet, die einen bekannten Bezug zum Ursprung des Roboterkoordinatensystems und somit zur Mitte der Aufstellplatte haben. Die Transformation T für dieses Koordinatensystem wurde unter Anwendung der Vektorrechnung und des Kreuzproduktes entsprechend Gl. (2) ermittelt (3. Methode, M3).

$$T_{API-Rob} = \begin{pmatrix} x \\ \frac{u}{|u|} \frac{v}{|v|} \frac{w}{|w|} y \\ z \\ 0 \ 0 \ 0 \ 1 \end{pmatrix}_{4 \times 4}$$

Zur Bestimmung der Ebene der Aufstellplatte wurde die Oberfläche der Platte mit einem SMR gescannt. Daraus wurde die z-Koordinate für die direkten Methoden berechnet. Abbildung 4 zeigt die Messpositionen auf der Aufstellplatte.

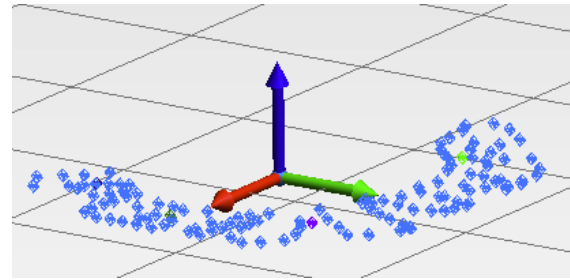


Abb. 4: Messpositionen auf der Ebene.

Im letzten Schritt wurde die zuletzt ermittelte Ebene mit den Kreisbewegungen (1. Methode) der indirekten Methode kombiniert (4. Methode, M4). Die These ist, dass der Ursprung des Koordinatensystems idealerweise durch den Mittelpunkt der Achse-1-Bewegung abgebildet wird. Ebenso wird die y-Richtung des Koordinatensystems durch die Achse-2-Bewegung am besten repräsentiert.

Zur Messdatenerfassung wurde die Achse A1 maximal bewegt und dabei der Roboterflansch mit dem "Active Target" erfasst. Dadurch konnten 240° des Kreises gemessen werden. Ebenso wurde die Achse 2 über einen Bereich von 120° verfahren und gemessen. Aus beiden Datensätzen wurden mit Hilfe eines Kreisfittings [11] der Mittelpunkt und der Radius berechnet.

Ergebnisse

Die maximale absolute Abweichung aufgrund der mechanischen Komponenten (Pins auf der Roboteraufstellplatte) ergibt sich bei einem mittleren Arbeitspunkt P_1 von $x = 1314$ mm und einem $TCP_z = 150$ mm zu 0,06 mm.

Der Arbeitspunkt P_1 wurde auch für die weiteren Vergleiche herangezogen.

Die max. mögliche Abweichung durch die Aufstellung addiert sich zu der absoluten Positioniergenauigkeit des Roboters am Arbeitspunkt.

Aus den drei Pin-Positionen kann der Ursprung und die Ausrichtung des Roboterkoordinatensystems berechnet werden. Abb. 5 zeigt die Lage dieses Koordinatensystems mit den verwendeten Konstruktionselementen.

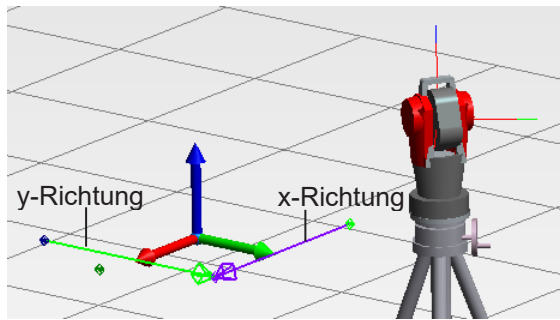


Abb. 5: Lage des Roboterkoordinatensystems auf der Basis von Methode 3. Die Pin-Positionen mit den entsprechenden Vektoren sind auch eingezeichnet. Die z-Koordinate wird bei beiden Methoden 3 und 4 über die gemessene Ebene festgelegt.

Im Folgenden werden die Ergebnisse aus den Fitting-Funktionen [11] der Ebene und den Kreisen dargestellt. Die Ebene wurde mit einem SMR abgetastet, der in einem Nest lag, um Kratzer auf der Oberfläche des SMRs zu vermeiden. Durch Berücksichtigung der Nesthöhe kann z bestimmt werden. Der „Root-Mean-Square“ (RMS) Fehler ergab sich zu: 0,012 mm.

Ebenso wurde der RMS des Kreis-Geometrie-Fits mit 0,012 mm für die Achse A1 und 0,011 mm für die A2 ermittelt.

Die Ergebnisse der vier Methoden resultieren aus den Messungen, Berechnungen und dem Fitting der Kreise bzw. Ebene, wie in Tab. 1 dargestellt. Bezüglich der Methode 4 sei darauf hingewiesen, dass es sich um die Bezugsmethode handelt.

Tab. 1: Ursprung der unterschiedlichen Methoden, Bezug ist die Methode M4, bei der die Achse 1 (Drehachse) dem Ursprung entspricht. Die Darstellung (*) erfolgt im Rahmen des RMS-Fehlers der Kreise und der Ebene.

mm	Δx	Δy	Δz
M1	0,233	0,56	0,415
M2	1,048	-0,669	0,281
M3	0,236	0,524	0*
M4	0*	0*	0*

Im Gegensatz zu [5], bei der beliebig verteilte Positionen verwendet wurden, bezieht sich dieser Beitrag für die Methoden 3 und 4 auf den ISO-Würfel. Dort wurden die absoluten Abweichungen entsprechend [9] an den 8

Würfecken, zuzüglich P_1 , entsprechend Gl. (3) berechnet.

$$AP = \sqrt{(x_c - \bar{x})^2 + (y_c - \bar{y})^2 + (z_c - \bar{z})^2}$$

Tabelle 2 zeigt die Ergebnisse der absoluten Positionierung des Roboters in Abhängigkeit von den beiden Verfahren direkt (M3) und kombiniert (M4).

Tab. 2: Absolute Positionierung des Roboters mit Bezug auf den ISO-Würfel. Inklusive P_1 sind dies neuen Positionen.

mm	Mittelwert	Max-Wert
M3	1,53	2,03
M4	1,35	1,82

Diskussion

Das kommerzielle Softwarepaket RoboDK ermittelt die Roboterbasis (Methode 1) über zwei Kreisbewegungen von 60° bzw. 55°. In Bezug auf die z-Koordinate findet das Nominalmodell des Roboters Anwendung. Es stellen sich robuste und wiederholbare Ergebnisse ein, die in der Regel zu einem Offset führen.

Bei der Methode 2, wird mit einem Soll-Ist-Vergleich eine möglichst gute Übereinstimmung erreicht. In diesem Beitrag wurde die Singulärwertzerlegung eingesetzt. Je genauer der Roboter selbst positioniert, umso präziser kann auch die Transformation entsprechend Gl. 2 ermittelt werden.

Die Methode 3 basiert auf geometrischen Elementen, die sich auf einer bearbeiteten Fläche befinden. Diese entsprechen den Führungsbolzen der Roboteraufstellung. Sowohl die Genauigkeit der Aufstellplatte als auch die Passung der Zentrierstifte führen zu einem Offset bezüglich Ursprung.

Die letzte untersuchte Methode kombiniert die Messung der Ebene der Aufstellplatte mit den Kreisbewegungen des Roboters. Da die aufgezeichneten Kreissegmente große Bereiche abdecken, ergibt sich bereits die umfangreichste Datenbasis. Physikalisch liegt der Ursprung des Roboterkoordinatensystems im Ursprung der Achse-1-Bewegung. Dieser Vorteil und auch der große Teilkreis werden bei der vorgeschlagenen Methode verwendet.

In Übereinstimmung mit den Ergebnissen in Tabelle 2 resultiert die Methode 4 in der

geringsten mittleren Abweichung von 1,35 mm, unter der Voraussetzung, dass ein Standardroboter mit Nullpunkt-Kalibrierung der bewegten Achsen zum Einsatz kommt. Die Methoden 1 und 2 weisen bereits im Ursprung eine z-Abweichung auf, da die Ebene nicht gemessen, sondern über die Kinematik bzw. durch eine Optimierung bestimmt wird. Für die Methode 2 wurde zudem die größte Abweichung in einer Richtung mit $\Delta x = 1,05$ mm ermittelt.

Fazit und Ausblick

Die Methode 4, die Kreismessungen mit der direkten Messung der Ebene kombiniert, führt bei der Prüfung der absoluten Referenzpositionen (9 Positionen des ISO-Würfels) zu den geringsten Abweichungen. Da durch die Rotation der Achse 1 der Ursprung ermittelt wird, kann hier von keiner weiteren Verbesserungsmöglichkeit ausgegangen werden.

Potential bietet eventuell die Ausrichtung der y – Achse. Erste Untersuchungen der L-Methode und deren y-Richtung führten zu keiner Verbesserung.

Von weiterem Interesse ist der Einfluss der gewählten Basisermittlung auf die Kalibrierung eines Roboters. Hierzu sind umfangreiche Untersuchungen notwendig, insbesondere da es auch diverse Verfahren zur Kalibrierung eines Roboters gibt.

Für die Robotikanwendungen stellt sich die Frage, ob durch die optimierte Basis auch die Offline-Programme effizienter in Betrieb genommen werden können.

Abschließend wurden die Daten dieser Arbeit mit einer Roboterkinematik ermittelt. Die Übertragbarkeit auf weitere Roboter oder auch andere Kinematiken wurde hier nicht untersucht.

Literaturnachweis

- [1] A. Nubiola, I. A. Bonev, Absolute calibration of an ABB IRB 1600 robot using a lasertracker, Robotics and Computer-Integrated Manufacturing 29 (2013) 236–245, DOI: 10.1016/j.rcim.2012.06.004
- [2] Roman Haas, Thomas M. Wendt, A Novel Vision-Based Method for Accurate Tool Center Point Calibration of Robots, IEEE Sensors (2025); DOI: 10.1109/SENSOR59705.2025.11330685
- [3] W. Dahmen, A. Reusken, Numerik für Ingenieure und Naturwissenschaftler, Springer Verlag (2022), ISBN 978-3-662-65180-3, DOI: 10.1007/978-3-662-65181-0
- [4] Sergio Ferrarini, Pietro Bilancia, Roberto Raffaeli, Margherita Peruzzini, Marcello Pellicciari, A method for the assessment and compensation of positioning errors in industrial robots, Published by Elsevier Ltd. (2023), DOI: 10.1016/j.rcim.2023.102622
- [5] Maxime Sélingue, Adel Olabi, Stephane Thiery, and Richard Béarée, Experimental analysis of robot base frame identification Methods, Sixth Iberian Robotics Conference, (2023), DOI: 10.1007/978-3-031-58676-7_41
- [6] Fa. Krypton/Metris/Nikon, www.metris3d.hu/k400_500_600_cmm_de_0905.pdf, (abgerufen am 30.03.2026)
- [7] U. Wiest, Kinematische Kalibrierung von Industrierobotern, Réseau Scanning Kamera (S. 41), Dissertation, Universität Karlsruhe (2001), ISBN 978-3-8265-8609-5
- [8] Fa. Creaform, C-Track Stereokamera, <https://www.creaform3d.com>, (abgerufen am 30.03.2026)
- [9] ISO 9283:1998 Manipulating industrial robots – performance criteria and related test methods, 1998
- [10] RoboDK Hilfe, <https://www.robodk.com/doc/en/Robot-Calibration-LaserTracker.html> (abgerufen am 17.04.2026)
- [11] Spatial Analyzer, Users Manual, <https://www.spatialanalyzer.com/download/> (abgerufen am 08.05.2026)

Measurement Concept for the Deflection Behaviour of Shape Memory Alloy Actuators in Robotics

Alexander Dierle¹, Thomas M. Wendt¹ and Thomas Seifert¹

¹ Offenburg University of Applied Sciences, Badstraße 24, 77652 Offenburg, Germany
alexander.dierle@hs-offenburg.de

Summary: In this paper, a combined measurement concept for characterizing a structurally integrated shape memory alloy actuator is described. The temperature-induced bending movement is measured without contact using a 3D laser line sensor, while the activation and transformation state is determined in parallel via the electrical resistance of the embedded shape memory wire. Supplementary simulations are utilised to evaluate the anticipated actuator dynamics and the viability of the measurement concept.

