AHMT

XXV. Messtechnisches Symposium

des Arbeitskreises der Hochschullehrer
für Messtechnik e.V.

22.-24. September 2011
in Karlsruhe

Tagungsband

Messtechnik und Sensorik

Fernando Puente León, Jürgen Beyerer (Hrsg.)

XXV. Messtechnisches Symposium

des Arbeitskreises der Hochschullehrer
für Messtechnik e.V.

22.-24. September 2011
in Karlsruhe

Shaker Verlag
Aachen 2011

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Copyright Shaker Verlag 2011
Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 978-3-8440-0388-8
ISSN 1610-4773

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen
Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9
Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

Seit einem Vierteljahrhundert widmet sich der Arbeitskreis der Hochschullehrer für Messtechnik e.V. (AHMT) sowohl der Förderung der Messtechnik als universitäre, fachübergreifende Grundlagendisziplin in allen ingenieurwissenschaftlichen Studiengängen als auch der Förderung des wissenschaftlichen Nachwuchses auf diesem Gebiet. In diesem Kontext sind insbesondere die Bedeutung des jährlichen Messtechnischen Symposiums des AHMT und die Vergabe des Messtechnik-Preises hervorzuheben.

Im vorliegenden Jahr findet das XXV. Messtechnische Symposium des AHMT am Karlsruher Institut für Technologie (KIT) statt. Zuständig für die Ausrichtung dieser Fachtagung sind

- das Institut für Industrielle Informationstechnik (IIIT) des KIT,
- der Lehrstuhl für Interaktive Echtzeitsysteme (IES) des KIT sowie
- das Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung (IOSB).

Zusätzlich feiert der AHMT in diesem Jahr sein XXV. Jubiläum. Es ist uns daher eine besondere Ehre, die Leitung und Organisation dieses Messtechnischen Symposiums übernehmen zu dürfen. Im Rahmen der zweitägigen Fachtagung werden vor allem junge Nachwuchswissenschaftler über aktuelle Forschungsergebnisse berichten. Der wissenschaftliche Höhepunkt der Veranstaltung wird mit dem Gastvortrag von Herrn Prof. Dr.-Ing. Dr. h.c. mult. Eckehard Schnieder und der anschließenden Verleihung des renommierten Messtechnik-Preises erreicht. Ganz herzlich möchten wir Herrn Dr. rer. nat. Guido Bartl für den Erhalt des diesjährigen Messtechnik-Preises gratulieren.

Für das wissenschaftliche Programm wurden 28 Beiträge ausgewählt, die einen umfassenden Überblick über die laufenden Forschungsaktivitäten auf dem Gebiet der Messtechnik im deutschsprachigen Raum wiedergeben. Sie decken nahezu alle Bereiche der universitären Messtechnik ab. Auch diesmal findet eine Posterausstellung statt, um die thematische Bandbreite zu erhöhen.

Unser besonderer Dank gilt den Mitarbeitern des KIT, die uns mit der Organisation dieser Tagung tatkräftig unterstützt haben. Stellvertretend seien an dieser

Stelle Frau Dipl.-Ing. Kristine Back, Herr Dr.-Ing. Konrad Christ, Frau Manuela Koffler und Herr Dr.-Ing. Andreas Sandmair genannt. Bedanken möchten wir uns ebenfalls bei der Firma National Instruments für die großzügige Spende des diesjährigen Messtechnik-Preises.

Karlsruhe, im September 2011

Fernando Puente León
Karlsruher Institut für Technologie

Jürgen Beyerer
Fraunhofer IOSB und Karlsruher Institut für Technologie

Inhalt

1 Signalverarbeitung und modellgestützte Verfahren

Untersuchung von Körperschallsignalen im Zeit-Frequenzbereich für ein körperschallbasiertes Motormanagement.....**1**
Marco Decker, Karsten Hintz und Clemens Gühmann
FG Elektronische Mess- und Diagnosetechnik, TU Berlin
Institut für Mobile Systeme, Otto-von-Guericke-Universität Magdeburg

Erfassung menschlicher Bewegungen durch merkmalsbasiertes Tracking.....**15**
Kristine Back und Fernando Puente León
Institut für Industrielle Informationstechnik, Karlsruher Institut für Technologie

Magnetische Vektorhysterese: Messtechnik und Modellierung.....**27**
Alexander Sutor, Jan Kallwies und Reinhard Lerch
Lehrstuhl für Sensorik, Friedrich-Alexander-Universität Erlangen-Nürnberg

Messtechnische Hardwareplattform zur Entwicklung neuer Motorsteuergeräte-Algorithmen und Funktionen am Beispiel der Klopferkennung.....**39**
Matthias Biehl, Stefan Kempf und Andreas König
Robert Bosch GmbH, Schwieberdingen
Lehrstuhl für Integrierte Sensorsysteme, Technische Universität Kaiserslautern

Drahtlose Messtechnik zur Langzeitüberwachung von Vitalparametern.....**51**
Timo Klingeberg und Meinhard Schilling
Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik, TU Braunschweig

Quellentrennung für unterbestimmte Szenarien im Frequenzbereich.....**63**
Andreas Sandmair und Fernando Puente León
Institut für Industrielle Informationstechnik, Karlsruher Institut für Technologie

Pore Detection on Cylinder Bore Surfaces.....**75**
Limeng Wang und Fernando Puente León
Institut für Industrielle Informationstechnik, Karlsruher Institut für Technologie

2 Optische Messverfahren

Probabilistische Verbesserung der Kalibrierung deflektometrischer Systeme.....**83**
Holger Rapp und Christoph Stiller
Institut für Mess- und Regelungstechnik, Karlsruher Institut für Technologie

Materialidentifikation mittels optisch realisierter Kreuzkorrelation der Reflektanzspektren.....**91**
Miro Sauerland, Robin Gruna und Jürgen Beyerer
Institut für Anthropomatik, Lehrstuhl für Interaktive Echtzeitsysteme, Karlsruher Institut für Technologie
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung, Karlsruhe

Übertragungsverhalten von tiefenscannenden Weißlichtinterferometern.....**103**
Weichang Xie, Jan Niehues und Peter Lehmann
FB Elektrotechnik/Informatik, Lehrstuhl für Messtechnik, Universität Kassel

Phasenrauschen in optischen Messsystemen mit strukturierter Beleuchtung.....**115**
Marc Fischer, Marcus Petz und Rainer Tutsch
Institut für Produktionsmesstechnik, Technische Universität Braunschweig

Verfahren und Vorrichtung zur Inspektion von Glasoberflächen.....**127**
Sebastian Höfer, Masoud Roschani, Stefan Werling und Jürgen Beyerer
Institut für Anthropomatik, Lehrstuhl für Interaktive Echtzeitsysteme, Karlsruher Institut für Technologie
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung, Karlsruhe

Integrierte CMOS PIN-Fototransistoren für kurzwelliges infrarotes Licht.....**139**
Plamen Kostov und Horst Zimmermann
Institute of Electrodynamics, Microwave and Circuit Engineering, TU Wien

Zerstörungsfreie Bestimmung der chemischen Bindungszustände vergrabener Nanoschichten durch Röntgenspektrometrie mit Synchrotronstrahlung.....**151**
Beatrix Pollakowski
Physikalisch-Technische Bundesanstalt, Berlin

Einsatz von Messsystemen auf Basis der Streifenprojektions- und Streifenreflexionstechnik zur Konturmessung und Oberflächen-charakterisierung.....**163**
Marc Sandner, Wansong Li, Christoph von Kopylow und Ralf B. Bergmann
Bremer Institut für angewandte Strahltechnik, Bremen
Vereinigte Elektronik Werkstätten GmbH, Bremen

Miniaturisierter IR-Öl-Sensor (MIRÖS).....**175**
Benjamin R. Wiesent, Daniel G. Dorigo, Özlem Simsek und Alexander W. Koch
Lehrstuhl für Messsystem- und Sensortechnik, Technische Universität München

3 Präzisionsmesstechnik und Sensorik

Piezoresistives Silizium-Messelement zur Messung im Hochdruck-Bereich bei 5.000 bar.....**187**
Patrick Heinickel und Roland Werthschützky
Institut für Elektromechanische Konstruktionen,
Technische Universität Darmstadt

Online Trace Gas Detection - Pushing a Bruker MicrOTOF II Towards ppq-Detection Limits Using Simple Atmospheric Pressure Ionization Sources.....**199**
Jens Langejuergen, Philipp Cochems und Stefan Zimmermann
Institute of Electrical Engineering and Measurement Technology,
Leibniz University Hannover

Präzisionsbestimmung der Boltzmann-Konstante an der Physikalisch-Technischen Bundesanstalt.....**211**
Thorsten Zandt, Christof Gaiser und Bernd Fellmuth
Physikalisch-Technische Bundesanstalt, Berlin

Rückgeführte Größenmessung von Nanopartikeln im Transmissionsmodus eines Rasterelektronenmikroskops (TSEM).....**227**
Tobias Klein, Egbert Buhr, Klaus-Peter Johnsen und Carl Georg Frase
Physikalisch-Technische Bundesanstalt, Braunschweig

Neuartiges Kalibrierstrahlersystem zur Präzisionskalibrierung von Strahlungsthermometern.....**241**
Marc Schalles und Thomas Fröhlich
Technische Universität Ilmenau

4 Fertigungsmesstechnik

Unsicherheit bei der Datenfusion dimensioneller Messungen.....**253**
Maryna Galovska, Marcus Petz und Rainer Tutsch
Institut für Produktionsmesstechnik, TU Braunschweig

Ermittlung einer aufgabenspezifischen Messunsicherheit für Verzahnungsmessungen.....**265**
Kerstin Rost, Frank Härting und Gert Goch
Physikalisch-Technische Bundesanstalt, Braunschweig und Berlin
Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft, BIMAQ, Universität Bremen

Design und Fertigung hochminiaturisierter Fasersonden für interferometrische Messungen.....**277**
Robert Schmitt und Niels König
Fraunhofer-Institut für Produktionstechnologie, Aachen

Berührungslose Messung des Drehwinkels und der Axialbewegung an Rotorwellen von Windenergieanlagen.....**289**
Jan F. Westerkamp, Michael Sorg, Carsten Bredemeier und Gert Goch
Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft, BIMAQ, Universität Bremen

Auswertung flächenhafter Messungen an Großverzahnungen.....**299**
Jan F. Westerkamp, Dirk Stöbener, Axel von Freyberg, Martina Fuhrmann und Gert Goch
Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft, BIMAQ, Universität Bremen

Untersuchung von Körperschallsignalen im Zeit-Frequenzbereich für ein körperschallbasiertes Motormanagement

[1]Marco Decker, [2]Karsten Hintz, [1]Clemens Gühmann

[1]TU-Berlin, FG Elektronische Mess- und Diagnosetechnik, 10587 Berlin, Deutschland, E-Mail: marco.decker@tu-berlin.de
[2]Otto-von-Guericke-Universität Magdeburg, Institut für Mobile Systeme (IMS), 39106 Magdeburg, Deutschland, E-Mail: karsten.hintz@ovgu.de

Kurzfassung

Mit Hilfe eines zylinderdruckbasierten Motormanagements werden eine zylinderselektive Optimierung, Regelung und Überwachung des Heizverlaufs möglich. Das Ziel hierbei ist es, die resultierenden Emissionen sowie deren Streubreite zu verringern. Im Beitrag werden Zusammenhänge zwischen Körperschall- und Zylinderdrucksignalen mit Hilfe der Kohärenz und der Wigner-Ville-Transformation aufgezeigt. Mit der anschließenden Merkmalsextraktion ist es möglich, Merkmale der Verbrennung und der Einspritzung zu extrahieren, welche im Motormanagement eingesetzt werden können. Damit ist es möglich, die kostenintensive Zylinderdrucksensorik durch günstigere und robustere Körperschallsensoren zu ersetzen. Abschließend werden erste Ergebnisse der Verbrennungsschwerpunktregelung vorgestellt und ein Ausblick auf weiterführende Arbeiten gegeben.

1 Einleitung

Zukünftige Abgasnormen erfordern zielgerichtete Maßnahmen, um die Abgasemissionen zu senken. Bei allen Maßnahmen sollen der Verbrauch und die Geräuschemissionen nicht ansteigen. Im Rahmen der Forschungsvereinigung Verbrennungskraftmaschinen (FVV) wird zu dieser Zielsetzung das For-

schungsprojekt „Geräuschgeregelter Dieselmotor“ unter Beteiligung des Fachgebietes Elektronische Mess- und Diagnosetechnk durchgeführt. Im Projekt soll geklärt werden, welche Verbesserungen sich durch die Integration von Körperschallsignalen in das Diesel-Motormanagementsystem bezüglich des Verbrauchs, des Geräuschs und der Abgasemissionen erzielen lassen. Zur Erreichung dieser Ziele ist eine detaillierte Signalanalyse notwendig. Hierbei wird untersucht, welche Informationen über die Verbrennung, den Zylinderdruck, die Einspritzung und das Motorgeräusch sich in einem Körperschallsignal befinden und wie diese extrahiert werden können.

Gelingt es durch Signalverarbeitung signifikante Merkmale über die Verbrennung bzw. den Zylinderdruck und die Geräuschemissionen des Motors unter Verwendung des Körperschalls zu bestimmen, kann ein geräuschbasiertes Motormanagementsystem realisiert werden, das ähnliche Vorteile wie ein zylinderdruckbasiertes Management besitzt, ohne die bezüglich Preis, Genauigkeit, Zuverlässigkeit und Lebensdauer kritischen Zylinderdrucksensoren einsetzen zu müssen (Bild 1).

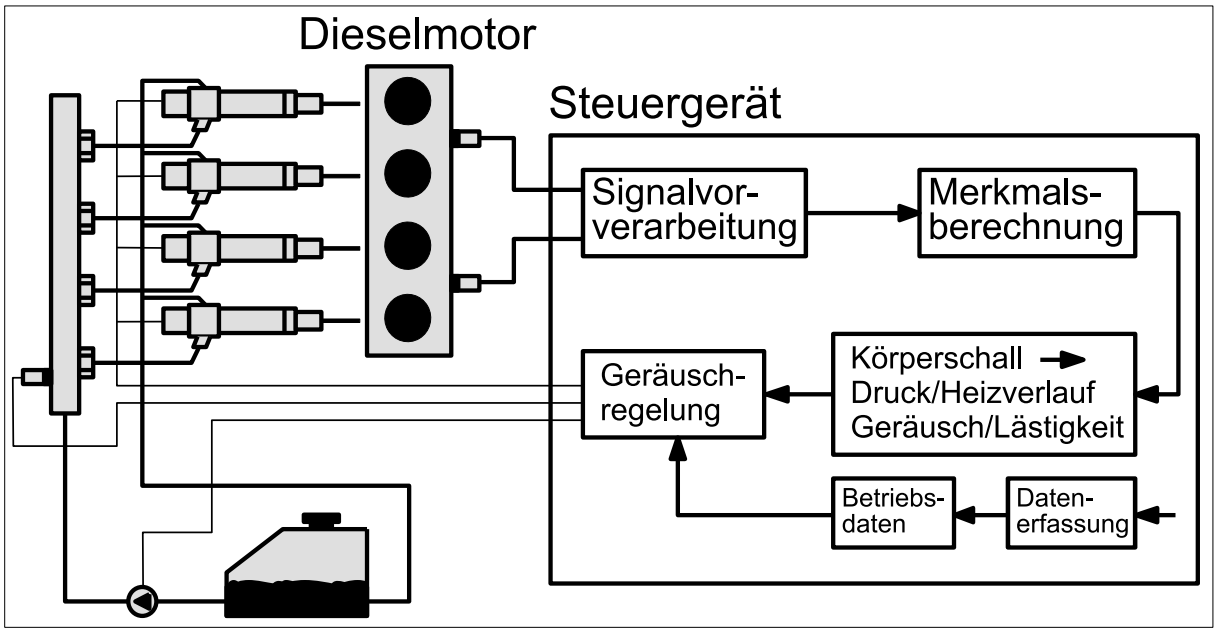


Bild 1: Prinzipschaltbild des geräuschgeregelten Dieselmotors mit der Integration der Geräuschkennwerte in die elektronische Steuerung eines Dieselmotors

In diesem Beitrag wird die Regelung der Verbrennungsschwerpunktlage auf der Basis der gemessenen Körperschallsignale dargestellt. Die Ansätze zur Bestim-

mung der Geräuschemissionen sind in [1] zu finden. Nach der Beschreibung des Versuchsaufbaus wird im Abschnitt 3 aufgezeigt, wie geeignete Körperschallsensorpositionen mit Hilfe der Kohärenz bestimmt werden können. Mit Hilfe der Smoothed-Pseudo-Wigner-Ville-Verteilung werden anschließend geeignete Merkmale zur Bestimmung der Verbrennungslage gefunden. Nach der Extraktion der Verbrennungslagemerkmale wird in Abschnitt 4 die Verbrennungsschwerpunktlagenregelung beschrieben. Anschließend erfolgt der Vergleich zwischen körperschall- und zylinderdruckbasierter Verbrennungsschwerpunktregelung. Abschließend werden die Ergebnisse zusammengefasst und der Ausblick auf weiterführende Arbeiten gegeben.

2 Versuchsaufbau

2.1 Sensorik und Messtechnik

Für die Versuche wurde auf dem Akustikprüfstand des Institutes für Mobile Systeme (IMS) der Universität Magdeburg ein aktueller Serien-4-Zylinder-Dieselmotor mit direkteinspritzenden Common-Rail-System aufgestellt und mit Messtechnik bestückt. Es wurden aus Voruntersuchungen 18 Positionen für Körperschallsensoren bestimmt, an denen ein gutes Übertragungsverhalten des Zylinderdrucks zu erwarten war. Diese Positionen am Zylinderkurbelgehäuse auf der Einlass- und Auslassseite sowie auf dem Motorlager an der Stirnseite und in der Nähe der Einspritzventile auf der Oberseite des Zylinderkopfes wurden für die weiteren Untersuchungen genutzt. In Bild 2 sind die Körperschallpositionen am Versuchsmotor abgebildet. Auf der Einlassseite (K 1-5), Auslassseite (H 1-6) und auf der Oberseite des Zylinderkopfes (O 1-4) befinden sich die eindimensionalen Beschleunigungsaufnehmer. Auf dem Motorlager wurde ein dreiachsiger Sensor angebracht (S 1-3).

Zur Erfassung der Zylinderdrücke und zur Erzeugung eines Referenzsignals für die Einspritzung wurde der Motor mit vier Zylinderdrucksensoren und einer Strommesszange am Injektor des Zylinders 1 ausgestattet.

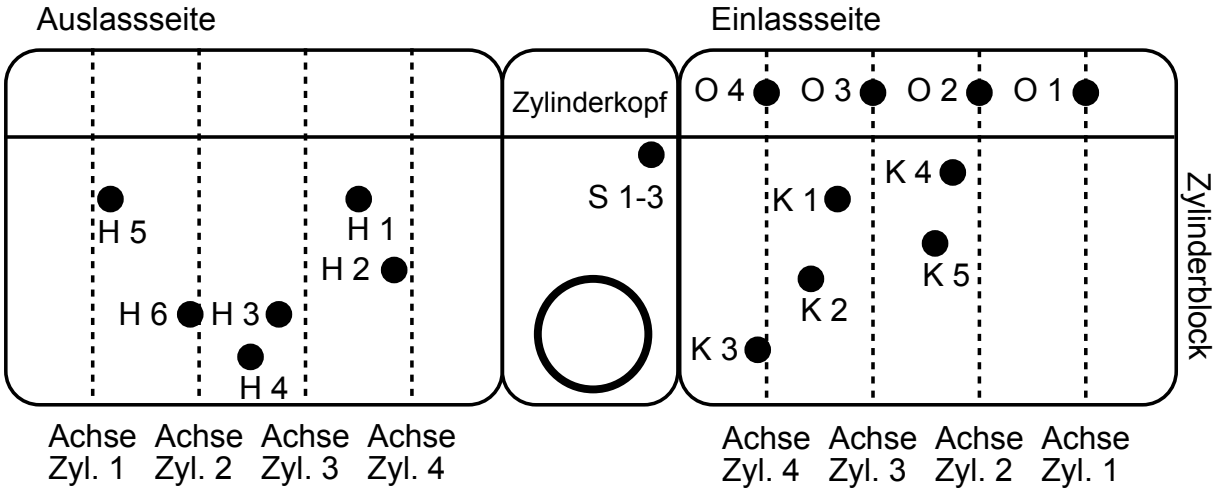


Bild 2: Übersicht der Beschleunigungsaufnehmer am Versuchsmotor

2.2 Zylinderdruckbasiertes Motormanagement

Um über eine größtmögliche Flexibilität bzgl. der Dieselmotorsteuerung zu verfügen und eine schnelle Integrierbarkeit neuer Algorithmen zu ermöglichen, wurde anschließend das mit MATLAB®/Simulink® programmierbare Rapid Prototyping System IAV-MPEC (Modular Prototyping Engine Controller) der IAV GmbH [2] für die Steuerung des Versuchsmotors in den Prüfstandsaufbau integriert.

Ein mit MATLAB®/Simulink® programmierbarer DSP führt zunächst die thermodynamische Auswertung der Zylinderdrucksignale und die Merkmalsberechnung der akustischen Signale in Echtzeit durch. Im Rapid-Prototyping-Steuergerät fließen die zylinderdruckbasierten Verbrennungsmerkmale bzw. akustischen Merkmale zyklusaufgelöst in die Regelkreise der Dieselmotorsteuerung ein. Der Zylinderdruckverlauf gilt allgemein als eine zentrale Größe zur Beschreibung innermotorischer Vorgänge. Mit Hilfe eines zylinderdruckbasierten Motormanagements werden eine zylinderselektive Optimierung, Regelung und Überwachung des Heizverlaufs möglich. Das Ziel hierbei ist es, die resultierenden Emissionen nochmals zu verringern. Bild 3 stellt die Struktur des zylinderdruckgeführten Motormanagements, wie es im MPEC-System umgesetzt wird, schematisch dar (IAV-AC3-Regelung).

Die Verbrennungsschwerpunktregelung dient dazu, den Heizverlauf zu stabilisieren. Anhand der berechneten Schwerpunktlage und einem vorgegebenen

Sollwert wird der Ansteuerbeginn der Einspritzventile zylinderselektiv eingeregelt. Für die Berechnung der Schwerpunktlage wurde im ersten Schritt der Heizverlauf nach Hohenberg ermittelt:

$$\Delta Q_H(\alpha) = \frac{c_V}{R}\left(p(\alpha)\cdot\Delta V(\alpha) + V(\alpha)\cdot\Delta p(\alpha)\right) + p(\alpha)\cdot\Delta V(\alpha), \quad (1)$$
$$\text{mit: } \frac{c_V}{R} = \frac{1}{\kappa-1}$$

Der Isotropenexponent wurde während der Berechnung konstant gehalten ($\kappa = 1,37$). Im nächsten Schritt erfolgt durch numerische Integration die Berechnung des integralen Heizverlaufs:

$$Q_H(\alpha) = \sum_{\alpha=-80°KW}^{\alpha=120°KW} \Delta Q_H(\alpha). \quad (2)$$

Abschließend wird die Verbrennungsschwerpunktlage wie folgt berechnet:

$$\alpha_{q50} = \frac{max\left(Q_H(\alpha)\right) - min\left(Q_H(\alpha)\right)}{2} + min\left(Q_H(\alpha)\right). \quad (3)$$

Die wesentliche Aufgabe der Geräuschregelung besteht darin, das Maximum des Zylinderdruckgradienten zu begrenzen bzw. zu reduzieren, was insbesondere der Geräuschemission zugutekommt. Der Eingriff des Geräuschreglers erfolgt über eine Korrektur des Sollwertes für die Verbrennungsschwerpunktregelung. Im Rahmen des Beitrags wird nur auf die Verbrennungsschwerpunktregelung eingegangen.

2.3 Versuchsdurchführung

Als Stellgröße für die AC3-Regelung dient der Ansteuerbeginn der Haupteinspritzung (MI_t). Mit Hilfe der Methode Design of Experiments (DoE) erfolgte eine statistische Versuchsplanung. Das Ziel hierbei ist es, mit einer vorgegebenen Anzahl an Messungen einen möglichst großen Versuchsraum zu beschreiben. Im Experiment wurden in 37 Messungen Drehzahl ($n = 1000...1500\ 1/min$), Last ($M = 50...200\ Nm$) und der Beginn der Haupteinspritzung ($MI_t = -8...8\ °KW$) variiert. Durch die Abdeckung eines großen Kennfeldbereiches ist es möglich, genaue Aussagen zwischen Merkmalen aus Zylinderdruck- und Körperschallsignalen zu treffen.

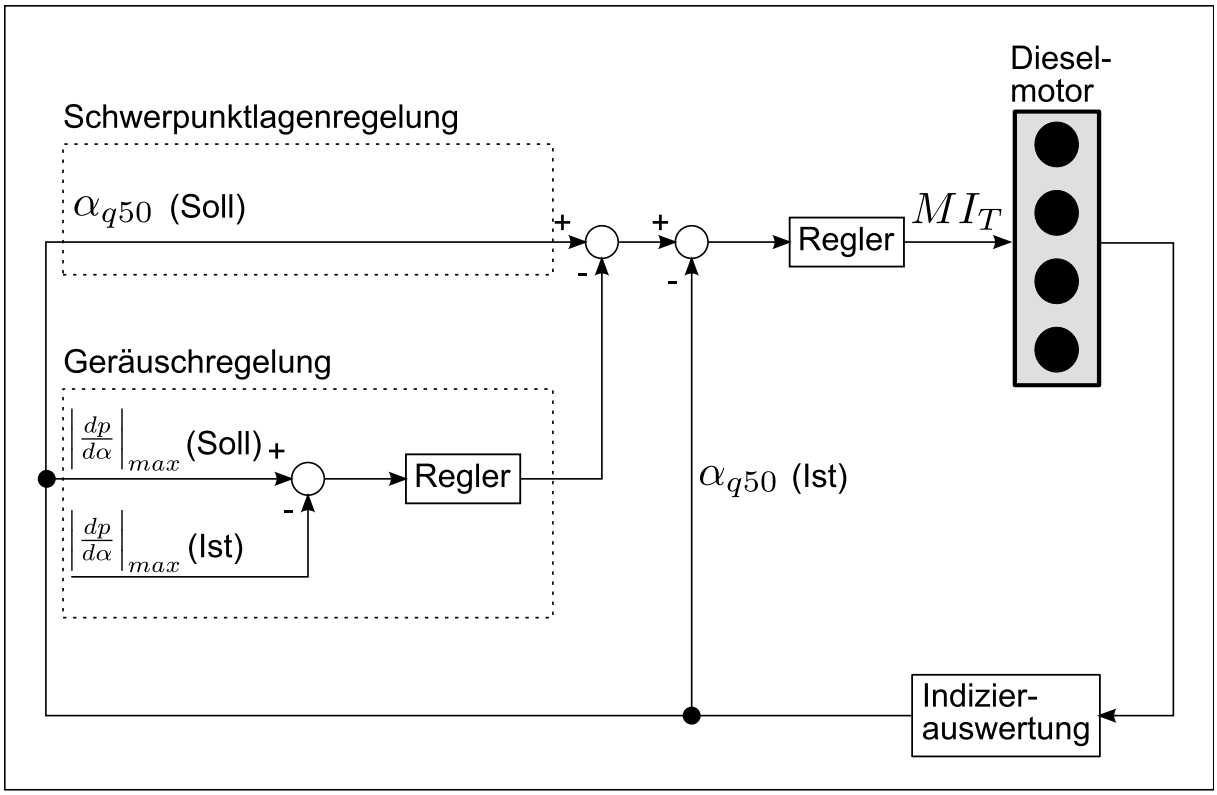


Bild 3: Schematische Darstellung der AC3-Regelung im zylinderdruckgeführten Motormanagement

3 Voruntersuchungen im Zeit- und Frequenzbereich

Das Ziel der Voruntersuchungen ist mit Hilfe der Kohärenzanalyse die besten Orte am Motor für die Sensoren zu bestimmen. Die Kohärenzanalyse zeigt Zusammenhänge zwischen zwei Signalen im Frequenzbereich auf:

$$\gamma^2(f) = \frac{|G_{KZ}(f)|^2}{G_{KK}(f) \cdot G_{ZZ}(f)}, \tag{4}$$

mit den Autoleistungsspektren vom Körperschall $(G_{KK}(f))$- und Zylinderdrucksignal $(G_{ZZ}(f))$, sowie dem Kreuzleistungsspektrum $(G_{KZ}(f))$ beider Signale. Der Wertebereich der Kohärenzanalyse liegt zwischen 0 und 1. Ein Wert von 1 bedeutet vollständige Abhängigkeit zwischen zwei Signalen und 0 zeigt keinen Zusammenhang auf. Um den Zusammenhang zwischen Körperschall- und Zylinderdrucksignalen bei Dieselmotoren aufzuzeigen, wurde die Kohärenzfunktion aus [3–5] genutzt. Die Kohärenz wurde in einem Kurbelwinkelfenster von $-90\ °KW$ bis $90\ °KW$, über die Anzahl der Arbeitsspiele einer Messung gemittelt. In [5] wird im Anschluss die mittlere Kohärenz über

einen definierten Frequenzbereich berechnet:

$$\gamma_M^2 = \frac{1}{k-j} \sum_{i=j}^{k} \gamma^2 \left(f \left(i \right) \right). \tag{5}$$

Aus jeder der insgesamt 37 Messungen wurde die gemittelte Kohärenz über ein Frequenzband von $200 - 2500\ Hz$ bestimmt. Diese Vorgehensweise erfolgte für alle 18 Beschleunigungsaufnehmer seperat für jeden Zylinder. Abschließend werden die Kohärenzen über die Anzahl der Messungen [6] gemittelt:

$$\overline{\gamma_M^2} = \frac{1}{n} \sum_{i=1}^{n=37} \gamma_M^2 \left(i \right). \tag{6}$$

In Bild 4 sind die Ergebnisse der Untersuchungen dargestellt. Die Sensoren auf der Auslassseite (H1-H6) und auf der Einlassseite (K1-K5) des Zylinderblocks weisen Kohärenzwerte zwischen 0,65 und 0,82 auf. Bei kurzer Distanz wischen den Zylindern und den Sensoren sind höhere Kohärenzwerte zu erzielen. Für die Sensoren auf dem Zylinderkopf in unmittelbarer Nähe zum Injektor ergaben sich die besten Werte. Hierbei zeigt sich, dass die Sensoren O1 und O2 jeweils sehr gute Werte für Zylinder 1 und Zylinder 2 aufweisen. Für die Zylinder 3 und 4 zeigen die Sensoren O3 und O4 die besten Ergebnisse auf. Der S-Sensor auf der Stirnseite auf dem Motorlager zeigt auch gute Kohärenzwerte für die Zylinder 3 und 4 auf. Die Untersuchungen zeigen insgesamt gute Tendenzen für die entsprechenden Sensorpositionen auf. Allerdings ist damit noch kein Verbrennungslagemerkmal aufgezeigt, welches sich aus dem Beschleunigungssignal bestimmen lässt. In der Literatur werden diese Zusammenhänge z.B. mit der Wigner-Ville-Verteilung aufgezeigt [7–10]:

$$W_{xx}(t,f) = \int_{-\infty}^{\infty} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau. \tag{7}$$

Für eine bessere Zeit- u. Frequenzdarstellung wird hier die Smoothed-Pseudo-Wigner-Ville-Verteilung (SPWV) verwendet:

$$SPWV_{xx}(t,f) = \int_{-\infty}^{\infty} g(t-t') \int_{-\infty}^{\infty} h\left(\tau\right) \cdot x\left(t' + \frac{\tau}{2}\right) x^*\left(t' - \frac{\tau}{2}\right) e^{-j\omega\tau} d\tau dt'. \tag{8}$$

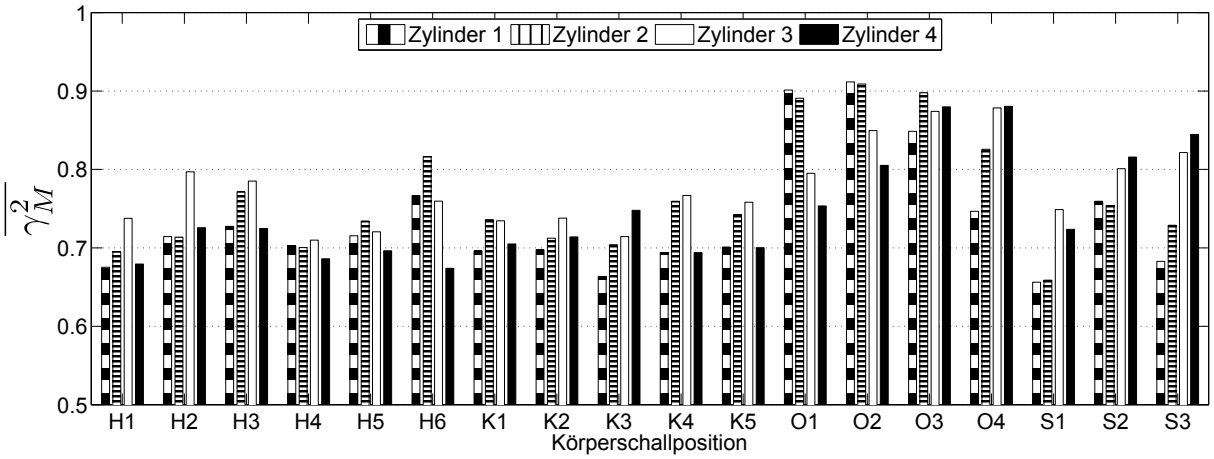


Bild 4: Gemittelte Kohärenz über 37 Messungen im Frequenzband von $200 - 2500\ Hz$ für alle Körperschallpositionen berechnet für jeden Zylinder

Die verbesserte Darstellung wird durch die Fensterfunktionen $h(\tau)$ im Frequenzbereich und durch $g(t)$ im Zeitbereich erreicht. Bild 5 zeigt die Auswertung von drei Messungen aus dem Design of Experiments mit einer starken Verbrennungsanregung. Die linken Bilder zeigen die Auswertungen für den Betriebspunkt $N = 1280\ \frac{1}{min}$, $M = 50\ Nm$, in der Mitte ist der Betriebspunkt $N = 1340\ \frac{1}{min}$, $M = 60\ Nm$ dargestellt und rechts der Betriebspunkt $N = 1500\ \frac{1}{min}$, $M = 100\ Nm$. Die Zeitsignale der Beschleunigungssensoren in den oberen Bildern wurden mit einem Rechteckfenster in Abhängigkeit vom Einspritzbeginn der Haupteinspritzung bewertet. Im Anschluss wurde vom gefensterten Beschleunigungssignal die Smoothed-Pseudo-Wigner-Ville-Verteilung berechnet (Bild Mitte). Die unteren Bilder stellen die Zeitsignale vom Zylinderdruck, Zylinderdruckgradienten und dem Heizverlauf dar. Dabei zeigen sich Zusammenhänge zwischen der Lage des maximalen Zylinderdruckgradienten sowie der Lage des maximalen Heizverlaufs. Diese korrelieren weiterhin mit dem unteren Frequenzbereich des Beschleunigungsaufnehmers O1. Mit den Prinzipuntersuchungen können die Frequenzbereiche der Nutzinformationen aufgezeigt werden. Um eine genaue Information über den korrelierten Frequenzbereich des Beschleungigungssignals zur Verbrennung zu treffen, wurden weiterführende Untersuchungen mit der SPWV unternommen [11]. Für jede Messung wird die SPWV:

$$SPWV_{xx}(n, \alpha, f), \text{ mit: } n = Messung \tag{9}$$

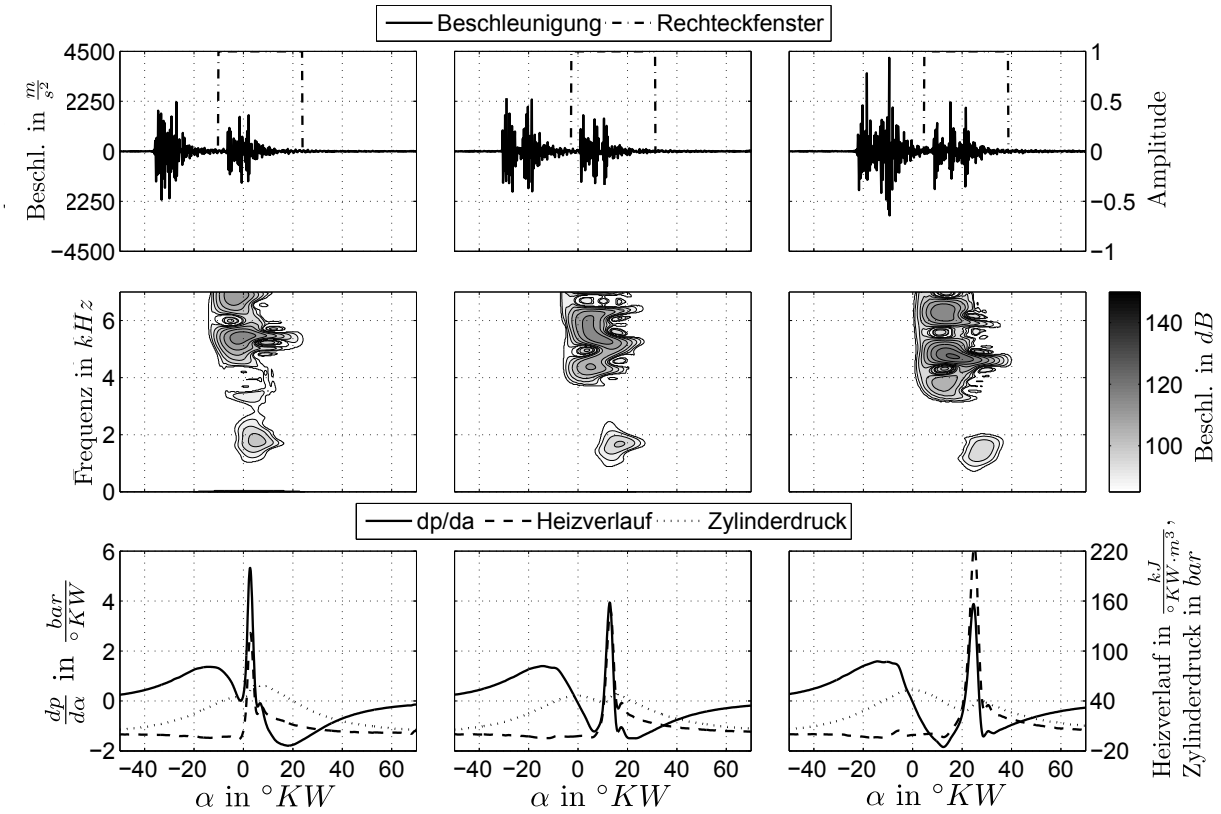


Bild 5: Zeit-Frequenzanalysen mit Hilfe der Smoothed-Pseudo-Wigner-Ville Verteilung, oben: Zeitsignal vom Beschleunigungssignal und Rechteckfenster, Mitte: SPWV vom Zylinderdruckgradienten, unten: Zeitsignale vom Zylinderdruck, Zylinderdruckgradienten und Heizverlauf links: Betriebspunkt $N = 1280\ \frac{1}{min}$, $M = 50\ Nm$, mitte: Betriebspunkt $N = 1340\ \frac{1}{min}$, $M = 60\ Nm$, rechts: Betriebspunkt $N = 1500\ \frac{1}{min}$, $M = 100\ Nm$

berechnet. Im nächsten Schritt wurde das Maximum für jede Frequenzstützstelle aus der SPWV ermittelt:

$$SPWV_{xx}^{max}(n, f) = \max_{-90<\alpha<90} SPWV_{xx}(n, \alpha, f). \tag{10}$$

Abschließend erfolgte die Berechnung des Korrelationskoeffizienten zwischen der Lage des maximalen Heizverlaufs und den Maxima der SPWV für jede Frequenzstützstelle. Somit kann der Korrelationskoeffizient als Funktion der Frequenz $r(f)$ dargestellt werden. In Bild 6 rechts ist der Korrelationskoeffizient über der Frequenz aufgetragen. Die höchsten Korrelationswerte liegen zwischen $f = 500...600\ Hz$. Das linke Bild zeigt den Korrelationsplot zwischen der Lage des maximalen Heizverlaufs und der Lage aus dem Maximum der SPWV des Beschleunigungssignals, an der Freqenzstützstelle $f \approx 500\ Hz$. Der Frequenzbereich zwischen $f = 250 - 1800\ Hz$ enthält die Nutzinformation über die

Verbrennung. Auch im Frequenzbereich zwischen $f = 5 - 10\ kHz$ sind hohe Korrelationskoeffizienten zu verzeichnen. Dieser Bereich weisst Informationen über die Einspritzung auf, die auch zu Verbrennungslagemerkmalen korrelieren. Die Untersuchungen beziehen sich auf den Beschleunigungsaufnehmer O1 und den Zylinder 1. Mit der vorgestellten Methode können Abhängigkeiten im Frequenzbereich zwischen Merkmalen aus Körperschall- und Zylinderdrucksignalen aufgezeigt werden.

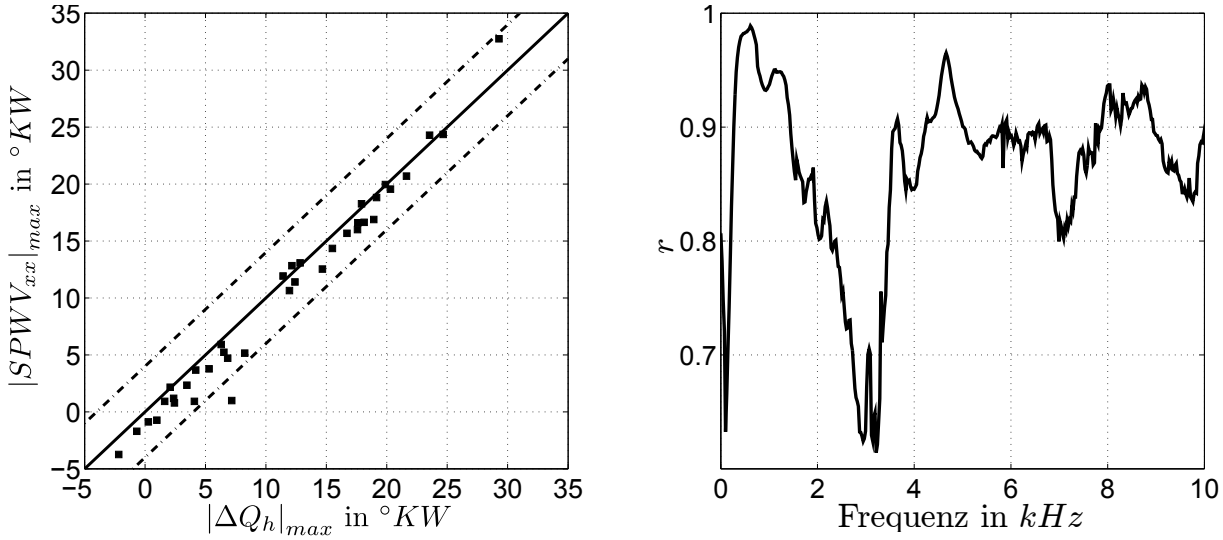


Bild 6: Darstellung des Korrelationskoeffizienten über der Frequenz (rechts), Korrelationsplot zwischen der Lage des maximalen Heizverlaufs und der Lage des Maximums aus der Smoothed-Pseudo-Wigner-Ville Verteilung bei $f \approx 600\ Hz$ (links)

4 Regelung der Verbrennungsschwerpunktlage

Nachdem der Frequenzbereich und die Sensorposition bestimmt wurden, in denen sich signifikante Merkmale über die Verbrennung abbilden, muss nun ein echtzeitfähiger Algorithmus zur Bestimmung der Merkmale implementiert werden. Die Grundlage des echtzeitfähigen Algorithmus zur Bestimmung der Verbrennungsschwerpunktlage bildet das mit einem Bandpass gefilterte und kurbelwinkelsynchron aufgezeichnete Zeitsignal des Beschleunigungssensors an der Oberseite des ersten Zylinders. Der zur Berechnung relevante Bereich des Zeitsignals, bezogen auf die Position in einem Arbeitsspiel, wird mithilfe eines Fensters herausgeschnitten. Die Fensterung erfolgt zeitlich nach der Hauptein-

spritzung und verschiebt sich in Abhängigkeit vom Einspritzbeginn. Innerhalb dieses Fensters werden die maximale Beschleunigungsamplitude und der zugehörige Kurbelwinkel $\alpha_{KS_{max}}$ ermittelt (Bild 7). Zur Detektion des Einspritzbeginns der Haupteinspritzung wird das mit einem Hochpass gefilterte Körperschallsignal verwendet. Anhand der Einhüllenden des Signals können neben dem Einspritzbeginn/ ende der Haupteinspritzung (α_{MI}) auch die Einspritzzeitpunkte der beiden Voreinspritzungen $(\alpha_{PI1}, \alpha_{PI2})$ bestimmt werden.

Mit Hilfe eines Modells konnte bei konstanten Lastpunkten bei variabler Drehzahl die Verbrennungsschwerpunktlage mit guten Ergebnissen modelliert werden [1, 12]:

$$\widehat{\alpha_{q50}} = a \cdot \alpha_{KS_{max}} + b \cdot n, \tag{11}$$

mit: $a, b-$ experimentell ermittelte Konstanten, $n -$ Drehzahl.

Die Implementierung der Verbrennungsschwerpunktregelung schloss sich an. Anhand der berechneten Schwerpunktlage und einem vorgegebenen Sollwert wird der Ansteuerbeginn der Einspritzventile eingeregelt. Die Untersuchungen des Körperschallmodells beziehen sich nur auf einen Zylinder, deshalb erfolgte die Regelung nicht zylinderselektiv.

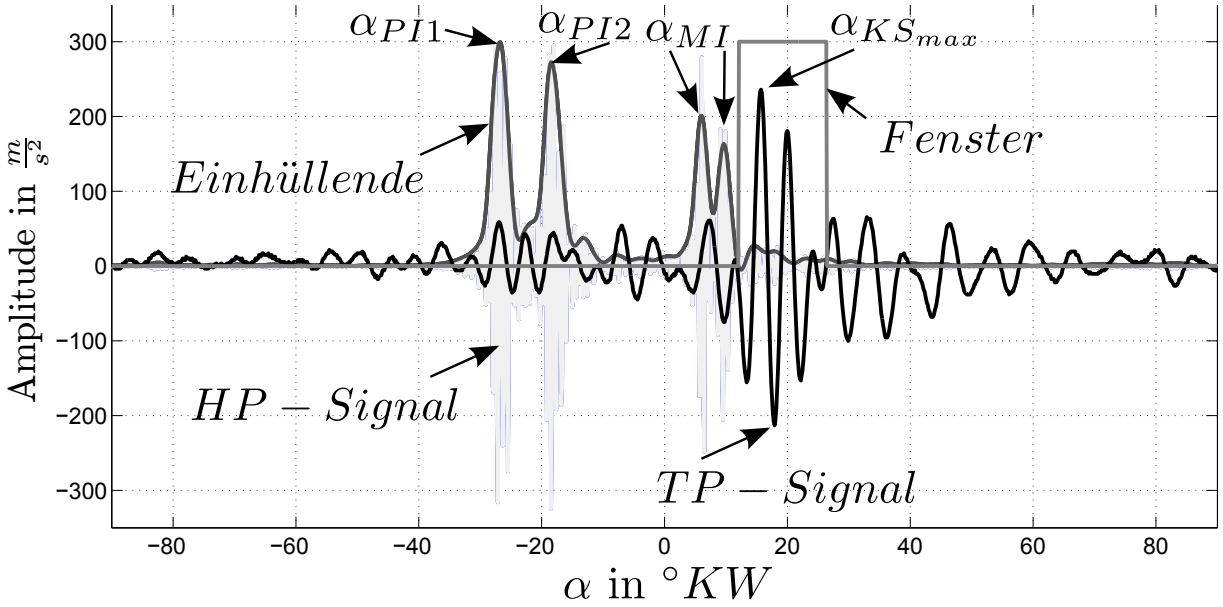


Bild 7: Zeitsignale des Beschleunigungssignals zur Bestimmung von $\alpha_{KS_{max}}$

In Bild 8 ist der Vergleich zwischen körperschall- (Bild 8 links) und zylinderdruckbasierter (Bild 8 rechts) Verbrennungsschwerpunktregelung aufgezeigt.

Im Rahmen der Untersuchungen wurde ein Sollwertsprung $\alpha_{q50} + 2\ ^\circ KW$ im Betriebspunkt $N = 1000\ \frac{1}{min}$, $M = 50\ Nm$ aufgeschaltet. Der Einspritzbeginn der Haupteinspritzung (Stellgröße) wird durch die Regelung in Richtung spät verschoben. Der Vergleich der Schwerpunktlagen zwischen Messungen aus Zylinderdrucksignalen und Körperschallmodell sowie dem vorgegebenen Sollwert ist in Bild 8 unten dargestellt. Auf der linken Seite erfolgte die Verbrennungsschwerpunktlagenregelung auf der Basis von Körperschallsignalen, auf der rechten Seite auf der Basis von Zylinderdrucksignalen. Die geregelten Schwerpunktlagen erreichen in beiden Fällen den Sollwert und zeigen auf, dass die Schwerpunktlage mit dem Einspritzbeginn geregelt bzw. korrigiert werden kann. Die Regelung konnte erfolgreich auf Basis von Körperschall- oder Zylinderdrucksignalen angewendet werden. Weitere Untersuchungen in einem erweiterten Betriebsbereich müssen die Allgemeingültigkeit des Ansatzes aufzeigen.

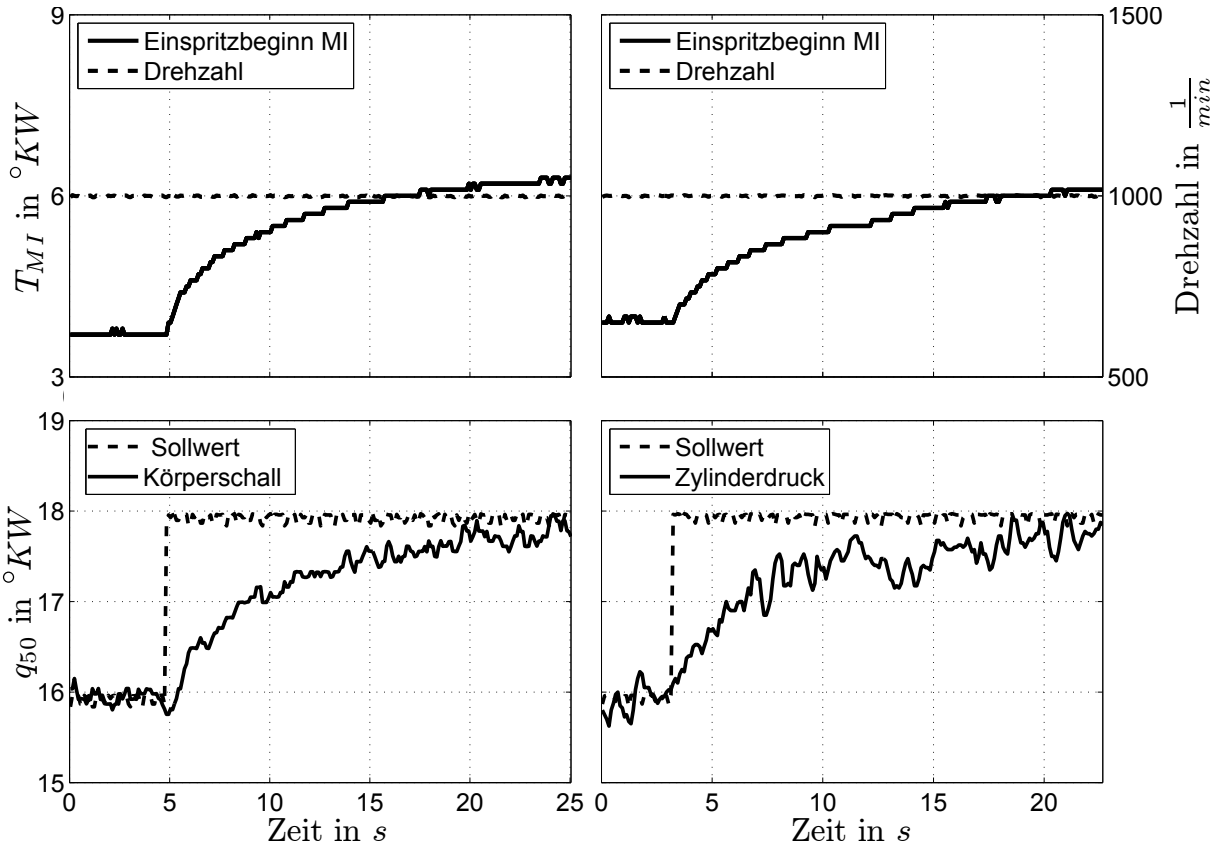


Bild 8: Regelung mit Sollwertsprung: $\alpha_{q50} + 2\ ^\circ KW$ im Betriebspunkt $N = 1000\ \frac{1}{min}$, $M = 50\ Nm$,links: Regelung mit Körperschallsignal, rechts: Regelung mit Zylinderdruck

5 Zusammenfassung und Ausblick

Im Beitrag wurde ein körperschallbasiertes Merkmalssystem zur echtzeitfähigen Bestimmung von Verbrennungsparametern vorgestellt. Es konnte mit Hilfe der Kohärenz aufgezeigt werden, welche Sensorpositionen für die Merkmalsbestimmung geeignet sind. In einer tiefergehenden Untersuchung der Sensorsignale mit Hilfe der Smoothed-Pseudo-Wigner-Ville-Verteilung konnten Zusammenhänge zwischen Zylinderdruckmerkmalen und Körperschallsignalen aufgezeigt werden. Insbesondere der für die Merkmalsbildung relevante Frequenzbereich wurde hiermit bestimmt. Weiterhin konnten die besten Sensorpositionen für die Merkmalsfindung gefunden werden. Nach der Umsetzung des Merkmalssystems konnte abschließend die Regelung in Betrieb genommen werden. Der Vergleich zwischen körperschall- und zylinderdruckgeführter Regelung der Verbrennungsschwerpunktlage zeigt für die ausgewählten Versuche keine signifikanten Unterschiede.

Im nächsten Schritt müssen die Merkmale im erweiterten Drehzahl- u. Teillastbereich untersucht und bewertet werden. Dabei ist eine Erweiterung des aktuellen Modells notwendig, um von der maximalen Heizverlaufsposition auf den Verbrennungsschwerpunkt zu schließen. Gelingt es für die anderen drei Zylinder die Verbrennungsschwerpunktlage aus Körperschallsignalen zu bestimmen, kann eine zylinderselektive Regelung der Schwerpunktlage vorgenommen werden. Dadurch können z. B. Alterungseffekte der Bauteile durch die Anpassung der Schwerpunktlage kompensiert werden. Weiterhin besteht die Möglichkeit, im Vergleich zur Serienapplikation die Streubreite der Abgastests sowie die Emissionen zu verringern. Abschließend muss das gesamte Motormanagementsystem bezüglich des Verbrauchs, des Geräuschs und der limitierten Abgasemissionen bewertet werden.

Literatur

[1] Decker, M., Hintz, K., Nobis, J., Joerres, M. und Gühmann, C., “Geräuschgeregelter Dieselmotor”, *Motorentechnische Zeitschrift (MTZ)* **3**, S. 226–233, 2011.

[2] Rempel, A., Stölting, E., Predelli, O. und Gratzke, R., "Flexible Motorprozessregelung für neue Brennverfahren", *at - Automatisierungstechnik* **57:1**, S. 14–22, 2009.

[3] Arnone, L., Manelli, S., Chiatti, G. und Chiavola, O., "In-Cylinder pressure analysis through accelerometer signal processing for diesel engine combustion optimization", in *SAE 2009 Noise and Vibration Conference and Exhibition*, (2009-01-2079), (St. Charles, Illinois, United States), 2009.

[4] Chiavola, O., Chiatti, G., Arnone, L. und Manelli, S., "Combustion characterization in diesel engine via block vibration analysis", in *SAE 2009 Noise and Vibration Conference and Exhibition*, (2010-01-0168), (St. Charles, Illinois, United States), 2010.

[5] Polonowski, C. J., Mathur, V. K., Naber, J. D. und Blough, J. R., "Accelerometer based sensing of combustion in a high speed HPCR diesel engine", in *SAE 2007 World Congress*, (2007-01-0972), (Detroit, Michigan, United States), 2007.

[6] Buono, S., "Rekonstruktion des Zylinderdruckverlaufs aus Körperschallsignalen und motorischen Kenngrößen basierend auf Künstlichen Neuronalen Netzen", Diplomarbeit, FG MDT, TU-Berlin.

[7] Carstens-Behrens, S., Wagner, M. und Boh J. F., "Improved knock detection by time variant filtered structure-borne sound", in *Proceedings of the Acoustics, Speech, and Signal Processing, 1999. IEEE International Conference - Volume 04*, *ICASSP '99*, S. 2255–2258, 1999.

[8] Lachmann, S., *Reduzierung des Einflusses von Störgeräuschen in Klopfregelsystemen mittels Blinder Quellentrennung*, dissertation.de, 2007.

[9] Chauvin, J., Grondin, O., Nguyen, E. und Guillemin, F., "Real-Time Combustion Parameters Estimation for HCCI-Diesel Engine Based on Knock Sensor Measurement", in *Proceedings of the 17th IFAC World Congress*, M. Chung (Hrsg.), S. 8501–8507, 2008.

[10] Grondin, O., Chauvin, J., Guillemin, F., Nguyen, E. und Corde, G., "Combustion parameters estimation and control using vibration signal : Application to the Diesel HCCI engine", in *47th IEEE Conference on Decision and Control*, S. 5621–5627, Dec. 2008.

[11] Pasch-Sannapiu, C., "Untersuchung von Körperschallsignalen im Zeit-Frequenzbereich sowie Merkmalsextraktion für ein Zylinderdruckbasiertes-Motormanagementsystem", Diplomarbeit, FG MDT, TU-Berlin.

[12] Decker, M., Gühmann, C., Lucas,S., Hintz, K., Tschöke, H., Nobis, J., Stölting, E. und Joerres, M., "Diesel engine management with center-of-heat-release control based on structure-borne sound signals", in *11th Stuttgart International Symposium*, **1**, S. 639–653, 2011.

Erfassung menschlicher Bewegungen durch merkmalsbasiertes Tracking

Kristine Back, Fernando Puente León

Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik,
Hertzstraße 16, 76187 Karlsruhe, E-Mail: {back, puente}@kit.edu

Kurzfassung

In diesem Beitrag wird ein Verfahren zur Erfassung des zeitlichen Verhaltens menschlicher Bewegungen vorgestellt, welches auf dem Verfolgen spärlicher Bewegungsmerkmale basiert. Mittels örtlicher und zeitlicher Filterung werden Regionen in Videos detektiert, an denen eine starke Bewegungsenergie vorliegt. Lokale Maxima dieser Bewegungsdetektoren werden mit dem Mean-Shift-Verfahren bei zusätzlicher Verwendung von Bildinformationen verfolgt. Es wird gezeigt, dass die erhaltenen zeitlichen Verläufe der Merkmalspunkte charakteristische Informationen enthalten, die eine Aktionserkennung ermöglichen.

1 Einleitung

Die videobasierte Erkennung und Analyse der Bewegungen von Menschen ist ein Forschungsbereich, der seit einigen Jahren immer größere Aufmerksamkeit auf sich zieht. Die Fähigkeit, menschliche Aktivitäten zu erfassen und zu verstehen, ist für intelligente technische Systeme, welche in menschlichen Umgebungen eingesetzt werden, von entscheidender Bedeutung.

Bei der visuellen Aktivitätsanalyse müssen zunächst Bildinformationen über die beobachtete Bewegung aus vorliegenden Sensordaten extrahiert und in einer bestimmten Form repräsentiert werden. Darauf basierend wird die Bewegung durch ein gewähltes Modell dargestellt. Dieses Modell wird dann in weiteren Verarbeitungsschritten für die anwendungsspezifische Aufgabe, beispielsweise

eine Aktivitätserkennung, verwendet. Je nach Anwendung existieren verschiedene Herangehensweisen von unterschiedlicher Komplexität in den einzelnen Verarbeitungsschritten.

Die Extraktion von Bewegungsinformationen aus Videodaten kann mit unterschiedlich ausgeprägten A-priori-Annahmen über die menschliche Anatomie und die Dynamik menschlicher Bewegungen vorgenommen werden. Es kann ein Modell des menschlichen Körpers erstellt und, basierend auf vorliegenden Bilddaten, die wahrscheinlichste Konfiguration des Modells geschätzt werden. Wird diese Posenschätzung im Laufe der Zeit verfolgt, erhält man detaillierte Informationen über den Bewegungsverlauf einzelner Körperteile. Solche Verfahren sind allerdings mit hohem Aufwand verbunden und weniger geeignet für Szenarien mit einer geringen Anzahl an Kameraansichten. Die Bewegungserfassung kann dagegen auch rein merkmalsbasiert erfolgen, z. B. indem man Silhouetten von Personen extrahiert und deren Variationen im Laufe der Zeit betrachtet. Ein solches Vorgehen liefert ebenso sehr umfassende Informationen, jedoch müssen auch hier starke Grundannahmen getroffen werden.

Im Bereich der Aktivitätserkennung, d. h. der Einordnung von Bewegungen in bestimmte Klassen, erfreuen sich seit einiger Zeit Methoden, die lediglich spärliche Merkmale verwenden, großer Beliebtheit. Dabei werden sog. Raum-Zeit-Merkmale aus Videodaten extrahiert, die aussagekräftige Informationen über die vorliegende Aktion enthalten. Zur Detektion solcher Merkmale existiert eine Vielzahl von Methoden. Häufig sind dies Erweiterungen von Detektoren für Bildmerkmale auf den dreidimensionalen Fall, wobei die Zeit die dritte Dimension darstellt. Die Stärke dieser Herangehensweise liegt darin, dass Merkmale nur im Bildbereichen detektiert werden, an denen eine Bewegung stattfindet und die somit eine hohe Aussagekraft besitzen. Solche Raum-Zeit-Merkmale werden weitläufig zur Aktivitätserkennung verwendet, oft ohne starke Berücksichtigung räumlicher und zeitlicher Zusammenhänge zwischen einzelnen Merkmalen („Bag of features“-Ansatz). Bei solchen Methoden zur Bewegungserfassung erhält man daher keine oder nur wenige Informationen über den dynamischen Verlauf von Bewegungen. In jüngerer Zeit kommen jedoch Methoden auf, die zeitliche Zusammenhänge spärlicher Merkmale verstärkt

berücksichtigen. Einen umfassenden Überblick über den Stand der Technik im Bereich der Aktivitätsanalyse kann man sich in beispielsweise in [10] und [9] verschaffen.

Um eine flexible und detaillierte Darstellungsform von Bewegungen zu erhalten, welche sich auf verschiedene Anwendungsbereiche übertragen lässt, ist die Erfassung des zeitlichen Bewegungsverlaufes wünschenswert. Hiervon können beispielsweise Verfahren zur zeitlichen Segmentierung oder Prädiktion von Bewegungen profitieren. In diesem Beitrag wird daher ein Verfahren zum Tracking von spärlichen Bewegungsmerkmalen vorgestellt. Die Motivation besteht darin, die Vorzüge modellbasierter Trackingverfahren und merkmalsbasierter Bewegungserfassung zu kombinieren. Dazu werden Detektoren für Raum-Zeit-Merkmale verwendet um charakteristische Punkte zu lokalisieren und über der Zeit zu verfolgen. Zum Tracking der Merkmalspunkte wird das Mean-Shift Verfahren herangezogen. Es resultieren Punktverläufe, die Informationen über die Bewegungsdynamik liefern.

In Abschnitt 2 werden zunächst die theoretischen Grundlagen erläutert, die zum Verständnis des Trackingverfahrens benötigt werden. In Abschnitt 3 wird die entwickelte Methode zur Bewegungserfassung dargestellt. Um die Aussagefähigkeit der Punktverläufe zu demonstrieren, werden eine Bewegungserkennung durchgeführt und Ergebnisse in Abschnitt 4 vorgestellt.

2 Theoretische Grundlagen

2.1 Das Mean-Shift-Verfahren

Das Mean-Shift-Verfahren [2, 4] ist eine kernel-basierte Methode zur Bestimmung lokaler Maxima einer Verteilungsdichte $f(\mathbf{x})$. Das Verfahren basiert auf der Verwendung von Kernel-Dichteschätzern. Dabei wird eine nichtparametrische Schätzung des Gradienten von f durchgeführt.

Die Kernel-Dichteschätzung von f in einem Punkt $\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^d$ bei n_h Beobachtungen $\{\mathbf{x}_i\}$, $i \in \{1 \dots n_h\}$ in einem Suchfenster $S_h(\mathbf{x})$ lautet

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n_h} k\left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) . \tag{1}$$

Dabei ist $k(x)$ das Profil des verwendeten Kernels $K(\mathbf{x})$, mit

$$K(\mathbf{x}) = c_{k,d}\, k(\|\mathbf{x}\|^2) \; . \tag{2}$$

Der Parameter h wird als Bandbreite bezeichnet und $c_{k,d}$ ist eine Normierungskonstante. Bei Verwendung eines differenzierbaren Kernels ergibt sich die Schätzung des Dichtegradienten durch den Gradienten des Dichteschätzers

$$\hat{\nabla} f_{h,K}(\mathbf{x}) = \nabla \hat{f}_{h,K}(\mathbf{x}) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^{n_h} (\mathbf{x} - \mathbf{x}_i) k' \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right) . \tag{3}$$

Wird eine neue Funktion

$$g(x) = -k'(x) \tag{4}$$

definiert, welche das Profil eines Kernels

$$G(\mathbf{x}) = c_{g,d}\, g(\|\mathbf{x}\|) \tag{5}$$

darstellt, ergibt sich

$$\hat{\nabla} f_{h,K} = \frac{2c_{k,d}}{h^2 c_{g,d}} \hat{f}_{h,G}(\mathbf{x})\, \mathbf{m}_{h,G}(\mathbf{x}) \; . \tag{6}$$

Dabei ist

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_h} g \left(\left\| \frac{\mathbf{x} - \mathbf{x}_i}{h} \right\|^2 \right)} - \mathbf{x} \tag{7}$$

der d-dimensionale Mean-Shift-Vektor. Dieser entspricht der Differenz zwischen dem Mittelpunkt $\mathbf{x}$ und dem gewichteten Mittel der Punkte $\mathbf{x}_i$ im Suchfenster $S_h(\mathbf{x})$. Der Mean-Shift-Vektor zeigt in Richtung des Dichtegradienten [2,4]. Er stellt somit einen nicht-parametrischen Gradientenschätzer dar und kann verwendet werden, um zu einem lokalen Maximum der geschätzten Verteilungsdichte aufzusteigen.

Beim Mean-Shift-Verfahren wird iterativ in Gradientenrichtung zu einem lokalen Maximum von $\hat{f}(\mathbf{x})$ fortgeschritten. Dabei wird in jedem Schritt der Mean-Shift-Vektor an der aktuellen Position berechnet und anschließend das Suchfenster um diesen Vektor verschoben. Die aufeinanderfolgenden Mittelpunkte des

Suchbereichs werden mit $\{\mathbf{y}_j\}$, $j \in \{0, 1, \ldots\}$ bezeichnet. Die Iterationsvorschrift lautet

$$\mathbf{y}_{j+1} = \mathbf{y}_j + \mathbf{m}_{h,G}(\mathbf{y}_j) = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i \, g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)} \, . \tag{8}$$

Wenn der Kernel K ein konvexes und monoton abnehmendes Profil besitzt, ist das Verfahren konvergent [4].

Ein häufig verwendeter Kern ist der Epanechnikov-Kern mit dem Profil

$$k_{\mathrm{E}}(x) = \begin{cases} 1 - x & 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases} \, . \tag{9}$$

Dabei ergibt sich für $g_{\mathrm{E}}(x)$ das Einheitsprofil und Gleichung (8) vereinfacht sich zu [3]

$$\mathbf{y}_{j+1} = \frac{1}{n_h} \sum_{i=1}^{n_h} \mathbf{x}_i \, , \tag{10}$$

was dem Mittelwert der Punkte $\mathbf{x}_i$ in $S_h(\mathbf{y}_j)$ entspricht.

2.2 Tracking von Objekten mittels Mean-Shift

Das Mean-Shift-Verfahren kann zum Tracking von Objekten verwendet werden [5]. Das zu verfolgende Objekt wird dabei durch seine Wahrscheinlichkeitsdichte q in einem Merkmalsraum repräsentiert. Als Merkmale können beispielsweise Farben aus dem RGB-Raum verwendet werden. In einem neuen Frame wird ein Zielkandidat an einer Position $\mathbf{y}$ definiert, welcher durch seine positionsabhängige Dichte $p(\mathbf{y})$ dargestellt wird. Basierend darauf kann ein Ähnlichkeitsmaß zwischen diesen beiden Dichten definiert werden und die Ziellokalisierung entspricht der Maximierung dieses Ähnlichkeitsmaßes.

Die Wahrscheinlichkeitsdichten von Ziel und Zielkandidat müssen aus den vorliegenden Bilddaten geschätzt werden. Als Zielmodell wird ein Histogramm mit m Abschnitten verwendet:

$$\hat{\mathbf{q}} = \{\hat{q}_u\}, \; u \in \{1 \ldots m\} \, , \quad \text{mit} \quad \sum_{u=1}^{m} \hat{q}_u = 1. \tag{11}$$

Entsprechend lautet das Kandidatenmodell an der Position $\mathbf{y}$:

$$\hat{\mathbf{p}}(\mathbf{y}) = \{\hat{p}_u(\mathbf{y})\},\ u \in \{1 \dots m\}, \quad \text{mit} \quad \sum_{u=1}^{m} \hat{p}_u = 1. \tag{12}$$

Die Modelle werden in elliptischen Regionen geschätzt, welche auf den Einheitskreis normalisiert werden. Für die Schätzung des Zielmodells ergibt sich [6]

$$\hat{q}_u = C \sum_{i=1}^{n} k(||\mathbf{x}_i||^2)\, \delta(b(\mathbf{x}_i) - u), \tag{13}$$

wobei $\{\mathbf{x}_i\}$, $i \in \{1 \dots n\}$ die normalisierten Koordinaten der Bildpunkte im Bereich des Zielmodells sind. $\delta(u)$ bezeichnet die Delta-Funktion und C eine Normierungskonstante für $\sum_{u=1}^{m} \hat{q}_u = 1$. Durch Verwendung des Kernels K werden Bildpunkte am Rande des Objektbereichs schwächer gewichtet. Entsprechend lautet das Kandidatenmodell

$$\hat{p}_u(\mathbf{y}) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) \delta(b(\mathbf{x}_i) - u) \tag{14}$$

mit den normalisierten Koordinaten $\{\mathbf{x}_i\}$, $i \in \{1 \dots n_h\}$ der Bildpunkte im Suchfenster $S_h(\mathbf{y})$.

Gesucht ist die Position $\mathbf{y}$, an der die beiden Modelle am ähnlichsten sind. Als Ähnlichkeitsmaß wird der Bhattacharyya-Koeffizient verwendet. Dessen diskrete Schätzung aus den Modellhistogrammen lautet

$$\hat{\rho}(\mathbf{y}) \equiv \rho\left[\hat{\mathbf{p}}(\mathbf{y}), \hat{\mathbf{q}}\right] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(\mathbf{y}) \hat{q}_u}\ . \tag{15}$$

Zur Maximierung des Bhattacharyya-Koeffizienten wird das Mean-Shift Verfahren angewandt um eine Gradientenschätzung $\hat{\nabla}\rho(\mathbf{y})$ zu erhalten. Daraus resultiert die Iterationsvorschrift für die Berechnung der neuen Zielposition $\mathbf{y}_{j+1}$

$$\hat{\mathbf{y}}_{j+1} = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i\, w_{ij}\, g\left(\left\|\frac{\hat{\mathbf{y}}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_{ij}\, g\left(\left\|\frac{\hat{\mathbf{y}}_j - \mathbf{x}_i}{h}\right\|^2\right)} \tag{16}$$

mit den Gewichtungsfaktoren

$$w_{ij} = w_i(\hat{\mathbf{y}}_j) = \sum_{u=1}^{m} \delta(b(\mathbf{x}_i) - u) \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{\mathbf{y}}_j)}} \ . \tag{17}$$

Der genaue Ablauf des Verfahrens sowie Details zur Implementierung und mögliche Erweiterungen können in [5] und [6] nachgelesen werden.

3 Bewegungserfassung durch bildbasiertes Tracking

Nun wird ein Verfahren zur Erfassung menschlicher Bewegungen durch bildbasiertes Tracking vorgestellt. Dies wird erreicht durch das Verfolgen einzelner Punkte am menschlichen Körper, die charakteristische Informationen über den Bewegungsverlauf liefern. Durch eine dreidimensionale Filterung werden Bildpunkte detektiert, an denen eine hohe Bewegungsenergie vorliegt und diese Punkte werden dann mit Hilfe der Verfahren aus Abschnitt 2 verfolgt.

3.1 Bewegungsdetektion

Zur Detektion von Bewegungsmerkmalen in einer Bildfolge wird die Methode aus [7] angewandt. Es wird eine Filterung einer Bildfolge $i(\mathbf{x}, t)$ in Orts- und Zeitrichtung durchgeführt. Als Filterantwort ergibt sich

$$\begin{aligned} r_{\sigma,\tau}(\mathbf{x}, t) = &\left(i(\mathbf{x}, t) * g(\mathbf{x}; \sigma) * h_{\mathrm{ev}}(t; \tau)\right)^2 \\ &+ \left(i(\mathbf{x}, t) * g(\mathbf{x}; \sigma) * h_{\mathrm{od}}(t; \tau)\right)^2 \ . \end{aligned} \tag{18}$$

Dabei ist $g(\mathbf{x}; \sigma)$ ein örtliches Gauß-Filter, in Zeitrichtung werden eindimensionale Gabor-Filter eingesetzt:

$$\begin{aligned} h_{\mathrm{ev}} &= -\cos(8\pi t/\tau) \cdot \exp\left(-\frac{t^2}{\tau^2}\right) \\ h_{\mathrm{od}} &= -\sin(8\pi t/\tau) \cdot \exp\left(-\frac{t^2}{\tau^2}\right) \ . \end{aligned} \tag{19}$$

σ und τ bezeichnen den örtlichen bzw. zeitlichen Skalenfaktor. Lokale Maxima der Filterantwort $r_{\sigma,\tau}(\mathbf{x}, t)$ geben Bildbereiche an, an denen eine hohe Bewegungsenergie vorliegt. Im Folgenden sollen solche Merkmalspunkte über der Zeit verfolgt werden.

3.2 Verfolgen von Bewegungsmerkmalen

Die Methode zum Verfolgen lokaler Maxima der Filterantwort des Bewegungsdetektors läuft in zwei Schritten ab. Zu jedem Zeitpunkt wird zunächst das Mean-Shift-Verfahren entsprechend Abschnitt 2.1 angewandt, um – ausgehend von der Position eines Merkmalspunktes im vorhergehenden Zeitschritt – zu einem lokalen Maximum von $r_{\sigma,\tau}(\mathbf{x}, t)$ fortzuschreiten. Im zweiten Schritt wird das Schätzergebnis durch Hinzuziehen weiterer Informationen verfeinert. Dazu erfolgt die Maximierung des Bhattacharyya-Koeffizienten zwischen Ziel- und Kandidatenmodell entsprechend Abschnitt 2.2 in einer Umgebung um die neuen Positionen der Punkte.

Es werden N Punkte $\mathbf{y}^k(t=0)$, $k = 1 \ldots N$, an lokalen Maxima von $r_{\sigma,\tau}(\mathbf{x}, 0)$ initialisiert.

Seien $\hat{\mathbf{y}}^k(t-1)$ die Positionen der Punkte zur Zeit $t-1$. Zum Zeitpunkt t werden die Startwerte der Punkte zu

$$\hat{\mathbf{z}}_0^k(t) = \hat{\mathbf{y}}^k(t-1) \tag{20}$$

gesetzt. Als nächstes wird mittels Mean-Shift-Iterationen in Gradientenrichtung zu einem lokalen Maximum von $r_{\sigma,\tau}(\mathbf{x}, t)$ aufgestiegen. Im Iterationsschritt j werden in einem Suchfenster $S_{h_1}(\hat{\mathbf{z}}_j^k(t))$ jeweils n_1 Punkte $\{\mathbf{x}_i^k(t)\}$, $i \in \{1 \ldots n_1\}$ abhängig von der Bewegungsintensität $r_{\sigma,\tau}(\mathbf{x}, t)$ ausgewählt. Bei Verwendung des Epanechnikov-Kernels lautet die Iterationsgleichung

$$\hat{\mathbf{z}}_{j+1}^k(t) = \frac{1}{n_1} \sum_{i=1}^{n_1} \mathbf{x}_i^k(t) \ . \tag{21}$$

Der Konvergenzpunkt des Verfahrens wird mit $\hat{\mathbf{z}}_c^k(t)$ bezeichnet.

