

Optimale Interferenzfilter für die chromatisch konfokale 3D Messtechnik im Sinne von Bayes

Miro Taphanel, Jürgen Beyerer

Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung,
e-mail: miro.taphanel@iosb.fraunhofer.de
e-mail: juergen.beyerer@iosb.fraunhofer.de

Kurzfassung

Mittels statistischer Versuchsplanung können Sensoren für ihren Anwendungszweck optimiert werden. Methodisch werden dabei die frei einstellbaren Sensorparameter auf optimale Weise gewählt, mit dem Ziel die resultierende Messunsicherheit zu reduzieren. In dieser Arbeit werden Filtertransmissionen von Interferenzfilter durch *Bayesian Experimental Design* auf optimale Weise gewählt. Eingesetzt werden diese optischen Filter zur Schätzung der 3D Oberflächentopographie in einem chromatisch konfokalen 3D Messverfahren.

1 Einleitung

Das chromatische Messprinzip basiert auf einer spektralen Kodierung der Messgröße. Im Falle des hier betrachteten chromatisch konfokalen 3D Mikroskops weist das an der Objektoberfläche reflektierte Lichtspektrum einen monochromatischen Konfokalpeak auf [1]. Jede Wellenlänge, die die Lage des Konfokalpeaks bezeichnet, korrespondiert mit einer Objekthöhe.

Klassischer Weise beinhaltet ein chromatisch konfokales 3D Mikroskop ein Spektrometer, welches das an der Objektoberfläche reflektierte Lichtspektrum abtastet. Digitale Signalverarbeitung kann dann die Lage des Konfokalpeaks bestimmen, z.B. indem das Intensitätsmaximum gesucht wird [1]. In dieser Arbeit wird eine grundsätzlich andere Signalverarbeitung vorgeschlagen, indem anstatt eines Spektrometers eine multispektrale Kamera Verwendung findet [2].

Im Fall einer multispektralen Kamera mit sechs Kanälen wird das an der Objektoberfläche reflektierte Licht in Form von sechs Grauwerten beobachtet. Jeder Kamerakanal weist dabei eine andere spektrale Empfindlichkeit auf, welche jeweils durch einen Interferenzfilter technisch realisiert werden kann. Während einer Messung hat die angeschlossene digitale Signalverarbeitung die Aufgabe, dem beobachteten Grauwertvektor die zugrundeliegende Objekthöhe zuzuweisen. Die Filtertransmissionen der einzelnen Kamerakanäle bestimmen dabei im Wesentlichen die Messunsicherheit dieses Messverfahrens. Ziel dieser Arbeit ist die Optimierung optimaler Interferenzfilter, wobei als Gütefunktional die Messunsicherheit betrachtet wird und methodisch das *Bayesian Experimental Design* angewandt wird.

Zur Optimierung eines Sensors mittels statistischer Versuchsplanung wird eine Messung als ein physikalisches Experiment betrachtet, dessen Designparameter in optimaler Weise gewählt werden sollen. Durch die messtechnische Anwendung kann die Güte eines Sensors durch seine Messunsicherheit bewertet werden, welche als die Varianz der Schätzung der Messgröße definiert ist. Nach dem Stand der Forschung können drei wesentliche Ansätze zur Quantifizierung der Messunsicherheit unterschieden werden, wobei alle der Parameter-Schätztheorie entstammen. Der erste Ansatz basiert auf einer analytischen Bestimmung der Schätzvarianz für eine betrachtete Schätzfunktion [3]. Dieses Vorgehen ist jedoch nur in Spezialfällen möglich, z.B. wenn ein lineares Sensormodell vorliegt oder der Fehler durch eine Linearisierung toleriert werden kann. Die zweite Methodik basiert auf der *Cramér-Rao* Ungleichung, welche eine untere Schranke für die Schätzvarianz darstellt [4]. Anstatt die Schätzvarianz direkt zu minimieren, wird stattdessen eine untere Schranke minimiert [5]. Der Vorteil dieses Ansatzes ist die kompakte analytische Form der *Cramér-Rao* Ungleichung. Jedoch werden aufgrund des Charakters einer unteren Schranke nicht alle Phänomene erfasst, die sich negativ auf die Messunsicherheit auswirken können. Besonders im Falle eines nichtlinearen Sensormodells ist dies problematisch [6]. Als drittes methodisches Vorgehen können numerische Ansätze genannt werden, welche die Schätzvarianz für eine konkrete Schätzfunktion numerisch ermitteln. Dazu wird in dieser Arbeit das Bayes'sche experimentelle Design betrachtet [7]. In dieser Arbeit wird ein vierter Ansatz vorgestellt,