Keywords: Shape memory alloys, laser line sensor, actuator, additive manufacturing, robotics

Introduction and motivation

Shape memory alloys (SMAs) are considered to be among the so-called 'smart materials' and are increasingly being used in robotics and automation. This is due to the reversible martensite-austenite transformation that these alloys exhibit, which enables compact actuator designs with high energy density [1], [2], [3]. When compared with conventional drives, SMA actuators enable lightweight, low-noise and miniaturized systems that are particularly suitable for integrated gripping mechanisms, medical applications and micro-robotic structures [1]. The technological and economic relevance of these actuators is reflected in a projected annual market growth of 13.4% until 2031 [4].

The temperature-dependent bending behaviour of the structure is of central importance for the development of reliable actuator systems, as it directly reflects the transformation state of the alloy [1]. Despite the growing prevalence of SMA actuators, there are no standardised procedures that enable reproducible measurement of both deflection and phase state (cf. [5], [6], [7]). Current measurement methods, such as marker-based optical systems or point displacement sensors, often only provide local information, are susceptible to interference, or do not allow time-resolved reconstruction of the entire bending shape (cf. [5], [6], [7]).

This study provides a proof-of-concept through numerical simulation, analyzing the expected performance and sensitivity of the integrated SMA actuator measurement system before physical implementation. The combined measurement strategy presented addresses existing shortcomings in terms of resolution, reproducibility and temporal dynamics and creates the basis for a standardisable characterisation of SMA actuators. The measurement setup, data evaluation and supplementary simulations describing the expected bending behaviour and material-

specific characteristic curves for the integrated actuator are described later in the paper.

Method

Based on the previous work of the authors [8], the actuator is modelled as an additively manufactured structure made of an elastic, temperature-resistant material (TPU 85A). A 40 mm long shape memory wire with a diameter of 0.5 mm is embedded in this structure during the manufacturing process. The wire is positioned 0.75 mm above the neutral axis. The temperature-induced contraction of the shape memory wire by 3.25% results in a defined bending of the structure due to the eccentric wire position. The structure has a width of 5.0 mm, a height of 4.5 mm and a length of 40 mm. The chosen geometry represents a compromise: thinner wires (<0.5 mm) enable shorter activation times but are mechanically unstable, while thicker wires (>0.5 mm) react more slowly thermally and are more difficult to activate. Lengths greater than 40 mm increase the absolute deflection, but result in lower bending stiffness and increased deflection due to their own weight. The actuator described is shown in Figure 1.

The following describes the proposed experimental measurement concept. The thermal activation of the SMA wire or actuator would be achieved by resistance-based heating with a constant heating current of 4 A. The electrical resistance of the wire would be determined by measuring the voltage drop at constant current using a four-wire measurement and would be used as a material-specific indicator of the temperature and transformation state. The choice of a heating current of 4 A is based on common recommendations and manufacturer specifications for NiTi shape memory wires (cf. [9], [10]). In this way, the temperature- and transformation-dependent state of the material can be characterised in a material-specific manner and without

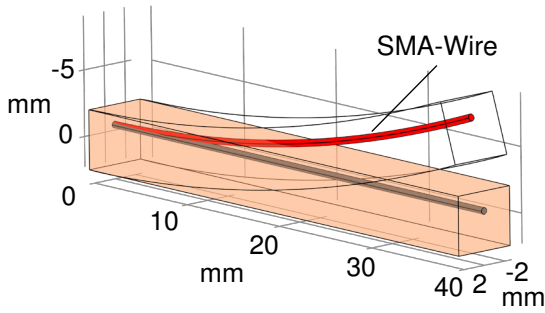


Fig. 1: Representation of the SMA actuator (red SMA-wire: activated state; grey SMA-wire: idle state)

additional sensor embedding [11], [12].

Electrical resistance is not understood as a direct temperature measurement variable, but rather as a material-specific state variable that maps the activation and transformation state via a previously determined resistance-temperature or resistance-phase characteristic curve.

A SmartRay ECCO 75.100 3D laser line sensor is proposed for the experimental acquisition of the resulting deflection profile. Its lateral resolution of 42–70 μm and vertical resolution of 5–12 μm enable a complete, time-resolved reconstruction of the bending shape with sampling rates of up to 4 kHz [13]. By linking the modelled deflection curve with the measured change in resistance, a curvature-temperature characteristic curve specific to the respective shape memory alloy can be derived. The schematic measurement setup is shown in Figure 2.

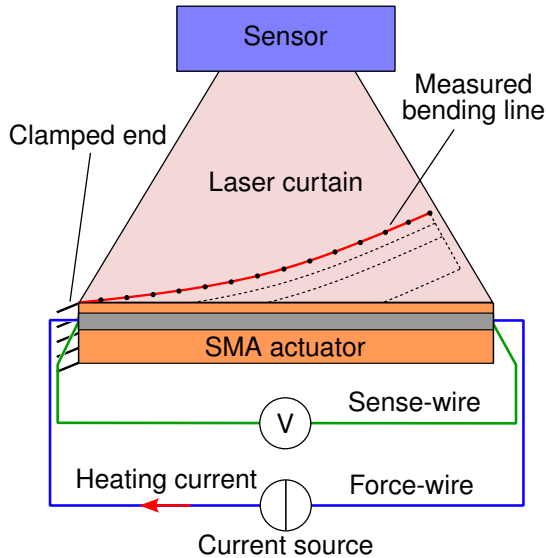


Fig. 2: Sketch of the described measurement setup

A finite-element simulation model of the considered actuator is employed to analyse deflection, curvature, and electrical resistance, as well

as to evaluate the proposed measurement concept. The model is constructed in COMSOL Multiphysics and is based on the implemented Lagoudas shape memory alloy model. Since COMSOL does not allow direct implementation of the two-way effect, this is approximated using a user-defined relationship that describes the transformation-induced strain ϵ^{tr} as a function of the martensite fraction ξ . The martensite fraction ξ evolves based on the temperature field and the total stress tensor [15]. $\epsilon^{tr,max}$ denotes the maximum transformation strain.

$$\epsilon^{tr} = (1 - \xi) \epsilon^{tr,max} \mathbf{e}_z \otimes \mathbf{e}_z \quad (1)$$

The martensite fraction ξ is dimensionless and takes values in the range $0 \leq \xi \leq 1$. When $\xi = 1$, the microstructure is completely martensitic, while when $\xi = 0$, it is completely austenitic.

The model described in [14] is used to estimate the electrical resistance. The stress-dependent term is neglected in the present design model. This is justified by the high compliance of the surrounding TPU structure compared to the shape memory wire

Young's modulus of structure $\ll$ Young's modulus of wire,

so that the thermally induced length change of the wire is predominantly absorbed by structural deformation (bending). The mechanical stresses arising in the wire are assumed to be small compared to the temperature- and phase-dependent contributions to resistivity. Thus, the following approximation applies:

$$\sigma \frac{\partial \rho_M}{\partial \sigma} = \sigma \frac{\partial \rho_A}{\partial \sigma} \approx 0. \quad (2)$$

The specific electrical resistance of the wire $\rho(T)$ is determined using a phase-weighted mixing rule,

$$\rho(T) = \rho_M(T) \xi + \rho_A(T) (1 - \xi) \quad (3)$$

where ρ_A and ρ_M are the specific electrical resistance of the austenite and the martensite phase, respectively.

The temperature-dependent specific resistances of the martensite and austenite phases are assumed to be linear in a first approximation and described as follows:

$$\rho_M(T) = \rho_{0M} + (T - T_{0M}) \frac{\partial \rho_M}{\partial T}, \quad (4)$$

$$\rho_A(T) = \rho_{0A} + (T - T_{0A}) \frac{\partial \rho_A}{\partial T}. \quad (5)$$

ρ_{0M} and ρ_{0A} represent the nominal specific resistances of the martensitic and austenitic phases at the corresponding measurement temperatures T_{0M} and T_{0A} .

The electrical resistance of the shape memory wire is described by its specific electrical resistivity. Literature values according to [14], which

Tab. 1: Material parameter Nitinol/ TPU 85A [8], [14]

Material parameter	Nitinol	TPU 85A
Reference temperature	20 °C	20 °C
Poisson's ratio (ν)	0.33	0.45
Density (ρ)	6450 kg/m ³	1114 kg/m ³
Young's modulus (E)	-	20 MPa
Young's modulus (martensite) (E_M)	28 GPa	-
Young's modulus (austenite) (E_A)	78 GPa	-
Maximum transformation strain ($\epsilon^{tr,max}$)	-3.25%	-
Martensite start/ finish temp. (M_s/ M_f)	66.42 °C/ 45.67 °C	-
Austenite start/ finish temp. (A_s/ A_f)	56.99 °C/ 74.48 °C	-
Thermal conductivity (k)	18 W/(m · K)	-
Linear temp. dep. spec. resistance (martensite) ($\frac{\partial \rho_M}{\partial T}$)	$1.099 \cdot 10^{-9}$ (Ω m)/K	-
Linear temp. dep. spec. resistance (austenite) ($\frac{\partial \rho_A}{\partial T}$)	$1.436 \cdot 10^{-10}$ (Ω m)/K	-
Nominal spec. resistance (martensite) (ρ_{0M})	$9.10 \cdot 10^{-7}$ Ω m	-
Nominal spec. resistance (austenite) (ρ_{0A})	$8.63 \cdot 10^{-7}$ Ω m	-
Meas. temp. spec. resistance (martensite) (T_{0M})	0 °C	-
Meas. temp. spec. resistance (austenite) (T_{0A})	225 °C	-

are originally specified for certain wire geometries, are converted into specific resistances for this purpose. This allows the electrical behaviour to be represented in the model independently of the wire geometry. However, it should be noted that these parameters are only valid for the alloy under consideration and must be experimentally redetermined for each shape memory alloy that is actually investigated. The material and model parameters used for the simulation are summarised in Table 1.

Results

To evaluate the feasibility of the proposed measurement concept, a simulation-based analysis was conducted. The results presented in this section are derived from the simulation model utilizing the model parameters defined in the previous section (see Table 1) to establish a baseline for the expected sensor characteristics and to quantify the coupling between deflection and transformation state. The heating of the SMA wire embedded in the structure occurs with a time delay compared to the manufacturer's specifications, as the surrounding structure must also be heated in addition to the wire. The temporal progression of the calculated average wire temperature is shown in Figure 3. After 0.68 s, the wire reaches the austenite start temperature $A_s = 57.0$ °C, thereby initiating the phase transformation from martensite to austenite. Complete austenitisation is achieved after 1.68 s at an average wire temperature of 74.6 °C, slightly above the stress-free value $A_f = 74.48$ °C from Table 1 due to the tensile stress in the con-

strained wire and the averaging over a non-uniform temperature field. The temperature rise across the transformation range thus extends over 1.0 s, which is reflected in the slope of the temperature curve.

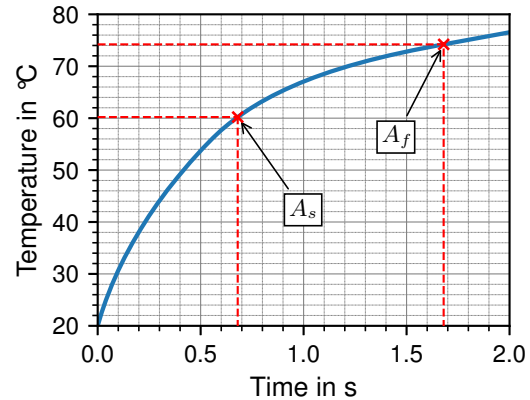


Fig. 3: The time-dependent change in the average temperature of the SMA wire embedded in the structure

The temperature-induced phase transformation results in an increasing bending of the actuator. The calculated maximum deflection at the end of the actuator is 7.05 mm and is reached at an average wire temperature of 74.6 °C, at which point the phase transformation is complete. The start of the bending coincides with the austenite

start temperature of 57.0 °C. The deflection profile of the activated SMA actuator, relative to the top of the actuator, is shown in Figure 4.