Im nächsten Schritt wird die Schätzung der Position durch Hinzuziehen zusätzlicher Informationen über die Bildbereiche um die Merkmalspunkte verfeinert. Dazu wird das visuelle Trackingverfahren aus Abschnitt 2.2 angewandt. Als Merkmale für die Zielmodelle $\hat{\mathbf{q}}^k(t)$ können Farbe, Grauwert und Bewegungsintensität $r_{\sigma,\tau}(\mathbf{x}, t)$ verwendet werden. In diesem Schritt wird eine Bandbreite h_2, $h_2 < h_1$, verwendet. Die Zielmodelle werden zu Beginn initialisiert und durch rekursives Update mit Hilfe eines Vergessensfaktors an Variationen angepasst. Ausgehend von $\hat{\mathbf{z}}_c^k(t)$ wird nun der Bhattacharyya-Koeffizient zwischen

Ziel- und Kandidatenmodell maximiert. Die Anfangswerte für die Iterationen werden zu

$$\hat{\mathbf{y}}_0^k(t) = \hat{\mathbf{z}}_c^k(t) \tag{22}$$

gesetzt. Die Gewichte werden entsprechend Gleichung (17) bestimmt. Die Iterationsvorschrift für die Positionsschätzung in diesem Schritt lautet schließlich:

$$\hat{\mathbf{y}}_{j+1}^k = \frac{\sum_{\mathbf{x}_i \in S_{h_2}(\hat{\mathbf{y}}_j^k)} \mathbf{x}_i \, w_{ij}}{\sum_{\mathbf{x}_i \in S_{h_2}(\hat{\mathbf{y}}_j^k)} w_{ij}} \tag{23}$$

Die resultierenden Zielpositionen zur Zeit t entsprechen den Konvergenzpunkten des Verfahrens nach (23):

$$\hat{\mathbf{y}}^k(t) = \hat{\mathbf{y}}_c^k(t) \ . \tag{24}$$

4 Ergebnisse

Nun soll gezeigt werden, dass die Punktverläufe charakteristische Informationen über die Dynamik der vorliegenden Aktion enthalten. Dazu wird die Eignung der Punktverläufe zur Bewegungserkennung untersucht. Als Merkmale für die Bewegungserkennung werden die Verläufe $\hat{\mathbf{y}}^k(t)$ sowie die Verläufe der Bewegungsenergie

$$\hat{e}^k(t) = \sum_{x_i \in S_{h_1}(\hat{\mathbf{y}}^k(t))} r_{\sigma,\tau}(\mathbf{x}_i, t) \tag{25}$$

herangezogen.

Die Evaluation des Verfahrens erfolgt hier am Beispiel des KTH-Datensatzes [8]. Es wurden $N = 100$ Punkte verfolgt. Als Objektmodelle $\hat{\mathbf{q}}$ bzw. $\hat{\mathbf{p}}(\mathbf{y})$ werden Histogramme der Grauwerte und der Filterantwort $r_{\sigma,\tau}(\mathbf{x}, t)$ verwendet. In Abbildung 1 sind Beispiele für die resultierenden Verläufe abgebildet.

Zur Aktionserkennung wird eine Methode ähnlich wie in [11] verwendet. Dabei werden im Trainingsschritt aus den Verläufen $\hat{\mathbf{y}}^k(t)$ und $\hat{e}^k(t)$ Prototypkurven für die unterschiedlichen Aktionen erzeugt. Die Erkennung einer beobachteten Aktion erfolgt durch Vergleich mit den Prototypen. In Abbildung 2 ist das Ergebnis der Aktionserkennung zu sehen. Die extrahierten Verläufe enthalten

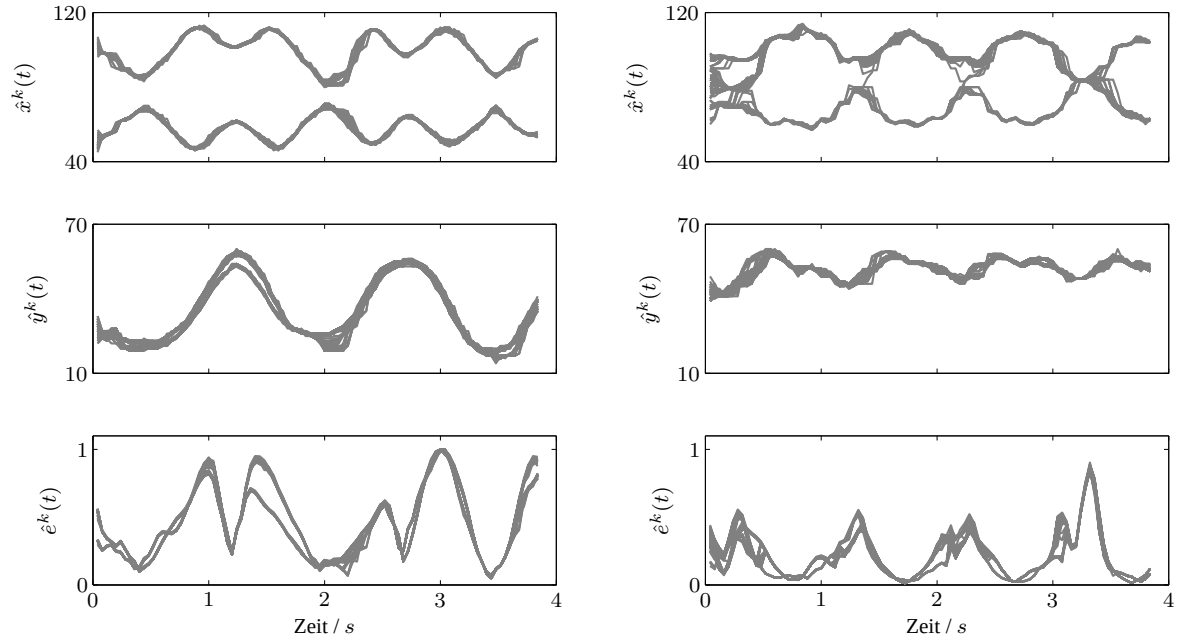


Bild 1: Beispiele der Verläufe $\hat{\mathbf{y}}^k(t)$ und $\hat{e}^k(t)$ für die Aktionen „Winken“ (a) und „Klatschen“ (b).

charakteristische Informationen über die vorliegende Aktion, sodass eine Erkennung möglich ist. Die erzielten Ergebnisse bei alleiniger Verwendung der dynamischen Bewegungsinformationen liegen in einem vergleichbaren Bereich wie beispielsweise [7], [8] und [1]. Es ist zu erwarten, dass eine Verbesserung der Erkennungsrate durch das Hinzuziehen weiterer Bildmerkmale erreicht werden kann.

5 Zusammenfassung und Ausblick

Es wurde ein Verfahren entwickelt, das den dynamischen Verlauf menschlicher Bewegungen erfasst, indem charakteristische Bewegungsmerkmale am menschlichen Körper verfolgt werden. Das Tracking der Bewegungsmerkmale kann mit Hilfe von effizienten Methoden realisiert werden. Es müssen keine starken Annahmen über das Erscheinungsbild von Personen oder die Umgebungsbedingungen getroffen und kein Zustandsvektor hoher Dimension geschätzt werden. Es wurde gezeigt, dass diese Verläufe Informationen bein-

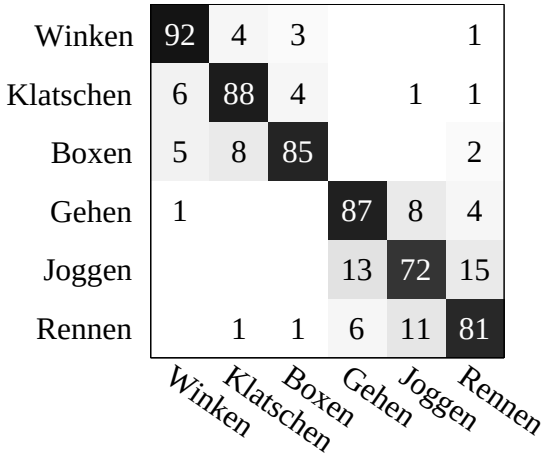


Bild 2: Ergebnis der Aktionserkennung anhand des KTH-Datensatzes in Prozent.

halten, die eine Bewegungsklassifikation erlauben.

Aspekte, die künftig behandelt werden, beinhalten die Integration von spezifischen A-priori-Dynamikmodellen in den Tracking-Algorithmus und die Verbesserung der Klassifikation durch Hinzunahme von zusätzlichen Informationen, wie z. B. Bildmerkmalen. Außerdem wird die Eignung der Trajektorien zur zeitlichen Segmentierung von Bewegungen untersucht.

Literatur

[1] M. Bregonzio, S. Gong, and T. Xiang. Recognising action as clouds of space-time interest points. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1948–1955, 2009.

[2] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:790–799, 1995.

[3] D. Comaniciu and P. Meer. Mean shift analysis and applications. *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2:1197–1203, 1999.

[4] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:603–619, 2002.

[5] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2:142–149, 2000.

[6] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:564–575, 2003.

[7] P. Dollár, V. Rabaud, G. Cottrell, and S. Belongie. Behavior recognition via sparse spatio-temporal features. pages 65–72, 2006.

[8] C. Schuldt, I. Laptev, and B. Caputo. Recognizing human actions: A local svm approach. *Proceedings of the 17th International Conference on Pattern Recognition*, 3:32–36, 2004.

[9] P. Turaga, R. Chellappa, V. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey. *IEEE Transactions on Circuits and Systems for Video Technology*, 18:1473–1488, 2008.

[10] D. Weinland, R. Ronfard, and E. Boyer. A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding*, 115:224–241, 2010.

[11] Y. Yacoob and M. Black. Parameterized modeling and recognition of activities. *Computer Vision and Image Understanding*, 73:232–247, 1999.

Magnetische Vektorhysterese: Messtechnik und Modellierung

Alexander Sutor, Jan Kallwies und Reinhard Lerch

Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl für Sensorik,
Paul-Gordan-Str. 3/5, D-91052 Erlangen, E-mail: sutor@lse.eei.uni-erlangen.de

Kurzfassung

Die Qualität numerischer Simulationen ist in entscheidender Weise von den verwendeten Materialmodellen und deren Parametern abhängig. Nur mit guten Modellbeschreibungen kann das Verhalten von Bauteilen, wie Sensoren und Aktoren, zuverlässig vorausberechnet werden. Werden magnetische Materialien sehr stark ausgesteuert, was bei einem effizienten Einsatz des Materials im Aktorbetrieb notwendig ist, so sind Hystereseerscheinungen nicht mehr zu vernachlässigen [1]. Auf der einen Seite benötigt man also Modellbeschreibungen für die verwendeten Materialien. Ein bekanntes numerisches Modell basiert auf dem Preisach-Operator [2]. Auf der anderen Seite sind Messplätze für die Parameterbestimmung notwendig.
In diesem Beitrag wird ein Messplatz zur vektoriellen Erfassung von magnetischer Hysterese vorgestellt, das sogenannte Vector-Vibrating-Sample-Magnetometer (VVSM). Die Messergebnisse werden mit anderen Messmethoden verglichen. Weiterhin wird ein neues, vektorielles Hysteresemodell vorgestellt, das auf dem Preisach-Hysteresemodell basiert. Die Verifizierung des Modells erfolgt durch einen Vergleich mit den VVSM-Messungen.

1 Einleitung

Am Lehrstuhl für Sensorik wurden in den letzten Jahren drei Magnetmessplätze mit unterschiedlichen Eigenschaften aufgebaut. Ein bekanntes Messprinzip ist der **Epstein-Rahmen**. Der komplette Kern besteht aus dem zu vermessenden Material. Daher kann aus dem eingeprägten Strom auf der Primärseite die ma-

gnetische Feldstärke direkt berechnet werden. Im Gegensatz dazu wird beim **Doppel-C-Joch** das Feld durch eine Streufluss-Spule gemessen. Das erhöht zwar den Einfluss von Wirbelströmen, bietet aber den Vorteil, dass auch kleinere Materialproben bzw. lokal begrenzte magnetische Eigenschaften erfasst werden können [3]. Im Gegensatz zu diesen beiden Messmethoden, die nach dem Übertragerprinzip arbeiten, wird mit dem **Vibration-Sample-Magnetometer (VSM)** die statische Magnetisierung einer kleinen Probe mit hoher Präzision über Bewegungsinduktion gemessen. Damit ist u.U. der Zeitaufwand für diese Messungen erheblich, es können aber sehr kleine magnetische Momente ohne die Verfälschung durch Wirbelstromeffekte erfasst werden. Eine entscheidende Erweiterung stellt schließlich das **Vector-VSM** dar, welches sowohl Betrag als auch Richtung der Magnetisierung bestimmt. Die Kombination dieser Messgeräte erlaubt die vollständige magnetische Charakterisierung unterschiedlichster Materialien bis hin zur Aufnahme vektorieller Hysteresekurven.

2 Das Vector-Vibration-Sample-Magnetometer (VVSM)

Vektor-Magnetometer werden in der Literatur behandelt [4], daher soll hier nur kurz auf die prinzipielle Funktionsweise und die besonderen Eigenschaften des von uns entwickelten Geräts eingegangen werden. Bei einem VSM wird zunächst ein geregeltes magnetisches Gleichfeld erzeugt. Die eigentliche Messeinheit ist ein System von Pickup-Spulen, die in unmittelbarer Nähe der Probe im Luftspalt angebracht sind. Die magnetisierte Probe erzeugt in ihrer Umgebung ein magnetisches Streufeld in der x-y-Ebene (siehe Bild 1), das sich mit dem wesentlich größeren Magnetfeld des Elektromagneten überlagert. Es ergibt sich eine gesamte magnetische Flussdichte im Luftspalt von

$$\vec{B} = \vec{B}_{\text{E-Magnet}} + \vec{B}_{\text{Probe}}, \tag{1}$$

die auch die Pickup-Spulen senkrecht durchsetzt. In den Bildern ist aus Gründen der Übersichtlichkeit lediglich das Streufeld der Probe eingezeichnet.
Wird die Probe zu einer mechanischen Schwingung in z-Richtung angeregt,

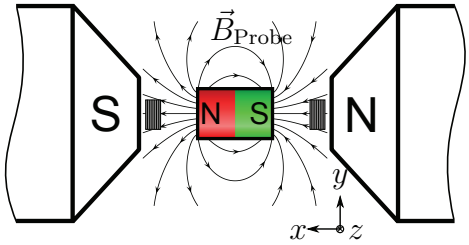


Bild 1: Prinzip eines VSM: Probe und Pickup-Spulen

wird in den Pickup-Spulen nach dem Induktionsgesetz

$$u_{\text{ind}}(t) = -\iint_A \frac{\partial \vec{B}}{\partial t}\,\mathrm{d}\vec{A} + \underbrace{\oint_C (\vec{v} \times \vec{B})\,\mathrm{d}\vec{s}}_{=0} \tag{2}$$

durch die zeitliche Änderung des magnetischen Flusses durch die Spulen eine Spannung induziert.

Da sich das gesamte $\vec{B}$-Feld im Luftspalt aus zwei Komponenten zusammen setzt (vgl. Gl. (1)), gilt:

$$\frac{\partial \vec{B}}{\partial t} = \underbrace{\frac{\partial \vec{B}_{\text{E-Magnet}}}{\partial t}}_{=0} + \frac{\partial \vec{B}_{\text{Probe}}}{\partial t} \tag{3}$$

Dabei ist zu beachten, dass $\vec{B}_{\text{E-Magnet}}$ über die Zeit konstant ist, die zeitliche Ableitung also 0 ist. Das statische Magnetfeld des Elektromagneten spielt also für das Wandlerprinzip eines VSM keine Rolle und ist deshalb in den Bildern 1 und 2 nicht eingezeichnet.

Der zweite Term in Gl. (2) mit $\vec{v} \times \vec{B}$ ist bei dem Funktionsprinzip eines VSM nicht von Bedeutung, da die Spulen selbst keine Bewegung ausführen, also $\vec{v} = 0$ ist. Somit vereinfacht sich Gl. (2) zu

$$u_{\text{ind}}(t) = -\iint_A \frac{\partial \vec{B}_{\text{Probe}}}{\partial t}\,\mathrm{d}\vec{A} \tag{4}$$

Da der Betrag der magnetischen Flussdichte $\vec{B}_{\text{Probe}}$ um die Probe von der Magnetisierung der Probe abhängt, ist die induzierte Spannung letztlich proportional zur Magnetisierung des Materials. Mit einer entsprechenden Kalibrierung des Messaufbaus kann so die Magnetisierung der Probe bestimmt werden.

2.1 Erweiterung zu einem Vector Vibrating Sample Magnetometer

Die Erweiterung eines VSM zu einem VVSM, das in der Lage ist die Magnetisierung vektoriell zu messen, besteht im Wesentlichen in der Verwendung von zwei weiteren Spulen-Paaren, die senkrecht zu den bereits vorhandenen angebracht werden [5].

Die Spulen zur Messung der Magnetisierung in Richtung des äußeren Magnetfeldes werden im Folgenden als Längs-Spulen (LS), die senkrecht dazu als Quer-Spulen (QS) bezeichnet, vgl. Bild 2.

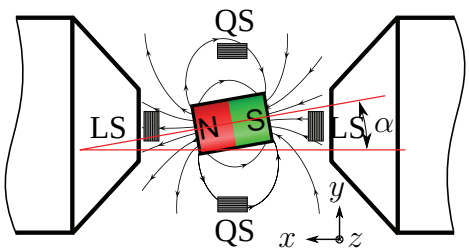


Bild 2: Prinzip eines VVSM mit einem gedrehten Magnetisierungsvektor

Angenommen, der Magnetisierungsvektor schließt mit den Feldlinien des H-Feldes den Winkel α ein (siehe Bild 2, so kann dieser Vektor in seine beiden Komponenten – $M_{\parallel}$ parallel zum Magnetfeld, sowie $M_{\perp}$ senkrecht dazu – zerlegt werden:

$$\vec{M} = \vec{M}_{\perp} + \vec{M}_{\parallel}. \tag{5}$$

Da die Pickup-Spulen in derselben Richtung[1] angeordnet sind, in der auch die Magnetisierung gemessen werden soll, ergeben sich die induzierten Spannung zu

$$u_{\text{ind,Längs}} \sim M_{\parallel} \quad \text{und} \quad u_{\text{ind,Quer}} \sim M_{\perp}. \tag{6}$$

Bild 3 zeigt den Aufbau im Luftspalt mit den vier Pickup-Spulen-Paaren, zwei Hallsensoren und der Probe. Die induzierten Spannungen werden mit Lock-In-Verstärkern *SR830* der Firma Stanford Research erfasst. Angetrieben wird die Probe durch einen Shaker *Tira Vib TV50018*. Die Betriebsfrequenz liegt bei 80 Hz. Der Shaker ist auf einem Drehtisch *M-037* der Firma PI montiert. Da der Messplatz vollständig automatisiert ist, bietet er die Möglichkeit beliebige zeitliche Abfolgen der Magnetfeldamplitude und der Orientierung der Probe

[1] Die Längs-Spulen liegen in der x-z-Ebene, messen also genau in x-Richtung. Die Quer-Spulen liegen in der y-z-Ebene, messen also genau in y-Richtung.

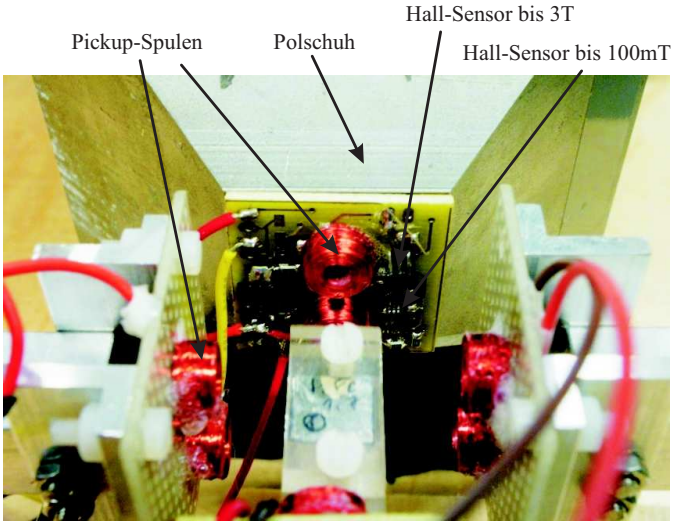


Bild 3: Platine auf einem Polschuh mit Pickup-Spulen und Magnetfeldsensoren

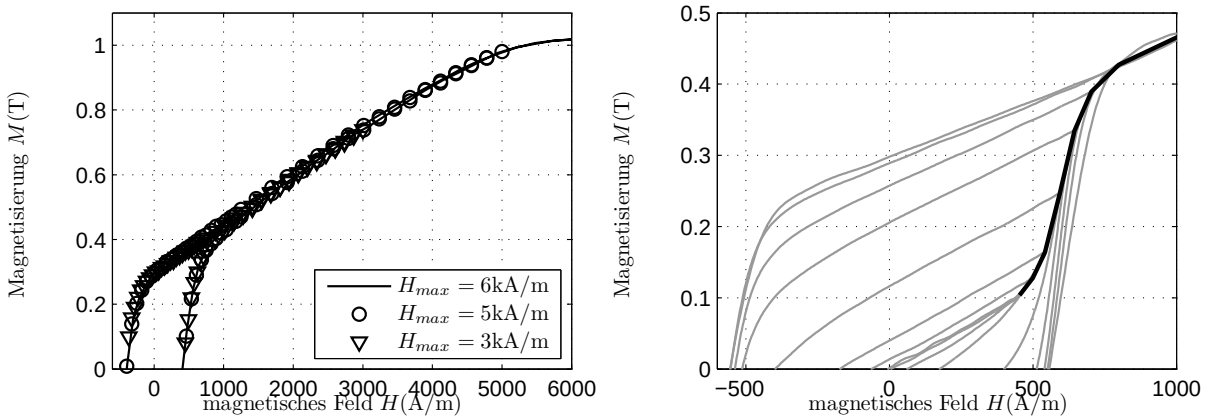


Bild 4: Familie von Hysteresekurven einer 1,7 µm gesputterten NiFe-Dünnschicht auf Silizium. Links: Messung mit dem Vibrating-Sample-Magnetometer, man beachte die gute Reproduzierbarkeit der Messungen. Rechts: Messung mit dem Doppel-C-Joch.

zu durchfahren. Insbesondere können die Magnetisierungsbewegungen in stark anisotropen Materialien wie dünnen magnetischen Schichten oder Transformatorblechen charakterisiert werden.

3 Messergebnisse

Bild 4 zeigt eine Familie von Hysteresekurven einer NiFe-Dünnschicht auf Silizium. Die Kurven in der linken Grafik sind mit dem VVSM bei Anregung mit unterschiedlicher H-Feld-Amplitude aufgenommen worden und stellen die Reproduzierbarkeit der VVSM-Messungen unter Beweis. Zum Vergleich sind in der rechten Grafik Kurven gezeigt, die mit dem Doppel-C-Joch gemessen

wurden. Gezeigt sind symmetrische Minor Loops und die Major Loop (grau) sowie die resultierende Kommutierungskurve (schwarz). Die Kurven stimmen gut überein, wobei zu beachten ist, dass eine Messung an einer Dünnschicht von 1,7 µm Dicke eine beachtliche Herausforderung für einen Doppel-C-Joch-Messplatz darstellt [3]. Unphysikalische Artefakte, wie sich berührende oder überlappende Kurven, resultieren daraus. Eine Darstellung vektorieller Messungen ist in Abschnitt 6 zu finden.

4 Vektor-Preisach-Modell

In der Literatur finden sich unterschiedliche Ansätze für skalare und vektorielle Hysteresemodelle [6–8]. Am Lehrstuhl für Sensorik werden verschiedene, auf dem Preisach-Operator basierende Modelle untersucht und eingesetzt [9, 10]. Im Folgenden wird zunächst der skalare Preisach-Operator skizziert und dann die vektorielle Erweiterung diskutiert.

4.1 Skalares Modell

Die allgemeine Formulierung des skalaren Preisach-Hysterese-Operators $\Gamma[\circ]$ lautet

$$f(t) = \Gamma[u](t) = \iint\limits_{\alpha \geq \beta} \mu(\alpha, \beta) \gamma_{\alpha\beta}[u](t) \, \mathrm{d}\alpha \, \mathrm{d}\beta \, , \tag{7}$$

wobei die Zeitsignale $u(t)$ und $f(t)$ den Eingang und den Ausgang bezeichnen. Der Definitionsbereich des Operators ist grundsätzlich endlich. Er wird hier festgelegt auf $u(t) \in [-0,5; 0,5]$. Die sogenannte Gewichtsfunktion $\mu(\alpha, \beta)$ bestimmt die Form der Hysteresekurve. Beispiele für einer analytischen Gewichtsfunktion zeigen die Bilder 5 und 6 für die in dieser Arbeit verwendeten Materialien FeCo und NiFe. Die Operatoren $\gamma_{\alpha\beta}[\circ]$ nehmen lediglich die diskreten Werte -1 und 1 an. Sie spiegeln das Gedächtnis des Hysteresoperators wider und werden oft als **Schaltoperatoren** bezeichnet. Das Integral wird über der zweidimensionalen, dreieckigen Preisach-Ebene ausgeführt, in der auch $\mu(\alpha, \beta)$ und $\gamma_{\alpha\beta}[\circ]$ definiert sind. Die Schaltoperatoren arbeiten dabei in Abhängigkeit des

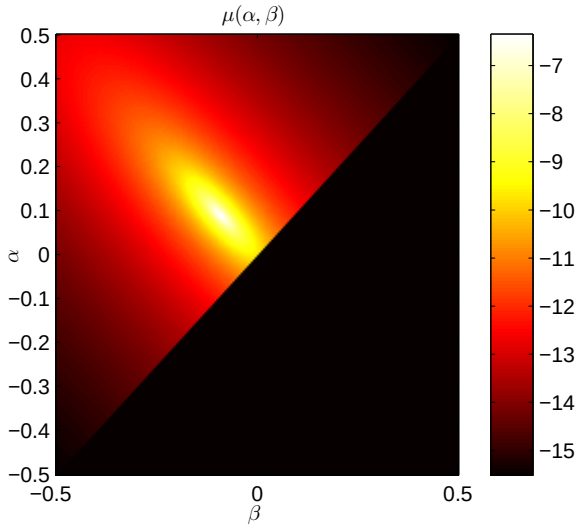


Bild 5: Besetzung der Preisach-Gewichtsfunktion von FeCo in logarithmischer Darstellung.

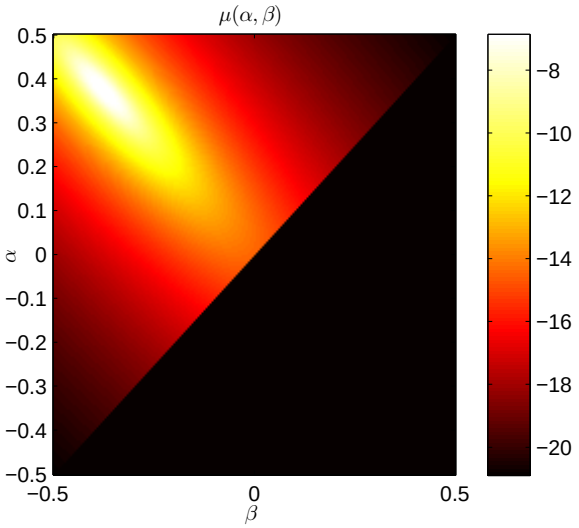


Bild 6: Als Vergleich die Preisach-Gewichtsfunktion von NiFe.

Eingangsverlaufs $u(t)$ nach folgendem Algorithmus

$$\gamma_{\alpha\beta}[u](t) = \begin{cases} +1 & : \quad u(t) > \alpha \\ \gamma_{\alpha\beta}[u](t^-) & : \quad \beta < u(t) < \alpha \\ -1 & : \quad u(t) < \beta\,, \end{cases} \tag{8}$$

siehe auch Bild 7.

4.2 Vektorielles Modell

Die jüngste Entwicklung ist ein effizientes, vektorielles Preisach-Modell $\vec{\Gamma}[\circ]$

$$\vec{v}(t) = \vec{\Gamma}[\vec{u}](t) = \iint\limits_{\alpha\geq\beta} \mu(\alpha,\beta)\cdot\gamma_{\alpha\beta}[\vec{u}](t)\cdot\vec{\chi}_{\alpha\beta}[\vec{u}](t)\,\mathrm{d}\alpha\,\mathrm{d}\beta\,. \tag{9}$$

$\vec{u}(t)$ und $\vec{v}(t)$ sind nun die Eingangs- und Ausgangssignale als vektorielle Zeitverläufe. Der Definitionsbereich ist auch hier beschränkt auf $\|u(t)\| \in [-0,5;0,5]$. Die Gewichtsfunktion $\mu(\alpha,\beta)$ beschreibt weiterhin die Form der Hysteresekurve. Den klassischen skalaren Schaltoperator $\gamma_{\alpha\beta}[\circ]$ nennen wir nun **Klappoperator** und ergänzen ihn mit dem vektoriellen Schaltoperator $\vec{\chi}_{\alpha\beta}$, der im Folgenden als **Drehoperator** bezeichnet wird. Als Schaltzustände sind für $\vec{\chi}[\circ]$ die Einheitsvektoren im $\mathbb{R}^2$ oder $\mathbb{R}^3$ definiert ($\|\vec{\chi}_{\alpha\beta}[\circ]\| = 1$). Eine weitere Unterscheidung zwischen zwei- oder drei dimensionaler Modellierung ist im

Algorithmus nicht notwendig.
Das Integral in Gl. (9) ist nicht als vektorielles Oberflächenintegral zu verstehen. man könnte es mit $\vec{v} = (v_i)$ und $\vec{\chi}_{\alpha\beta}[\circ] = (\chi_{\alpha\beta,i}[\circ])$ auch komponentenweise schreiben

$$v_i(t) = \iint\limits_{\alpha\geq\beta} \mu(\alpha,\beta)\cdot\gamma_{\alpha\beta}[\vec{u}](t)\cdot\chi_{\alpha\beta,i}[\vec{u}](t)\mathrm{d}\alpha\mathrm{d}\beta\ . \tag{10}$$

4.2.1 Das Schaltverhalten des Klappoperators $\gamma_{\alpha\beta}[\circ]$

Beim skalaren Modell wird als Argument für den Schaltoperator die (vorzeichenbehaftete) Eingangsgröße verwendet. Beim Vektormodell wird die Projektion des Eingangsvektors auf das Ergebnis des Winkeloperators $\vec{\chi}_{\alpha\beta}[\circ]$ als Argument verwendet. Als physikalische Motivation gilt die Annahme, dass nur dieser projizierte Anteil $u_{||}(t)$ für das Klappen der Domänen verantwortlich ist, der Betrag und die Richtung des Eingangsvektors eher für die Drehung

$$\gamma_{\alpha\beta}[u](t) = \begin{cases} +1 & : \quad u_{||}(t) > \alpha \\ \gamma_{\alpha\beta}[u_{||}](t^-) & : \quad \beta < u_{||}(t) < \alpha \\ -1 & : \quad u_{||}(t) < \beta\,, \end{cases} \tag{11}$$

Mathematisch wird diese Projektion durch das Skalarprodukt (•) ausgedrückt

$$u_{||}(t) = \vec{u}(t)\bullet\vec{\chi}_{\alpha\beta}[\vec{u}](t) = \|\vec{u}(t)\|\cdot\cos\left[\angle\left(\vec{u}(t),\ \vec{\chi}_{\alpha\beta}[\vec{u}](t)\right)\right]\ . \tag{12}$$

4.2.2 Das Schaltverhalten des Drehoperators $\vec{\chi}_{\alpha\beta}[\circ]$

Das Schaltverhalten des Drehoperators wird symmetrisch modelliert, d.h. die Gerade $\alpha = -\beta$ ist eine Symmetrieachse der $\vec{\chi}_{\alpha\beta}$-Ebene, siehe auch Bild 8. Das ist gerechtfertigt, da das Vorzeichen im Klappoperator enthalten ist. Die Multiplikation der beiden Operatoren in Gl. (9) ermöglicht sozusagen eine parallele oder antiparallele Ausrichtung zum Eingangssignal $\vec{u}(t)$. Die Schaltvorschrift für den Drehoperator lautet also

$$\vec{\chi}_{\alpha\beta}[u](t) = \begin{cases} \vec{e}_u(t) & : \quad u_\chi(t) > \alpha \vee -u_\chi(t) < \beta \\ \vec{\chi}_{\alpha\beta}[u](t^-) & : \quad \beta < -u_\chi(t) \wedge u_\chi(t) < \alpha \end{cases} \tag{13}$$

mit der Schaltschwelle $u_\chi(t)$

$$u_\chi(t) = \frac{1}{2}\left(2\|\vec{u}(t)\|\right)^k\ . \tag{14}$$

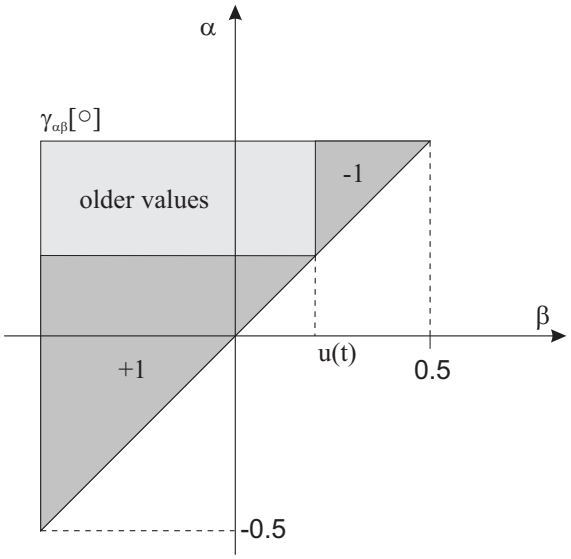


Bild 7: Verdeutlichung der Schaltvorgänge von $\gamma_{\alpha\beta}[u](t)$.

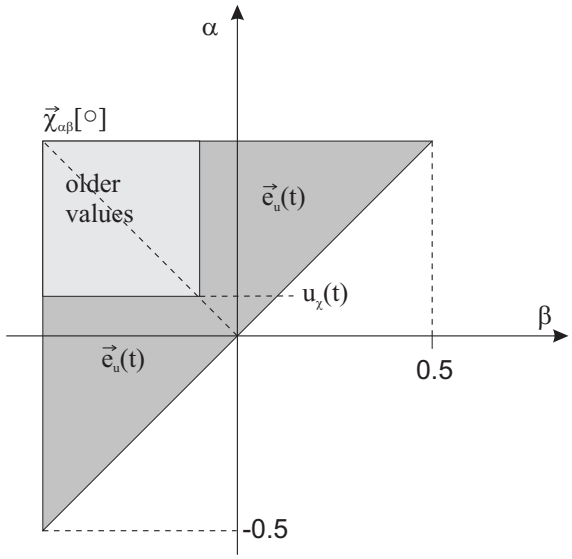


Bild 8: Besetzung der χ-Ebene.

Hier bezeichnet $\vec{e}_u(t)$ den Einheitsvektor in Richtung von $\vec{u}(t)$

$$\vec{e}_u(t) = \frac{\vec{u}(t)}{\|\vec{u}(t)\|} \, . \tag{15}$$

Der Parameter k ist der einzige (bis dato) freie Parameter des Vektormodells. Er beschreibt die Geschwindigkeit der Drehprozesse im Vergleich zu den Klapp-Prozessen. Für $k = 1$ sind die Schaltschwellen für beide Prozesse vom Betrag her gleich. $k > 1$ resultiert in einer langsameren (späteren) Drehung, was bei vielen Materialien der Fall ist. Siehe dazu Bild 10. Wir bezeichnen k als *Drehwiderstand*.

4.2.3 Eigenschaften des vektoriellen Preisach-Operators

Beim vektoriellen Preisach-Operator handelt es sich um eine Erweiterung des skalaren Preisach-Operators. Wendet man den vektoriellen Operator auf eindimensionale Signale an, so verhält er sich exakt wie der skalare Operator. Die Eigenschaften des skalaren Operators wie Kongruenz und Wipe-Out-Rule [1] bleiben also erhalten.
Ein weiterer großer Vorteil dieser Verallgemeinerung ist, dass Entwicklungen am klassischen Operator, wie die Modellierung von Temperatur- oder Frequenzverhalten, auf den Vektor-Operator übertragen werden können.

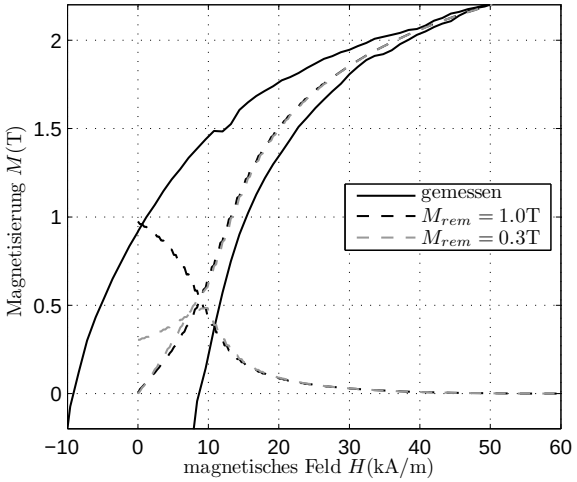


Bild 9: Simulationsergebnisse bei Variation der anfänglichen Remanenz, verglichen mit einer gemessenen äußeren Hysteresekurve. Material: FeCo

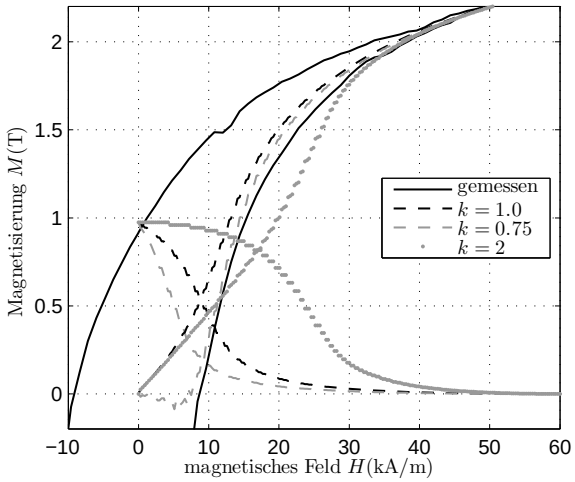


Bild 10: Simulationsergebnisse bei Variation des Drehwiderstands k.

5 Simulationen mit dem Vektormodel

Zunächst sollen Plausibilitätstests mit dem Modell durchgeführt werden. Es wird jeweils eine bestimmte Remanenzmagnetisierung M_{rem} eingestellt, dann wird senkrecht dazu das H-Feld von Null weg auf einen Maximalwert erhöht. Die senkrechte Komponente der Magnetisierung muss bei diesem Vorgang zu Null werden, die parallele Komponente muss sich dem Verlauf der skalaren, äußeren Hysteresekurve annähern.

Die Bilder 9 und 10 demonstrieren dies für die gesputterte FeCo-Dünnschicht. In Bild 9 sind Verläufe für unterschiedliche Anfangsmagnetisierungen zu sehen. In Bild 10 wird der Einfluss des Drehwiderstands k illustriert. Bei einem größeren Drehwiderstand können auch Verläufe außerhalb der äußeren (skalaren) Hysteresekurve auftreten. Es ist jeweils ein Vergleich mit der gemessenen Kurve gezeigt. Die Bilder 11 und 12 enthalten dieselbe Information für die gesputterte NiFe-Schicht.

6 Verifikation des Modells durch Messungen

Vektorielle Hysteresekurven sind in Bild 13 dargestellt, wobei die Magnetisierung in Komponenten senkrecht und parallel zum magnetischen Feld H aufgeteilt ist. Um die Darstellungen vergleichbar zu machen, wurde stets eine Remanenzmagnetisierung senkrecht zum anzulegenden Feld erzeugt, dann das Feld

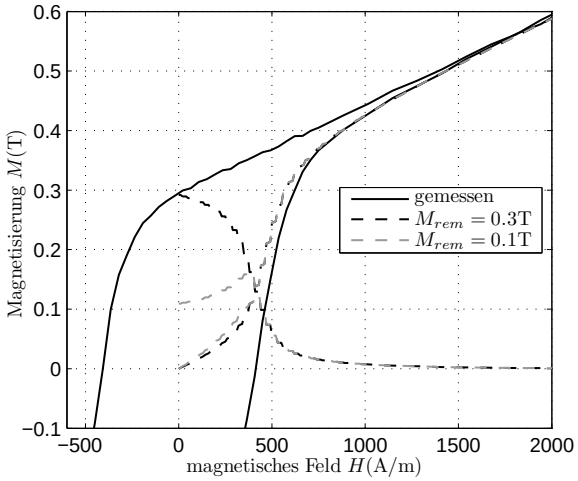


Bild 11: Simulationsergebnisse bei Variation der anfänglichen Remanenz. Material: NiFe.

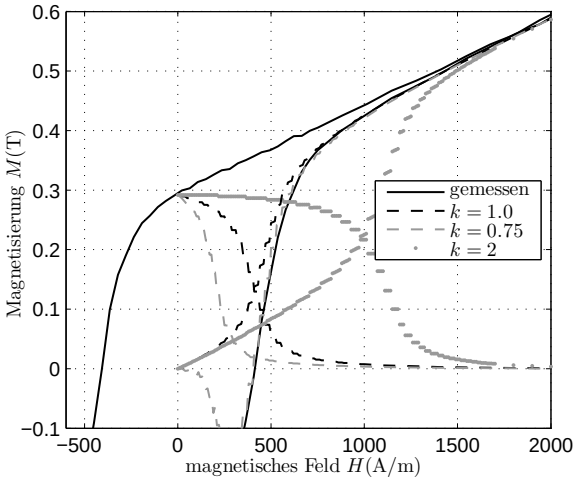
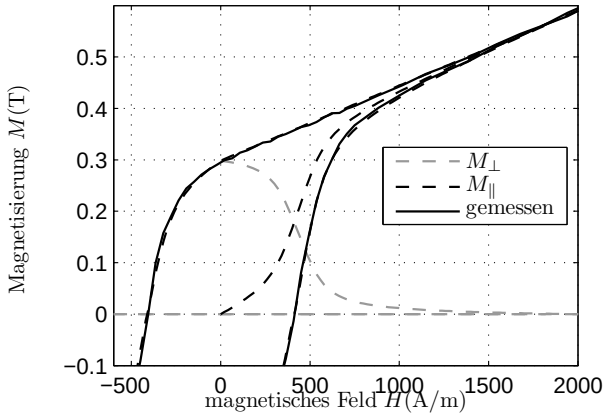


Bild 12: Simulationsergebnisse bei Variation des Drehwiderstands k.



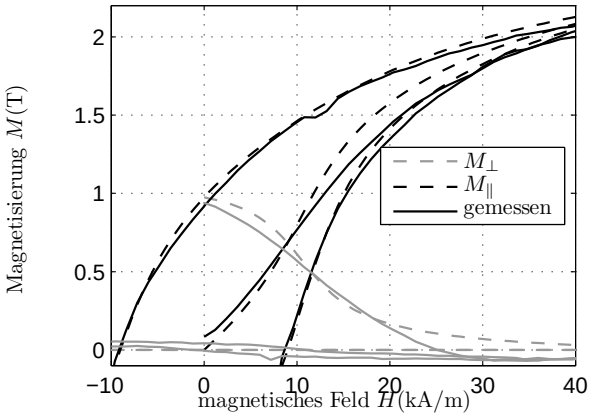


Bild 13: Komponenten vektorieller Hysteresekurven von NiFe (links) und FeCo (rechts). Gezeigt sind die Komponenten der Magnetisierung in Feldrichtung (schwarz) und senkrecht dazu (grau). Gestartet wird immer mit einer senkrechten Remanenzmagnetisierung und einem ansteigenden H-Feld.

erhöht und eine äußere Hysteresekurve durchfahren. In der unteren Darstellung ist schließlich die vektorielle Messung an einer FeCo-Dünnschicht mit dem Simulationsergebnis verglichen.

7 Ausblick

Mit dem VVSM steht nun ein hochempfindliches Messgerät für die vollständige, dreidimensionale Charakterisierung verschiedenster magnetischer Materialien zur Verfügung. Diese Messergebnisse sind beispielsweise für die vollständige Charakterisierung von magnetischen Dünnschichten, wie sie in Mikrosystemen Verwendung finden, wichtig. Ein weiteres Anwendungsgebiet

ist die Charakterisierung von Stahlblechen für Transformatoren und andere Magnetkreise.

Das Vektormodell muss an weiteren Materialien verifiziert und mit anderen Modellen verglichen werden. Der Algorithmus für die Auswertung des Modells muss so effizient werden, dass ein Einsatz in numerischen Berechnungsprogrammen möglich wird.

Literatur

[1] G. Bertotti und I. D. Mayergoyz, *The Science of Hysteresis*, Elsevier, 2005.

[2] A. Sutor, S. J. Rupitsch und R. Lerch, "A Preisach based hysteresis model for magnetic and ferroelectric hysteresis", *Applied Physics A* **100**(2), S. 425–430, 2010. doi: 10.1007/s00339-010-5884-9.

[3] S. Bi, A. Sutor, R. Lerch und Y. Xiao, "A Measurement Setup for Acquiring the Local Magnetic Properties of Plastically Deformed Soft Magnetic Materials", *J. Appl. Phys* , 2011. in press.

[4] L. Benito, J. I. Arnaudas und A. del Moral, "High-sensitivity vector magnetometer for measuring magnetic torque at low temperatures", **77**(2), S. 025101, 2006.

[5] J. Kallwies, "Planung, Aufbau und Evaluierung eines Vector Vibrating Sample Magnetometers", 2010. Bachelorarbeit, Lehrstuhl für Sensorik, Friedrich-Alexander-Universität Erlangen-Nürnberg.

[6] A. A. Adly und H. H. Hanafy, "Incorporating core hysteresis properties in three-dimensional computations of transformer inrush current forces", *Journal of Applied Physics* **105**(7), S. 07A329, 2009.

[7] E. Dlala, A. Belahcen, K. Fonteyn und M. Belkasim, "Improving loss properties of the mayergoyz vector hysteresis model", *Magnetics, IEEE Transactions on* **46**, S. 918 –924, march 2010.

[8] G. Kahler, E. Della Torre und E. Cardelli, "Implementation of the preisach-stoner-wohlfarth classical vector model", *Magnetics, IEEE Transactions on* **46**, S. 21 –28, jan. 2010.

[9] T. Hegewald, B. Kaltenbacher, M. Kaltenbacher und R. Lerch, "Efficient Modeling of Ferroelectric Behavior for the Analysis of Piezoceramic Actuators", *Journal of Intelligent Material Systems and Structures* **19**(10), S. 1117–1129, 2008.

[10] F. Wolf, A. Sutor, S. Rupitsch und R. Lerch, "Modeling and measurement of hysteresis of ferroelectric actuators considering time-dependent behavior", *Procedia Engineering* **5**, S. 87 – 90, 2010.

Messtechnische Hardwareplattform zur Entwicklung neuer Motorsteuergeräte-Algorithmen und Funktionen am Beispiel der Klopferkennung

[1]Biehl Matthias, [1]Kempf Stefan, [2]König Andreas

[1]Robert Bosch GmbH, Entwicklung Applikation Klopf- und Ladedruckregelung, 71701 Schwieberdingen,
[2]Lehrstuhl für Integrierte Sensorsysteme, Technische Universität Kaiserslautern, 67663 Kaiserslautern,

E-Mail: {matthias.biehl, stefan.kempf}@de.bosch.com, koenig@eit.uni-kl.de

Kurzfassung

Es wird die Entwicklung einer Hardwareplattform für den Bereich der Messdatenerfassung, der Applikation und der Entwicklung neuer Steuergerätefunktionen für den Automobilbereich vorgestellt. Das System besitzt die für diesen Bereich typischen Kommunikationsschnittstellen. Besonderheit sind die damit verbundenen leistungsfähigen Recheneinheiten und A/D (Analog/Digital)-Umsetzer, mit im Automobilbereich atypisch hohen Abtastraten von 500kHz. Die erfassten Daten werden in Echtzeit verbrennungsweise mit den Steuergerätemessdaten synchronisiert und eignen sich für Spezialanwendungen und Analysezwecke. Durch die frei programmierbaren Systemkomponenten ist ein vielfältiger Einsatz möglich. Neben den daraus resultierenden Möglichkeiten für eine (teil-) automatisierte Messdatenerfassung und Applikation ist der Einsatz als Entwicklungssystem gewährleistet. Dies wird anhand der Vorteile für die effiziente Entwicklung von Steuergerätefunktionen, durch die echtzeitfähige Hardwareplattform zur frühzeitigen Verifikation und Validierung im offenen und geschlossenen Regelkreis vorgestellt. Der Einsatz des Systems zur hardwarebasierten Simulation und Gütebewertung einer SVM (Support Vector Machines)-basierten Klopferkennung wird in den einzelnen Entwicklungsphasen dargestellt. Ausgehend von den Simulationsergebnissen und nach einer umfassenden Testphase ist der Einsatz des Systems zur Verifikation im offenen

und geschlossenen Regelkreis am Motorprüfstand geplant.

1 Einleitung

Die Aufgabengebiete der Messdatenerfassung, der Applikation und der Entwicklung von Steuergerätefunktionen stellen ein wichtiges Themengebiet im Entwicklungsprozess eines modernen Verbrennungsmotors dar. Mit steigender Komplexität der Systeme und dem dadurch erhöhten Aufwand in der Entwicklungsphase werden neue Anforderungen an zukünftige (teil-) automatisierte Messsysteme gestellt. Eine wichtige Rolle spielt die Echtzeitfähigkeit der Aufzeichnung, Verarbeitung und Auswertung von Messdaten und -signalen parallel zum eingesetzten Motorsteuergerät.
Im Arbeitsgebiet der Klopfregelung gewinnen Störgeräusche im klopfrelevanten Frequenzbereich von 5 bis 30 kHz der komplexer werdenden Motoren zunehmend an Bedeutung. Zur Unterdrückung dieser unerwünschten Störgeräusche können mit dem KID2 (Klopf Intensitäts Detektor)-Hardwaremess- und Applikationssystem die aktuellen Steuergerätefunktionen effizient bedatet werden [1]. Um neue Algorithmen für die Klopferkennung zu implementieren und zu verifizieren ist eine erweiterte Testumgebung notwendig. Die Entwicklung einer für diese Aufgaben angepassten Hardwareplattform, welche im Besonderen den hohen Anforderungen der Klopfregelung an die Datenerfassung, -verarbeitung und -analyse entspricht, wird im Folgenden vorgestellt. Die Leistungsfähigkeit wird durch die Verwendung von Klopferkennungsalgorithmen mit SVM aufgezeigt.

2 Die Klopfregelung

Im folgenden Kapitel werden die wichtigsten Grundlagen der Klopfregelung zusammengefasst.

2.1 Das Phänomen klopfende Verbrennung

Mit klopfenden Verbrennungen werden unerwünschte Verbrennungsvorgänge in einem Ottomotor bezeichnet die eine physikalische Grenze für die Leistungs-

und Wirkungsgrad-Steigerung darstellen. Zusätzlich führt ein andauernder klopfender Motorbetrieb, insbesondere im oberen Drehzahlbereich, zu einer irreparablen Schädigung des Motors, wodurch eine Klopferkennung und -regelung in modernen Verbrennungsmotoren unverzichtbar ist.
Hauptmerkmal einer klopfenden Verbrennung ist die Ausbreitungsgeschwindigkeit der Flammenfront, die bei klopfenden Verbrennungen um den Faktor 10 bis 100 mal höher ist als bei normalen Verbrennungen. Die durch die hohen Flammenfrontgeschwindigkeiten enstehenden Druckwellen sind die Ursache für das charakteristisch metallische Klopfgeräusch, im typischen Frequenzbereich, welches beim Auftreffen an den Zylinderwänden entsteht.
Die in Hinsicht auf die Klopfregelung wichtigste Einflussgröße ist der Zündzeitpunkt, beziehungsweise der Zündwinkel. Weitere Grundlagen wie eine genaue Theorie über das Entstehen von klopfenden Verbrennungen können [2] entnommen werden.

2.2 Methoden zur Klopferkennung

In diesem Abschnitt werden die wichtigsten Verfahren für die Detektion klopfender Verbrennungen vorgestellt.
Das geeignetste Verfahren zur Beurteilung und Erkennung von Klopfereignissen ist die Auswertung des Druckverlaufs im Inneren des Brennraums. Die Druckmesstechnik hat sich hauptsächlich als Referenz für die Beurteilung der Erkennungsgüte anderer Auswertungsstrategien durchgesetzt.
Das heute verbreitetste Verfahren im Serieneinsatz ist die Auswertung des Beschleunigungsverlaufs am Motorblock mit einem Körperschallsensor. Diese kostengünstige und meist robuste Lösung bietet für herkömmliche Verbrennungsmotoren eine ausreichende Erkennungsgüte. Die zunehmende Überlagerung des eigentlichen Nutzsignals mit den verschiedensten mechanischen Störeinflüssen auf das Kurbelgehäuse und somit auf den Klopfsensor ist eine bekannte Herausforderung.

3 Messtechnische Hardwareplattform

In diesem Kapitel werden die allgemeinen und klopfspezifischen Hardwarean-

forderungen anhand der Anwendungsgebiete und den Eigenschaften klopfender Verbrennungen aus Kapitel 2 definiert. Die Lösungsstrategie wird näher beschrieben und die resultierenden Neuerungen und Fortschritte vorgestellt.

3.1 Anforderungen an das Hardwaresystem

Die allgemeinen Anforderungen für die Messdatenerfassung, Applikation und Entwicklung von Steuergerätefunktionen wird durch das automobile Umfeld und deren Standards bestimmt.
Für die Messdatenerfassung sind das die Schnittstellen zum Steuergerät, über das CAN (Controller Area Network) Bussystem mit XCP (Universal Measurement and Calibration Protocol) oder CCP (CAN Calibration Protocol), sowie die proprietäre ETK (EmulatorTastKopf) Schnittstelle der Firma ETAS, die unterstützt werden müssen. Zur Akquise von Messdaten über aufgabenspezifische externe Sensoren, zusätzlich zur Fahrzeugsensorik, werden ergänzend zum CAN Bus, der LIN (Local Interconnect Network) Bus, sowie A/D-Umsetzer benötigt. Die dabei notwendige Abtastrate der A/D-Umsetzer ist stark abhängig vom auszuwertenden Signal und muss dementsprechend variabel gewählt werden können. Das Themengebiet der Klopfregelung stellt an explizit dieser Stelle eine für den Automobilbereich hohe Anforderung, durch eine Abtastrate für die Zylinderdruck- und Klopfsensorsignale von mindestens 200kHz. Die über die verschiedenen Messpfade gewonnenen Daten müssen, bedingt durch ihre unterschiedlich lange Auswerte- und Übertragungsdauer, zeitlich beziehungsweise verbrennungsweise synchronisiert werden. Des Weiteren muss die Übertragung der Daten an einen Computer und das Aufzeichnen selektiv gewählter Messkanäle mit der im Automobilbereich weit verbreiteten Software INCA möglich sein.
Das Aufgabengebiet der Applikation hat den Anspruch an ein Messsystem, dass die erfassten Daten, beziehungsweise deren Auswertung, und die spätere Darstellung an einem Computer in einer engen zeitlichen Korrelation von maximal 800 Millisekunden erfolgen muss. Die sich so ergebenden Echtzeitanforderungen betreffen somit die Hardware in Bezug auf Rechenleistung und -zeit, aber auch die maximale Übertragungsrate und -geschwindigkeit.
Für den Fall des geschlossenen Regelkreises muss eine Anwendungsfall spe-

zifische, echtzeitfähige Rückkopplung von Berechnungsergebnissen der Hardwareplattform in den Funktionsablauf des Steuergeräts bis zum nächsten Arbeitsspiel über einen Steuergeräte-Bypass gewährleistet werden. Die minimal hierfür zur Verfügung stehende Zeit ergibt sich für eine maximale Drehzahl des Motors von 7000 Umdrehungen pro Minute zu etwa 5 Millisekunden. Die Klopfregelung stellt in diesem Bereich keine weiteren spezifischen Anforderungen. Jedoch ist durch die Abtastrate der Rohsignale und der komplexen Filter-, Analyse- und Auswerteverfahren eine im Vergleich zu anderen Steuergerätefunktionen hohe Rechenleistung erforderlich.

3.2 Lösungsstrategie

Die unter Abschnitt 3.1 genannten Anforderungen sind in die drei Bereiche Kommunikation, Signalerfassung und Aufbereitung, sowie die Auswertung bzw. Berechnung von Algorithmen zu unterteilen. Da eine komplette Eigenentwicklung, durch Anforderungen wie die proprietäre ETK Schnittstelle, nicht umsetzbar ist, werden verschiedene bereits auf dem Markt erhältliche Hardwaresysteme so kombiniert und angepasst, dass der Entwicklungsaufwand in diesem Bereich minimiert wird.
Für die Erfassung von Signalen über einen A/D-Umsetzer wird ein Steuergerät von Bosch Motorsport eingesetzt, dessen maximale Abtastrate 500kHz beträgt. Darüber hinaus verfügt diese Hardware über eine variable Signalvorverarbeitung, umgesetzt auf einem FPGA (Field Programmable Gate Array), und ein frei programmierbaren MPC (Mikroprozessor) für die Implementierung und Verifikation von Algorithmen. Für die Kommunikation mit einem Computer unterstützt die Hardware das XCP Protokoll über Ethernet und verfügt über CAN Schnittstellen für den Datenaustausch mit anderen Hardwaresystemen.
Für den Zweck der Kommunikation kommt das Prototyping- und Schnittstellenmodul ES910 der Firma ETAS zum Einsatz, dessen Schnittstellen alle Anforderungen an den Bereich Steuergeräte- und Sensorkommunikation, bis auf einen anforderungsgerechten A/D-Umsetzer, abdeckt. Diese Hardware erfüllt die Anforderungen an einen Steuergerätebypass und somit ebenfalls an die Verifikation von Funktionen und Algorithmen im offenen und geschlossenen Regelkreis.

Da beide Systeme über eine gute Rechenleistung verfügen, die Kommunikation aber durch die Übertragungsrate des CAN Busses eingeschränkt wird, ist bei der Implementierung von Algorithmen darauf zu achten, dass Daten vor einer möglichen Übertragung aufbereitet und auf ein Minimum reduziert werden. Der schematische Hardwareaufbau der beschriebenen Lösungsstrategie kann Bild 1 entnommen werden.

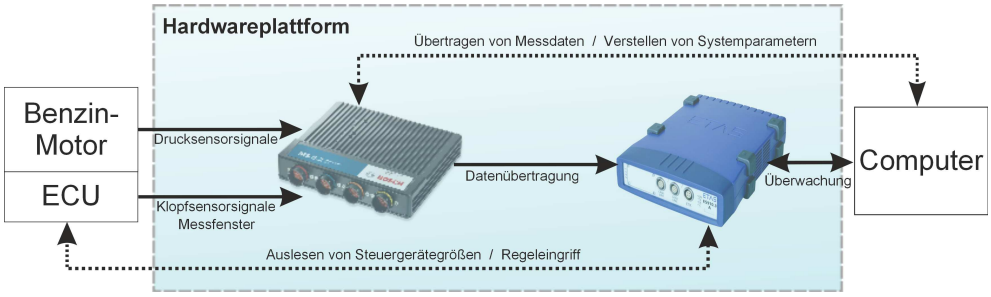


Bild 1: Schematischer Aufbau der Hardwareplattform

3.3 Ergebnisse der Hardware-Umsetzung

Durch die beschriebene Lösungsstrategie können Signale mit einer für den Automobilbereich ausreichenden Abtastrate erfasst und mit den Messgrößen aus dem Steuergerät synchronisiert werden. Ergänzt wird dies um eine echtzeitfähige Analyse und Auswertung, sowie der Besonderheit eines auf hochabgetasteten A/D-Signalen basierenden Steuergeräte-Bypasses, wodurch ein gezielter Regeleingriff in den Funktionsablauf durchgeführt werden kann.
Durch die Kombination und Anpassung der einzelnen Hardwarekomponenten ist eine Implementierung neuer Funktionen zur Klopferkennung basierend auf dem Klopfsensorsignal möglich. Mit der höhere Rechenleistung des Hardwaresystems, gegenüber aktuellen Steuergeräten, können so auch komplexere und zukunftsweisende Funktionen entwickelt und frühzeitig im realen System erprobt werden. Der Einsatz des Test- und Entwicklungssystems zur (teil-) automatisierten Applikation und Messdatenerfassung ist ebenfalls möglich. Ein Einsatzgebiet ist die kontinuierliche Überwachung des Klopfverhaltens in der gesamten Applikationsphase, durch Implementierung einer druckbasierten Regelung, welche unabhängig vom Körperschallsignal eine stets zuverlässige Erkennung bietet [3].
Für den Einsatz als Messsystem ist die limitierende Größe die Übertragungsra-

te von Sensorrohsignalen zu einer Messdatenerfassungssoftware, wie zum Beispiel INCA.

4 Klopferkennung mit Support Vector Machines

Die in Kapitel 2.2 genannten Herausforderungen in der Klopferkennung führen bei modernen Motoren zur Untersuchung neuer Erkennungsverfahren, welche in der Lage sind, auch durch Störgeräusche überlagerte Nutzsignale zuverlässig auswerten zu können.
Mustererkennungsverfahren mit Support Vector Machines bieten hierfür einen geeigneten Ansatz. Dieses Verfahren wird bereits bei der Analyse von geologischen und seismischen Signalen zur Detektion von Anomalien eingesetzt. Da diese Signale prinzipiell mit einem Klopfsensorssignal vergleichbar sind, können daraus folgernd auch für die Klopferkennung gute Ergebnisse angenommen werden. Erste Robustheitsuntersuchungen wurden bereits durch Offline-Simulationen im Rahmen einer Kooperation zwischen dem Lehrstuhl Integrierte Sensorsysteme der Technischen Universität Kaiserslautern und der Robert Bosch GmbH durchgeführt [4].

4.1 Grundlagen der Klopferkennungsmethode

Die folgende Beschreibung des Erkennungsverfahrens mit SVM ist an den Besonderheiten der Klopfregelung ausgerichtet. Die Grundlage hierfür bildet die SVM wie sie [5] entnommen werden kann.
Als Basis der Klassifikationsmethode werden verschiedene Merkmale genutzt, die offline aus Messdateien der jeweiligen Motorvariante gewonnen werden. Hierzu gehören spezifische Frequenzspektren des Klopfsensorsignals, sowie Steuergerätekenngrößen einer jeden Verbrennung. Als Referenz für die spezifische Klopfintensität und die Klassenzugehörigkeit, zu klopfenden oder nicht klopfenden Verbrennungen, wird der gefilterte Druckspitzenwert genutzt.
Die Klopferkennung mit SVM unterteilt sich in die beiden Bereiche Anlernen und Anwenden. In der Anlernphase werden die Muster der Merkmale für klopfende und nicht klopfende Verbrennungen gelernt. Die Herausforderung einer nicht existierenden exakten Grenze zwischen klopfenden und nicht klopfenden

Verbrennungen kann während des Anlernprozesses durch eine geschickte Selektion von Verbrennungen gelöst werden. Unter Verwendung des minimalen gefilterten Druckspitzenwerts und dem Ergebnis des Anlernvorgangs, der Hyperebene, ist es möglich eine direkte Umrechnung zwischen der wahrscheinlichen Klassenzugehörigkeit, welche durch den Abstand zwischen den Merkmalen einer Verbrennung und der angelernten Hyperebene definiert ist, und dem Ausgangsmerkmal, dem Regressionswert, herzustellen. Die Güte der daraus entstehenden Regressionsberechnung lässt sich direkt über ihren Korrelationskoeffizienten, ihre positive und negative Standardabweichung, sowie die Steigung ihrer Ausgleichsgeraden bewerten. Die Druckschwelle ab der Verbrennungen in der Anwendungsphase als klopfend erkannt werden sollen ist als Applikationsparameter projektspezifisch zu ermitteln.

4.2 Implementierung auf dem Hardwaresystem

Ausgehend von den Algorithmen der Robustheitsuntersuchungen [4] wurde die Softwarearchitektur an die Zielhardware angepasst und das Verfahren weiterentwickelt.
In der ersten Phase wurden die Ergebnisse der Robustheitsuntersuchungen anhand von Computersimulationen nach Bild 2 überprüft, bestätigt und das Verfahren optimiert. Die Algorithmen dieser Simulationen bilden die Grundlage für die weiteren Entwicklungsschritte.

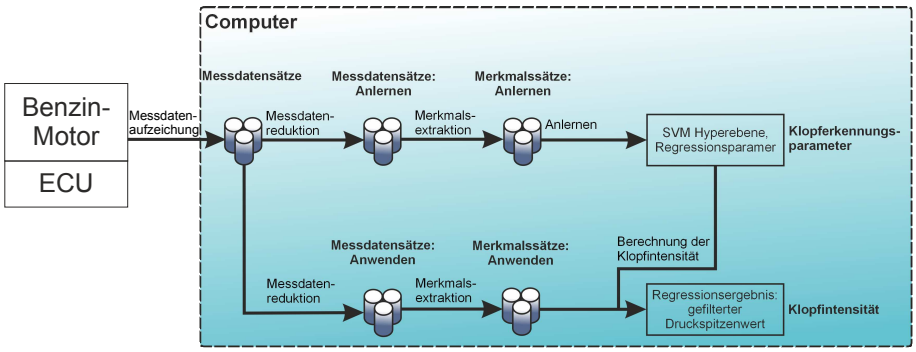


Bild 2: Schematische Darstellung des Datenflusses bei der Simulation

In der zweiten Phase wurden die Algorithmen zur Abstandsberechnung der Merkmale zur Hyperebene, die Regressionsberechnung und die Klopferkennung auf dem Hardwaresystem, wie in Bild 3 dargestellt, realisiert. Die

Berechnung der Merkmale erfolgt dabei weiterhin auf einem Computer. Die Simulationsergebnisse der ersten und zweiten Phase sind verhaltensäquivalent. Zur Bestätigung wurden die Simulationen an 3, 4, 6 und 8 Zylinder Motoren mit Ein- und Mehrklopfsensorsystemen durchgeführt und verglichen.

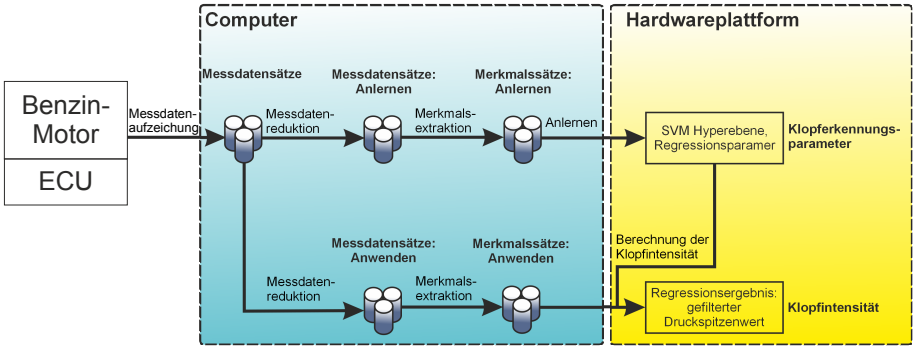


Bild 3: Schematische Darstellung des Datenflusses bei der hardwaregestützten Simulation

In der aktuellen dritten Entwicklungsphase ist der Anlernprozess weiterhin auf dem Computer implementiert. Auf dem beschriebenen Hardwaresystem wird die komplette Messdatenerfassung und -aufbereitung, die Merkmalsextraktion, die Abstandsberechnung der Merkmale zur Hyperebene, die Regressionsberechnung und die Entscheidung auf klopfen oder nicht klopfen, wie in Bild 4 dargestellt, realisiert. Bei der Aufteilung der Rechenoperationen und Funktionen wird darauf geachtet, dass eine möglichst kleine Menge an Daten übertragen werden muss. Bei der Wahl der Algorithmenimplementierung wird beachtet, dass die jeweils ausführende Hardware eine möglichst geringe Berechnungszeit benötigt. Aus diesem Grund erfolgt die Implementierung der Filterfunktionen und die Berechnung der Frequenzspektren als Soft-Cores auf dem FPGA des Motorsportsteuergeräts und die Bestimmung des Abstands der Merkmale zur Hyperebene als C-Code beziehungsweise Simulink Modelle auf dem MPC. Für die hardwarenahe Umsetzung der Algorithmen aus den Simulationsphasen wird der Rechenaufwand weiter optimiert.
Mit Hilfe des berechneten Klopferkennungsmerkmals, sowie weiterer Steuergeräteparameter, erfolgt die abschließende Regressionsberechnung der Klopfintensität auf der ES910. Beim Überschreiten des applizierten Schwellwerts kann ein direkter Steuergeräteeingriff erfolgen, welcher den Regelkreis schließt.
Nach weiteren Simulationen und einer ausführlichen Testphase ist die Verfikati-

on des Systems und der Algorithmen im offenen und geschlossenen Regelkreis aus Resourcengründen erst in der letzten Phase am Motorprüfstand geplant. Das Ziel ist der abschließende Vergleich von Simulations- und Messergebnissen im offenen Regelkreis bei einer einheitlichen Datenbasis und die Gütebewertung des Erkennungsverfahrens im geschlossenen Regelkreis im realen Motorbetrieb.

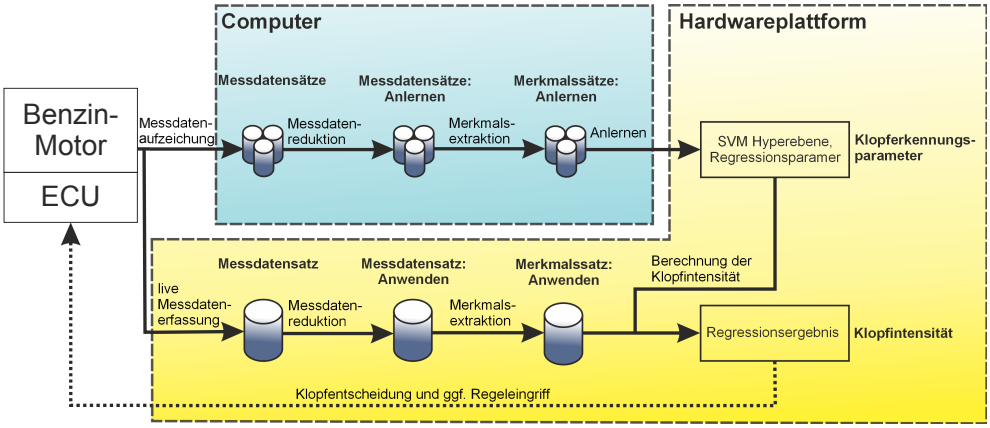


Bild 4: Schematische Darstellung des offenen/geschlossenen Regelkreises

4.3 Ergebnisse der Klopferkennung mit Support Vector Machines

Die Klopferkennungsgüte aus den Robustheitsuntersuchungen mit einer SVM-basierten Klopferkennung konnten mit den durchgeführten Simualtionen nach Bild 2 und 3 bestätigt werden. Es wurden die Messdaten aus [4] mit dem optimierten Verfahren, wie in Bild 6 dargestellt, simuliert und die Korrelationskoeffizienten berechnet.

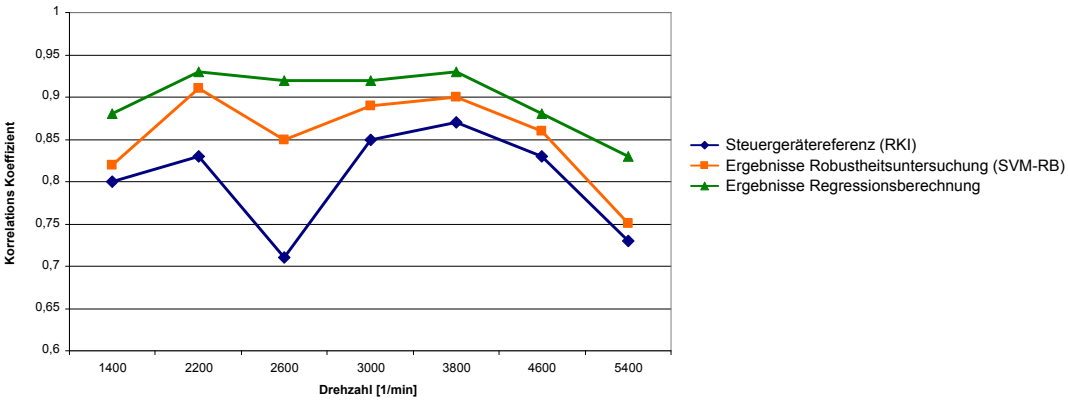


Bild 5: Vergleich der Verfahren anhand des Korrelationskoeffizients

Ein Vergleich verschiedener Methoden ist aufgrund der unterschiedlichen Wertebereiche der Klopferkennungsmerkmale nur eingeschränkt möglich. Jedoch ist die Bewertung eines Klopferkennungsverfahrens über die Anzahl der Falsch- und Nichterkennungen, sowie über die Steigung und Korrelationskoeffizienten, wie in Bild 5 dargestellt, aussagekräftig.

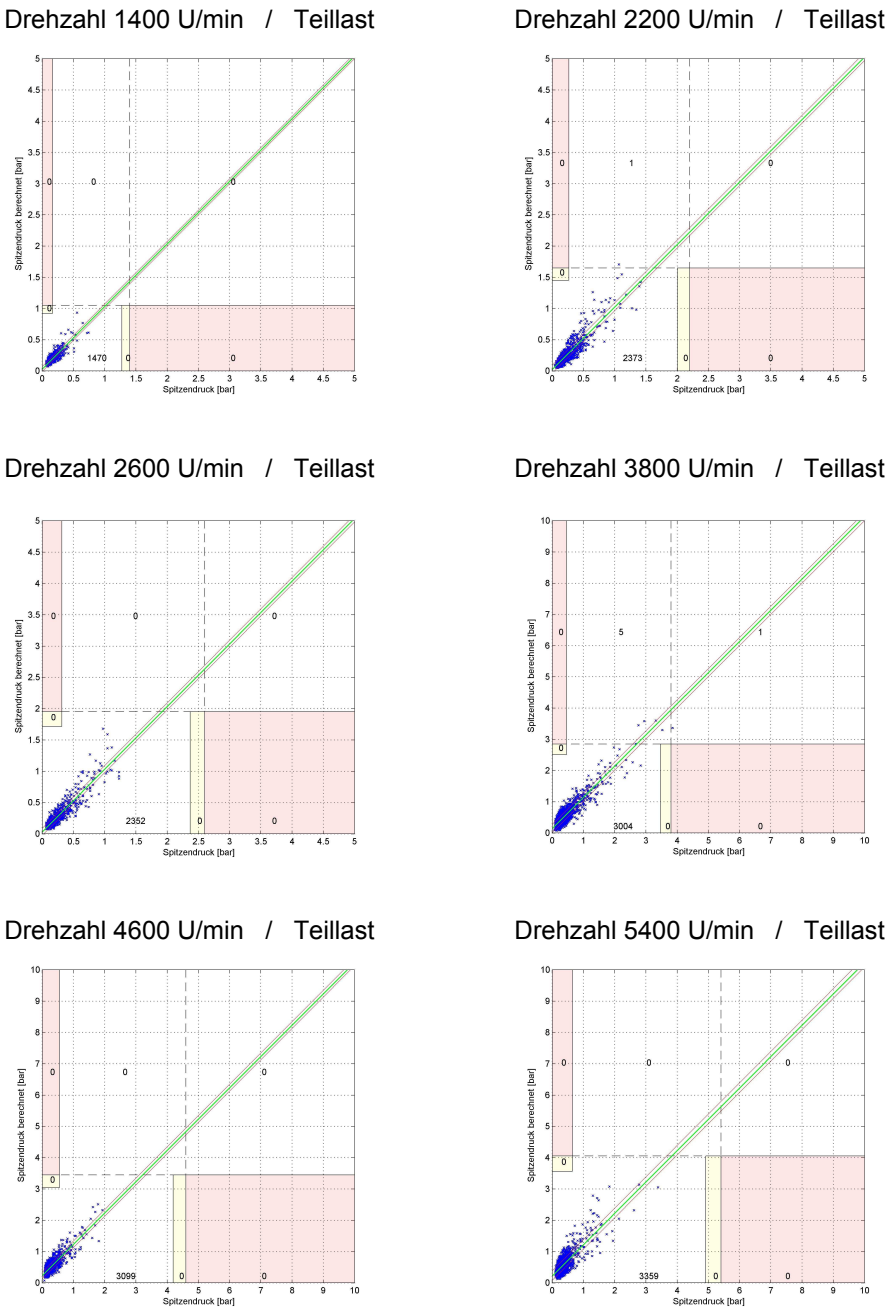


Bild 6: Ergebnisse der SVM basierten Klopferkennung

5 Zusammenfassung und Ausblick

Die Hardwareplattform, bestehend aus einem angepassten, programmierbaren Motorsportsteuergerät und einer ES910, wird in der aktuellen Entwicklungsphase als Test- und Entwicklungssystem eingesetzt. Die Umsetzung umfasst die offline und hardwaregestützte Simulation von beliebigen Funktionen. Die Leistungsfähigkeit des Systems wird anhand einer SVM-basierten Klopferkennung und -regelung bewertet. Auf Basis bestehender Algorithmen wurde die Softwarearchitektur in mehreren Phasen an die Zielhardware angepasst, das Verfahren zur Klopferkennung weiter optimiert und der Rechenaufwand reduziert. Die Simulationsergebnisse aus den ersten beiden Phasen sind verhaltensäquivalent.
In der aktuellen Entwicklungsphase wird die vollständige Funktionalität auf dem Test- und Entwicklungssystem umgesetzt und das Klopferkennungsverfahren weiter optimiert. Nach einer ausführlichen Simulationsphase ist der Betrieb am Motorprüfstand geplant. Eine abschließende Bewertung der Rechenleistung und der SVM-basierten Klopferkennungsgüte im offenen und geschlossenen Regelkreis ist vorgesehen.
Ein weiteres mögliches Einsatzgebiet der Hardwareplattform ist die Unterstützung einer (teil-) automatisierten Applikation von Steuergerätefunktionen durch eine druckbasierte Klopferkennung.