welcher die rechenaufwändige numerische Simulation approximiert und auf der *Bhattacharyya* Ungleichung basiert.

Der Beitrag gliedert sich wie folgt: Zu Beginn wird ein chromatisch konfokales 3D Mikroskop modelliert. In Abschnitt 3 wird das Optimierungsfunktional formuliert und die Approximation zur schnelleren Auswertung vorgestellt. Exemplarische Optimierungsergebnisse sind in Abschnitt 4 zu finden. Der Beitrag schließt mit einer Zusammenfassung der wesentlichen Aussagen.

2 Sensormodell

Ein beobachteter Grauwertvektor $\mathbf{g} = (g_1, \dots, g_n)^\top$ einer n kanaligen multispektralen Kamera unterliegt Rauschen und wird als normalverteilter Zufallsvektor $G \sim \mathcal{N}(\mathbf{g}; \mathbf{g}_\mu, \Sigma)$ modelliert. Der Erwartungswert

$$\mathbf{g}_\mu(h, \mathbf{x}, \mathbf{p}) = r(\mathbf{x}) \int_{\lambda_{\min}}^{\lambda_{\max}} \mathbf{f}(\lambda, \mathbf{p}) s(\lambda, \lambda_f(h)) d\lambda$$

beschreibt einen Grauwertvektor als Funktion der skalaren Objektreflektanz $r(\mathbf{x})$ als Funktion des Ortes \mathbf{x} und als Funktion des Spektrums $s(\lambda, \lambda_f(h))$. Dieses kann als *Gauß*-Kurve modelliert werden, mit einer zentralen Wellenlänge $\lambda_f(h)$, welche je nach Messgröße in einer anderen Objekthöhe h fokussiert ist. Des Weiteren sind die optischen Filter der Kamerakanäle in einem Vektor $\mathbf{f}(\lambda, \mathbf{p})$ organisiert. Die Transmissionen $\mathbf{f}(\lambda, \mathbf{p})$ hängen dabei von der Wellenlänge λ ab und können durch die Designparameter \mathbf{p} eingestellt werden. Im Falle von Interferenzfilter enthält \mathbf{p} die Schichtdicken der einzelnen Filter, welche später optimiert werden. Abschließend wird noch ein breitbandiger Kamerapixel angenommen über den spektral integriert wird. Quanteneffizienz, Beleuchtungsspektrum, Kameraquantisierung und weitere Aspekte werden hier aus Gründen der Einfachheit vernachlässigt. Mit einem detaillierteren Sensormodell wurde z.B. in [2] gearbeitet. Die Kovarianzmatrix

$$\Sigma = \text{diag}(\sigma^2(g_{\mu,1}), \dots, \sigma^2(g_{\mu,n})) \text{ mit } \sigma^2(g_{\mu,i}) = \sigma_d^2 + k g_{\mu,i}$$

als Diagonalmatrix zeigt die angenommene stochastische Unabhängigkeit der einzelnen Kamerakanäle. Die Abhängigkeit der einzelnen Varianzen $\sigma^2(g_{\mu,i})$

vom Erwartungswert ist auf eine Approximation der Poisson Verteilung durch die Normalverteilung zurück zu führen. Die Poisson Verteilung modelliert dabei das Photonenrauschen der Lichtquelle. Ebenso berücksichtigt wird ein Rauschterm für das sogenannte Dunkelrauschen σ_d und k stellt die sogenannte Systemverstärkung dar.