The corresponding calculated resistance behaviour of the modelled SMA alloy is shown in Figure 5. Before the phase transformation begins, the resistance increases with temperature due to the temperature dependence of the electrical resistivity of the martensitic phase. At the start of the phase transformation at the austenite start temperature $A_s = 57.0$ °C, a maximum resistance of 198.1 mΩ is reached. The subsequent characteristic drop in resistance marks the beginning of the phase transformation. As austenitisation progresses, the resistance continues to decrease until a minimum of 171.3 mΩ is reached at complete transformation (A_f). As the temperature continues to rise, the resistance then increases linearly again (to 172.2 mΩ at 100 °C), which is due to the temperature-dependent electrical conductivity of the austenitic phase.

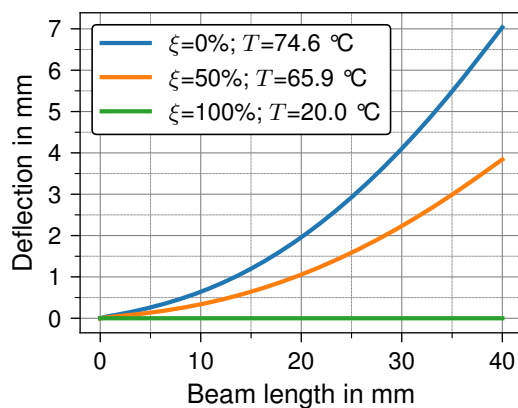


Fig. 4: Actuator deflection at different wire temperatures

Discussion

The results show that the temperature-induced phase transition of the embedded shape memory wire is consistently converted into a mechanical bending movement of the structure. The temporal course of the transition and the resulting deflection are within realistic ranges and confirm the basic functionality of the selected actuator concept.

On this thermo-mechanical basis, the simulation allows the electrical resistance measurement to be evaluated as an indirect state variable. The characteristic curve of the resistance signal in the transformation range shows the start of activation and the completed transformation state. The decisive factor here is not the absolute resistance value, but the reproducible signal shape in relation to the transformation state,

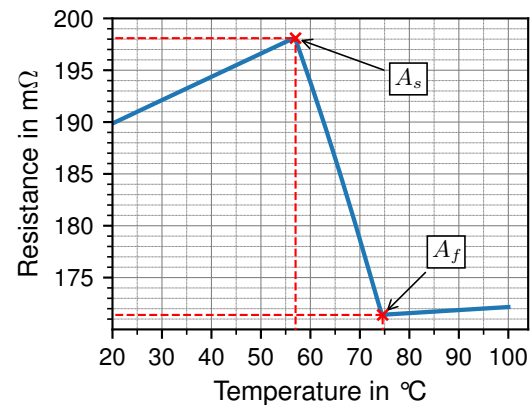


Fig. 5: Phase- and temperature-dependent resistance of the SMA wire

whereby the resistance can be interpreted as a functional indicator of the global actuator state.

The time scales of the simulated activation range from sub-hertz to low hertz, corresponding to the heating and cooling processes typical for shape memory wires. Against this background, the selected sensor technology is suitable for consistently recording both the mechanical deflection and the electrical signal over time. The simulation is not intended to predict the actual system behaviour exactly, but rather to estimate relevant measuring ranges, sampling rates and signal magnitudes for the designed test setup.

At the same time, it must be taken into account that the measured electrical resistance represents an effective average value over the length of the wire and cannot resolve local temperature or conversion states. This limitation is acceptable for the pursued goal of a conceptual, sensorless measurement approach, since the resistance is used as a calibrated state variable to describe the global activation and conversion state.

Conclusion and further work

The simulations carried out show that the temperature-induced phase transition of a shape memory wire integrated into a structure leads to reproducible mechanical actuation and that the activation and transition state is also reflected in the electrical resistance. The numerical results confirm that the selected sensor technology and the designed measurement setup are suitable for reliably recording the relevant mechanical and electrical variables in the dynamic range expected in the experiment.

The next step is to implement and validate the proposed measurement setup experimentally. In particular, the relationship between electrical resistance, temperature and mechanical deflection will be calibrated experimentally, and the influence of temperature gradients and hysteresis effects will be investigated.

References

- [1] S. Langbein and A. Czechowicz, *Formgedächtnistechnik: Entwickeln, Testen und Anwenden*. Wiesbaden: Springer Fachmedien, (2021)
- [2] M. N. Islam, M. B. Gonzalez, S. Tabassum, and K. M. M. Billah, "3d printed shape memory alloy wire embedded actuator," in *2024 IEEE 17th Dallas Circuits and Systems Conference (DCAS)*, ISSN: 2766-5186, (2024), pp. 1–5; doi: 10.1109/DCAS61159.2024.10539916
- [3] G. Stano, S. M. A. I. Ovy, J. R. Edwards, M. Cianchetti, G. Percoco, and Y. Tadesse, "One-shot additive manufacturing of robotic finger with embedded sensing and actuation," *The International Journal of Advanced Manufacturing Technology*, vol. 124, no. 1, pp. 467–485, (2023); doi: 10.1007/s00170-022-10556-x
- [4] Metastat Insight and Technologies Pvt. Ltd. "Shape memory alloy market size & share report by 2031," Metastat Insight, Accessed: Feb. 21, 2025. [Online]. Available: <https://metastatinsight.com>
- [5] M. Hübler, *Methodik zur Auslegung und Herstellung von aktiven SMA-FKV-Hybridverbunden*. Technische Universität Kaiserslautern, (2016)
- [6] A. Hein, *Auslegung und Entwurf adaptiver Strukturen und Systeme mit Formgedächtnislegierungen*. Stuttgart : Institut für Konstruktionstechnik und Technisches Design, (2021)
- [7] T. M. Wendt, A. Dierle, T. Seifert, and A. Schnebel, "C4.2 - novel characterisation and modeling of the thermal behavior of multi-wire SMA actuators for robot grippers," *Lectures*, pp. 153–154, (2025); doi: 10.5162/SMSI2025/C4.2
- [8] A. Dierle, T. M. Wendt, and T. Seifert, "Integration of shape memory alloy wires into additive fabricated structures for robotic grippers," in *Proceedings of the XI ECCOMAS Thematic Conference on Smart Structures and Materials (SMART 2025)*, in press, (2025)
- [9] DYNALLOY Inc. "Technical characteristics of FLEXINOL® actuator wires," DYNALLOY Inc. Accessed: Dec. 15, 2025. [Online]. Available: <https://dynalloy.com/wp-content/uploads/2025/03/TCF1140.pdf>
- [10] "Nitinol linear actuator (2 way) shape memory alloy wire," Nexmetal Corporation, Accessed: Dec. 23, 2025. [Online]. Available: <https://nexmetal.com/products/nitinol-linear-actuator>
- [11] S. Sławski, M. Kciuk, and W. Klein, "Change in electrical resistance of SMA (NiTi) wires during cyclic stretching," *Sensors (Basel, Switzerland)*, vol. 22, no. 9, p. 3584, (2022); doi: 10.3390/s22093584
- [12] D. Scholtes, S. Seelecke, and P. Motzki, "Electro-thermo-mechanical characterization of shape memory alloy wires for actuator and sensor applications—part 1: The effects of training," *Engineering Reports*, vol. 6, no. 10, e12867, (2024); doi: 10.1002/eng2.12867
- [13] "3D goes HD - ECCO 75 Series high resolution laser sensors," SmartRay, Accessed: Sep. 1, 2025. [Online]. Available: <https://www.smartray.com/de/3d-sensoren/ecco-75-100/>
- [14] H. Song, E. Kubica, and R. Gorbet, "Resistance modelling of shape memory alloy wires," *e-Journal of Nondestructive Testing*, vol. 17, no. 2, (2012)
- [15] "COMSOL 6.3 - shape memory alloy," Accessed: Mar. 25, 2026. [Online]. Available: https://doc.comsol.com/6.3/doc/com.comsol.help.sme/sme_ug_theory.06.041.html#3685019

Eye-Tracking Accuracy in Robot Controls: Separating Genuine Issues from Artificial Barriers

Anke Fischer-Janzen¹, Julia Roederer¹, Thomas M. Wendt¹
Work-Life Robotics Institute, Offenburg University of Applied Science
Max-Planck-Str. 1, 77656 Offenburg

Abstract

Eye-tracking technology has been extensively researched for its potential as an input modality in human-robot interaction. It has undergone significant advancements in recent years, leading to enhanced accuracy. However, despite these improvements, the narrative of eye-tracker inaccuracy, which originated in 2001, remains prevalent in current literature. This critical review aims to investigate the validity of these longstanding claims and determine whether they are supported by substantial evidence or simply perpetuated through repeated citation. By examining influential and representative studies that state limited accuracy, we aim to determine the contemporary relevance of this assertion in various contexts. While not exhaustive, this review identifies key areas in which the inaccuracy claim may still be significant, particularly in specific robotic control applications. Our objective is to provide a nuanced understanding of eye-tracking accuracy and advocate for a contextual approach to evaluating it in modern robotic systems.

Keywords: Robotics, Eye-Tracking, Metrology.

Introduction

The accuracy of eye tracking has improved significantly, reducing errors from over 1° [1] to less than 0.6° with modern devices such as the Tobii Pro Glasses 3. This advancement has mitigated the impact of head movements, slippage and involuntary eye movements, such as lens jitter, which were a significant issue in earlier systems used to control robotic arms. Consequently, the idea that limited accuracy is a general problem with eye-tracking technology is being questioned.

Our previous work indicated that the narrative of insufficient or limited accuracy is frequently echoed by other researchers. In light of the diverse research backgrounds influencing eye-tracking-based robotic controls, it is crucial to adopt distinct perspectives rather than merging these statements into a generalized assumption. The present paper explores specific viewpoints that contribute to claims of limited accuracy in eye-tracking-driven controls, assesses the current validity of these claims, and estimates the accuracy deficiencies reported in the reviewed studies.

Analysis of Accuracy Claims

The present analysis identified 20 such publications in which limited accuracy was

cited as a motivator for research. In four studies it was found that the claim for insufficient eye-tracking accuracy is not valid due to outdated nature of the works cited. The references used are between 10 and 20 years old. The studies in question addressed the technological challenges of that era [2-5].

In the context of systematic reviews and circumstances involving environmental and user-specific influences, user intent recognition, and 3D-gaze estimation, concerns regarding limited accuracy remain valid. It is mandatory that reviews state findings without initial interpretations. Therefore, in the included reviews the statement of insufficient eye tracking accuracy was found [6,7]. Important for new works is that these statements should be evaluated critically on their applicability.

Eye tracking outcomes are influenced by environmental factors, including illumination, as well as user-specific parameters such as ocular morphology and visual impairments [8,9]. Additionally, insufficient accuracy can result from errors in the calibration process [10,11]. Even though calibration might be good, the accuracy of the measurement may deteriorate over time due to the occurrence of slippage [12]. Slippage has been observed to occur at a higher frequency in cases where the eye-tracking glasses are not properly fitted to the participant's head. Accuracy errors in eye tracking glasses are worsened when the user

is forced to look through the periphery of the glasses [13].

In the context of user intent recognition, the accuracy of the eye tracker directly impacts the performance of trained machine learning models. These models often exhibit limited accuracy in predicting correct tasks or objects [14-18]. Given the reduced gaze accuracy, the gaze is consequently tracked next to the original fixation point, resulting in an inability to precisely match gaze patterns with objects or the environment. This mismatch can potentially lead to incorrect indications regarding the designated objects, which, in turn, may result in erroneous task predictions. In the majority of cases, the authors refer to prediction accuracy. Finally, issues with the accuracy of 3D-gaze estimation can be found in the depth measurements [19]. These accuracy errors can occur due to calibration drift and head movements [20]. Especially in robot control, such errors are complex to handle, particularly in the context of 3D-gaze estimation, which is used to track gaze behavior on small object features, obtaining information for user intent recognition based on hand-eye coordination. The depth information is used to instruct the robot to plan grasps or to move to a distinct location. Errors have been shown to result in unsuccessful task performance.

These rationales suggest that while the statement of insufficient or limited accuracy is applicable, it must be contextualized within specific scenarios. This review is not exhaustive and acknowledges that the scope of evaluated contexts does not encompass all potential influences on eye-tracking accuracy. It is important to note that additional biases may exist due to experimental setup, device specifications, and the diverse range of application settings.