Literatur

[1] R. Sloboda, W. Häming und W. Fischer, “Toolkette zur effektiven Applikation von Klopferkennungssystemen”, *Motortechnische Zeitschrift* (01), S. 26–34, 2004.

[2] M. Fischer und Mitautoren, *Klopfregelung für Ottomotoren - Trends für Serienentwickler*, expert verlag, 2004.

[3] M. Fischer und Mitautoren, *Klopfregelung für Ottomotoren II - Trends für Serienentwickler*, expert verlag, 2006.

[4] I. Kuncup, S. Kempf, A. König und R. Sloboda, “Robustheitsuntersuchung eines SVM-Basierten Klopferkennungsverfahrens”, *Motortechnische Zeitschrift* (07-08), S. 486–491, 2010.

[5] B. Schölkopf und A. J. Smola, *Support Vector Machines, Regulations, Optimizations, and Beyond*, 2004.

Drahtlose Messtechnik zur Langzeitüberwachung von Vitalparametern

Timo Klingeberg, Meinhard Schilling

TU Braunschweig
Institut für Elektrische Messtechnik und Grundlagen der Elektrotechnik
Hans-Sommer-Straße 66, 38106 Braunschweig
Tel.: 0531-391-3867

Email: t.klingeberg@tu-bs.de, m.schilling@tu-bs.de

Kurzfassung

Dieser Beitrag stellt ein Gerät zur Langzeitüberwachung von Vitalparametern vor, welches im Bereich der Prävention und Rehabilitation eingesetzt werden wird. Eine spezielle Zielgruppe stellen dabei ältere Menschen dar, deren Gesundheitsüberwachung von besonderer Bedeutung ist [1]. Bedingt durch den kleinen Aufbau und die einfache Benutzung hat das Gerät eine hohe Akzeptanz. Es werden zeitabhängig die Vitalwerte von EKG, Blutdruck, Blutsauerstoffsättigung, Bewegung und Temperatur aufgenommen und können sowohl lokal gespeichert, als auch über eine Bluetooth Schnittstelle an einen stationären Rechner übertragen werden.

1 Einleitung

Die Aufzeichnung von nicht-invasiv überwachbaren Vitalparametern wie EKG, Blutdruck, Blutsauerstöffsättigung, Bewegung und Temperatur erlaubt aussagekräftige Rückschlüsse auf den Gesundheitszustand sowohl in der Prävention als auch in der Rehabilitation. Durch den universellen Aufbau kann ein solches Gerät in einem breiten Anwendungsfeld eingesetzt werden und ist nicht auf spezielle Krankheitsbilder beschränkt. Im Rahmen des Niedersächsischen Forschungsverbundes „Gestaltung Altersgerechter Lebenswelten“ (GAL) [2] entwickeln wir ein Gerät zur Langzeitüberwachung von Vitalparametern, das wir als „VitalSign Watch“ bezeichnen.

1.1 Themenverwandte Arbeiten

Die Thematik tragbarer Geräte zur Überwachung von Vitalparametern wurde bereits in einer Vielzahl ähnlicher Projekte und Untersuchungen [3–8] bearbeitet. Neben Entwicklungen, die eine konfigurierbare Plattform darstellen und an die universell Sensoren angebunden werden können (ShimmerTM [9], Nexus-10 [10]), gibt es auch Ansätze, die die Sensorik direkt in die Kleidung integrieren [11]. Unser Gerät ist am ehesten vergleichbar mit dem AMON Projekt [12], in dem ebenfalls ein am Handgelenk tragbares Multisensorgerät entwickelt wurde. Im Unterschied zu AMON ist unser Ziel, neben der Aufnahme der einzelnen Werte von EKG, Blutdruck, Blutsauerstoffsättigung, Bewegung und Temperatur, Algorithmen zu entwickeln, mit denen die über einen langen Zeitraum aufgenommenen Vitaldaten untereinander verknüpft werden können, um daraus präzisere Aussagen über den Gesundheitszustand bzw. Gesundheitsverlauf (Verbesserung/Verschlechterung) geben zu können. Auch die Bewertung auf Gültigkeit der Sensorsignale, die aufgrund von Bewegungen durch Artefakte überlagert werden können, ist durch die Datenfusion möglich. Darüber hinaus wird eine Reduzierung des Strombedarfs angestrebt, um eine möglichst lange Nutzungsdauer ohne Batteriewechsel zu erzielen. Hierdurch kann die Compliance des Gerätes gerade bei älteren Menschen (z. B. solche mit feinmotorischen Einschränkungen) erhöht werden.

2 Gerätebeschreibung

Der zentrale Kern des Geräts ist ein Atmel AVR Mikrocontroller ATmega1284P der durch analoge und digitale Ein- und Ausgänge die gesamte Steuerung übernimmt. Das Bluetooth Modul WT12 von Bluegiga ist seriell mit dem Mikrocontroller verbunden und enthält das vollständige Übertragungsprotokoll. Die Energieversorgung wird von zwei AAA-Batterien bereitgestellt und mit Hilfe von Step-Up-Reglern auf die benötigte Spannung von 3,3 V (Bluetooth Modul) bzw. 5 V (Mikrocontroller) geregelt. Weitere Linearregler erzeugen die für die analoge Elektronik benötigten Referenzspannungen von 3,0 V und 2,0 V. Die Sensorik wird im Folgenden näher beschrieben.

2.1 Drahtlose Datenübertragung

Das in dem Gerät integrierte Bluetooth Modul der Klasse 2 ist verantwortlich für die Funkübertragung. Auf der Hardwareseite steht eine serielle Schnittstelle zur Verfügung, die direkt an die UART-Schnittstelle des Mikrocontroller angeschlossen werden kann. Auf der PC-Seite stellt ein virtueller COM-Port den gegenüberliegenden Endpunkt für die Datenübertragung dar. Das Übertragungsprotokoll ist in diesem Fall transparent für den Benutzer. Bei Bluetoothgeräten der Klasse 2 wird eine Übertragungsreichweite von 30 m angegeben, die Versuche mit dem Prototypen zeigten bei Übertragungsraten von 15 kB/s innerhalb eines 25 m^2 großen Raumes keine Beeinträchtigungen.

2.2 Zeitbasis

Der Echtzeituhrbaustein Maxim DS1394 stellt eine Zeitbasis zur Verfügung, die durch einen Pufferkondensator auch während des Batteriewechsels ununterbrochen weiterläuft. Die Markierung der abgespeicherten Vitalparameter mit einem zusätzlichen Zeitstempel ermöglicht bei nachträglicher Analyse der Daten eine absolute Zuordnung zur jeweiligen Tages- oder Nachtzeit.

2.3 Elektrokardiogramm

Die Aufzeichnung des EKG erfolgt klassisch mit zwei im Brustbereich eingesetzten Klebeelektroden und ermöglicht die Messung der Herzfrequenz durch die Bestimmung des zeitlichen Intervalls zwischen zwei aufeinander folgenden QRS-Komplexen.

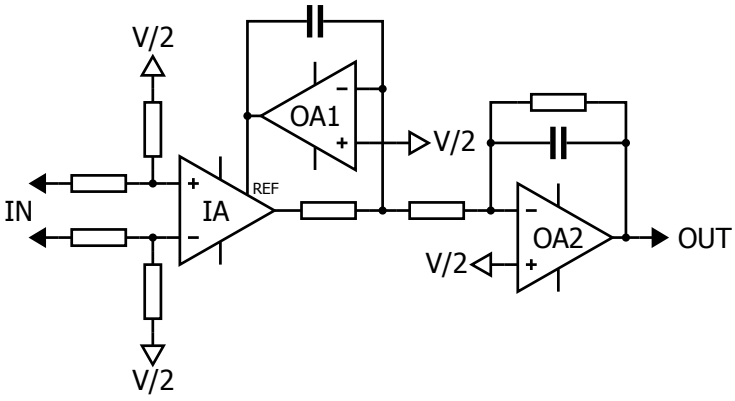


Bild 1: Verstärkerschaltung des EKG-Signals nach [13]

Das elektrische Signal der Elektroden, welches üblicherweise im einstelligen Millivoltbereich liegt, wird über eine Verstärkerschaltung (siehe Bild 1) erhöht und mit dem Analog-Digital-Umsetzer des Mikrocontrollers mit einer Frequenz von 200 Hz abgetastet. Eine einfache Bestimmung der Herzfrequenz kann durch eine Schwellwertdetektion erfolgen, da die R-Zacke die größte Amplitude des QRS-Komplexes aufweist. Eine Alternative zur Verbesserung der Erkennung stellt beispielsweise der Algorithmus von Pan & Tompkins [14] dar, welcher durch geringen Rechenaufwand auch innerhalb des Mikrocontrollers in Echtzeit angewandt werden kann.

2.4 Blutdruck

Die Blutdruckmessung in dem vorgestellten Gerät arbeitet nach der oszillometrischen Methode [15] über eine aufblasbare Manschette am Handgelenk. Der zur Berechnung der Werte für systolischen und diastolischen Blutdruck benötigte Algorithmus erfordert zwei Signale, die aus einem einzelnen Drucksensor analog vorverarbeitet werden.

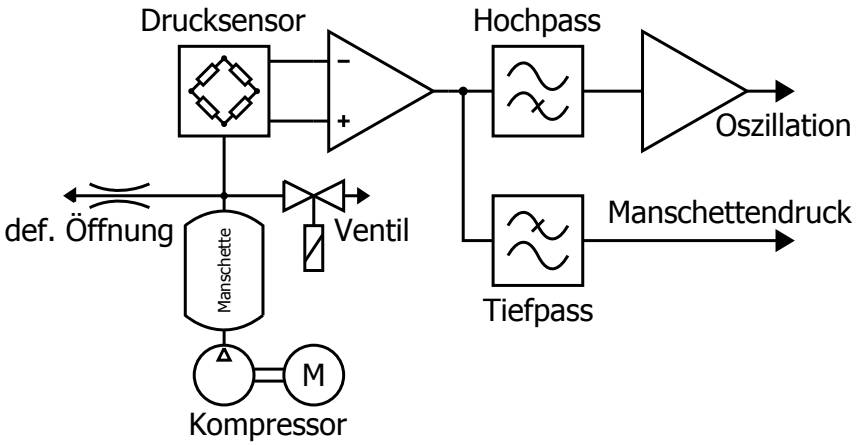


Bild 2: Verstärkerschaltung des Drucksensors, Blockdiagramm

Zunächst wird über einen piezoresistiven Drucksensor ein elektrisches Signal gewonnen, welches von einem Instrumentenverstärker vorverstärkt und anschließend in zwei Pfade aufgeteilt wird (Bild 2). Im ersten Pfad erfolgt eine analoge Tiefpassfilterung, nach der das Signal den mittleren Druck in der Manschette in einem Messbereich von 0 – 300 mmHg repräsentiert. Der zweite Signalpfad durchläuft einen analogen Hochpass und wird ein weiteres Mal verstärkt, um die durch den Herzschlag hervorgerufenen Druckänderungen in

der Manschette zu erhalten, die sich typischerweise im Amplitudenbereich von ±3 mmHg bewegen. Der Wechselspannungsanteil des Signals wird vom Mikrocontroller mit einer Frequenz von 200 Hz abgetastet, die langsamere Änderung des mittleren Drucks erfordert hingegen nur eine geringere Abtastfrequenz. Beim Prototypen wurde hier 20 Hz gewählt. Die Manschette selbst wird über einen Miniaturkompressor aufgeblasen, das Verringern des Drucks von etwa 3 mmHg pro Sekunde erfolgt durch eine definierte Öffnung. Um die Restluft zum Ende der Messung vollständig abzulassen, wird ein kleines Magnetventil eingesetzt, welches wie der Kompressor auch vom Mikrocontroller angesteuert wird.

2.5 Blutsauerstoffsättigung

Zur Messung der Blutsauerstoffsättigung mit Hilfe der Photoplethysmographie (PPG) [16] ist eine Kombination aus analoger und digitaler Signalverarbeitung erforderlich. Die in Bild 3 dargestellte Schaltung ist an die Application Note SLAA274A [17] von Texas Instruments angelehnt, in der der Aufbau und die Software eines Pulsoxymeters erläutert wird. Zur Messung wird der Finger des Probanden mit zwei verschiedenen Wellenlängen (roter und infraroter Bereich) durchleuchtet, das Maß der Absorption wird mit einer Photodiode erfasst.

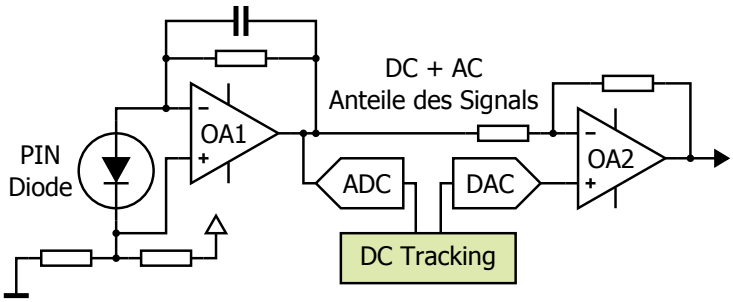


Bild 3: Verstärkerschaltung der Photoplethysmographie

Die erste Stufe bildet ein Transimpedanzverstärker, dessen Ausgang einerseits an den invertierenden Eingang der nächsten Stufe gekoppelt ist, andererseits auch zur Bestimmung des Gleichspannungsanteils von einem analogen Eingang des Mikrocontrollers abgetastet wird. Dieses Signal wird digital tiefpassgefiltert und über einen Digital-Analog-Umsetzer an den positiven Eingang der nächsten Verstärkerstufe geleitet. Diese Lösung für die Offsetkompensation ermöglicht

den zeitgemultiplexten Betrieb der Leuchtdioden, die mit einer Frequenz von 500 Hz umgeschaltet werden. Unter Berücksichtigung der Einschwingzeiten der Verstärkerstufen kann das am Ausgang des zweiten Verstärkers anliegende Signal schließlich synchron zu der Aktivierung der Leuchtdioden abgetastet werden. Aus dem Verhältnis der Absorption beider Wellenlängen wird die Blutsauerstoffsättigung (SpO_2) ermittelt.

2.6 Bewegung

Der im Gerät eingebaute 3-achsige Beschleunigungssensor MMA7455 von Freescale Semiconductor mit einem Messbereich von ± 8 g (g = $9.81 \frac{m}{s^2}$) verfügt über einen integrierten A/D-Umsetzer und ist digital an den Mikrocontroller angebunden. Der Sensor ermöglicht Abtastfrequenzen von bis zu 250 Hz bei einer Auflösung von 10 Bit. Die Messung der Bewegung gibt Aufschluss über die Aktivität des Nutzers und kann beispielsweise als Plausibilitätsprüfung genutzt werden (erhöhte Bewegungsaktivität kann Ursache für erhöhten Herzschlag oder Blutdruck sein). Darüber hinaus können Bewegungsartefakte bei anderen Sensoren leichter identifiziert werden, wenn zeitgleich Bewegungsdaten am Beschleunigungssensor erkannt werden. Da der Sensor ebenfalls die Gravitationsbeschleunigung aufnimmt, kann eine absolute Aussage über die Lage des Sensors im Raum getroffen werden. Am Handgelenk eingesetzt bedeutet dies beispielsweise die Erkennung, ob der Arm während der Blutdruckmessung in die richtige Position gebracht wurde.

2.7 Temperatur

Die Temperaturmessung erfolgt über einen Pt1000-Widerstand in einer Brückenschaltung mit Ausschlagverfahren (Bild 4). Die Brückenspannung wird über einen Differenzverstärker an den analogen Eingang des Mikrocontrollers angepasst und dort digitalisiert. Durch Veränderung des parallelen Widerstandes R_{set} in der Brücke und des Verstärkungsfaktors, kann der zu messende Temperaturbereich ausgewählt werden. Gemessen werden die Außentemperatur, die Handgelenkstemperatur sowie ein nahe am Körperkern liegender Wert durch Anbringen des Pt1000 direkt unter einer der Klebeelektroden der EKG-Messung.

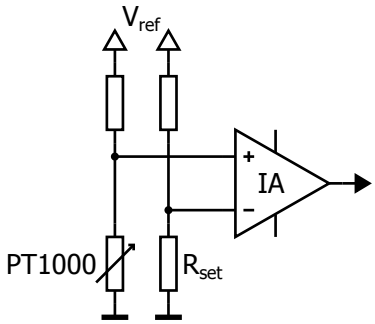


Bild 4: Temperaturmessbrücke mit Pt1000

3 Funktionsweise, „VitalSign Watch“

Um die Funktion der Sensorik zu testen, wurde zunächst ein Prototyp entworfen (siehe Bild 5), der bis auf die Photoplethysmographie alle Sensoren implementiert hat. Integriert wurde der Schaltungsaufbau in ein Gehäuse eines handelsüblichen Blutdruckmessgerätes, da hier bereits Manschette, Kompressor und Magnetventil enthalten waren.

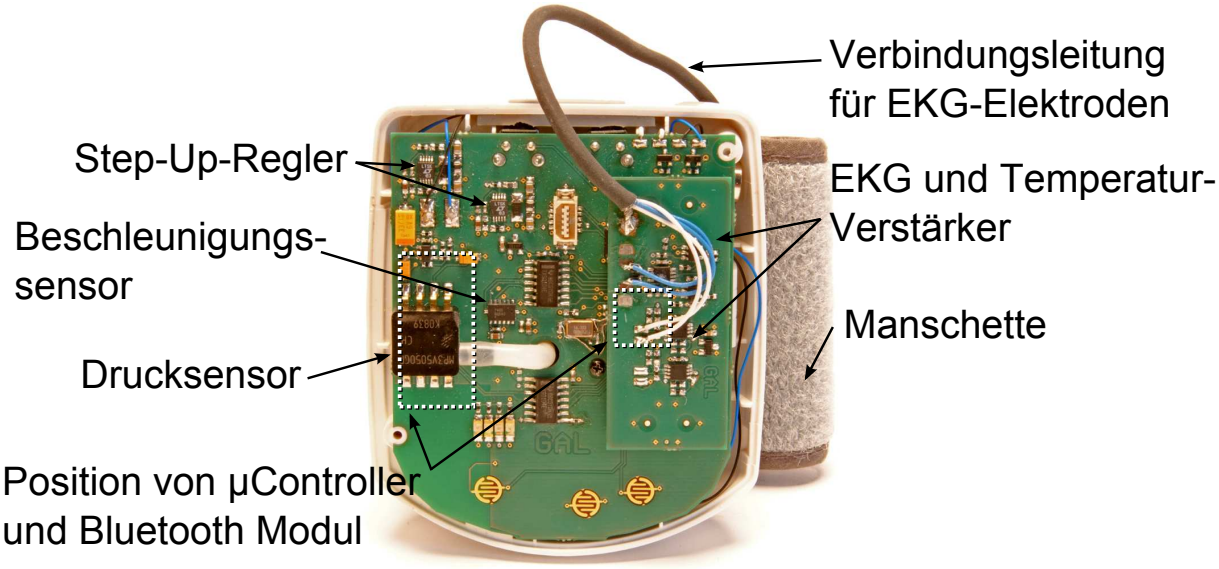


Bild 5: VitalSign Watch, geöffnete Darstellung

Um die Messwerte der Sensoren vollständig untersuchen zu können, wird vom Mikrocontroller des Prototyps zunächst keine Datenvorverarbeitung durchgeführt. Über einen Zeitgeber werden alle analogen und digitalen Werte mit einer fest definierten Frequenz abgefragt und für das Versenden über die Bluetooth Schnittstelle in einem seriellen Datenstrom vorbereitet. Beim Prototyp

wurde die Übertragung auf eine einfache Dekodierung optimiert, um die Rohsensordaten am PC zunächst ohne großen Aufwand visualisieren oder sichern zu können und die Weiterverarbeitung mit MATLAB [18] zu ermöglichen.

3.1 Energiebedarf

Im Prototypen wurde die grundsätzliche Funktion der Sensorik und der Datenübertragung realisiert. Daher können nur allgemeine Aussagen zum Energiebedarf gegeben werden, die sich aus Messungen und auf der Basis der jeweiligen Datenblätter ergeben.

Komponente	benötigter Strom
Miniaturkompressor	300 mA
Magnetventil	150 mA
Bluetooth Modul (Übertragung)	30 mA
Bluetooth Modul (Standby)	5 mA
Leuchtdioden (PPG)	20 mA
Mikrocontroller (16 MHz, 5 V)	10 mA
Display	< 1 mA
Displaybeleuchtung	80 mA
EKG, Temperatur, Bewegung	< 1 mA

Tabelle 1: Strombedarf einzelner Komponenten

Tabelle 1 gibt einen Überblick über den Strombedarf der einzelnen Komponenten, während diese aktiv sind. Beim Mikrocontroller bedeutet das z. B. einen kontinuierlichen Bedarf von 10 mA während jeder Art von Messung. Das Bluetooth Modul hat ein integriertes Energiemanagement und benötigt die angegebenen 30 mA nur während der Übertragung von Daten. Das Datenblatt gibt einen Strombedarf von 5 mA an, wenn das Modul mit dem Rechner verbunden ist, jedoch keine Daten übertragen werden.

Ausgehend von einer oszillometrischen Blutdruckmessung, die alle 10 Minuten durchgeführt wird, jeweils 30 Sekunden dauert und die Komponenten Kompressor, Ventil, Bluetooth Modul, Mikrocontroller und Display benötigt, kann ein Strombedarf von etwa 35 mA/h errechnet werden. Beim Einsatz von handelsüblichen AAA-Batterien mit einer Kapazität von 1000 mA/h ergibt sich eine Laufzeit von nicht viel mehr als einem Tag. Durch Einsparung von Energie, z. B. durch Sleep-Modi des Mikrocontrollers und einer Datenreduktion durch

Vorverarbeitung kann die Batterielebensdauer bzw. damit das Batteriewechselintervall vergrößert werden, was sich ebenfalls positiv auf die Compliance bei älteren Menschen auswirkt. Zur weiteren Optimierung ist darüber hinaus der Einsatz der Pulswellenlaufzeitmessung zur Blutdruckbestimmung denkbar, wodurch der Kompressor und das Magnetventil lediglich zur Kalibiermessung eingesetzt würden.

3.2 Messungen

Zur Demonstration der Funktionsweise sollen an dieser Stelle einige Beispielmessungen dargestellt werden. Bild 6 zeigt den vollständigen Verlauf der Rohdatenaufzeichnung für die Blutdruckbestimmung. Im oberen Teil A ist der mittlere Manschettendruck über der Zeit aufgetragen, der bedingt durch das Aufblasen zu Beginn ansteigt. Im unteren Teil B ist der Amplitudenverlauf des hochpassgefilterten Signals dargestellt. In der Phase des Druckverlustes durch die definierte Öffnung zeichnet sich im Teil B ein für diese Messung typischer Amplitudenverlauf ab. Der am Ende des Manschettendrucks noch sichtbare schnelle Abfall auf Null wird durch das finale Öffnen des Magnetventils hervorgerufen.

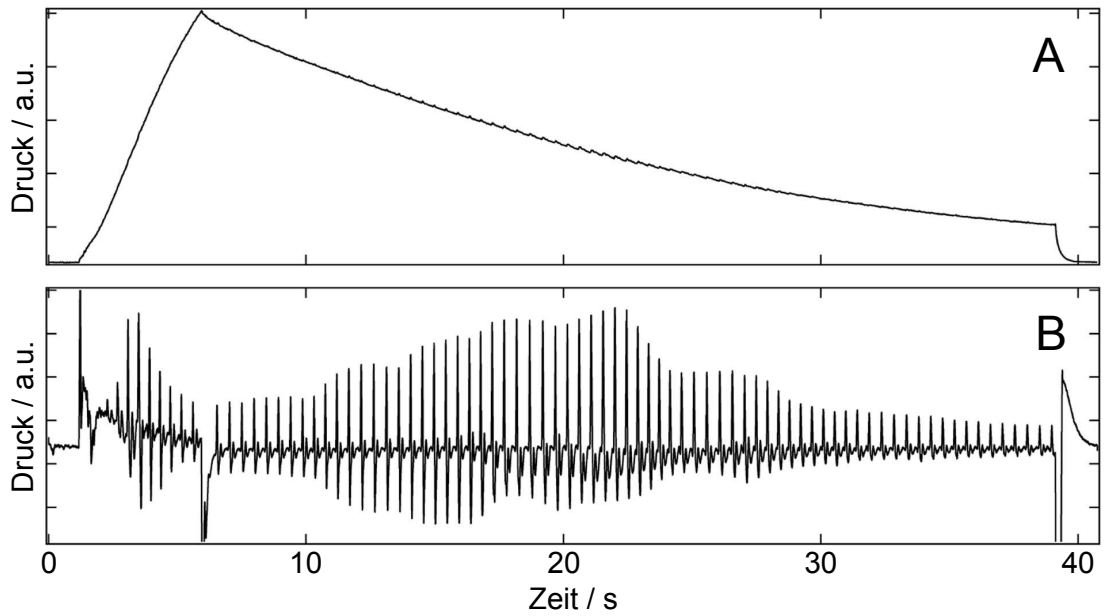


Bild 6: Rohdaten der Blutdruckmessung, A: mittlerer Manschettendruck, B: Amplitudenverlauf des hochpassgefilterten Signals

Bild 7 stellt eine Beispielmessung für die Sensorfusion dar. Im oberen Teil (A) sind zwei QRS-Komplexe zu erkennen, die mit dem 1-Kanal EKG über die

Klebeelektroden abgeleitet wurden. Der untere Teil (B) stellt einen vergrößerten Ausschnitt des Wechselanteils des Drucksensors dar (vergl. Bild 6B). Die Pulswellenlaufzeit (Pulse Transition Time) von etwa 300 ms zwischen R-Zacke im EKG-Signal und dem Amplitudenmaximum beim Drucksensor lässt Rückschlüsse auf den Blutdruck zu.

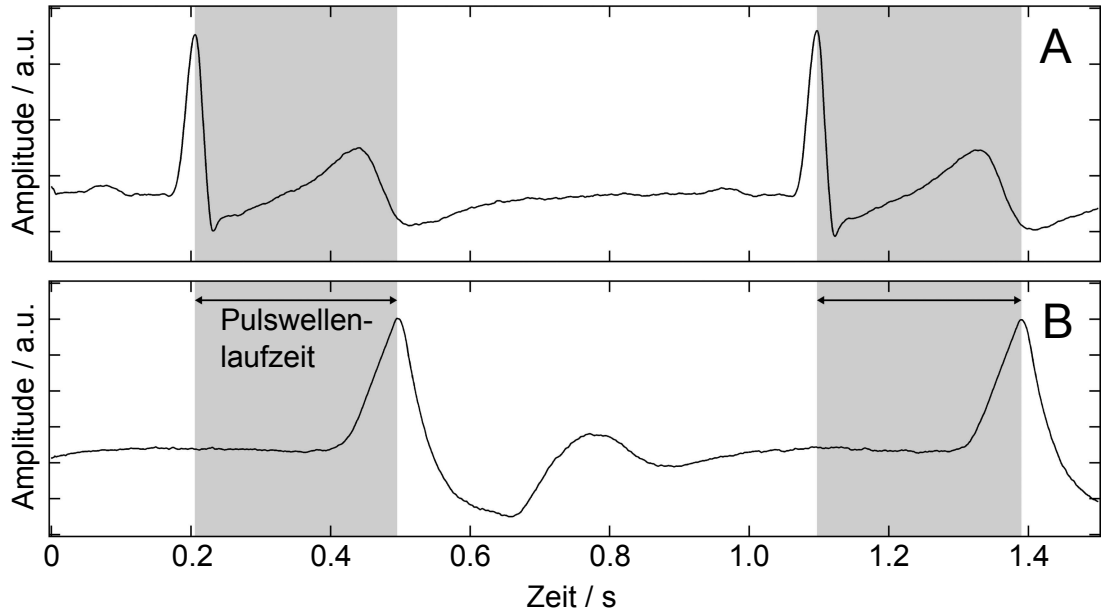


Bild 7: A: EKG-Signal, B: Manschettendrucksensor

4 Zusammenfassung und Ausblick

Es wurde ein drahtloses Sensorsystem zur Langzeitüberwachung von Vitalparametern vorgestellt. Mit dem Prototypen wurde gezeigt, dass die Aufnahme und die drahtlose Übertragung der zahlreichen Vitalparameter über das Bluetooth Modul funktioniert. Die Beispielmessungen zeigen darüber hinaus die Möglichkeit der Datenfusion, um fehlerbereinigte Daten zu erhalten. Aus den im Ruhezustand gewonnenen Messwerten konnten bereits als abgeleitete Größen Herzrate, Blutdruckwerte und Aktivitätsdaten ermittelt werden.

Im nächsten Schritt ist es notwendig, den Fokus auf eine wesentliche Senkung des Energiebedarfs zur Erhöhung der Batterielaufzeit zu konzentrieren, um die „VitalSign Watch“ mit hoher Compliance einsetzen zu können. Ansätze hierfür bieten beispielsweise die Wahl eines verbesserten Mikroncotrollers der Reihe ATxmega, der bei sowohl geringerer Spannungsvorsorgung als auch nied-

rigerem Strombedarf eine höhere Taktfrequenz aufweist. Der Stromverbrauch von Kompressor und Ventil kann durch die Pulslaufzeitmessung über EKG und Photoplethymographie zur Bestimmung des Blutdrucks reduziert werden, eine Vorverarbeitung und Auswertung der Sensordaten (z. B. Implementierung der Berechnungsalgorithmen im Mikrocontroller) ermöglicht niedrigere Datenraten über die drahtlose Schnittstelle.

5 Danksagungen

Besonderer Dank gilt dem Niedersächsischen Forschungsverbund „Gestaltung Altersgerechter Lebenswelten“ (GAL) für die Finanzierung des Projektes.

Literatur

[1] Haux et al., “The lower saxony research network design of environments for ageing: Towards interdisciplinary research on information and communication technologies in ageing societies”, *Informatics for Health and Social Care* **35**(3-4), S. 92–103, 2010.

[2] GAL, “Niedersächsischer Forschungsverbund Gestaltung altersgerechter Lebenswelten”. http://www.altersgerechte-lebenswelten.de/, 07 2011.

[3] T. Degen, H. Jaeckel, M. Rufer und S. Wyss, “SPEEDY: A Fall Detector in a Wrist Watch”, *Proc. Seventh IEEE International Symposium on Wearable Computing* , S. 184–187, 2003.

[4] J. Luprano, J. Sola, S. Dasen, J. M. Koller und O. Chetelat, “Combination of Body Sensor Networks and On-Body Signal Processing Algorithms: the practical case of MyHeart project”, in *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks*, 2006.

[5] T. Yilmaz, R. Foster und Y. Hao, “Review: Detecting Vital Signs with Wearable Wireless Sensors”, *Sensors* **10**, S. 10837–10862, 2010.

[6] M. Schilling und U. Tegtbur, “Informationsfusion in Multisensorsystemen für die Sportmedizin”, *Expertenforum (GMA/VDE) Informationsfusion in der Mess- und Sensortechnik* **Juli 2006**, S. 251–260, 2006.

[7] M. A. D. Raya und L. G. Sison, “Adaptive Noise Cancelling of Motion Artifact in Stress ECG Signals Using Accelerometer”, in *Proceedings of the Second Joint EMBS/BMES Conference*, 2002.

[8] J. Ottenbacher, M. Kirst, L. Jatobá, M. Huflejt, U. Großmann und W. Stork,

"Reliable Motion Artifact Detection for ECG Monitoring Systems with Dry Electrodes", in *30th Annual International IEEE EMBS Conference*, 2008.

[9] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu und V. Cionca, "Shimmer - A Wireless Sensor Platform for Noninvasive Biomedical Research", *IEEE Sens* **10**, S. 1527–1534, 2010.

[10] Mind Media B.V., "Mind Media B.V. Information about the NeXus-10.". http://www.mindmedia.nl/english/nexus10.php/, 07 2011.

[11] R. Paradiso, G. Loriga und N. Taccini, "A Wearable Health Care System Based on Knitted Integrated Sensors", *IEEE Transactions on Information Technology in Biomedicine* **9**, S. 337–344, 2005.

[12] U. Anliker, J. A. Ward, P. Lukowicz, G. Tröster, F. Dolveck, M. Baer, F. Keita, E. B. Schenker, F. Catarsi, L. Coluccini, A. Belardinelli, D. Shklarski, M. Alon, E. Hirt, R. Schmid und M. Vuskovic, "AMON: A Wearable Multiparameter Medical Monitoring and Alert System", *IEEE Transactions on Information Technology in Biomedicine* **8**, S. 415–427, 2004.

[13] Texas Instruments, "SBOS168D: INA321 microPower, Singly-Supply, CMOS Instrumentation Amplifier", 01 2006.

[14] J. Pan und W. J. Tompkins, "A Real-Time QRS Detection Algorithm", *IEEE Transactions on Biomedical Engineering* **3**, S. 230–236, 1985.

[15] M. Ramsey III, "Noninvasive automatic determination of mean arterial pressure", *Medical and Biological Engineering and Computing* **17**, S. 11–18, 1979.

[16] I. Yoshiya, Y. Shimada und K. Tanaka, "Spectrophotometric monitoring of arterial oxygen saturation in the fingertip", *Med. & Biol. Eng. & Comput.* **18**, S. 27–32, 1980.

[17] Texas Instruments, "SLAA274A: A Single-Chip Pulsoximeter Design Using the MSP430", 11 2005.

[18] MathWorks, "MathWorks Deutschland - MATLAB and Simulink for Technical Computing". http://www.mathworks.de/, 07 2011.

Quellentrennung für unterbestimmte Szenarien im Frequenzbereich

Andreas Sandmair, Fernando Puente León

Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik,
Hertzstr. 16, 76187 Karlsruhe, E-Mail: {sandmair, puente}@kit.edu

Kurzfassung

Die Fähigkeit des intelligenten, selektiven Hörens ermöglicht es dem Menschen, auch in akustisch anspruchsvollen Umgebungen eine Unterhaltung führen zu können. Bei der Aktivität mehrerer Sprecher können wir durch eine Zerlegung der Signale die relevanten Sprachanteile extrahieren. Eine Trennung von Sprachsignalen ist auch aus technischer Sicht sehr interessant. Es wird ein Verfahren zur Quellentrennung im Frequenzbereich vorgestellt, das insbesondere in unterbestimmten Szenarien (Anzahl der Quellen $>$ Anzahl der Sensoren) verwendet werden kann. Durch die zusätzliche Berücksichtigung von Nachbarschaften in Zeit- und Frequenzrichtung lässt sich die Qualität der rekonstruierten Signale verbessern.

1 Einleitung

Der Mensch besitzt die Fähigkeit, auch in geräuschvoller Umgebung (viele Störquellen) seinem Gesprächspartner folgen zu können. Durch eine intelligente Auswertung der Laufzeit- und Pegeldifferenzen zwischen den beiden Ohren kann der Mensch eine Separation der Quellsignale durchführen und störende Schallquellen unterdrücken. Dieses Phänomen wird durch den Cocktail-Party-Effekt beschrieben [1]. Personen, die in ihrem Hörvermögen eingeschränkt sind, haben in derartigen Szenarien Verständnisprobleme. Somit könnte eine technische Umsetzung der Quellentrennung [2] beispielsweise im Bereich der Hörgeräteakustik sinnvoll eingesetzt werden.

Um ein entsprechendes Verfahren zu realisieren, ist als Erstes eine mathematische Beschreibung der Problemstellung (siehe Bild 1) notwendig. Die Schallsignale der einzelnen Quellen $s_i(t)$ ($i = 1, ..., M$) werden durch den Einfluss der spezifischen Impulsantworten $a_{ji}(t)$ verändert und überlagern sich entsprechend dem Superpositionsprinzip an den Mikrofonen. Die Sensorsignale lassen sich somit durch

$$x_j(t) = \sum_{i=1}^{M} a_{ji}(t) * s_i(t) \tag{1}$$

beschreiben. Die Impulsantworten sind von der Position der Quelle bzw. des Sprechers und den Umgebungsbedingungen abhängig. Die Schallausbreitung im Freifeld kann durch $a_{ji}(t) = b_{ji}\,\delta(t - t_{ji})$ beschrieben werden. Die Parameter sind die Dämpfung b_{ji} und die Laufzeit t_{ji} entlang des direkten Pfades. In reflexionsbehafteter Umgebung ist eine Berücksichtigung der Reflexionspfade notwendig. Die resultierende Impulsantwort $a_{ji}(t) = \sum_{k=1}^{K} b_{ji}^k\,\delta(t - t_{ji}^k)$ trägt diesem Aspekt Rechnung, indem die spezifischen Dämpfungen und Zeitverzögerungen aller K Ausbreitungspfade berücksichtigt werden.

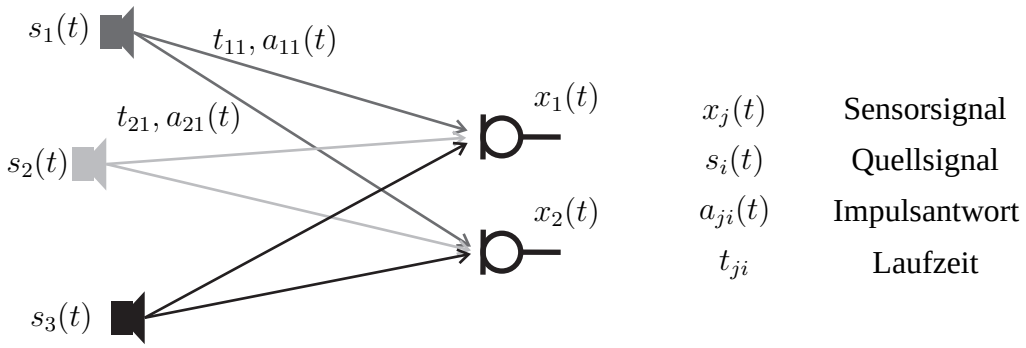


Bild 1: Abstrahierte Darstellung eines Szenarios für $M = 3$ Quellen und zwei Sensoren. Die direkten Ausbreitungspfade sind eingezeichnet.

Um eine Separation der Signale durchführen zu können, müssen mindestens zwei Sensoren zur Verfügung stehen. Aus den beiden Sensorsignalen können charakteristische Merkmale für jede Quelle extrahiert werden, die eine Trennung der Signale ermöglichen. Analog zum Menschen eignen sich die Laufzeitunterschiede der einzelnen Quellen als spezifische Eigenschaften.

Für den reflexionsfreien Fall lassen sich die Merkmale sehr einfach beschreiben. Die Laufzeitdifferenz für den i-ten Sprecher ergibt sich zu $\Delta\, t_i = t_{1i} - t_{2i}$ aus

den Laufzeiten der direkten Pfade. Mit dem Zeitunterschied ist eine frequenzabhängige Phasendifferenz $\Delta\varphi_i(f) = 2\pi f \Delta t_i$ verbunden. Im Zeitbereich ist eine Detektion der quellspezifischen Merkmale und somit die Trennung der Signale schwierig. Eine Lösung stellt die Transformation in den Zeit-Frequenz-Bereich dar. Der Faltungsoperator wird durch eine Multiplikation ersetzt, wodurch sich für (1) das Signalmodell ergibt:

$$X_j(\tau, f) = \sum_{i=1}^{M} A_{ji}(f) \cdot S_i(\tau, f). \tag{2}$$

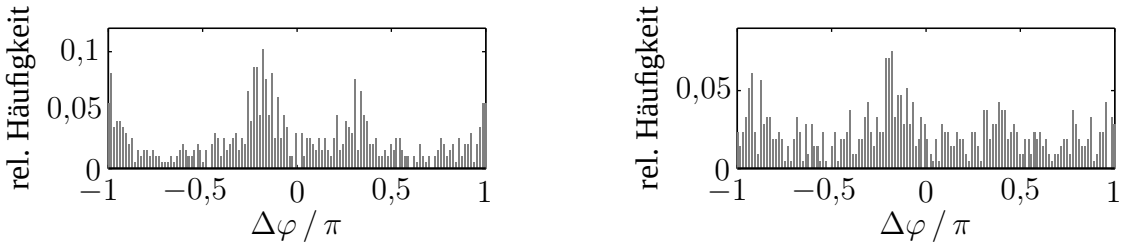
Die Phasenlage zwischen den korrespondierenden Koeffizienten $A_{ji}(f)$ entspricht der frequenzabhängigen Phasendifferenz des i-ten Sprechers gemäß

$$\Delta\varphi_i(f) = \arg[A_{1i}(f)] - \arg[A_{2i}(f)]. \tag{3}$$

Die Detektion der Phasendifferenz zwischen den Sensorsignalen $\Delta\varphi(\tau, f) = \arg[X_1(\tau, f) \cdot X_2(\tau, f)^*]$ in jedem Zeitschritt bietet somit die Möglichkeit zur Rekonstruktion der Signale. Anhand des Wertes kann eine Schätzung der Anteile einzelner Quellen erfolgen.

Die Separation in reflexionsbehafteter Umgebung basiert auf dem selben Grundgedanken. Für jede Frequenz können quellspezifische Merkmale ermittelt werden, auf deren Basis eine Zerlegung der Signale erfolgt. Die Charakteristika werden anhand von Bild 2 erläutert. Die Histogramme über die Phasendifferenzen zeigen jeweils drei Anhäufungen, welche den einzelnen Sprechern ($M = 3$) zugeordnet werden können. Die Klassen lassen sich durch die Erwartungswerte charakterisieren, die folglich quellspezifische Merkmale darstellen. Die Erwartungswerte weichen auf Grund der Mehrwegeausbreitung von den Phasendifferenzen des direkten Pfades ab. Mit zunehmenden Reflexionen wird die Detektion der Merkmale schwieriger.

Für die Separation im Frequenzbereich werden die beschriebenen Eigenschaften genutzt. Detailliertere Betrachtungen zur Signaltrennung im Frequenzbereich sind unter anderem in Yilmaz et al. [3] zu finden. Im Folgenden wird ein Verfahren zur Quellentrennung in reflexionsbehafteter Umgebung vorgestellt und gezeigt, wie durch die Berücksichtigung von Nachbarschaften in Zeit- und



(a) Nachhallzeit: 130 ms (b) Nachhallzeit: 250 ms

Bild 2: Werteverteilung der Phasendifferenzen (für $f = 0{,}3\, f_\mathrm{A}$, $f_\mathrm{A} = 16\,\mathrm{kHz}$) im reflektionsbehafteten Fall für $M = 3$ Sprecher und zwei unterschiedliche Nachhallzeiten (allgemeines Maß für die Stärke der Reflexionen).

Frequenzrichtung eine Verbesserung der Ergebnisse erzielt werden kann. Anschließend werden die Resultate vorgestellt.

2 Quellentrennung

Die Vorgehensweise bei der Quellentrennung orientiert sich an der Beschreibung im vorhergehenden Kapitel. Entsprechend Bild 3 sind vier grundlegende Schritte notwendig. Nahezu alle Verfahren zur Separation im Frequenzbereich

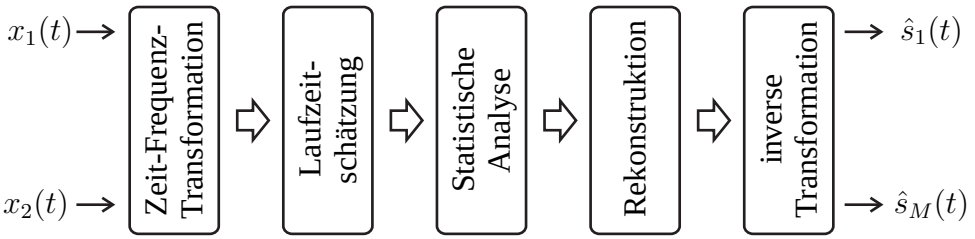


Bild 3: Blockdiagramm zur Beschreibung des Separationsverfahrens. Die einzelnen Verarbeitungsschritte können blockweise durchgeführt werden.

zeigen einen ähnlichen Aufbau. Ein Beispiel ist in [4] zu finden. Die einzelnen Verarbeitungsschritte werden in diesem Kapitel näher beschrieben.

2.1 Zeit-Frequenz-Transformation

Für die Berechnung der Zeit-Frequenz-Darstellung wird die Kurzzeit-Fourier-Transformation (‘short-time Fourier transform’, STFT) verwendet. Die STFT ist durch

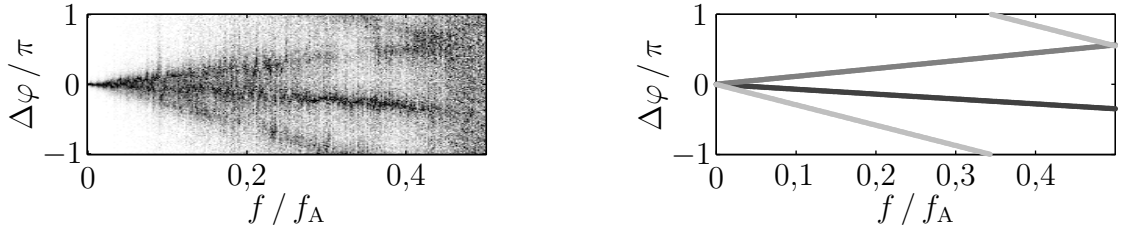
$$X_j(\tau, f) = \int_{-\infty}^{\infty} x_j(t)\, \gamma^*(t-\tau)\, \mathrm{e}^{-\mathrm{j}2\pi f t}\, \mathrm{d}t \tag{4}$$

definiert. Durch die Fensterfunktion $\gamma(t)$ wird das Signal vor der Transformation auf kompakte Abschnitte beschränkt. Diese Begrenzung ist notwendig, um quasistationäre Bereiche des Sprachsignals zu ermitteln.

2.2 Laufzeitschätzung

Die Bestimmung der quellspezifischen Laufzeitunterschiede / Phasendifferenzen (jeweils für den direkten Pfad) stellt den nächsten Schritt dar. Obwohl die Ausbreitungseigenschaften in reflexionsbehafteter Umgebung nicht durch die Phasenwerte des direkten Pfades beschrieben werden können, liefert die Schätzung wichtige Informationen für die weiteren Verarbeitungsschritte.

Die Laufzeitschätzung erfolgt durch die Auswertung der frequenzabhängigen Phasendifferenzen für alle Zeit- und Frequenzschritte [5, 6]. Im Folgenden wird die Umsetzung nach [6] kurz erläutert. Ausgangspunkt des Verfahrens bildet die Frequenz-Phasendifferenz-Darstellung (siehe Bild 4 (a)). Die Histogramme werden für alle Frequenzbänder ermittelt und mit aufsteigender Frequenz in einer Matrix spaltenweise angeordnet. Der Grauwert gibt die Häufigkeit der Phasenwerte an. In dem Bild sind drei geradenhafte Verläufe erkennbar, wobei



(a) Frequenz-Phasendifferenz-Darst. (b) Laufzeitschätzung

Bild 4: Ermittlung der Laufzeit auf der Basis einer modifizierten Radon-Transformation. Die Transformation wird auf die Frequenz-Phasendifferenz-Darstellung angewendet.

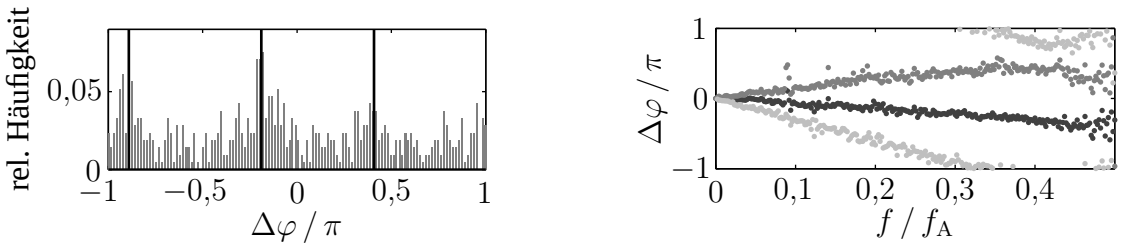
die zugrundeliegende lineare Struktur dem direkten Pfad zugeordnet ist und die Abweichungen aus der Mehrwegeausbreitung resultieren. Mit der Detektion der Geraden erhält man faktisch eine Schätzung der Laufzeit. Für die Bestimmung von geradenähnlichen Strukturen wird in der Bildverarbeitung unter anderem die Radon-Transformation verwendet. Nach einigen Anpassungen lässt sich die Transformation auch zur Liniendetektion in der Frequenz-Phasendifferenz-Darstellung anwenden. Beispielsweise müssen die Sprünge der Phasenwerte

(bei $f = 0{,}35\, f_\mathrm{A}$) berücksichtigt werden, welche aus der Periodizität der Exponentialfunktion resultieren. Die detektierten Geraden sind in Bild 4 (b) eingezeichnet. Aus den Steigungen können gemäß $\Delta\,\varphi_i(f) = (2\,\pi\,\Delta\,t_i)\cdot f$ die Laufzeitdifferenzen der einzelnen Sprecher ermittelt werden.

2.3 Statistische Analyse

Ziel der statistischen Analyse ist die Ermittlung der quellspezifischen Ausbreitungseigenschaften in jedem Frequenzband. Die Werte der Phasendifferenzen ordnen sich tendenziell um die sprecherspezifischen Erwartungswerte (siehe Bild 5 (a)) an. Die konzentrierte Verteilung um die Erwartungswerte resultiert insbesondere aus der Charakteristik der menschlichen Sprache. In einem quasistationären Sprachabschnitt eines Sprechers sind normalerweise nur wenig dominante Frequenzen enthalten. Sind mehrere Sprecher gleichzeitig aktiv, besteht eine hohe Wahrscheinlichkeit, dass diese dominanten Sprecheranteile im Frequenzbereich disjunkt sind und nur eine Person einen hohen Beitrag pro Frequenzband liefert [3]. Die Phasenwerte zwischen zwei sprecherspezifischen Erwartungswerten resultieren aus zeitgleichen Beiträgen der Sprecher.

Für eine Bestimmung der sprecherabhängigen Kenngrößen eignen sich Ansätze zur Fuzzy-Clusteranalyse (FZC) [7]. Bei diesen Verfahren werden die Phasenwerte $\Delta\,\varphi(\tau, f)$ anteilig den Sprechern / Klassen zugeordnet. Im Folgenden soll die Vorgehensweise zur Clusteranalyse (‘fuzzy-c-means’) kurz erläutert werden [8].



(a) $f = 0{,}3 f_\mathrm{A}$ (b) Alle Frequenzbänder

Bild 5: Ermittlung der Erwartungswerte für $M = 3$ Sprecher. Die Werte werden mit einem Clusterverfahren in jedem Frequenzband bestimmt.

Bei der Clusteranalyse wird allgemein versucht, durch Minimierung einer Bewertungsfunktion die ideale Zuordnung der Messwerte zu den Sprechern / Klas-

sen zu finden. Die Funktion hat im Bereich des FZC die Form

$$J(u_i(\tau, f), \Delta\varphi_i^{\mathrm{E}}(f)) = \sum_{\tau} \sum_{i} u_i(\tau, f) \cdot D_i(\tau, f). \tag{5}$$

Die Variable $u_i(\tau, f) \in [0, 1]$ kennzeichnet die Zugehörigkeit des Messwertes $\Delta\varphi(\tau, f)$ zum i-ten Sprecher, $D_i(\tau, f)$ ist ein beliebiges Abstandsmaß zwischen Messwert und sprecherspezifischem Erwartungswert $\Delta\varphi_i^{\mathrm{E}}(f)$. Häufig wird die euklidische Distanz gewählt. Mit Hilfe des Lagrangeverfahrens können Bestimmungsgleichungen für die Variablen ermittelt werden. Es bietet sich folgende Vorgehensweise an:

1. Die Initialisierungswerte werden aus der Laufzeitschätzung zu $\Delta\varphi_i^{\mathrm{E},0}(f) = \Delta\varphi_i(f)$ gewählt.
2. Die Zugehörigkeiten lassen sich nach

   $$u_i(\tau, f) = \frac{(1/D_i(\tau, f))^{\frac{1}{1-m_{\mathrm{E}}}}}{\sum_i 1/D_i(\tau, f))^{\frac{1}{1-m_{\mathrm{E}}}}}. \tag{6}$$

   berechnen. Über den Parameter m_{E} kann die Schärfe der Zuordnung eingestellt werden.
3. Die Mittelwerte im l-ten Iterationsschritt werden durch

   $$\Delta\varphi_i^{\mathrm{E},l}(f) = \arg\left[\frac{\sum_\tau u_i(\tau, f) \cdot \mathrm{e}^{\mathrm{j}\Delta\varphi(\tau,f)} \cdot |X_1(\tau, f)|}{\sum_\tau u_i(\tau, f) \cdot |X_1(\tau, f)|}\right] \tag{7}$$

   bestimmt. Durch die Berücksichtigung der Amplitudenwerte wird eine schnellere Konvergenz erreicht und die Schätzung der Erwartungswerte verbessert. Die Verwendung der Exponentialfunktion berücksichtigt die Sprünge der Phasenwerte bei π bzw. $-\pi$.
4. Das Verfahren wird beendet, wenn die Erwartungswerte konvergieren (Sonst: weiter mit Schritt 2).

In Abbildung 5 sind im linken Bild die Erwartungswerte im Histogramm eingetragen. Im rechten Bild sind die Erwartungswerte über alle Frequenzbänder aufgetragen. Die Abweichungen vom direkten Pfad sind deutlich erkennbar. Auf Grund der Initialisierung mit den Werten aus der Laufzeitschätzung kann bereits eine konkrete Zuweisung der Erwartungswerte zu den einzelnen Sprechern

erfolgen. Die sprecherspezifischen Werte sind farblich gekennzeichnet. Mit Hilfe der Zugehörigkeiten können weitere Kenngrößen zur Charakterisierung der einzelnen Sprecher ermittelt werden (z. B. die Varianzen $\sigma_i(f)$ der einzelnen Klassen).

2.4 Rekonstruktion

Die Informationen über die sprecherspezifischen Merkmale ($\Delta\varphi_i^{\mathrm{E}}(f)$, $\sigma_i(f)$) lassen sich im Folgenden zur Rekonstruktion der Quellsignale verwenden. Aus den Kenngrößen wird für jeden Sprecher eine Wahrscheinlichkeitsverteilung $\mathcal{P}(\Delta\varphi_i^{\mathrm{E}}(f), \sigma_i(f))$ für die Phasenwerte ermittelt. Als Verteilungsfunktion eignet sich die Laplaceverteilung. Es hat sich gezeigt, dass die resultierende Gesamtverteilung

$$\mathcal{P}(f) = \sum_i \alpha_i \cdot \mathcal{P}(\Delta\varphi_i^{\mathrm{E}}(f), \sigma_i(f)) \tag{8}$$

die Werteverteilung der Phasendifferenzen gut approximiert. Die Koeffizienten α_i werden aus den Zugehörigkeiten ermittelt und beschreiben die Sprecheranteile in dem jeweiligen Frequenzband.

Auf der Basis der Verteilungen wird für jeden Messwert eine sprecherabhängige Wahrscheinlichkeit

$$p_i(\tau, f) = p(\Delta\varphi(\tau, f); \alpha_i, \Delta\varphi_i^{\mathrm{E}}(f), \sigma_i(f))$$

berechnet. Der relative Wert

$$M_i(\tau, f) = \frac{p_i(\tau, f)}{\sum_i p_i(\tau, f)}$$

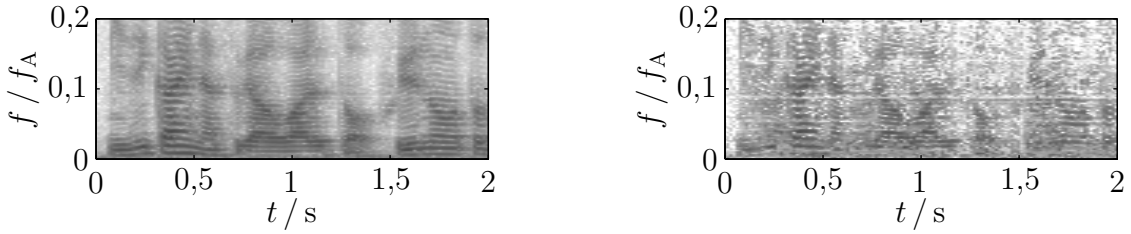
wird zur Ermittlung der Quellsignale verwendet. Eine Schätzung ergibt sich zu

$$\hat{S}_{ji}(\tau, f) = M_i(\tau, f) \cdot X_j(\tau, f). \tag{9}$$

Es stehen nur die Mischsignale an den Sensoren zur Verfügung. Aus diesem Grund lassen sich ausschließlich die Quellsignale an den Sensoren rekonstruieren, welche durch $\hat{S}_{ji}(\tau, f)$ gekennzeichnet sind.

3 Nachbarschaften in Zeit und Frequenz

Durch einen Vergleich der Quellsignale $S_{ji}(\tau, f)$ und der rekonstruierten Koeffizienten $\hat{S}_{ji}(\tau, f)$ lassen sich die Eigenschaften der Rekonstruktion genauer untersuchen. Die Zeit-Frequenz-Darstellungen sind in Bild 6 skizziert. Die Struktur des Sprachsignals ist prinzipiell auch in der rekonstruierten Darstellung erkennbar. Die Rekonstruktion zeigt jedoch zwei Nachteile.



(a) Originalsignal $S_{ji}(\tau, f)$ (b) Rekonstruktion $\hat{S}_{ji}(\tau, f)$

Bild 6: Vergleich der ursprünglichen und geschätzten Koeffizienten im Zeit-Frequenz-Bereich. Die Darstellung ist zur besseren Sichtbarkeit in Zeit und Frequenz beschränkt.

Im unteren Frequenzbereich treten vermehrt Ungenauigkeiten auf. Diese resultieren aus dem geringen Abstand der Erwartungswerte der Phasendifferenzen (siehe Bild 5 (b)), wodurch eine eindeutige Zuordnung zu einzelnen Sprechern erschwert wird. Um eine Verbesserung der Ergebnisse zu erzielen, wird zusätzlich die Sprecherwahrscheinlichkeit in einem Zeitschritt ermittelt:

$$p_i(\tau) = \frac{\sum_{f^\mathrm{K}} p_i(\tau, f)}{\sum_i \sum_{f^\mathrm{K}} p_i(\tau, f)}. \tag{10}$$

Für die Berechnung werden nur die Werte im Frequenzbereich $f^\mathrm{K} \in [f^\mathrm{A}, f^\mathrm{B}]$ verwendet. Diese Einschränkung ist notwendig, um einerseits die Charakteristik der menschlichen Sprache zu berücksichtigen, andererseits Bereiche mit einem großen Abstand der Erwartungswerte zu wählen. Diese Wahrscheinlichkeit wird verwendet, um neue Werte zu berechnen. Eine Anpassung erfolgt nach

$$p_i^\mathrm{T}(\tau, f) = \begin{cases} p_i(\tau, f) \cdot p_i(f) & \text{für} \quad f \in [1, f_\mathrm{C}] \\ p_i(\tau, f) & \text{sonst} \end{cases} \tag{11}$$

nur im unteren Bereich bis zu einer Grenzfrequenz f_C.

Aus der elementweisen Rekonstruktion resultiert ein weiterer Nachteil. Im Originalsignal ist ein kontinuierlicher Verlauf der Koeffizienten über die Zeit erkennbar, der typisch für die menschliche Sprache ist. In der Rekonstruktion treten Unterbrechungen im Verlauf auf. Durch die Berücksichtigung mehrerer Zeitschritte kann eine Glättung der Werte in einem Frequenzband erfolgen. Die modifizierte Wahrscheinlichkeit

$$p_i^{\mathrm{F}}(\tau, f) = \sum_{n_v=0}^{N_{\mathrm{V}}} p_i(\tau - n_v\,\Delta\tau, f) \cdot \beta^{n_v} \tag{12}$$

wird durch Berücksichtigung der N_{V} vorhergehenden Werte erzielt. Durch den Vergessensfaktor β wird der Einfluss der älteren Wahrscheinlichkeiten festgelegt.

Zur Ermittlung der relativen Sprecherwahrscheinlichkeit können die neuen Werte $p_i^{\mathrm{X}}(\tau, f)$ verwendet werden. Eine Kombination der beiden Ansätze ist ebenfalls möglich, wird im Folgenden jedoch nicht betrachtet.

4 Simulation und Ergebnisse

Für die Evaluation der Verfahren ist die Wahl der Testsignale und der Bewertungsverfahren elementar. Die verwendeten Sprachsignale wurden im Rahmen der *Signal Separation Evaluation Campaign*[1] als 'Development'-Datensatz zur Verfügung gestellt. Aus den Quellsignalen an den Sensoren lassen sich bis 32 unterschiedliche Mischsignale ($M = 3$ Sprecher) definieren.

Die Evaluation erfolgt anhand spezifischer Kenngrößen, welche für die Bewertung von Verfahren zur Quellentrennung entwickelt wurden [9]. Steht das Quellsignale $s_{ji}(t)$ zur Verfügung, kann mittels einer 'Least-squares'-Projektion die Zerlegung

$$\hat{s}_{ji}(t) = s_{ji}(t) + e_{ji}^{\mathrm{spat}}(t) + e_{ji}^{\mathrm{interf}}(t) + e_{ji}^{\mathrm{artif}}(t) \tag{13}$$

ermittelt werden. Es erfolgt eine Aufspaltung in Nutz- und Fehlersignale. Die Fehlersignale beschreiben den Einfluss räumlicher Störungen ('spat'), anderer Sprecher ('interf') und künstlicher Artefakte bei der Rekonstruktion ('ar-

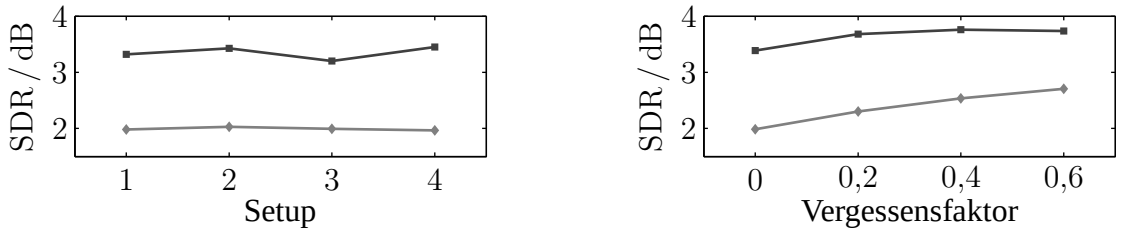
[1] http://sisec2010.wiki.irisa.fr/tiki-index.php?page=Audio+source+separation

tif'). Eine Bewertung erfolgt anhand von Energieverhältnissen. Das 'signal-to-distortion ratio' ist beispielsweise ein Maß für die allgemeine Qualität der Rekonstruktion und wird gemäß

$$\mathrm{SDR}_i = 10 \log_{10} = \frac{\sum_j \sum_t s_{ji}(t)^2}{\sum_j \sum_t \left(e_{ji}^{\mathrm{spat}}(t) + e_{ji}^{\mathrm{interf}}(t) + e_{ji}^{\mathrm{artif}}(t) \right)^2} \tag{14}$$

für jeden Sprecher berechnet. Informationen über weitere Bewertungsmaße sind in [9] zu finden.

Um die Erweiterungen zu bewerten, wird für alle 32 Testsignale die Quellentrennung durchgeführt und die durchschnittliche Separationsqualität ermittelt. Die Ergebnisse sind für beide Fälle in Bild 7 angegeben. Für die Ermittlung der



(a) Zeitschritt $p_i^{\mathrm{T}}(\tau, f)$ (b) Frequenzband $p_i^{\mathrm{F}}(\tau, f)$

Bild 7: Ergebnisse unter Berücksichtigung der Nachbarschaften in Zeit und Frequenz. Die Versuche wurden für zwei verschiedene Nachhallzeiten durchgeführt: $130\,\mathrm{ms}$ (dunkelgrau) und $250\,\mathrm{ms}$ (hellgrau).

Sprecherwahrscheinlichkeit in einem Zeitschritt (Bild 7 (a)) wurden die Grenzen folgendermaßen gewählt:

$f_{\mathrm{A}} = f_{\mathrm{C}}$	1. ohne	2. 1 kHz	3. 2 kHz	4. 1 kHz
f_{B}	1. ohne	2. 4 kHz	3. 4 kHz	4. 2 kHz

Der Vergessensfaktor β zur Berücksichtigung mehrerer Zeitschritte wurde im Bereich von $0 - 0{,}6$ gewählt. Es gilt zu berücksichtigen, dass die Werte für β in Abhängigkeit der Fensterlänge $\gamma(t)$ festgelegt werden müssen. Aus den Resultaten ist ersichtlich, dass mit beiden Ansätze eine Verbesserung der Ergebnisse erzielt werden kann.

5 Zusammenfassung

Bei der Trennung von Sprachsignalen ist die Separation im Frequenzbereich ein vielversprechender Ansatz. Das vorgestellte Verfahren liefert bei vergleichsweise geringem Rechenaufwand bereits gute Ergebnisse. Durch die Verwendung von Nachbarschaften kann die Qualität der Separation weiter verbessert werden. Insbesondere die Berücksichtigung der Koeffizienten über mehrere Zeitschritte ermöglich eine deutliche Erhöhung der Signalqualität.

Literatur

[1] E. C. Cherry, "Some experiments on the recognition of speech, with one and with two ears", *Journal of the acoustical society of America* **25**(5), S. 975–979, 1953.

[2] P. D. O'Grady, B. A. Pearlmutter und S. T. Rickard, "Survey of sparse and non-sparse methods in source separation", *International Journal of Imaging Systems and Technology* **15**(1), S. 18–33, 2005.

[3] O. Yilmaz und S. Rickard, "Blind separation of speech mixtures via time-frequency masking", *IEEE Transactions on Signal Processing* **52**(7), S. 1830–1847, 2004.

[4] H. Sawada, S. Araki und S. Makino, "A two-stage frequency-domain blind source separation method for underdetermined convolutive mixtures", in *2007 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, S. 139–142, IEEE, 2007.

[5] F. Nesta, M. Omologo und P. Svaizer, "A novel robust solution to the permutation problem based on a joint multiple tdoa estimation", in *Proc. International Workshop for Acoustic Echo and Noise Control, IWAENC*, 2008.

[6] A. Sandmair, M. Lietz, J. Stefan und F. Puente León, "Time delay estimation in the time-frequency domain based on a line detection approach", in *ICASSP*, (Prague), 22-27 May 2011.

[7] J. A. Hartigan und M. A. Wong, "A k-means clustering algorithm", *Journal of the Royal Statistical Society. Series C, Applied statistics* **28**, S. 100–108, 1979.

[8] K. Kroschel, *Statistische Informationstechnik: Signal-und Mustererkennung, Parameter-und Signalschätzung*, Springer, 2003.

[9] E. Vincent, R. Gribonval und C. Févotte, "Performance measurement in blind audio source separation", *IEEE Transactions on Audio, Speech, and Language Processing* **14**(4), S. 1462–1469, 2006.

Pore Detection on Cylinder Bore Surfaces

Limeng Wang, Fernando Puente León

Institute of Industrial Information Technology, Karlsruhe Institute of Technology
76187 Karlsruhe, E-Mail: {limeng.wang, fernando.puente-leon}@kit.edu

Abstract

Nowadays cast iron with vermicular graphite is used as the material of cylinder bodies. By honing cylinder bores, the graphite components are opened as pores on the surface. Pores contribute to the engine lubrication and further influence engine performance in the form of microscopic oil reservoirs. To evaluate the pore structure, the images of cylinder bore surfaces are acquired by an optical microscope. This paper proposes a feature-based method for pore detection. The features are derived from a pair of morphological scale spaces. The experiment manifests that our method has good robustness against illuminating and textural disturbances in the background.

1 Introduction

The manufacturing technology of cylinder liners keeps developing in order to reduce noxious emissions (e.g. particles and exhaust). Nowadays cast iron with vermicular graphite is used as the material of cylinder bodies. By honing cylinder bore surfaces the graphite components are opened as pores, which construct a specially designed tribological system (“micro-pressure chamber system”) [1]. Pores contribute to the engine lubrication in the form of microscopic oil reservoirs. On the running process of engines, the pistons “swim” on the pores where motor oil is retained. Utilizing this principle, the friction between the piston and the cylinder wall can be heavily reduced. Thus, the pore structure is of high significance for the engine performance such as the fuel and oil consumption, running resistance and piston wear.

The current evaluation methods for the pore structure are based on the analysis of 2D or 2.5D grayscale images. On the one hand an optical microscope is applied to the visual inspection of lateral surface features. Due to the highly insta-

ble illumination and the unexpected background texture, at present pore detection is still performed visually. The evaluation results are subjective and not reproducible. On the other hand pores can be detected by measuring the surface topology with a white light interferometer (WLI). The surface data coded with gray values are known as 2.5D images, in which pores are segmented in terms of the depth information [2]. Due to the low speed of image acquisition with a WLI, the method is not suitable for the fast evaluation of overall surface appearance. This paper proposes a feature-based method in optical microscope images. The features are derived from a pair of morphological top-hat and bottom-hat scale spaces. Through the investigation of scale-space behavior, a feature persistency measure is developed for the pore detection.

The remainder of this paper is organized as follows: Section 2 introduces the basic idea of our method. Section 3 describes the multiscale morphological filtering. Strategies allowing for feature extraction and pore segmentation are proposed in Section 4. The performance analysis is carried out in Section 5. Finally, we conclude our work.

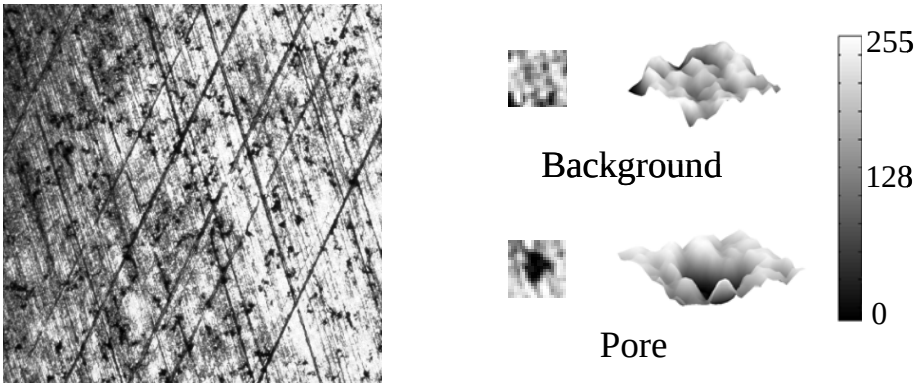


Fig. 1 Cylinder bore surface and its components.

2 Basic idea

With an optical microscope, pores are always imaged as dark regions, since very little light can escape from such deep structures. The same situation also occurs at deep grooves. In Fig. 1 the image intensity is taken as the third extent except the spatial coordinates. The foreground objects, that is, pores and deep honing grooves, are shown as large valleys, whereas background disturbances like metal shining and fine grooves are composed of clustered small peaks. The intensity variation of fore- and background can be seen as the image roughness of differ-

ent scales. Our method starts with the foreground segmentation based on the morphological scale-space behavior. The final detection is achieved by eliminating grooves according to their geometrical features.

3 Multiscale morphological filtering

Mathematical morphology [3] is a nonlinear methodology for image analysis. The basic morphological operations are dilation and erosion. More complicated operators can be designed by connecting a number of these basic operators in a specified order. The simplest case is the combination of a dilation operator and an erosion operator. Sequential calculation of dilation and erosion leads to the morphological opening and closing. Consider an image $i(x)$ as a function mapping a discrete grid E into $\mathbb{R} \cup \{-\infty, \infty\}$. Then, dilation and erosion of a gray-level image $i(x)$ by a structuring element $B(x)$ can be notated as:

$$(i \oplus B)(x) = \sup_{y \in E} \{ i(y) + B(x-y) \}, \tag{1}$$

$$(i \ominus B)(x) = \inf_{y \in E} \{ i(y) - B(y-x) \}, \tag{2}$$

where “sup” denotes supremum and “inf” denotes infimum.
Furthermore, opening and closing are defined as:

$$y_o(x) \triangleq (i \circ B)(x) = (i(x) \ominus B(x)) \oplus B(x), \tag{3}$$

$$y_c(x) \triangleq (i \bullet B)(x) = (i(x) \oplus B(x)) \ominus B(x). \tag{4}$$

Subtracting the opened image from the original one, we obtain the top-hat transform:

$$t(x) \triangleq (i - (i \circ B))(x). \tag{5}$$

Similarly, the bottom-hat transform is calculated by subtracting the original image from the closed one (in order to obtain a positive result):

$$b(x) \triangleq ((i \bullet B) - i)(x). \tag{6}$$

Note that only the peaks (valleys) smaller than the structuring element can be detected by the top-hat (bottom-hat) transform. Therefore, we incorporate the scale of structuring element as an additional attribute. The multiscale top-hat and bottom-hat transforms can be defined as:

$$t(x,n) \triangleq (i - (i \circ nB))(x), \tag{7}$$

$$b(x,n) \triangleq ((i \bullet nB) - i)(x), \tag{8}$$

where n is an integer representing the scale factor of the structuring element. $nB(x)$ is obtained by dilating $B(x)$ recursively $n-1$ times with $B(x)$ itself [4]:

$$nB(x) = \underbrace{(B \oplus B \oplus B \cdots \oplus B)}_{n-1 \text{ times}}(x) . \tag{9}$$

In the top-hat transform, the extracted peaks are known as bright features; in the bottom-hat transform, dark features are found at valleys. Observing each pixel position the traced path through a scale space creates a feature curve, which serves as an aid in visualization of the feature variation. The scale-space behavior is discussed in the next section.

4 Feature based detection

4.1 Morphological scale-space behavior

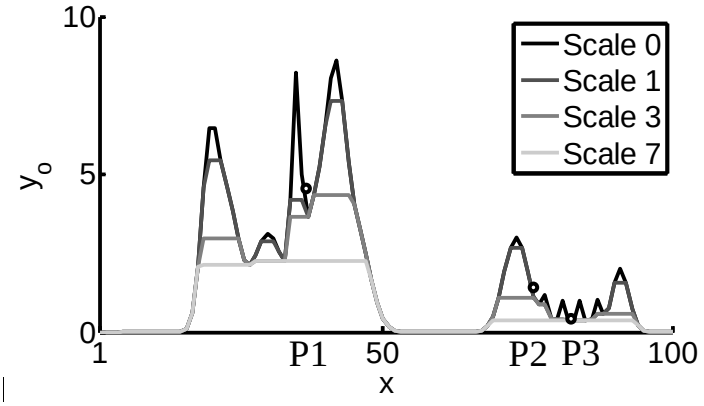
Fig. 2(a) shows an example of opening a discrete 1D signal with a series of line-like structuring elements. The length of the structuring element is sequentially selected from the set {3, 5, 7, 9, … ,.21}, i.e., the scale factor *n* ranges from 1 to 10. An opening smoothes the signal from below by cutting down its peaks. Thus, the opened signals can be viewed as cutting lines. The peaks are preserved in the top-hat scale space, Fig. 2(b). Analogously, a closing smoothes the signal from above by filling up its valleys.

In Fig. 2(a) the cutting line varies with the length of the structuring element. It drops off to a local minimum and then stops to change. The falling process can be restarted at some positions as lower local minima are involved in the filtering domain by enlarging the structuring element. The behavior of the cutting line leads to stair-like feature curves in Fig. 2(c). The increased amplitude of the bright features corresponds to the falling distance of the cutting line.