Zur Schätzung der Messgröße h aus einer Beobachtung \mathbf{g} muss auch die nicht interessierende Oberflächenreflektanz $r(\mathbf{x})$ als Störparameter mit geschätzt werden. Nur dann kann mittels des Sensormodells eine Beobachtung erklärt werden. Alternativ kann die Betrachtung auch, wie in der Bildverarbeitung üblich, intensitätsnormiert erfolgen. Dazu wird ein normierter Zufallsvektor $G_{\text{norm}} = G / \|\mathbf{g}_\mu\|_2$ eingeführt, welcher zur Folge hat, dass sich $r(\mathbf{x})$ aus dem Erwartungswert herauskürzt. Jedoch gilt $\text{Var}\{aX + b\} = a^2 \text{Var}\{X\}$ und

$$\Sigma_{\text{norm}} = \text{diag}\left(\sigma_{\text{norm}}^2(\mathbf{g}_{\mu,1}), \dots, \sigma_{\text{norm}}^2(\mathbf{g}_{\mu,n})\right) \text{ mit } \sigma_{\text{norm}}^2(\mathbf{g}_{\mu,i}) = \frac{\sigma_d^2 + k \mathbf{g}_{\mu,i}}{\|\mathbf{g}_\mu\|_2^2}.$$

Folglich kürzt sich die Oberflächenreflektanz aus der Kovarianzmatrix nicht heraus. Der Grund hierfür liegt in der approximierten Poisson-Verteilung, gemäß derer es einen Unterschied macht, wie viele Photonen eingesammelt werden. Somit wird für die folgenden Betrachtungen eine minimale Reflektanz $r(\mathbf{x}) = r_{\min} = 0.7$ angenommen.

3 Bayes'sche Sensoroptimierung

Das beste Experiment im Sinne von Bayes maximiert die erwartete Nützlichkeit

$$E\{u(\mathbf{p}^*)\} = \arg \max_{\mathbf{p}} \int \int u(\mathbf{g}, h, \mathbf{p}) p(\mathbf{g} | h, \mathbf{p}) p(h) d\mathbf{g} dh,$$

indem über alle möglichen Beobachtungsdaten integriert wird [7]. In diesem Formalismus stellt auch h eine Zufallsvariable dar, welche durch ihre Verteilung $p(h)$ charakterisiert ist. Da keine Information darüber vorliegt, welche Objekthöhen mit welcher Wahrscheinlichkeit vorkommen, wird $p(h)$ als nicht-informative Gleichverteilung angenommen. Die Nutzenfunktion $u(\mathbf{g}, h, \mathbf{p}) = (h - \hat{h}(\mathbf{g}))^2$ bewertet dabei die Messunsicherheit als quadratische Abweichung der Schätzung $\hat{h}(\mathbf{g})$ vom wahren Wert h . Die Wahrscheinlich-

keitsdichtefunktion $p(\mathbf{g}|h, \mathbf{p})$ beschreibt die Normalverteilung des normierten Sensormodells aus Abschnitt 2. Zur Optimierung muss das Funktional

$$J(\mathbf{p}) = \iint (h - \hat{h}(\mathbf{g}))^2 p(\mathbf{g}|h, \mathbf{p}) d\mathbf{g} dh$$

minimiert werden. Aus Sicht einer Implementierung ist dabei die mehrdimensionale Integration $\int (\cdot) d\mathbf{g}$ problematisch, da keine effiziente Lösung bekannt ist; eine Übersicht gängiger Integrationsverfahren ist in [7] zu finden. Im folgenden Abschnitt wird eine auf der *Bhattacharyya* Ungleichung basierende approximative Lösung des obigen Optimierungsfunktional vorgestellt. Durch diese konnte in einer Referenzimplementierung die Rechenzeit um den Faktor 800 reduziert werden.