Conclusion

The present analysis demonstrates that the extensive claim regarding the lack of accuracy in eye-tracking technology is not universally applicable. Our findings emphasize the importance of acknowledging and accounting for differing contexts and backgrounds in future reviews, in order to derive meaningful conclusions from accuracy metrics. The objective of this work is to emphasize the importance of differentiation by providing specific examples to illustrate the necessity for it. Further research is needed to identify additional contextual differences. It is assumed that analogous observations may be applicable to other research domains involving eye-tracking, where these results could be extrapolated.

It is therefore essential to define what is meant by 'inadequate accuracy'. In consideration of the extensive range of applications, it is not possible to establish a universal threshold. However, evidence suggests that in fields such as user intent recognition and 3D-gaze estimation, sub-millimetre accuracy is essential, as retrieving precise information from small object details is critical for effective control strategies. Conversely, in domains where less nuanced interaction is sufficient, errors up to a few centimetres are tolerable (e.g., in scenarios where gaze needs to align only with coarse bounding boxes to infer intent, supplemented by LLM techniques).

References

- [1] D. Yoo et al., *IEEE/RSJ*, 1196-1201 (2002); doi: 10.1109/IRDS.2002.1043896
- [2] P. M. Tostado, W. W. Abbott, A. A. Faisal, *IEEE ICRA*, 3295-3300 (2016); doi: 10.1109/ICRA.2016.7487502
- [3] W. W. Abbott, A. A. Faisal, *J. Neural Eng.* 9, 046016 (2012); doi: 10.1088/1741-2560/9/4/046016
- [4] A. Graser et al., *IEEE RAM* 20, 148-159 (2013); doi: 10.1109/MRA.2013.2275695
- [5] D. Yoo, J. Kim, D. Kim, M. Chung, *IEEE/RSJ IROS*, 1196-1201 (2002); doi: 10.1109/IRDS.2002.1043896
- [6] E. Suryadarma et al., *JRC* 5 (2024); doi: 10.18196/jrc.v5i3.21686
- [7] A. Fischer-Janzen et al., *Front. Rob. AI* 11 (2024); doi: 10.3389/frobt.2024.1326670
- [8] M. Dahmani et al., *Sensors* 20, 3936 (2020); doi: 10.3390/s20143936
- [9] S. Stalljann et al., *Sensors* 20, 7162 (2020); doi: 10.3390/s20247162
- [10] H. Edughele et al., *IEEE Access* 10, 41952-41972 (2022); doi: 10.1109/ACCESS.2022.3164075
- [11] L. Loke et al., *Wearable Technologies* 6 (2025); doi: 10.1017/wtc.2024.27
- [12] D. C. Niehorster et al., *Behav Res* 52, 1140-1160 (2020); doi: 10.3758/s13428-019-01307-0
- [13] L. Scalera et al., *Robotics* 10, 54 (2021); doi: 10.3390/robotics10020054
- [14] A. Aronson et al., *IEEE/RSJ IROS*, 7307-7314 (2021); doi: 10.1109/IROS51168.2021.9636551
- [15] I. Gonzalez-Diaz, *IEEE J. Biomed. Health Inform* 28, 7517-7530 (2024); doi: 10.1109/JBHI.2024.3430810
- [16] Koochaki et al., *IEEE TNSRE* 29, 974-984 (2021); doi: 10.1109/TNSRE.2021.3083815
- [17] E. Nunez Sardinha et al., arXiv:2401.03944 (2024); doi: [10.48550/arXiv.2401.03944](https://arxiv.org/abs/2401.03944)
- [18] C.-M. Huang, B. Mutlu, *ACM/IEEE HRI*, 83-90 (2016); doi: 10.1109/HRI.2016.7451737
- [19] S. Li, X. Zhang, J. D. Webb, *IEEE Trans. Biomed. Eng.* 64, 2824-2835 (2017); doi: 10.1109/TBME.2017.2677902
- [20] A. Aronson, H. Admoni, *ACM ETRA* 2, 1-9 (2019); doi: 10.1145/3314111.3319840

Waferintegrierte PowderMEMS®-Mikromagnete: CMOS-kompatible Stützfeldquellen für moderne Sensoranwendungen

Ole Behrmann^{1*}, Thomas Lisec¹, Mani Teja Bodduluri¹, Daniel Cichon², Björn Gojdka¹

¹Fraunhofer Institut für Siliziumtechnologie ISIT, Itzehoe

²Fraunhofer Institut für Integrierte Schaltungen IIS, Erlangen

*ole.behrmann@isit.fraunhofer.de

Kurzfassung

Die Miniaturisierung und kosteneffiziente Fertigung magnetischer Sensorsysteme erfordert neue Ansätze zur lokalen Magnetfelderzeugung. Konventionelle, diskret montierte Permanentmagnete sind mit zusätzlichem Aufwand sowie Unsicherheiten in der präzisen Positionierung verbunden. Die Integration von Mikromagneten direkt in das Sensorsubstrat bietet hierfür eine skalierbare und präzisere Lösung. Eine technologische Umsetzung dieser Integration stellt die PowderMEMS®-Technologie dar.

Die PowderMEMS®-Technologie ermöglicht die Wafer-Level-Integration dreidimensionaler NdFeB-Mikromagnetstrukturen in MEMS/CMOS-Prozesse bei Temperaturen unter 150 °C. Die resultierenden Mikromagnete weisen eine Remanenz von über 0,5 T sowie eine hohe thermische Stabilität auf und eignen sich damit als integrierte Stützfeldquellen für Hall-, AMR/GMR/TMR- und NV-Sensoren.

1. Einleitung

Für viele Magnetsensoren – darunter Back-biased-Hall-Sensoren, AMR/GMR/TMR (xMR) Sensoren und NV-Zentren-basierte Quantensensoren – ist ein definiertes magnetisches Bias- oder Stützfeld unerlässlich, um den optimalen Arbeitspunkt zu definieren, die Linearität sicherzustellen und die Robustheit gegenüber äußeren Störfeldern zu gewährleisten [1–3] oder um Vektorfähigkeit zu realisieren. Bisher werden diese Stützfelder überwiegend durch diskret montierte Permanentmagnete erzeugt, die außerhalb oder neben dem eigentlichen Sensorchip platziert werden. Dieses Konzept stößt bei zunehmender Miniaturisierung an grundlegende Grenzen: Die Positioniergenauigkeit bei der Montage liegt typischerweise im Bereich von $\pm 50 \mu\text{m}$, und jeder Montagesschritt erhöht die Produktionskosten sowie das Risiko von Bauteilausfällen. Darüber hinaus stellen die diskreten Magnete einen relevanten Kostenpunkt dar und beanspruchen wertvollen Bauraum, was insbesondere in kompakten Anwendungen wie z.B. Roboter Gelenken oder Kleinfahrzeugen problematisch ist. Zusätzlich ergeben sich zunehmend Probleme bei der Verfügbarkeit von seltenen Erden durch strategische Exportbeschränkungen der Produktionsländer, wodurch entsprechende Einsparmöglichkeiten schon während der Magnetherstellung an Relevanz gewinnen. Die PowderMEMS®-Technologie bietet eine elegante Lösung für diese Herausforderungen: Sie erlaubt die direkte, waferbasierte Integration von dreidimensionalen NdFeB-Mikromagnetstrukturen in Silizium- oder Glassubstrate bei minimalem Verbrauch von Magnetmaterial und Prozesstemperaturen, die mit dem CMOS-Back-End-of-Line (BEOL) kompatibel sind [4]. Dieser Beitrag gibt einen Überblick über das Technologieprinzip und stellt drei repräsentative Sensoranwendungen vor, in denen PowderMEMS®-Mikromagnete erfolgreich als Stützfeldquellen demonstriert wurden.

2. Die PowderMEMS®-Technologie

Die PowderMEMS®-Technologie basiert auf drei aufeinanderfolgenden Schritten, die auf 200 mm-Wafern durchgeführt werden [4]:

1. **Strukturierung von Kavitäten:** Mittels Fotolithografie und Deep Reactive Ion Etching (DRIE) werden präzise Hohlräume mit definierten Tiefen von bis zu mehreren 100 μm in das Siliziumsubstrat geätzt. Dieses Verfahren ermöglicht Positioniergenauigkeiten im Bereich weniger Mikrometer.
2. **Trockenes Befüllen:** Trockenes NdFeB-Pulver (z. B. $d_{50} = 25 \mu\text{m}$) wird durch niederfrequente Vibration und Ultraschall in die Kavitäten eingebracht.
3. **Pulververfestigung:** Die losen Partikel werden durch Atomlagenabscheidung (ALD) von Al_2O_3 (~70 nm, bei 75-150 °C) miteinander zu mechanisch stabilen Mikroformkörpern verbunden.

Nach der Verfestigung kann die Substratoberfläche von überschüssigem Pulver gereinigt werden, sodass die PowderMEMS®-Wafer in Reinraumumgebungen weiterverarbeitet werden können. Die Magnetisierung des gesamten 200 mm-Wafers erfolgt bei Raumtemperatur out-of-plane in einem gepulsten Magnetfeld von 3,5 T, bei einer in-plane Magnetisierung erfolgt diese für individuelle Chips. Das Verfahren ist flexibel einsetzbar: Es kann an nahezu beliebiger Stelle in der MEMS-Fertigung eingebunden werden und ist nicht auf NdFeB beschränkt. Auch die Verwendung von alternativen hart- oder weichmagnetischen Materialien ist möglich [5].

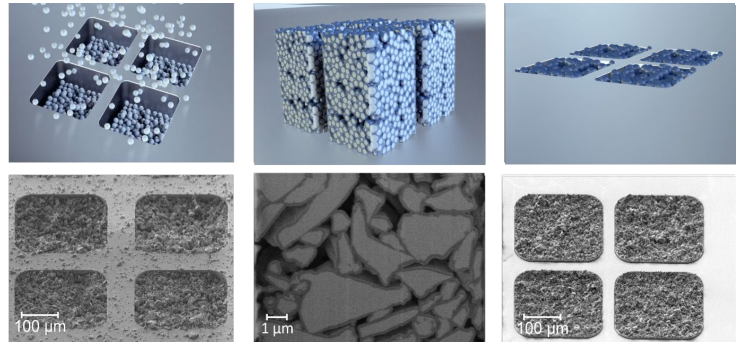


Abbildung 1: Der PowderMEMS® Prozess für die Herstellung von waferintegrierten Mikromagneten [4]. Links: Trockenes Magnetpulver (z. B. NdFeB) wird in Mikrokanalitäten eingebracht. Mitte: Das lose Pulver wird durch ALD von Al_2O_3 zu einem mechanisch stabilen Mikroformkörper verfestigt. Rechts: Bei Bedarf kann die Waferoberfläche gereinigt werden. PowderMEMS® Wafer können hierdurch im MEMS-Reinraum weiterprozessiert werden.

2.1 Magnetische Eigenschaften

Die gefertigten NdFeB-Mikromagnete erreichen eine Remanenz B_r von etwa 0,55 T wobei die intrinsische Koerzitivfeldstärke dem Bulkwert des verwendeten Pulvers entspricht [6]. Die thermische Stabilität wurde für Temperaturen bis 125 °C über mehr als 2000 Stunden nachgewiesen: Nach einem initialen Remanenzverlust von ca. 4,2% in den ersten 500 Stunden, vergleichbar mit gesinterten NdFeB-Magneten, verhalten sich die Mikromagnete langfristig stabil [1, 6].

3. Anwendungen in modernen Sensorarchitekturen

3.1 Back-biased CMOS 3D-Hall-Sensoren

In Kooperation mit dem Fraunhofer IIS wurden die weltweit kleinsten Back-biased CMOS 3D-Hall-Sensoren realisiert. PowderMEMS®-Mikromagnete werden dabei direkt in das Siliziumsubstrat unterhalb der CMOS-prozessierten Hall-Elemente integriert. Die resultierende Gesamtbauhöhe liegt unter 1 mm, was den Einsatz in bauraumkritischen Anwendungen wie Robotergelenken oder kleinen Antriebseinheiten ermöglicht. Die Funktionsfähigkeit wurde anhand der kontaktlosen Drehzahlmessung an einem Zahnrad demonstriert, einer typischen Anwendung in der Industriearbeit und dem Automobilbereich. Im Vergleich zu konventionellen Modulen reduziert der integrierte Ansatz den Bedarf an Seltenen Erden, da wesentlich weniger Magnetmaterial benötigt wird [3].