For a fixed signal point, only a limited number of local minima can be included in the investigated scale range. The shape of the bright feature curve is determined by the alignment of those relevant local minima which are lower than the observed signal point. Suppose that a signal point is near to the bottom of a valley. Three situations may appear: (1) if the relevant local minima are flatly arranged in the neighborhood, the bright feature curve includes few steps, like the signal point P2; (2) if the signal point is the lowest local minimum, such as P3,

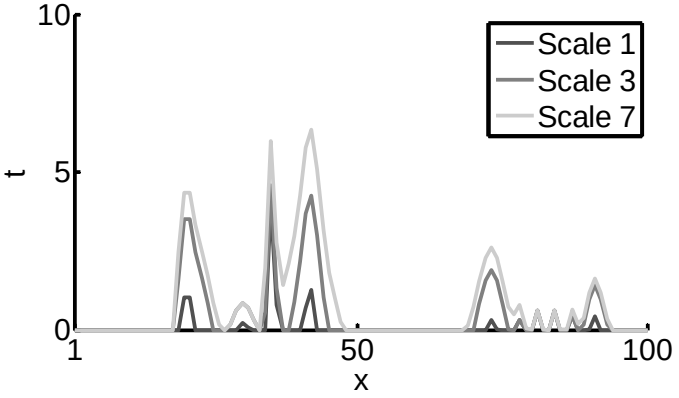
the bright feature is always equal to zero; (3) if the signal point is surrounded by the unevenly leveled local minima, its bright feature curve shows large steps similar to P1. The analog properties can also be observed in the bottom-hat scale space. The only difference lies in that local minima are replaced by local maxima in the discussion. The points far from local minima and maxima have no flat feature curves in both scale spaces.

(a) Opening



Scale 0 = original signal, P1 = 38, P2 = 83, P3 = 89.

(b) Top-hat transform



(c) Bright feature curve

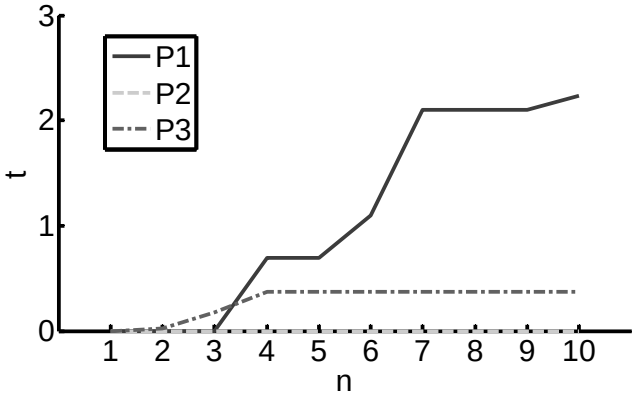


Fig. 2 Morphological scale spaces.

In summary, the shape of the bright feature curves describes the local signal flatness from below. The shape of the dark feature curves describes the local signal flatness from above. Combining both scale spaces enables a comprehensive description of the signal roughness.

4.2 Foreground segmentation

In the 2D case, we filter the image with disk-like structuring elements whose radii are increased from 1 to 20 pixels. Larger structuring elements are not necessary, since they tend to extract global features. Similar to the 1D case, the bright features of deep structures keep near zero in the top-hat scale space, whereas their dark features increase in the bottom-hat scale space. In the textured background, local minima and maxima are leveled randomly. Therefore, stair-like feature curves are created in both scale spaces. In addition, large areas of reflected light imply intensity peaks. Because of the duality of the top-hat and bottom-hat transforms, the scale-space behavior of shining areas is opposite to deep structures. Through the pairwise investigation of the top-hat and bottom-hat scale spaces, the surface components can be detected with a feature persistency measure:

$$m(x)=\frac{\max\limits_{n}\left\{\dfrac{\partial t}{\partial n}(x,n)\right\}-\max\limits_{n}\left\{\dfrac{\partial b}{\partial n}(x,n)\right\}}{\max\limits_{n}\left\{\dfrac{\partial t}{\partial n}(x,n)\right\}+\max\limits_{n}\left\{\dfrac{\partial b}{\partial n}(x,n)\right\}}. \tag{10}$$

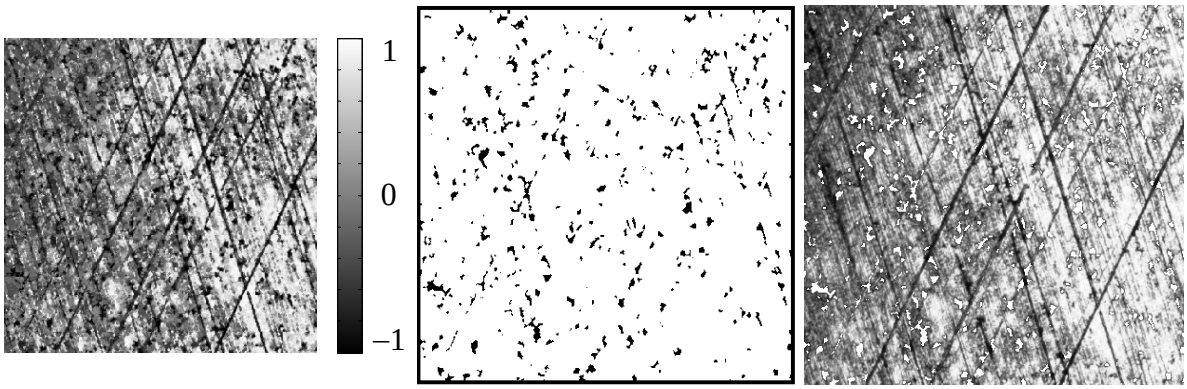
Considering the maximum slopes is advantageous for finding the most prominent variation in feature curves. As seen in Fig. 3(a), the values in foreground objects are distributed near -1, whereas the textured background has the values in the proximity of 0. The shining areas approximate to 1. Thus, the segmentation is based on a threshold of -0.5. In the postprocessing, the fragments with an area less than 5 pixels are removed as noise.

4.3 Groove elimination

The final detection of pores needs to eliminate grooves from the foreground. Groove detection follows the algorithm proposed in [5], which can accurately estimate groove parameters such as position, angle and width by iteratively using the Radon transform. The final result is demonstrated in Fig. 3(b).

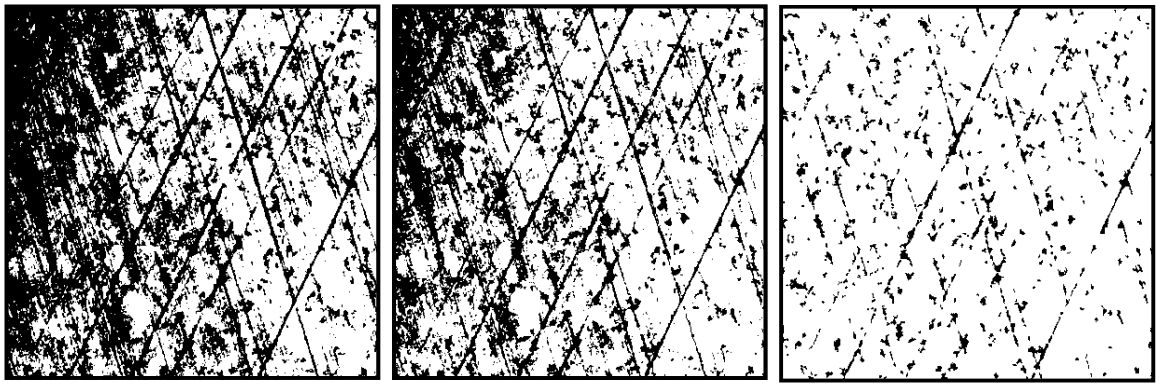
5 Experiment result

The performance analysis is divided into two parts. First, our method for the foreground segmentation is compared with the maximum entropy (ME) threshold [6] and the multilevel threshold method [7]. Fig. 4 illustrates that our method is superior to the histogram-based methods. The automatic thresholds are not capable to tackle the problem of textural disturbance. By comparison our method shows good robustness against inhomogeneous illumination and background textures. Second, we verify the correctness of the final detection by removing the pores from the original image. Fig. 3(c) illustrates that a small amount of pores overlapped with grooves are missed. This results in false negative errors. Moreover, false positive errors are visible in Fig. 3(b). Some fine groove segments remain in the pore map. Nevertheless, our method has provided a satisfactory accuracy.



(a) (b) (c)

Fig. 3 (a) Feature persistency of Fig. 1. (b) Pores. (c) Background and grooves.



(a) (b) (c)

Fig. 4 Foreground segmentation: (a) ME, (b) the darkest region with 3-level thresholds, (c) our method.

6 Conclusion

In this paper, a morphological approach for pore detection has been presented. We found that the feature curves derived from the top-hat and bottom-hat scale spaces were tightly related with image roughness. In the grey-level images, pores were shown as large valleys. This characteristic was exploited to reject other surface components like honing textures. Our method relaxed the dependence on image intensities by comparing a pair of scale spaces. Experiments manifested that the detection result can serve as a basis for an automated evaluation.

Literature

[1] Gehring Technologies GmbH: Laser structuring – Optimum Tribosystems through Functionally Optimized Surfaces, www.gehring.de, 2004.

[2] Werdner, A., Seewig, J., Reithmeier, E.: 3D Roughness Evaluation of Cylinder Liner Surfaces Based on Structure-oriented Parameters, Measurement Science and Technology 17, pp. 477-482, 2006.

[3] Maragos, P.: Morphological Filtering for Image Enhancement and Feature Detection, The Image and Video Processing Handbook (2nd ed.), pp. 135-156, Elsevier Academic Press, 2005.

[4] Mukhopadhyay, S., Chanda, B.: A Multiscale Morphological Approach to Local Contrast Enhancement, Signal Processing 80, pp. 685-696, 2000.

[5] Zhang, Q.P., Couloigner, I.: Accurate Centerline Detection and Line Width Estimation of Thick Lines Using the Radon Transform, IEEE Transaction on Image Processing, Vol. 16, No.2, pp. 310-316, 2007.

[6] Kapur, J.N., Sahoo P.K., Wong A.K.C.: A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram, Graphical Models and Image Processing 29 (3), pp. 273-285, 1985.

[7] Liao, P.S., Chen, T.S., Chung, P.C.: A Fast Algorithm for Multilevel Thresholding, Journal of Information Science and Engineering 17, pp. 713-727, 2001.

Probabilistische Verbesserung der Kalibrierung deflektometrischer Systeme

H. Rapp und C. Stiller

Karlsruher Institut für Technologie, Institut für Mess- und Regelungstechnik,
Engler-Bunte-Ring 21, D-76131 Karlsruhe, E-Mail: holger.rapp@kit.edu

Kurzfassung

Dieser Aufsatz befasst sich mit dem Problem der Feinkalibrierung eines deflektometrischen Versuchstandes zur drei-dimensionalen Vermessung voll reflektierender Oberflächen. Hierzu wird die Deflektometrie als drei-dimensionales Messverfahren erläutert, es wird dabei besonders auf das Verfahren der Stereodeflektometrie eingegangen. Das Problem der Kalibrierung wird anhand einer Simulation verdeutlicht und ein probabilistischer Ansatz vorgestellt der in der Simulation Fehler in der Kalibrierung erfolgreich herausrechnen kann. Der Aufsatz schließt mit einem Ausblick auf zukünftige Arbeiten.

1 Einleitung

Die Deflektometrie ist eine etablierte Methode um geometrische Fehlstellen auf voll reflektierenden Bauteilen zu detektieren. Sie findet Anwendung bei der Qualitätssicherung in vielen Bereichen der Industrie: Lackteile beim Automobil, Keramikgeschirr so wie Badamaturen können damit geprüft werden.

Das Messprinzip dabei ist simpel und basiert auf der Bestimmung der Neigung der zu untersuchenden Oberfläche. Bild 1 zeigt den schematischen Aufbau eines Deflektometriesystems: Ein Schirm oder Monitor zeigt ein definiertes Muster an, dieses wird von der Oberfläche reflektiert und das verzerrte Muster wird von der Kamera beobachtet. Durch aus der Interferometrie bekannten

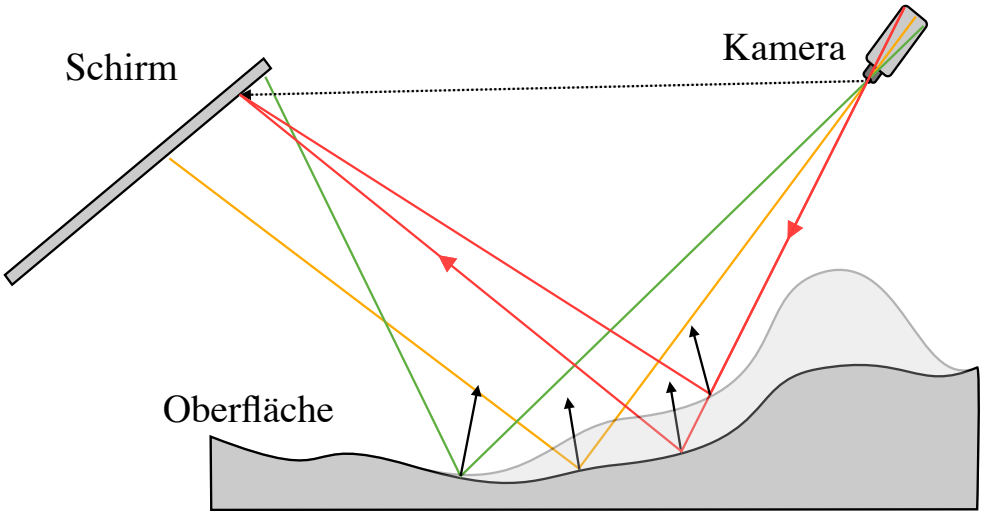


Bild 1: Messprinzip und Eindeutigkeitsproblem.

Phasenschubverfahren oder ähnlicher Kodierungsmechanismen kann eine Zuordnung zwischen beobachtetem Schirmpixel und Kamerapixel hergestellt werden; diese Abbildung wird einfache Abbildungsfunkion l genannt [1–3].

Die Deflektometrie ist bereits ein Industriestandard für die optische Sichtprüfung hoch reflektiver Teile. Hier interessiert jedoch oft nur die qualitative Aussage über die Oberfläche. Sofern gewisse Randbedingungen beim Design des Messaufbaus berücksichtigt werden kann die einfache Abbildungsfunktion als proportional zur lokalen Neigung der Oberfläche angenommen werden; ihre Ableitung ist demnach proportional zur lokalen Krümmung [2, 4]. Durch eine geschickte Visualisierung sind geometrische Defekte für den Menschen deutlich erkennbar zu machen. Eine quantitative Aussage über Tiefe und Ausprägung der Defekte ist allerdings so nicht möglich.

Der momentane Forschungsschwerpunkt liegt auf der hoch genauen drei-dimensionalen Vermessung reflektiver Teile mit deflektometrischen Methoden. Sind Kamera- und Schirmposition relativ zueinander bekannt, ist es möglich aus einer deflektometrischen Messung und Zusatzinformationen oder aus mehreren deflektometrischen Messungen die drei-dimensionale Geometrie des Targets zu errechnen. Hier sind reichhaltig Verfahren wie zum Beispiel Propagationsvefahren [5], Kombinationsansätze, die z.T. zumindest partiell diffus reflektierende Körper benötigen [6–8], und numerische Ausgleichsverfahren mehrerer Messungen [9] untersucht. Eine Übersicht über

die Funktionsweise verbreiteter Verfahren gibt [10].

Viele interessante Fragestellungen sind allerdings noch offen: Gibt es optimale Kamera- und Schirmpositionen? Was ist die minimale Anzahl an Messungen, die nötig sind, um eine bestimmte Form zu vermessen? Was ist die ideale Schirmgeometrie?

Grundvoraussetzung für jegliche quantitative Messung ist jedoch eine genaue Information über Schirm- und Kameraposition. Für deren Kalibrierung sind Verfahren Stand der Technik, die eine bekannte Targetgeometrie nutzen, zum Beispiel einen flachen Spiegel [11]. Diese Kalibrierung muss jedoch für jedes Schirm- und Kamerapositions-Paar neu durchgeführt werden. Eine Autokalibrierung des Messaufbaus ist daher wünschenswert. Der hier vorgestellte Ansatz konzentriert sich hierbei auf eine Verbesserung einer groben Vorkalibrierung, die bereits als gegeben vorausgesetzt wird.

Im nächsten Abschnitt wird vorerst von perfekten Messungen ausgegangen und das zum Einsatz kommende Rekonstruktionsverfahren erläutert.

2 Stereodeflektometrie

Gegeben sind nun zwei perfekt kalibrierte deflektometrische Messungen eines Objektes mit zwei verschiedenen relativen Kamera- und Schirmpositionen, aber überschneidenden Messvolumina. Eine drei-dimensionale Rekonstruktion kann nun mit Hilfe der Stereodeflektometrie erfolgen [12]. Dieses Verfahren ist in Bild 2 visualisiert.

Die beiden Messungen werden als induzierte potentielle Normalenfelder über dem jeweiligen Messvolumen M_i mit $i \in 1,2$ aufgefasst. D. h. jede Messung definiert für jeden Raumpunkt x,y,z im eigenen Messvolumen M_i einen potentiellen Normalenvektor, also die Normale die die Oberfläche hätte, würde sie in diesem Punkt liegen. Im gemeinsamen Messvolumen $M = M_1 \cap M_2$ erhält man gegensätzliche Aussagen in jedem Punkt, außer der Punkt liegt tatsächlich auf der gesuchten Oberfläche.

Um also die Oberfläche zu finden bietet es sich an, ein Ähnlichkeitsmaß S zwischen den beiden Normalen zu definieren und dieses auf einem kartesischen

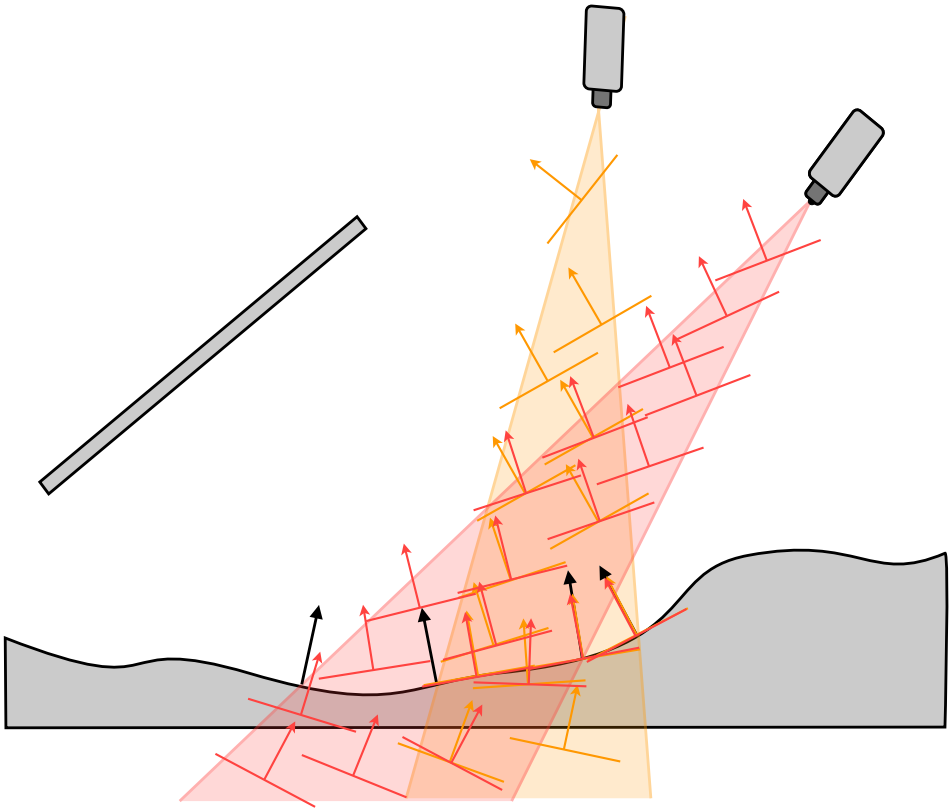


Bild 2: Funktionsprinzip der Stereodeflektometrie

Grid entlang einer Raumachse zu minimieren. Ein einfacher Vorschlag ist

$$S(n_1(x,y,z),n_2(x,y,z)) = 1 - |n_1(x,y,z) \cdot n_2(x,y,z)|. \tag{1}$$

Hierdurch erhält man eine Approximation der Raumpunkte der Oberfläche und deren Normalenvektoren. Durch eine Hermite Interpolation, beispielsweise mit radialen Basisfunktionen [13] oder auch durch Lösen der zugehörigen Poisson Gleichung erhält man eine Repräsentation der Oberfläche.

Zwischenergebnisse aus einer simulierten Rekonstruktion sind in Abbildungen 3 - 5 zu sehen. In Bild 6 ist die selbe Rekonstruktion gezeigt, hierbei ist allerdings bei der Kalibrierung einer Kamera ein Fehler von 2 ° simuliert worden. Dies zeigt eindrucksvoll wie stark ein kleiner Fehler in der Kalibrierung das Endergebnis verfälscht. Diese hohe Sensitivität der Deflektometrie kann man sich zu nutze machen um eine Verbesserung der Kalibrierung zu erreichen.

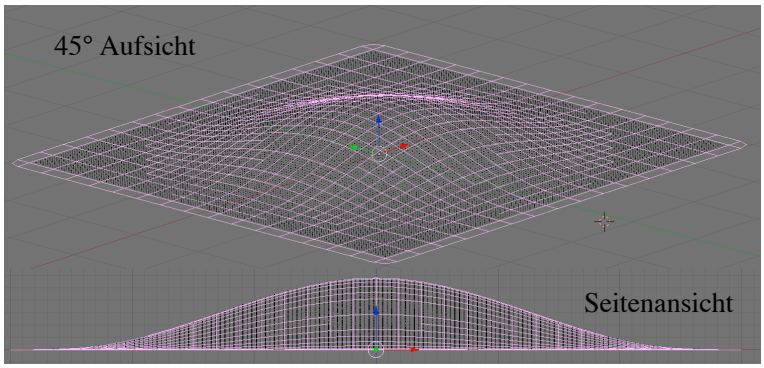


Bild 3: Orthografische Ansicht des Simulationstargets. Die Maße sind $6 \times 6 \times 0,6$ Einheiten.

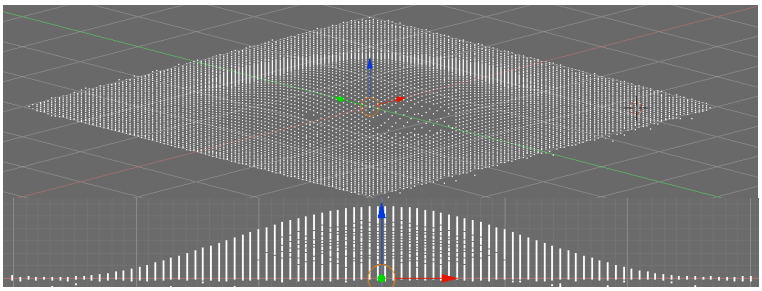


Bild 4: Punkte mit minimalen Normalenunterschied S auf einem Grid. Minimiert wurde entlang der z-Achse.

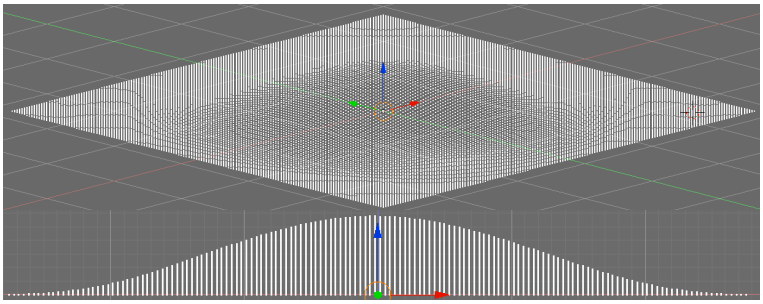


Bild 5: Endergebnis der Rekonstruktion. Die Normalen der Punkte aus Abbildung 4 wurden mit Hilfe von radialen Basisfunktionen interpoliert.

3 Verbesserung der Kalibrierung

Unsere Problemstellung kann also folgendermaßen formuliert werden: Gegeben ist eine Anzahl $m \geq 3$ von deflektometrischen Messungen, die als potentielle Normalenfelder in einem jeweiligen Messvolumen M_i aufgefasst werden. Alle Messungen sind dabei bezogen auf des selben raumfeste kartesische Koordinatensystem. Die Menge aller Messungen definieren wir als

$$N = \{n_1(x,y,z),\ n_2(x,y,z)\ \ldots\ n_m(x,y,z)\} \tag{2}$$

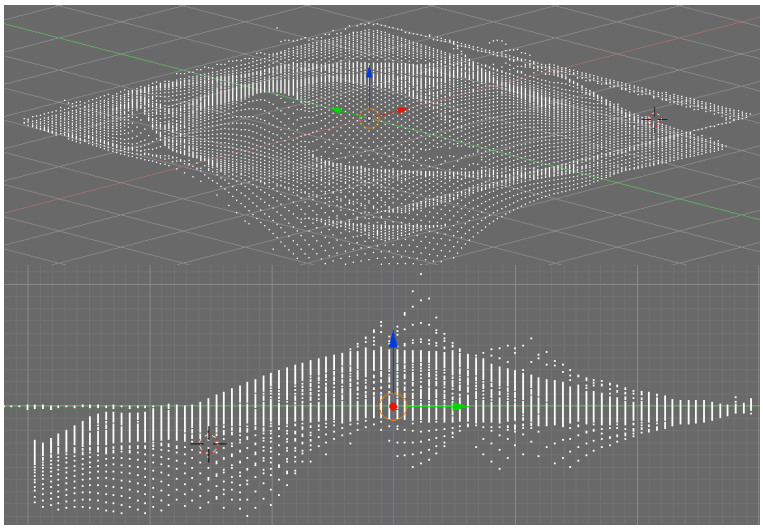


Bild 6: Wie Bild 4. Hier ist bei der Kalibrierung einer der beiden Kameras ein Fehler von 2 ° simuliert worden.

Außerdem sind die dazugehörigen Kamera- und Schirmpositionen und deren jeweiligen Sichtvektoren bis zu einer gewissen Genauigkeit bekannt. Modelliert werden sie als Normalverteilung mit Mittelwerten $x^C_{1\ldots m}$, $x^S_{1\ldots m}$, $t^C_{1\ldots m}$ und $t^S_{1\ldots m}$ und Varianzen $\sigma^2_{x^C_{1\ldots m}}$, $\sigma^2_{x^S_{1\ldots m}}$, $\sigma^2_{t^C_{1\ldots m}}$ und $\sigma^2_{t^S_{1\ldots m}}$. Die Oberfläche wird als Normalverteilung über das Höhenprofil auf einem kartesischen Grid modelliert

$$z(x,y) = \mathcal{N}(\mu_z(x,y), \sigma^2_z(x,y)) \tag{3}$$

Der Algorithmus arbeitet nun in zwei Schritten. Im ersten Schritt werden mit den momentanen Parametern durch paarweises Vergleichen der Normalenfeldern - wie in Abschnitt 2 beschrieben - aktuelle Vorschläge für die Oberfläche gefunden. Dies entspricht einem Samplen aus der Population der gesamten möglichen Oberflächen, wobei die Parameter als Beobachtungen gesehen werden. Mit der empirischen Varianz und dem Mittelwert dieser Sampels wird die Verteilungsdichte für z upgedated.

Nun wird zufällig ein Schirm oder eine Kamera einer Messung ausgewählt und dort dann auch zufällig die Position oder der Sichtvektor ausgesucht. Von dem ausgewählten Wert werden Mittelwert und 2-σ-Punkte gesampelt und der erste Schritt mit diesen neuen Werten wiederholt. Derjenige Punkt, der die kleinste Summe über die Varianzen der Oberfläche

$$E = \sum_{x,y \in \bigcup_{i=1}^{m} M_i} \sigma^2_z(x,y) \tag{4}$$

ergibt wird als neuer Mittelwert festgelegt.

Diese zwei Schritte werden so lange wiederholt bis E eine festgelegte Schranke unterschreitet.

In der obigen Simulation bringt der Algorithmus gute Ergebnisse: bei $m = 3$ und einem Kalibrationsfehler von maximalen 1 cm in der Position bzw. 3° im Winkel ergibt sich mit diesem Verfahren die korrekte Geometrie nach Minimierung von E. Die gewählten Kalibrationsfehler sind deutlich konservativer als jene, die von modernen Handhabungsgeräten erreicht werden können. Experimentelle Ergebnisse stehen allerdings noch aus.

4 Weiterführende Arbeiten

Die Simulationsergebnisse lassen hoffen, dass dieser Ansatz in der Praxis funktionieren kann. Dadurch wird es möglich, schnell eine Vielzahl von deflektometrischen Messungen mit verschiedenen Kamera- und Schirmpositionen aufzunehmen und so auch komplexe Geometrien deflektometrisch zu vermessen. Ein entsprechender Versuchsaufbau ist vorhanden und praktische Experimente sind als nächster Schritt geplant.

Der momentan eingesetzte Algorithmus zur Optimierung der Schirm und Kamerapositionen ist $O(m!)$ mit der Anzahl an deflektometrischen Messungen m. Eine Reduktion der Komplexitätsklasse scheint möglich indem man die Anzahl an Samples für z in Schritt eins reduziert und eine höhere empirische Varianz akzeptiert. Ein RANSAC Sampling bietet sich hier an. Eine reduzierte Komplexität erlaubt mehr Messungen in die Optimierung mit aufzunehmen, was geringere Fehler im Endergebnis zu erwarten lässt. Ob hier tatsächlich Verbesserungen zu erreichen sind soll noch untersucht werden.

Literatur

[1] D. Pérard, *Automated Visual Inspection of Specular Sufaces with Structured-Lighting Reflection Techniques*. Dissertation, Universität Karlsruhe (TH), Karlsruhe, 2001. Fortschritt-Berichte, Reihe 8, Nr. 869, VDI-Verlag, Düsseldorf.

[2] S. Kammel, *Deflektometrische Untersuchung spiegelnd reflektierender Freiformflächen*. Dissertation, Universität Karlsruhe (TH), Karlsruhe, 2005. Schriftenreihe Institut für Mess- und Regelungstechnik, Universitätsverlag Karlsruhe, Nr. 004.

[3] M. Petz und R. Tutsch, "Rasterreflexions-photogrammetrie zur messung spiegelnder oberflächen", *TM - Technisches Messen* (71), S. 389–397, 2004.

[4] S. Kammel und F. P. Leon, "Deflectometric measurement of specular surfaces", *IEEE Transactions on Instrumentation and Measurement* **57**, S. 763–769, Apr. 2008.

[5] S. Kammel und J. Horbach, "Topography reconstruction of specular surfaces", in *Videometrics VIII*, J.-A. Beraldin, S. F. El-Hakim, A. Gruen und J. S. Walton (Hrsg.), **5665**(1), S. 59–66, SPIE, 2005.

[6] J. Balzer, S. Werling und J. Beyerer, "Deflektometrische Rekonstruktion teilspiegelnder Freiformflächen (Deflectometric Reconstruction of Partially Specular Free-Form Surfaces)", *tm-Technisches Messen* **74**(11), S. 545–552, 2007.

[7] K. Donner, "Stereo reconstruction of reflecting surfaces", *Pattern Recognition and Image Analysis* **14**(2), S. 168–172, 2004.

[8] R. Kickingereder, *Optische Vermessung partiell reflektierender Oberflächen*. PhD thesis, Universität Passau, November 2006.

[9] J. Balzer, S. Werling und J. Beyerer, "Regularization of the deflectometry problem using shading data", *Two- and Three-Dimensional Methods for Inspection and Metrology IV* **6382**(1), S. 63820B, SPIE, 2006.

[10] H. Rapp und C. Stiller, "Deflektometrische Methoden zur Sichtprüfung und 3D-Vermessung voll reflektierender Freiformflächen", in *Forum Bildverarbeitung 2010, Regensburg*, S. 217, KIT Scientific Publishing, 2010.

[11] S. Hofer, S. Werling und J. Beyerer, "Neuartige Strategie zur vollständigen Kalibrierung eines Sensorsystems zur automatischen Sichtprüfung spiegelnder Oberflächen", in *Forum Bildverarbeitung 2010, Regensburg*, S. 25, KIT Scientific Publishing, 2010.

[12] M. C. Knauer, *Absolute Phasenmessende Deflektometrie*. PhD thesis, Universität Erlangen-Nürnberg, Mai 2006.

[13] S. Ettl, J. Kaminski und G. Häusler, "Generalized hermite interpolation with radial basis functions considering only gradient data", in *Curve and Surface Fitting: Avignon 2006*, S. 141–149, Nashboro Press, 2007.

Materialidentifikation mittels optisch realisierter Kreuzkorrelation der Reflektanzspektren

Miro Sauerland, Robin Gruna, Jürgen Beyerer

Karlsruher Institut für Technologie, Fakultät für Informatik, Institut für Anthropomatik, Lehrstuhl für Interaktive Echtzeitsysteme, 76187 Karlsruhe, E-Mail: miro.sauerland@kit.edu, robin.gruna@kit.edu, Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung, 76131 Karlsruhe, E-Mail: juergen.beyerer@iosb.fraunhofer.de

Kurzfassung

Dieser Artikel stellt ein neuartiges, zum Patent angemeldetes Messverfahren [1] vor, welches die Ähnlichkeit zweier Spektren in Form eines normierten Kreuzkorrelationskoeffizienten (nKK) bemisst. Die Besonderheit liegt in der optische Realisierung des Verfahrens. Auf Grundlage des nKK wird eine Messgleichung entwickelt, die dann zur Herleitung des Signal-zu-Rausch Verhältnisses genutzt wird. Als Experiment wird die Aufgabenstellung behandelt, metamere Farben anhand ihrer Spektren zu unterscheiden. Gewonnene Messergebnisse werden abschließend mit einem vorhergesagten Ergebnis verglichen.

1 Einleitung

Ein wichtiger Aspekt des maschinellen Sehens ist die Aufgabe Objekte zu identifizieren. Neben Farbe und Textur [2] [3] können auch winkelabhängige Reflektanzeigenschaften [4] für eine Materialklassifikation genutzt werden. Das Reflektanzspektrum einer Oberfläche, auch als „Fingerabdruck“ bezeichnet [5], ist dabei ein sehr diskriminatives Merkmal. Zur Gewinnung dieser Information ist jedoch ein bildgebendes Spektrometer notwendig, welches einige Nachteile mit sich bringt:

- Komplexer optischer Aufbau und damit teuer

- Keine flächenhafte Erfassung einer Szene
- Frameraten von wenigen 100 Zeilen pro Sekunde

Für konkrete Anwendungen der Materialidentifikation kann auf eine dichte Abtastung des Spektrums verzichtet werden [6], da nicht alle Wellenlängen gleichermaßen informationstragend sind. Die Wahl geeigneter Stützwellenlängen, z.B. durch wenige schmalbandige Interferenzfilter, ist jedoch nicht eindeutig und muss für jede Aufgabenstellung neu beantwortet werden.

Einen weiteren Ansatz zur Auswertung des Spektrum stellt die Anwendung sogenannter *Multivariater Optischer Elemente, MOE* [7] oder das sogenannte *Molecular Factor Computing, MFC* [8] dar. Hierbei werden breite Filterfunktionen realisiert; bei MOEs in Form von Interferenzfiltern, im Falle von MFC als flüssige Lösungen. Eine Filterkurve wird dabei als Vektor aufgefasst, der mit dem Spektrum komponentenweise multipliziert wird. In Kombination mit einem breitbandigen Sensor kann der Messwert als das Skalarprodukt aus Filtervektor und Reflektanzspektrum [9] interpretiert werden. Für eine konkrete Aufgabenstellung werden dann Filterkurven realisiert, die z.B. den Eigenvektoren einer Hauptkomponentenanalyse entsprechen [10]. Nachteilig bei diesem Vorgehen ist der hohe technische Aufwand zum Design geeigneter Filterkurven. Einen Überblick über diese Ansätze ist in [11] zu finden.

Der in diesem Artikel vorgestellte Ansatz für die Materialerkennung folgt im wesentlichen dem Grundgedanken einer speziell designten Filterfunktion. Der Hauptvorteil des vorgestellten Aufbaus liegt in der einfachen Adaptierbarkeit an eine Aufgabenstellung, da das Reflektanzspektrum des Materials selbst als Filterfunktion eingesetzt wird.

2 Messprinzip

Die Problemstellung, die gelöst wird, ist ein unbekanntes Material anhand seines Spektrums zu identifizieren. Die Identifikation einer Probe erfolgt durch einen Vergleich mit bekannten Referenzspektren. Somit wird die Aufgabenstellung dahingehend eingeschränkt, dass das unbekannte Material sicher einem der Referenzmaterialien entspricht. Als Messung wird dazu die Ähnlichkeit

des Spektrums s_u des unbekannten Materials kombinatorisch mit allen Referenzspektren s_r bestimmt. Als Ähnlichkeitsmaß wird der normierte Kreuzkorrelationskoeffizient ϕ_{ur} genutzt, welcher die Ähnlichkeit in Form einer Zahl $(0 \leq \phi_{ur} \leq 1)$ bemisst, wobei 1 der maximalen Ähnlichkeit entspricht, bzw. dem normierten Autokorrelationkoeffizienten (nAK). Der nKK ist definiert als:

$$\phi_{ur} := \frac{\int s_u(\lambda) s_r(\lambda)\, \mathrm{d}\lambda}{\sqrt{\int s_u^2(\lambda)\, \mathrm{d}\lambda}\sqrt{\int s_r^2(\lambda)\, \mathrm{d}\lambda}}. \tag{1}$$

Die Normierung stellt sicher, dass das Ergebnis ϕ_{ur} invariant gegenüber skalierten Spektren ist. Somit ist das vorgestellte Verfahren unabhängig gegenüber Helligkeitsschwankungen bei zu untersuchenden Proben.

2.1 Messaufbau

Anstatt das Spektrum des unbekannten Materials mittels Spektrometer zu vermessen und die Berechnung mittels Mikroprozessor durchzuführen, wird im Folgenden ein Aufbau vorgestellt (Abbildung 1), welcher die Kreuzkorrelation optisch realisiert. Aus Gleichung (1) werden die für eine optische Realisierung wesentlichen mathematischen Operatoren ersichtlich: der Integration über λ und die Multiplikation der Spektren. Die Integration leistet ein breitbandiger CCD-Sensor. Um die beiden Spektren zu multiplizieren, werden die Proben optisch scharf aufeinander abgebildet. Hierbei wird ausgenutzt, dass dem Licht multiplikativ das Reflektanzspektrum des Materials aufgeprägt wird, wenn es an dessen Oberfläche reflektiert wird.

2.2 Messvorgang

Zur einfacheren Argumentation im Sinne von Licht als Photonentransport werden kontinuierliche, von λ abhängige Variablen diskretisiert. Dazu wird der betrachtete Wellenlängenbereich in Intervalle $[\lambda_i - \frac{\Delta}{2}, \lambda_i + \frac{\Delta}{2})$ aufgeteilt (vgl. Abbildung 2). Das diskrete Reflektanzspektrum ergibt sich somit zu $s(\lambda_i) := \int_{\lambda_i - \frac{\Delta}{2}}^{\lambda_i + \frac{\Delta}{2}} s(\lambda)\mathrm{d}\lambda$. Des Weiteren wird das Kameramodell

$$g = \sum_i n(\lambda_i) q(\lambda_i) \tag{2}$$

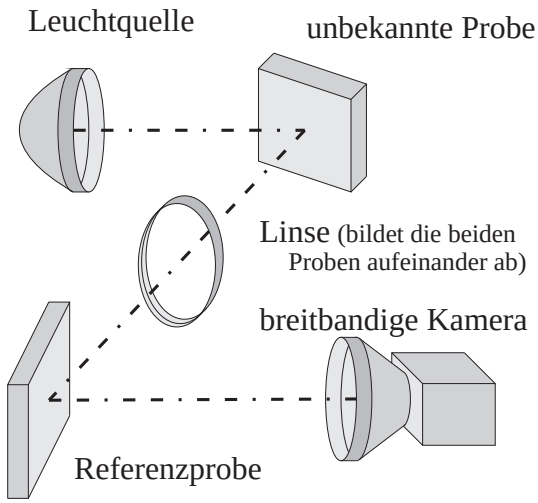


Bild 1: Prinzip einer Messanordnung.

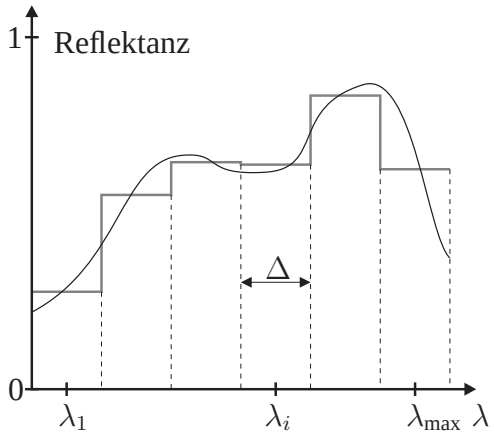


Bild 2: Diskretisierung der Spektren.

eingeführt, welches den Grauwert g auf die Anzahl Photonen $n(\lambda_i)$ zurückführt, welche gemäß der Sensorempfindlichkeit $q(\lambda_i)$ auch ein Elektron-Loch-Paar erzeugen. Rauschterme sind nicht Gegenstand dieser Betrachtung.

Zum Identifizieren eines unbekannten Materials muss die Ähnlichkeit und damit der nKK in Kombination mit jedem relevanten Referenzmaterial bestimmt werden. In Gleichung (1) wurde der Einfluss der spektralen Ausprägung der Beleuchtung $n(\lambda_i)$, bzw. der Sensorempfindlichkeit $q(\lambda_i)$ nicht betrachtet. $n(\lambda_i)$ beschreibt die Anzahl Photonen die im Intervall $[\lambda_i - \frac{\Delta}{2}, \lambda_i + \frac{\Delta}{2})$ emittiert werden. Die Berücksichtigung dieses Einflusses führt zu der folgenden Messgleichung (3), welche aus dem nKK (1) unter Berücksichtigung des Kameramodells (2) hervorgeht:

$$m_{ur} := \frac{\sum_i n(\lambda_i) s_u(\lambda_i) s_r(\lambda_i) q(\lambda_i)}{\sqrt{\sum_i n(\lambda_i) s_r^2(\lambda_i) q(\lambda_i)}}. \tag{3}$$

Der Zähler beschreibt das Messsignal, welches in quantisierter Form vom Kamerasensor ausgelesen wird. Neben dem Verhältnis von reflektierter Leistung zu eingebrachter Leistung, kann das Reflektionsspektrum $s(\lambda_i)$ auch als die Wahrscheinlichkeit angesehen werden, dass ein auftreffendes Photon auch reflektiert wird. Somit kann das Produkt $s_u(\lambda_i) s_r(\lambda_i) q(\lambda_i)$ als die Wahrscheinlichkeit interpretiert werden, dass ein emittiertes Photon auch ein Elektron-Loch-Paar erzeugt. Die beiden Wurzeln in (1) stellen Normierungskonstanten dar, die den $\phi_{ur} \in [0, 1]$ normieren. Gemäß der Aufgabenstellung ist eines der

Spektren unbekannt und somit fehlt dessen Normierungskonstante in der Messgleichung (3). Die fehlende Normierung skaliert das Ergebnis auf unbekannte Weise. Bleibt jedoch die Aufnahmekonstellation der unbekannten Probe unverändert, so werden alle Messungen in gleicher Weise skaliert und die Vergleichbarkeit bleibt erhalten. Eine Messung setzt sich somit aus einem Grauwert der Kamera und einer softwareseitigen Normierung auf das verwendete Referenzmaterial zusammen. Wie der Kalibrierungsprozess aussieht, in dem die Normierungskonstanten bestimmt werden, ist in Abschnitt 4.2 erläutert.

Eine maximale Messung identifiziert dann ein unbekanntes Spektrum als das der entsprechenden Referenzprobe. Hierbei wird ausgenutzt, dass der nAK als Grenzfall des nKK immer maximal wird. Dass die Messgleichung (3) den Gesetzmäßigkeiten des nKK folgt, wird sofort ersichtlich, wenn anstatt $s_j(\lambda)$ das Spektrum $s'_j(\lambda) = \sqrt{n(\lambda)}\sqrt{q(\lambda)}s_j(\lambda)$ in Gleichung (1) eingesetzt wird. Diese Betrachtung zeigt auch, welchen Einfluss der spektrale Verlauf der Beleuchtung bzw. der Sensorempfindlichkeit auf das Ergebnis des nKK hat.

3 Signal-Rausch-Verhältnis

Werden die Spektren der verschiedenen zu unterscheidenden Materialien immer ähnlicher, wird das Verfahren ab einem gewissen Punkt versagen. Eine quantitative Aussage darüber soll das im Folgendem entwickelte Signal-Rausch-Verhältnis (SNR)[1] liefern.

Das SNR wird für ein zwei Klassenproblem s_A, s_B motiviert, für den Fall dass s_A als unbekanntes Spektrum vorliegt. Zur Identifikation dieser Probe sind somit zwei Messungen notwendig:

$$m_{AA} = \frac{\sum_i n(\lambda_i)s_A(\lambda_i)s_A(\lambda_i)q(\lambda_i)}{\sqrt{\sum_i n(\lambda_i)[s_A(\lambda_i)]^2 q(\lambda_i)}}, \quad m_{AB} = \frac{\sum_i n(\lambda_i)s_A(\lambda_i)s_B(\lambda_i)q(\lambda_i)}{\sqrt{\sum_i n(\lambda_i)[s_B(\lambda_i)]^2 q(\lambda_i)}}.$$

Eine Messung wird dabei durch vier Rauschquellen gestört: Das Photonenrauschen der Lichtquelle, die zufällige Auswahl von Photonen z.B. bei der Reflektion, das Rauschen der Kameraelektronik und die Quantisierung. Die

[1] SNR steht für Signal-to-Noise Ratio.

kameraspezifischen Rauscheinflüsse werden aus Platzgründen hier nicht behandelt; für ein tieferes Verständnis sei auf [12] verwiesen.

Wird eine thermische, breitbandige Beleuchtung eingesetzt, so schwankt die Anzahl emittierter Photonen pro Zeitintervall gemäß einer Poisson-Statistik $P(\overline{n})$. Diese Aussage ist sowohl gültig wenn alle Photonen interessieren, als auch sowie wenn nur die Anzahl Photonen $n(\lambda_i)$ im i-ten Intervall betrachtet werden. Zur Modellierung wird die Anzahl Photonen als Zufallsvariable $n(\lambda_i) \sim P(\overline{n}(\lambda_i))$ eingeführt. Gemäß der Poisson-Verteilung ist der Erwartungswert $\mathrm{E}\{n(\lambda_i)\} = \overline{n}(\lambda_i)$ gegeben durch die mittlere Anzahl Photonen, bzw. die $\mathrm{Var}\{n(\lambda_i)\} = \overline{n}(\lambda_i)$. Handelt es sich um die mittlere Anzahl Photonen während der Messung, wird diese mit $\overline{n}_0(\lambda_i)$ bezeichnet, sowie mit $\overline{n}_k(\lambda_i)$, wenn die Kalibrierung mit einer anderen Anzahl Photonen durchgeführt wurde.

Der zweite Rauscheinfluss, der hier betrachtet werden soll, beschreibt einen zufälligen Auswahlprozess. Ein emittiertes Photon erzeugt mit der mittleren Wahrscheinlichkeit[2] $s_{u,i}s_{r,i}q_i$ ein messsignalwirksames Elektron-Loch-Paar. Diese zufällige und stochastisch unabhängige Auswahl kann durch einen Bernoulli-Prozess modelliert werden [12]. Durchläuft ein Poisson-verteilter Photonenstrom solch einen Bernoulli-Prozess, so bleibt die Verteilung erhalten. Photonen, die auch ein Elektron-Loch-Paar erzeugen, sind folglich $n_i \sim P(\overline{n}_{0,i}s_{u,i}s_{r,i}q_i)$ verteilt.

Das Signal-zu-Rausch-Verhältnis wird wie folgt eingeführt:

$$\mathrm{SNR} := \frac{[\mathrm{E}\{m_{AA} - m_{AB}\}]^2}{\mathrm{Var}\{m_{AA} - m_{AB}\}} = \frac{[\mathrm{E}\{m_{AA}\} - \mathrm{E}\{m_{AB}\}]^2}{\mathrm{Var}\{m_{AA}\} + \mathrm{Var}\{m_{AB}\}}. \tag{4}$$

Hierbei wurde ausgenutzt, dass die Messungen als stochastisch unabhängig angesehen werden können, z.B. indem sie zu unterschiedlichen Zeitpunkten erfolgen. Im Folgenden werden Signal und Rauschen getrennt voneinander betrachtet. Signal:

[2] Auf die Abhängigkeiten der Variablen von (λ) wird im Folgendem zur besseren Übersichtlichkeit verzichtet. Folgende Schreibweisen sind äquivalent $s_u(\lambda_i) = s_{u,i}$.

$$[\mathrm{E}\{m_{AA} - m_{AB}\}]^2 = \left[\frac{\sum_i \mathrm{E}\{n_{1,i}\} s_{A,i} s_{A,i} q_i}{\sqrt{\sum_i \overline{n}_{k,i} [s_{A,i}]^2 q_i}} - \frac{\sum_i \mathrm{E}\{n_{2,i}\} s_{A,i} s_{B,i} q_i}{\sqrt{\sum_i \overline{n}_{k,i} [s_{B,i}]^2 q_i}}\right]^2 \quad (5)$$

$$= \left[\left(\frac{\sum_i \overline{n}_{0,i} s_{A,i} s_{A,i} q_i}{\sqrt{\sum_i \overline{n}_{k,i} \left(\frac{\overline{n}_{0,i}}{\overline{n}_{0,i}}\right) [s_{A,i}]^2 q_i}} - \frac{\sum_i \overline{n}_{0,i} s_{A,i} s_{B,i} q_i}{\sqrt{\sum_i \overline{n}_{k,i} \left(\frac{\overline{n}_{0,i}}{\overline{n}_{0,i}}\right) [s_{B,i}]^2 q_i}}\right) \frac{\sqrt{\sum_i \overline{n}_{0,i} [s_{A,i}]^2 q_i}}{\sqrt{\sum_i \overline{n}_{0,i} [s_{A,i}]^2 q_i}}\right]^2$$

Dieser Ausdruck kann vereinfacht werden, wenn das Verhältnis $\frac{\overline{n}_{k,i}}{\overline{n}_{0,i}} = f \; \forall i$ in allen Intervallen i den gleichen Wert annimmt. Ein von i unabhängiger Faktor lässt sich somit aus der Summe ausklammern. $\overline{n}_i$ beschreibt den Verlauf der erwarteten mittleren Photonenzahl über die diskreten Wellenlängenintervalle λ_i. Die Forderung $\frac{\overline{n}_{k,i}}{\overline{n}_{0,i}} = f \; \forall i$ ist dann erfüllt, wenn sich der Verlauf von $\overline{n}_{k,i}$ und $\overline{n}_{0,i}$ nur in einem skalaren Faktor f unterscheidet. Am einfachsten wird diese Annahme erfüllt, wenn die Beleuchtungsparameter zwischen Kalibrierung und Messbetrieb nicht verändert werden. Mit den beiden Erweiterungen kann der obige Ausdruck dann auf den nKK zurückgeführt werden:

$$[\mathrm{E}\{m_{AA} - m_{AB}\}]^2 = \frac{\overline{n}_{0,i}}{\overline{n}_{k,i}} (1 - \phi_{AB})^2 \sum_i \overline{n}_{0,i} [s_{A,i}]^2 q_i. \quad (6)$$

Mit einem analogen Vorgehen kann das Rauschen in folgender Weise auf den nKK zurückgeführt werden. Rauschen:

$$\mathrm{Var}\{m_{AA} - m_{AB}\} = \sqrt{\frac{\overline{n}_{0,i}}{\overline{n}_{k,i}}} (1 + \phi_{AB}) \sqrt{\sum_i \overline{n}_{0,i} [s_{A,i}]^2 q_i}. \quad (7)$$

Die Zwischenergebnisse für Signal und Rauschen führen dann zu folgendem SNR:

$$\mathrm{SNR}(\phi_{AB}) = \sqrt{\frac{\overline{n}_{0,i}}{\overline{n}_{k,i}}} \overbrace{\frac{(1 - \phi_{AB})^2}{1 + \phi_{AB}}}^{(2)} \overbrace{\sqrt{\sum_i \overline{n}_{0,i} [s_{A,i}]^2 q_i}}^{(3)}. \quad (8)$$

Die erste Wurzel aus dem Verhältnis der Menge Photonen während der Kalibrierung zur Messsituation wird in der Praxis zu 1 gewählt. Zum einen muss die Farbtemperatur der Beleuchtung gleichbleiben, zum anderen wird die Kalibrierung selbst mit maximalem SNR durchgeführt und somit mit einer möglichst hohen Anzahl Photonen $\overline{n}_k$. Der zweite Term beschreibt die Auswirkung von ϕ_{AB} auf das SNR, dargestellt in Abbildung 4. Dem Verlauf kann entnommen werden, dass der zweite Term für immer ähnlichere Spektren ($\phi_{AB} \rightarrow 1$) bis auf Null abfällt. Identische Spektren haben somit einen SNR von 0 und können erwartungsgemäß nicht unterschieden werden. Der dritte Term bestraft unbekannte Spektren (hier s_A), welche viel Licht absorbieren. Eine Verdopplung der Photonenanzahl verbessert das SNR um den Faktor $\sqrt{2}$.

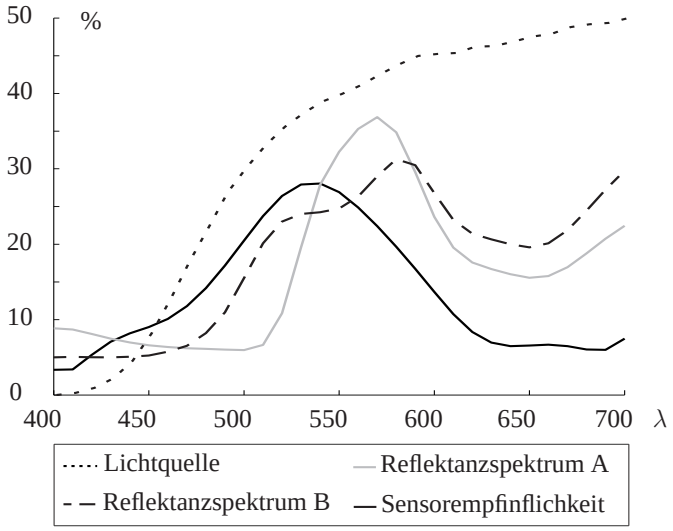


Bild 3: Verschiedene spektrale Verläufe; der Verlauf der Beleuchtung ist rein qualitativ aufgetragen.

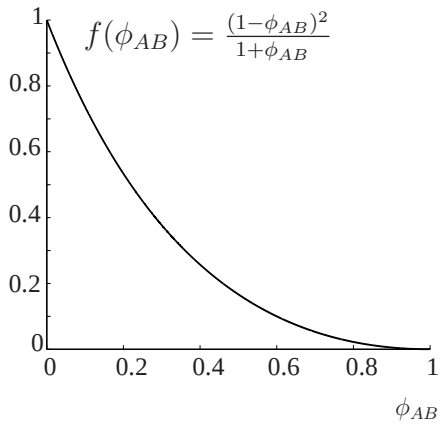


Bild 4: Einfluss des nKK auf das SNR.

4 Experiment

Ziel des Experiments ist es zu zeigen, dass Proben anhand ihres Reflektanzspektrums identifiziert werden können, indem der nKK in beschriebener Weise gebildet wird. Neben einer Messung wird das Ergebnis auch simuliert und verglichen. Als Aufgabenstellung sollen zwei metamere Farbproben A und B unterschieden werden. D.h. die Farbproben, die Beleuchtung und die Kamera sind so gewählt, dass sich bei direkter Betrachtung ein identischer Grauwert einstellt. Somit ist sichergestellt, dass eine Unterscheidung nur dann möglich

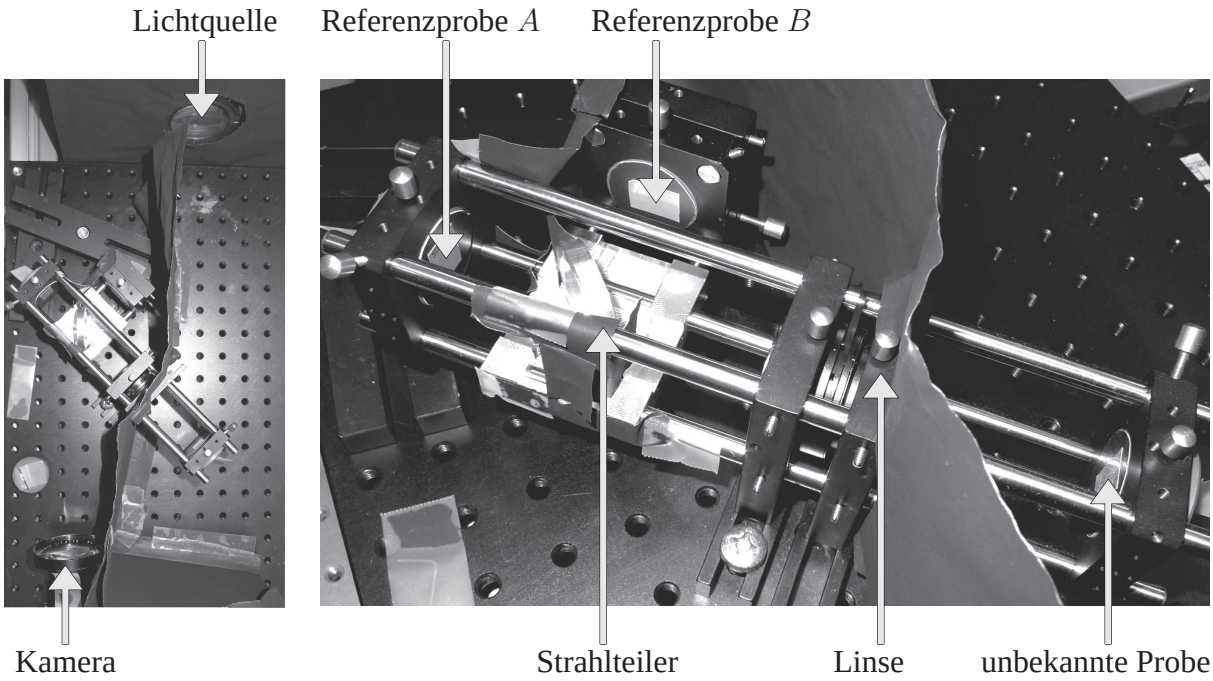


Bild 5: Experimenteller Aufbau mit Strahlteiler und zwei Referenzproben.

ist, wenn die Information der unterschiedlichen Reflektanzspektren ausgenutzt wird.

4.1 Experimenteller Aufbau

Der Aufbau ist in Abbildung 4 dargestellt. Durch die Verwendung eines Strahlteilers[3] können beide Referenzmaterialien gleichzeitig betrachtet werden. Damit die Farbproben metamer erscheinen, müssen Beleuchtung und Kamera auf die Farbproben abgestimmt sein. Als Leuchtquelle wird ein Halogenstrahler[4] eingesetzt der nach Normlicht D50 spezifiziert ist. Im Beleuchtungsstrahlengang ist zusätzlich ein Glasfilter[5] eingebracht. Die Farbproben sind jeweils einem Streifen eines Farbtemperatur Indikatorstreifens[6] entnommen, wie sie in der Druckindustrie zur Beleuchtungskontrolle eingesetzt werden. Die Farbkamera[7] wird als Grauwertkamera genutzt, indem die einzelnen RGB-Kanäle entsprechend der Gewichtungsfaktoren $\{0.16, 0.68, 0.16\}$ gewichtet werden. Die Kennlinien wurden aus Datenblättern zusammengetragen und die Farbproben

[3] Polka Dot Strahlteiler mit konstantem Teilungsverhältnis für unterschiedliche Wellenlägen; Edmund Optics.
[4] SoLux 4700 Kelvin, betrieben bei 14V sodass 5000K Farbtemperatur erreicht wird.
[5] Schott Glasfilter GG10.
[6] UGRA Metameriekarte 5000K.
[7] Stingray F-146C von Allied Vision Technologies GmbH.

mit einem Spektrometer[8] vermessen. Auf Basis dieser Daten wird ein nKK von $\phi_{AB} = 0.96$ vorhergesagt. Im Vergleich zu einer fiktiven Aufgabenstellung in es gilt Rot und Blau als Proben zu unterschieden, ist die Unterscheidung der vorgestellten metamere Farben deutlich schwieriger. Unter der Annahme die Farben Rot und Blau wären maximal unähnlich ($\phi_{\text{Rot,Blau}} = 0$), muss für dieses Experiment ($\phi_{AB} = 0.96$) mit einem um Faktor 1000 geringeren SNR gerechnet werden. Der Unterschied ergibt sich, wenn $\phi_{\text{Rot,Blau}}$ und ϕ_{AB} in den zweiten Term des SNR eingesetzt werden (Gleichung (8)).

4.2 Durchführung

Der eigentlichen Messung ist ein einmaliger Kalibrierprozess vorgeschaltet. In diesem werden die Kalibrierkonstanten bestimmt, wie sie im Nenner der Messgleichung (3) angegeben sind. Dazu wird die unbekannte Probe identisch zur Referenzprobe gewählt und bis auf eine Wurzel-Operation entspricht der gemessene Grauwert direkt der Normierungskonstante. Dieser Zusammenhang wird ersichtlich wenn die Normierungskonstanten mit dem Kameramodell (2) verglichen werden. Die Konstante wird als Mittelwert aus 100 Messungen gewonnen und somit näherungsweise als rauschfrei betrachtet.

Da das Experiment als unbekannte Probe nur zwei Möglichkeiten A und B zulässt, kann die Auswertung sehr anschaulich durch die Verhältnisbildung $\frac{m_{uB}}{m_{uA}}$ der Einzelmessungen erfolgen, da gilt:

$$\frac{m_{AB}}{m_{AA}} = \frac{\phi_{AB}}{\phi_{AA}} = \phi_{AB} \leq 1, \qquad \frac{m_{BB}}{m_{AB}} = \frac{\phi_{BB}}{\phi_{BA}} = \frac{1}{\phi_{BA}} \geq 1. \tag{9}$$

Wären mehr als zwei Referenzmaterialien involviert, müssten zur Auswertung die Absolutwerte z.B. m_{uA}, m_{uB}, m_{uC} auf ein Maximum hin untersucht werden.

Die eigentliche Messreihe besteht aus jeweils 100 Messungen $\frac{m_{AB}}{m_{AA}}$ und $\frac{m_{BB}}{m_{BA}}$ bei unterschiedlichen Belichtungszeiten. Über die Belichtungszeit lässt sich sehr komfortabel die Anzahl Photonen beeinflussen und somit das SNR. Die Messwerte sind als Whisker-Plot in Abbildung 4.2 dargestellt. Die zwei Niveaus, in denen die Messwerte liegen, identifizieren jeweils eine unbekann-

[8] Minolta Spektrophotometer CM-3610d

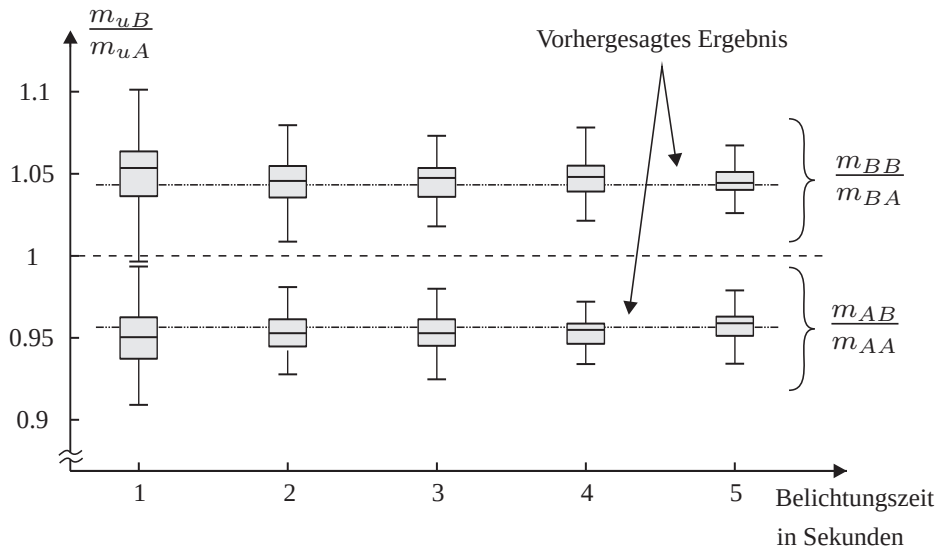


Bild 6: Belichtungsreihe für Probe A und B.

te Probe. Es ist zu sehen, dass für das beschriebene Experiment eine sichere Unterscheidung erst ab 2 Sekunden Belichtungszeit möglich ist. Qualitativ lässt sich sagen, dass wie erwartet das SNR mit steigender Belichtungszeit zunimmt. Durch eine hellere Beleuchtung kann die Aufnahmezeit verkürzt werden. Der Mittelwert über die Einzelmessungen liegt nahe am vorhergesagten Wert $\phi_{AB} = 0.96$, bzw. $\frac{1}{\phi_{BA}} = \frac{1}{\phi_{AB}} \approx 1.04$.

5 Zusammenfassung und Ausblick

Die bildgebende Spektroskopie macht optische Materialeigenschaften für die Bildverarbeitung zugänglich, die geeignet sind, Materialien zu identifizieren. Für inline-Industrieprozesse sind schnelle Sensoren notwendig, die sich leicht an eine Problemstellung adaptiert lassen. Das vorgestellte Messprinzip erfüllt prinzipiell die Forderung an hohe Geschwindigkeiten, zum anderen lässt es sich leicht an Aufgabenstellungen anpassen. Des Weiteren kann es als ein Ähnlichkeitsmaß zwischen Spektren interpretiert werden, d.h. in Form eines optisch realisierten normierten Kreuzkorrelationskoeffizienten. Besonders attraktiv sind Anwendungen im NIR, da hier zum einen die Spektren von der chemischen Struktur der Materialien abhängen (Schwingungsspektren), zum anderen geeignete Kamerasensoren verfügbar sind.

Literatur

[1] J. Beyerer, M. Sauerland, and R. Gruna, "Anordnung zur optischen Bestimmung einer Probe und entsprechendes Verfahren." Deutsches Patent- und Markenamt, May 2011. Fraunhofer-Gesellschaft.

[2] C. Liu, L. Sharan, E. H. Adelson, and R. Rosenholtz, "Exploring features in a bayesian framework for material recognition," in *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 239–246, IEEE Computer Society, (Los Alamitos, CA, USA), 2010.

[3] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **31**(11), pp. 2032–2047, 2009.

[4] O. Wang, P. Gunawardane, S. Scher, and J. Davis, "Material classification using BRDF slices," in *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, pp. 2805–2811, IEEE Computer Society, (Los Alamitos, CA, USA), 2009.

[5] D. L. Pavia, G. M. Lampman, G. S. Kriz, and J. A. Vyvyan, *Introduction to Spectroscopy*, Brooks Cole, 4 ed., Mar. 2008.

[6] M. Michelsburg, R. Gruna, and F. L. Puente, "Spektrale Bandselektion beim Entwurf automatischer Sortieranlagen," in *Forum Bildverarbeitung*, p. 426, KIT Scientific Publishing, (Regensburg), Dec. 2010.

[7] M. L. Myrick, O. Soyemi, J. Karunamuni, D. Eastwood, H. Li, L. Zhang, A. E. Greer, and P. Gemperline, "A single-element all-optical approach to chemometric prediction," *Vibrational Spectroscopy* **28**, pp. 73–81, Feb. 2002.

[8] B. Dai, A. Urbas, C. C. Douglas, and R. A. Lodder, "Molecular factor computing for predictive spectroscopy," *Pharmaceutical Research* **24**, pp. 1441–1449, Mar. 2007.

[9] S. E. Bialkowski, "Species discrimination and quantitative estimation using incoherent linear optical signal processing of emission signals," *Analytical Chemistry* **58**, pp. 2561–2563, Oct. 1986.

[10] M. P. Nelson, J. F. Aust, J. A. Dobrowolski, P. G. Verly, and M. L. Myrick, "Multivariate optical computation for predictive spectroscopy," *Analytical Chemistry* **70**, pp. 73–82, Jan. 1998.

[11] J. Medendorp and R. A. Lodder, "Applications of integrated sensing and processing in spectroscopic imaging and sensing," *Journal of Chemometrics* **19**, pp. 533–542, Oct. 2005.

[12] B. E. A. Saleh and M. C. Teich, *Grundlagen der Photonik*, Wiley-VCH, Mar. 2008.

Übertragungsverhalten von tiefenscannenden Weißlichtinterferometern

Weichang Xie, Jan Niehues, Peter Lehmann

Universität Kassel, FB Elektrotechnik/Informatik, Lehrstuhl für Messtechnik
Wilhelmshöher Allee 71, 34109 Kassel, E-Mail: p.lehmann@uni-kassel.de

Kurzfassung

Tiefenscannende Weißlichtinterferometer haben sich in der Mikro- und Nanotechnologie als Messgeräte zur 3D-Geometrieerfassung etabliert. Während eine axiale Auflösung im Sub-Nanometerbereich erreicht werden kann, ist die laterale Auflösung durch die Beugung auf die Größenordnung der Lichtwellenlänge begrenzt. Zudem treten bei der Messung von Kantenstrukturen Abweichungen von einer korrekten Übertragung der zu messenden Topographie auf. Die Beugung führt hier zu dem Batwing-Effekt. Ein Ansatz, das Übertragungsverhalten eines optischen Profilometers zu charakterisieren, basiert auf der ITF (Instrument Transfer Function), die sich von der für Abbildungssysteme gebräuchlichen MTF (Modulation Transfer Function) ableitet. Aufgrund von Diskrepanzen zwischen dem theoretischen Verlauf der ITF und realen Messergebnissen wird eine numerische Simulation zur Untersuchung der ITF in Abhängigkeit der Periode und der Amplitude eines Rechteckgitters durchgeführt. Die Simulationsergebnisse zeigen ein nichtlineares Übertragungsverhalten und eine gute Übereinstimmung mit realen Messergebnissen.

1 Einleitung

Besonders in den Mikro- und Nanotechnologien ist die Weiterentwicklung von Produktionsprozessen eng verknüpft mit der Verfügbarkeit von geeigneten Messinstrumenten zur Erfassung dreidimensionaler Strukturen. Dabei bekommt neben der häufig zufriedenstellenden Höhenauflösung die laterale Auflösung

einen immer höheren Stellenwert.
Dieser Beitrag beschäftigt sich mit der Charakterisierung des Übertragungsverhaltens von Weißlichtinterferometern. Die Weißlichtinterferometrie (WLI) ist eine schnelle und präzise Messmethode zur Erfassung von 3D-Topographien [1]. Sie bietet den Vorteil, dass die Höhenauflösung unabhängig davon ist, ob es sich um ein mikroskopisches oder makroskopisches Messfeld handelt. Bild 1 zeigt den schematischen Aufbau eines Weißlicht-Interferenzmikroskops.

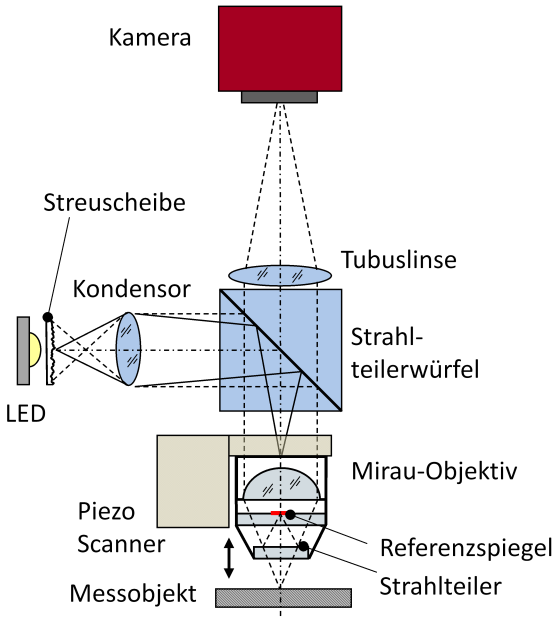


Bild 1: Aufbau eines Weißlichtinterferometers mit Mirau-Objektiv.

Das Messprinzip basiert auf einem sogenannten Tiefenscan, bei dem sich die optische Weglänge im Messarm des Interferometers ändert, während sie im Referenzarm konstant bleibt. Dabei nimmt eine Kamera in äquidistanten Schritten Bilder auf [2]. Der Tiefenscan liefert somit für jedes Kamerapixel ein separates Interferenzsignal, das als Korrelogramm bezeichnet wird. Bild 2 zeigt den typischen Verlauf eines solchen Korrelogramms. Die Breite der Einhüllenden verhält sich proportional zur Kohärenzlänge des verwendeten Lichtes.

Abhängig von der gewünschten Vergrößerung und Auflösung kommen in der WLI verschiedene Interferenzobjektive zum Einsatz. Wegen ihrer kompakten und robusten Bauweise werden in Industrieanwendungen vorzugsweise Mirau-Objektive mit integriertem Strahlteiler und Referenzspiegel verwendet. Ein Nachteil dieser Objektive ist, dass die numerische Apertur (*NA*) und der Ar-

beitsabstand durch den Strahlteiler vor dem Objektiv eingeschränkt sind (siehe Bild 1).

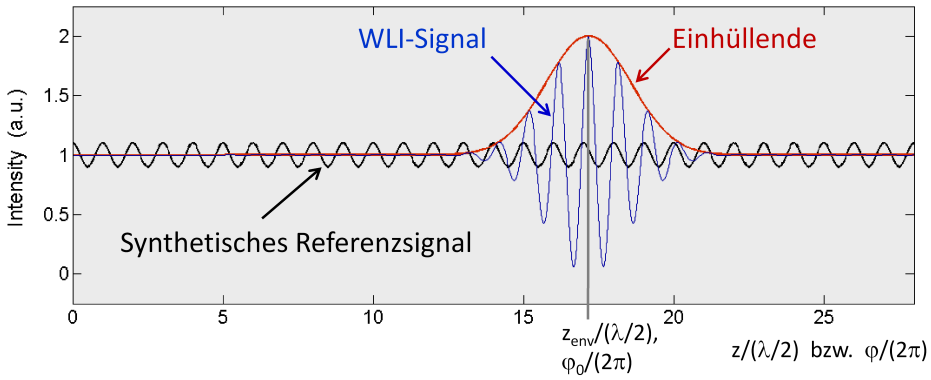


Bild 2: Bei einer WLI-Messung resultierendes Korrelogramm mit Hüllkurve und synthetischem Referenzsignal zur Phasenauswertung.

Ziel der WLI-Signalverarbeitung ist es, aus den gemessenen Korrelogrammen die 3D-Topographie zu berechnen. Üblicherweise wird in einem ersten Schritt die Position der in Bild 2 dargestellten Einhüllenden bestimmt. Dies kann über den Interferenzkontrast oder im Ortsfrequenzbereich über den Anstieg der Phase erfolgen [1]. Um Höhenauflösungen im Sub-Nanometerbereich zu erreichen, muss die Phasenlage des Interferenzsignals in die Auswertung einbezogen werden [1, 3, 4]. In Bild 2 wird die Phase des Korrelogramms auf ein synthetisch generiertes Referenzsignal bezogen. Die aus der Position der Einhüllenden gewonnene Lage des Interferenzstreifens 0-ter Ordnung löst das Problem der 2π Mehrdeutigkeit der Ergebnisse der Phasenauswertung. Damit sind die in der WLI erreichbaren Höhenauflösungen denen der phasenschiebenden Interferometrie (PSI) vergleichbar. Zusätzlich können jedoch Stufenhöhen größer der halben Lichtwellenlänge korrekt gemessen werden [1, 4, 5].

An Stufen mit Höhen, die maximal der Kohärenzlänge des Lichtes entsprechen, kommt es zu dem bekannten Batwing-Effekt [5]. Dieser Effekt wirkt sich u. U. auch auf das gemessene Übertragungsverhalten aus. Die Einhüllende kann sich so stark verschieben, dass eine fehlerfreie Bestimmung der Interferenzstreifenordnung nicht mehr möglich ist, was zu Phasensprüngen im Ergebnis führt [5–7].

Die laterale Auflösung in der 3D-Mikroskopie ist wie bei jedem optischen Ab-

bildungssystem beugungsbegrenzt. In einer Simulationsuntersuchung wurden die Übertragungseigenschaften der WLI analysiert. Die Ergebnisse zeigen, dass auf Grundlage der Phaseninformationen das Profil einer rechteckigen Gitterstruktur bis zur Auflösungsgrenze korrekt gemessen werden kann.
Um die laterale Auflösung zu verbessern, wurde ein Linnik-Interferometer entwickelt, das im nahen UV-Bereich arbeitet. Es stellt einen Kompromiss zwischen hoher *NA* und großem Arbeitsabstand dar.