3.1 Approximation des Optimierungsfunktional

Als erste Vereinfachung wird das kontinuierliche Integral $\int (\cdot) dh$ durch eine Summe ersetzt

$$J(\mathbf{p}) \approx \frac{1}{m} \sum_{i=1}^m \int (h_i - \hat{h}(\mathbf{g}))^2 p(\mathbf{g}|h_i, \mathbf{p}) d\mathbf{g}$$

und diskretisiert. Im nächsten Schritt wird zur Motivation der Approximation die Schätzfunktion

$$\hat{h}(\mathbf{g}) = \arg \min_h \|\mathbf{g} / \|\mathbf{g}\|_2 - \mathbf{g}_{\mu, \text{norm}}(h, \mathbf{p})\|_2^2$$

als *Least-Square*-Schätzer festgelegt. Diese wählt immer denjenigen Schätzwert $\hat{h}(\mathbf{g}) = h_j$, welcher eine minimale euklidische Distanz zwischen Beobachtung \mathbf{g} und modellierter Beobachtung $\mathbf{g}_{\mu, \text{norm}}(h_j)$ zur Folge hat. Um jede diskrete Beobachtung $\mathbf{g}_{\mu, \text{norm}}(h_j)$ kann somit eine Region $R_j = \{\mathbf{g} \in R_j \mid \hat{h}(\mathbf{g}) = h_j\}$ definiert werden, in welcher alle Beobachtungen dem Schätzwert h_i zugeordnet werden. Diese Aufteilung des Beobachtungsraumes entspricht einem Voronoi-Diagramm. Somit kann auch das Integral $\int (\cdot) d\mathbf{g}$ als Summe einzelner Integrale über diese Regionen geschrieben werden.

$$\begin{aligned} J(\mathbf{p}) &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n \int_{R_j} (h_i - \hat{h}(\mathbf{g}))^2 p(\mathbf{g}|h_i, \mathbf{p}) d\mathbf{g} \\ &= \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n (h_i - h_j)^2 \int_{R_j} p(\mathbf{g}|h_i, \mathbf{p}) d\mathbf{g} \end{aligned}$$

Durch diese Darstellung muss nur das verbleibende Integral über die Wahrscheinlichkeitsdichteverteilung gelöst werden. Dieses kann als eine Fehlerwahrscheinlichkeit interpretiert werden, welche die Wahrscheinlichkeit quantifiziert, dass h_i zugrunde liegt und h_j geschätzt wird. Im Fall einer Normalverteilung und bei der Betrachtung eines Zweiklassenfalls, existiert für dieses Integral eine geschlossene Lösung in Form einer oberen Schranke als sogenannte *Bhattacharyya* Ungleichung [8]. Diese Ungleichung stellt selbst wiederum eine Approximation der *Chernoff* Ungleichung dar. Der Zweiklassenfall ignoriert dabei die Existenz aller möglichen Höhenwerte außer h_i und h_j . Dadurch ergibt sich das approximative Optimierungsfunktional

$$J(\mathbf{p}) = \frac{1}{m} \sum_{i=1}^m \sum_{j=1}^n (h_i - h_j)^2 e^{-k} \quad \text{mit}$$

$$k = \frac{1}{8} (\mathbf{g}_{\mu,i} - \mathbf{g}_{\mu,j})^\top \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mathbf{g}_{\mu,i} - \mathbf{g}_{\mu,j}) + \frac{1}{2} \ln \frac{\left| \frac{\Sigma_i + \Sigma_j}{2} \right|}{\sqrt{|\Sigma_i| |\Sigma_j|}}$$