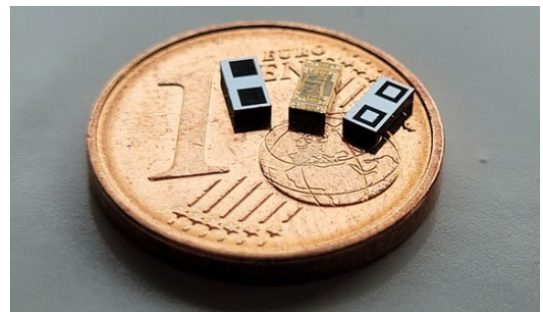


Abbildung 2: Weltweit kleinste Back-biased CMOS Hall-Sensoren mit integrierten PowderMEMS® Mikromagneten [3].

3.2 AMR-Stromsensoren mit substratintegriertem Stützfeld

Anisotrope magnetoresistive (AMR) Sensoren benötigen für ihre Funktion ein definiertes in-plane-Biasfeld, das die magnetoresistiven Schichten in einem definierten Arbeitspunkt hält. In einem kommerziell eingesetzten AMR-Sensorsystem (ADK769, Sensitec GmbH, Wetzlar) werden bislang zwei diskrete Hartferrit-Magnete per Pick-and-Place in das Sensormodul eingebracht (**Abbildung 3**, links).

Im PowderMEMS®-Ansatz werden NdFeB-Mikromagnete direkt in die Rückseite des fertig prozessierten ADK769-Wafers integriert (**Abbildung 3**, rechts). Das damit realisierte In-Plane-Biasfeld beträgt $H_x = 3,109 \pm 68 \text{ kA/m}$ und erfüllt damit die für den zuverlässigen Betrieb des Sensors notwendigen $3 \pm 1 \text{ kA/m}$. Die mit PowderMEMS®-Magneten modifizierten sowie unmodifizierten ADK769 Chips wurden nachfolgend in Stromsensormodule

(CMS2025, Sensitec) eingebaut und verglichen. Beide Chipvarianten zeigen ein nahezu identisches Messverhalten, wodurch die Tauglichkeit des PowderMEMS® Ansatzes demonstriert wird [1].

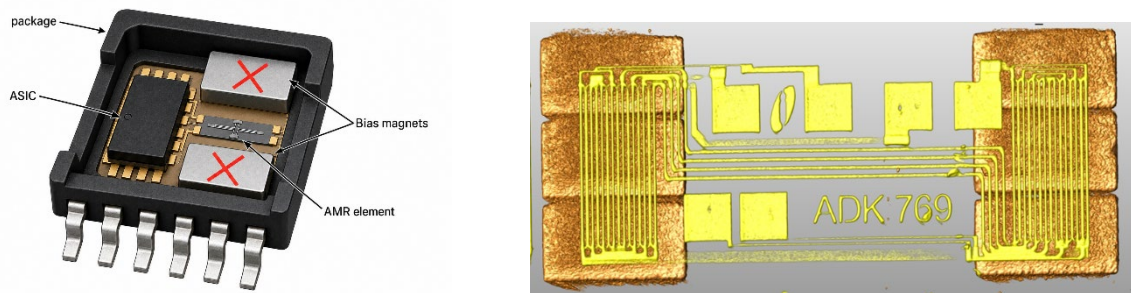


Abbildung 3: Modifikation eines handelsüblichen AMR-Sensors mit PowderMEMS®-Mikromagneten. Links: Bestehender Aufbau mit diskret eingebrachten Stützmagneten (markiert durch rote Kreuze), welche durch PowderMEMS® Mikromagnete ersetzt werden [1]. Rechts: µCT-Aufnahme eines PowderMEMS® modifizierten ADK769 AMR Chips. Die Mikromagnete sind direkt unterhalb der AMR-Elemente angeordnet.

3.3 NV-Zentren-basierte Quantensensoren

Stickstoff-Fehlstellen-Zentren (NV-Zentren) in Diamant gelten als vielversprechende Plattform für hochempfindliche Quantensensoren, da sie bei Raumtemperatur kohärente Spin-Zustände aufweisen. Für die Vektorisierung der Zeeman-Aufspaltung der magnetischen Unterzustände, sowie der Einstellung eines besonders empfindlichen Arbeitspunktes, ist ein definiertes statisches Stützfeld notwendig. Im Rahmen einer Kooperation wurden PowderMEMS®-Mikromagnete unterschiedlicher Geometrien ($700\text{ }\mu\text{m} \times 700\text{ }\mu\text{m}$ und $800\text{ }\mu\text{m} \times 800\text{ }\mu\text{m}$) in Si-Interposer-Chips integriert und als Stützfeldquellen für Diamant-NV-Zentren verwendet. Die Flussdichte der PowderMEMS® Mikromagnete wurde mittels ODMR-Spektroskopie durch Analyse der Zeeman-Aufspaltung vermessen. Es wurde zudem gezeigt, dass die Feldstärke lateral bis zu einem Abstand von 6 mm zum Magnetmittelpunkt homogen ist. Die Integration auf dem Si-Interposer ermöglicht eine drastische Volumenreduktion des bisher makroskopischen Permanentmagneten für NV-Sensoren und ist ein wesentlicher Schritt in Richtung massenfertigungstauglicher Quantensensoren [2].

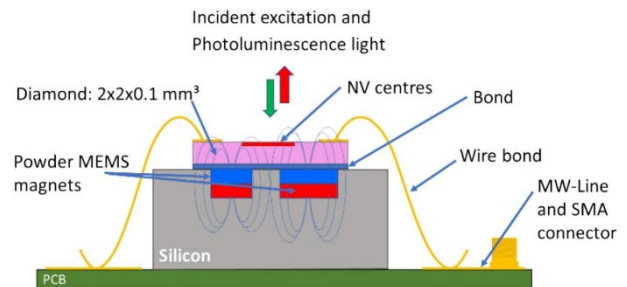


Abbildung 4: Integrationsbeispiel für PowderMEMS Mikromagneten in NV-Sensoren [2].

Quellen

- [1] B. Gojdka et al., "AMR Sensor With Substrate-Integrated PowderMEMS Permanent Magnets for In-Plane Biasing," *IEEE Sensors Journal*, vol. 25, no. 21, pp. 39289–39295, 2025, doi: 10.1109/JSEN.2025.3609635.
- [2] M. Bähr et al., "PowderMEMS® magnets as enabler for miniaturized NV based quantum sensors and quantum processor architectures," *Micro and Nano Engineering*, vol. 28, p. 100316, 2025, doi: 10.1016/j.mne.2025.100316.
- [3] B. Gojdka et al., "Fully Integrated Back-Biased 3d Hall Sensor with Wafer-Level Integrated Permanent Micromagnets," in *Proceedings of the IEEE MEMS*, Munich, Germany, 2023, pp. 1099–1102.
- [4] T. Lisec, O. Behrmann, and B. Gojdka, "PowderMEMS—A Generic Microfabrication Technology for Integrated Three-Dimensional Functional Microstructures," *Micromachines*, vol. 13, no. 3, p. 398, 2022, doi: 10.3390/mi13030398.
- [5] M. Paesler, T. Lisec, and H. Kapels, "High Temperature Magnetic Cores Based on PowderMEMS Technique for Integrated Inductors with Active Cooling," *Micromachines*, vol. 13, no. 3, 2022, doi: 10.3390/mi13030347.
- [6] M. T. Bodduluri, B. Gojdka, N. Wolff, L. Kienle, T. Lisec, and F. Lofink, "Investigation of Wafer-Level Fabricated Permanent Micromagnets for MEMS," *Micromachines*, vol. 13, no. 5, p. 742, 2022, doi: 10.3390/mi13050742.

Sensorbasierte Zustands- und Positionsüberwachung von Umformwerkzeugen im Bereich der Innenhochdruckumformung – Erste Ergebnisse

Thomas Reuter¹, Pierre Fischer¹, Thomas Frank², Kai Nestler³, Andreas Grundmann¹

¹ICM e.V., Otto-Schmerbach-Straße 19, 09117 Chemnitz, Deutschland

²CiS Forschungsinstitut für Mikrosensorik GmbH, Konrad-Zuse-Straße 14, 99099 Erfurt, Deutschland

³KOMITEC electronics GmbH, Albin-Trommler-Straße 2, 08297 Zwönitz, Deutschland

Zusammenfassung

Zustandsüberwachung und vorausschauende Instandhaltung gewinnen in der Produktionsindustrie zunehmend an Bedeutung. Industrielle Condition-Monitoring-Systeme basieren überwiegend auf der Analyse großer Datenmengen, die durch umfangreiche Sensorik an Maschinen und Anlagen erfasst werden. Diese Ansätze ermöglichen zwar eine zuverlässige Überwachung des Anlagenzustands sowie die Ableitung von Verschleißmodellen, erlauben jedoch keine direkte, prozesszyklusbezogene Erfassung der tatsächlichen Werkzeugbelastung. Insbesondere in der Innenhochdruckumformung (IHU), die durch sehr hohe Flächenbelastungen und Schließkräfte von bis zu 50.000 kN gekennzeichnet ist, stellt dies eine zentrale Herausforderung dar. Zur Adressierung dieser Problematik wird ein sensorbasiertes Messkonzept vorgestellt, das die Integration eines Silizium-Dehnungsmessstreifens (Si-DMS) in die Matrize eines IHU-Werkzeugsystems umfasst. Der Sensor ermöglicht eine indirekte Erfassung der Werkzeugbelastung und liefert Rückschlüsse auf die Bauteilumformung. Ergänzend wurde ein Messwertverstärker zur Signalaufbereitung sowie ein Werkzeugpositionierungssystem zur eindeutigen Lokalisierung integriert. Die Validierung des Sensorkonzepts erfolgte anhand der Umformung eines T-Stücks aus dem Werkstoff 1.4301 bei variierenden Umformdrücken zwischen 500 und 2000 bar. Die Auswertung umfasste Prozessgrößen, Sensorsignale sowie geometrische Bauteilmerkmale. Mittels statistischer Analysen und Clusterverfahren konnten eindeutige Zusammenhänge zwischen Sensorsignal und Umformgrad identifiziert werden. Die Ergebnisse zeigen, dass ab einem Umformdruck von 1750 bar ein Ausformungsgrad von 94 % und bei 2000 bar nahezu vollständige Ausformung erreicht wird. Das vorgestellte Sensorkonzept ermöglicht somit eine prozessintegrierte Zustandsüberwachung und bietet Potenzial zur Echtzeitbewertung der Bauteilumformung, zur Optimierung der Prozessführung sowie zur Steigerung der Effizienz und Prozesssicherheit in der IHU.

Keywords: Innenhochdruckumformung (IHU), Umformwerkzeug, Instandhaltung, Silizium-Dehnungsmessstreifen (Si-DMS), Werkzeugmonitoring, Umformsimulation.