2 Laterale Auflösung und Übertragungsverhalten

Für ein auf mikroskopischer Abbildung basierendes optisches Profilometer geht man in der Regel davon aus, dass die laterale Auflösung der optischen Auflösung entspricht, die durch das Rayleigh-Kriterium: $\delta = 0,61\lambda/NA$, oder das Sparrow-Kriterium: $\delta = 0,47\lambda/NA$ definiert ist, wobei δ den kleinsten Abstand zweier noch auflösbarer Punkte auf dem Messobjekt darstellt [8].
Während bei einer 2D-Abbildung nur lateral versetzte Punkte betrachtet werden, sollen in der 3D-Profilometrie jedoch Punkte unterschiedlicher Höhe aufgelöst werden. Die laterale Auflösung muss sich also auf einen Punkt auf dem oberen Niveau und einen auf dem unteren Niveau beziehen, wie durch Bild 3 veranschaulicht wird. Bei der dargestellten Rechteckstruktur entspricht die auflösbare Periode Λ deshalb dem Doppelten der optischen Auflösung δ eines reinen Abbildungssystems.

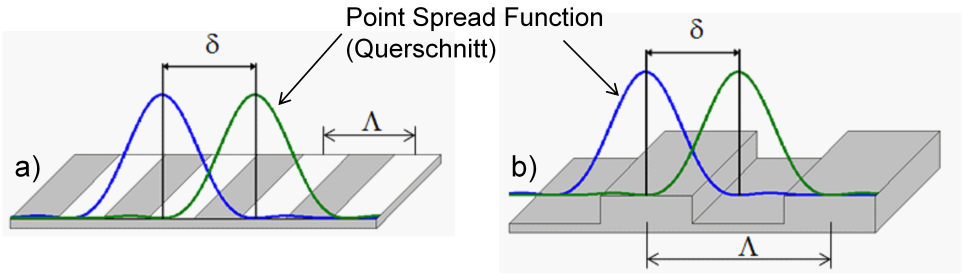


Bild 3: a) Gemäß Rayleigh-Kriterium lateral aufgelöstes 2D-Amplitudengitter der Periode Λ bei mikroskopischer Abbildung, b) aufgelöste Höhenstruktur eines Rechteckgitters der Periode Λ im Fall der 3D-Profilometrie.

Nach der VDI/VDE Richtlinie 2655 [9] liegt die kurze Grenzwellenlänge eines zu messenden Profils, bei der die mittels WLI gemessene Modulationstiefe

noch 50% der realen Modulationstiefe beträgt, bei dem 2- bis 3-fachen Wert der lateralen Auflösung nach dem Rayleigh-Kriterium. Diese Überlegung geht von einem näherungsweise linearen Übertragungsverhalten des Profilometers aus. Um das Übertragungsverhalten theoretisch zu beschreiben, definieren de Groot et al. [10] die Instrument Transfer Function (ITF) auf der Grundlage der aus der Abbildungstheoerie bekannten Modulation Transfer Function (MTF).
Die ITF stellt den Quotienten der gemessenen Amplitude eines Sinusnormals zur realen Amplitude in Abhängigkeit der Ortsfrequenz dar. Die ITF hat den Wert Eins, wenn bei einer gegebenen Ortsfrequenz die volle Modulationstiefe erfasst wird. Bei steigender Ortsfrequenz des Messobjektes nimmt die ITF ab. Sie nähert sich dem Wert Null, wenn die Ortsfrequenz bis hin zur Auflösungsgrenze ansteigt. Die ITF stimmt mit der MTF des entsprechenden mikroskopischen Abbildungssystems überein, wenn die Oberflächeamplitude sehr viel kleiner ist als ein Viertel der mittleren Wellenlänge des verwendeten inkohärenten Lichts [10]. Bild 4 zeigt ein Beispiel der MTF für ein Objektiv mit einer *NA* von 0,55. Obwohl die Auflösungsgrenze nach dem Rayleigh-Kriterium ungefähr 1500 Zyklen pro mm entspricht, kann die volle Modulationstiefe nur bei Ortsfrequenzen unter 10 Zyklen pro mm erreicht werden. Für die Praxis ist diese Definition der ITF nur von eingeschränkter Bedeutung, da die Voraussetzung bezüglich der Oberflächenamplituden häufig nicht erfüllt ist.

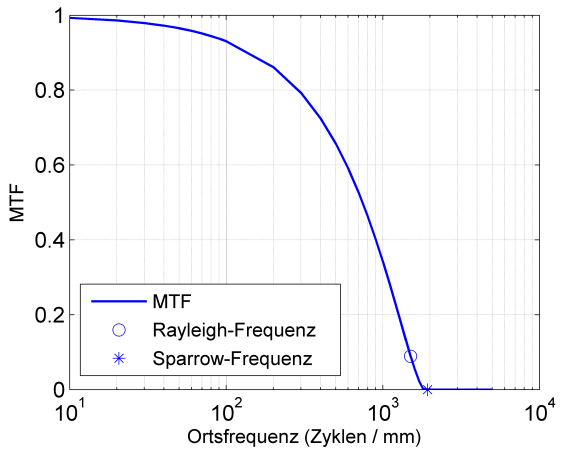


Bild 4: Verlauf der MTF für *NA* = 0,55 und $\overline{\lambda}$ = 600 nm; die Rayleigh und Sparrow-Frequenz entsprechen jeweils dem Kehrwert der Rayleigh und Sparrow-Auflösung; die MTF entspricht der ITF eines Profilometers unter der Voraussetzung sehr geringer Oberflächenhöhen.

3 Simulationsuntersuchung

Anders als bei der analytischen Ermittlung der ITF gemäß Bild 4, bei der die ITF aus der normierten Autokorrelationsfunktion der Amplitude Transfer Function resultiert [11], können durch eine numerische Simulation sowohl der tiefenscannende Messprozess als auch die Phasenauswertung in der WLI berücksichtigt werden.
Das Interferenzsignal, das während des Tiefenscans mit einer konstanten Schrittweite Δz für das Höhenprofil $h(x, y)$ aufgenommen wird, lässt sich wie folgt beschreiben:

$$I_n(x,y) = I_0 \left(1 + \exp\left\{-4\left[\frac{h(x,y)+n\,\Delta z}{l_c}\right]^2\right\} \cos\left\{\frac{4\pi}{\overline{\lambda}}\left[h(x,y)+n\,\Delta z\right]\right\}\right) \quad (1)$$

Hierbei ist $I_0 = I_{\text{obj}} + I_{\text{ref}}$ die Summe der Intensität I_{obj} im Objekt- und I_{ref} im Referenzarm. l_c ist die Kohärenzlänge des Lichtes der mittleren Wellenlänge $\overline{\lambda}$ und $n = \cdots -2, -1, 0, 1, 2 \cdots$ stellt die einzelnen Schritte des Tiefenscans dar.
Um die durch das optische System bedingte Beugung zu berücksichtigen, muss $I_n(x, y)$ mit der Punktverwaschungsfunktion (Point Spread Function) gefaltet werden:

$$I_{n,\text{image}}(x,y) = I_n(x,y) * \text{psf}(x,y) \quad (2)$$

Mit Hilfe des Faltungstheorems kann Gl. (2) umgeschrieben werden:

$$I_{n,\text{image}}(x,y) = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{I_n(x,y)\right\} \cdot \mathcal{F}\left\{\text{psf}(x,y)\right\}\right\}, \quad (3)$$

wobei $\mathcal{F}\{\ldots\}$ und $\mathcal{F}^{-1}\{\ldots\}$ die Fouriertransformation bzw. die Fourier-Rücktransformation bezeichnen.
Als zu untersuchendes Messobjekt wird ein Rechteckgitter der Höhe h_0 und der Periode T angenommen:

$$h(x,y) = h(x,y=0) = \begin{cases} h_0 & mT \le x < mT + \frac{T}{2} \\ 0 & mT + \frac{T}{2} \le x < mT + T \end{cases} \quad (4)$$

mit $m = 0, 1, 2....$
Für jede Position n des Tiefenscans werden die ungefilterte Intensität $I_n(x, y)$ und die Bildintensität $I_{n,\text{image}}(x, y)$ berechnet. Da das Höhenprofil des Messobjektes ausschließlich in x Richtung verläuft, sind $I_n(x, y)$ und $I_{n,\text{image}}(x, y)$

zweidimensionale Matrizen, die von x und der Scanposition z abhängen. Bild 5 zeigt das Simulationsergebnis der ungefilterten Intensität und der Bildintensität für das in Gl. (4) definierte Messobjekt mit $h_0 = \overline{\lambda}/4$ und $T = 2\overline{\lambda}$.

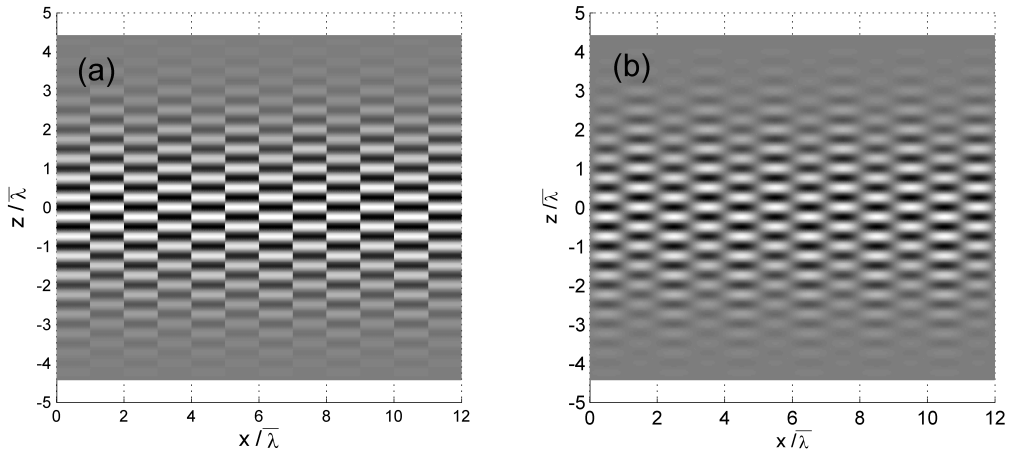


Bild 5: Simulierte Intensitätsverläufe für ein Rechteckgitter mit $h_0 = \overline{\lambda}/4$ und $T = 2\overline{\lambda}$ in Graustufendarstellung (dunkle Stellen entsprechen geringer Intensität, helle Stellen hoher Intensität); a) Intensität $I_n(x)$ ohne Berücksichtigung der Beugungsbegrenzung, b) Bildintensität $I_{n,\mathrm{image}}(x)$ unter Berücksichtigung der Beugung.

Während die Grenzen zwischen dem oberen und dem unteren Niveau in Bild 5 a) deutlich zu erkennen sind, sind sie im Bild 5 b) unscharf. Da die Faltung mit der Punktverwaschungsfunktion einer Tiefpass-Filterung entspricht, gehen die Anteile der hohen Ortsfrequenzen bei der Abbildung verloren.
Die simulierten Bildintensitäten $I_{n,\mathrm{image}}(x, y)$ lassen sich nun mit denselben Signalverarbeitungsalgorithmen auswerten wie experimentell ermittelte Korrelogramme und das Höhenprofil des Messobjektes wird auf der Grundlage der Simulationsergebnisse bestimmt. Durch die Wiederholung solcher Simulationsberechnungen für unterschiedliche Periodenlängen T ergeben sich simulierte ITF-Verläufe, indem die aus den simulierten Oberflächenprofilen bestimmte Profilhöhe zur vorgegebenen Profilhöhe h_0 ins Verhältnis gesetzt wird. Der Einfluss der vorgegebenen Profilhöhe h_0 lässt sich auf diese Weise ebenfalls untersuchen.

Bild 6 zeigt simulierte ITF-Verläufe im Vergleich zur analytischen ITF gemäß Bild 4 bei unterschiedlichen Amplituden h_0 [12]. Die Ermittlung der ITF im Bild 6 erfolgte mittels der Phasenauswertung. Die Ergebnisse zeigen, dass die

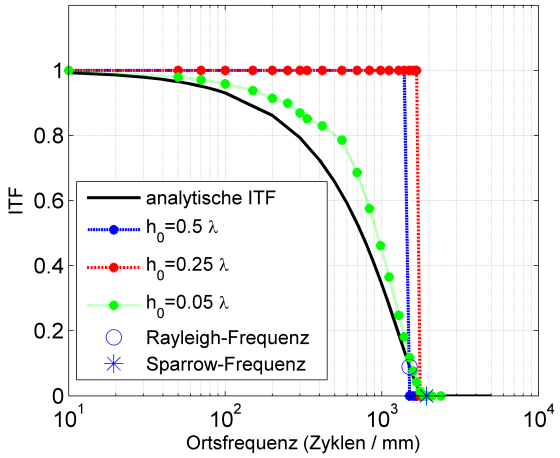


Bild 6: Vergleich der simulierten ITF mit der analytischen ITF gemäß Bild 4 für unterschiedliche Amplituden h_0 des Rechteckgitters.

ITF nicht nur von der Ortsfrequenz abhängig ist, sondern dass auch die Amplitude des Rechteckgitters einen erheblichen Einfluss hat. Die laterale Auflösung ist optimal, wenn die Amplitude des Rechteckgitters exakt ein Viertel der mittleren Wellenlänge des verwendeten Lichtes beträgt. Bis die Rayleigh-Frequenz, die sich aus dem Rayleigh-Kriterium für die mikroskopische 2D-Abbildung ergibt, erreicht ist, wird die volle Modulationstiefe gemessen. Dieses Ergebnis steht weder in Übereinstimmung zu der Überlegung nach Bild 3 noch mit der VDI/VDE-Richtlinie [9], nach der die Periode eines Rechteckgitters mindestens dem 4-fachen der Rayleigh-Auflösung entsprechen muss, damit die volle Modulationstiefe des Profils erfasst wird. Dieser Unterschied lässt dadurch erklären, dass bei der Simulation sowohl der für die WLI typische Tiefenscan als auch die Phasenlage der Korrelogramme berücksichtigt werden.
Die Simulationsergebnisse zeigen, dass die WLI sich stark nichtlinear verhält. Bei bestimmten Verhältnissen aus der Amplitude des Rechteckgitters zur mittleren Lichtwellenlänge ist das Übertragungsverhalten bei hohen Ortsfrequenzen wesentlich besser als es nach der analytischen ITF zu erwarten wäre.

4 Vergleich von Simulations- und Messergebnissen

Im Rahmen des EU-Projektes „NanoCMM“ wurde ein Weißlichtinterferometer mit hoher lateraler Auflösung und großem Arbeitsabstand entwickelt. Das

Interferometer wurde als kompakter Sensor ausgelegt, um es in kommerzielle Koordinatenmessmaschinen für Mikroobjekte integrieren zu können. Für einen sicheren Einsatz innerhalb der Maschine muss der Arbeitsabstand mindestens 5 mm betragen. Der Tiefenscan erfolgt in diesem Fall durch Bewegen des Messobjektes z.B. mittels einer Achse des Koordinatenmessgerätes.
Um unter diesen Voraussetzungen eine laterale Auflösung unter 0,5 μm zu realisieren, ist das Interferometer auch für den blauen und nahen UV-Spektralbereich ausgelegt. Die höchstmögliche *NA*, die unter diesen Bedingungen zu erreichen ist, bietet ein Linnik-Aufbau, in dem im Mess- und im Referenzarm des Interferometers jeweils separate Mikroskopobjektive eingesetzt werden. Im konkreten Fall wurden Objektive mit einer *NA* von 0,55, einer 50x Vergrößerung und einem Arbeitsabstand von 7 mm verwendet.

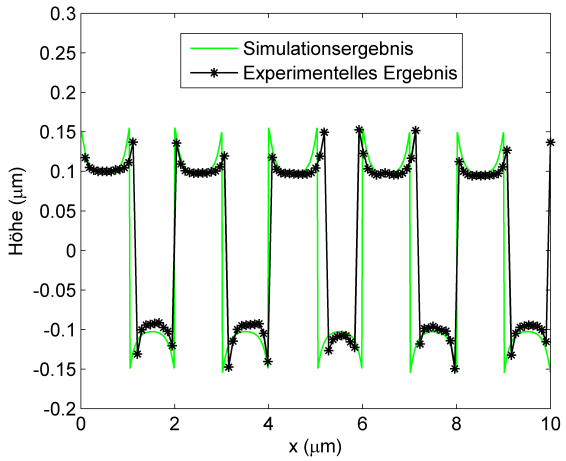


Bild 7: Vergleich von Messergebnissen des Linnik-Interferometers bei einer mittleren Wellenlänge von 620 nm an einem Rechtecknormal mit einer Höhe von 192 nm und einer Periodenlänge von 2 μm mit Profilen, die anhand simulierter Korrelogramme bestimmt wurden.

Die nachfolgend dargestellten Resultate wurden erneut mit Hilfe der Phasenauswertung der Korrelogramme bestimmt. Bild 7 zeigt Messergebnisse, die mit dem beschriebenen Linnik-Interferometer an in Silizium geätzten Rechteckstrukturen (RS-N Normal der Fa. Simetrics) einer Höhe von 192 nm und einer Periodenlänge von 2 μm aufgenommen wurden. Das bei einer mittleren Wellenlänge von 620 nm betriebene Interferometer löst die Rechteckstruktur auf. Die zusätzlich in dem Diagramm dargestellten Simulationsergebnisse zeigen eine sehr gute Übereinstimmung mit den experimentellen Resultaten. Dies

gilt insbesondere auch für den Batwingeffekt, der an den Kanten sichtbar ist.
Für Bild 8 wurde eine UV-LED zur Beleuchtung verwendet und die mittlere Wellenlänge damit auf 395 nm reduziert. Außerdem wurde ein Rechteck-Normal mit einer Periodenlänge von 0,6 μm verwendet, das eine Profilhöhe von ca. 160 nm aufweist. Die Rechteckstruktur kann eindeutig aufgelöst werden. Auch hier stimmt das gemessene Profil sowohl in der Höhe als auch in der Profilform gut mit dem Simulationsergebnis überein. Weitere experimentelle Ergebnisse sind in [13] veröffentlicht.

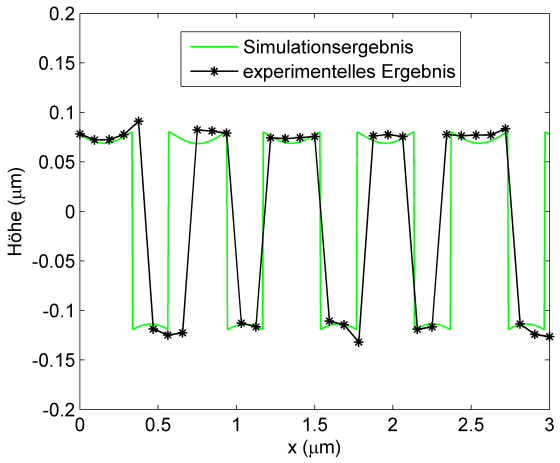


Bild 8: Vergleich von Messergebnissen des Linnik-Interferometers bei einer mittleren Wellenlänge von 395 nm an einem Rechtecknormal mit einer Höhe von 160 nm und einer Periodenlänge von 0,6 μm mit Profilen, die anhand simulierter Korrelogramme bestimmt wurden.

In diesem Zusammenhang ist zu erwähnen, dass die Strukturbreite im letzten Fall nur 300 nm beträgt, was nach dem Rayleigh-Kriterium unterhalb der Auflösungsgrenze von 440 nm liegt. Dies steht im Gegensatz zu der entsprechenden VDI-Richtlinie [9] und den Betrachtungen gemäß Bild 3.
Die theoretische ITF aus Bild 4 lässt sogar erwarten, dass die volle Modulationstiefe eines periodischen Profils erst dann resultiert, wenn die optische Auflösung mindestens zwei Größenordnungen geringer ist als die Strukturbreite des Objektes. Diese Voraussetzung steht ebenfalls im Widerspruch zu den Messergebnissen aus Bild 8. Jedoch wird bei der ITF gemäß Bild 4 vorausgesetzt, dass die Höhendifferenz des Rechteckgitters im Vergleich zur mittleren Wellenlänge des Lichtes gering ist, was in Bild 8 nicht erfüllt ist.

5 Zusammenfassung und Ausblick

Tiefenscannende Weißlichtinterferometer sind weit verbreitete Messgeräte zur Erfassung von 3D-Mikrotopographien. Bisherige Ansätze zur Beschreibung ihres Übertragungsverhaltens gehen von der linearen optischen Abbildungstheorie aus. Dabei wird vorausgesetzt, dass die Höhendifferenz der zu messenden Strukturen sehr klein sind, so dass die Übertragungsfunktion, die ITF, mit der MTF eines inkohärenten mikroskopischen Abbildungssystems übereinstimmt.
Bereits bei Strukturhöhen von 100 nm sind die Voraussetzungen der linearen Beschreibungsweise nicht mehr erfüllt. Zudem bleibt der auf einem Tiefenscan basierende Messprozess, bei dem die Phasenlage einzelner Interferenzsignale ausgewertet wird, unberücksichtigt. Folglich zeigen reale Messergebnisse deutliche Abweichungen gegenüber den theoretischen Voraussagen. Dies gilt insbesondere für Strukturen nahe der lateralen Auflösungsgrenze.
Unser Simulationsmodell hebt die Beschränkung auf geringe Oberflächenhöhen auf und berücksichtigt zudem den tiefenscannenden Messprozess. Die auf diese Weise simulierten Signale lassen sich mit denselben Algorithmen auswerten wie reale Messsignale. Die Simulationsergebnisse zeigen eine gute Übereinstimmung mit experimentell gewonnenen Daten. Sie offenbaren, dass sich tiefenscannende Weißlichtinterferometer nichtlinear verhalten. Die Übertragungsfunktion hängt nicht nur von der Periodenlänge eines gemessenen Profils, sondern auch von dessen Amplitude und Form ab. Entgegen den bisherigen Annahmen konnte bei Rechteckprofilen geeigneter Amplitude der Profilverlauf bis zu einer Periodenlänge, die der optischen Auflösung gemäß dem Rayleigh-Kriterium entspricht, korrekt erfasst werden.
Die Simulationsergebnisse spiegeln den Batwing-Effekt ebenfalls realistisch wider. Damit eröffnet das Simulationsmodell die Möglichkeit, gerätetechnische Varianten simulativ zu untersuchen und Optimierungen vorzunehmen.

Danksagung

Ein Teil der hier vorgestellten Arbeiten wurde im Rahmen des Projektes „KalWint“ (Nr. 01FS10045) mit Mitteln des BMWi gefördert.

Literatur

[1] L. Deck und P. de Groot, "High-speed noncontact profiler based on scanning white-light interferometry", *Appl. Opt.* **33**, S. 7334–7338, 1994.

[2] B. S. Lee und T. C. Strand, "Profilometry with a coherence scanning microscope", *Appl. Opt.* **29**, S. 3784–3788, 1990.

[3] P. de Groot und L. Deck, "Surface profiling by analysis of white-light interferograms in the spatial frequency domain", *J. of Mod. Opt.* **42**, S. 389–401, 1995.

[4] M. Fleischer, R. Windecker und H. J. Tiziani, "Fast algorithms for data reduction in modern optical three-dimensional profile measurement systems with MMX technology", *Appl. Opt.* **39**, S. 1290–1297, 2000.

[5] A. Harasaki und J. C. Wyant, "Fringe modulation skewing effect in white light vertical scanning interferometry", *Appl. Opt.* **39**, S. 1290–1297, 2000.

[6] P. de Groot und X. C. de Lega, "Determination of fringe order in white-light interference microscopy", *Appl. Opt.* **41**, S. 4571–4578, 2002.

[7] P. Lehmann, "Systematic effects in coherence peak and phase evaluation of signals obtained with a vertical scanning white-light mirau interferometer", *SPIE Proceedings* **6188**, S. 618811–1–618811–11, 2006.

[8] R. Leach, *Optical Measurement of Surface Topography*, Springer, 2011.

[9] "Optische Messtechnik an Mikrotopografien: Kalibrieren von Interferenzmikroskopen und Tiefeneinstellnormalen für die Rauheitsmessung", *VDI/VDE Richtlinien 2655, Blatt 1.1* , 2008.

[10] P. de Groot und X. C. de Lega, "Interpreting interferometric height measurements using the instrument transfer function", *Proceedings Fringe 2005, Spinger* , S. 30–37, 2005.

[11] J. W. Goodman, *Introduction to Fourier Optics*, Roberts & Company Publishers, third ed., 2005.

[12] W. Xie, P. Lehmann und J. Niehues, "Lateral resolution and transfer characteristics of vertical scanning white-light interferometers", *Appl. Opt.* , submitted.

[13] J. Niehues und P. Lehmann, "Improvement of lateral resolution and reduction of batwings in vertical scanning white-light interferometry", *SPIE Proceedings* **8082**(80820W), 2011.

Phasenrauschen in optischen Messsystemen mit strukturierter Beleuchtung

Marc Fischer, Marcus Petz, Rainer Tutsch

Institut für Produktionsmesstechnik, Technische Universität Braunschweig
Schleinitzstraße 20, 38106 Braunschweig
Tel.: 0531-391-7023, e-mail: marc.fischer@tu-braunschweig.de

Kurzfassung

Der Schwerpunkt wissenschaftlicher Arbeiten zum Thema Phasenrauschen lag bisher auf dem Gebiet der Interferometrie. Die Untersuchungen basierten häufig auf der Annahme von normalverteiltem signalunabhängigem Intensitätsrauschen und resultierten in der Vorhersage einer umgekehrt proportionalen Abhängigkeit des Phasenrauschen vom relativen Streifenkontrast (Sichtbarkeit) und der mittleren Intensität. Da das Rauschverhalten moderner Kameras bei sinnvoller Aussteuerung im Allgemeinen von signalabhängigem Poissonverteiltem Photonenrauschen dominiert wird, sind in der Praxis deutliche Abweichungen von den Vorhersagen zu beobachten. Im Rahmen dieser Arbeit wird ein Phasenrauschmodell auf der Grundlage des EMVA 1288 Kameramodells vorgestellt, welches deutlich besser in der Lage ist, das Phasenrauschen für optische Messsysteme mit strukturierter Beleuchtung vorherzusagen.

1 Einleitung

Die Berechnung eines unbekannten Phasenwinkels aus einer Serie von phasenverschobenen Sinusmustern ist die Grundlage vieler optischer Messverfahren, sowohl mit kohärentem als auch mit inkohärentem Licht. Bei Messverfahren auf der Grundlage von strukturierter Beleuchtung wird diese Phaseninformation genutzt, um über einen Triangulationsprozess Positionswerte zu bestimmen. Dafür wird ein Streifenmuster auf das Objekt projiziert (Streifenprojektion) oder auf einer Referenzfläche angezeigt (Deflektometrie). Zum Einsatz kommen dabei häufig digitale Projektoren (DLP) bzw. Computermonitore (LCD). Im Gegensatz zur Interferometrie haben Vibrationen und Phasenschiebefehler nur gerin-

gen Einfluss auf die Genauigkeit der Phasenmessung, da diese Systeme mit inkohärentem Licht arbeiten und die Genauigkeit der Phasensprünge durch die Fertigungstoleranzen des Pixelrasters festgelegt ist. Außerdem können Aliasing-Effekte, welche die Hauptquelle für systematische Fehler darstellen, durch Defokussierung erheblich reduziert werden. Neben diesen systematischen Einflüssen hängt die Genauigkeit der Phasenberechnung daher wesentlich von den Rauscheigenschaften der eingesetzten Kamera ab.

Bisherige wissenschaftliche Untersuchungen zum Thema Phasenrauschen basierten häufig auf der Annahme einer normalverteilten signalunabhängigen Rauschquelle [1][2]. Die Bilder moderner Kameras werden jedoch im Allgemeinen von Poisson-verteiltem signalabhängigem Photonenrauschen dominiert. Für die allgemeine Charakterisierung von Kameras fand diese Erkenntnis Eingang in die Norm "EMVA Standard 1288 – Standard for Characterization of Image Sensors and Cameras" (EMVA 1288) [3] der European Machine Vision Association in Form eines erweiterten linearen Kameramodells. Der aktuelle Normentwurf wurde im November 2010 veröffentlicht und ermöglicht eine objektive und vergleichbare Charakterisierung von Bildsensoren und Kameras. Darüber hinaus kann das erweiterte Kameramodell der EMVA 1288 genutzt werden, um die Leistungsfähigkeit und Genauigkeit von optischen Messsystemen zu analysieren [4][5].

Im Rahmen dieser Arbeit erfolgt eine systematische Analyse des Phasenrauschens bei Messsystemen mit strukturierter Beleuchtung. Für den symmetrischen M-Schritt Algorithmus wird ein mathematischer Ausdruck hergeleitet, mit dem das Phasenrauschen σ_φ aus den Parametern eines EMVA 1288 konformen Datenblatts und zwei normierten signalbeschreibenden Größen vorhergesagt werden kann. Die Gültigkeit wird mittels Experimenten belegt.

2 Grundlagen

2.1 Optische Messsysteme mit strukturierter Beleuchtung

Der Begriff „strukturierte Beleuchtung" bezeichnet eine Klasse von optischen Messverfahren bei welchen das Messobjekt oder eine Referenzstruktur durch Intensitätsmuster örtlich kodiert und dieses Muster anschließend mit einer Kamera aufgezeichnet wird. Ein häufig verfolgter Ansatz basiert dabei auf dem

Einsatz von mehreren phasenverschobenen sinusförmigen Streifenmustern. Die Intensität I_i in einem Pixel der Kamera bei Beobachtung des Musters mit der Nummer i kann dabei als Summe einer konstanten mittleren Intensität I' und einer modulierten Intensität I'' beschrieben werden:

$$I_i = I' + I'' cos(\varphi + \psi_i); \quad i = 1,...,M. \tag{1}$$

Hierbei bezeichnet φ den zu berechnenden Phasenwinkel, ψ_i die Phasenverschiebung des entsprechenden Musters und M die Gesamtanzahl der verwendeten Muster. Für eine allgemeingültige Diskussion ist es sinnvoll, den absoluten Bezug der Intensitätsgrößen I' und I'' zu eliminieren. Durch die Definition von

$$I' = \beta I_{sat} \tag{2}$$

und

$$I'' = \gamma I' \tag{3}$$

kann (1) durch folgenden Zusammenhang ausgedrückt werden:

$$I_i = I_{sat} \beta [1 + \gamma cos(\varphi + \psi_i)]. \tag{4}$$

Hierin bezeichnet I_{sat} die Sättigungskapazität des Pixels, β die relative mittlere Intensität (Belichtung) und γ den relativen Streifenkontrast (Sichtbarkeit).

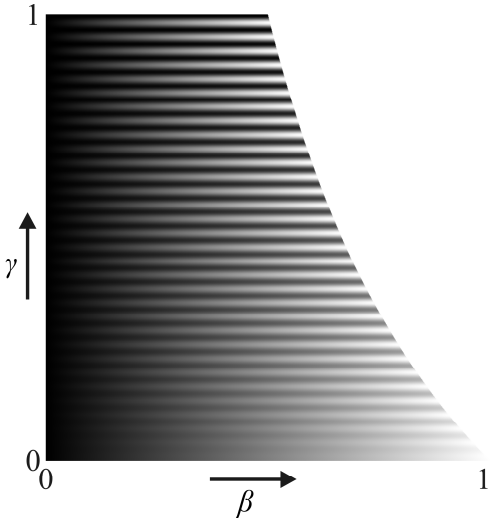


Bild 1: Streifenmuster in Abhängigkeit von der Sichtbarkeit (γ) und der Belichtung (β). Auf der rechten Seite des Diagramms sind Punkte welche die Bedingung (5) für die maximale Intensität verletzen weiß gezeichnet.

Der Vorteil dieser Darstellung ist die Trennung von skalierenden Effekten, welche die Belichtung β beeinflussen, und Kontrast-Effekten, welche die Sichtbarkeit γ verändern. Darüber hinaus kann (4) durch Einsetzen des entsprechenden Wertes für I_{sat} sehr einfach an unterschiedliche Intensitätsgrößen, wie zum Beispiel Anzahl der Photonen oder digitale Grauwerte, angepasst werden. Die fundamentalen Kenngrößen β und γ werden dadurch nicht beeinflusst. Der Wert für die Belichtung β wird im Wesentlichen durch das Zusammenspiel von maximaler Helligkeit des Projektors bzw. Displays und der Belichtungszeit und Blendeneinstellung der Kamera sowie von winkelabhängigen Effekten wie z.B. der Streucharakteristik eines Objekts oder der Helligkeitskennlinie eines Displays beeinflusst. Die Sichtbarkeit γ hingegen hängt vorrangig vom maximalen Kontrast des Projektors bzw. Displays, den Kontrastübertragungsfunktionen der eingesetzten Objektive sowie dem Anteil an Umgebungslicht ab. Beide Parameter besitzen Werte im Bereich [0…1] mit einer zusätzlichen Beschränkung der maximalen Belichtung β bei gegebener Sichtbarkeit γ, da die tatsächliche Intensität des Pixels I_i immer kleiner sein muss als die Sättigungskapazität I_{sat}:

$$\beta \leq \frac{1}{1+\gamma} \tag{5}$$

In Bild 1 ist der Einfluss der beiden Parameter auf die Ausprägung des Streifenmusters dargestellt. Der optimale Streifenkontrast wird mit $\beta = 0.5$ und $\gamma = 1$ erreicht.

2.2 Phasenschiebe-Algorithmen

Aus den aufgezeichneten Intensitäten eines Pixels I_i gemäß (4) kann der in diesem Bildpunkt beobachtete Phasenwert φ anschließend mit Hilfe von Phasenschiebe-Algorithmen berechnet werden. Eine systematische Analyse der zufälligen Abweichungen bei dieser Berechnung wurde von Rathjen [1] auf der Grundlage von bivariaten Verteilungsfunktionen durchgeführt. Ausgehend von einer allgemeinen Beschreibung eines Phasenschiebe-Algorithmus in Form eines von den Intensitäten I_i abhängigen Bruches mit einem Zähler Z und einem Nenner N:

$$\tan\varphi = \frac{N(I_1,...,I_M)}{Z(I_1,...,I_M)} \tag{6}$$

konnte gezeigt werden, dass die Form der resultierenden Verteilungsfunktion für den Phasenwinkel φ neben dem Signal-Rausch-Verhältnis (*SNR*) im Wesentlichen vom Korrelationskoeffizient ρ und dem Verhältnis α der Standardabweichungen von Zähler σ_Z und Nenner σ_N abhängt. Unter Verwendung der Kovarianz von Zähler und Nenner σ_{ZN} gilt dabei:

$$\rho = \frac{\sigma_{ZN}}{\sigma_Z \sigma_N} \tag{7}$$

und

$$\alpha = \frac{\sigma_N}{\sigma_Z}. \tag{8}$$

Nur für $\rho = 0$ und $\alpha = 1$ ist die Verteilungsfunktion phasenunabhängig und der Phasenschiebe-Algorithmus ein erwartungstreuer Schätzer für den unbekannten Phasenwinkel. Daher sind viele der in der Interferometrie eingesetzten Algorithmen für optische Messverfahren mit strukturierter Beleuchtung nicht geeignet. Diese Schlussfolgerung deckt sich mit den Erkenntnissen von Bothe [4] bei ähnlichen Untersuchungen von Phasenauswertungen. Eine Gruppe von Phasenschiebe-Algorithmen die beide Bedingungen erfüllen, sind symmetrische *M*-Schritt Algorithmen. Diese basieren auf gleichmäßig über eine volle Periode verteilten Phasenverschiebungen:

$$\tan\varphi = \frac{-\sum_{i=1}^{M} I_i \sin\psi_i}{\sum_{i=1}^{M} I_i \cos\psi_i}; \quad \psi_i = \frac{i2\pi}{M} \tag{9}$$

mit dem bekannten 4-Schritt Algorithmus für $M = 4$:

$$\tan\varphi = \frac{I_3 - I_1}{I_4 - I_2}; \quad \psi_i = \left(\frac{\pi}{2}; \pi; \frac{3\pi}{2}; 2\pi\right). \tag{10}$$

Für diese Algorithmen kann gemäß [2] ein einfacher Zusammenhang für das Phasenrauschen angegeben werden:

$$\sigma_\varphi = \sqrt{\frac{2}{M}} \frac{1}{\gamma} \frac{\sigma_I}{I'} = \sqrt{\frac{2}{M}} \frac{1}{\gamma} \frac{1}{SNR} \tag{11}$$

Dieser Ausdruck basiert auf der Annahme von signalunabhängigem normalverteiltem Intensitätsrauschen mit konstantem σ_I, wodurch das Signal-Rausch-Verhältnis linear mit der mittleren Intensität I' steigt. In Abschnitt 3 wird gezeigt, dass (11) darüber hinaus ebenfalls für das Signal-Rausch-Verhältnis bei komplexeren Kameramodellen gilt.

2.3 Erweitertes Kameramodell der EMVA 1288

Für die allgemeine Charakterisierung von Bildsensoren und Kameras wurde von der European Machine Vision Association die Norm EMVA 1288 herausgegeben. Sie basiert auf einem erweiterten Kameramodell für zeitliches Intensitätsrauschen, welches neben normalverteiltem Ausleserauschen auch Photonen- und Quantisierungsrauschen berücksichtigt. Darüber hinaus sind in der Norm auch Auswerteverfahren für örtliches Rauschen und Dunkelstrom beschrieben. Da die Phasenauswertung für jedes Pixel unabhängig möglich ist und bei den üblicherweise vorliegenden Belichtungszeiten der Einfluss des Dunkelstroms im Allgemeinen vernachlässigt werden kann, werden diese beiden Effekte im Weiteren nicht weiter berücksichtigt.

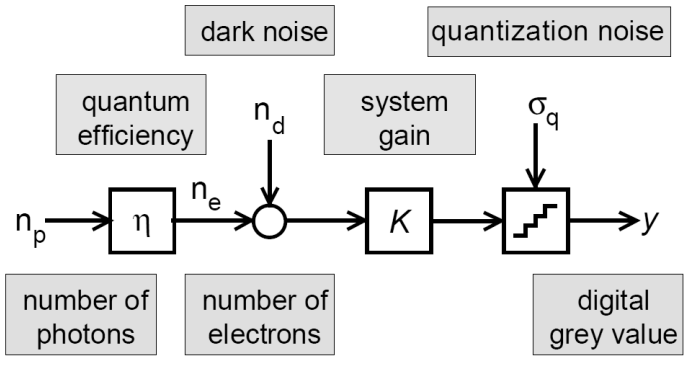


Bild 2: Erweitertes Kameramodell der EMVA 1288. Der gesamte Bildentstehungsprozess wird als linear angenommen. *Quelle: EMVA 1288*

Das grundlegende Kameramodell der EMVA 1288 ist in Bild 2 dargestellt. Der Prozess beschreibt die Umwandlung der Anzahl Photonen n_p in eine Anzahl Elektronen n_e mit der absoluten Quanteneffizienz η als Umwandlungswahrscheinlichkeit und der anschließenden Konvertierung in einen digitalen Grauwert in der Ausleseelektronik mit der Systemverstärkung K als Skalierungsfaktor. Drei Rauschquellen werden im Laufe der Umwandlung wirksam: die erzeugten Elektronen sind Poisson-verteilt ($\sigma_e^2 = \mu_e$), das Ausleserauschen ist normalverteilt (μ_d, σ_d) und die Quantisierung erzeugt gleichverteiltes Rauschen ($\sigma_q^2 = 1/12$). Aufgrund der vorausgesetzten Linearität des Konvertierungsprozesses können die Varianzen der Rauschquellen addiert werden um die resultierende Varianz des Ausgangssignals zu berechnen:

$$\sigma_y^2 = K^2\left(\sigma_e^2 + \sigma_d^2\right) + \sigma_q^2. \tag{12}$$

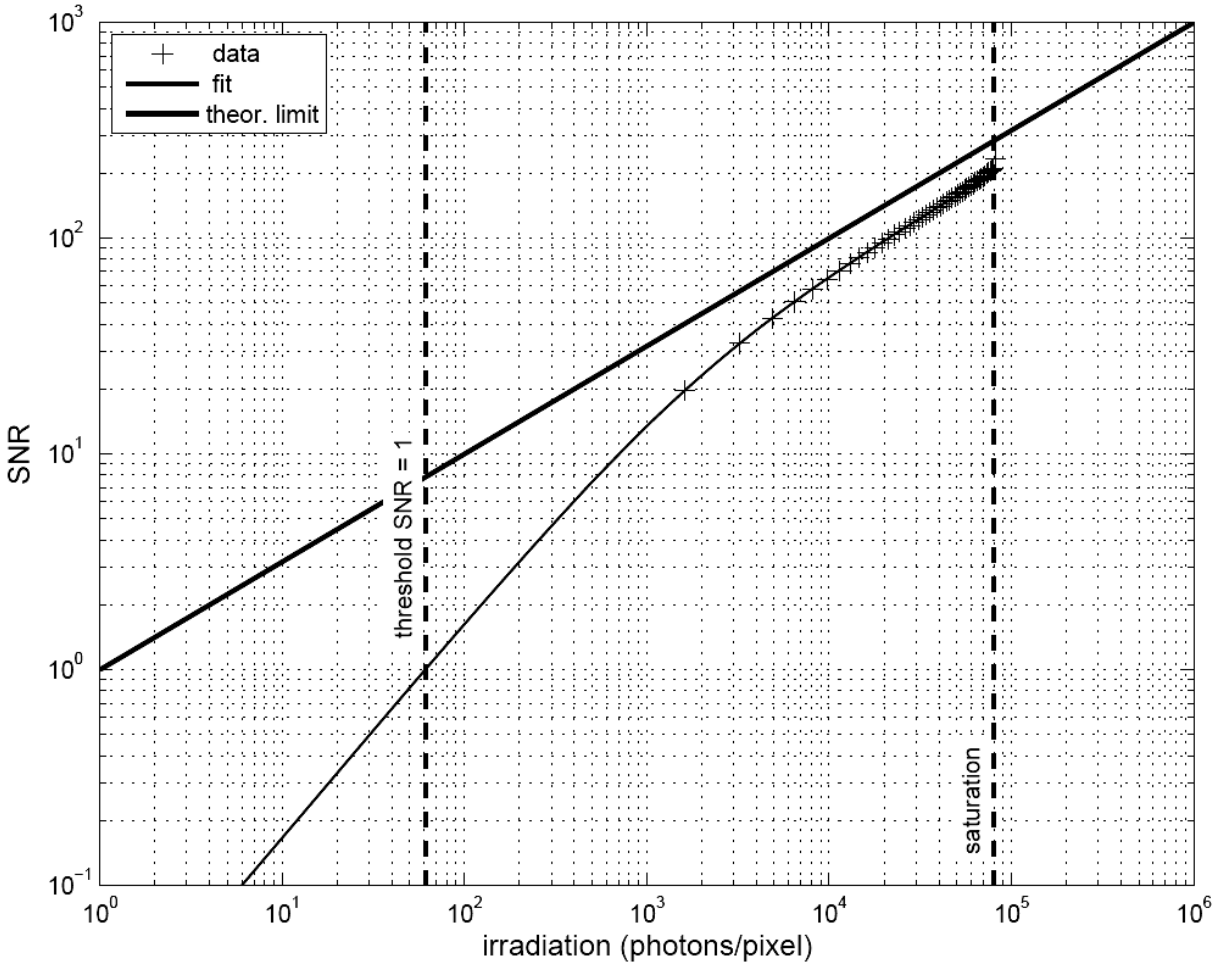


Bild 3: SNR-Diagramm einer typischen Kamera. Die Asymptoten entsprechen einem linearen Verlauf (Steigung 1) für geringe und einem Wurzelzusammenhang (Steigung ½) für hohe Beleuchtungsstärken. *Quelle: EMVA 1288*

Auf der Grundlage von (12) wird das Signal-Rausch-Verhältnis definiert als:

$$SNR = \frac{\mu_y - \mu_{y,d}}{\sigma_y} = \frac{\eta\mu_p}{\sqrt{\eta\mu_p + \sigma_d^2 + \frac{1}{12}\frac{1}{K^2}}} \tag{13}$$

Es ist erkennbar, dass die Beleuchtungsstärke μ_p auch unter der Wurzel im Ausdruck für σ_y vorhanden ist, so dass das Intensitätsrauschen signalabhängig wird. Für kleine Werte von μ_p ist das Signal-Rausch-Verhältnis daher proportional zur Beleuchtungsstärke und es dominiert der Einfluss des Ausleserauschens. Für große Werte von μ_p ist es proportional zur Wurzel aus der Beleuchtungsstärke und wird vom Photonenrauschen dominiert. Dieser Zusammenhang ist im Bild 3 beispielhaft dargestellt.

3 Phasenrauschmodell auf der Grundlage der EMVA 1288

Durch Kombination der in den Abschnitten 2.1, 2.2 und 2.3 vorgestellten Ergebnisse ist es möglich, ein erweitertes Phasenrauschmodell auf der Grundlage des EMVA 1288 Kameramodells zu formulieren. Im Ausdruck für das Signal-Rausch-Verhältnis (13) kann in Anlehnung an (4) ein Bezug zur Belichtung β und der Sättigungskapazität I_{sat} (hier Elektronensättigungskapazität $\mu_{e,sat}$) hergestellt werden durch den Zusammenhang:

$$\eta\mu_p = \mu_e = \beta\mu_{e,sat}. \tag{14}$$

Durch Einsetzen in den Ausdruck für das Phasenrauschen bei symmetrischen M-Schritt Algorithmen (11) ergibt sich eine Formel zur Vorhersage des Phasenrauschens, welche sowohl Auslese-, als auch Photonen- und Quantisierungsrauschen berücksichtigt:

$$\sigma_\varphi = \sqrt{\frac{2}{M}}\frac{1}{\gamma\beta\mu_{e,sat}}\sqrt{\beta\mu_{e,sat} + \sigma_d^2 + \frac{1}{12K^2}} \tag{15}$$

Hierin sind alle Parameter entweder fundamentale Größen welche die Messung beschreiben (β, γ, M) oder Kameraparameter die einem EMVA 1288 konformen Datenblatt entnommen werden können ($\mu_{e,sat}$, σ_d, K). Das erweiterte Modell wurde sowohl durch Monte-Carlo-Simulationen als auch experimentell vali-

diert. Für die Experimente wurde ein LCD-Monitor von der untersuchten Kamera beobachtet. Auf dem Display erfolgte die Darstellung mehrerer Muster für verschiedene Ausprägungen von β und γ mit einer anschließenden statistischen Auswertung von jeweils 1000 Phasenberechnungen. Die Ergebnisse der Untersuchungen zeigen eine deutlich bessere Übereinstimmung mit realen Messdaten als das einfache lineare Modell (11). Dies ist beispielhaft in Bild 4 für eine Kamera vom Typ „Vosskühler CCD-4000“ dargestellt. Die Abhängigkeit von der Sichtbarkeit γ wurde dabei durch Berücksichtigung auf der Ordinatenachse eliminiert. Dadurch liegen die Punkte für beliebige Werte von γ auf der gleichen Modellkurve. Zur Erhöhung der Lesbarkeit ist der Kehrwert des Phasenrauschens dargestellt, wodurch der Verlauf der Modellkurve annähernd wurzelförmig ist.

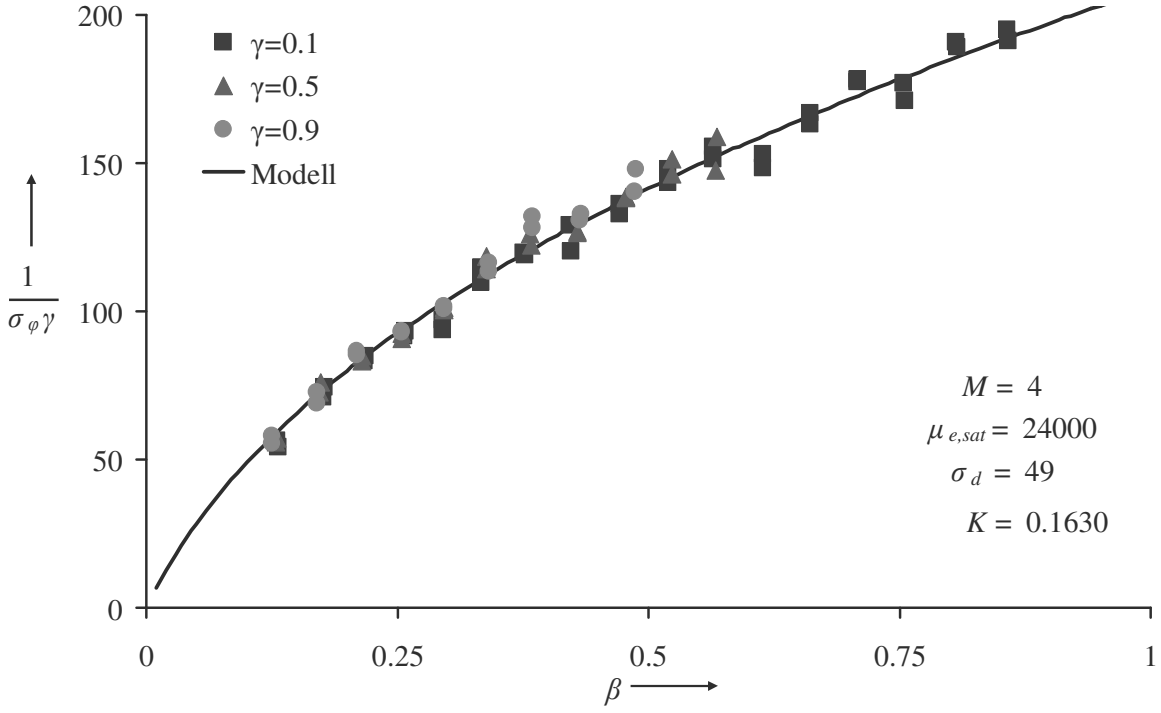


Bild 4: Experimentelle Validierung des erweiterten Phasenrauschmodells (15) für eine Kamera vom Typ “Vosskühler CCD-4000”. Das einfache Modell (11) würde einen in der Realität nicht zu beobachtenden linearen Verlauf vorhersagen.

Auf der Grundlage dieses erweiterten Phasenrauschmodells können die für eine gegebene Kamera zu erwartenden zufälligen Abweichungen bei der Phasenauswertung bereits im Vorfeld relativ genau anhand der im Datenblatt spezifizierten Kameraparameter abgeschätzt werden.

4 Zusammenfassung

Im Rahmen dieser Arbeit wurde ein erweitertes Phasenrauschmodell für optische Messsysteme mit strukturierter Beleuchtung vorgestellt. Bisherige einfache Modelle basierten häufig auf der Annahme von konstantem signalunabhängigem normalverteiltem Intensitätsrauschen. Ausgangspunkt war eine allgemeine Signalbeschreibung mit zwei fundamentalen Parametern, der relativen mittleren Intensität (Belichtung) und dem relativen Streifenkontrast (Sichtbarkeit). Es wurde gezeigt, dass symmetrische M-Schritt Algorithmen mit gleichmäßig über eine volle Periode verteilten Phasenverschiebungen die besten Ergebnisse in Bezug auf das Phasenrauschen liefern, da sie erwartungstreue Schätzer ohne phasenabhängige zufällige Abweichungen sind. Auf der Grundlage des erweiterten Kameramodells der Norm EMVA 1288 konnte für diese Klasse von Phasenschiebe-Algorithmen ein mathematischer Ausdruck zur Vorhersage des Phasenrauschens hergeleitet werden, der neben den fundamentalen Größen nur von Kameraparametern abhängt, die einem EMVA 1288 konformen Datenblatt entnommen werden können. Das erweiterte Phasenrauschmodell berücksichtigt neben normalverteiltem Ausleserauschen auch Poisson-verteiltes Photonenrauschen und gleichverteiltes Quantisierungsrauschen. Es sagt für die in typischen Messanwendungen vorkommenden Belichtungen, im Gegensatz zur linearen Vorhersage des einfachen Modells, einen annähernd wurzelförmigen Zusammenhang zwischen dem Kehrwert des Phasenrauschens und der Belichtung voraus. Dies konnte sowohl durch Monte-Carlo-Simulationen als auch experimentell validiert werden. Das Modell ermöglicht somit die Abschätzung der zufälligen Abweichungen bei der Phasenauswertung mit Hilfe des EMVA 1288 konformen Datenblatts einer Kamera.

Danksagung

Die Autoren danken der Deutschen Forschungsgemeinschaft (DFG) für die Förderung dieser Arbeit im Rahmen des Projekts Pe1402/2-1.

Literatur

[1] Rathjen, C.: Statistical properties of phase-shift algorithms. Journal of the Optical Society of America A, 12 (1995) 9, S. 1997-2008

[2] Surrel, Y.: Additive noise effect in digital phase detection. Applied Optics 36 (1997) 1, S. 271-276

[3] European Machine Vision Association: EMVA Standard 1288 – Standard for Characterization of Image Sensors and Cameras, Release 3.0, November 29, 2010

[4] Bothe, T.: Grundlegende Untersuchung zur Formerfassung mit einem neuartigen Prinzip der Streifenprojektion und Realisierung in einer kompakten 3D-Kamera, Dissertation, Vollertsen, F., Bergmann, R. (Hrsg.): Strahltechnik, Bd. 32, Bremen, BIAS Verlag, 2008, S. 52-60

[5] Erz, M.: Charakterisierung von Laufzeitkamerasystemen für Lumineszenzlebensdauermessungen, Dissertation, Heidelberg, Universität Heidelberg, 2011, S. 58-62

Verfahren und Vorrichtung zur Inspektion von Glasoberflächen

Sebastian Höfer[1], Masoud Roschani[1], Stefan Werling[2], Jürgen Beyerer[2]

[1] Karlsruher Institut für Technologie (KIT), Institut für Anthropomatik,
Lehrstuhl für Interaktive Echtzeitsysteme (IES),
Adenauerring 4, 76131 Karlsruhe,
E-Mail: {sebastian.hoefer, masoud.roschani}@kit.edu

[2] Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung,
Fraunhoferstraße 1, 76131 Karlsruhe,
E-Mail: {stefan.werling, juergen.beyerer}@iosb.fraunhofer.de

Kurzfassung

In dieser Arbeit wird ein Verfahren zur Inspektion der Oberflächengeometrie transparenter Oberflächen vorgestellt. Hierzu wird die Inspektionsmethode der Deflektometrie verwendet. Es wird gezeigt, wie sich die Probleme bei der Prüfung mit sichtbarem Licht durch den Einsatz von langwelligem Infrarotlicht umgehen lassen. Die Ergebnisse werden anhand realer Messungen veranschaulicht. Abschließend werden zwei Ansätze vorgestellt, die mittels Arrays aus Heizelementen oder Absorption von Laserlicht die dynamische Erzeugung thermischer Muster ermöglichen.

1 Einleitung

Computergestützte visuelle Inspektion ist ein wichtiger Bestandteil industrieller Fertigungsanlagen. Die vollständige und zerstörungsfreie Qualitätskontrolle erlaubt die stetige Überwachung und Optimierung des Produktionsablaufs. Insbesondere transparente Materialien stellen für die visuelle Prüfung eine Herausforderung dar.

Bisher werden bei der Vermessung transparenter Oberflächen hauptsächlich Durchlicht- und Polarisationsverfahren angewendet. Bei den Durchlichtverfah-

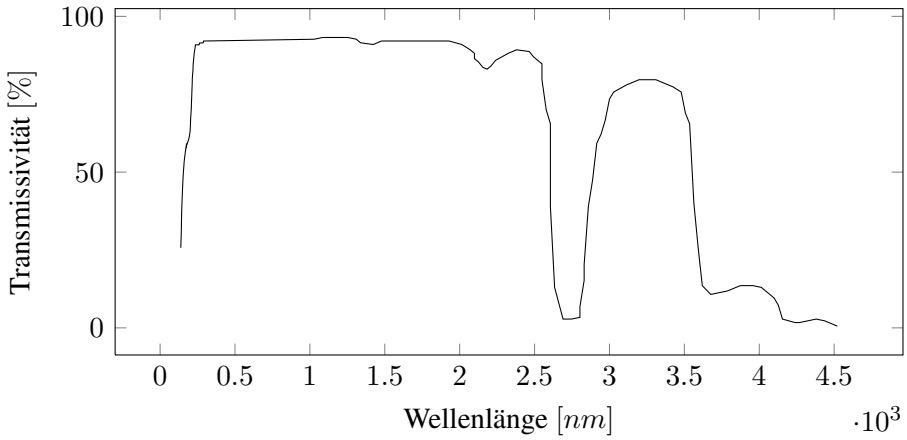


Bild 1: Transmissionskurve für herkömmliches Glas bei 10mm Dicke. Im langwelligen Infrarot- und kurzwelligen UV-Spektrum wird Glas opak. (Quelle: Almaz Optics, Inc.)

ren [1] wird die Verzerrung eines bekannten Musters durch das transparente Prüfobjekt mit einer Kamera beobachtet. Die Verzerrungen lassen Rückschlüsse auf die Form des Objekts zu, wobei allerdings nicht zwischen den Einflüssen der verschiedenen Oberflächen differenziert werden kann. Beim Polarisationsverfahren [2] wird ausgenutzt, dass Licht bei der Reflexion abhängig von der Neigung der Oberfläche polarisiert wird. Der Polarisationszustand dient als Indikator für die Normalenrichtung der Oberfläche. Ein Verfahren zur Vermessung von Fahrzeugverglasungen mittels Deflektometrie wird in [3] beschrieben. Hierbei tritt die Schwierigkeit auf, dass sich die Reflexionen von Vorder- und Rückseite der Scheibe überlagern. Dieses Problem wird durch eine Dekodierung, die aus einer Kombination aus Gray-Code und Phasenschiebeverfahren besteht, gelöst. In diesem Fall lassen sich dann die überlagerten Muster der Rückseitenspiegelung und der Oberflächenspiegelung trennen. Allerdings funktioniert dieses Verfahren nur bei Doppelreflexionen und eignet sich nicht für komplex geformte, transparente Objekte, bei denen Mehrfachreflexionen auftreten können.

Bei dem hier vorgestellten Verfahren werden die störenden, optischen Effekte bei transparenten Materialien vermieden, indem die spektralen Eigenschaften des Materials ausgenutzt werden. Glas verliert beispielsweise seine transparenten Eigenschaften im langwelligen Infrarotspektrum (siehe Bild 1). Es wird gezeigt, wie sich das deflektometrische Messprinzip im thermischen Infrarot anwenden lässt, um auch die Oberflächen von Glas und anderen transparenten Materialien zu untersuchen, ohne die Oberflächen speziell vorzubehandeln.

Der Artikel ist wie folgt gegliedert: In Kapitel 2 wird kurz das deflektometrische Verfahren vorgestellt. Die Ergebnisse der Experimente zur Vermessung transparenter Objekte mittels Deflektometrie im infraroten Spektrum werden im Kapitel 3 gezeigt. Anschließend werden in Kapitel 4 mögliche technische Umsetzungen thermischer Bildgeber zur Erzeugung dynamischer Infrarotmuster präsentiert.

2 Deflektometrie

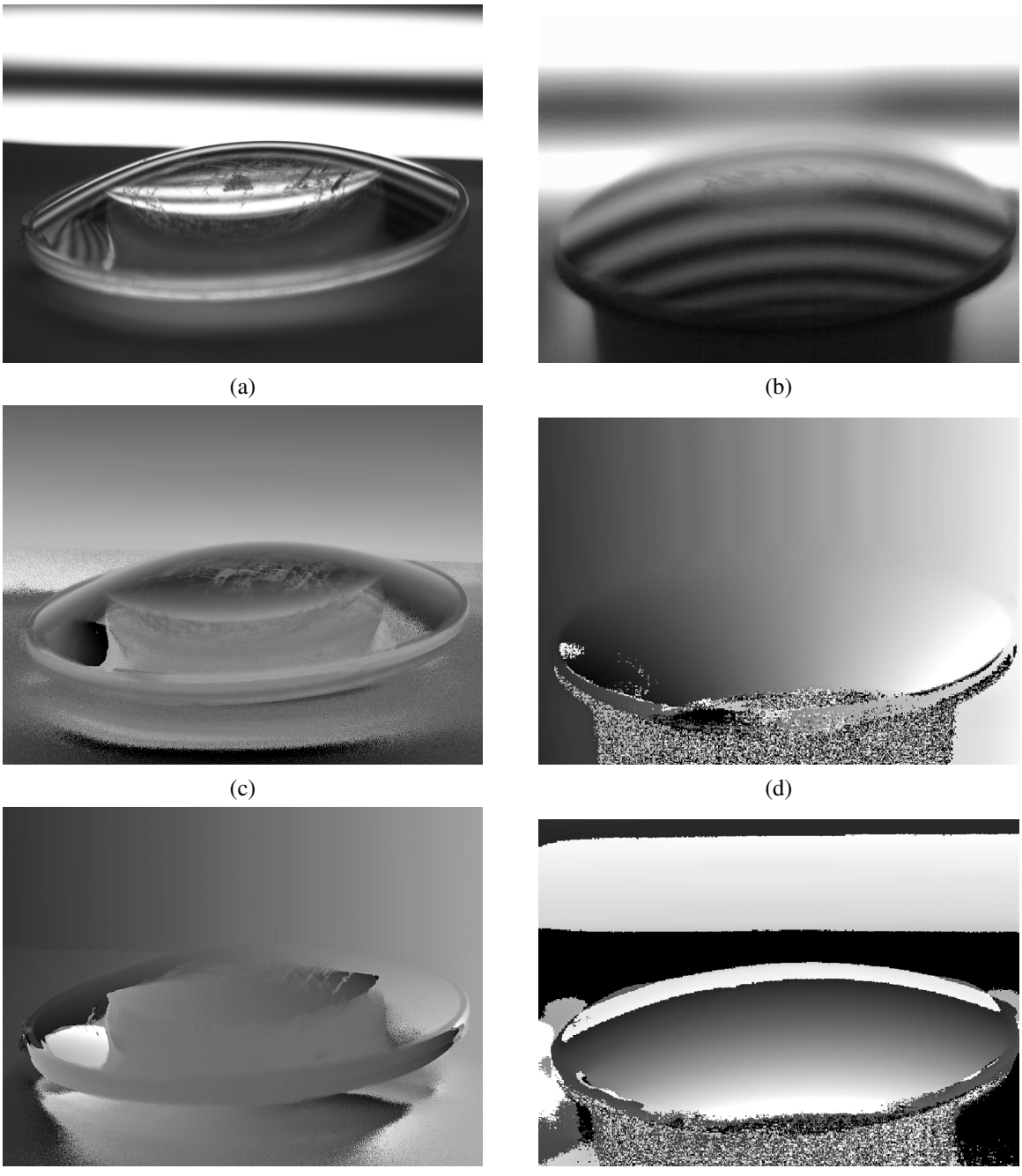
Zur Vermessung von Oberflächengeometrien eignet sich sehr gut das Verfahren der Deflektometrie [4, 5]. Hierbei wird eine Sequenz von Mustern zur Positionscodierung angezeigt. Mit einer Kamera wird die Spiegelung dieses Musters an der zu untersuchenden Oberfläche aufgenommen und ausgewertet. Deformationen der Oberfläche führen zu Verzerrungen in den dekodierten Mustern. Anhand dieser Daten lassen sich schon viele Oberflächendefekte erkennen. Nach einer Regularisierung der gewonnenen Daten ist sogar eine vollständige Rekonstruktion der Oberfläche möglich [6]. Dieses Verfahren bietet eine sehr hohe Empfindlichkeit gegenüber der Oberflächenneigung, die auch maßgeblich für die menschliche Wahrnehmung spiegelnder Oberflächen ist. Deflektometrische Verfahren eignen sich daher sehr gut für Sichtprüfaufgaben, bei denen ein objektiver Vergleich zur menschlichen Wahrnehmung spiegelnder Oberflächen erforderlich ist.

Die Positionscodierung kann mittels eines Phasenschiebeverfahrens gelöst werden [6]. Dabei wird ein Sinusmuster angezeigt und mit jeder Aufnahme phasenverschoben. Um die Genauigkeit der Messung zu erhöhen werden die Aufnahmen mit mehreren Wellenlängen wiederholt und bei der Phasenentfaltung der Sinusmuster wieder kombiniert. Phasenschiebeverfahren eignen sich besonders zum Einsatz bei der Deflektometrie wegen ihrer Robustheit gegenüber unscharfer Abbildung. Dadurch kann die Kamera auf die Oberfläche des Prüfobjektes fokussiert werden, während die Spiegelung des Musters nicht im Schärfebereich liegen muss.

3 Experimente

In diesem Abschnitt werden die Ergebnisse der Deflektometrie im Infrarotspektrum für die Prüfung transparenter Oberflächen gezeigt. Dazu wird die deflektometrische Messung transparenter Prüfobjekte im sichtbaren und im infraroten Spektrum miteinander verglichen. Der experimentelle Aufbau für die Infrarotdeflektometrie gleicht einem üblichen Aufbau zur Deflektometrie im sichtbaren Spektrum: ein Schirm zur Mustererzeugung, eine Kamera und Prüfobjekt. Eine Wärmebildkamera mit Mikrobolometersensor dient zur Bildaufnahme. Der Kamerasensor weist eine Auflösung von 640×480 Bildpunkten auf, bei einer thermischen Auflösung (*Noise equivalent temperature difference*) von 0,03 Kelvin. Zur Mustererzeugung im infraroten Spektrum wird eine beheizte Kupferplatte mit einem aufgedruckten Muster verwendet (siehe auch Kapitel 4.1). Mittels einer geregelten Wärmezufuhr wird eine gleichbleibende Temperatur der Platte sichergestellt. Für die Experimente wurde eine Temperatur von 20 Kelvin über der Umgebungstemperatur gewählt. Zur Positionscodierung wird das schon bei der herkömmlichen Deflektometrie etablierte Phasenschiebeverfahren benutzt. Hierzu wurden Aufnahmen mit zwei Wellenlängen und vier Phasenpositionen aufgenommen, jeweils in horizontaler und vertikaler Richtung. In den Rohdaten, siehe Bild 2b, erkennt man bereits die ungestörte Spiegelung des infraroten Musters an der Oberfläche des Prüfobjekts. Das Ergebnis der dekodierten Mustersequenz ist in den Bildern 2d, 2f zu sehen. Zum Vergleich wurde das gleiche Objekt mit sichtbarem Licht deflektometrisch vermessen. Schon in den Rohbildern, siehe Bild 2a, wird deutlich, dass die Rekonstruktion des Musters aufgrund der sich überlagernden optischen Effekte nicht möglich ist und somit die deflektometrische Auswertung fehl schlägt.

Die Ergebnisse zeigen, dass sich eine vollständige deflektometrische Messung auch mit thermischem Infrarotlicht durchführen lässt, wodurch sich die störenden Effekte im sichtbaren Spektrum vermeiden lassen. Hiermit bietet die Infrarotdeflektometrie eine geeignete Prüfmethode, um transparente Glasoberflächen berührungslos auf Defekte der Oberflächengeometrie zu überprüfen.



(a) (b) (c) (d) (e) (f)

Bild 2: Aufnahme einer Linse im sichtbaren Spektrum zwischen $400-700nm$ (a) und infraroten Spektrum zwischen $8 - 14\mu m$ (b). Die Dekodierung der deflektometrischen Messung zeigt im sichtbaren Spektrum deutliche Störungen (c),(e) während im Infraroten eine Messung der tatsächlichen Oberfläche erfolgt (d),(f).

4 Technische Umsetzung

Obwohl schon in [7] von Horbach gezeigt wurde, dass sich das Konzept der Deflektometrie auch im thermischen Infrarotspektrum anwenden lässt, wurde bis jetzt noch keine praktische Anwendung realisiert. Dies begründet sich hauptsächlich in dem Mangel an dynamischen Bildgebern im thermischen Infrarot. Derzeit etablierte Technik besteht hauptsächlich aus IR-Projektoren zum Test von IR-Sensorik [8], deren Verfügbarkeit sich jedoch auf militärische Anwendungen beschränkt. Solche Systeme basieren auf Mikrospiegelaktoren oder Mikrowiderstandsarrays und erreichen hohe Bildwiederholraten. Allerdings ist die erzielte Leuchtstärke nicht ausreichend für die Deflektometrie.

In diesem Kapitel werden einige Möglichkeiten gezeigt, wie sich ein geeigneter Bildgeber zur Anwendung für die Infrarotdeflektometrie realisieren lässt. Bei der praktischen Umsetzung im thermischen Infrarot stehen den Anforderungen an einen solchen Bildgeber, wie ein hoher Kontrast, hohe Auflösung und schneller Bildwechsel, den thermischen Eigenschaften der verwendeten Materialien gegenüber. Bei den einzelnen Verfahren wird darauf eingegangen, wie sich diese jeweils auswirken.

4.1 Statische Muster

Bei den Versuchen zu dieser Arbeit als auch in [7] wurden statische Wärmemuster benutzt. Solche Muster lassen sich vergleichsweise einfach herstellen, indem der Emissionsgrad unterschiedlicher Materialien ausgenutzt wird.

Polierte Metalloberflächen weisen im infraroten Spektrum einen sehr niedrigen Emissionsgrad auf. Dadurch strahlen sie kaum Wärmestrahlung an die Umgebung ab und erscheinen für eine Wärmebildkamera kalt. Gleichzeitig sorgt die gute Wärmeleitfähigkeit von Metallen für eine gleichmäßige Verteilung der Wärme im Material. Dadurch eignen sie sich sehr gut als Trägermaterial für statische Muster, ohne selbst das Wärmebild zu stören. Für die Versuche wurde das anzuzeigende Wärmemuster auf Kupferplatten gedruckt. Durch den höheren Emissionsgrad der Druckfarbe erscheinen diese Flächen im Vergleich zur Kupferoberfläche im Infrarotbild mit hoher Intensität, obwohl beide Flächen bei einer Kontaktmessung die gleiche Temperatur aufweisen. Auf diese Weise

lassen sich thermische Muster mit hoher Auflösung und hohem Kontrast herstellen.

Der offensichtliche Nachteil von statischen Mustern ist der Bildwechsel, bei dem das Muster entweder mechanisch ausgetauscht oder bewegt werden muss. Dabei steigt die Komplexität des Aufbaus mit der Anzahl an benötigten Mustern und Bewegungen. Eine Sequenz von Mustern zur Phasenschiebecodierung, wie sie bei der Deflektometrie verwendet werden, lässt sich dabei durch ein Muster für jede Wellenlänge und deren Verschiebung erzeugen. Für solche sich wiederholenden Mustersequenzen, die sich hauptsächlich in einer linearen Bewegung unterscheiden, eignet sich auch die Anordnung auf einem Endlosband, um den mechanischen Aufwand zu minimieren.

Bei Metallen als Träger für thermische Muster ist zu beachten, dass Metalle eine hohe Reflexivität im infraroten Spektrum aufweisen. Daher entstehen leicht Störungen, wenn sich Wärmequellen in der Umgebung oder die Kamera selbst in den metallischen Flächen spiegeln.

4.2 Heizelement Arrays

Ein naheliegender Ansatz zur dynamischen Bilderzeugung im thermischen Infrarot ist der Aufbau eines Arrays aus Thermoelementen. Im einfachsten Fall handelt es sich hierbei um eine ebene Anordnung von einzeln ansteuerbaren Widerständen. Jeder Widerstand entspricht dabei einem Pixel des Infrarotbildgebers. Abhängig von der durchfließenden Strommenge erwärmen sich die Widerstände, sodass sich mittels einer geeigneten Ansteuerung die abgestrahlte Wärme jedes einzelnen Pixels gezielt steuern lässt.

Bild 3a zeigt einen Versuchsaufbau bestehend aus einem Array aus SMD-Widerständen. Diese weisen eine kompakte Bauweise auf und ermöglichen eine einfache Montage. Dadurch lässt sich ein modularer Aufbau realisieren, sodass sich auch großflächige Bildgeber aufbauen lassen.

Während die Bildauflösung eines solchen Bildgebers nur von der Packungsdichte der verarbeiteten Widerstände abhängt, werden Kontrast und Bildfrequenz durch thermische Eigenschaften des Aufbaus begrenzt. Bildbereiche, die von hoher Intensität auf kalt wechseln, müssen zunächst abkühlen. Die-

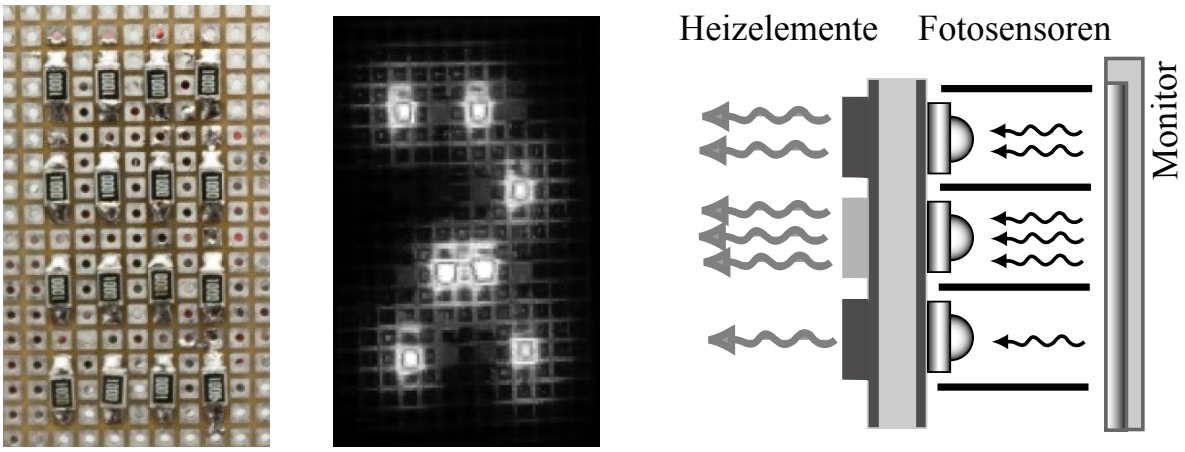


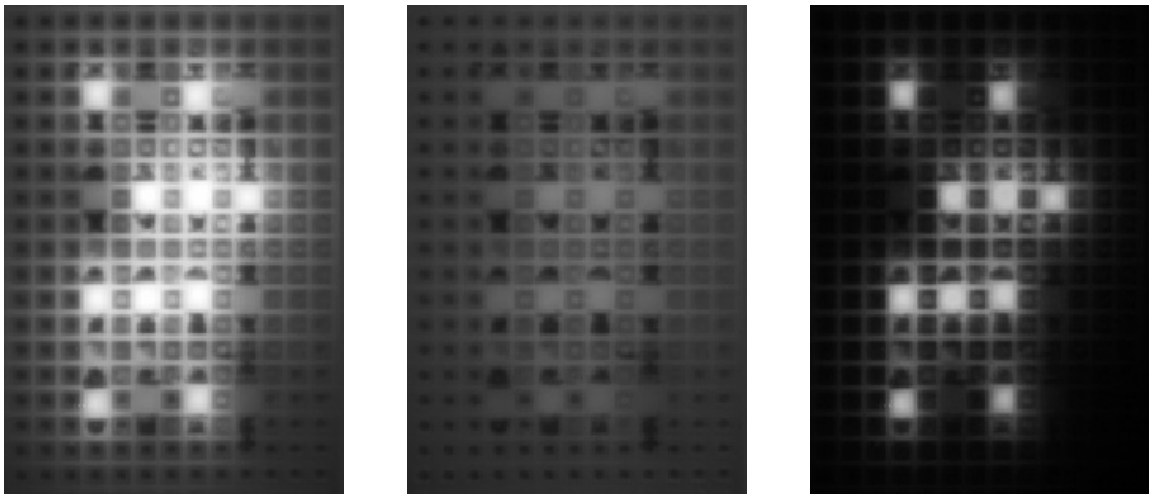
Bild 3: Array aus SMD-Widerständen als Heizelemente im sichtbaren (a) und infraroten (b) Spektrum. Stromdurchflossene Widerstände sind im thermischen Infrarot deutlich sichtbar. Bild (c) zeigt den Aufbau als spektralen Konverter, wobei jeder Widerstand einzeln über die Helligkeit eines Monitors angesteuert werden kann. Die lichtempfindlichen Elemente sorgen über geeignete Verstärker für ausreichend Ströme, um die jeweiligen Widerstände zu erwärmen.

ser Prozess ist außer von der Wärmekapazität der Widerstände zusätzlich davon abhängig, wie gut die Wärme bei dem gegebenen Aufbau abgeführt werden kann. Dieser Wärmetransfer führt jedoch auch dazu, dass sich die Umgebung eines aktiven Pixels mit erwärmt, was zu einem Übersprechen zwischen einzelnen Bildpunkten führt.

Damit die zurückbleibende Wärme nicht den Bildkontrast mindert und um das Löschen eines angezeigten Bildes zu beschleunigen, kann mittels Differenzbildung mit einem Referenzbild vor der eigentlichen Anzeige des Musters, das Muster hervorgehoben werden (siehe Bild 4). Mit einer aktiven Kühlung des Aufbaus können der Kontrast und die Zeit zum Löschen eines Bildes zusätzlich verbessert werden.