Bei Untersuchungen hat sich herausgestellt, dass diese Approximation die Messunsicherheit zum Teil erheblich unterschätzt. Als heuristische Erweiterung wurde dazu eine nichtlineare Gewichtung der Kovarianzmatrix hinzugefügt, welche das besagte Verhalten kompensieren kann:

$$\Sigma_{i,j} = \frac{2}{1 + \exp(|h_i - h_j| / a_1)} \Sigma_i$$

Nicht benachbarte Messpunkte werden dann durch eine größere Varianz gewichtet, wobei die ursprünglich in Σ_i modellierte Varianzen maximal verdoppelt werden. Die Nachbarschaft bezieht sich dabei auf den relativen Höhenunterschied $|h_i - h_j|$. Der heuristische Parameter a_1 muss entsprechend für jedes Rauschmodell geeignet gewählt werden. In dem approximierten Optimierungsfunktional wird Σ_i durch $\Sigma_{i,j}$ ersetzt und muss entsprechend für jede Kombination von i und j ausgewertet werden. Analog wird Σ_j durch $\Sigma_{j,i}$ ersetzt.

4 Optimierungsergebnisse

In Bild 1 ist ein Optimierungsergebnis mit 4 Interferenzfiltern dargestellt. Die Transmissionscharakteristiken der einzelnen Filter realisieren immer wieder

steile Flanken. Daraus resultiert eine hohe Empfindlichkeit für die Schätzung der Wellenlänge von annähernd monochromatischem Licht. Es sei darauf hingewiesen, dass der dargestellte spektrale Ausschnitt nur 50 nm beträgt.

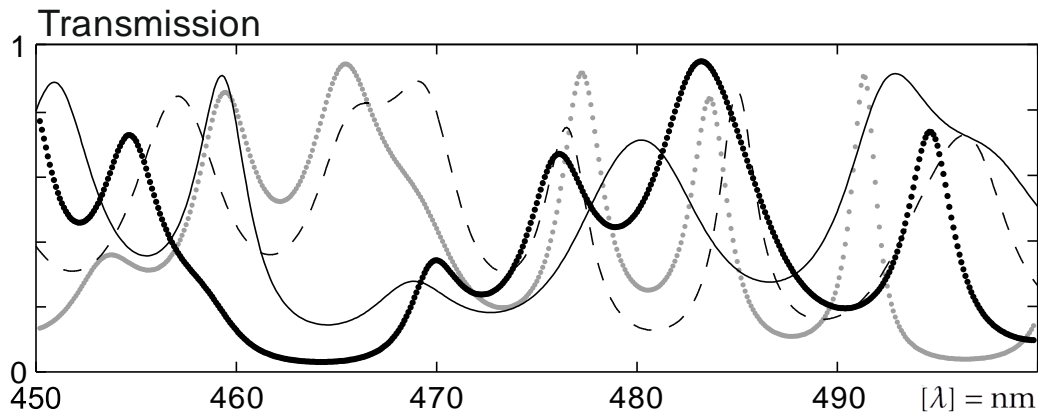


Bild 1: Filtertransmissionen von 4 Filtern als Funktion der Wellenlänge. Die Filter wurden für einen CCT Sensor zur Schätzung der Wellenlänge von fast monochromatischem Licht (FWHM 2,5 nm) optimiert.

In Bild 2 ist die zugehörige Messunsicherheit als Funktion der Wellenlänge dargestellt. Diese könnte alternativ auch als Funktion der Objekthöhe dargestellt werden, erlaubt aber in dieser Darstellung eine bessere Zuordnung zu den Filtertransmissionen. Die Messunsicherheit quantifiziert die Standardabweichung der Wellenlängenschätzung. Die nicht unterbrochene Funktion beschreibt dabei die vorgestellte Approximation mittels *Bhattacharyya* Ungleichung. Die gestrichelte Funktion in Bild 2 stellt eine numerische Lösung der Messunsicherheit mittels *Monte-Carlo* Simulation dar. Beide Verläufe stimmen gut überein, wobei die Approximation charakteristische Peaks aufweist. An diesen Stellen nähern sich die Beobachtungen für nicht benachbarte Objekthöhen an und die Messunsicherheit wird überschätzt.