Einleitung

Zustandsüberwachung (Condition Monitoring) und vorausschauende Instandhaltung (Predictive Maintenance) stellen in der Produktionsindustrie zunehmend bedeutende Entwicklungs- und Anwendungsfelder dar [1-3]. Industriell eingesetzte Condition-Monitoring-Systeme basieren auf der Analyse umfangreicher Datenmengen, die in der Regel durch die Integration zusätzlicher Sensorik sowie entsprechender Auswerteeinheiten an Maschinen und Anlagen erfasst werden [4]. Die daraus resultierende hohe Sensordichte führt zu großen Datenvolumina, die als Grundlage für eine datengetriebene Modellierung von Produktionssystemen dienen

[2,5,6]. Auf Basis dieser Modelle werden anschließend die Prozessüberwachung sowie die Fehlerdetektion und -diagnose realisiert [2,6,7]. Im Bereich der IHU existieren verschiedene Ansätze, die sich mit dem Qualitätsmanagement, der Prozessüberwachung sowie der vorausschauenden Instandhaltung befassen. Diese Ansätze zielen darauf ab, durch den Einsatz geeigneter Methoden und Technologien eine verbesserte Prozessstabilität, eine frühzeitige Erkennung von Abweichungen sowie eine zustandsorientierte Planung von Instandhaltungsmaßnahmen zu ermöglichen. Dabei kommen insbesondere datenbasierte Verfahren zum Einsatz, die eine kontinuierliche

Analyse relevanter Prozess- und Zustandsgrößen erlauben und somit zur Steigerung der Produktqualität und Anlagenverfügbarkeit beitragen [8-10]. Die Überwachung ist jedoch ausschließlich auf den Anlagenzustand ausgerichtet. Auf Grundlage der erfassten Daten können Muster identifiziert und Modelle abgeleitet werden, die Rückschlüsse auf einen potenziellen Werkzeugverschleiß zulassen. Eine direkte, prozesszyklusbezogene Erfassung bzw. Quantifizierung der tatsächlich auf das Werkzeug wirkenden Belastungen während eines einzelnen Umformvorgangs ist mit derartigen Systemen jedoch nicht möglich [2,4]. Insbesondere bei Werkzeugen, die sehr hohen Flächenbelastungen ausgesetzt sind, wie sie in der IHU mit Schließkräften von bis zu 50.000 kN auftreten, steht derzeit keine hinreichend zuverlässige Sensorik zur Verfügung, um die reale Werkzeugbelastung messtechnisch direkt zu erfassen. Neben der Schließkraft wirken zusätzlich der Wirkmediendruck im Bauteil sowie axiale Vorschubbewegungen des Halbzeugs als maßgebliche Belastungsgrößen auf das Werkzeug [11,12]. Darüber hinaus ist neben der fehlenden direkten Messbarkeit der Werkzeugbeanspruchung auch keine präzise, prozessimmanente Überwachung der Bauteilumformung möglich. Die eingestellten Prozessparameter üben jedoch einen signifikanten Einfluss auf den Werkzeugverschleiß aus. Diese Parameter werden in der Regel im Vorfeld mithilfe von FEM-Simulationen bestimmt, müssen jedoch im realen Prozess häufig angepasst werden, da nicht alle relevanten Einflussgrößen hinreichend genau abgebildet werden können [12,13]. Zur Minimierung dieser Unsicherheiten werden die Prozessparameter häufig mit konservativen Sicherheitsaufschlägen ausgelegt, sodass deutlich höhere Werte gewählt werden, als für den eigentlichen Umformvorgang erforderlich wären. Dies führt zu verlängerten Prozesszeiten und bewirkt eine signifikante Reduktion der Effizienz des IHU-Prozesses.

Vor diesem Hintergrund kommt der Prozess- und Zustandsüberwachung während der Bauteilumformung eine zentrale Bedeutung zu. Insbesondere eine Echtzeitüberwachung der Ausformung in Kombination mit der Möglichkeit zur prozessbegleitenden Optimierung sowie zur präzisen Bestimmung der Werkzeugbelastung würde es erlauben, die gewonnenen Daten unmittelbar in die Prozessauslegung und -regelung zu integrieren [11,12]. Dies eröffnet das Potenzial für eine ganzheitliche Digitalisierung des IHU-

Fertigungsprozesses. Daraus ergeben sich insbesondere Vorteile hinsichtlich der frühzeitigen Erkennung und Behebung von Produktionsabweichungen, eines prozesssicheren Werkzeugeinsatzes, einer zustandsorientierten Instandhaltung, einer durchgängigen Qualitätsüberwachung sowie einer Steigerung der Prozesseffizienz [9,14]. Ein möglicher Lösungsansatz wird in diesem Beitrag vorgestellt. Dieser basiert auf einem sensorbasierten Messmodul, das speziell für IHU-Werkzeugsysteme entwickelt wurde und eine präzise Erfassung der Werkzeugbelastung ermöglicht. Dadurch wird ein effektives Werkzeugmonitoring realisiert. Ergänzend wird ein Werkzeugpositionierungssystem integriert, das eine eindeutige und reproduzierbare Bestimmung der Werkzeugposition im Unternehmen sicherstellt.

Theoretische Grundlagen

In den folgenden Abschnitten werden die theoretischen Grundlagen der Instandhaltung, der Innenhochdruckumformung (IHU) sowie der Werkzeugüberwachung erläutert.

Instandhaltung

Die Instandhaltung dient dem Erhalt bzw. der Wiederherstellung der Funktionsfähigkeit von Maschinen und Anlagen. Hierfür stehen verschiedene Instandhaltungskonzepte zur Verfügung, die in unterschiedliche Klassen unterteilt werden können, darunter reaktive (schadensabhängige), vorbeugende (planmäßige), zustandsorientierte, prädiktive (vorausschauende), aufgeschobene sowie periodische Instandhaltung [15].

Die Zustandsüberwachung bildet dabei die Grundlage zustandsorientierter und prädiktiver Instandhaltungskonzepte. Sie umfasst die kontinuierliche Erfassung und Auswertung von Maschinen- und Prozessparametern zur Beurteilung des aktuellen Anlagenzustands.

Bei der reaktiven Instandhaltung werden keine präventiven Maßnahmen zur Sicherung der Funktionsfähigkeit durchgeführt. Erst beim Auftreten eines Defekts erfolgt eine Instandsetzung, wobei eine möglichst schnelle Wiederherstellung des Betriebs zur Minimierung von Folgekosten angestrebt wird.

Im Gegensatz dazu basiert die prädiktive Instandhaltung auf der fortlaufenden Zustandsüberwachung während des Betriebs. Durch die statistische Auswertung der erfassten Daten können Zustandsprognosen erstellt und optimale Wartungszeitpunkte bestimmt werden, wodurch sich die Lebensdauer von Anlagenkomponenten verlängern lässt [15, 16].

Innenhochdruckumformung (IHU)

Der IHU-Prozess unterscheidet sich von konventionellen Umformverfahren dadurch, dass die Umformkraft nicht durch starre Werkzeuggeometrien, sondern durch ein unter Druck stehendes Wirkmedium erzeugt wird. In den Abbildungen 1 und 2 sind der Prozess am Beispiel der Herstellung eines T-Stücks schematisch dargestellt sowie das IHU-Werkzeug mit dem umgeformten Bauteil abgebildet. Dabei werden Hohlkörper (Halbzeuge), in der Regel gerade oder gebogene Rohre, in ein Umformwerkzeug, bestehend aus Ober- und Untermatrize, eingelegt und anschließend durch eine Presse geschlossen.

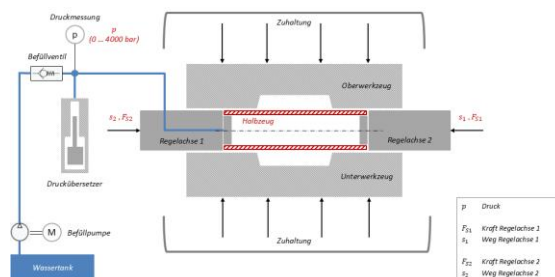


Abb. 1: Funktionsprinzip der IHU am Beispiel eines T-Stücks.



Abb. 2: IHU-Werkzeug mit umgeformtem Bauteil (T-Stück).

Im weiteren Verlauf werden die Enden des Halbzeugs mittels Dichtstempeln abgedichtet und über mindestens eine in den Dichtstempeln integrierte Bohrung mit einem Umformmedium gefüllt. Dieses Medium, üblicherweise eine Wasser- oder Ölemulsion, wird unter hohen Druck gesetzt, sodass eine plastische Verformung des im Werkzeug befindlichen Halbzeugs einsetzt. Zusätzlich wird während des Umformprozesses eine axiale Kraft über die Dichtstempel eingebracht, wodurch das Halbzeug gestaucht und Material gezielt in die Hauptumformzone nachgeführt

wird. Auf diese Weise können hohe Umformgrade bei gleichzeitig weitgehend homogener Wanddickenverteilung realisiert werden [12,13].

Werkzeugüberwachung

In der industriellen Praxis erfolgt die Werkzeugzustandsüberwachung überwiegend indirekt über Maschinen- und Prozessgrößen wie Presskraft-, Druck- oder Wegsignale. Diese ermöglichen jedoch nur eingeschränkte Rückschlüsse auf die lokale Werkzeugbelastung und den tatsächlichen Verschleißzustand. Eine direkte messtechnische Erfassung der im Werkzeug wirkenden Spannungs- und Dehnungszustände ist aufgrund der hohen Belastungen und der eingeschränkten Integrationsmöglichkeiten von Sensorik nur begrenzt realisierbar [11,17]. Zur Erweiterung der Informationsbasis werden daher zunehmend sensorbasierte und modellgestützte Ansätze eingesetzt, bei denen beispielsweise Dehnungsmessstreifen, piezoelektrische Sensoren oder FEM-basierte Simulationen zur Abschätzung nicht direkt messbarer Größen genutzt werden. Insbesondere in der IHU-Forschung werden hybride Ansätze aus Simulation und experimenteller Sensorintegration untersucht, um eine verbesserte Zustandsbewertung des Werkzeugs zu ermöglichen [18,19].

Material und Methoden

In den nachfolgenden Abschnitten werden der experimentelle Versuchsaufbau sowie die Methodik der Versuchsdurchführung, der Umformsimulation und der Datenauswertung detailliert beschrieben.

Experimentelles Setup

Grundlage für die Erfassung der Werkzeugbelastung bildet ein Si-DMS, der als sensorischer Stift ausgeführt und in die Matrize des IHU-Werkzeugsystems integriert wurde (Abbildung 3).



Abb. 3: Sensorstift mit Si-DMS (Fixierung: links geklebt, rechts Glasfritte).

Der Si-DMS zeichnet sich durch eine hohe Empfindlichkeit (k-Faktor 80; $0,4 \text{ mV}/(\text{V} \cdot \text{MPa})$), eine miniaturisierte Bauform sowie eine hohe Langzeitstabilität aus. Er ermöglicht die indirekte Erfassung der im IHU-Werkzeug wirkenden Belastung und liefert damit Rückschlüsse auf die plastische Verformung des Bauteils innerhalb der Matrize, ohne in der direkten Umformzone positioniert zu sein. Der Sensor ist im Bereich der maximalen Umformung, in unmittelbarer Nähe des Radius der Domausbildung, angeordnet (Abbildung 4).

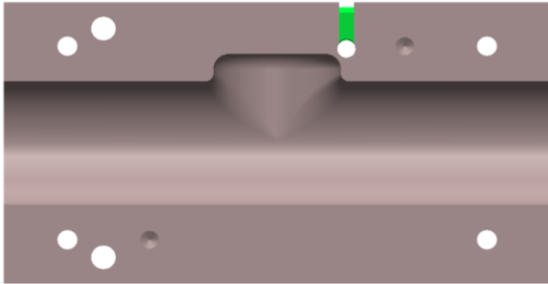


Abb. 4: Umformwerkzeug Obergesenk mit Bohrung (grün) für den Sensorstift.

Zur Signalaufbereitung wurde ein Messwertverstärker entwickelt, der das Sensorsignal erfasst, mittels gezielter Filterung und Differenzverstärkung von Störanteilen bereinigt und zu einem rauscharmen Ausgangssignal verstärkt. Das resultierende DC-Signal wird über eine integrierte M12-Rundsteckverbindung an ein Messdatenerfassungssystem mit AD-Wandler übertragen. Über dieselbe Schnittstelle erfolgt die Spannungsversorgung des Messverstärkers im Bereich von 12–24 V. Das Ausgangssignal des DMS wird als prozentualer Anteil der Referenzspannung von 10 V ausgegeben. Ergänzend ist ein Werkzeugpositionierungssystem integriert, das eine präzise Richtungsbestimmung sowie eine zuverlässige Indoor-Ortung in geschlossenen Produktionsumgebungen ermöglicht. Hierzu sendet das System kontinuierlich Signale an umliegende Antennenplatinen, in denen sowohl Einfallswinkel als auch Entfernung ausgewertet werden. Durch den Einsatz von mindestens drei Antennenplatinen kann mittels Triangulation eine eindeutige Positionsbestimmung des Werkzeugs realisiert werden.

Versuchsdurchführung

Zur Validierung des Sensorkonzepts wurde ein IHU-typisches Bauteil untersucht. Dabei wurde die Ausformung eines T-Stücks mit definierter Geometrie betrachtet. Als Ausgangsmaterial diente ein Halbzeug aus dem Werkstoff 1.4301 mit den Abmessungen $70 \times 2 \times 135 \text{ mm}$. Die

Zielgeometrie bestand aus einer Domausbildung mit einem Plateau-durchmesser von 52,8 mm in Längs- und Querrichtung. Die Umformversuche wurden bei variierenden Umformdrücken im Bereich von 500 bar (keine Umformung) bis 2000 bar (nahezu vollständige Ausformung) durchgeführt. Hierbei wurde eine ausreichende Anzahl an Versuchsproben berücksichtigt ($N = 5$ pro Umformdruck). In Abbildung 5 ist der Versuchsaufbau in der IHU-Anlage dargestellt.