Eine Erweiterung dieses Konzepts ist die Verwendung von kleinen Peltierelementen anstelle von Widerständen. Die einzelnen Pixel ließen sich dadurch nicht nur heizen, sondern auch aktiv kühlen. Hierdurch wären schnellere Bildwechsel möglich, indem die Wärme gezielt von der Bildfläche abgeführt wird.

Eine alternative elektrische Ansteuerung für das Array lässt sich realisieren, indem jedes Thermoelement eine eigene Ansteuerung in Form eines lichtempfindlichen Widerstandes oder Fototransistors erhält. Der Aufbau entspricht ei-



(a) (b) (c)

Bild 4: Im Betrieb erwärmt sich das Trägermaterial des Widerstandarrays und mindert den Kontrast des thermischen Bildes (a). Durch Differenzbildung mit einem Referenzbild (b) vor dem Anzeigen des Musters, lässt sich das Muster besser hervorheben (c).

nem Konverter, der auf der einen Seite auf sichtbares Licht reagiert und auf der anderen Seite ein entsprechendes Wärmemuster erzeugt (siehe Bild 3c). Die Ansteuerung geschieht mittels eines herkömmlichen Monitors, dessen Bild direkt in ein thermisches Bild umgewandelt wird. So kann ein deflektometrischer Aufbau für das sichtbare Spektrum mittels Konverter und Austausch der Kamera für den Einsatz im infraroten Spektrum umgerüstet werden.

4.3 Erzeugung thermischer Muster mittels Laser

Im sichtbaren Spektrum bietet der Einsatz von Projektoren als Bildgeber den Vorteil, dass sich große Anzeigeflächen relativ einfach realisieren lassen. Für das thermische Infrarot existiert keine vergleichbare Projektortechnik. Ein Ansatz zur Realisierung einer Projektion thermischer Muster ist der Einsatz eines leistungsstarken Lasers zur Mustererzeugung. Solche Laser werden in vielen Bereichen zur Materialbearbeitung eingesetzt und sind Leistungsklassen von wenigen Watt bis mehreren tausend Watt gebräuchlich. Die Position des Laserstrahls lässt sich hierbei schnell und präzise mittels eines Spiegelscanners steuern. Dadurch ergeben sich zwei mögliche Ansätze thermische Muster zu generieren.

Einerseits können die Muster direkt mit dem Laser geschrieben werden, wobei

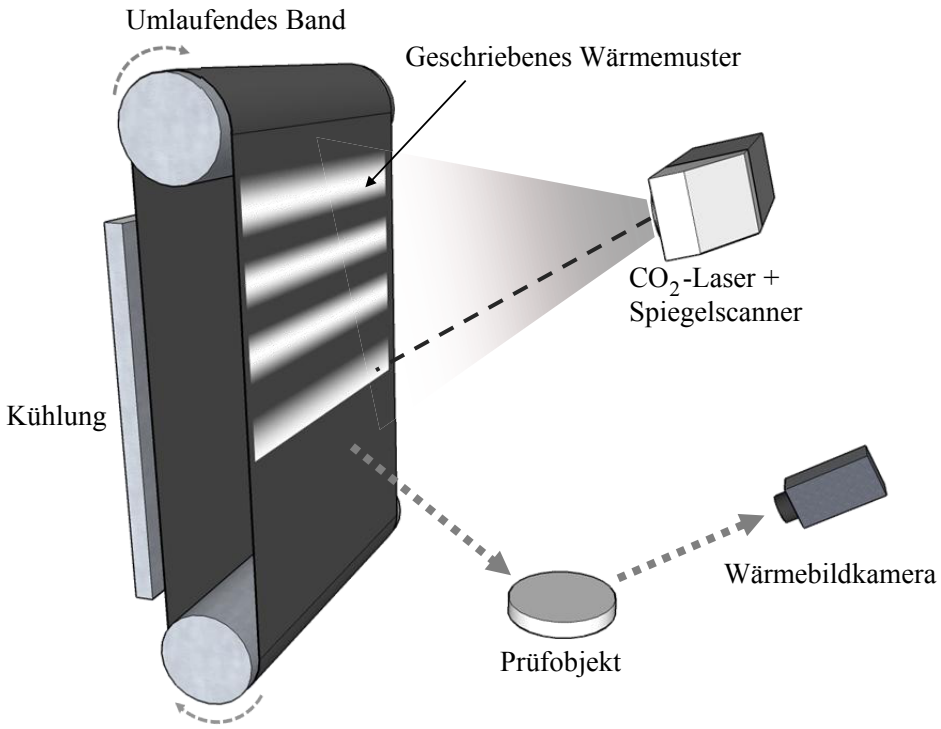


Bild 5: Thermische Mustererzeugung durch Absorption von Laserlicht. Der Laser "schreibt" das anzuzeigende Bild auf eine Projektionsfläche, die als Endlosband ausgeführt ist. Das Band kann weiterbewegt werden, um ein neues Muster zu schreiben, während das vorherige auf der Rückseite abkühlt.

der Laser eine Wellenlänge im spektralen Empfindlichkeitsbereich der Wärmebildkamera aufweisen muss. Hierzu eignen sich im besonderen CO_2-Laser, die mit einer Wellenlänge von $10,6\mu m$ genau in dem Empfindlichkeitsbereich von Wärmebildkameras für das langwellige Infrarotspektrum zwischen $8 - 14\mu m$ liegen. Zur Darstellung eines Bildes schreibt der Laser das Muster auf eine diffus streuende Projektionsfläche. Während des Schreibvorganges bleibt die Kamera geöffnet und nimmt die zeitliche Integration vor. Gegebenenfalls müssen hierbei mehrere Kameraaufnahmen fusioniert werden, falls die Zeit für einen Bildaufbau des Laserscanners über der Belichtungszeit der Wärmebildkamera liegt. Effektiv erhält man so eine Projektion im thermischen Infrarot, sodass über die Reflexion des Musters in der Oberfläche eines zu untersuchenden Objektes eine deflektometrische Messung erfolgen kann. Durch die hohe Intensität des Laserlichts ist hierbei ein Laser mit wenigen Watt Leistung ausreichend.

Eine weitere Möglichkeit ist, die thermischen Muster mit einem Laser indi-

rekt zu erzeugen. Wie auch bei dem vorherigen Ansatz wird das Muster dabei von dem Laser auf eine Projektionsfläche geschrieben. Hierbei wird die Projektionsfläche durch Absorption des Laserlichts erhitzt. Das thermische Muster entsteht wiederum durch die Emission von Wärmestrahlung von der lokal erhitzten Projektionsfläche. Für ein hinreichend sichtbares Muster wird das Material dabei um etwa 10 Kelvin erwärmt. Für die Eigenschaften der Projektionsfläche bedeutet dies, dass sie einerseits der kurzzeitigen, intensiven Erhitzung durch den Laserstrahl standhalten muss. Andererseits soll die Wärmeleitfähigkeit möglichst gering sein, damit ein geschriebenes Muster nicht durch Wärmediffusion zu schnell zerfließt. Schnelle Bildwechsel sind mit einer festen Projektionsfläche jedoch nicht möglich, da das Löschen eines Bildes dem Abkühlen des Musters entspricht. Um diesen Prozess zu beschleunigen, bietet es sich an, die Projektionsfläche als ein Endlosband zu realisieren, sodass die beschriebene Fläche wegbewegt werden kann, um auf der nicht sichtbaren Seite abzukühlen, während der Laser das nächste Bild schreibt (siehe Bild 5). Das Abkühlen lässt sich weiter beschleunigen, indem das Band der Projektionsfläche rückseitig aktiv gekühlt wird.

Da die geschriebenen Muster zumindest für einen kurzen Zeitraum persistent sind, hat dies den Vorteil, dass ein Muster für mehrere Aufnahmen einer Phasenschiebesequenz verwendet werden kann. Die Phasenverschiebung des Musters wird hierbei durch die Verschiebung des Bandes ermöglicht.

5 Fazit und Ausblick

Es wurde gezeigt, dass sich das deflektometrische Messverfahren auch auf transparente Materialien anwenden lässt, indem die spektralen Eigenschaften des Materials ausgenutzt werden. Dazu ist keine spezielle Beschichtung der Prüfobjekte notwendig, da Glas im thermischen Infrarot vollständig opak ist. Dadurch lassen sich Störungen, wie interne Reflexionen und Lichtbrechung, vollständig eliminieren. Nachdem vorangegangene Arbeiten [7] schon die Anwendbarkeit bei rauen, metallischen Oberflächen gezeigt haben, eröffnet sich mit der Inspektion transparenter Materialien ein weiteres praktisch wichtiges Anwendungsgebiet für die Infrarotdeflektometrie.

Zusätzlich wurden Ansätze zur Erzeugung der dynamischen thermischen Infrarotmuster vorgestellt. Sie bilden die Grundlage für die praktische Umsetzung eines deflektometrischen Prüfsystems für das thermische Infrarot [9].

Literatur

[1] J. Wang und A. K. Asundi, "A computer vision system for wineglass defect inspection via gabor-filter-based texture features", *Inf. Sci.* **127**, S. 157–171, August 2000.

[2] M. Saito, Y. Sato, K. Ikeuchi und H. Kashiwagi, "Measurement of surface orientations of transparent objects using polarization in highlight", in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, **1**, S. 2 vol. (xxiii+637+663), 1999.

[3] J. Xu, N. Xi, C. Zhang, Q. Shi und J. Gregory, "A geometry and optical property inspection system for automotive glass based on fringe patterns", *Optica Applicata* **40**, S. 827–841, 2010.

[4] S. Werling, J. Balzer und J. Beyerer, "A new approach for specular surface reconstruction using deflectometric methods", in *37. Jahrestagung der Gesellschaft für Informatik e.V. (GI): INFORMATIK 2007 - Informatik trifft Logistik*, **1**, S. 44–48, Sept. 2007.

[5] S. Werling, M. Mai, M. Heizmann und J. Beyerer, "Inspection of specular and partially specular surfaces", *Metrology and Measurement Systems* **16**(3), S. 415–431, 2009.

[6] S. B. Werling, *Deflektometrie zur automatischen Sichtprüfung und Rekonstruktion spiegelnder Oberflächen*. PhD thesis, 2011.

[7] J. Horbach und S. Kammel, "Deflectometric inspection of diffuse surfaces in the far–infrared spectrum", in *Machine Vision Applications in Industrial Inspection XIII*, J. R. Price und F. Meriaudeau (Hrsg.), *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **5679**, S. 108–117, 2005.

[8] P. Bryant, S. Solomon und J. James, "Bolometers running backward: the synergy between uncooled ir sensors & dynamic ir scene projectors", *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVII* **6207**(1), S. 62070J, SPIE, 2006.

[9] J. Beyerer, M. Heizmann und S. Werling, "Konzept zur Erzeugung eines räumlich und/oder zeitlich veränderbaren thermischen Strahlungsmusters", *European Patent EP 10191409.1* , 11 2009.

Integrierte CMOS PIN-Fototransistoren für kurzwelliges infrarotes Licht

Plamen Kostov, Horst Zimmermann

TU Wien, Institute of Electrodynamics, Microwave and Circuit Engineering
Tel.: 0043-1-58801-354624, e-mail: plamen.kostov@tuwien.ac.at,
horst.zimmermann@tuwien.ac.at

Kurzfassung

Dieser Artikel berichtet über PIN PNP Fototransistoren mit einer zusätzlichen intrinsischen Schicht, zur Erhöhung der Bandbreite und dadurch zu einer reduzierten Anstiegszeit. Die Fototransistoren wurden in Größen von 10×10 μm^2 und 100×100 μm^2 in einem 0,6 µm OPTO ASIC CMOS Prozess hergestellt. Durch die Realisierung verschiedener Designs der Basis und des Emitters, entstanden Fototransistoren mit unterschiedlichen Eigenschaften. Die Fototransistoren erzielten Empfindlichkeiten bis zu 35,4 A/W und Bandbreiten bis zu 15,4 MHz. Weiters wurden minimale Anstiegszeiten von bis zu 23 ns gemessen. Durch die Herstellung in einem CMOS Prozess können billige optoelektronische integrierte Schaltungen hergestellt werden. Mögliche Anwendungsgebiete sind dabei Lichtschranken, optische Empfänger, Optokoppler, usw.

1 Einleitung

Die Konvertierung von optischen Signalen in elektrische ist in den letzten Jahrzehnten zu einem wichtigen Bestandteil in der Elektrotechnik geworden. Diese Konvertierung wird mithilfe von Fotodetektoren durchgeführt. Heutzutage stellen Fotodioden (PN und PIN), Lawinenfotodioden (APDs) sowie Fototransistoren die wichtigsten Fotodetektoren dar.

1.1 PN-Fotodioden

PN-Fotodioden werden durch die Implementierung einer p- bzw. n-Wanne in einem n- bzw. p-Substrat oder durch ein p^+/n^+-Gebiet in einer n/p-Wanne hergestellt. Durch die Dotierung der einzelnen Gebiete, stellt sich eine relativ dünne Raumladungszone (RLZ) von < 2 µm bei Sperrspannung < 5V ein. Da bei kurzwelligem Infrarotlicht die Eindringtiefe relativ groß ist (850 nm hat eine 1/e Eindringtiefe von ca. 16,6 µm [1]), erfolgt der Großteil der Elektronen-Loch-Paar Generation bei PN-Fotodioden im Substrat, also außerhalb des schnellen Driftbereiches (RLZ-Bereich). Dadurch sind PN-Fotodioden mit einem relativ großen Diffusionsanteil des Stromes gekennzeichnet, der wiederum eine kleine Bandbreite bewirkt.

1.2 PIN-Fotodioden

Bei PIN-Fotodioden wird die Bandbreitenlimitierung der PN-Fotodioden durch eine zusätzliche 10-15 µm dicke, niedrig dotierte Epi-Schicht erhöht. Aufgrund dieser Epi-Schicht vergrößert sich die Raumladungszone der Fotodiode. Beim Anlegen einer Sperrspannung von einigen Volt vergrößert sich die Raumladungszone weiter, wodurch der Driftanteil des Fotostromes steigt und dadurch auch die Bandbreite der Fotodiode vergrößert wird. Mit Bandbreiten bis zu einigen GHz eignet sich die PIN-Fotodiode perfekt für Anwendungen wie z.B. Abstandsmessung [2], schnelle optische Empfänger [3], usw.

1.3 Lawinenfotodioden - APD

Da Fotodioden keine interne Verstärkung haben, ist ihre Empfindlichkeit, ohne Antireflektionsbeschichtung des fotosensitven Bereichs, auf ca. 0,33 A/W bei 850 nm limitiert [4]. Konventionelle Lawinendurchbruchfotodioden (APDs) und Fototransistoren können durch ihren internen Verstärkungsmechanismus höhere Empfindlichkeiten als Fotodioden erreichen. APDs verstärken dabei das Signal mittels ihres internen Lawinendurchbrucheffekts, durch welchen die generierten Ladungsträger multipliziert werden. Ein großer Nachteil bei APDs sind die für die Verstärkung benötigten hohen Spannungen, da sie nur schwer auf integrierten Schaltungen zu bewältigen sind [5]. Ein weiterer Nachteil ist die starke Hintergrundlichtverstärkung, die zu einer schnellen Sättigung der

Schaltung führen kann. In [6] ist eine APD mit einer Empfindlichkeit von 4,6 A/W bei 430 nm und 19,5 V Sperrspannung beschrieben.

1.4 Fototransistoren

Im Gegensatz zu APDs brauchen Fototransistoren für ihre interne Verstärkung keine hohen Spannungen. Niedrigere Spannungen sind ein wesentlicher Vorteil bei integrierten System-on-Chip Applikationen.

Ein Fototransistor besteht aus einer Fotodiode, welche mittels der Basis-Kollektor Diode des Fototransistors gebildet wird und einem internen Bipolartransistor zur Verstärkung des generierten Fotostromes (Bild 1). Hierbei werden die Ladungen, welche in der Basis-Kollektor Diode generiert wurden, vom herrschenden elektrischen Feld getrennt und gelangen zur Basis bzw. zum Kollektor. Bei einem PNP Fototransistor lädt sich die Basis durch die Ansammlung von Elektronen negativ auf. Wird die Diffusionsspannung erreicht, beginnt der p^+ Emitter positiv geladene Ladungsträger (Löcher) in die Basis zu injizieren. Dieser Prozess führt zur Verstärkung des primären Fotostroms. Durch die Dotierstoffkonzentration der einzelnen Schichten haben konventionelle Fototransistoren die gleichen Nachteile wie PN-Dioden und sind somit in der Bandbreite begrenzt. In [7] sind Fototransistoren mit einer Empfindlichkeit von 2,7 A/W bei 850 nm beschrieben. Der Grund für diese geringe Empfindlichkeit ist die relativ dünne Basis-Kollektor RLZ von ca. 1 µm.

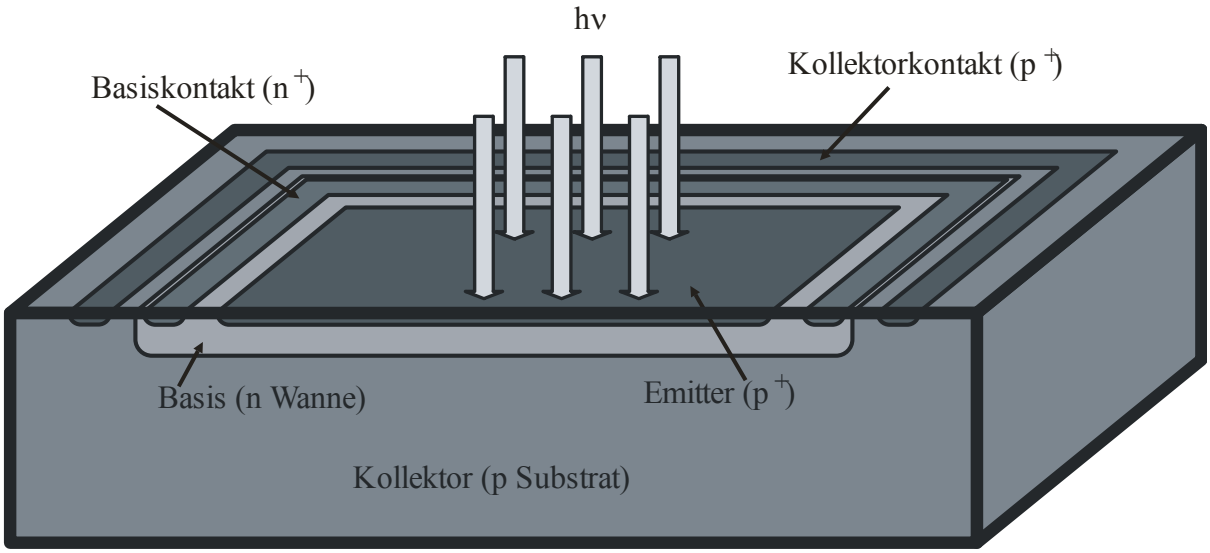


Bild 1: Querschnitt eines konventionellen PNP Fototransistors

2 Aufbau Fototransistor

Um die Vorteile des Fototransistors und der PIN-Fotodiode zu kombinieren und dadurch schnelle und hoch empfindliche optoelektrische Konverter zu erzeugen, wurden einige 10×10 μm^2 und 100×100 μm^2 PIN-PNP-Fototransistoren in einem 0,6 µm OPTO ASIC CMOS Prozess hergestellt. Die Implementierung einer ca. 15 µm dicken, niedrig dotierten p^- Epi-Schicht (2×10^{13} cm^{-3}) und einer ca. 1 µm dicken, niedrig dotierten n^- Epi-Schicht (1×10^{14} cm^{-3}) zwischen Basis und Kollektor führt zu einer dickeren Basis-Kollektor Raumladungszone und dadurch zu einem schnelleren Fototransistor. Diese dicke Raumladungszone ist notwendig um Ladungsträger, die durch kurzwelliges infrarotes Licht (z.B. 850 nm) tief im Fototransistor generiert werden, schnell trennen zu können. Um Fototransistoren mit unterschiedlichen Eigenschaften zu bekommen wurde bei den Fototransistoren das Layout der Basis und des Emitters variiert. Durch die Verwendung von Standard CMOS Technologie besteht die Möglichkeit der billigen Implementierung dieser PIN-Fototransistoren in optoelektronische-integrierte-Schaltungen. Mögliche Anwendungsgebiete sind hierbei hochempfindliche optische Sensoren, Bildsensoren, System-on-Chip Applikationen für Abstandsmessungen, Optokoppler, usw.

2.1 Kollektor Layout

Der Kollektor wird vom p Substrat gebildet und mittels einer großen Anzahl von Substratkontakten entlang der Kante des Fototransistors kontaktiert. Da der Kollektor auf Substratpotential liegt, können die beschriebenen PNP Fototransistoren nur als Emitterfolger in integrierten Schaltungen verwendet werden.

2.2 Basis Layout

Die Basis wird von der 1 µm dicken n^- Epi-Schicht gebildet. Durch die Implementierung verschieden dicker n-Wannenstreifen mit unterschiedlichen Abständen zueinander, kann die Basisdotierung eingestellt werden. Ein Querschnitt der implementierten PNP Fototransistoren ist in Bild 2 dargestellt. Die realisierten Fototransistoren wurden mit zwei unterschiedlichen Basisdotierungen entworfen:

- 33_B: In der n^- Epi-Schicht implantierte N-Wannenstreifen mit 1 µm Breite und 2 µm Abstand zu einander bilden die Basis dieser Fototransistoren.

- 50_B: N-Wannenstreifen mit 1 µm Breite und 1 µm Abstand zu einander bilden die Basis dieser Fototransistoren.

2.3 Emitter Layout

Der Emitter wird durch eine p^+ Schicht realisiert. Es wurden Fototransistoren mit zwei unterschiedlichen Designs entworfen (Bild 3):

- S_E: Der Emitter besteht aus einer 1,4×1,4 μm^2 kleinen p^+ Fläche. Bei diesem Design liegt der Emitter im Zentrum des fotosensitiven Bereichs.

- F_E: Bei diesem Design ist der Emitter als vollflächigen p^+ Bereich über die gesamte fotosensitive Fläche implementiert.

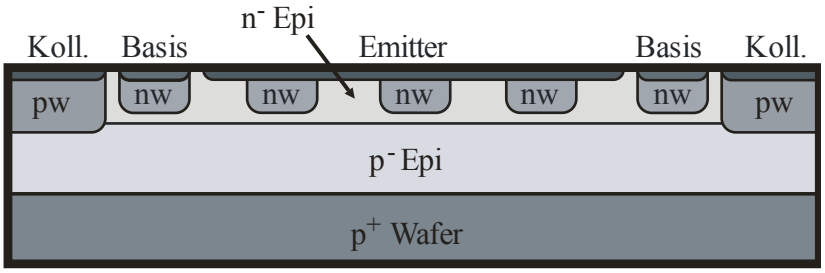


Bild 2: Querschnitt der implementierten PNP Fototransistoren

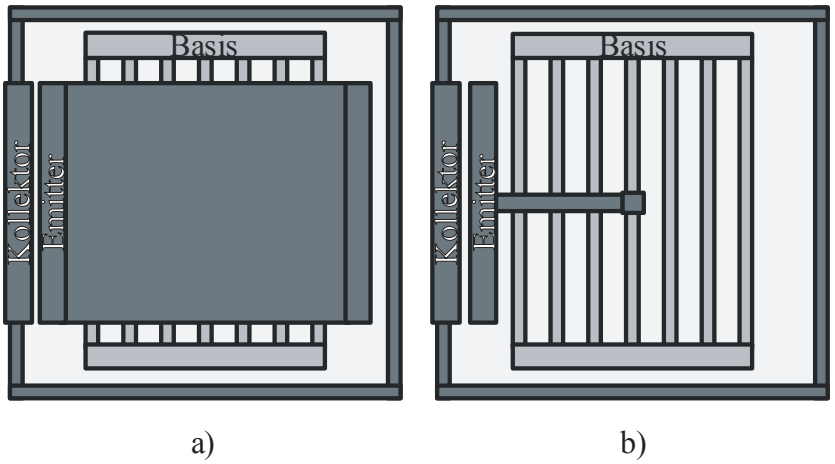


Bild 3: Aufsicht der PNP Fototransistoren mit unterschiedlichen Emitterdesign
a) vollflächiger Emitter F_E, b) kleiner zentraler Emitter S_E

2.4 Frequenzverhalten

Durch die Struktur des Fototransistors entstehen bei den Übergängen zwischen p- und n-Schicht Sperrschichtkapazitäten. Ein Fototransistor enthält somit eine Basis-Kollektor (C_{BC}) und eine Basis-Emitter (C_{BE}) Sperrschichtkapazität. Abhängig von den Dicken der Raumladungszonen haben die Sperrschichtkapazitäten dementsprechende Kapazitätswerte. Die Dicken der Raumladungszonen hängen wiederum von den Dotierungen der p- bzw. n-Schicht ab. Da zwischen Basis und Kollektor die niedrig dotierte p^- Epi-Schicht liegt, ist die Basis-Kollektor Sperrschichtkapazität C_{BC} meisten kleiner als die Basis-Emitter Sperrschichtkapazität C_{BE}. Die Sperrschichtkapazitäten sind der Hauptfaktor für das Frequenzverhalten des Fototransistors. Die -3 dB Grenzfrequenz f_{-3dB} entspricht nach [8]

$$f_{-3dB} = \frac{1}{2\pi\beta \cdot \left(\tau_B + \frac{k_B T}{q I_E} \left(C_{BE} + C_{BC} \right) \right)} . \tag{1}$$

Dabei ist β die Stromverstärkung, τ_B die Transitzeit der Minoritätsladungsträger durch die Basis, k_B die Boltzmannkonstante, T die absolute Temperatur, q die Elementarladung, I_E der Emitterstrom und C_{BE} und C_{BC} die Sperrschichtkapazitäten des Fototransistors. Wie in Gleichung (1) ersichtlich, sind kleine Sperrschichtkapazitäten notwendig um eine hohe -3 dB Grenzfrequenz zu erreichen. Eine Verkleinerung der Sperrschichtkapazitäten kann mittels Änderung der Dotierung der Basis bzw. durch einen kleineren Emitter realisiert werden. Kleinere Emitter führen jedoch zu kleineren Stromverstärkungen wie unter 2.5 beschrieben ist.

2.5 Stromverstärkung

Die Stromverstärkung ist stark von der effektiven Basisdicke W_B abhängig. Die effektive Basisdicke W_B ist der Abstand zwischen beiden Raumladungszonen in der Basis. Die Ausbreitung beider Raumladungszonen in der Basis ist dabei von der Dotierung der Basis abhängig. Diese kann, wie unter 2.2 beschrieben, durch unterschiedliches Design der Basis variiert werden. Da die Stromverstärkung indirekt proportional zur effektiven Basisdicke ist, ist für eine hohe Stromverstärkung eine niedrige Dotierung der Basis notwendig. Die Stromverstärkung β des PNP Fototransistors ist wie folgt definiert [9]:

$$\beta = \frac{1}{\frac{W_B^2}{2\tau_b D_p} + \frac{D_n}{D_p}\frac{W_B}{L_n}\frac{N_D}{N_A}}. \qquad (2)$$

Dabei ist β die Stromverstärkung, W_B die effektive Basisdicke, τ_b die Lebensdauer der Minoritätsladungsträger in der Basis, D_p der Diffusionskoeffizient der Löcher in der Basis, D_n der Diffusionskoeffizient der Elektronen in der Basis, L_n die Diffusionslänge der Elektronen im Emitter, N_D die Donatordichte in der Basis und N_A die Donatordichte im Emitter.

Bei vollflächigen Emitter (F_E) herrscht ein nahezu homogenes elektrisches Feld zwischen Basis und Emitter über die gesamte fotosensitive Fläche. Dadurch müssen die generierten Ladungsträger nur kurze Distanzen überwinden um den Emitter zu erreichen. Hingegen müssen die meisten generierten Ladungsträger bei kleinflächigen Emittern (S_E) längere Distanzen zurücklegen um zum Emitter zu gelangen. Dieser Umstand erhöht die Wahrscheinlichkeit für Rekombination in der Basis und bewirkt dadurch eine Reduktion der Empfindlichkeit.

3 Messergebnisse

Die elektrischen und optischen Eigenschaften der Fototransistoren wurden mittels Gummelplotmessungen bzw. optischen AC Messungen bei verschiedenen Arbeitspunkten eruiert.

3.1 Gummel Messungen

Gummelmessungen wurden durchgeführt um die elektrische Stromverstärkung des Fototransistors zu charakterisieren. In Bild 4 ist die Stromverstärkung der 100×100 µm^2 Fototransistors 33_BS_E (33% dotierte Basis mit kleinen zentralen Emitter) und 50_BS_E (50% dotierte Basis mit kleinen zentralen Emitter) dargestellt. Hierbei wurden maximale Verstärkungen zwischen 107 und 176 bei Kollektorströmen $I_C < 11$ nA erreicht. Wie bereits im Absatz 2.5 erklärt, haben Fototransistoren mit einer niedrig dotierten Basis und die daraus resultierende dünne effektive Basisdicke eine höhere Stromverstärkung als Fototransistoren mit höherer Basisdotierung.

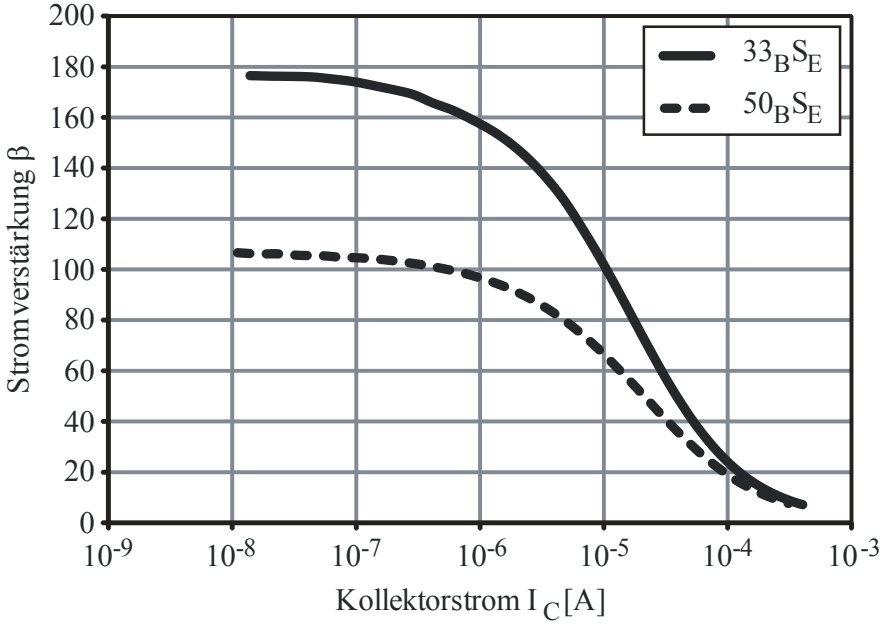


Bild 4: Stromverstärkung der 100×100 µm² 33_BS_E und 50_BS_E Fototransistoren

3.2 Optische AC Messungen

Optische AC Messungen wurden bei 850 nm durchgeführt. Dabei wurde ein Laser mit einer optischen Leistung von -15,9 dBm und einem Extinktionsverhältnis von 1,48 verwendet. Das Laserlicht wurde über eine 9 µm dicke Lichtwellenfaser auf den Fototransistor geleitet. Die Fototransistoren wurden bei drei unterschiedlichen Kollektor-Emitter Spannungen (V_{CE} = -2 V, V_{CE} = -5 V, V_{CE} = -10 V) charakterisiert. Weiters wurde bei jeder eingestellten Kollektor-Emitter Spannung der Arbeitspunkt mittels eines zusätzlichen Basisstroms (I_B = 0 A, I_B = -1 µA, I_B = -2 µA, I_B = -5 µA, I_B = -10 µA) verändert. In Bild 5 ist die Empfindlichkeit der 100×100 µm² Fototransistoren 33_BF_E (33% dotierte Basis und vollflächigen Emitter) und 33_BS_E (33% dotierte Basis und kleinen zentralen Emitter) dargestellt. Der große Unterschied in der Empfindlichkeit ergibt sich durch die unterschiedliche Größe des Emitters.

Bild 6 zeigt die Bandbreiten der beiden oben erwähnten Fototransistoren. Die kleine Kapazität der Basis-Emitter Diode C_{BE} beim 33_BS_E Fototransistor bewirkt eine höhere -3 dB Grenzfrequenz als beim 33_BF_E Fototransistor. Ein Vergleich der Bandbreite zwischen 10×10 µm² und 100×100 µm² großen 33_BS_E Fototransistoren ist in Bild 7 dargestellt. Hierbei hat die kleinere Struktur wegen der kleineren Basis-Kollektor Kapazität im Vergleich zur größeren Struktur eine

höhere Bandbreite. Bild 8 stellt das Empfangssignal des 10×10 µm^2 großen 50_BS_E Fototransistors mit einer minimalen Anstiegszeit von 40 ns dar. Tabelle 1 zeigt die Anstiegszeiten für diesen und den oben genannten Fototransistoren bei unterschiedlichen Kollektor-Emitter Spannungen V_{CE} und bei einem Basisstrom I_B = 10 µA. Minimale Anstiegszeiten von bis zu 23 ns wurden gemessen.

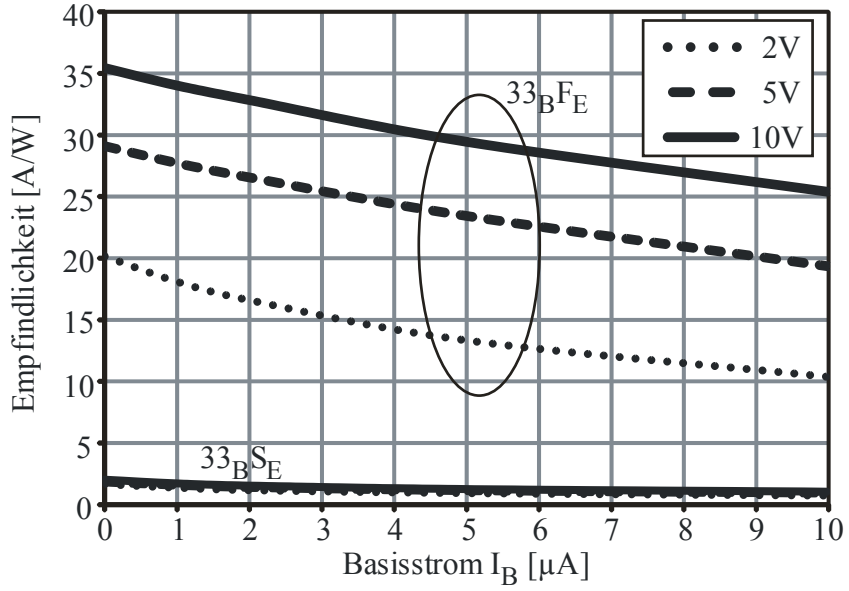


Bild 5: Empfindlichkeiten in A/W der 100×100 µm^2 großen Fototransistoren 33_BF_E und 33_BS_E bei unterschiedlichen Arbeitspunkten

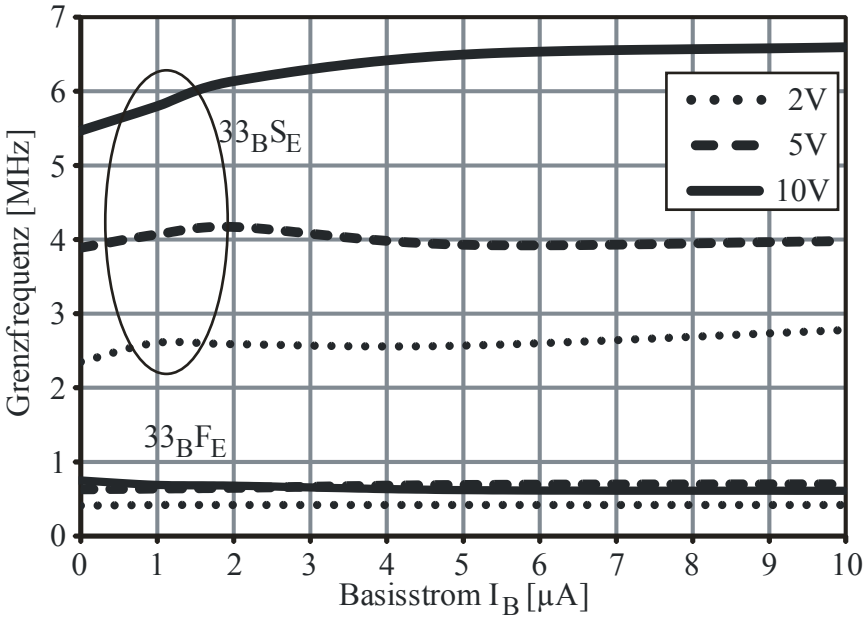


Bild 6: -3 dB Grenzfreuqenzen in MHz der 100×100 µm^2 großen Fototransistoren 33_BF_E und 33_BS_E bei unterschiedlichen Arbeitspunkten

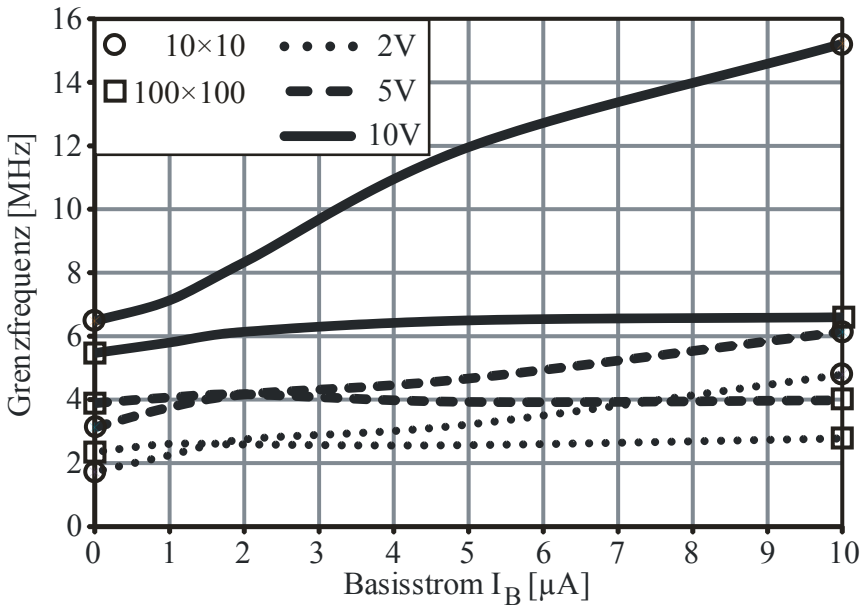


Bild 7: -3 dB Grenzfreuqenzen in MHz des 10×10 µm² und 100×100 µm² großen 33_BS_E Fototransistoren bei unterschiedlichen Arbeitspunkten

Tabelle 1: Anstiegszeiten in ns von vier Fototransistoren bei unterschiedlichen Kollektor-Emitter Spannungen V_{CE} und bei einem Basisstrom I_B = 10 µA

		V_{CE} = -2V	V_{CE} = -5V	V_{CE} = -10V
33_BS_E	10×10	73	57	23
50_BS_E	10×10	93	67	40
33_BS_E	100×100	126	88	53
33_BF_E	100×100	835	501	571

4 Schlusswort

In diesem Artikel wurden Silizium PIN-PNP-Fototransistoren präsentiert, welche in einem 0,6 µm CMOS OPTO ASIC Prozess hergestellt wurden. Aufgrund einer ca. 15 µm dicken, niedrig dotierten Epi-Schicht zwischen Basis und Kollektor sowie den unterschiedlichen Designs von Basis und Emitter, erzielen diese Fototransistoren -3 dB Grenzfrequenzen von bis zu 15,4 MHz.

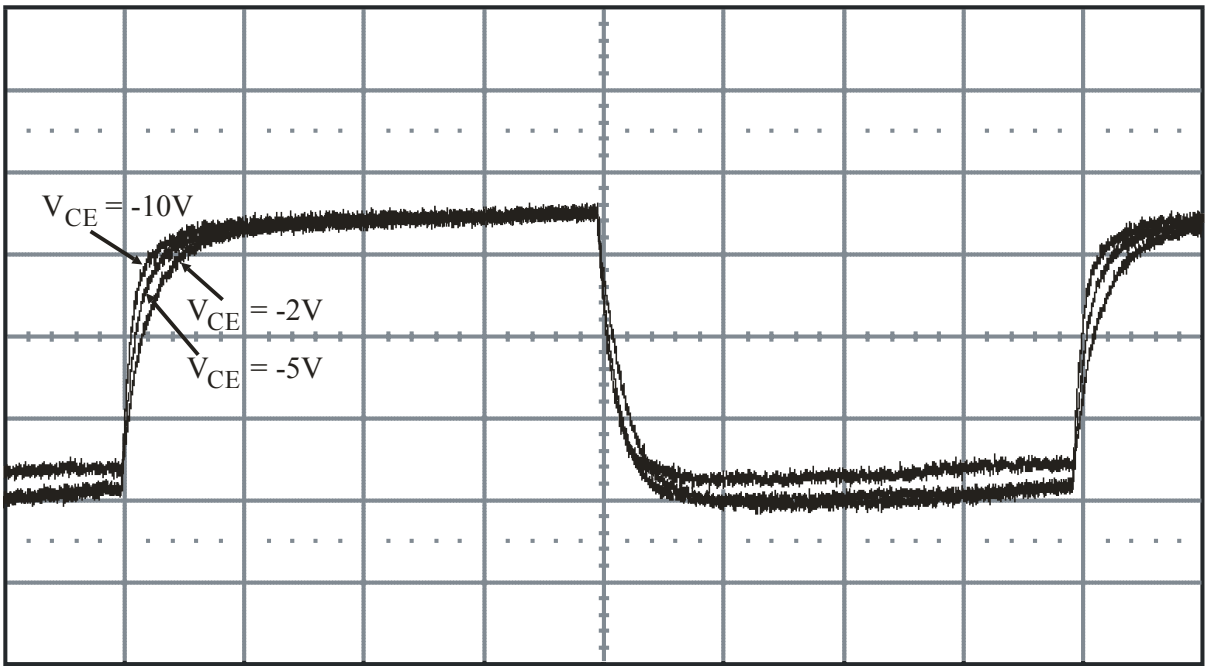


Bild 8: Sprungantwort des 10×10 µm² $50_B S_E$ Fototransistors bei drei Kollektor-Emitter-Spannungen. Arbeitspunkt: Popt = -15,9 dBm @ 850 nm, I_B = 10 µA. (Zeitbasis: 200 ns/div)

Empfindlichkeiten bis zu 35,4 A/W und minimale Anstiegszeiten von bis zu 23 ns wurden gemessen. Durch den Einsatz von billigen CMOS Fertigungs-Technologien besteht die Möglichkeit der kostengünstigen Implementierung dieser PIN-Fototransistoren in optoelektrische-integrierte-Schaltungen. Mögliche Anwendungsgebiete sind hierbei hochempfindliche optische Sensoren, System-on-Chip Applikation für Abstandsmessungen, usw.

Literatur

[1] Palik, E.D.: Hand Book of Optical Constants of Solids, Orlando: Academic Press, Inc., 1985

[2] Zach, G.; Davidovic, M.; Zimmermann, H.: Extraneous-light resistant multipixel range sensor based on a low-power correlating pixel-circuit, Proceedings of ESSCIRC, pp. 236-239, 2009

[3] Swoboda, R.; Zimmermann, H.: 11 Gb/s monolithically integrated silicon optical receiver for 850 nm wavelength, IEEE ISSCC, vol. 49, pp. 240-241, 2006

[4] Zimmermann, H.: Integrated Silicon Optoelectronics, 2nd ed., Berlin, Heidelberg: Springer-Verlag, 2010

[5] Cova, S.; Ghioni, M.; Lacaita, A.; Samori, C.; Zappa, F.: Avalanche photodiodes and quenching circuits for single-photon detection, Applied Optics, vol. 35, no. 12, pp. 1956-1976, 1996

[6] Pauchard, A.; Rochas, A.; Randjelovic, Z.; Besse, P.A.; Popovic, R.S.: Ultraviolet avalanche photodiode in CMOS technology, IEEE IEDM, pp. 709-712, 2000

[7] Yin, T.; Pappu, A.M.; Apsel, A.B.: Low-cost, high-efficiency, and high-speed SiGe phototransistors in commercial BiCMOS, IEEE Photonics Technology Letters, vol. 18, no. 1, pp. 55-57, 2006

[8] Gray, P.; Hurst, P.; Lewis, S.; Meyer, R.: Analysis and Design of Analog Integrated Circuits, p. 8ff., New York: Wiley, 2008

[9] Winstel, G.; Weyrich, C.: Optoelektronik II, p. 97ff., Berlin, Heidelberg: Springer-Verlag, 1986

Zerstörungsfreie Bestimmung der chemischen Bindungszustände vergrabener Nanoschichten durch Röntgenspektrometrie mit Synchrotronstrahlung

Beatrix Pollakowski

Physikalisch-Technische Bundesantalt, PTB, Abbestr. 2-12, 10587 Berlin, E-Mail: beatrix.pollakowski@ptb.de

Kurzfassung

Unter der Zielsetzung einer Rückführbarkeit ist angestrebt worden, eine zerstörungsfreie und nicht-präparative analytische Methode entwickelt, die es erlaubt funktionelle und nanostrukturierte Materialien zu charakterisieren. Neben der Elementverteilung und der geometrischen Anordnung beeinflusst die chemische Bindung Materialeigenschaften oft signifikant. Eine besondere Herausforderung liegt darin, die Art der chemischen Bindung auch in Abhängigkeit der Tiefe von ausgedehnten Objekten zu bestimmen. Als Ansatz dazu wurde hier eine “Photon-in Photon-out” Methodik gewählt, deren Entwicklung vorgestellt wird. Dazu wurde Röntgenabsorptionsspektroskopie, welche Zugang zum chemischen Bindungszustand ermöglicht, mit Röntgenspektrometrie unter streifendem Einfall kombiniert. Der Winkelbereich des streifenden Einfalls ist durch Interferenzeffekte von einfallender und reflektierter Strahlung charakterisiert, so dass die Eindringtiefe durch Variation von Einfallswinkel und Energie dediziert eingestellt werden kann. Anhand von Testsystemen, bestehend aus nanoskaligen Schichten leichter Elemente und Übergangsmetallen, wurde die vorgeschlagene Methodik validiert. Die Ergebnisse der Messungen an diesen Schichtsystemen zeigen, dass es möglich ist, den chemischen Bindungzustand tiefensensitiv zu bestimmen.

1 Einleitung

Funktionelle und nanostrukturierte Materialien gewinnen immer mehr an Bedeutung. Dies erfordert, dass deren Herstellungsprozesse weiterentwickelt, optimiert und auch kontrolliert werden. Neben der stetig wachsenden Bedeutung solcher Materialien nimmt auch ihre Komplexität zu, so dass sich ein Wandel von Oberflächen zu ausgedehnten Nanostrukturen beobachten lässt. In diesem Sinne kann man von einer Weiterentwicklung der Nanotechnologie sprechen. Begleitend dazu wächst auch der Bedarf an einer zerstörungsfreien und nichtpräparativen Analytik, die vor dem Hintergrund fehlender Referenzmaterialien darüber hinaus die Rückführbarkeit gewährleisten kann. Mit oberflächensensitiven Methoden, wie der Mikroskopie oder elektronenspektroskopischen Ansätzen lassen sich Oberflächenstrukturen hinreichend gut charakterisieren. Diffraktion dagegen gibt nur Auskunft über das Gesamtsystem. Derzeit sind etwa 700 Oberflächenstrukturen in verschiedenen Datenbanken erfasst. Jedoch sind die neueren Nanosysteme nicht nur durch die Grenzfläche zwischen Vakuum und Material gekennzeichnet, sondern auch durch Grenzflächen verschiedener Materialien. Dabei hat sich gezeigt, dass die bisherigen auf Oberflächen eingeschränkten Methoden nur bedingt geeignet sind, um auch tiefer liegende Grenzflächen zu analysieren.

Neben dem Aspekt der geometrischen Anordnung und der Elementverteilung kann die chemische Bindung der vorliegenden Elemente in hohem Maße die Materialeigenschaften beeinflussen. Die Charakterisierung der chemischen Bindung von in der Tiefe ausgedehnten Strukturen stellt dabei eine große Herausforderung dar. Im Folgenden wird die Entwicklung diese Ansatzes vorgestellt, der auf einer “Photon-in-Photon-out” Methodik beruht. Sie ist durch eine hohe Dynamik der Informationstiefe gekennzeichnet, gewährleistet dabei Zerstörungsfreiheit, erfordert jedoch keine Präparation der Proben. Hierzu wurde Absorptionsspektroskopie, die eine Bestimmung des Bindungszustands ermöglicht, mit der referenzprobenfreien Röntgenfluoreszenzanalyse [1] unter streifendem Einfall GIXRF-NEXAFS (Grazing Incidence X-Ray Fluorescence - Near Edge X-ray Absorption Fine Structure) kombiniert [2, 3]. Insbesondere der GIXRF-Bereich ist durch das Auftreten von Interferenzeffekte von einfal-

lender und reflektierter Strahlung charakterisiert. Dieses Phänomen wird auch als stehendes Wellenfeld bezeichnet. Die Eindringtiefe der Strahlung kann dadurch dediziert mithilfe des Einfallswinkels und der Photonenenergie eingestellt werden. Der Dynamikbereich der mittleren Eindringtiefe reicht dabei von einigen wenigen bis zu einigen hundert Nanometern. Diese winkelabhängige Absorptionsspektroskopie bietet dabei die Möglichkeit tiefensensitiv den Bindungszustand vergrabener Nanoschichten zu bestimmen.

Die Methodik GIXRF-NEXAFS wurde insbesondere für den weichen Röntgenbereich (etwa 150 eV bis 1900 eV) entwickelt, um Zugang zu den leichten Elementen, wie Kohlenstoff, Stickstoff, Sauerstoff und Übergangsmetallen zu erhalten. Anhand eines Systems, bestehend aus einer sehr dünnen (10 nm) Titanoxidschicht, die zwischen zwei Kohlenstoffschichten (jeweils 5 nm) angeordnet ist, und einem Siliziumsubstrat, ist diese Methodik erfolgreich validiert worden.

2 Experimentelles Vorgehen

2.1 Referenzprobenfreie Röntgenfluoreszenzanalyse

Die Bestimmung von Massendepositionen bzw. die Schichtdicke von komplexen neuartigen Materialien stellt sich zum Teil als Herausforderung dar. Eine Vielzahl an analytischen Methoden benötigen dabei gut charakterisierte Referenzproben, die den zu analysierenden Systemen sehr ähnlich sein müssen. Die Herstellung solcher Referenzmaterialien ist sehr aufwendig und auch langwierig. Im Gegensatz dazu werden zunehmend neue komplexe Systeme entwickelt, für die keine adäquaten Referenzproben existieren. Die referenzprobenfreie Röntgenfluoreszenzanalyse (XRF) bietet hier die Möglichkeit, die Elementgehalte basierend auf dem Fundamentalparameteransatz [4, 5] zu bestimmen, ohne das Referenzproben zum Einsatz kommen. Der referenzprobenfreie Ansatz wurde hier dazu verwendet, die Schichtdicken von Titanverbindungen zu validieren.

Die referenzprobenfreie Röntgenfluoreszenzanalyse setzt voraus, dass experimentelle Parameter, wie der Raumwinkel der Detektion, die einfallende Strah-

lungsleistung, der Einfalls- und der Detektionswinkel, sowie die Zählraten und die Detektoreffizienz wohl bekannt sind. Des Weiteren müssen die relevanten atomaren Fundamentalparameter bekannt sein. In der vorliegenden Arbeit wurden dazu die Datenbanken von Elam et al. [6] und Ebel et al. [7] herangezogen. Die Massendeposition eines Elements wurde durch den Ansatz von Sherman [8] berechnet. Der Einfalls- und der Detektionswinkel θ betragen in der konventionellen XRF $\theta = 45\,^\circ$ und der Zusammenhang zwischen der Zählrate N_l und der Massendeposition m/A nimmt dabei die folgende Form an:

$$\frac{m}{A} = \rho d = -\ln\left[1 - \frac{N_l[\mu_M(E_P) + \mu_M(E_l)]}{I_0 \epsilon_{Det} \frac{\mathrm{d}\Omega}{4\pi} \tau_{Xi}(E_P) g_{Xi} \omega_{Xi}}\right] \left[\frac{\mu_M(E_P)}{\sin 45\,^\circ} + \frac{\mu_M(E_l)}{\sin 45\,^\circ}\right]. \quad (1)$$

In der Tabelle 1 ist eine Übersicht über sämtliche Größen gegeben.

Tabelle 1: Übersicht über alle experimentellen Größen und Fundamentalparameter.

Symbol	Bezeichnung	Einheit
I_0	einfallender Photonenfluss	s^{-1}
ϵ_{Det}	Detektoreffizienz für die entsprechende Energie	-
$\mathrm{d}\Omega$	Raumwinkel der Detektion	sr
$\theta_{in,out}$	Einfalls- und Detektionswinkel	Grad (°)
$\mu_M(E)$	Massenabsorptionskoeffizient des Probenmaterials für eine Energie E	$cm^2 g^{-1}$
ω_{Xi}	Fluoreszenzausbeute der Schale Xi	-
g_{Xi}	Übergangswahrscheinlichkeit für die Fluoreszenzlinie l der von der i-ten in die j-ten Schale Xi	-
$\tau_{Xi}(E)$	photoelektrischer Wechselwirkungsquerschnitt der Schale Xi für eine Energie E	$cm^2 g^{-1}$
S_{Xi}	Sprungverhältnis an der Absorptionskante	-
ρ	Dichte der Probe	gcm^{-3}

Die Fluoreszenzstrahlung wurde hier mittels eines fensterlosen kalibrierten Siliziumdriftdetektor (SDD) detektiert, dessen Antwortverhalten und dessen Effizienz sehr gut bekannt ist. Es existiert insbesondere auch ein physikalisches Modell, das die Detektorantwort beschreiben kann. Die einfallende Strahlungleistung bzw. der Fluss wurde mittels kalibrierter Instrumentierung bestimmt. Zur Beschreibung des effektiven Raumwinkel existiert ein Modell, welches die bestrahlte Fläche auf der Probe mitberücksichtigt [9].

2.2 Absorptionsspektroskopie unter streifendem Einfall

Bei der Methode GIXRF-NEXAFS handelt es sich um eine kombinierte Methode, bei der die Dynamik in der Eindringtiefe im GIXRF-Regime ausgenutzt wird, um tiefenaufgelöst mit hochauflösender Absorptionspektroskopie den Bindungszustand des Materials zu ermitteln. Der streifende Einfall ist charakterisiert durch Interferenzeffekte von einfallender und reflektierter Strahlung, die auf der Oberfläche und in den darunter liegenden Schichten auftreten. Diese Interferenzeffekte beruhen auf dem sogenannten stehenden Wellenfeld (X-Ray Standing Waves - XSW) (siehe Bild 1 links). Die Intensitätsverteilung des XSW-Feldes wird als Tiefensensor verwendet, der in Abhängigkeit von Einfallswinkel und Photonenenergie die Eindringtiefe in das Material definiert. Die Tiefenverteilung des elektromagnetischen Feldes wird dabei durch die optischen Eigenschaften n und k (Real- und Imaginärteil des komplexen Brechungsindex) maßgeblich bestimmt. Insbesondere wird die XSW-Intensität durch das Verhalten an den Grenzflächen beeinflusst. Der Imaginärteil k des Brechungsindex zeigt in der Nähe von Absorptionskanten eine resonantes Verhalten, das heißt, dass sich damit die Lage von Extrema des XSW-Feldes in Abhängigkeit der Photonenenergie deutlich verändert, und die Eindringtiefe stark variiert. Ein deutliche Variation in der Eindringtiefe der Strahlung unterhalb und oberhalb einer Absorptionskante eines Matrixelements muss daher bei den Messungen berücksichtigt sein. Bei einer NEXAFS-Messung muss gewährleistet werden, dass die Eindringtiefe konstant bleibt. Da bei der Untersuchung in der Nähe einer Absorptionskante die Änderung der einfallenden Photonenenergie vorgegeben ist, kann darum nur entsprechend der Einfallswinkel angepasst werden. Im Spektralbereich weicher Röntgenstrahlung sind die Minima und Maxima der XSW-Intensität im Allgemeinen nicht sehr stark lokalisiert. Die Tiefenauflösung liegt bei einigen wenigen Nanometern. Zur Berechnung des Einfallswinkels wird die integrale XSW-Intensität des relevanten Tiefenbereichs betrachtet. Der Einfallswinkel, bei dem die Intensität maximal ist, wird als Einfallswinkel für die Energie eingestellt. Für die weiteren Photonenenergien wurde der Winkel ermittelt, bei dem die integrale XSW-Intensität bzw. die mittlere Eindringtiefe den selben Wert annimmt. Daraus lässt sich für jede Ener-

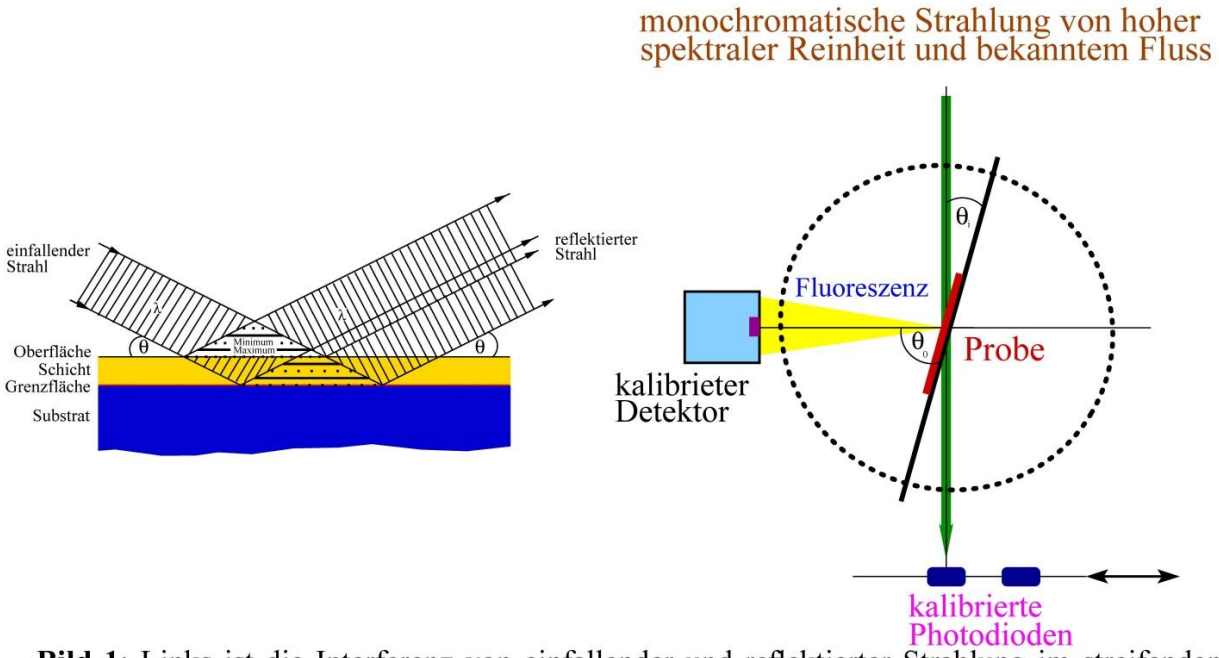


Bild 1: Links ist die Interferenz von einfallender und reflektierter Strahlung im streifendem Einfall dargestellt. Das Bild rechts zeigt skizzenhaft den experimentellen Aufbau.

gie der Einfallswinkel bestimmen, der für die darauf folgende Messung einstellt werden muss.

Das Experiment ist im Bild 1 (rechts) dargestellt. Weitere Anforderungen sind eine sehr kleine Oberflächen- und Grenzflächenrauigkeit, sowie eine longitudinale Kohärenz der Strahlung.

2.3 Proben: Tief vergrabene Titanoxide

Zur Entwicklung der Methodik GIXRF-NEXAFS wurde ein Probensystem benötigt, das zum einen eine komplexe Schichtstruktur aufweist und zum anderen erfordert, dass dessen Rauigkeiten an den jeweiligen Grenzflächen unter einem halben Nanometer rms liegen. Zur Herstellung von sehr glatten, nichtporösen nanoskaligen Schichten eignet sich zum Beispiel eine Ionenstrahl-Abscheidetechnik (Ion Beam Sputtering Deposition - IBSD). IBSD erlaubt ein reaktives Ionenätzen, das sehr gut für das Abscheiden von Oxiden geeignet ist. Die Probenpräparationen wurden vom Fraunhofer Institut für Werkstoff- und Strahltechnik (IWS) in Dresden vorgenommen, das sehr viel Erfahrung darin hat, sehr dünne Multischichtsysteme mit geringen Rauigkeiten bezüglich der Oberfläche und der jeweiligen Grenzschichten abzuscheiden. Die dort befind-

liche Ionenstrahlabscheideanlage “IonSys1600” wird schon seit Jahren erfolgreich verwendet [10].

Für die Testsysteme wurde ein Siliziumwafer mit 50 nm Molybdän und danach mit 5 nm Kohlenstoff beschichtet. Im Anschluß wurden dann 10 nm Titan abgeschieden und direkt danach wiederum mit 5 nm Kohlenstoff abgedeckt. Für die weiteren Proben wurden beim Abscheiden des Titans Sauerstoffionen mit eingelassen. In Abhängigkeit von der Sauerstoffflussrate Q konnten unterschiedliche Titanoxide herstellt werden. Die Flussrate wird in $sccm$ angegeben, welche 0.0018124 Pa $\cdot$ m^3/s entspricht. Bei den Dickenangaben handelt es sich um nominell vom IWS angegebene Werte.

3 Ergebnisse

3.1 Bestimmung der optischen Konstanten

Die optischen Konstanten, welche der CXRO-Datenbank [11] entnommen wurden, beschreiben die Intensitätsverteilung des stehenden Wellenfeldes in der Probe, insbesondere in der Nähe von Absorptionskanten, nur bedingt richtig. Als Konsequenz wurden die optischen Konstanten der jeweiligen Titanoxidverbindungen in den relevanten Energiebereichen mit Röntgenreflektometrie (XRR) an den Einzelschichten dediziert neu bestimmt. Im Energiebereich von 450 eV bis 475 eV (d.h. um die Ti-L_3- und Ti-L_2-Kanten herum) wurden Messungen durchgeführt, wobei die jeweilige Reflektivität in Abhängigkeit des Einfallswinkels ermittelt wurde. Im Bild 2 ist die Reflektivität als Funktion des steifendem Einfallswinkel für ein vergrabenes Titanoxidschichtsystem dargestellt. Als Photonenenergien wurden relevante Energien unterhalb und oberhalb der Titan $L_{3,2}$ Absorptionskanten gewählt. Die Auswertung der XRR-Messungen erfolgte mithilfe der Software IMD [12]. Daraus konnten neben den optischen Eigenschaften n und k, die Schichtdicke und die Rauigkeiten bestimmt werden. Die Schichtdicke der Titanoxidschicht aus Bild 2 kann mit (11.8 $\pm$ 0.6) nm angegeben werden.

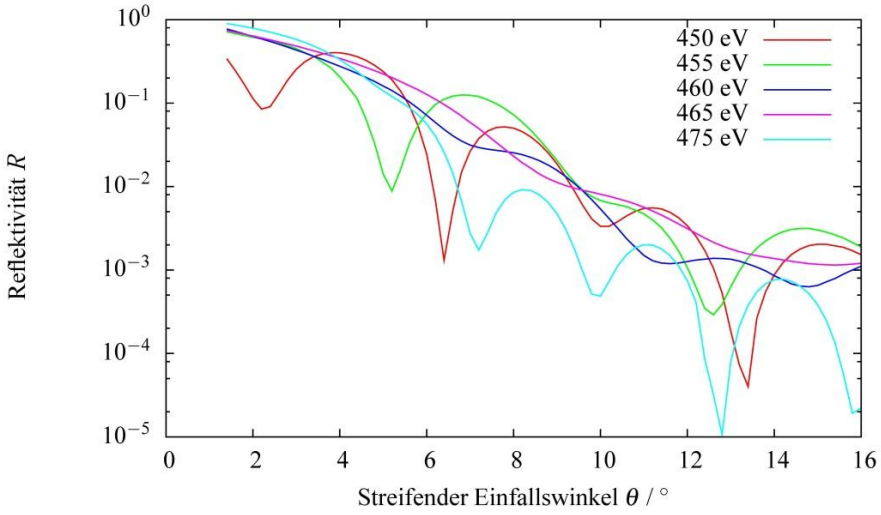


Bild 2: Reflektivität eines vergrabenen Einzelschichtsystems beschichtet mit einer geringen Sauerstoffflussrate, welches TiO entspricht, in Abhängigkeit vom Einfallswinkel. Unterschiedliche Energien zeigen veränderte Reflektivitäten.

3.2 Speziation dünner vergrabener Titanoxidfilme

Für eine verlässliche Speziation mittels Absorptionsspektroskopie von tief vergrabenen Schichten muss gewährleistet sein, dass die Eindringtiefe hinreichend ist, um die Fluoreszenzstrahlung des vergrabenen Materials beobachten zu können und um während des Messprozesses nach Möglichkeit gleiche Massenbelegungen einer Spezies zum Messsignal beitragen zu lassen. Das bedeutet, dass der Einfallswinkel entsprechend eingestellt werden muss. Im Folgenden werden eine Auswahl von Ti-$L_{3,2}$ GIXRF-NEXAFS Spektren vorgestellt, gemessen an vergrabenen Titanoxidschichten bei variablen Einfallswinkel θ.

In Bild 3 sind die Ti-$L_{3,2}$ GIXRF-NEXAFS Spektren dargestellt, bei der Winkelkorrekturen verwendet wurden. Deutlich ist dabei zu erkennen, dass die beiden Absorptionskanten mit steigender Sauerstoffflussrate sich zu höheren Energien verschieben. Die Feinstruktur an der Ti-L_3 Kante als auch an der Ti-L_2-Kante verändert sich dabei auch signifikant, von der spezifischen Feinstruktur des metallischen Titan über die von TiO (bis 2 sccm) und Ti_2O_3 (3 sccm bis 4.5 sccm) bis hin zu TiO_2 (ab 9 sccm). Für die dazwischen liegenden Flussraten kann man Mischsysteme beobachten. In diesen NEXAFS-Spektren ist auch sehr gut zu erkennen, dass sich die kristalline Struktur ändert, also die Anordnung des Sauerstoffs bezüglich des Titans und entsprechend dessen Bindungs-

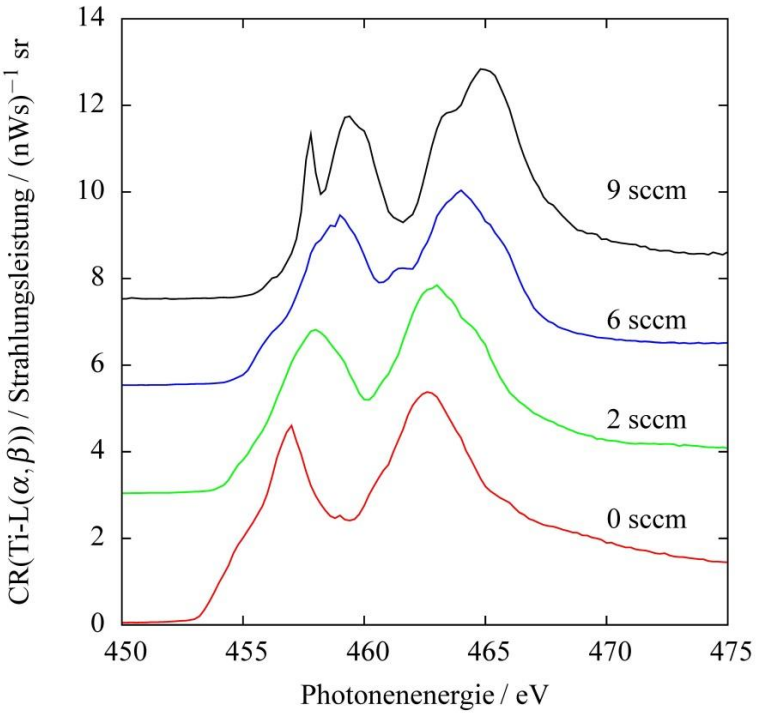


Bild 3: Ti-$L_{3,2}$ GIXRF-NEXAFS Spektren von vergrabenen Titanoxidschichten, die bei variablen Einfallswinkel θ gemessen wurden. Unterschiedliche Sauerstoffflussraten beim Beschichtungsprozess ergeben unterschiedliche Titanoxidverbindungen.

zustand. Von Ti_2O_3 zu TiO_2 gewinnt das so genannte “Crystal Field” an Bedeutung; durch die oktaedrische Anordnung des Sauerstoffs um das Titan sorgt es für eine weitere Aufspaltung in der Feinstruktur der jeweiligen Absorptionskante in drei Komponenten.

3.3 Quantifizierung der Massenbelegung

Des Weiteren wurden die vergrabenen Nanoschichtsysteme bezüglich der Elementzusammensetzung und der Schichtdicke analysiert. Dazu wurde die referenzprobenfreie Röntgenfluoreszenzanalyse herangezogen. Zur Berechnung der Massendeposition wurden ausschließlich primäre Prozesse herangezogen, da im weichen Röntgenbereich relevante sekundäre Anregungseffekte wegen sehr kleiner Fluoreszenzausbeuten und Streuquerschnitte ausgeschlossen werden können. Für die Photonenenergie der einfallenden Strahlung wurde 1280 eV gewählt, um die Hauptbestandteile des Schichtsystems, wie Kohlenstoff, Sauerstoff und Titan anzuregen. Ferner wurde diese Photonenenergie verwendet, um neben den Hauptbestandteilen mögliche Kontaminationen zu identifizieren.

Mithilfe des durch Gleichung 1 dargestellten Modells lässt sich ein Zusammenhang zwischen Massendeposition und emittierter Fluoreszenzstrahlung herstellen. Mit der Kenntnis aller experimentellen Parameter (einfallende Strahlungleistung, Raumwinkel und Nachweiseffizienz der Detektion) und aller atomaren Fundamentalparametern (Wechselwirkungsquerschnitte, Fluoreszenzausbeuten und Übergangswahrscheinlichkeiten) kann aus der Fluoreszenzstrahlung auf eine Massendeposition geschlossen werden. In der Tabelle 2 sind die entsprechenden Schichtdicken und die Massendepositionen von Titan und Sauerstoff dargestellt.

Tabelle 2: Überblick: Aus der Primärfluoreszenzintensität berechneten Massenbelegungen m/F und Schichtdicken d. Die nominelle Schichtdicke wurde vom IWS mit 10 nm angegeben. Die abgeschätzte relativen Unsicherheit (k=1) der hier bestimmten Schichtdicken beträgt 15 %.

Flussrate Q	d / nm	m/F Ti / μg/cm^2	m/F O / μg/cm^2
0 sccm	9.0	3.84	0.32
2 sccm	11.0	3.94	1.08
6 sccm	13.8	3.84	2.31
9 sccm	11.8	2.99	1.98

4 Zusammenfassung

Die hier vorgestellte Methode zur zerstörungsfreien und nicht-präparativen Bestimmung des chemischen Bindungszustands weist ein hohes Potential auf, auch einen Zugang zum Bindungszustand von tief vergrabenen Schichten und zu vergrabenen Grenzschichten zu erhalten. Die Methode ist allerdings noch keine Standardmethode, da sie bisher auf eine Synchrotronstrahlungsquelle sowie auf die Bestimmung oder a-priori Kenntnis zuverlässiger optischer Konstanten angewiesen ist. Zur Zeit sind die Energiebereiche von Laborröntgenquellen durch die charakteristischen Linien bestimmt und die auftretende Kontinuumstrahlung weist einen sehr niedrigen Fluss auf, der zu gering ist, um eine leistungsfähige, durchstimmbare Quelle zu darzustellen.

Durch GIXRF-NEXAFS-Messungen an verschiedenen Oxiden des Titans wurde im PTB-Labor am Elektronenspeicherring BESSY II ein “Proof-of-

principle" einer zerstörungsfreien und nicht-präparativen Analysemethode für die Bestimmung der chemischen Bindungszustände vergrabener Nanoschichten erbracht. Die Tiefenauflösung im Bereich weicher Röntgenstrahlung war nur moderat (wenige Nanometer) im Gegensatz zum harten Röntgenbereich, der aber keinen optimalen Zugang zu den leichteren Elementen ermöglicht. Es müssen daher in Zukunft noch erweiterte Messstrategien entwickelt werden. Die Verbesserung der lateralen Auflösung stellt eine instrumentelle Herausforderung dar, da die Fokusierung der Anregungsstrahlung auf einen kleinen Bereich mit einer deutlichen Steigerung der Strahldivergenz einhergeht. Optimal wäre hier die Nutzung kleinerer Quellgrößen und perfekterer Röntgenoptiken.

5 Danksagung

Die Autorin möchte der Deutschen Forschungsgemeinschaft für finanzielle Unterstützung des Forschungsprojekts "Nanoschichtspeziation", der AG "Röntgenspektrometrie" (PTB) und den Herren Stefan Braun und Peter Gawlitza vom Fraunhofer IWS Dresden für die Herstellung des Probensystems danken.

Literatur

[1] B. Beckhoff, P. Hoenicke, D. Giubertoni, G. Pepponi und M. Bersani, "GIXRF in the soft x-ray range used for the characterization of ultra shallow junctions", *AIP Conference Proceedings* **1173**(1), S. 29–33, 2009.

[2] B. Pollakowski, B. Beckhoff, F. Reinhardt, S. Braun und P. Gawlitza, "Speciation of deeply buried TiO_x nanolayers with grazing-incidence x-ray fluorescence combined with a near-edge x-ray absorption fine-structure investigation", *Phys. Rev. B* **77**(23), S. 235408, 2008.

[3] M. Pagels, F. Reinhardt, B. Pollakowski, M. Roczen, C. Becker, K. Lips, B. Rech, B. Kanngießer und B. Beckhoff, "GIXRF – NEXAFS investigations on buried ZnO/Si interfaces: A first insight in changes of chemical states due to annealing of the specimen", *Nuclear Instruments and Methods in Physics Research B* **268**(3–4), S. 370–373, 2010.

[4] M. Kolbe, B. Beckhoff, M. Krumrey und G. Ulm, "Thickness determination for Cu and Ni nanolayers: Comparison of completely reference-free fundamental parameter-based x-ray fluorescence analysis and x-ray reflectometry", *Spectrochim. Acta B* **60**(4), S. 505–510, 2005.

[5] B. Beckhoff, R. Fliegauf, M. Kolbe, M. Müller, B. Pollakowski, J. Weser und G. Ulm, "X-ray spectrometry for wafer contamination analysis and speciation as well as for reference-free nanolayer characterization", *Solid State Phenomena* **134**, S. 277–280, 2008.

[6] W. T. Elam, B. Ravel und J. R. Sieber, "A new atomic database for x-ray spectroscopic calculations", *Rad. Phys. Chem.* **63**, S. 121–128, 2002.

[7] H. Ebel, R. Svagera, M. F. Ebel, A. Shaltout und J. H. Hubbell, "Numerical description of photoelectric absorption coefficients for fundamental parameter programs", *X-Ray Spectrom.* **32**(6), S. 442–451, 2003.

[8] J. Sherman, "The theoretical derivation of fluorescent x-ray intensities from mixtures", *Spectrochim. Acta* **7**, S. 283–306, 1955.

[9] B. Beckhoff, R. Fliegauf, M. Kolbe, M. Müller, J. Weser und G. Ulm, "Reference-free total reflection x-ray fluorescence analysis of semiconductor surfaces with synchrotron radiation", *Anal. Chem.* **79**(20), S. 7873–7882, 2007.

[10] P. Gawlitza, S. Braun, A. Leson, S. Lipfert und M. Nestler, "Precise thin film synthesis by ion beam sputter deposition", *Vakuum in Forschung und Praxis* **19**(2), S. 37–43, 2007.

[11] B. Henke, E. Gullikson und J. Davis, "X-ray interactions: photoabsorption, scattering, transmission, and reflection at $\mathrm{E} = 50 - 30000\mathrm{eV}, \mathrm{Z} = 1 - 92$", *At. Data and Nucl. Data Tables* **54**, S. 181–342, 1993.

[12] D. L. Windt, "IMD - software for modeling the optical properties of multilayer films", *Computers in Physics* **12**, 1998.