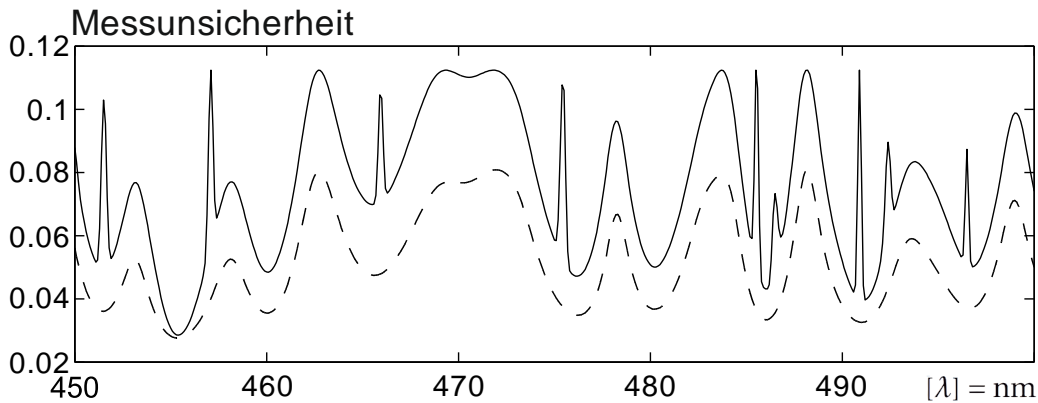


Bild 2: Messunsicherheit als Standardabweichung der Wellenlängenschätzung. Ermittelt per *Monte-Carlo* Simulation (unterbrochene Linie) und approximativ mittels *Bhattacharyya* Ungleichung (kontinuierliche Linie).

5 Zusammenfassung

Die Methodik des *Bayesian Experimental Design*'s kann in einer Optimierungsberechnung direkt genutzt werden, die Messunsicherheit von Sensoren zu reduzieren. Problematisch ist jedoch der resultierende hohe Rechenaufwand der Optimierungsberechnung, da ein mehrdimensionales Integral numerisch gelöst werden muss. In dieser Arbeit wurde eine approximative Lösung vorgestellt, welche eine geschlossene Lösung des Optimierungsfunktional bereitstellt. Dadurch konnte die Optimierung um einen Faktor 800 beschleunigt werden. Das Ergebnis der Optimierung zeigt eine gute Übereinstimmung zwischen numerischer und approximativer Lösung.

Literatur

- [1] VDI/VDE 2655, Blatt 1.2, Beuth Verlag, 2010
- [2] Taphanel, M., Hovestreydt, B., Beyerer, J.: Speed-up chromatic sensors by optimized optical filters. In: *SPIE Optical Metrology 2013*. International Society for Optics and Photonics, 2013. S. 87880S-87880S-10.
- [3] Miedema, M. A. O., Van den Bos, A., Buist, A. H.: Experimental design of exit wave reconstruction from a transmission electron microscope defocus series. *Instrumentation and Measurement, IEEE Transactions on*, 1994, 43. Jg., Nr. 2, S. 181-186.

- [4] van den Bos, A.: Parameter estimation for scientists and engineers, John Wiley & Sons, 2007.
- [5] van Aert S.: Optimal experimental design of STEM measurement of atom column positions. *Ultramicroscopy*, 2002, 90. Jg., Nr. 4, S. 273-289.
- [6] Taphanel, M.: Frequentistic and a Bayesian Approach for Optimal Optical Filter Design, Proceedings of the 2012 Joint Workshop of Fraunhofer IOSB and Institute for Anthropomatics, Vision and Fusion Laboratory, Jürgen B., Pak, A., 2013
- [7] von Toussaint, U.: Bayesian inference in physics, *Reviews of Modern Physics*, 2011, 83. Jg., Nr. 3, S. 943.
- [8] Duda, R. O., Hart, P. E., Stork, D. G.: Pattern classification, John Wiley & Sons, 2012.