Abb. 5: IHU-Anlage mit integriertem Versuchsaufbau (1 Umformwerkzeug mit Unter- und Obergesenk, 2 Messtechnik, 3 Zugbügel, 4 Hydraulik, 5 T-Stück).

Zur Bewertung der Umformergebnisse wurden die gefertigten Bauteile geometrisch mit dem Atos Q System vermessen (Auflösung 12 MP, Messvolumen $170 \times 130 \times 130 \text{ mm}$). Die ermittelten Geometriedaten wurden anschließend mit den Ergebnissen der Umformsimulation verglichen, um den jeweiligen Umformungsgrad quantitativ zu bestimmen. Auf dieser Grundlage konnte eine Einordnung der Prozesszustände in Abhängigkeit des applizierten Umformdrucks erfolgen.

Umformsimulation

Der Innenhochdruck-Umformprozess wurde mittels numerischer Simulation virtuell abgebildet. Dabei wurde die Werkzeugbewegung unter Annahme ideal starrer, reibungsbehafteter Werkzeuge beschrieben. Auf Basis des empirisch ermittelten Druck - Zeit - Verlaufs sowie geometrischer, tribologischer und werkstoffspezifischer Parameter wurde in einer

ersten idealisierten Simulationsphase das Umformverhalten des Halbzeugs ausschließlich unter Einwirkung des Innendrucks analysiert, wobei axiale Stauchkräfte zunächst nicht berücksichtigt wurden. Zur Reduktion des Rechenaufwands wurden vereinfachte Modellannahmen durch den Einsatz von Halbmodellen getroffen. Als Werkstoff wurde der Edelstahl 1.4301 mit hinterlegter Fließkurve verwendet. Der Werkzeugkontakt wurde durch einen kombinierten Reibwert ($\mu_{\text{Coulomb}} = 0,04$, $\mu_{\text{Schub}} = 0,06$) beschrieben. Zur verbesserten geometrischen Abbildung der Domausbildung erfolgte eine lokale Verfeinerung der Modellauflösung in den relevanten Umformbereichen. Die Parameter maximaler Umformgrad sowie die geometrischen Abmessungen der Domausbildung wurden ausgewertet.

Datenauswertung

Die statistische Auswertung der Prozesskurven, bestehend aus Umformdruck, Nachschiebeweg und Zuhaltekraft, sowie der Sensorsignale und der geometrischen Abmessungen der umgeformten T-Stücke erfolgte mit dem Ziel, Korrelationen zu identifizieren und Cluster zur Beschreibung des Zusammenhangs zwischen Sensorsignal und Umformprozess zu bilden.

Ergebnisse und Diskussion

In den Abbildungen 6 bis 8 sind die Ergebnisse der DMS-Messung exemplarisch für einen Umformdruck von 2000 bar den IHU-Prozessparametern (Umformdruck, Zuhaltekraft und Nachschiebeweg) gegenübergestellt. Die Kurvenverläufe zeigen, dass das DMS-Signal den Prozessverlauf eindeutig abbildet und eine klare Korrelation zu den betrachteten Prozessparametern aufweist. Auf dieser Grundlage kann der IHU-Prozess zuverlässig beschrieben und in charakteristische Prozessphasen segmentiert werden. Am Beispiel der T-Stückumformung ergibt sich eine Gliederung in 13 zeitlich aufeinanderfolgende Prozessphasen, die den vollständigen Ablauf von der Werkstückhandhabung über die Umformung bis hin zur Werkzeugöffnung abbilden (Abbildung 9). Dazu zählen das Schließen der Schutztür (1), das Schließen des Werkzeugs (2), das Schwenken des Bügels (3) sowie das Spannen (4) und die Stempelbewegung (5). Im Anschluss erfolgen die prozessrelevanten Umformschritte, beginnend mit dem Vorfüllen (5), gefolgt vom Druckaufbau (6), der Phase des maximalen Drucks (8) sowie dem anschließenden Druckabbau (9). Daraufhin

folgen die Rückfahrt der Stempel (10) und das Entspannen des Systems (11). Abschließend werden der Bügel erneut geschwenkt (12) und das Werkzeug geöffnet (13), womit der Prozesszyklus beendet wird.

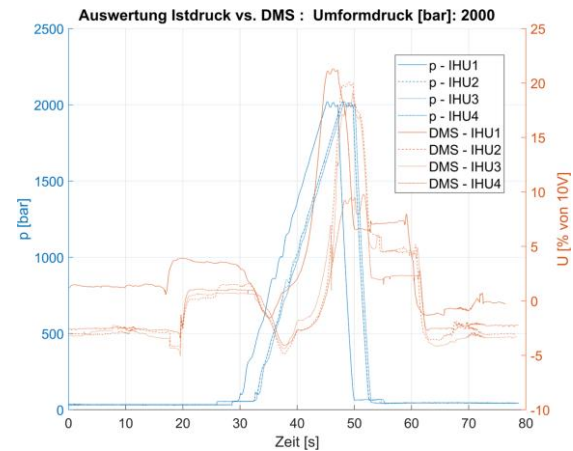


Abb. 6: Vergleich DMS-Signal (rot) mit dem Umformdruckverlauf (blau).

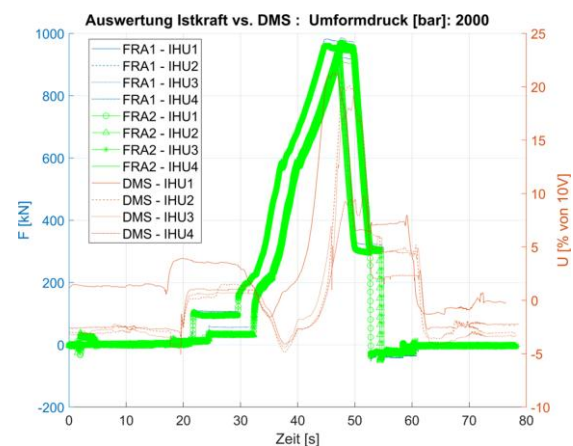


Abb. 7: Vergleich DMS-Signal (rot) mit dem Zuhaltekraftverlauf (grün).

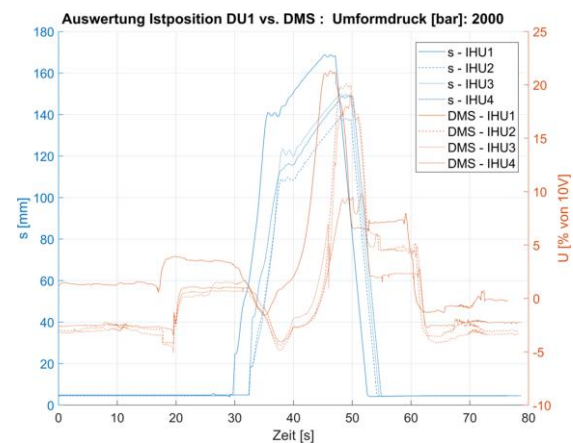


Abb. 8: Vergleich DMS-Signal (rot) mit dem Nachschiebewegverlauf (blau).

Für die Analyse der Werkzeugbelastung und der Bauteilumformung sind insbesondere die Prozessphasen 4 bis 8 von Relevanz.

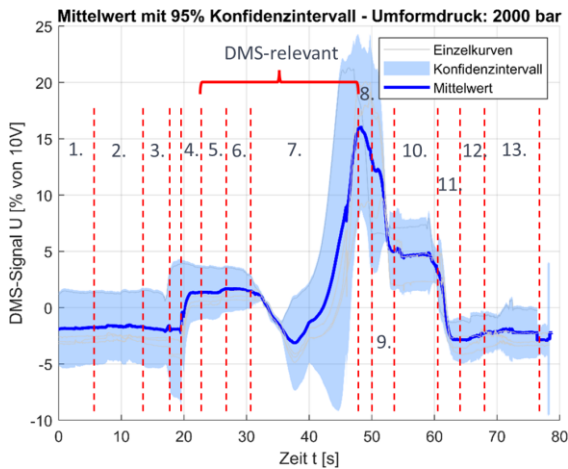


Abb. 9: DMS-Signal (Mittelwert und Konfidenzintervall) bei 2000 bar Umformdruck mit den charakteristischen Prozessphasen.

Der Vergleich der mittleren DMS-Signale in den Prozessphasen 4 bis 8 ist in Abhängigkeit des Umformdrucks im Bereich von 500 bis 2000 bar in Abbildung 10 dargestellt.

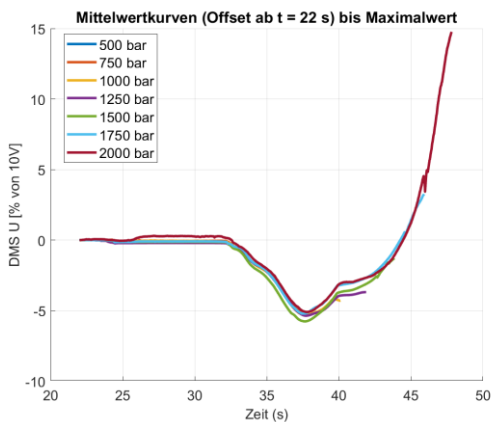


Abb. 10: Vergleich der mittleren DMS-Signale für Umformdrücke im Bereich von 500 bis 2000 bar in den Prozessphasen 4 bis 8.

Der Kurvenverlauf ist für alle untersuchten Umformdrücke grundsätzlich konsistent und zeigt ab dem Ende von Phase 6 einen Abfall des Signals bis zu einem Minimum in der Mitte von Phase 7, gefolgt von einem anschließenden erneuten Anstieg. Die Signalsteigung sowie die Signalthöhe variieren dabei in Abhängigkeit des eingestellten Umformdrucks. Der charakteristische Kurvenverlauf ist auf die Position des SI-DMS zurückzuführen, der ein resultierendes

Summensignal der überlagerten Kraftkomponenten erfasst. Für die Herstellung eines Bezugs zur Bauteilumformung sind die experimentell ermittelten Domabmessungen sowie die Ergebnisse der Umformsimulation in Tabelle 1 zusammenfassend dargestellt. Die Abweichungen zwischen Experiment und Simulation liegen im Bereich von 4 % (2000 bar) bis 35 % (750 bar). Es ist ersichtlich, dass bei einem Umformdruck von 500 bar keine Bauteilumformung stattfindet, während ab 1750 bar eine Ausformung von über 94 % erreicht wird. In Abbildung 11 ist die Umformsimulation des T-Stücks für einen Umformdruck von 2000 bar dargestellt. Der maximale Umformgrad beträgt in der Simulation zwischen 0,13 (750 bar) und 0,33 (1750 und 2000 bar).

Tab. 1: Experimentelle und simulativ ermittelte Geometriedaten im Dombereich des T-Stücks.

p/ bar	Domabmessung längs [mm]		Domabmessung quer [mm]	
	exp.	sim.	exp.	sim.
500	-	-	-	-
750	29,3	21,1	21,6	13,9
1000	46,7	42,8	39,3	34,8
1250	48,8	47,7	44,2	41,2
1500	52,2	49,8	49,0	45,2
1750	53,8	52,1	52,1	49,3
2000	54,4	52,3	53,3	52,2

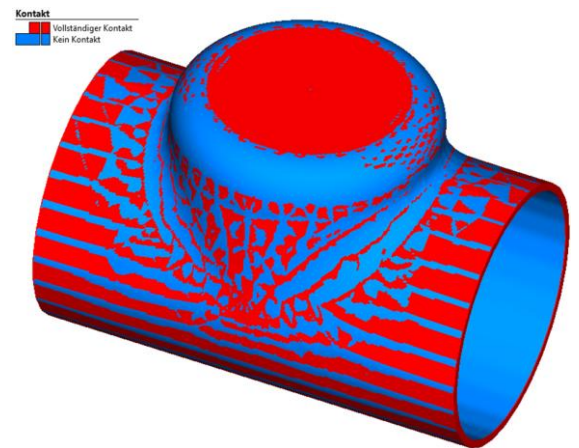


Abb. 11: Umformsimulation bei 2000 bar Umformdruck (Werkzeugkontakt: rot – vollständig, blau – kein Kontakt).