Einsatz von Messsystemen auf Basis der Streifenprojektions- und Streifenreflexionstechnik zur Konturmessung und Oberflächencharakterisierung

Marc Sandner[1], Wansong Li[2], Christoph von Kopylow[1], Ralf B. Bergmann[1]

[1] BIAS – Bremer Institut für angewandte Strahltechnik, Bremen
Tel.: 0421-21858010, e-mail: msandner@bias.de
[2] VEW – Vereinigte Elektronik Werkstätten GmbH, Bremen

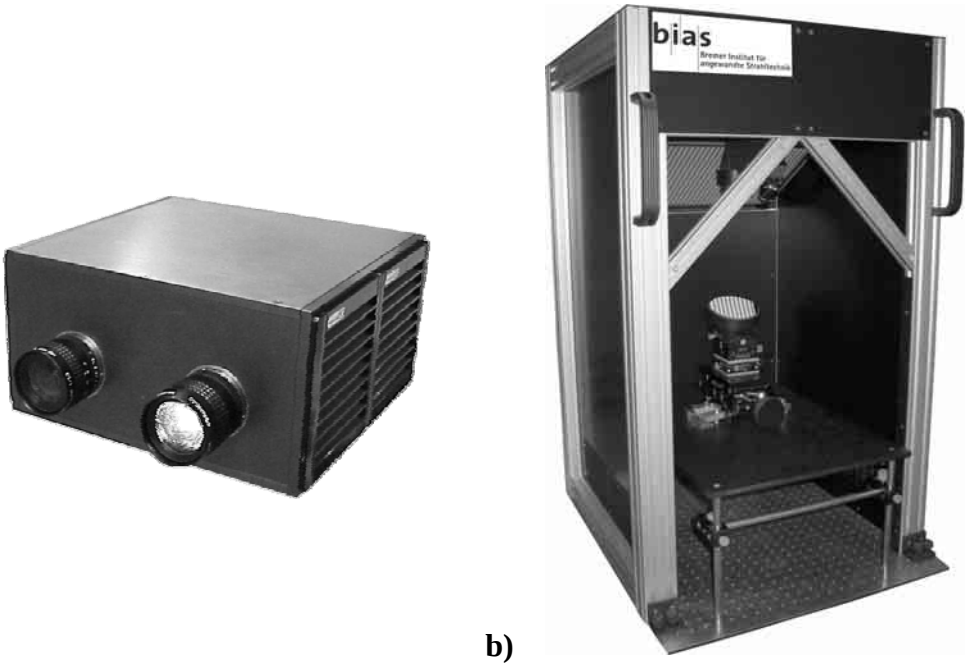
Kurzfassung

In diesem Beitrag wird der Stand der Technik von zwei berührungslosen optischen Messmethoden dargestellt. Für optisch raue Oberflächen kommt die Streifenprojektionstechnik zum Einsatz, die sich besonders zur Form- und Verformessung in industriellen Anwendungen eignet; spiegelnde Oberflächen können mit der Streifenreflektionstechnik vermessen werden, welche besonders empfindlich auf Oberflächenneigungen anspricht. Beide Techniken werden anhand von Beispielen dargestellt.

1 Einleitung

Bei der 3D-Vermessung von Konturen und Oberflächen, und in der Zerstörungsfreien Prüfung [1] werden zunehmend berührungslose optische Messtechniken eingesetzt, die gegenüber taktilen Verfahren teilweise deutlich schneller hochaufgelöste Messdaten erzeugen. Eine Untergruppe der optischen Messtechniken stellen hierbei die Streifenreflexions- (SRT) und Streifenprojektionstechnik (SPT) dar, die mit inkohärentem, räumlich amplitudenmoduliertem Licht als Signalquelle arbeiten. Durch die Skalierbarkeit der Messfeldgröße können die beiden Techniken in verschiedensten Einsatzgebieten verwendet werden, so beispielsweise in der Formvermessung an Kraftfahrzeugen und Mikrobauteilen. Während die SPT zur Formmessung an Objekten mit optisch rauer Oberfläche verwendet wird, findet die SRT ihr Haupteinsatzgebiet in der Charakterisierung

von spiegelnden Oberflächen. Als Oberflächengradienten-Messtechnik ist die SRT auch für die Detektion von extrem kleinen Oberflächendetails wie Defekten auf Lackoberflächen und Polierspuren auf Linsen einsetzbar. Im Folgenden werden die beiden Messtechniken vorgestellt und Messbeispiele aus verschiedenen Einsatzgebieten präsentiert.



a) **b)**

BIAS ID 110830

Bild 1: Fotos der am BIAS entwickelten Aufbauten:
a) kompakte 3D-Kamera (auf SPT basierend, Größe: 20x22x13 cm³)
b) SRT-Messaufbau (Größe: 100x60x60 cm³)

Tabelle 1: Eigenschaften der Streifenprojektionstechnik (SPT) und der Streifenreflexionstechnik (SRT)

Messtechnik	**Signalgeber**	**Oberflächentyp**	**Messgrößen**
Streifenprojektions-technik (SPT)	Projektor	Rau, diffus reflektie-rend	Form
Streifenreflexions-technik (SRT)	Monitor	Glatt, gerichtet re-flektierend	Lokale Gradienten, Krümmungen; Form

2 Beschreibung der Messtechniken

2.1 Streifenprojektionstechnik (SPT)

Bei einer Messung mittels der Streifenprojektionstechnik (SPT) werden gewöhnlich mit einem Streifenprojektor linienförmige sinusförmige oder binäre Streifen auf ein zu untersuchendes Objekt projiziert [2] (siehe Bild 2). Das sich unter einem veränderten Blickwinkel ergebende deformierte Bild der Streifen auf der Objektoberfläche wird mit einer Kamera aufgezeichnet. Der Kern der Messtechnik besteht in der Zuordnung eines jeden Kamerapixels (i,j), das einen Objektpunkt $(x,y,z)_{ij}$ beobachtet, zu den Koordinaten eines Projektorpixels (l,m), von dem es indirekt beleuchtet wird. Dieser Prozess wird Absolutphasenmessung genannt. Hierzu werden verschiedene Muster auf das Objekt projiziert und dann die entsprechenden Kamerabilder aufgezeichnet und anschließend verrechnet.

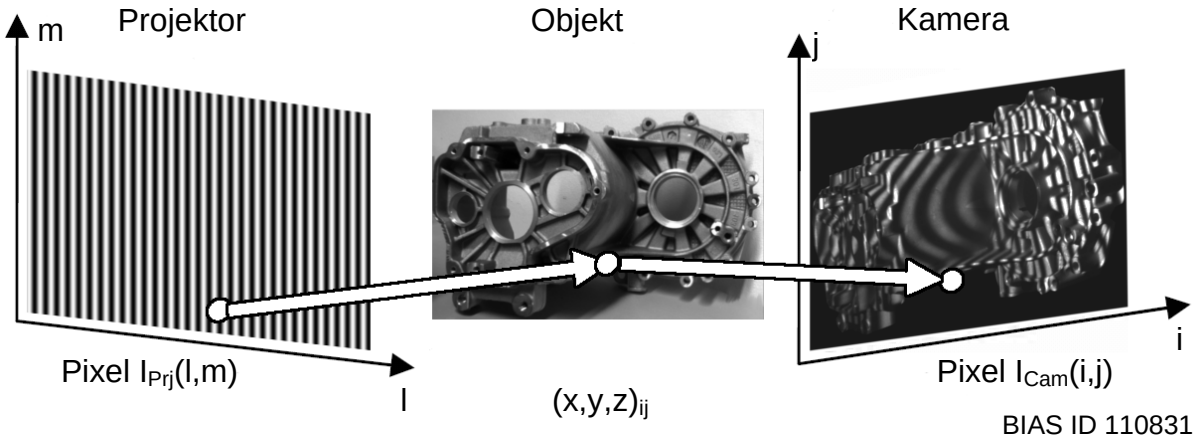


Bild 2: Messung mittels Streifenprojektionstechnik, Absolutphasenmessung: Bestimmung der Korrespondenz zwischen Positionen der Projektor- und Kamerapixel. Nach: T. Bothe, 2004 [6]

Der von uns verwendete Ansatz zur Absolutphasenmessung bedient sich der Methode des hierarchischen Verstetigens mit synthetischen Wellenlängen [3]. Die entwickelten Algorithmen sind sehr robust und an die geforderte Messsituation anpassbar, so dass zwischen Messauflösung und -geschwindigkeit gewichtet werden kann. Ist die Orientierung von Projektor zu Kamera bekannt, so kann eine geometrisch - strahlenoptische Verbindung von Projektorpixel (l,m) zu Objektpunkt $(x,y,z)_{ij}$ und weiter zum Kamerapixel (i,j) modelliert und aus dem sich ergebenden Dreieck die Koordinate $(x,y,z)_{ij}$ bestimmt werden. Die Güte der ge-

messenen 3D-Koordinaten hängt primär von einer robusten Kalibrierung der Parameter ab, die die Beziehungen der Objektkoordinaten zu den Bildkoordinaten von Projektor und Kamera beschreiben. Zur Kalibrierung von SPT-Systemen wird das am BIAS entwickelte Sichtstrahl-Kalibrierungsverfahren [4],[5] verwendet.

2.1.1 3D-Kamera: SPT mit paralleler Ausrichtung von Projektor und Kamera

Bei Standard-SPT-Systemen sind Kamera und Projektor derart zueinander geneigt, dass sich deren optische Achsen in der Messebene treffen (Bild 3a). Eine Messung ist hierbei nur im Überlappbereich der beiden Bildfelder möglich, und das Messvolumen wird zudem durch die begrenzte Schärfentiefe der Systeme eingeschränkt. Es ergibt sich ein relativ kleines Messvolumen, in welchem das zu vermessende Objekt platziert werden muss. Da die SPT auf dem Konzept der Triangulation beruht, gilt auch hier, dass die Empfindlichkeit mit dem verwendeten Triangulationswinkel zwischen den optischen Achsen von Projektor und Sensor zunimmt. Bei zunehmenden Winkel ergeben sich jedoch durch Abschattungen von Projektor und/oder Kamera (Bild 3b) Probleme bei stärkeren Oberflächengradienten, so z.B. bei Einbuchtungen und Schrägen.

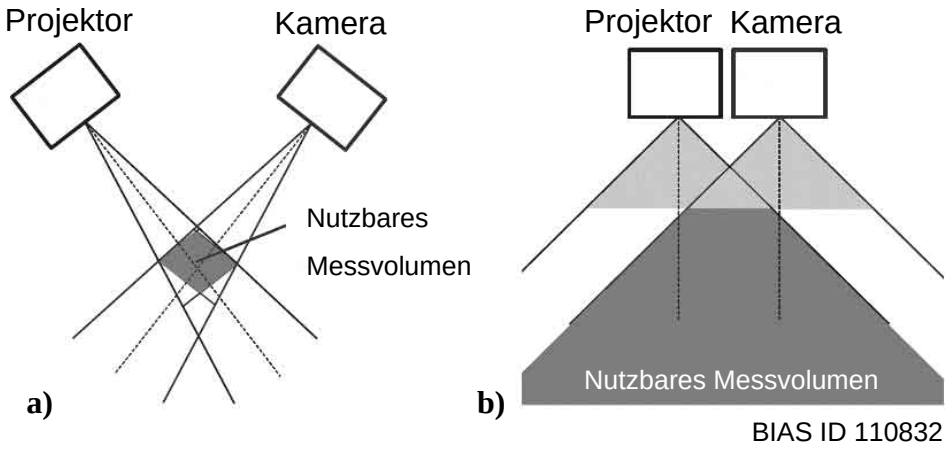


Bild 3: a) gewöhnliche Streifenprojektionssysteme, b) 3D-Kamera: parallele optische Achsen, 90° Öffnungswinkel, große Schärfentiefe → Messfeld vergleichbar mit Weitwinkelkamera. In den hier hellgrau markierten Bereichen ist aufgrund fehlender Bildschärfe keine Messung möglich. Nach: T. Bothe, 2004 [6]

Bei der am BIAS entwickelten 3D-Kamera [6] wurde ein anderes Konzept gewählt, Projektor und Kamera werden hier mit parallelen optischen Achsen nahe

beieinander montiert (siehe Bild 3c). Der Öffnungswinkel beider Komponenten wurde auf mehr als 90° ausgelegt und der Schärfentiefe-Bereich so eingestellt, dass Objekte ab einem Abstand zum Messsystem von ca. 30-40 cm vermessen werden können. Im Vergleich zu den Messfeldern üblicher Systeme ist dasjenige der 3D-Kamera erheblich weiter (Bild 3c), wodurch das System für eine Vielzahl an Messaufgaben verwendbar wird, siehe Bild 4.

Für eine Messung benötigt das System 1 bis 20 Sekunden, abhängig von den Benutzervorgaben für die zu erreichende Auflösung. Ist eine Auflösung von 100 µm verlangt und der Abstand des Objekts vom Messsystem auf 50 cm eingestellt, benötigt eine Messung beispielsweise 4 Sekunden. Das System kann wie eine Kamera auf einem Stativ benutzt werden.

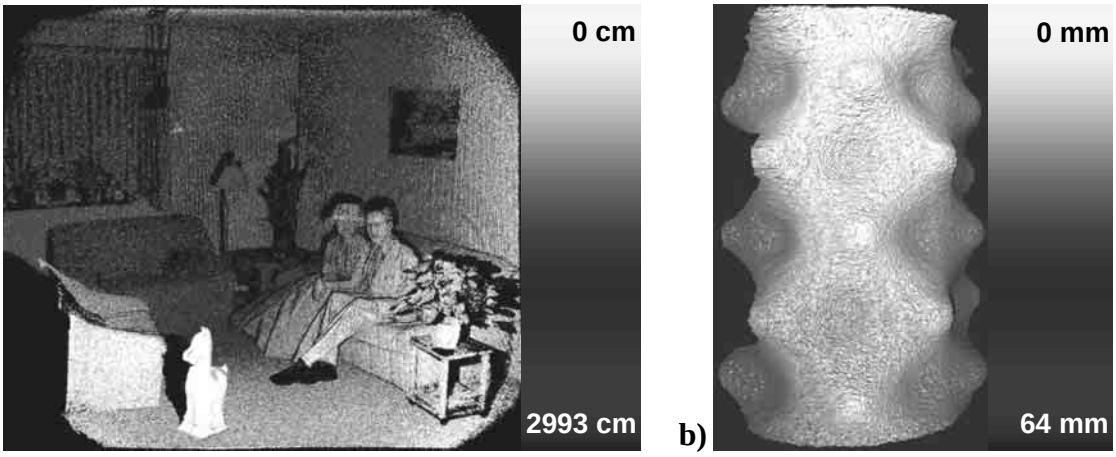


Bild 4: Mit der auf SPT basierenden kompakten 3D-Kamera erzeugte Punktwolken, Helligkeitscodierung für Tiefe in z-Richtung: a) Szene mit großem Messvolumen: Oberflächendaten eines vermessenen Wohnzimmers; b) Szene mit geringem Messvolumen: Oberflächendaten aus Messung eines Kunststoffschaums, Wiederholungenauigkeit hier 50 µm. Aus: T. Bothe, 2004 [6], mit freundlicher Genehmigung des Wichmann Verlags

2.2 Streifenreflexionstechnik (SRT)

Während die SPT zur Vermessung von optisch rauen Objektoberflächen verwendet wird, können mit der Streifenreflexionstechnik (SRT) optisch glatte bis spiegelnde Oberflächen vermessen werden.

Das Streifenreflexionsverfahren beruht auf der Auswertung eines bekannten Referenzmusters, welches sich in der Objektoberfläche spiegelt und durch dabei

entstehende Verzerrungen Rückschlüsse über das Objekt erlaubt. Im täglichen Leben kann leicht festgestellt werden, dass die Verzerrung des Spiegelbildes eine sehr empfindliche Messgröße ist: wenn es z.B. darum geht, die Lackoberfläche eines Autos zu untersuchen, schaut man sich in der Regel intuitiv die Zerrspiegeleffekte an und kann Unebenheiten im Lackauftrag und kleinste Dellen problemlos erkennen. Auch auf großen Fensterflächen führt bereits eine Durchbiegung von wenigen Mikrometern zu deutlich sichtbaren Verzerrungen der in der Scheibe gespiegelten Szene. Aufmerksame Beobachter werden zusätzlich feststellen, dass die Stärke der Verzerrungen (also die Messempfindlichkeit) von Parametern wie Blickwinkel und Abstand zum Objekt abhängt, wie in Bild 5 dargestellt wird. Die Fensterfront zeigt eine bei kurzem Abstand zur Scheibe nur ansatzweise erkennbare Durchbiegung. Diese sind bei einem größeren Abstand deutlich an den Verzerrungen des Spiegelbildes zu erkennen.

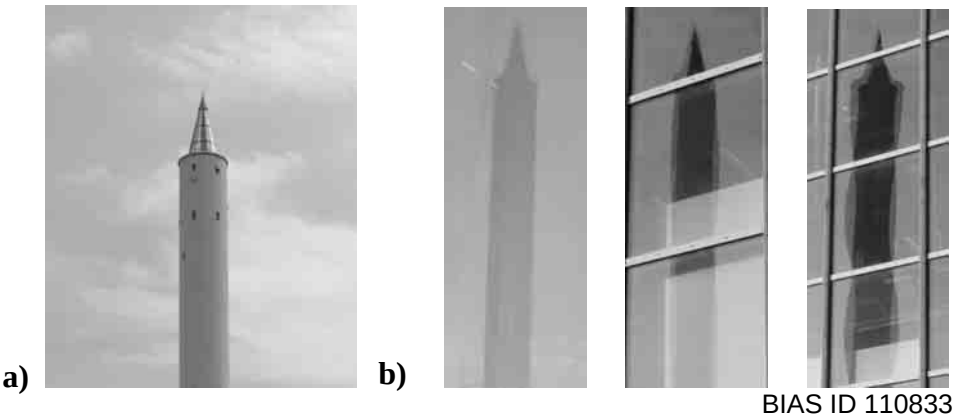


Bild 5: a) Foto der Spitze des Fallturms (ZARM) am Campus der Uni Bremen
b) Verzerrtes Spiegelbild des Fallturms aus verschiedenen Abständen zur verglasten Rückseite des BIAS-Gebäudes

Bei der Messung mit einem SRT-Messaufbau wird analog zur SPT eine Absolutphasenmessung vorgenommen, bei der ein Streifenmuster auf einem Mustergeber (TFT-Monitor) dargestellt und die Verformung des Spiegelbildes quantitativ bestimmt wird. Ergebnis der Absolutphasenmessung ist eine Karte mit Koordinatenverknüpfungen von jedem Pixel in der Kamerasensorebene zu einer entsprechenden Position in der Mustergeberebene, die auf dem Sichtstrahl des Pixels liegt. Solche Koordinatenpaare erlauben aber die Ermittlung der Objektoberfläche noch nicht, denn die Zuordnung zwischen dem Abstand der Kamera zum vermessenen Bereich auf der Objektoberfläche und der lokalen Neigung desselben ist nicht eindeutig [9]. Zur Auflösung dieser Uneindeutigkeit sind verschiedene Verfahren bekannt [10]. Bei dem verwendeten Aufbau wird das Ob-

jekt so im Bildfeld zentriert, dass seine Oberfläche einen Punkt im Raum enthält, der auf der optischen Achse der Kamera liegt und entlang derselben einen definierten Abstand *d* zum Kamerazentrum hat. Durch Kenntnis des Abstands *d* kann dann mittels der korrespondierenden Sensor-/ Mustergeber-Koordinaten für den entsprechenden Oberflächenpunkt dessen Normalenwinkel berechnet werden. Davon ausgehend können in einem iterativen Verfahren Gradienten und Formdaten für die gesamte Oberfläche gewonnen werden.

Die Positionierung des Objekts auf einen Sollabstand vom Messkopf d_0 kann von Hand entweder mit visueller Kontrolle mittels zweier sich im Raum kreuzender Laserstrahlen aus handelsüblichen Laserzeigern durchgeführt werden (Genauigkeit ±0,25 mm), oder es wird ein in den Messkopf integrierter Distanzsensor verwendet, der den Istwert mit einer Genauigkeit von 1 µm messen kann. Bei Verwendung eines Distanzsensors kann bei einer Messfeldgröße von 47 mm x 50 mm der systematische Restfehler im Höhenbild auf etwa 110 nm reduziert werden [11].

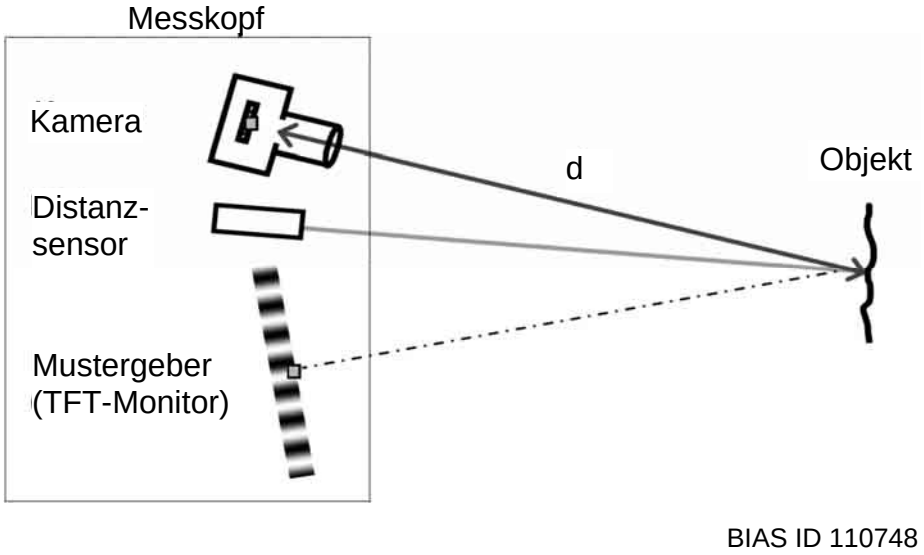


Bild 6: Skizze des SRT – Messaufbaus mit Distanzsensor. Die Kamera ist auf das Objekt ausgerichtet und nimmt das virtuelle Bild des auf dem TFT-Monitor dargestellten Streifenmusters auf. Aus: M. Sandner, 2011 [11]

Viele mögliche Anwendungsfelder für die SRT wurden z. B. in **[**12**]** beschrieben. Die wichtigste Voraussetzung ist eine zumindest teilweise spiegelnde Oberfläche: Metall, Glas, Kunststoff, Lack oder Flüssigkeitsoberflächen wurden erfolgreich durch die SRT gemessen.

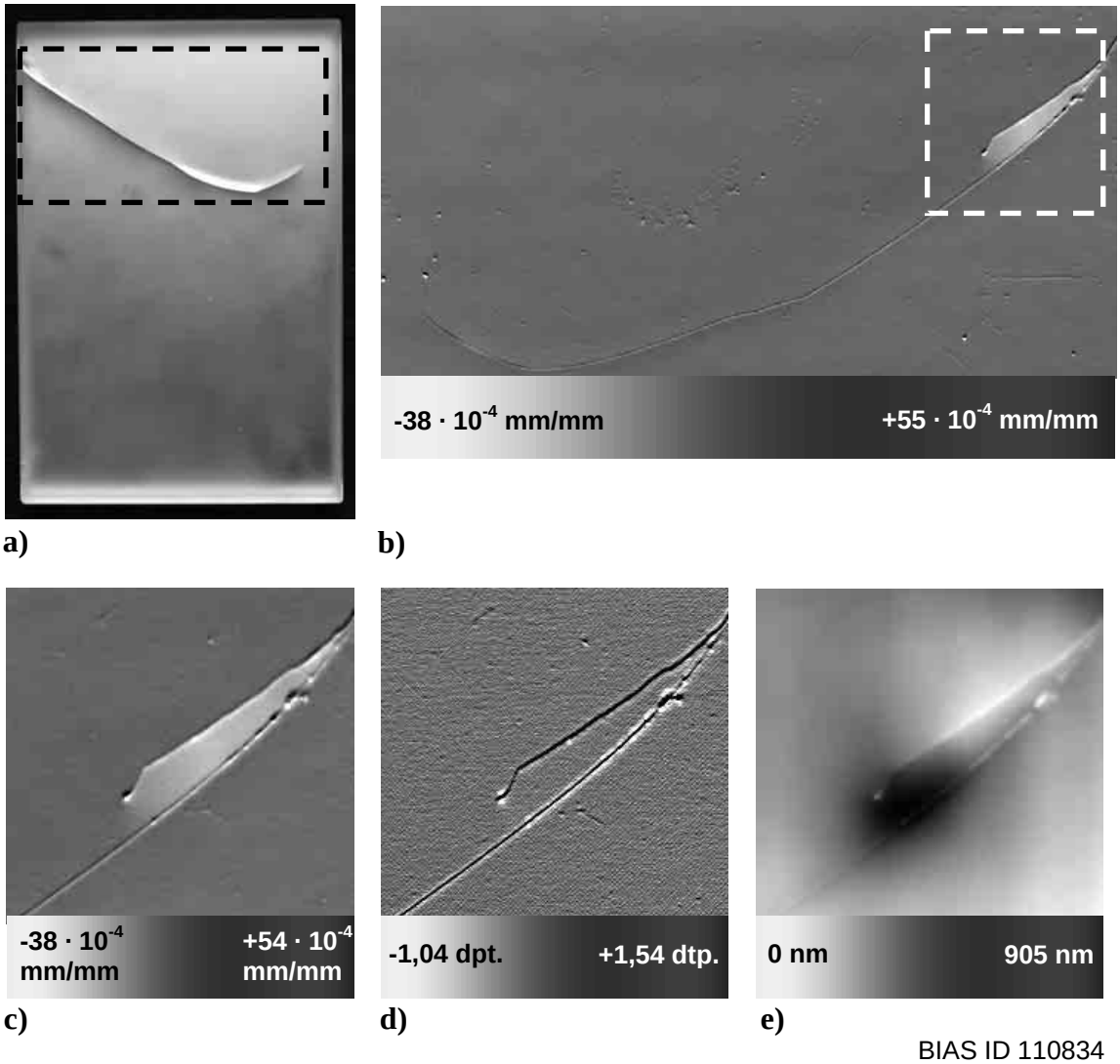
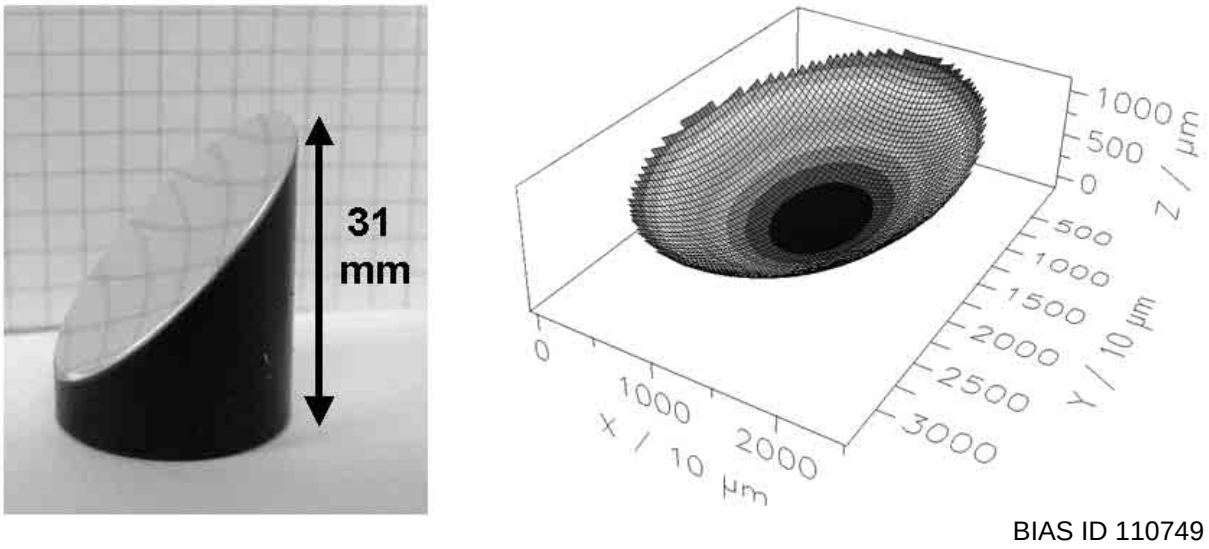


Bild 7: Vermessung eines Spiegel mit gebrochenem Substrat mittels der SRT: a) Foto des Spiegels (Substratseite), Ausschnitt für b) in schwarz eingezeichnet; b) Mit SRT-Aufbau gemessener lokaler Oberflächengradient in vertikaler Richtung (Spiegelseite), Ausschnitt für c)-e) in weiß eingezeichnet; c) Detail aus Oberflächengradientenbild in vertikaler Richtung; d) Detail aus Bild der lokalen Krümmung (in Einheit Dioptrien; Ableitung der Gradienten) in vertikaler Richtung; e) Detail aus Höhenbild (Integral der Gradienten), Helligkeitscodierung für Tiefe in z-Richtung

Beispielhafte Ergebnisse für eine Messung eines ebenen Spiegels sind in Bild 7 gezeigt. Der in Bild 7 a erkennbare Riss im Substrat des vermessenen Spiegels, aber auch Kratzer und Staub werden in Gradienten- und besonders gut im Krümmungsbild erkennbar. Es ist bekannt, dass sich die lokale Krümmung einer Fläche hervorragend eignet, um hochfrequente Oberflächenstörungen zu detektieren; in vielen Fällen sogar besser als die Höhendaten selbst. So lassen sich Größen wie „Welligkeit“ oder „Beulen und Dellen“ in bestimmten Ortsfrequenzbereichen definieren, wo nicht nur die Höhe einer Struktur ein Ausfallkri-

terium darstellt, sondern die örtliche Änderung der Oberflächenneigung, d.h. die Krümmung.

Ein weiteres Beispiel für Messungen mit dem FRT-System ist die Messung eines sog. 90°-Off-Axis Parabolspiegels, siehe Bild 8 a. Das Objekt wurde hierfür mit korrektem Abstand d_0 zum Messkopf ausgerichtet. Die aus der Messung berechneten Höhendaten sind in Bild 8 b dargestellt.



a) **b)**

Bild 8: Vermessener Off-Axis Parabolspiegel: a) Foto des Spiegels, b) Höhendaten, gemessen mit SRT-Aufbau unter Verwendung eines Distanzsensors. Aus: M. Sandner, 2011 [11]

Der Parabolspiegel wurde bereits bei der CSIRO in Sydney mit zwei unterschiedlichen interferometrischen Nulltest-Aufbauten vermessen ([13], [14]). Um die aus diesen Messungen erhaltenen Formabweichungen (Bild 9, linke Bilder) mit den Höhendaten aus der SRT-System-Messung vergleichen zu können, wurde von letzteren ein gefittetes Rotationsparaboloid abgezogen. Die Brennweite desselben ergibt sich zu 25,5 mm, was innerhalb der Herstellerangaben liegt (f = 25,4 mm ± 0,254 mm). Nach Abzug der berechneten Modellform von den SRT-Messdaten tritt dann die Feinstruktur zu Tage wie bei den interferometrischen Nulltests, welche zu absoluter Formmessung nicht ohne weiteres in der Lage sind, aber mit hoher Genauigkeit die optischen Fehler liefern. Die aus den SRT-Daten erhaltene Feinstruktur ist in Bild 9, rechtes Bild zu sehen und zum besseren Vergleich entlang der x-Achse gestaucht. Das Ergebnis stimmt in Dynamik und Struktur gut mit den interferometrisch gemessenen Formabweichun-

gen überein. Die unterschiedliche Durchbiegung der vertikal verlaufenden Rillen resultiert aus der unterschiedlichen Perspektive bei den Aufnahmen [13], [14].

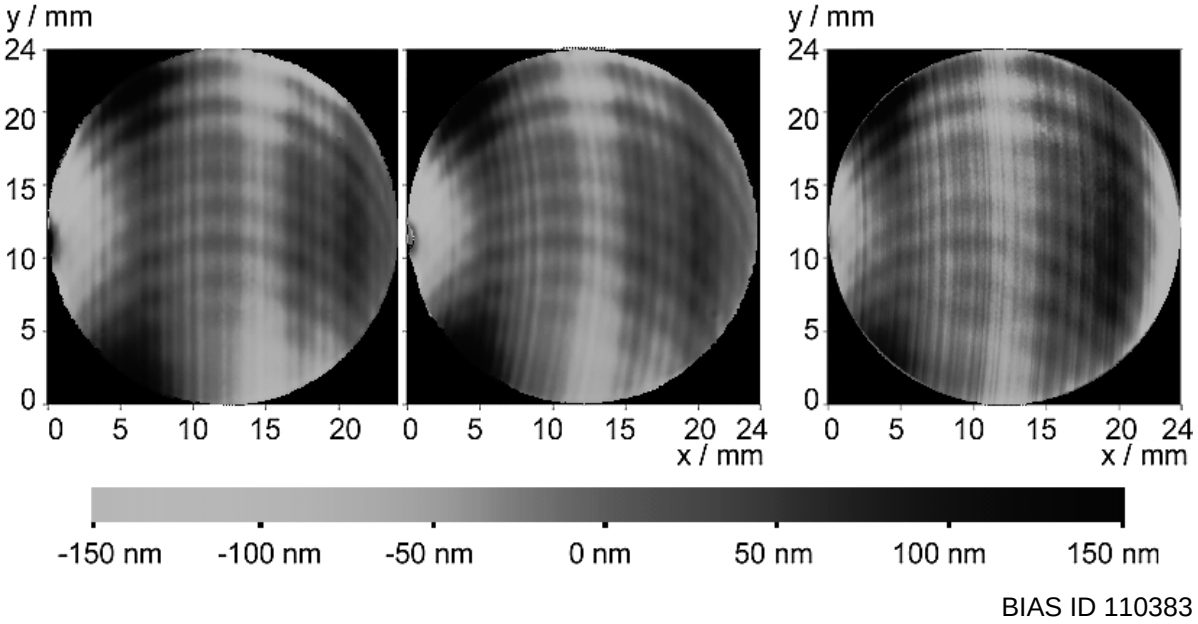


Bild 9: Bilder der Formabweichung eines 90° Off-Axis Parabolspiegels in z-Richtung. Linke Bilder: Ergebnisse der interferometrischen Nulltest-Messungen. Rechtes Bild: Berechnete Feinstruktur aus Streifenreflexionsmessung (in x-Richtung gestaucht). Aus: M. Sandner, 2011 [11]. Linke Bilder aus J. Burke, 2009 [13], [14].

3 Fazit

Die hier vorgestellten berührungslosen und schnellen Messmetechniken Streifenprojektionstechnik und Streifenreflexionstechnik sind in einem weiten Bereich von Umgebungsbedingungen nutzbar und werden bereits in zahlreichen industriellen Entwicklungs- und Überprüfungsanwendungen eingesetzt. Die Techniken liefern Oberflächendaten mit hoher Auflösung und Genauigkeit. Ein Vergleich zwischen Messdaten aus einer Streifenreflexionstechnik-Messung und solchen aus einer interferomterischen Messung zeigen, dass die Auflösung der Messungen vergleichbar ist. Neuere Entwicklungen am BIAS weisen auf weiteres Entwicklungspotenzial der Streifenreflexionstechnik hin.

4 Danksagung

Wir bedanken uns bei den Institutionen, die die Entwicklung der 3D-Kamera und der Streifenreflexionstechnik gefördert haben: dem Bundesministerium für Bildung und Forschung (BMBF), der Deutschen Forschungsgemeinschaft (DFG), der Bremer Innovationsgesellschaft, der Europäischen Kommission, sowie bei den Firmen VEW Bremen und Satisloh.

Des Weiteren danken wir dem Australian Centre for Precision Optics am CSIRO, und insbesondere Jan Burke für die Bereitstellung des vermessenen Off-Axis Parabolspiegels, der interferomterischen Messdaten und Hilfe bei der Visualisierung der Vergleichsdaten aus unserer SRT-Messung.

Literatur

[1] Bergmann, R. B., Bothe, T., Falldorf, C., Huke, P., Kalms, M., von Kopylow, C.: “Optical metrology and optical non-destructive testing from the perspective of object characteristics“, Invited Paper, Proceedings SPIE Interferometry XV: Applications, Vol. 7791, 779102, S. 1-15, 2010

[2] Chen, F., Brown, G.M., Song, M.: “Overview of three-dimensional shape measurement using optical methods”, Opt.Eng. 39, S. 10-22, 2000

[3] Burke, J., Bothe, T.: “Reverse Engineering by fringe projection”, Proc. SPIE Vol.4778, S. 312-324, 2002

[4] Schulte, M., Bothe, T., Li, W., Gesierich, A., von Kopylow, C., Jüptner, W.: “Hochauflösende Sichtstrahlkalibrierung für optisch abbildende Systeme am Beispiel Streifenprojektion”, Proc. Oldenburger 3D-Tage 2006, S. 130-137, 2006

[5] Bothe, T., Li, W., Schulte, M., von Kopylow, C., Bergmann, R. B.: “Vision Ray Calibration for the Quantitative Geometric Description of General Imaging and Projection Optics in Metrology“, Applied Optics 49(30), S. 5851-5860, 2010

[6] Bothe, T., Gesierich, A., von Kopylow, C., Jüptner, W.: "3D-Kamera - ein miniaturisiertes Streifenprojektionssystem zur Formerfassung", Tagungsband der Oldenburger 3D-Tage, in "Photogrammetrie, Laserscanning, Optische 3D-Messtechnik", S. 38-47, Wichmann Verlag, Augsburg, 2004

[7] Li, W., Bothe, T., Osten, W.: "Object adapted Pattern Projection – Part I: Generation of Inverse Patterns", Optics and Lasers in Engineering 41, S. 31-50, 2002

[8] Li, W., Bothe, T., Kalms, M., von Kopylow, C.: "Applications for Inverse Pattern Projection", Proc. SPIE Vol. 5144, S. 493-503, 2003

[9] Knauer, M., Kaminski, J., Häusler, G. (2004): Phase measuring deflectometry: a new approach to measure specular free-form surfaces Proc. SPIE Vol. 5457, S. 366-376

[10] Heizman, M.: „Moderne Methoden für bildaufnehmende 3-D-Inspektion für Oberflächen“, 37. Heidelberger Bildverarbeitungsforum, Vortrag, 2008

[11] Sandner, M., v. Kopylow, Ch., Bergmann, R. B.: „Absolut-Abstandsreferenz für die Streifenreflexionstechnik zur Verringerung systematischer Messfehler“, DGaO-Proceedings 111, A2, 2011

[12] Bothe, T., Li, W., Kopylow, C., Jüptner, W.: “High Resolution 3D Shape Measurement on specular surfaces by fringe reflection”, Proc. SPIE Vol. 5457, S. 411-422, 2004

[13] Burke, J., Wang, K., Bramble, A.: “Null test of an off-axis parabolic mirror. I. Configuration with spherical reference wave and flat return surface”, Optics Express 17(5), S. 3196-3210, 2009

[14] Burke, J.: “Null test of an off-axis parabolic mirror. II. Configuration with planar reference wave and spherical return surface”, Optics Express17(5), S. 3242-3254, 2009

Miniaturisierter IR-Öl-Sensor (MIRÖS)

Benjamin R. Wiesent, Daniel G. Dorigo, Özlem Şimşek und Alexander W. Koch

Lehrstuhl für Messsystem- und Sensortechnik,Technische Universität München, Prof. Koch, 80333 München, E-Mail: b.wiesent@tum.de

Kurzfassung

Infrarot basierte Ölzustandssensoren liefern wertvolle Informationen über den Zustand von Getrieben und ermöglichen die langfristige Planung von Wartungsarbeiten. Derzeit verfügbare Systeme basieren meist auf nicht-dispersiven Verfahren und liefern lediglich Information über einen Ölalterungsparamter. Dabei sind sie oftmals nur auf einen Öltyp angepasst. Miniaturisierte Infrarotspektrometer bieten die Möglichkeit, viele relevante Ölalterungsparameter simultan zu erfassen und sind dabei gleichzeitig auf einen großen Kreis von Öltypen anwendbar. In diesem Beitrag präsentieren wir die Realisierung eines miniaturisierten Infrarotspektrometers auf Basis eines Linear Variablen Filters (LVF) zur Getriebeölüberwachung von Offshore-Windkraftgetrieben. Dabei zeigen und diskutieren wir Einschränkungen solcher Systeme im Hinblick auf spektrale Bandbreite, Auflösung und SNR auf die Güte der spektralen Datenauswertung.

1 Einleitung

1.1 Motivation

Die Online-Überwachung des Ölzustands in großen Offshore-Windkraftgetrieben bietet eine Vielzahl neuer Möglichkeiten für die Wartung der Anlagen. Der Prozess der Ölalterung verläuft durch die großen Schmierstoffmengen relativ langsam. Durch eine kontinuierliche Online-Überwachung kann die Restbetriebszeit des Öls bereits lange im Vorfeld bestimmt werden. Anstehende Wartungen, gerade im Offshore Bereich, können bedarfsabhängig geplant und in Schönwetterphasen ausgeführt werden. Lange Stillstandszeiten

oder ein zu frühes Wechseln des Öls und die damit verbundenen Kosten lassen sich somit minimieren.

1.2 Ölzustandsanalyse

Zur Bestimmung des Ölzustands wird derzeit meist im halbjährlichen Zyklus eine Probe entnommen und in einem spezialisierten Labor chemischen und physikalischen Tests unterzogen. Dabei umfasst eine Laboruntersuchung den Test auf Verschleißmetalle, Verunreinigungen, Ölzustand, Additivabbau und auf Partikel [1]. Durch Methoden der IR-Spektroskopie können die Parameter Verunreinigung, Ölzustand und Additivabbau gemessen werden, auf welche wir im Folgenden näher eingehen.

Unter Verunreinigungen von Öl versteht man z.B. das Eindringen von Wasser oder festen Fremdstoffen. Feste Fremdstoffe beschleunigen durch eine katalytische Wirkung verstärkt den chemischen Verschleiß [2]. Kontamination mit Wasser beschleunigt die Öloxidation und kann manche hydrolyseempfindliche Additive verändern, wodurch ölunlösliche, teilweise korrosive Spalt- und Folgeprodukte entstehen [2].

Die chemisch-physikalischen Schmierstoffparameter Viskosität bei $40\ ^\circ C$ Oxidation und Total Acid Number (TAN) bilden den Ölzustand ab. Bei der Schmierstoffalterung kommt es in vielen Fällen zu einer Erhöhung der Viskosität gegen Ende der Betriebszeit. Dies hat einen Anstieg der Temperatur und letztendlich den Verlust der Schmierwirkung zur Folge [3]. Die TAN korreliert direkt mit dem Anstieg der Oxidation und gibt die Gesamtheit aller sauren Bestandteile des Öls an [3].

Die Oxidation wird aus dem Infrarot-Transmissionsspektrum berechnet und gilt in vielen Fällen als Synonym für die klassische Ölalterung. Durch Oxidation des Kohlenwasserstoffs wird das Öl sauer, wodurch ölunlösliche, teilweise korrosive Spalt- und Folgeprodukte entstehen [4]. Eine Messung der Oxidation ist aber nur bei Getriebeölen ohne Ester-Komponente möglich, da hier die Oxidationsbande bei $1710\ cm^{-1}$ nicht überlagert wird.

Die chemische Veränderung der im Schmieröl befindlichen Additive kann ein erwünschter Effekt sein, z.B. wenn sich eine schützende Salzschicht auf einer

Metalloberfläche bildet. Sie kann aber auch zu einer unerwünschten Zersetzung führen, wenn sie durch äußere Einflüsse verursacht wird. Durch eine gemeinsame Betrachtung all dieser Parameter lassen sich aussagekräftige Schlüsse über den Öl- und Maschinenzustand ableiten.

1.3 Stand der Technik

Die derzeit verfügbaren Online- und Vor-Ort Analysemethoden beruhen auf Sensorik, mit der oftmals nur ein spezieller Alterungsparameter überwacht werden kann. Dabei werden vorwiegend nicht-dispersive (ND), IR basierte Verfahren angewandt. Die auf diesen Verfahren basierenden Sensoren bestehen aus einer Kombination von IR-Detektor und spektralen Bandpassfiltern. Dabei wird die Filtertransmission speziell auf eine Bande angepasst, die sich dem zu messenden Parameter eindeutig zuordnen lässt. Es sind mehrere Systeme mit bis zu vier spektralen Kanälen publiziert [5]. Diese Systeme sind robust gegen Umwelteinflüsse, günstig und erfüllen ihre Messaufgabe für das jeweilige Öl. Ein Einsatz solcher Ölzustandssensoren für eine große Zahl an Ölen ist jedoch nicht möglich, da dabei die spektralen Bandpassfilter an jedes Öl neu angepasst werden müssten.

1.4 Lösungsansatz

Ziel unserer Forschung ist die Entwicklung, eines kleinen, robusten IR-Spektrometers zur kontinuierlichen, schnellen Detektion der Parameter Oxidation, Neutralisationszahl, Additivabbau und den Eintrag von Fremdstoffen, wie z.B. Wasser, oder Falschöle [6]. Bei einer Verkürzung des Messintervalls von aktuell sechs Monaten auf eine tägliche Messung lassen sich schnell verändernde Parameter, wie z.B. der Anstieg des Wassergehalts, zeitnah erfassen. Das System muss zudem für unterschiedliche Öltypen verwendbar sein. Dazu verwenden wir ein robustes miniaturisiertes IR-Spektrometer ohne bewegliche Teile (Bild 1) in Kombination mit einem multivariaten Kalibrationsalgorithmus. Eine Anpassung des Messsystems an den jeweils zu messenden Öltyp erfolgt über die Anpassung der Kalibration. Dabei sind keine Änderungen an der Hardware nötig.

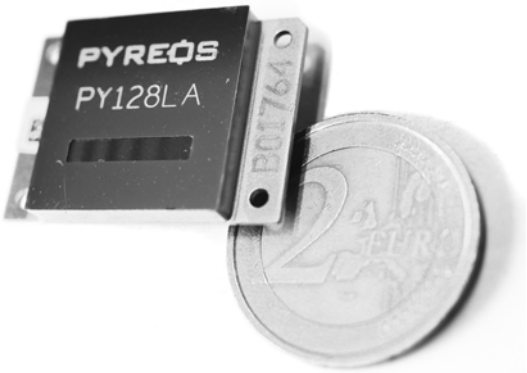


Bild 1: *Spektralapparat des IR-Miniaturspektrometers, bestehend aus 128-Element-Zeilendetektor und LVF.*

2 Linear Variable Filter (LVF) IR-Miniaturspektrometer

2.1 Linear Variable Filter (LVF)

Als dispersives Element verwenden wir Linear Variable (Bandpass-) Filter (LVF). Dabei handelt es sich um Fabry-Pérot Etalone mit einem, über der Länge des Filters linear verlaufenden Transmissionsverhalten. Sie bestehen aus einem keilförmig ansteigenden Resonator und zwei dielektrischen Spiegeln (Bragg-Reflektoren) auf der Ober- und Unterseite [7]. Die spektrale Auflösung ist dabei geringfügig niedriger als bei einem korrespondierenden Fabry-Pérot-Filter, aufgrund der nicht parallelen Reflektoren. Dieser Effekt ist als ”walk-off effect” bekannt und beschreibt die Veränderung des Transmissionsverhaltens [8]. Die jeweilige Filtertransmission weist dabei Gauß-förmiges Verhalten auf [9] und kann wie folgt beschrieben werden:

$$T_{LVF} = e^{-\frac{1}{2}(\frac{\lambda-\lambda_c(x)}{\sigma})^2} \tag{1}$$

Dabei ist T_{LVF} die LVF Transmission an der zur jeweiligen Zentralwellenlänge $\lambda_c(x)$ gehörenden Position x des Filters. Die intrinsische spektrale Auflösung wird dabei in $\%$ der Zentralwellenlänge λ_c angegeben, welche linear über den Filter zunimmt und wie folgt berechnet werden kann:

$$\Delta\lambda(x) = c \cdot \lambda_c(x) \tag{2}$$

Dabei ist $\Delta\lambda$ die spektrale Auflösung üblicherweise in μm, c ist üblicherweise die Full width at half maximum (FWHM) des LVF in % der Zentralwellenlänge. Wir verwenden ein LVF im Spektralbereich $5{,}5 - 11\ \mu m$ und einer FWHM von

$c = 1{,}8\ \%$. Dieses 2:1 Verhältnis zwischen den Grenzen des Transmissionsbandes ist auf den freien spektralen Bereich (FSR) von Fabry-Pérot Resonatoren in erster Ordnung zurückzuführen.

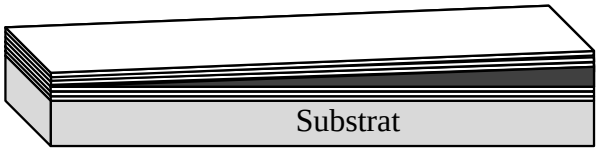


Bild 2: *Struktur eines Linear Variablen Filters (LVF).*

2.2 Systemaufbau

Die zentrale IR-Komponente des Messsystems besteht aus einer Kombination von LVF und einem pyroelektrischen Zeilendetektor des Herstellers *Pyreos* mit 128 Detektorelementen. Pyroelektrische Detektoren reagieren auf eine Änderung des Strahlungsflusses, was eine Modulation der IR-Lichtquelle erfordert, und liefern ein differentielles Ausgangssignal. Die Detektorelemente bestehen aus dem modifizierten keramischen Werkstoff PZT (Blei-Zirkonat-Titanat) mit einer Pixelbreite von $65\ \mu m$ und einem Mittenabstand von $100\ \mu m$. Der Zeilendetektor verfügt über eine, im Gehäuse untergebrachte, anwendungsspezifische integrierte Schaltung (ASIC), bestehend aus einem Ladungs-Spannungswandler, einem Tiefpassfilter, einem Verstärker, einer Sample & Hold-Stufe und einem Analog-Multiplexer. Der LVF wird wenige μm über dem Zeilendetektor in das Gehäuse integriert. Damit entsteht ein sehr kompaktes und robustes Miniaturspektrometer ohne bewegliche Teile. Um die große Apertur des LVF Spektrometers möglichst optimal zu beleuchten und somit das Signal-Rausch-Verhältnis (SNR) zu erhöhen verwenden wir eine Multi-Emitter-IR-Lichtquelle bestehend aus zwei IR-Strahlern mit parabolischem Reflektor (Bild 3). Dazu verwenden wir die elektronisch modulierbare thermische IR-Quelle *IR-55* von *Hawkeye Technologies* und betreiben sie bei einer Temperatur von $750\ ^\circ C$, einer Modulationsfrequenz von $10\ Hz$ und einem Tastgrad von $25\ \%$. Eine in MEMS-Technologie gefertigte Membran bestehend aus einer dünnen Schicht von diamantähnlichem, nanostrukturiertem amorphem Kohlenstoff des Herstellers *Intex* bildet den Kern dieser Quelle. Die Abstrahlcharakteristik ist dabei ähnlich der eines schwarzen Strahlers mit einem Emissi-

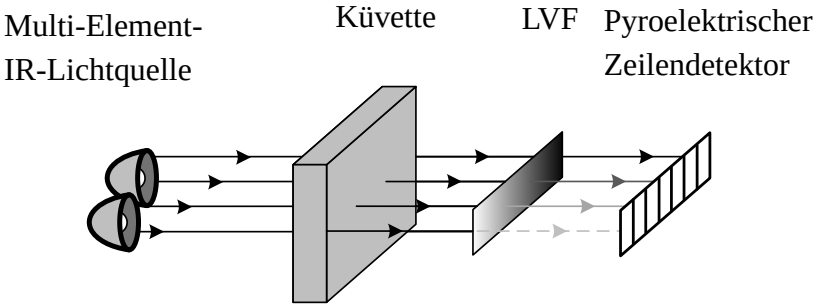


Bild 3*: IR-Miniaturspektrometer zur Online-Detektion von Ölalterung, bestehend aus einem 128 Element pyroelektrischen Zeilendetektor, einem Linear Variablen Filter (LVF) im Spektralbereich* $5{,}5 - 11\ \mu m$*, einer Küvette bestehend aus Zinkselenid (ZnSe) und einem* $100\ \mu m$ *Distanzhalter und einer Multi-Element-IR-Lichtquelle bestehend aus zwei elektrisch modulierten IR-Strahlern mit parabolischem Reflektor.*

onsgrad von $0{,}8$. Durch die geringe thermische Masse der Membran ist eine schnelle Modulation mit bis zu $10\ Hz$ bei maximaler Modulationstiefe möglich. Die Küvette besteht aus zwei Zinkselenid (ZnSe) Kristallfenstern und einem $100\ \mu m$ dicken Distanzhalter. Das realisierte Messsystem verfügt über vier

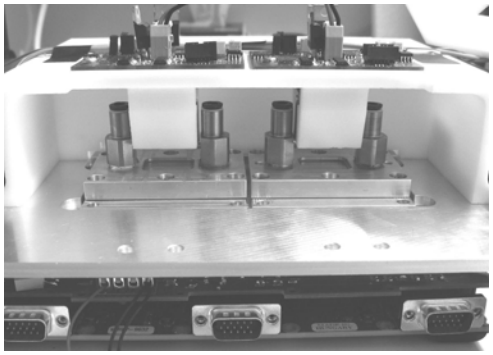


Bild 4*: Messsystem mit zwei LVF-Spektrometer Modulen.*

Messkanäle (LVF Spektrometer), wovon zwei in Bild 4 zu sehen sind. Das Messsystem wurde auf Basis eines *sbRIO-9632* von *National Instruments* für den Einsatz in einem Prüfstand entwickelt. Die nötigen Schaltkreise des LVF-Spektrometers wurden auf einer Tochterplatine realisiert, welche auf das *sbRIO* aufgesteckt wird. Die zeitkritische Ansteuerung der Messsysteme erfolgt auf dem FPGA Teil des *sbRIO*. Ablaufsteuerung und Datenauswertung erfolgen hingegen auf dem Realzeitprozessor des *sbRIO*. Das System ist autark betriebsfähig und kommuniziert über eine Ethernet-Verbindung zur in LabVIEW entwickelten Fernwartsoftware.

3 Multivariate Datenanalyse

Um die sich im Laufe der Ölalterung einstellenden Varianzen im Transmissionsspektrum mit den über andere analytische Methoden bestimmten Stoffkonzentrationen zu korrelieren, verwenden wir Methoden der multivariaten Datenanalyse. Im ersten Schritt wird mit Hilfe des Lambert-Beerschen Gesetzes das gemessene Transmissionsspektrum in ein Absorptionsspektrum umgerechnet. Damit wird ein linearer Zusammenhang zwischen Spektrum und chemischer Konzentration in der Probe hergestellt:

$$A_\lambda = -\log(T_\lambda) = \varepsilon \cdot k \cdot d \tag{3}$$

Dabei steht ε für den molaren Extinktionskoeffizienten, k für die Konzentration und d für die Schichtdicke der verwendeten Küvette. A_λ bzw. T_λ bezeichnen die bei einer bestimmten Wellenlänge auftretende Absorption bzw. Transmission [10]. Als nächstes wird eine Hauptkomponentenanalyse (Principal Component Analysis, PCA) verwendet, die mittels Datenreduktion die enthaltene Information vom Rauschen trennt. Zur Erstellung eines Kalibrationsmodells verwenden wir die Partial Least Squares Regression (PLS). Für das Modell wurde ein Datensatz bekannter Proben und zugehöriger Zielgrößen $S = (x_i, y_i)$ erstellt. Diese wurden aus einer Gebrauchtöldatenbank eines großen, unabhängigen Ölanalyselabors gefiltert. Dabei wurden sämtliche auftretenden realen Alterungszustände abgedeckt und normalverteilt in den Kalibrationsdatensatz aufgenommen. Aus diesem Datensatz wurde ein lineares Kalibrationsmodell $\hat{y} = f(x)$ berechnet und mit einem externen, ebenfalls normalverteilten Datensatz validiert. Der sich ermittelte Zusammenhang kann wie folgt formuliert werden:

$$\hat{y} = \mathbf{x}^T \mathbf{b}_1 + b_0 \tag{4}$$

wobei $\hat{y}$ die geschätzte Zielgröße, $\mathbf{x}^T$ das gemessene Spektrum, $\mathbf{b}_1$ den in der Kalibration erzeugten Regressionskoeffizienten und b_0 den statischen Gleichanteil der Regressionsgeraden angibt [10]. Zur Bewertung der Modellqualität verwenden wir die Güteparameter, mittlerer quadratischer Fehler (Root Mean Square Error, RMSE) und Bestimmtheitsmaß (r^2). Als RMSE wird gemäß Gleichung 5 die Wurzel aus dem mittleren quadratischen Fehler bzw. aus der Rest-

varianz bezeichnet, wobei n die Anzahl der Proben angibt.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \tag{5}$$

Das sogenannte Bestimmtheitsmaß (Gleichung 6) gibt den Anteil der durch die Regression erklärten Varianz an der Gesamtvarianz der Zielgröße an. Der Parameter $\overline{y}$ bezeichnet die Mittelwerte der gemessenen Zielgrößen. [10]

$$r^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{6}$$

In unserer Untersuchung verwenden wir FTIR-Spektren zweier weit verbreiteter Windgetriebeöle (ein mineralisches und ein synthetisches) im Spektralbereich $4000 - 500\ cm^{-1}$ bei einer Auflösung von $1\ cm^{-1}$. Dabei werden die Spektralbereiche $1330\ cm^{-1}$ bis $1500\ cm^{-1}$ und $2758\ cm^{-1}$ bis $3067\ cm^{-1}$ ausgeblendet, da in diesen Bereichen Absorptionswerte größer als $2{,}5$ gemessen werden und somit Nichtlinearitäten auftreten.

4 Untersuchung der Einflüsse realer Messaufbauten

Das hier vorgestellte Messsystem unterscheidet sich von FTIR-Spektrometern hinsichtlich der drei entscheidenden Kriterien spektrale Bandbreite, spektrale Auflösung und Signal-Rausch-Verhältnis (SNR).

Die spektrale Bandbreite und Auflösung wird maßgeblich durch den verwendeten LVF beschränkt. Wir haben zwei kommerziell erhältliche LVF und dessen Einfluß auf die Modellgüte untersucht (Tabelle 1). Bei dem untersuchten

Anordnung	spektrale Bandbreite		c	spektrale Auflösung
	Wellenlänge	Wellenzahl		
LVF 1	5.5 µm - 11 µm	$1818\,\mathrm{cm}^{-1}$ - $909\,\mathrm{cm}^{-1}$	1,0 %	$9.3\,\mathrm{cm}^{-1}$ - $18.5\,\mathrm{cm}^{-1}$
			1,8 %	$16.7\,\mathrm{cm}^{-1}$ - $33.3\,\mathrm{cm}^{-1}$
LVF 2	7.0 µm - 14 µm	$1429\,\mathrm{cm}^{-1}$ - $715\,\mathrm{cm}^{-1}$	1,0 %	$7.2\,\mathrm{cm}^{-1}$ - $14.3\,\mathrm{cm}^{-1}$
			1,8 %	$12.9\,\mathrm{cm}^{-1}$ - $25.7\,\mathrm{cm}^{-1}$

Tabelle 1*: Übersicht über die untersuchten linear variablen Filter*

mineralischen Öl wird exemplarisch auf die Prädiktion der TAN eingegangen,

Öltyp	LVF	Zielgröße	r^2	RMSE	Anzahl PCs	Kalibrier-proben	Validier-proben
mineralisch	1		0,847	0,024	4	55	139
	2	TAN	0,838	0,025	6	42	152
	1 + 2	[mgKOH/g]	0,895	0,022	6	50	144
	FTIR		**0,899**	**0,020**	**6**	**50**	**144**
synthetisch	1		0,825	111,9	5	50	147
	2	Molybdän	0,892	125,3	4	45	152
	1 + 2	[ppm]	0,905	119,4	5	47	150
	FTIR		**0,875**	**92,61**	**5**	**57**	**139**

Tabelle 2*: Übersicht über die Detektionsgüten der Zielgrößen TAN und Molybdän (mineralisches bzw. synthetisches Getriebeöl) bei Verwendung eines Spektrometers auf Basis eines Linear Variablen Filters mit* $c = 1{,}8\ \%$.

während bei dem synthetischen Polyalphaolefinöl (PAO) der Additivabbau anhand der Molybdän-Konzentration verdeutlicht wird. Dabei wurde der Datensatz $S = (x_i, y_i)$ auf den Spektralbereich von LVF 1, LVF 2, einer Kombination aus LVF 1 und 2 beschränkt und die spektrale Auflösung dabei auf jeweils $c = 1\ \%$ bzw. $c = 1{,}8\ \%$ reduziert. Dieser LVF-Datensatz wurde mittels der PLS-Methode unter Verwendung einer vollen Kreuzvalidierung ausgewertet. In Tabelle 2 sind die Ergebnisse dieser Analyse und ein Vergleich zu den Ergebnissen mit hochqualitativen FTIR-Spektren dargestellt.
Aus Tabelle 2 ist bei Vergleich der Güteparameter r^2 und RMSE zu erkennen. Sowohl bei LVF 1 als auch bei LVF 2 werden sehr hohe Prädiktionsgüten, ähnlich den Ergebnissen der FTIR Spektren erzielt. Eine Kombination von LVF 1 und 2 erzielt gleich gute Prädiktionsgüten, wie FTIR-Spektren. Die Beschränkung der spektralen Auflösung und Bandbreite, wie es bei LVF Spektrometern der Fall ist, hat somit keinen wesentlichen Einfluss auf die Bestimmung wichtiger Ölalterungsparameter.

Ein weiteres entscheidendes Kriterium ist das Signal-Rausch-Verhältnis (SNR). Dazu wurde der Einfluss von spektralem Rauschen auf die Güte der Regressionsmodelle quantifiziert und die Absorptionsspektren mit proportionalem weißen Rauschen in Höhe von $3\ \%$, $6\ \%$ und $9\ \%$ beaufschlagt. Bild 5 veranschaulicht den Einfluss des SNR bei der Bestimmung der TAN. Das SNR übt dabei

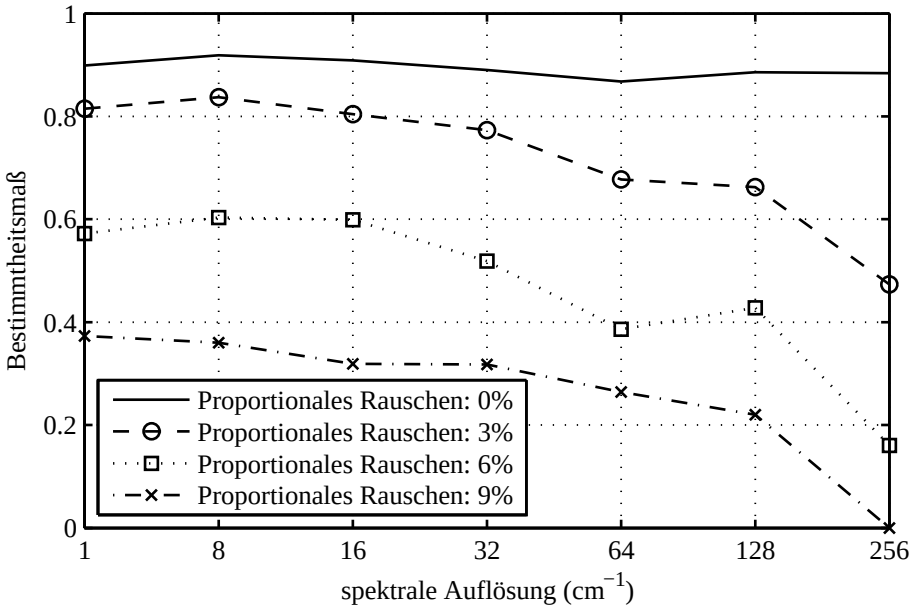


Bild 5*: Einfluss verschiedener Kombinationen aus begrenztem Auflösungsvermögen und weißem Rauschen auf das Bestimmtheitsmaß am Beispiel der Neutralisationszahl (TAN) eines mineralischen Getriebeöls.*

einen viel stärkeren negativen Einfluss auf die Prädiktionsqualität aus, als die spektrale Auflösung. Damit ist der primäre Optimierungsparameter für miniaturisierte IR-Spektrometer in der Applikation Ölzustandsanalyse identifiziert.

5 Messergebnisse

Zur Darstellung der Qualität der LVF-Spektren wurde jeweils eine Probe des untersuchten mineralischen und synthetischen Getriebeöls mit einem FTIR- und mit unserem LVF-Spektrometer gemessen (Bild 6). Die LVF-Spektren wurden bei einer Frequenz von $10\ Hz$ aufgenommen und über jeweils 100 Messungen gemittelt. Das gemessene LVF-Spektrum zeigt alle spektralen Details der FTIR-Messung in reduzierter Auflösung. Im Bereich um $8\ \mu m$ ist verstärkt spektrales Rauschen zu erkennen, was auf eine Dünnschichtinterferenz auf den Detektorpixeln (Dicke $1{,}4\ \mu m$) zurückzuführen ist. Die leicht nichtlineare Skalierung der Wellenlängenachse ist auf den Abstrahlkegel der IR-Quelle zurückzuführen, was die Resonatorbedingung im LVF verändert. Dieser Effekt hat auf die nachfolgende Multivariate Datenanalyse keinen Einfluss, da er statischer Natur ist und nicht die Varianz im Spektrum beeinträchtigt.

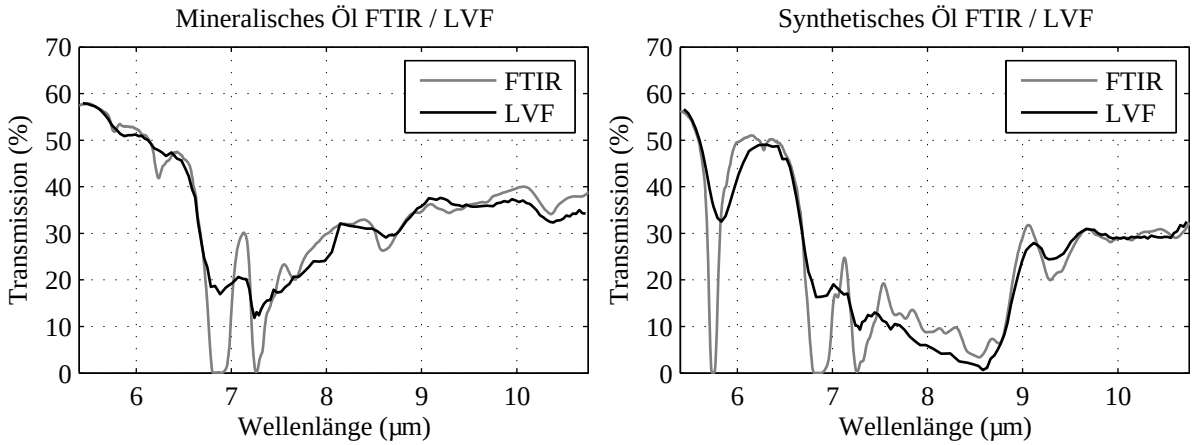


Bild 6*: Ein direkter Vergleich zwischen hochaufgelöstem FTIR-Spektrum und LVF-Spektrum anhand eines mineralischen und eines synthetischen Getriebeöls.*

6 Zusammenfassung & Ausblick

Zur Erfassung des Ölzustands in Windkraftgetrieben mittels Infrarot basierter Verfahren, sind konventionelle nicht-dispersive Verfahren oftmals nur unzureichend geeignet. Eine geeignete Alternative hierzu stellen miniaturisierte Infrarotspektrometer dar. Sie erlauben die simultane Erfassung vieler wichtiger Ölalterungsparameter und sind dabei gleichzeitig auf einen großen Kreis von Öltypen anwendbar. Im Rahmen des Projekts “MIRÖS“ wurde erstmals ein Ölzustandssensor auf Basis eines robusten, miniaturisierten, Linear Variablen Filter (LVF) Spektrometers realisiert. In diesem Beitrag wurde der Einfluss der beschränkten spektralen Bandbreite, Auflösung und SNR auf die Güte der spektralen Datenauswertung von LVF Miniaturspektrometern aufgezeigt und diskutiert. Dabei stellte sich das SNR als wichtigster Optimierungsparameter des Spektrometers heraus. Durch den Einsatz eines solchen Ölzustandssensors lassen sich schwere Maschinenschäden vermeiden und nötige Wartungsintervalle bereits lange im Vorfeld bedarfsorientiert planen. Hersteller von Getrieben können mit Hilfe unseres Messsystems überprüfen, ob ihre Vorgaben hinsichtlich der Schmiermittelwahl und Wartung eingehalten wurden. Die Größe, der niedrige Preis, die Vibrations- und Temperaturbeständigkeit des entwickelten miniaturisierten IR-Spektrometers stellen einen entscheidenden Schritt hin zum portablen, robusten MIR-Spektrometer für mannigfaltige Anwendungen dar.

7 Danksagung

Die Autoren bedanken sich bei Oelcheck GmbH und Klüber Lubrication München KG für die zur Verfügung gestellten Spektren, Analysewerte und vielen hilfreichen Diskussionen. Der Prüftechnik Condition Monitoring GmbH wird für die Unterstützung bei der Entwicklung des Messsystems gedankt, sowie der Renk AG für die Einblicke in moderne Windkraftgetriebe. Das Projekt Mirös wird durch das Bayerische Staatsministerium für Wirtschaft, Verkehr, Infrastruktur und Technologie gefördert und durch den VDI/VDE-IT fachlich und organisatorisch betreut.

Literatur

[1] R. Krethe, “Oil Condition Monitoring - Online, Onsite or Lab analysis?”, in *Proceedings of Oildoc, Rosenheim 2011*,

[2] R. M. Cresham und G. E. Totten, *Lubrication & Maintenance of Industrial Machinery*, CRC Press Taylor & Francis Group, 2009.

[3] H. Bloch, *Pracatical Lubrication For Industrial Facilities*, 2nd editio ed., 2009.

[4] K. Kudlaty, *Attenuated Total Reflection technique for on-line oil monitoring by means of a FTIR fiber-optic probe*. Phd thesis, TU München, 2004.

[5] B. R. Wiesent, D. G. Dorigo und A. W. Koch, “Suitability of tunable Fabry-Perot spectrometers for condition monitoring purposes of gear oils in offshore wind turbines”, in *Proceedings of 12th International Conference on Infrared Sensors & Systems*, S. 159, AMA Fachverband für Sensorik e.V., (Nürnberg), 2011.

[6] B. Wiesent, D. Dorigo, M. Schardt und A. Koch, “Gear oil condition monitoring for offshore wind turbines using band limited low resolution spectra”, in *Proceedings of Oildoc*, 2011.

[7] D. Morelli, *Interference Filter Handbook*, 2nd editio ed., 2006.

[8] J. Green, “The’walk-off’effect in Fabry-Perot etalons-limitations to’single-shot’linewidth measurements at longer wavelengths”, *Journal of Physics E: Scientific Instruments* **13**, S. 1302, 1980.

[9] B. Wiesent, D. Dorigo und A. Koch, “Limits of IR-spectrometers based on linear variable filters and detector arrays”, in *Proceedings of SPIE*, **7767**(77670L), S. 77670L, 2010.

[10] W. Kessler, *Multivariate Datenanalyse*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2007.

Piezoresistives Silizium-Messelement zur Messung im Hochdruck-Bereich bei 5.000 bar

Patrick Heinickel, Roland Werthschützky

Institut für Elektromechanische Konstruktionen, Technische Universität Darmstadt,
E-Mail: p.heinickel@emk.tu-darmstadt.de

Kurzfassung

In diesem Beitrag wird ein neuartiges piezoresistives Messelement zur Erfassung von Hochdruck bis 5.000 bar präsentiert. Dieses Verbundelement besteht aus einem festen Materialverbund eines prozessierten Siliziumchips und eines Glasgegenkörpers. Durch allseitige Druckbelastung des massiven Materialverbundes verspannt sich dieses aufgrund des mechanisch fehlangepassten Gegenkörpers. Die implantierten piezoresistiven Widerstände ändern sich daher infolge der mechanischen Verspannung.

Mittels der präsentierten analytischen Beschreibung des Verbundelementes auf Basis des Energieerhaltungssatzes in einem Mehrschichtkörper und dem übereinstimmenden Vergleich mit FEM-Strukturanalysen, werden Messergebnisse für ein optimiertes Layout des Messelementes vorgestellt. Durch Variation des Aufbaus werden erste erfolgreiche Messungen bei einem Nenndruck von 5.000 bar präsentiert. Die Empfindlichkeit des Verbundelementes ist mit $13,8\,\frac{\mu V}{V\,bar}$ höher als die von kommerziell erhältlicher Hochdrucksensoren. Die systematischen statischen Fehler entsrechen mit kleiner 0,5% den Zielstellungen.

1 Motivation

Die Weiterentwicklung der Technik erschließt neuartige Anwendungen im Hochdruckbereich. Hierzu zählen Massenanwendungen, wie beim Einspritzsystem moderner Common-Rail Diesel Motoren bis aktuell ca. 1.800 bar. In

der Entwicklung wird bereits mit Einspritzdrücken bis 3.000 bar gearbeitet. Außerdem erfordern stark zunehmende Hochdruck-Anwendungen, wie beispielsweise das Hydroforming mit Druckbereichen bis ca. 6.000 bar und die Lebensmittel- und Chemieindustrie sogar bis über 15.000 bar, zusätzliche kostengünstige und berstfeste Lösungen.

Auf dem Markt erhältliche Hochdruck-Sensoren nach dem Stand der Technik sind meistens aus einem metallischen oder keramischen Verformungskörper mit applizierten Dehnmesselementen in Dünnfilmtechnik bzw. konventionellen DMS aufgebaut. Diese zeichnen sich durch sehr hohe Listenpreise aus und sind außerdem nicht berstfest.

Die in der Druckmessung am weitesten verbreiteten konventionellen piezoresistiven Silizium-Messelemente nach dem Stand der Technik decken einen Nenndruckbereich von etwa 0,1 bar bis ca. 1.000 bar bei einer Messunsicherheit von 1% ab (Bild 1). Steigende Belastungen führen zu einem Bersten der Druckplatte aufgrund der hohen auftretenden Zugspannungen. Um die Vorteile dieser piezoresistiven Druck-Messelemente, wie niedrige Fertigungskosten und miniaturisierbare Abmessungen bei hoher Empfindlichkeit, auch im Hockdruck-Bereich größer 1.000 bar anwenden zu können, wurde an der Technischen Universität Darmstadt ein neuartiges Konzept eines Hochdruck-Messelements entwickelt [1, 2]. Der in diesem Projekt betrachtete Hochdruckbereich reicht von 100 bar bis 5.000 bar, dies entspricht 500 MPa bzw. ca. 72.500 psi.

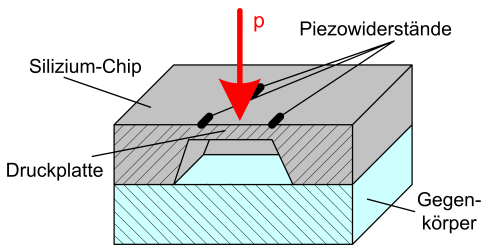


Bild 1: Konventionelles piezoresistives Silizium-Druck-Messelement; typische Kantenlänge ca. 2 mm

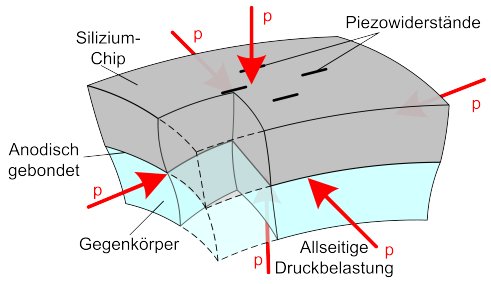


Bild 2: Neuartiges piezoresistives Silizium-Hochdruck-Messelement; Kantenlänge 1,7 mm

2 Neuartiges Messelement für den Hochdruckbereich

Der neuartige Ansatz für ein piezoresistives Silizium-Hochdruck-Messelement verzichtet auf die bisherige Schwachstelle einer Silizium-Druckplatte als Verformungskörper. Als neuartiger Ansatz wird die zur Messung erforderliche mechanische Differenzspannung im Bereich der piezoresistiven Widerstände [3] durch eine mechanische Verspannung des gesamten Messelementes erzeugt. Das sogenannte Verbundelement besteht aus einem prozessierten Siliziumchip ohne rückseitige Abdünnung zur Erzeugung einer Druckplatte, das anodisch an einen mechanisch fehlangepassten Gegenkörper gebondet ist. Die Abbildung 2 zeigt dieses Verbundelement, das eine Kantenlänge von ungefähr 1,7 mm und eine Gesamtdicke ca. 0,85 mm aufweist.

Der Verbund des Silizium-Chips als Substrat mit einem vergleichsweise weicheren Substratgegenkörper, wie bspw. Borosilikatglas, wird zugunsten der Vermeidung von Zugspannungen im Messelement und somit der gesteigerten Berstfestigkeit allseitig mit Druck belastet. Da sich der mechanisch weichere Substratgegenkörper infolge der Druckbelastung stärker komprimiert als das härtere Silizium, die Dehnung an der Zwischenschicht jedoch stetig verläuft, verwölbt sich das Verbundelement schalenförmig, wie in Bild 2 dargestellt. Die auftretenden mechanischen Differenzspannungen, als Überlagerung aus Kompression und Verbiegung, führen zu Änderungen der Piezowiderstände und somit zu einer messbaren Verstimmung der im Siliziumelement integrierten Wheatstone-Messbrücke [4].