Die Datenanalyse ergab, dass mithilfe der experimentell und simulativ ermittelten Geometriedaten sowie der DMS-Signale Cluster gebildet werden können, die einen eindeutigen Zusammenhang zwischen

Bauteilform und Umformdruck aufzeigen (lineare Korrelation mit $R^2 = 0,63$). Als Kennwerte wurden hierbei die Maximal- und Minimalwerte der DMS-Signale verwendet. Die für die Clusterbildung herangezogenen Daten wurden zuvor dimensionslos normiert. In Abbildung 12 ist die Clusterbildung unter den entsprechenden physikalischen Randbedingungen dargestellt.

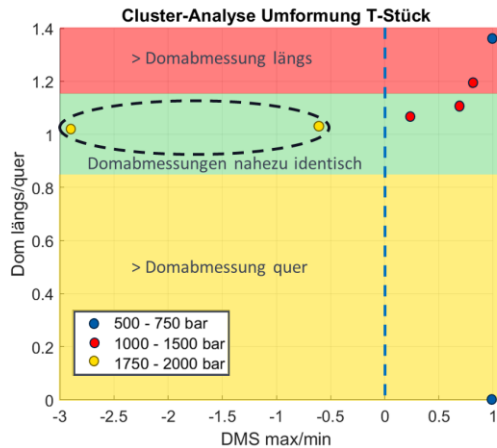


Abb. 12: Clusterbildung in Abhängigkeit von Domabmessung (Längsrichtung zu Querrichtung) und DMS-Signal (maximaler zu minimaler DMS-Signalwert) (schwarze Linie - >94 % ausgeformt, blaue Linie - keine bzw. <85 % Ausformung).

Aus der Clusterbildung wurde eine Kennfelddarstellung abgeleitet. Diese zeigt, dass Bauteile mit nahezu identischen Domabmessungen in Längs- und Querrichtung einem Verhältnis von etwa 1,0 entsprechen (grüner Bereich). Dieser Bereich ist als prozessseitig (nahezu) vollständig ausgeformt zu bewerten, da eine weitgehend homogene Domausbildung vorliegt. Höhere bzw. niedrigere Verhältniswerte weisen hingegen auf eine bevorzugte Ausformung in Längs- bzw. Querrichtung und damit auf eine unzureichende Gesamtformgebung hin (roter bzw. gelber Bereich). Zudem wird ersichtlich, dass unterschiedliche Druckniveaus charakteristische Signal- und Ausformzustände hervorrufen. Im hohen Druckbereich (1750–2000 bar) konzentrieren sich die Messpunkte überwiegend im Bereich nahezu identischer Domabmessungen, was auf eine stabile und reproduzierbare Bauteilformgebung schließen lässt; gleichzeitig liegen die DMS-Signalverhältnisse im negativen Max/Min-Bereich (schwarze Linie, Ausformung >94 %). Im mittleren Druckbereich (1000–1500 bar) treten hingegen erhöhte Streuungen auf, die auf Übergangszustände der Ausformung hindeuten. Im niedrigen Druckbereich (500–

750 bar) dominieren deutliche Abweichungen, die auf unvollständige Ausformungen zurückzuführen sind. Sowohl im mittleren als auch im niedrigen Druckbereich liegen die DMS-Signalverhältnisse im positiven Max/Min-Bereich (blaue Linie), was Ausformgrade von unter 85 % kennzeichnet.

Zusammenfassung und Ausblick

Basierend auf den erfassten Messwerten wurde ein systematisches Vorgehen entwickelt, das die Ableitung von Informationen zur Werkzeugverformung im kritischen Bereich, zur Ausformungserkennung sowie zur Bestimmung von Stückzahlen pro Werkzeug ermöglicht. Die Validierungsversuche zeigen eine hohe Wiederholgenauigkeit sowie eine ausgeprägte Sensitivität des entwickelten Sensorkonzepts. Mittels Cluster-Analyse konnten zudem eindeutige Zusammenhänge zwischen Prozessparametern und Sensorsignal identifiziert werden.

Die Ergebnisse belegen, dass ab einem Umformdruck von 1750 bar ein Ausformungsgrad von 94 % erreicht wird, während bei 2000 bar nahezu vollständige Ausformung (99 %) vorliegt. Damit ermöglicht das Sensorkonzept eine unmittelbare Ableitung des Umformgrades aus dem Prozesssignal.

Im Ausblick ist die Durchführung weiterer Versuche mit unterschiedlichen IHU-Bauteilgeometrien vorgesehen, um die Übertragbarkeit des Ansatzes zu erweitern. Darüber hinaus soll zum einen die Sensorpositionierung mithilfe numerischer Umformsimulationen weiter optimiert werden, zum anderen erfolgt eine Untersuchung der DMS-Signale in Abhängigkeit von der Messrichtung (x- bzw. y-Richtung) mit dem Ziel, den Einfluss der einzelnen Kraftkomponenten zu quantifizieren.

Aufbauend auf den gewonnenen Erkenntnissen wird ein Auswertealgorithmus entwickelt, der eine automatisierte Analyse und anwenderfreundliche Aufbereitung der Messdaten ermöglicht. Dieser soll den Abgleich mit vorgegebenen Kennwerten sowie einen prozessnahen Eingriff bei unvollständiger Ausformung oder Prozessabweichungen unterstützen. Ziel ist zudem die Integration einer automatisierten Fehlererkennung mit Anbindung an übergeordnete Systeme zur gezielten Ausschleusung fehlerhafter Bauteile im Rahmen einer durchgängigen Prozesskette.

Danksagung

Das Forschungs- und Entwicklungsprojekt „Entwicklung eines sensorgestützten

Messmoduls zur Zustands- und Positionsüberwachung für die Integration in ein Umformwerkzeug als Basis für ein effizientes Werkzeugmonitoring“ wird vom Bundesministerium für Wirtschaft und Energie (BMWi) aufgrund eines Beschlusses des Deutschen Bundestages gefördert (Förderkennzeichen 16KN104422) und vom Projektträger VDI/VDE Innovation + Technik GmbH betreut.

Literaturnachweis

- [1] T. Reuter, I. Taraschuk, S. Liebl, C. Luft, T. Modaleck, Entwicklung eines Überwachungs- und Servicemanagementsystems für Sterilisations- und Schredderanlagen: Methodik zur Erarbeitung eines dynamischen Wartungsregimes aus Expertenwissen und Maschinendaten, *ZWF* 117(5), 275–282 (2022); <https://doi.org/10.1515/zwf-2022-1065>
- [2] C. Tsallis, P. Papageorgas, D. Piromalis, R.A. Munteanu, Application-Wise Review of Machine Learning-Based Predictive Maintenance: Trends, Challenges, and Future Directions, *Appl. Sci.* 15, 4898 (2025); <https://doi.org/10.3390/app15094898>
- [3] T. Reuter, J. Schmidt, Á. Pérez, A. Grundmann, Vorausschauende Instandhaltung und Effizienzbewertung von Niederspannungsmotoren: Entwicklung eines universell einsetzbaren Condition-Monitoring-Systems Teil 2: Prozessüberwachung und Effizienzbewertung am Beispiel der mechanischen Glasbearbeitung, *ZWF* 119(9), 624–632 (2024); <https://doi.org/10.1515/zwf-2024-1124>
- [4] T. Reuter, J. Schmidt, A. Grundmann, Entwicklung eines Condition-Monitoring-Systems für Niederspannungsmotoren – Teil 1: Untersuchung von Belastungs- und Schadensfällen mittels Schwingungsmessung, *ZWF* 117(10), 659–666 (2022); <https://doi.org/10.1515/zwf-2022-1139>
- [5] M. Tshilidzi, Condition Monitoring Using Computational Intelligence Methods – Applications in Mechanical and Electrical Systems, Springer-Verlag, London (2012); <https://doi.org/10.1007/978-1-4471-2380-4>
- [6] R. Isermann, Fault-Diagnosis Systems – An Introduction from Fault Detection to Fault Tolerance, 1. Auflage, Springer-Verlag, Berlin, Heidelberg (2006); <https://doi.org/10.1007/3-540-30368-5>
- [7] M. K. Menon, R. Tuladhar, Asset Management decision-making through data-driven Predictive Maintenance – an overview, techniques, benefits and challenges, *Maintenance, Reliability and Condition Monitoring*, 4(2), (2024); <https://doi.org/10.21595/marc.2024.24232>
- [8] T. Reuter, K. Massalsky, K. Hoyer, G. Ivanov, T. Burkhardt, Qualitätsmanagement und vorausschauende Instandhaltung auf der Basis synthetischer Datensätze: Virtuelle Technologieentwicklung für den Mittelstand am Beispiel der Innenhochdruck-Umformung, *ZWF* 116(10), 673–680 (2021); <https://doi.org/10.1515/zwf-2021-0167>
- [9] T. Reuter, K. Massalsky, T. Burkhardt, Concept Development for Quality Management and Predictive Maintenance in the Area of hydroforming (IHU), In: 29. *Sächsische Fachtagung Umformtechnik*, Technische Universität Dresden, Artikelnummer 25, (2024); <https://doi.org/10.25368/2024.12>
- [10] T. Reuter, K. Massalsky, T. Burkhardt, Fertigungsprozessanalyse eines mittels Innenhochdruck-Umformung hergestellten Bauteils: Anwendung von Data Science für die vorausschauende Instandhaltung und Qualitätssicherung, *ZWF* 119(10), 742–748 (2024); <https://doi.org/10.1515/zwf-2024-1142>
- [11] T. Altan, G. Ngaile, G. Shen, Cold and hot forging: Fundamentals and applications, *ASM International*, (2005); <https://doi.org/10.31399/asm.tb.chffa.9781627083003>
- [12] R. Neugebauer, Hydro-Umformung, VDI-Buch, (2007), Springer, Berlin, Heidelberg; https://doi.org/10.1007/978-3-540-49013-5_2
- [13] M. Koc, T. Altan, An Overall Review of the Tube Hydroforming (THF) Technology, *Journal of Materials Processing Technology* 108, 384–393 (2001); [https://doi.org/10.1016/S0924-0136\(00\)00830-X](https://doi.org/10.1016/S0924-0136(00)00830-X)
- [14] T. Burkhardt, S. Liebl, K. Massalsky, R. Riedel, N. Göhlert, P. Wilsky, R. Prielipp, Neue Anwendungen für das Innenhochdruckumformen – Mit smarten Konzepten den Marktzugang zum Innenhochdruckumformen für KMU nachhaltig verbessern, *wt Werkstattstechnik online* 108(10), 698–702 (2018); <https://doi.org/10.37544/1436-4980-2018-10-62>
- [15] M. Schenk, Instandhaltung technischer Systeme: Methoden und Werkzeuge zur Gewährleistung eines sicheren und wirtschaftlichen Anlagenbetriebs, 1. Auflage, Springer-Verlag, Berlin, Heidelberg (2010); <https://doi.org/10.1007/978-3-642-03949-2>
- [16] G. Pawellek, Integrierte Instandhaltung und Ersatzteillogistik: Vorgehensweisen, Methoden, Tools, 2. Auflage, Springer-Vieweg-Verlag, Wiesbaden (2016); https://doi.org/10.1007/978-3-662-48667-2_4
- [17] G. Byrne, D. Dornfeld, I. Inasaki, G. Ketteler, W. König, R. Teti, Tool condition monitoring (TCM) - The status of research and industrial application, *CIRP Annals*, 44(2), 541–567 (1995); [https://doi.org/10.1016/S0007-8506\(07\)60503-4](https://doi.org/10.1016/S0007-8506(07)60503-4)
- [18] R. Teti, K. Jemielniak, G. O'Donnell, D. Dornfeld, Advanced monitoring of machining operations, *CIRP Annals*, 59(2), 717–739 (2010); <https://doi.org/10.1016/j.cirp.2010.05.010>
- [19] M. Papananias, Non-Intrusive Monitoring of Machining Processes for In-Process Product Health Prediction based on Machine Learning, *Procedia CIRP* 138, 346–351 (2026); <https://doi.org/10.1016/j.procir.2026.01.060>