3 Analytische Beschreibung des Verbundelements

Zur analytischen Beschreibung des Verbundelementes wird ein Energie-Ansatz gewählt, der auf der physikalischen Grundlage der *Erhaltung der Energie* innerhalb eines abgeschlossenen Systems beruht. Ähnliche Ansätze zur Lösung vergleichbarer Probleme von Mehrschichtkörpern sind in [5,6] zu finden. Diese Ansätze beschreiben jedoch nur den Fall einer eindimensionalen Verformung - zur Lösung der Beschreibung des Verbundelementes ist eine mehrdimensionale Betrachtung nötig. Das Bild 3 zeigt eine Skizze des schalenförmig verbogenen

Verbundelementes mit den für den analytischen Ansatz wichtigen Parametern.

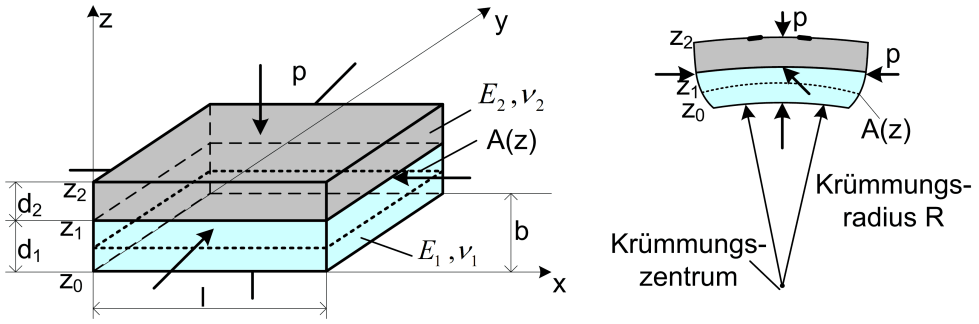


Bild 3: Schalenförmig verbogenes Verbundelement infolge hydrostatischer Druckbelastung für den analytischen Ansaz

Für ein allseitig belastetes Volumenelement ergibt sich für die mechanische Spannung T_i in i-Richtung eine Abhängigkeit von der mechanischen Dehnung S_i und von den mechanischen Koeffizienten nach Gl. (1). Für die hydrostatische Belastung wird das Hydrostatische Modul H_i als eine Kombination des Elastizitätsmoduls E_i und der Querkontraktionszahl ν_i eingeführt.

$$T_i = S_i \cdot \frac{E_i}{1 - 2\nu_i} = S_i \cdot H_i \tag{1}$$

Zur Bestimmung der mechanischen Dehnung im Verbundelement wird angenommen, dass die mechanische Dehnung S_i nur von der z-Richtung abhängig ist. Die gesamte im hydrostatisch belasteten Verbundelement gespeicherte Energie W berechnet sich nach Gl. (2) durch Integration über die von z abhängige stetige mechanische Dehnung $S(z)$ und der Summe der Energieinhalte der einzelnen Schichten [7].

$$W = \int w(z)\,\mathrm{d}V = \sum_{i=1}^{\mathrm{n}} \int_{z_{\mathrm{i}-1}}^{z_{\mathrm{i}}} \left\{ \frac{S^2(z) \cdot H_{\mathrm{i}}}{2\,(1 - 2\nu_{\mathrm{i}})} \right\} \mathrm{d}A\,\mathrm{d}z \tag{2}$$

Aufgrund der hydrostatischen Belastung des Verbundelementes wird dieses nun als frei betrachtet, d.h. es werden keine Einspannungsrandbedingungen angenommen. Der Körper nimmt somit die für ihn energetisch günstigste Lage ein. Dieses Minimum der Energie ist ein Gleichgewichtszustand, in dem keine äußeren Kräfte und Momente wirken. Es stellt sich eine Kombination aus

Kompression und schalenförmiger Verbiegung ein. Mit der BERNOULLI*'schen Hypothese der Biegetheorie* ergibt sich für die unbekannte Dehnung $S(z)$ eine Abhängigkeit von der Dehnung der Grundschicht $S_{z0} = S(z_0)$ und vom Biegeradius R nach Gl. (3):

$$S(z) = S_{\mathrm{z0}} + \frac{1}{R} \cdot z \tag{3}$$

Der Verbundkörper strebt den Zustand eines Minimums der Energie an. Daher wird Gl. (2) nach den beiden unbekannten Summanden der Gl. (3) S_{z0} und $\frac{1}{R}$ differenziert und gleich Null gesetzt.

$$\frac{\partial \frac{\partial W}{\partial A}}{\partial \frac{1}{R}} \stackrel{!}{=} 0 \quad \text{und} \quad \frac{\partial \frac{\partial W}{\partial A}}{\partial S_{\mathrm{z0}}} \stackrel{!}{=} 0 \tag{4}$$

Das partiellen Ableitungen führen zu folgendem Gleichungssystem Gl. (5):

$$\left.\begin{aligned} A \cdot S_{z0} + B \cdot \frac{1}{R} &= C \\ D \cdot S_{z0} + E \cdot \frac{1}{R} &= F \end{aligned}\right\} \quad \text{mit} \quad \left\{\begin{aligned} A &= \sum_{i=1}^{n} \left[\frac{z_i^2 - z_{i-1}^2}{2} \cdot K_i\right] \\ B &= \sum_{i=1}^{n} \left[\frac{z_i^3 - z_{i-1}^3}{3} \cdot K_i\right] \\ C &= -\sum_{i=1}^{n} \left[\frac{z_i^2 - z_{i-1}^2}{2} \cdot T_{v,i}\right] \\ D &= \sum_{i=1}^{n} \left[(z_i - z_{i-1}) \cdot K_i\right] \\ E &= \sum_{i=1}^{n} \left[\frac{z_i^2 - z_{i-1}^2}{2} \cdot K_i\right] = A \\ F &= -\sum_{i=1}^{n} \left[(z_i - z_{i-1}) \cdot T_{v,i}\right] \end{aligned}\right. \tag{5}$$

Die beiden Unbekannten ergeben sich durch Lösen des Gleichungssystems (5). Das Ergebnis liefert die Gl. (6):

$$R = \frac{DB - AE}{CD - AF} \quad \text{und} \quad S_{z0} = \frac{CE - BF}{AE - BD} \tag{6}$$

Mit Gl. (3) erhält man die mechanische Dehnung entlang der Mittelachse des infolge der allseitigen Belastung schalenförmig verbogenen Verbundelementes. Aufgrund der nicht eingespannten Seitenkanten des Messelementes können an den Rändern keine Lagerreaktionen auftreten. Dies bedeutet, dass an den Seitenkanten des Messelementes keine Momente und Querkräfte vorhanden sind. Die Normalkräfte der Seitenkanten sind aufgrund des geltenden Kräftegleichgewichts (*Reaktionsaxiom*) den orthogonal wirkenden Druckkräften entgegengerichtet. Daher ist an den Rändern des Verbundelementes die entsprechende Normalspannung $T_x(x = 0 \text{ bzw. } x = l) = -p$ bzw. $T_y(y = 0 \text{ bzw. } y = b) = -p$ vorhanden. Somit ergibt sich ein Spannungsverlauf auf der Oberseite des Silizium-Chips. Zum Rand des Verbundelementes baut sich die der Vorspannung überlagerte Biegespannung ab.

Das Bild 4 zeigt als Ergebnis der FEM-Analyse mit ANSYS Workbench die x Komponente der mechanischen Spannung T_x am allseitig mit Druck belasteten verbogenen Verbundelement. Die schalenförmige Verwölbung ist stark verstärkt dargestellt. Der errechnete Biegeradius R liegt bei einer Belastung von 1.000 bar in der Größenordnung von 10 m. Die zur Messung mittels des piezoresistiven Effektes interessante Differenzspannung $\Delta T = T_x - T_y$ ist in Bild 5 dargestellt. Die Auswertung der Strukturanalyse deckt sich mit den Ergebnissen des analytischen Ansatzes.

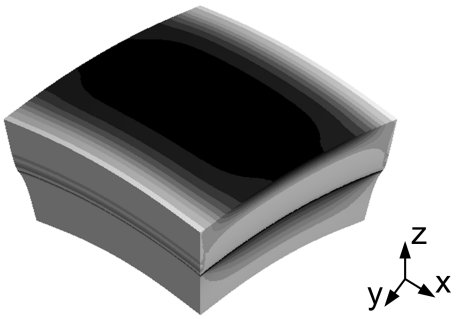


Bild 4: FEM Analyse am Verbundelement: x-Komponente der mechanischen Spannung T_x

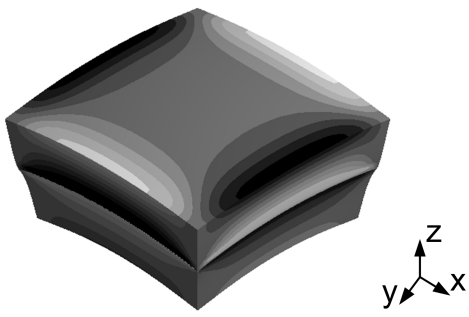


Bild 5: FEM Analyse am Verbundelement: Mechnaische Differenzspannung $\Delta T = T_x - T_y$

Analytische Berechnungen der mechanischen Spannung und Strukturanalysen nach der Finite-Elemente-Methode haben zu verschiedenen Optimierungsansätzen geführt. Die interessanten Parameter, die zur Steigerung des Übertragungsfaktors führen, sind unter anderem die Position und Orientierung der Piezowiderstände auf dem Siliziumchip und im Siliziumkristall, sowie die geometrische Gestaltung des Messelements einschließlich des Schichtdickenverhältnisses aus Silizium und fehlangepasstem Gegenkörpermaterial [8, 9].

4 Messergebnisse

4.1 Hochdruck-Messplatz und Musteraufbau

Eine Herausforderung stellt die messtechnische Charakterisierung der Verbundelemente im Hochdruck-Bereich bis 5.000 bar dar. Zur präzisen Druckerzeugung wird eine manuell getriebene Hochdruckspindel in Kombination mit einer Druckwaage eingesetzt. Als Druck-Kopplungs-Medium kommt ein spezielles Silikonöl zum Einsatz. Die maximale Messunsicherheit wird für die Druckwaage in Kombination mit dem vorhandenen geeichten Gewichtssatz auf 5 ppm beim Maximaldruck von 5.000 bar angegeben [10]. Ergänzt wird das Messsystem durch einen handelsüblichen Referenzsensor im Hochdruck-Bereich zu Vergleichszwecken.

Das neuartige Messelement wird direkt im Hochdruckbereich eingesetzt. Deshalb ermöglicht vor allem die speziell entwickelte elektrische Hochdruck-Keramik-Durchführung erst seit kurzem die eigentliche Charakterisierung der Verbundelemente bei dem angestrebten Nenndruck von 5.000 bar [9, 11]. Bild 6 zeigt ein Verbundelement, das auf einer Keramik-Durchführung aufgebaut und mittels Drahtbonden kontaktiert ist.

4.2 Messungen bei 5000 bar

Durch zahlreiche durchgeführte Verbesserungen - vor allem im Aufbau der Messelemente, wie eine Sicherstellung einer guten Kontaktierung durch Umstellung auf Drahtbonden und dem Einsatz von speziellen Durchführungen zur kontinuierlichen Steigerung des Nenndruckbereichs - werden inzwischen reproduzierbare Messergebnisse mit den Verbundelementen erzielt [4]. Die ersten

experimentellen Untersuchungen [1, 2] wiesen große Probleme durch den hohen Nenndruck auf, die zu nicht reproduzierbaren Ergebnissen führten, aber das neue Funktionsprinzip experimentell nachweisen konnten.

Optimierungen an der Gestaltung des Verbundelementes, wie die Position und Ausrichtung der piezoresistiven Widerstände [4], die Geometrie der Grundfläche [12] und das Verhältnis der Schichtdicken [8, 9], führen zu unterschiedlichen Messelement-Layouts.

In Bild 7 wird eine typische Messung eines optimierten Verbundelementes nach aktuellem Layout präsentiert. Die Schichtdicken des Messelementes betragen 390 µm für Silizium und ca. 480 µm für den Glasgegenkörper. Bild 7 zeigt die ersten Messergebnisse eines Verbundelements bei 5.000 bar Druckbelastung. Der präsentierte einzelne Lastzyklus mit einem Nenndruck von 5.000 bar und einer Zwischenstufe ist einer Serie gleicher Lastzyklen mit reproduzierbaren Ergebnissen entnommen.

Für eine Sensorspeisung von 2,5 V werden aus den Messergebnissen die in Tabelle 1 dargestellten charakteristischen Werte für den Nenndruck 5.000 bar berechnet. Für jeden in Tabelle 1 dargestellten Nenndruck werden gesonderte Messungen mit mehreren Lastzyklen analog zu Bild 7 am selben Messelement durchgeführt. Zur besseren Vergleichbarkeit der Ergebnisse untereinander wird die Empfindlichkeit B_0 der verschiedenen Messelemente durch die Nennspannung normiert auf die Versorgungsspannung und den Nenndruck angegeben: $B_0 = \frac{u_n}{U_0 \cdot p_0}$. Bei den systematischen statischen Fehlern handelt es sich um die re-

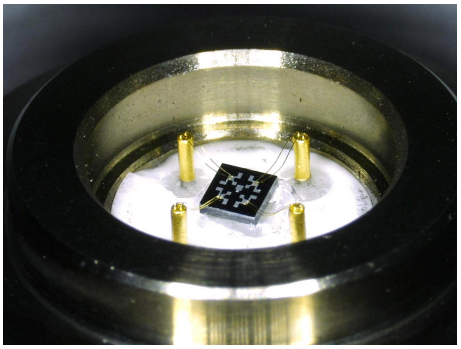


Bild 6: Verbundelement auf spezieller Keramik-Durchführung zur Messung bei 5.000 bar aufgebaut

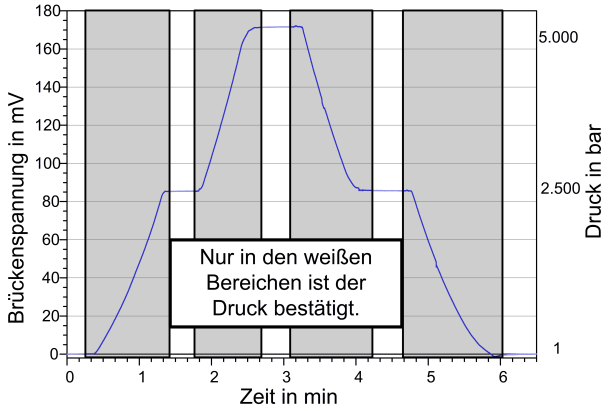


Bild 7: Messergebnis eines Verbundelementes bei einem Nenndruck von 5.000 bar

duzierten Fehler bezogen auf eine Sensorsollkennlinie nach einer Festpunkteinstellung.

Tabelle 1: Charakteristische Werte für das Verbundelement mit 2,5 V Spannungsspeisung bei unterschiedlichem Nenndruck

Nenndruck	bar	1.000	2.000	3.000	4.000	5.000
Nennspannung	mV	33,98	68,38	103,5	137,2	172,5
Empfindlichkeit	µV/Vbar	13,59	13,68	13,81	13,72	13,80
Linearitätsfehler	%	0,17	0,13	0,23	0,12	0,26
Hysterese Fehler	%	0,08	0,02	0,10	0,08	0,29
Nullpunkt Fehler	%	0,65	0,20	0,26	0,07	0,01

Der Vergleich zu Messergebnissen mit anderen Verbundelementen zeigt, dass die Ergebnisse reproduzierbar für jedes Layout sind. Zwei Elemente des gleichen Layouts unterschieden sich bei gleichen Messunsicherheiten in der Empfindlichkeit um bis zu 10%, was an der Fertigungstoleranz der Messelemente und vor allem an der Ungenauigkeit der Positionierung der Sägeschnitte zur Vereinzelungskante der Chips liegt.

Nach [8, 9] ist das Verhältnis der Schichtdicken von Silizium zu Gegenkörper für die Empfindlichkeit B_0 ein maßgebliches Optimierungskriterium. So kann anhand [8] die Empfindlichkeit weiter gesteigert werden, wenn die Dicke des Substratgegenkörpers auf ca. $250\,\mu m$ bei der vorhandenen Siliziumdicke von $390\,\mu m$ reduziert wird. Ein dickerer Substratgegenkörper hat nach [9] nicht nur eine niedrigere Empfindlichkeit zur Folge, sondern auch höhere systematische Fehler, die mit zunehmendem Druck weiter steigen. Dieses Verhalten wird durch einen nichtlinearen Einfluss des Glases als Gegenkörpermaterial bei hohem Druck vermutet.

Gerade der nichtlineare Einfluss des Glas unter Hochdruck-Belastung und ein mögliches Kriechen machen Untersuchungen zum Langzeitverhalten und der Dynamik des neuartigen Messelementes interessant. Der vorhandene Hochdruck-Messplatz bietet in dieser Hinsicht nur begrenzte Möglichkeiten. Eine systembedingte Undichtigkeit durch die Druckwaage machen lange Messungen ohne dauerhafte Interaktionen eines Benutzers zur Nachführung des Drucks unmöglich. Die Dynamik ist aufgrund der manuell getriebenen Hochdruckspindel ebenfalls stark begrenzt. Durch das Öffnen eines Ventils gegen

Umgebungsdruck fällt der Systemdruck schnell, es kann also ein „Drucksprung“ von Hochdruck zu Umgebungsdruck erzeugt werden.

Bild 8 zeigt eine Messung über einen Zeitraum von ca. 50 Minuten. Der mittels der Druckwaage referenzierte Hochdruck von 2.500 bar liegt für ca. 25 Minuten an. Die Abweichung des Messwerts beträgt in dieser Zeitspanne maximal $\pm 0,1\%$. Da der Messwert statistisch um den Mittelwert schwankt, ist für diesen Zeitraum kein Driften durch Hochdruck-Belastung des Verbundelementes erkennbar.

Durch das Öffnen des Ventils zur Erzeugung des negativen „Drucksprungs“ sollen erste Aussagen über die Dynamik des Messelementes und das Verhalten des Nullpunktes nach einer Druckbelastung geschlossen werden. Bild 8 zeigt, dass der Nullpunkt nach Öffnen des Ventils schnell erreicht ist. Die Nullpunktabweichung beträgt ab ca. 1 Sekunde nach Öffnen des Ventils maximal $\pm 0,1\%$. Bild 9 zeigt das Messsignal im Zeitraum des Ventilöffnens im Detail. Eine Aussage zur Dynamik lässt sich nur schwer ableiten, da derzeit der tatsächliche Druckverlauf im Messsystem nicht bekannt ist.

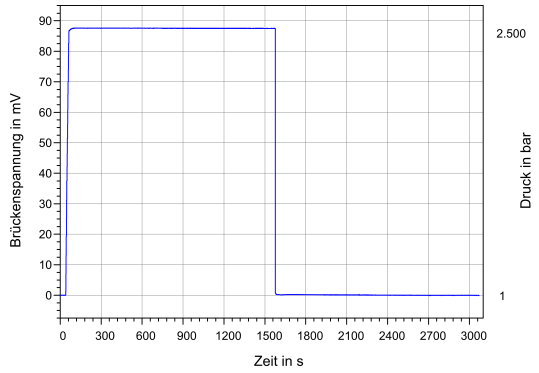


Bild 8: Messung zum Langzeitverhalten und Drucksprung am Verbundelement

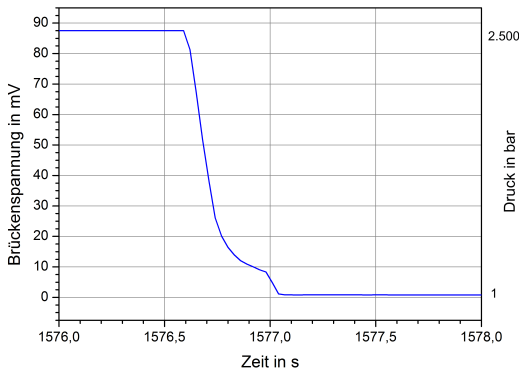


Bild 9: Negativer Drucksprung am Verbundelement; Detail zu Bild 8

Die charakteristischen Messwerte aus Tabelle 1 zeigen das Potential dieser überlastfesten piezoresistiven Messelemente. Ein weiterer Vorteil neben der günstigen Fertigungstechnologie ist vor allem die allseitige Druckeinleitung. Dadurch sind die Verbundelemente nahezu berstfest.

5 Ausblick

Interessant ist der Vergleich der Empfindlichkeit des präsentierten Verbundelementes mit kommerziell erhältlichen Hochdruckelementen des gleichen Nenndruckbereichs von 5.000 bar. Die Empfindlichkeit des Hochdrucksensors mit klassischen DMS ist mit $B_{DMS} = 1\,\frac{\mathrm{mV}}{\mathrm{V}}$ deutlich niedriger gegenüber der Empfindlichkeit des Verbundelementes nach Tabelle 1 mit $B_0 = 13,8\,\frac{\mu\mathrm{V}}{\mathrm{V\,bar}} = 69\,\frac{\mathrm{mV}}{\mathrm{V}}$. Nicht nur, dass durch das höhere Ausgangssignal stabilere Messergebnisse zu erwarten sind, ist der Umkehrschluss dieses Vergleichs noch interessanter. Ein Verbundelement mit der gleichen Empfindlichkeit wie Drucksensoren mit Stahl-Druckmessplatte und DMS $B_0 = B_{DMS}$ führt zu einem minimalen Nenndruck von ca. 72,5 bar bei einer Überlastfestigkeit von über 5.000 bar.

Gerade hinsichtlich der Überlegung des Einsatzes der Verbundelemente auch im Normaldruckbereich bis 100 bar bzw. im unteren Hochdruckbereich ist eine weitere Steigerung der Empfindlichkeit interessant, da die Überlastfestigkeit nicht wie bei piezoresistiven Messelementen nach Stand der Technik abhängig von der Druckplattendicke und somit von der Empfindlichkeit ist. Sie ist eine absolute Größe, die nach aktuellem Stand durch den Berstdruck der elektrischen Durchführung höher als 5.000 bar liegt.

Weitere Messungen an überarbeiteten Silizium-Messelementen mit unterschiedlichen Layoutvarianten werden in naher Zukunft vor allem im erhöhten Druckbereich bis 5.000 bar durchgeführt. Dabei liegt der Fokus neben einer statistisch aussagekräftigen Anzahl an Messungen auch auf der messtechnischen Untersuchung der verschiedenen Entwurfsansätze. Die Untersuchungen zum Langzeitverhalten und vor allem zur Dynamik des Messelementes werden weiter verfolgt.

Das Packaging ist ein weiteres Forschungsziel, das im Bereich der hohen Belastung enorme Anforderungen und somit Herausforderungen birgt.

Literatur

[1] S. Stavroulis und R. Werthschützky, “Neuartiger überlastfester piezoresistiver silizium-hochdrucksensor”, in *tm - Technisches Messen, Vol. 70*, S. 199–205, Ouldenbourg Wissenschaftsverlag, 2003.

[2] S. Stavroulis und R. Werthschützky, “Ein neuartiger piezoresistiver siliziumhochdrucksensor”, in *XVI. Messtechnisches Symposium des Arbeitskreises der Hochschullehrer für Messtechnik.*, W.-J. Becker (Hrsg.), Shaker Verlag, Aachen, 2002.

[3] M.-H. Bao, *Micro mechanical transducers : pressure sensors, accelerometers and gyroscopes*, Vol. 8 of *Handbook of Sensors and Actuators*, ELSEVIER, Amsterdam, 1. ed., 2004.

[4] P. Heinickel und R. Werthschützky, “Novel silicon high pressure sensing element.”, in *Sensor + Test Conference 2009. Proceedings Sensor, Volume I.*, S. 181–186, Nürnberg, 2009.

[5] A. Schroth, *Modelle für Balken und Platten in der Mikromechanik*, Dresdner Beiträge zur Sensorik, Dresden University Press, Dresden, 1 ed., 1996.

[6] R. Ballas, *Piezoelectric Multilayer Beam Bending Actuators*, Springer Verlag, Heidelberg, 2007.

[7] P. Heinickel und R. Werthschützky, “Functionality of a novel overload resistant silicon high pressure sensing element”, in *IEEE Transducers 2009 Conference Proceeding*, S. 252–255, Denver, CO, USA, 2009.

[8] P. Heinickel und R. Werthschützky, “Effect of the thickness - ratio of the piezoresistive composite element for measuring high - pressure”, in *Proc. Eurosensors XXIV*, S. 665–668, Linz, 2010.

[9] P. Heinickel und R. Werthschützky, “High-pressure measurements at 500 mpa of the novel piezoresistive composite element show the influence of glass counter body”, in *Proc. Eurosensors XXV*, Athen, 2011.

[10] DH Instruments, Inc., *PG7000 Piston Gauges - Operation and Maintenance Manual*, Phoenix AZ, USA, 1998.

[11] W. Hummel und E. Altstadt, “Spezial glasdurchführungen und mehr”, in *Sensor Report, 4-2007*, S. 27–28, Verlag Rek & Thomas Medien AG, St. Gallen, 2007.

[12] P. Heinickel, R. G. Ballas, T. Meiss und R. Werthschützky, “Siliziumhochdrucksensor-messelement”, in *8. Dresdner Sensor-Symposium*, S. 333–336, Dresden, 2007.

Online Trace Gas Detection - Pushing a Bruker MicrOTOF II Towards ppq-Detection Limits Using Simple Atmospheric Pressure Ionization Sources

Jens Langejuergen, Philipp Cochems, Stefan Zimmermann

Institute of Electrical Engineering and Measurement Technology,
Leibniz University Hannover
Phone: +49 511-7624680, e-mail: langejuergen@geml.uni-hannover.de

Abstract

In various applications trace gas concentrations in the ppt-range need to be detected in less than a second. Due to an increasing number of contaminants present at ppt-levels ultra-high sensitivity also requires excellent selectivity. Our approach to achieve this combination is to push highly selective mass spectrometers towards extreme sensitivities by coupling atmospheric pressure ionization sources (API), as known from ion mobility spectrometry (IMS), to time-of-flight mass spectrometers (ToF-MS). In this work we present a non-radioactive corona discharge ionization source and a radioactive Tritium (^3H) ionization source both coupled to a Bruker micrOTOF II leading to ppq-detection limits with a mass resolution of 3500 for protonated Acetone (MH^+ = 59.05 au). Using our Tritium ionization source at atmospheric pressure very soft ionization results with just a few fragments leading to quite simple mass spectra. Both, protonated monomers and molecular dimers e.g. of 1-Octanol, DMMP and Acetone can be detected. Furthermore, negative ion formation e.g. for 1,1,2-Trichlorethane has been investigated. In the positive mass spectra of dry clean air mainly protonated water clusters ($(H_2O)_nH^+$) are present. However, the electrical fields inside the ion transfer region (skimmer, lens, hexapole) of the MS can lead to fragmentation that cannot be completely avoided while demanding sufficient ion transfer efficiency. In addition, mass discrimination due to a limited hexapole bandwidth affects the spectrum. Aim of our future work is to develop miniaturized mass spectrometers using API sources for low detection limits and with sufficient mass resolution tuned for trace gas analysis in field applications.

1 Introduction

In applications such as the detection of explosives or highly toxic compounds in air or monitoring biomarkers in exhaled breath, trace gas concentrations in the ppt-range need to be detected in less than a second. Due to the complex matrix of different molecules present at concentrations from ppm to ppt ultra-high sensitivity also requires excellent selectivity to minimize cross sensitivities. One promising approach to manage this analytical challenge is to push highly selective mass spectrometers towards extreme sensitivities by coupling atmospheric pressure ionization sources (API) to mass spectrometers [1-4]. In this work we couple simple API sources, as known from ion mobility spectrometry (IMS) [5], to a commercial time-of-flight mass spectrometer (ToF-MS). As API sources we either use accelerated electrons (several keV) emitted from beta-decay (^{3}H), see right in Fig. 2, electron guns [6], not presented here, as well as corona discharge (Fig. 2, left). Due to molecular collisions at atmospheric pressure the primary build ions (reactant ions) collide with neutrals and, in case of either higher proton or electron affinity, transfer their charge to other molecules to form product ions. This leads to a very soft ionization with almost no fragmentation by reactions of e.g. H^+, H_3O^+, O_2^- or OH^- with the target molecules. A detailed description of this chemical gas phase ionization process can be found in [5]. A major advantage of API is the extreme gain in sensitivity. Most substances like amines, ketones, aromatics, esters and alcohols have higher proton affinity [5] compared to the ions formed in dry clean air resulting in a preferred ionization of these molecules. This, of course, is also one of the major disadvantages in API we still deal with. Since the percentage of ions of the target compound is usually not proportional to the percentage of their neutrals a quantitative measurement especially in complex mixtures can hardly be made. This also includes water which needs to be kept at a constant level during analysis. Thus, the impact of humidity on the ionization of target compounds has to be significantly reduced. Furthermore, most field applications require simple miniaturized MS tuned for specific tasks. Both aspects are part of our future work.

2 Experimental setup

Talking about ppt detection limits, we need a system providing carrier gas (dry clean air) with minimal contaminants, adjustable humidity and trace gas concentrations. Our experimental setup is displayed in Fig. 1. We use a zero-gas-generator including a pressure swing dryer and additional activated carbon

filters to constantly flow through the entire setup and to keep it at moderate excess pressure (5 mbar). Flowing through a permeation oven containing permeation tubes filled with liquid samples of the trace gases to be tested, trace gas concentrations in the ppb to ppm range can be reached depending on the permeation and carrier gas flow rates. By using mass flow controllers (MFC) we can further dilute this concentration by a factor of up to 1000.

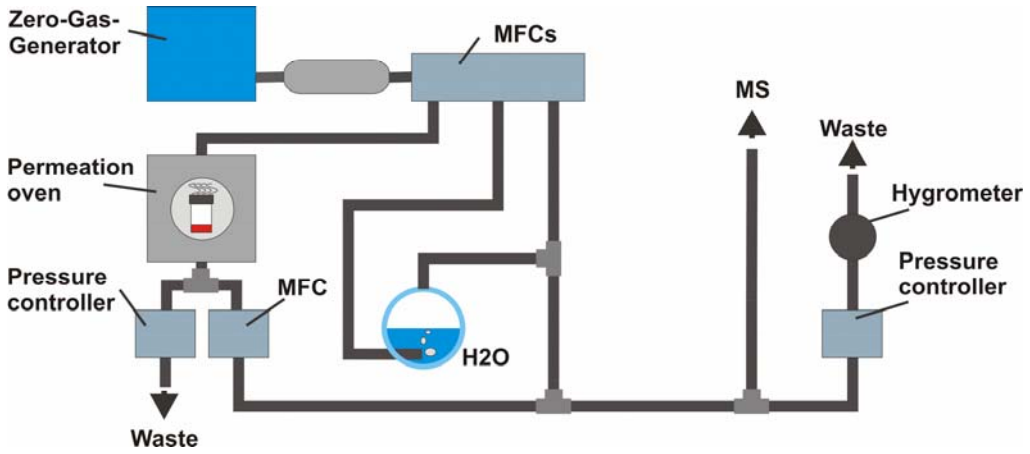


Fig. 1: Experimental setup to provide clean carrier gas with adjustable trace gas and water concentration. The gas flow through different parts of the setup is controlled by mass flow controllers (MFC). The whole setup is at controlled excess pressure of 5 mbar.

The initial concentration of trace gas permeated into the carrier gas can be calculated from the weight loss of the sample in the permeation tube after a sufficient amount of time.

In Fig. 2 the used ionization sources are shown. The Tritium source (left) has an activity of 300 MBq with average kinetic electron energy (beta-decay) of 5.6 keV and a maximum energy of 18.59 keV. We can substitute this radioactive source by a non-radioactive electron-gun [6] with almost identical specifications - not discussed here. Using these ionization sources both ion polarities are generated simultaneously. They can be separated by a weak electrical field (U_{Bias}) to reduce recombination, which would lead to reduced detection limits. Furthermore, a non-radioactive corona discharge ionization source (right) can be coupled to the MS. Here a voltage of some kV (U_{Corona}) is applied to a wolfram needle facing the counter electrode grid. This setup is comparably easy to manufacture but the high electrical field results in an unwanted fragmentation of the ionized molecules. Using this point-to-plane corona discharge ionization source just one ion polarity is present in the gas stream, so that ion-ion recombination is completely avoided. However, field induced ion losses at the

grid and transfer capillary cause significant ion losses as compared to the radioactive Tritium source. Therefore, a new construction is under development providing a sufficient extraction field but minimum electrical fields for keeping the ions within the gas stream.

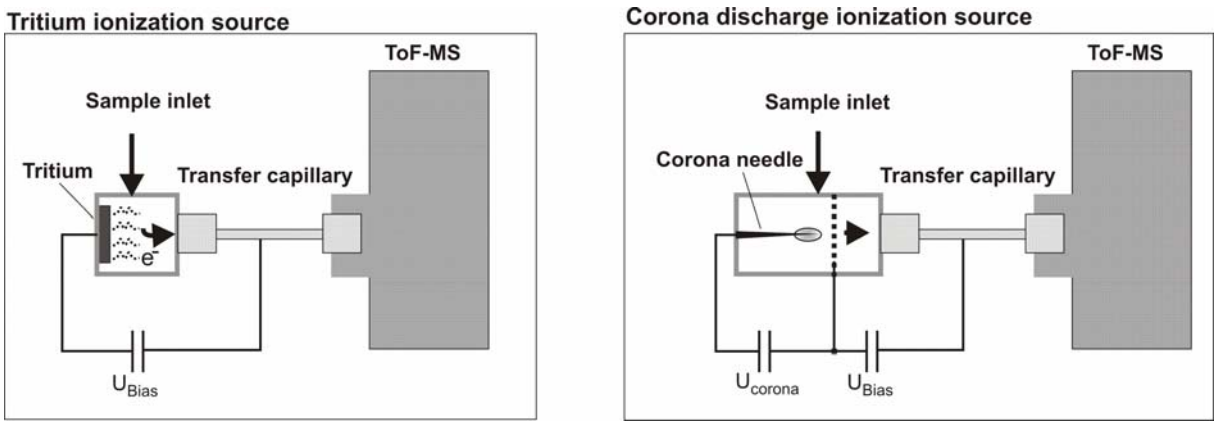


Fig. 2: Setup of the Tritium ionization source (left) and corona discharge ionization source (right). The sample inlet flow rate is 800 ml/min at 1 atm. The ionized sample gas is transferred through a transfer capillary coupled to the Bruker MicrOTOF II.

Due to high gas flow rates of about 800 ml/min a significant number of ions is simply transported from atmospheric pressure into the MS in the gas stream. Inside the MS the ions are further transferred and focused by a variety of electrical fields. Unfortunately, these fields cannot be reduced to zero, as it would be ideal for our application, without significant loss in ion transfer efficiency. This is mainly due to the different target application for this type of MS that significantly differs from our focus. The negative effect of high electrical fields on the mass spectrum is discussed later.

For trace gas detection in field applications miniaturized API-ToF-MS with sufficient mass resolution and high ion transfer efficiency for masses from 20 au to 500 au are needed. The development of such devices is one focus of our work.

2 Experimental results

2.2 Mass spectra of clean air at different MS parameters and humidities

In Fig. 3 the positive (top) and negative (bottom) mass spectra of clean dry air at a dew point of -53.0 °C are shown at different MS operating voltages. Here, our Tritium ionization source was used. The mass range is optimized for masses from 20 au to 160 au. Lighter or heavier ions cannot be transferred into the MS due to the limited hexapole (mass filter) bandwidth.

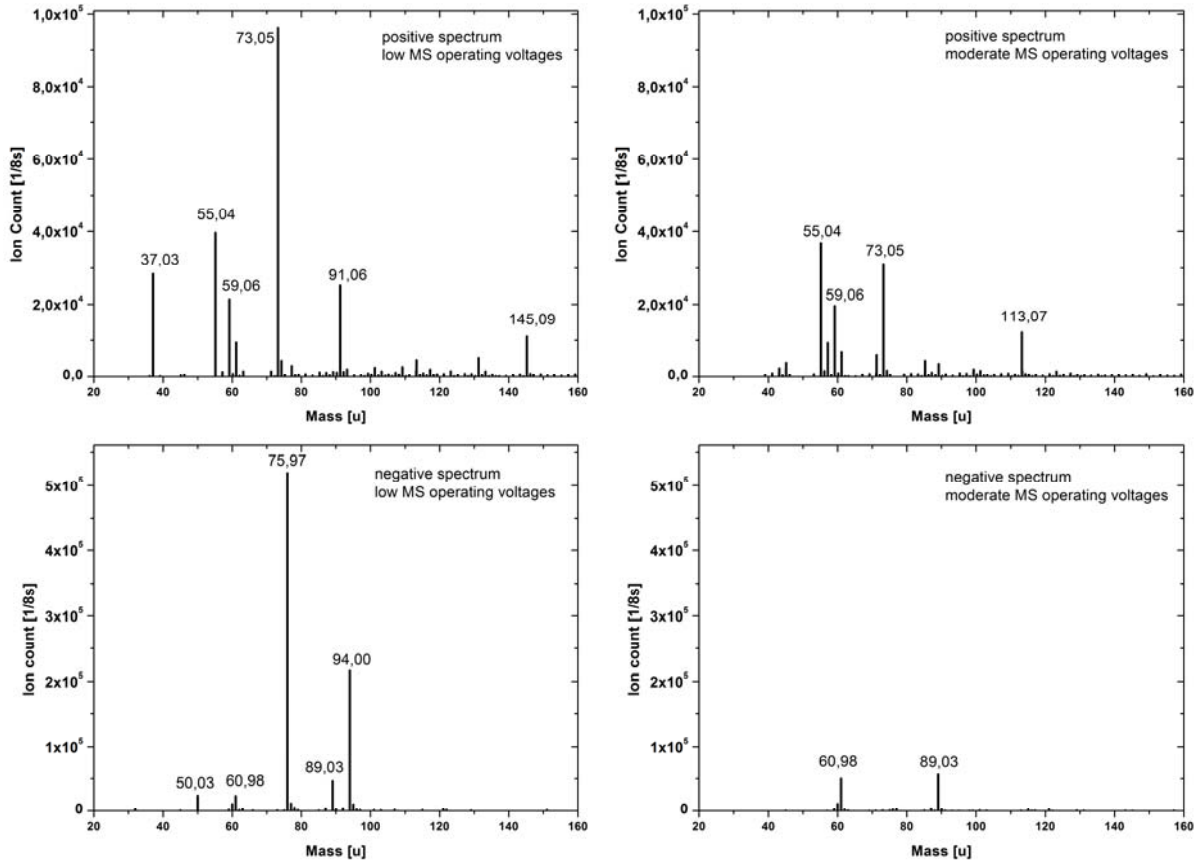


Fig 3: Mass spectra of dry clean air at a dew point if -53.0 °C ionized with our ^3H ionization source. The spectra for positive (top) and negative (bottom) ions are shown for different MS operating voltages.

The positive spectrum at low operation voltages (top left) shows that the detected ions are basically consisting of protonated H_2O clusters $(H_2O)_nH^+$. A single protonated water molecule H_3O^+ (19.02 au) could not be detected for the reason mentioned above. The highest abundance of water clusters is measured

for the $(H_2O)_4H^+$ ions (73.05 au). Even large water clusters such as $(H_2O)_8H^+$ (145.09 au) are clearly present in the spectrum. It is noteworthy to mention that clusters consisting of 8 water molecules have a higher abundance than clusters of 6 or 7 molecules. Since the peak for $(H_2O)_4H^+$ is most prominent in the spectrum it seems like pairs of 4-water clusters are particularly stable at these conditions.

At 59.06 au a small peak of the protonated monomer of Acetone is visible caused by a significant background concentration of Acetone present in all positive spectra of clean dry air. In comparison, the mass spectrum recorded under similar conditions but at higher MS operating voltages is shown (top right). A significant decrease in abundance of all protonated water clusters except $(H_2O)_3H^+$ (55.04 au) can be seen. This leads to the conclusion that this is the ion with the highest binding energy. Furthermore a peak at 113.07 au becomes visible which is most likely the protonated Acetone plus three water molecules attached to it.

At the bottom of Fig. 3 the negative ion mass spectra are displayed for the same MS operating voltages. The effect of fragmentation is clearly evident since the peaks at 50.03 au, 75.97 au and 94.00 au completely disappear at moderate electrical fields while the smaller fragments cannot be detected due to the limited bandwidth of the hexapole. The dominating peaks at 60.98 au, 75.97 au and 94.00 au could not be identified yet. The peaks at 50.03 au and 94.00 au are supposed to be $O_2^-(H_2O)$ and $OH^-(H_2O)_3$ clusters.

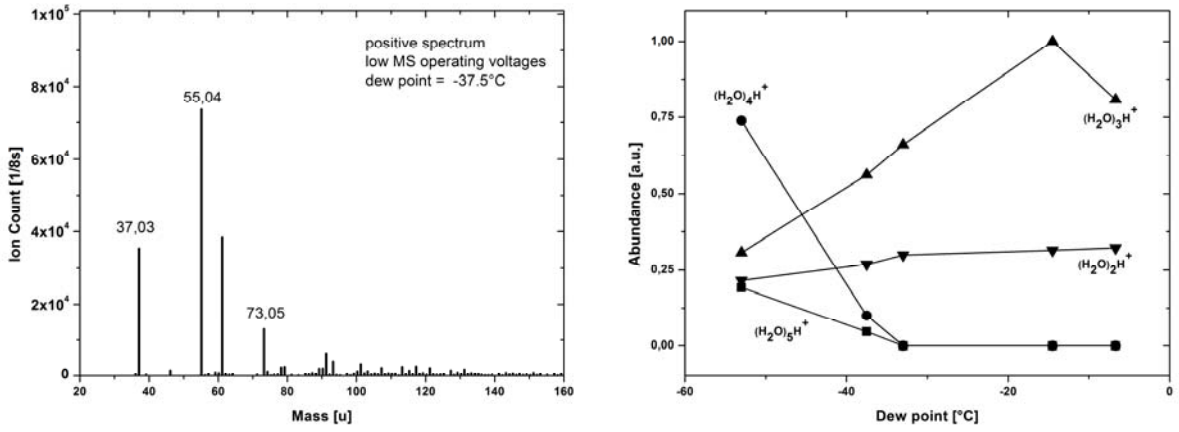


Fig. 4: Dependence of the positive mass spectrum of clean air on humidity using our ^3H ionization source.

In Fig. 4 the dependence of humidity on the positive ion spectrum is shown. Surprisingly, at higher dew points the relative and absolute abundance of smaller

ion clusters increase compared to Fig. 3 (top, left). A reasonable explanation could not be found yet.

2.2 Corona discharge

In Fig. 5 positive mass spectra using our corona discharge ionization source are shown. The spectrum of clean air (left) shows a significantly more complex structure than the spectrum measured with a ^{3}H ionization source. This is due to higher electrical fields and higher ionization density close to the corona needle tip. Nevertheless, protonated water clusters at 73.05 au and protonated Acetone at 59.06 au are clearly evident. For an Acetone concentration of 700 ppt (right) a strong peak significantly larger than all other peaks develops.

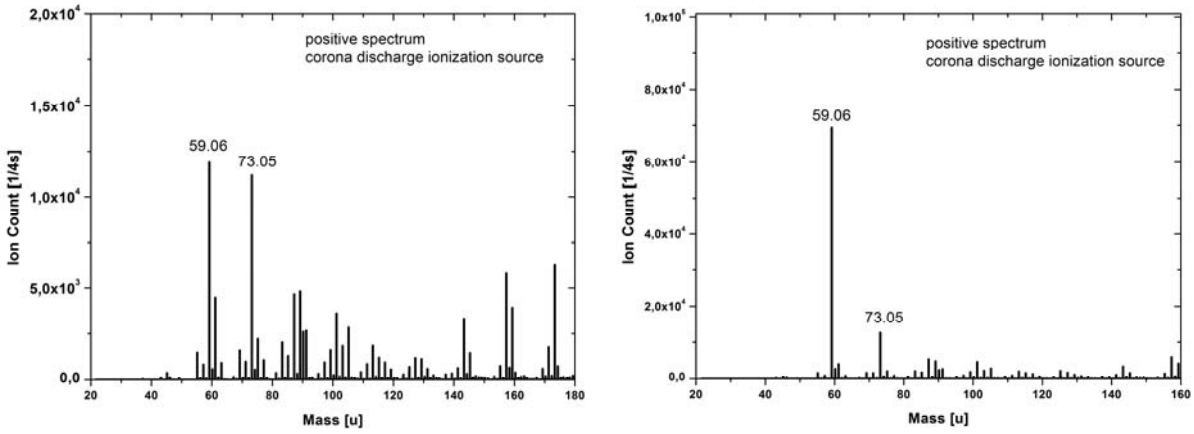


Fig. 5: Positive mass spectrum of clean dry air (left) and 0.7 ppb Acetone (right) using a corona discharge ionization source.

2.3 Acetone detection limit in dry clean air

In Fig. 6 the measured ion count at 59.06 au using a corona ionization source is shown for different Acetone concentrations up to 229 ppt. The signal is integrated for 4 seconds. On the right side of Fig. 6 the corresponding noise of 43 +/- 26 counts at 58.06 au is shown. Using this value we calculate the detection limits for Acetone in clean dry air at a dew point of -53 °C for different ionization sources and different integration times. In Fig. 7 (right) the formation of the protonated Acetone monomer and dimer can be seen at 59.06 au and 117.12 au. Both, the monomer and dimer have no additional water molecules but single H^+ attached. As can be seen from Fig. 7 (left), the sensitivity of the corona discharge ionization source is about five times lower than the sensitivity of the Tritium ionization source. The detection limits are

given in Tab. 1 for Acetone at different integration times. At 500 ms integration time, as required for online breath analysis, the detection limit is about 584 ppq

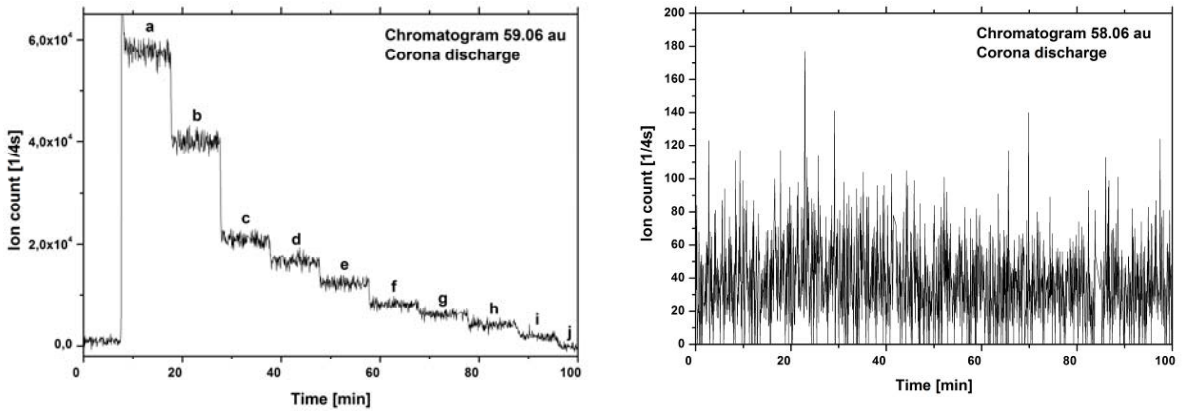


Fig. 6: Chromatogram at 59.06 au for different concentrations of Acetone (concentrations from 229 ppt (a) to 8 ppt (i)) using the corona discharge ionization source. The integration time is set to 4 s. The chromatogram on the left represents the corresponding noise at 58.06 au of 43 +/- 26 counts standard deviation.

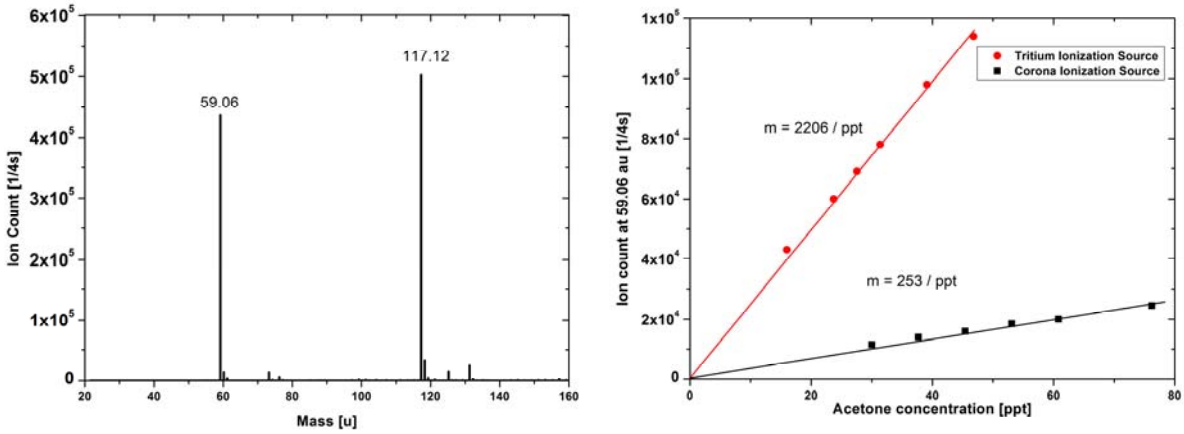


Fig. 7: Spectrum of the protonated Acetone monomer and dimer (left) using our ^{3}H ionization source and sensitivities (right). The slope m of the linear fit is given.

Tab. 1: Calculated detection limits ($3\delta_{Noise}$) for Acetone at different integration times. m is the sensitivity, δ_{Noise} the noise standard deviation.

	4 s	500 ms
Tritium	51 ppq (m=2206 /ppt, δ_{Noise}=38)	584 ppq (m=77/ppt, δ_{Noise}=15)
Corona	308 ppq (m=253/ppt, δ_{Noise}=26)	1083 ppq (m=36/ppt, δ_{Noise}=13)

2.4 DMMP detection limit

In Fig. 8 (left) the mass spectrum of Dimethyl-methylphosphonate (DMMP) at concentrations of 1.5 ppb using a Tritium ionization source is shown. The protonated monomer and dimer of DMMP are clearly visible. Both ions have no additional water molecules attached. In Fig. 8 (right) the ion count for both DMMP monomer and dimer are shown at different concentrations from 0.4 ppb to 1.8 ppb. The detection limit is 4 ppt using the dimer signal since the monomer signal is already saturated at these concentrations. However, when using the monomer signal for detecting DMMP even lower detection limits in the ppq-range are expected.

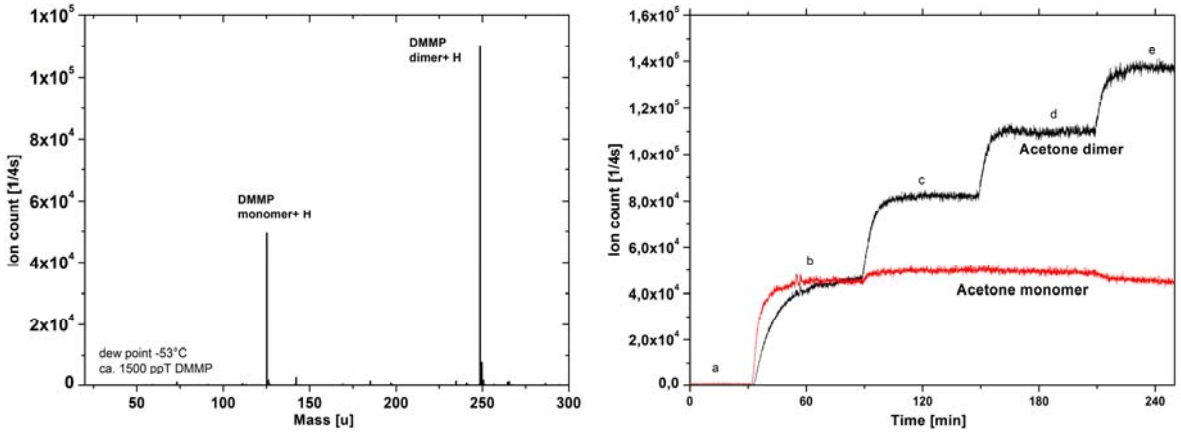


Fig. 8: Mass spectrum of DMMP using a tritium ionization source.

2.5 Mass spectrum of other positive ions

In Fig. 9 the positive mass spectra of 1-Octanol (left) and 2-Hexanone (right) are shown using our ^{3}H ionization source. Even at high 1-Oktanol concentrations no protonated monomer is visible in the spectrum which is expected to be at 131.14 au. Also no 1-Octanol with protonated water clusters attached has been

found. Nevertheless, 1-Octanol can be identified even at lower concentrations because of the very prominent protonated dimer which seems to be very stable at API conditions. Also a protonated Octene can be found in the spectrum formed from 1-Octanol by losing a single water molecule ($C_8H_{18}O \rightarrow C_8H_{16} + H_2O$).

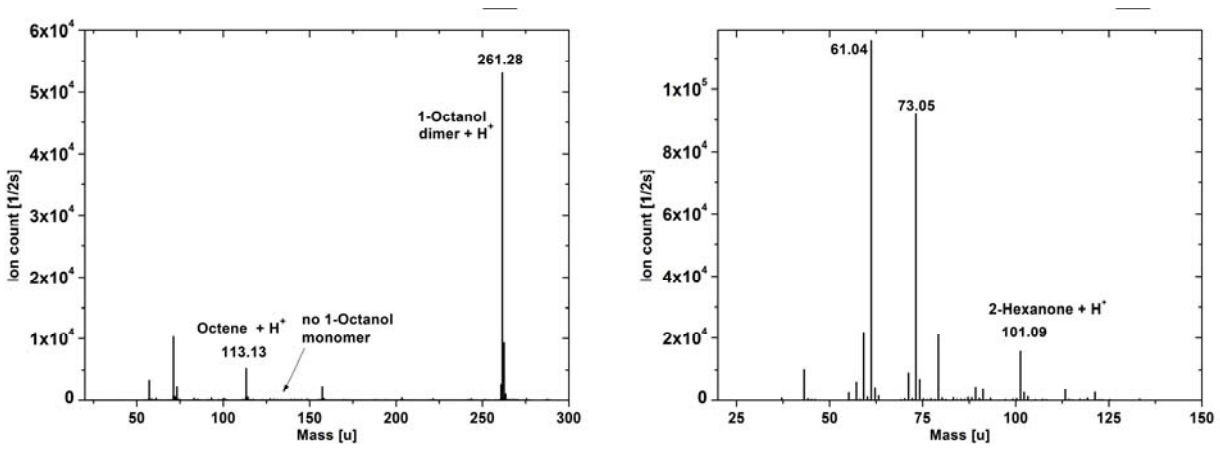


Fig. 9: Positive spectrum of 1-Octanol (left) and 2-Hexanone (right) using our ^3H ionization source.

Surprisingly, a relatively small peak of protonated 2-Hexanone can be found in the spectrum at 101.09 au, see Fig. 9 (left), although the proton affinity of 2-Hexanone is close to that of water [5]. The most dominant peak at 61.04 au seems to be an ionized fragment of this molecule we have not identified yet. The second largest peak is protonated water that can be also found in clean air (see 2.1).

2.5 Mass spectrum of negative ions

In Fig. 10 the negative mass spectra for 1,1,2-Trichlorethane (left) and 1,1,1-Trichlorethan (right) using our ^3H ionization source are shown. Even with minimum electrical fields no molecular peak can be found ($C_2H_3Cl_3$, monoisotopic mass 131.93 au), leading to the conclusion, that these ions are relatively unstable. Nevertheless, fragments and products of Trichlorethane can be easily detected when knowing the isotope pattern of Chlorine. In both spectra monoatomic ionized Chlorine is present. Furthermore, both isomers show significantly different mass spectra and detection limits.. Most peaks in the spectrum of 1,1,1-Trichlorethan could not be indentified yet. Due to a significant lower 1,1,1-Trichlorethane sensitivity we assume that also contaminants are present in the sample gas as a consequence of minimum sample dilution.

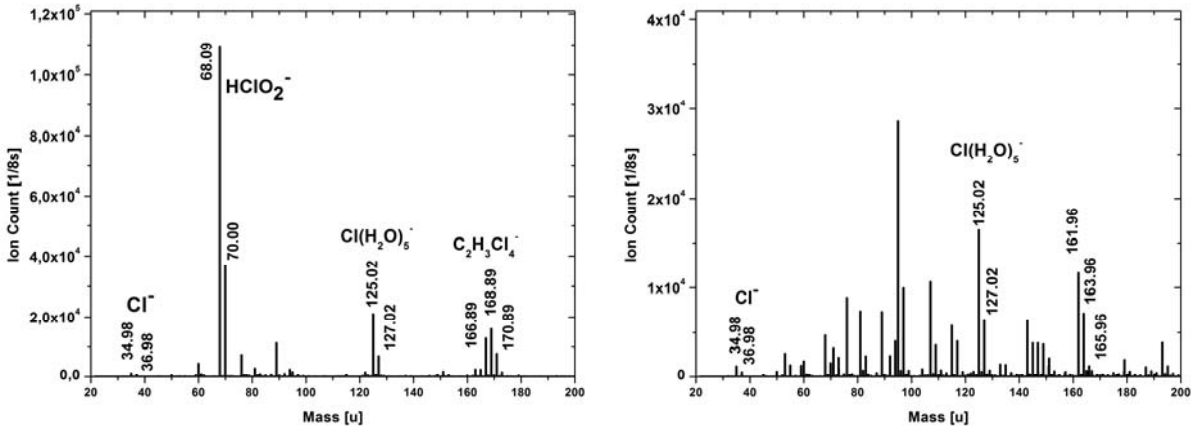


Fig 10: Negative spectrum of 1,1,2-trichlorethane (left) and 1,1,1-Trichlorethane using a ^{3}H ionization source. Carrier gas dew point is -53.0°C.

In the spectrum of 1,1,2-Trichlorethane we identified Chlorous Acid ($HClO_2$) as the most dominant peak. Also negatively charged Chlorine with 5 water molecules attached and Tetrachlorethane ($C_2H_2Cl_4$), which is build by substitution of H with Cl^- in Trichlorethane, is clearly visible. This leads to the conclusion that a relatively large amount of Cl^- is available during the ionization process. The spectrum of 1,1,1-Trichlorethane is much more complex and only a few peaks, such as free and hydrated Cl^- could be indentified yet.

3 Conclusion

Due to the soft but efficient ionization process at atmospheric pressure ultra-low detection limits in the ppq-range and simple mass spectra with almost no fragments can be obtained. However, one major disadvantage of API is that concurring ionization processes lead to an abundance of detected ions not necessarily representing the concentration of neutrals. Especially an increase of humidity in the sample gas leads to a significant change in the composition of the detected ions. To ensure quantitative measurements and minimize the influence of sample gas humidity the amount of primary ions has to be increased while lowering the operation pressure from 1 atm to some mbar. As ionization sources either radioactive Tritium, non-radioactive electron guns or corona discharge needles can be used. For Tritium the detection limits are lower compared to our corona discharge ionization source. The electrical fields within the ion transfer region of the MS lead to some fragmentation and mass discrimination that could not be completely avoided while demanding sufficient ion transfer efficiency. Also the dimensions of the used MS are not suited for

any field application. This is the motivation to develop small miniaturized mass spectrometers with API and sufficient mass resolution for trace gas analysis but minimal operating voltages within the ion transfer region.

Literature

[1] D.Smith, P.Španěl, "Selected Ion Flow Tube Mass Spectrometry, SIFT-MS for On-Line Trace Gas Analysis of Breath", World Scientific 2005

[2] A.Hansel, A.Jordan, R.Holzinger, P.Prazeller, W.Vogel, W.Lindinger, "Proton transfer reaction mass spectrometry: on-line trace gas analysis at ppb level", International Journal of Mass Spectrometry and Ion Processes, 149/150 (1995), 609-619

[3] L.M.Matz, G.R.Asbury and H.H.Hill, "Two-dimensional separations with electrospray ionization ambient pressure high-resolution ion mobility spectrometry/quadrupole mass spectrometry", Rapid Communications in Mass Spectrometry, 2002

[4] W.Lindinger, J.Hirber, H.Paretzke, "An ion/molecule-reaction mass spectrometer used for on-line trace gas analysis", International Journal of Mass Spectrometry and Ion Processes, 129 (1993), 79-88

[5] G.A. Eiceman, Z.Karpas, "Ion Mobility Spectrometry", CRC Press, (2005)

[6] W. Bäther, S. Zimmermann, F. Gunzer, "Investigation of the influence of voltage parameters on decay times in an ion mobility spectrometer with a pulsed non-radioactive electron source", Int. J. Ion Mobil. Spec. (2010) 13:95-101

Präzisionsbestimmung der Boltzmann-Konstante an der Physikalisch-Technischen Bundesanstalt

Thorsten Zandt, Christof Gaiser, Bernd Fellmuth

Physikalisch-Technische Bundesanstalt, Abbestr. 2-12, 10587 Berlin
Tel.: 030-3481-7796, e-mail: thorsten.zandt@ptb.de

Kurzfassung

Im Rahmen eines internationalen Projekts für die Neudefinition der Basiseinheit Kelvin hat sich die Physikalisch-Technische Bundesanstalt die Aufgabe gestellt, die Boltzmann-Konstante *k* mithilfe der Dielektrizitätskonstanten-Gasthermometrie mit einer Unsicherheit von etwa 2 ppm zu bestimmen. Dafür wurde ein neues Experiment aufgebaut, bestehend aus einem großvolumigen Thermostaten, einem vakuumisolierten Messsystem, zwei 10 pF Zylinderkondensatoren aus Edelstahl, einer Spartransformator-Kapazitätsmessbrücke, einem hochreinen Gas-Handling-System einschließlich eines Massenspektrometers und einem speziellen Kolbenmanometer mit einer Querschnittsfläche des Kolben-Zylindersystems von 2 cm². Mit diesem Dielektrizitätskonstanten-Gasthermometer konnte die Boltzmann-Konstante am Wassertripelpunkt mit einer Standardmessunsicherheit von 9,2 ppm bestimmt werden. Der Aufbau des Messsystems, die bisher vorliegenden Ergebnisse und weitere Schritte auf dem Weg zu einer Bestimmung der Boltzmann-Konstante am Wassertripelpunkt mit einer Unsicherheit von 2 ppm sollen in diesem Artikel beschrieben und diskutiert werden.

1 Einleitung

Seit 1954 ist die Temperatureinheit Kelvin als der 273,16te Teil der thermodynamischen Temperatur des Tripelpunktes von Wasser (TPW) definiert [1]. Diese Definition scheint auf den ersten Blick eindeutig und unveränderlich zu sein. Tatsächlich beinhaltet sie aber eine Maßverkörperung, denn die realisierte Tripelpunkttemperatur hängt von den Eigenschaften der benutzten Wasserprobe ab. Weil unterschiedliche Effekte die Tripelpunkttemperatur beeinflussen, ist

damit die Langzeitstabilität der Basiseinheit Temperatur nicht gesichert. Problematisch ist insbesondere die Abhängigkeit der Tripelpunkttemperatur von der Isotopenzusammensetzung, den Verunreinigungen und dem Druck in der die Wasserprobe enthaltenden Fixpunktzelle.

Die fehlende Absicherung der Langzeitstabilität der Einheit Kelvin stellt ein ernstes Problem dar, ist doch die Temperatur die in Industrie und Wissenschaft meistgemessene physikalische Größe. Deshalb hat das Comité International des Poids et Mesures (CIPM) 2005 den metrologischen Staatsinstituten empfohlen, für das Kelvin, wie auch für das Kilogramm, das Ampere und das Mol, geeignete Vorschläge vorzulegen, damit diese vier Basiseinheiten gemeinsam neu definiert werden können [2]. Ziel ist es, die Basiseinheiten künftig über Naturkonstanten zu definieren und sie damit unabhängig von Maßverkörperungen zu machen.

Praktisch bedeutet das, dass die Temperatur auf die Energieeinheit, das heißt, auf die mechanischen Einheiten zurückzuführen ist. Weil aber der Konversionsfaktor zwischen Temperatur und Energie, die Boltzmann-Konstante k, bisher nicht mit einer Unsicherheit bestimmt werden kann, die der Unsicherheit der Darstellung des Wassertripelpunktes gleichkommt ($u_{\mathrm{rel}} = 3 \cdot 10^{-7}$), stellt die Präzisionsbestimmung der Boltzmann-Konstante am Wassertripelpunkt die zentrale Aufgabe bei der Neudefinition der Basiseinheit Kelvin dar. Für diese Aufgabe führt man Primärthermometrie am Wassertripelpunkt durch. Als Primärthermometer bezeichnet man Thermometer, für die der Zusammenhang zwischen ihrer Anzeige und der thermodynamischen Temperatur mit genügender Genauigkeit theoretisch beschrieben werden kann. Zu den möglichen primärthermometrischen Verfahren [3] gehören z. B. das Akustische Gasthermometer (AGT), bei dem die Schallgeschwindigkeit des idealen Gases gemessen wird [4], das Brechungsindex-Gasthermometer (RIGT), bei dem der Brechungsindex in einem Gas gemessen wird [31], das Rauschthermometer (JNT), welches die Nyquist-Formel für die Beschreibung des thermischen Rauschens ausnutzt [6], das Doppler-Verbreiterungsthermometer (DBT), bei dem die Doppler-Verbreiterung in einem optischen Absorptionsspektrum gemessen wird [7] und das Dielektrizitätskonstanten-Gasthermometer (DCGT) [8, 9]. Letzteres wird an der Physikalisch-Technischen Bundesanstalt in der Arbeitsgruppe 7.43 „Grundlagen der Thermometrie“ betrieben und soll in diesem Artikel genauer beschrieben werden.

Der derzeitig international akzeptierte Wert [10] von $1{,}3806504 \cdot 10^{-23}$ J/K für die Boltzmann-Konstante besitzt eine relative Standardmessunsicherheit von $1{,}7 \cdot 10^{-6}$ und wurde mittels AGT ermittelt [4], wobei sich k gemäß der Beziehung $k = R / N_A$ ergibt (R molare Gaskonstante, N_A Avogadro-Konstante). Die für eine Festlegung des Wertes der Boltzmann-Konstante notwendige Genauigkeit ist damit noch nicht erreicht, denn es darf durch diese Festlegung keinen wesentlichen Rückschritt im Vergleich zur Unsicherheit der Darstellung des Wassertripelpunktes geben. Es ist daher wichtig, weitere Experimente zur k-Bestimmung durchzuführen, wobei zur Aufdeckung systematischer Fehlerquellen unterschiedliche Verfahren einzusetzen sind.

2 Das Prinzip des Dielektrizitätskonstanten-Gasthermometers

Die Dielektrizitätskonstanten-Gasthermometrie basiert auf der Messung der relativen Dielektrizitätskonstanten (DK) ε_r und des Druckes p eines Gases [3, 8, 9]. Primäre Thermometrie erfordert dabei die Kenntnis der Polarisierbarkeit des Gases. Der Zusammenhang zwischen der DK und dem Druck kann mit ausreichender Genauigkeit theoretisch nur für ein ideales Gas vorhergesagt werden. Im Experiment hat man es aber immer mit einem realen Gas zu tun, das heißt, man muss zur Eliminierung der Realgaseigenschaften Isothermen in Abhängigkeit vom Druck messen und dann auf den Druck Null extrapolieren, um die Eigenschaften des idealen Gases mit genügender Genauigkeit zu approximieren. Die aus der Statistik folgenden grundlegenden Relationen hierfür sind [9]:

$$(\varepsilon_r - 1) / \varepsilon_r = 3A_\varepsilon (1 + b^*/V_M + c^*/V_M^2 + \ldots) / V_M, \qquad (1)$$

$$p = RT (1 + B^*/V_M + C^*/V_M^2 + \ldots) / V_M \qquad (2)$$

mit $B^* = B - 3A_\varepsilon b^* D^2 / 2RT\varepsilon_0$ und $C^* = C - 3A_\varepsilon c^* D^2 / 2RT\varepsilon_0$. In den Beziehungen sind T die Temperatur, V_M das molare Volumen, B und C der zweite und dritte Dichtevirialkoeffizient, b^* und c^* der zweite und dritte dielektrische Virialkoeffizient, A_ε die molare Polarisierbarkeit, ε_0 die Dielektrizitätskonstante des Vakuums und D die dielektrische Verschiebung. Mit (1) erhält man die Virialentwicklung der Clausius-Mossotti-Gleichung:

$$(\varepsilon_r - 1) / (\varepsilon_r + 2) = A_\varepsilon \ (1 + b/V_M + c/V_M^2 + \ldots) / V_M \qquad (3)$$

mit $b = b^* + 2A_\varepsilon$ und $c = c^* + 4A_\varepsilon b^* + 4A_\varepsilon^2$.

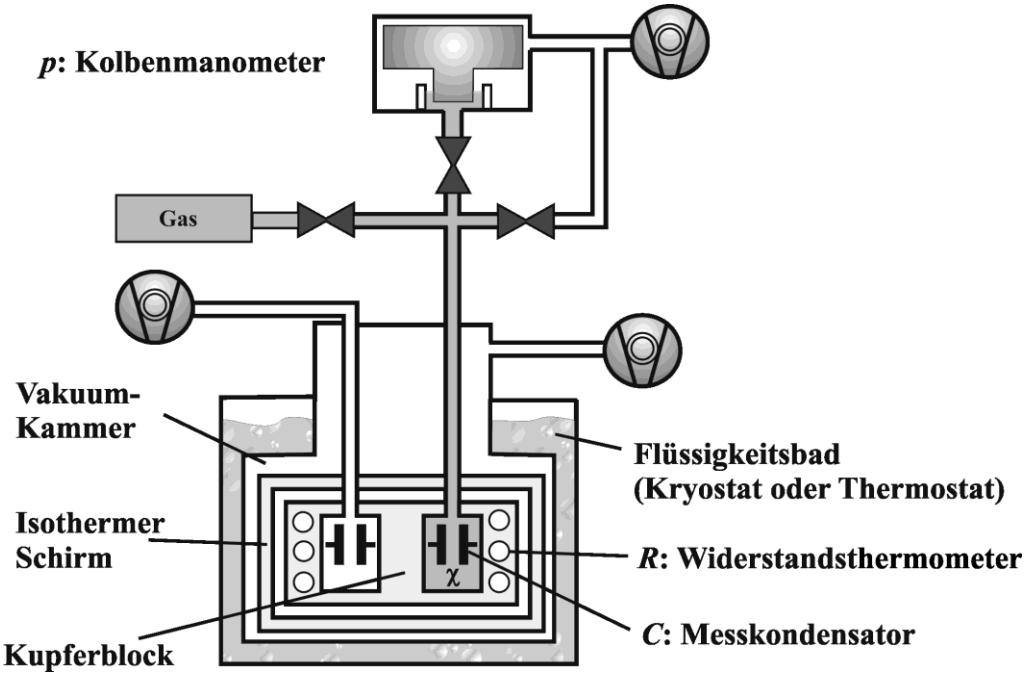


Bild 1: Schematische Darstellung des Aufbaus des DCGT-Experiments, wie es in der PTB verwendet wird (Referenz-Kondensator auf der linken Seite, Messkondensator auf der rechten Seite). Drei Größen sind zu messen: der Druck p mithilfe des Kolbenmanometers, die relative Kapazitätsänderung $\Delta C(p)/C(0)$ und über das Widerstandsthermometer die Temperatur T.

Die Kombination der Beziehungen (2) und (3) eliminiert das Molvolumen. Die Bestimmung der DK erfolgt über die Messung der relativen Kapazitätsänderung eines Kondensators durch das Messgas (vgl. Abbildung 1), gemäß der Beziehung

$$\varepsilon_r = C(p) / C(0) / (1 + \kappa_{eff} \cdot p), \qquad (4)$$

in der $C(p)$ und $C(0)$ die Kapazitätswerte des Messkondensators mit und ohne das Messgas und κ_{eff} seine effektive Kompressibilität sind. Die Beziehungen (2), (3) und (4) liefern den notwendigen Zusammenhang zwischen dem Druck und der dielektrischen Suszeptibilität $\chi = \varepsilon_r - 1$, der näherungsweise durch folgende Formel für die Isotherme gegeben ist:

$$p \approx \frac{\chi}{\frac{3A_\varepsilon}{RT} + \kappa_{eff}} \left[1 + \frac{B(T)}{3A_\varepsilon}\chi + \frac{C(T)}{(3A_\varepsilon)^2}\chi^2 + \ldots \right] \qquad (5)$$

mit

$$\chi \approx (\mathrm{C(p)} - \mathrm{C(0)}) / \mathrm{C(0)}. \qquad (6)$$

Für die Extrapolation auf $p = 0$ wird die Isotherme in einem bestimmten Bereich ausgemessen und Formel (5) wird an die erhaltenen Wertepaare $p_i(\chi_i)$ mittels Ausgleichsrechnung angepasst. Da die Messungen zur Bestimmung der Boltzmann-Konstante am Wassertripelpunkt durchgeführt werden, das heißt, die Temperatur bekannt ist ($T = 273{,}16$ K), kann dann der Quotient $R / 3A_\varepsilon$ aus dem Anstieg bei $p = 0$ bestimmt werden, der die Boltzmann-Konstante k liefert:

$$k = \frac{R}{3A_\varepsilon} \frac{\alpha_0}{\varepsilon_0}, \qquad (7)$$

wobei α_0 die statische Dipolpolarisierbarkeit eines Gasteilchens ist.

Primäre Thermometrie mit dem Dielektrizitätskonstanten-Gasthermometer erfordert, dass die Polarisierbarkeit α_0 des Messgases genügend genau theoretisch berechnet werden kann. Hier hat es in den letzten Jahren durch die Weiterentwicklung der theoretischen Methoden und der Rechentechnik ganz wesentliche Fortschritte bei der Ab-initio-Berechnung der Eigenschaften von Helium gegeben [11, 12], sodass für die Bestimmung der Boltzmann-Konstante mit dem DCGT die experimentellen Unsicherheitskomponenten des DCGTs dominierend sein werden.

Im Rahmen des iMERA Joint Research Project "Boltzmann-Constant" (JRP Nr. T1.J1.4), wurde an der PTB ein neues DCGT-Experiment aufgebaut. Die Entscheidung für ein Dielektrizitätskonstanten-Gasthermometer wurde im Hinblick auf die ausgezeichneten experimentellen Ergebnisse, die im Tieftemperaturbereich [9, 13-16] erhalten wurden, getroffen.

3. Aufbau des Experimentes

In Tabelle 1 sind konzeptionelle Abschätzungen für die wichtigsten Unsicherheitskomponenten vom Typ B bei der Messung der Boltzmann-Konstante am Wassertripelpunkt mittels der Dielektrizitätskonstanten-Gasthermometrie aufgelistet.

Tabelle 1: Konzeptionelle Unsicherheitsbudget für die Bestimmung der Boltzmann-Konstante mit dem DCGT (nur Komponenten vom Typ B)

Komponenten	**Hochrechnung Tieftemperaturexp.**	**Ziel**
$\chi(p = 0)$	3 ppm	1 ppm
Druck p	5 ppm	1 ppm
Kompressibilität κ_{eff}	13 ppm	0.5 ppm
Polarisierbarkeit α_0	2 ppm	0.5 ppm
Verunreinigungen	5 ppm	0.5 ppm
Adsorption	0.5 ppm	0.2 ppm
Temperatur T	2 ppm	1 ppm
Type B kombiniert	**15 ppm**	**2 ppm**

Der Hochrechnung bezieht sich auf die vor dem Start des Boltzmann-Projekts der PTB vorhandene Instrumentierung, einen Gasthermometerkryostaten mit einem Arbeitsbereich von 1,5 K bis 300 K (Tieftemperaturexperiment).

Zur notwendigen Verringerung der Unsicherheit kommt es vor allem darauf an, den Einfluss der effektiven Kompressibilität des Kondensators κ_{eff} zu senken. Im neu aufgebauten Experiment werden Zylinderkondensatoren mit beidseitigen Masseschirmen verwendet [17]. Bild 2 zeigt einen solchen Messkondensator. Dieser Kondensator bietet mehrere verschiedene Betriebsmodi. So kann der Kondensator z. B. ohne Masseschirme betrieben werden, wodurch κ_{eff} durch

die gesamte Längenänderung der Elektroden durch den Druck des Messgases bestimmt wird. Die in-situ Messung der Verformung des äußeren Gehäuses beim Zylinderkondensator mit beidseitigen Masseschirmen mittels eines Planspiegel-Interferometers gestattet dagegen, eine wesentlich genauere Korrektur anzubringen, sodass nur noch die Längenänderung der kurzen Masseschirme durch den Druckeinfluss übrig bleibt, was den Einfluss der Unsicherheit von κ_{eff} weiter senkt. Eine dritte Möglichkeit des Betriebs ist die Verwendung eines äußeren Druckbehälters, der die Verformung des äußeren Gehäuses beim Zylinderkondensator reduziert.

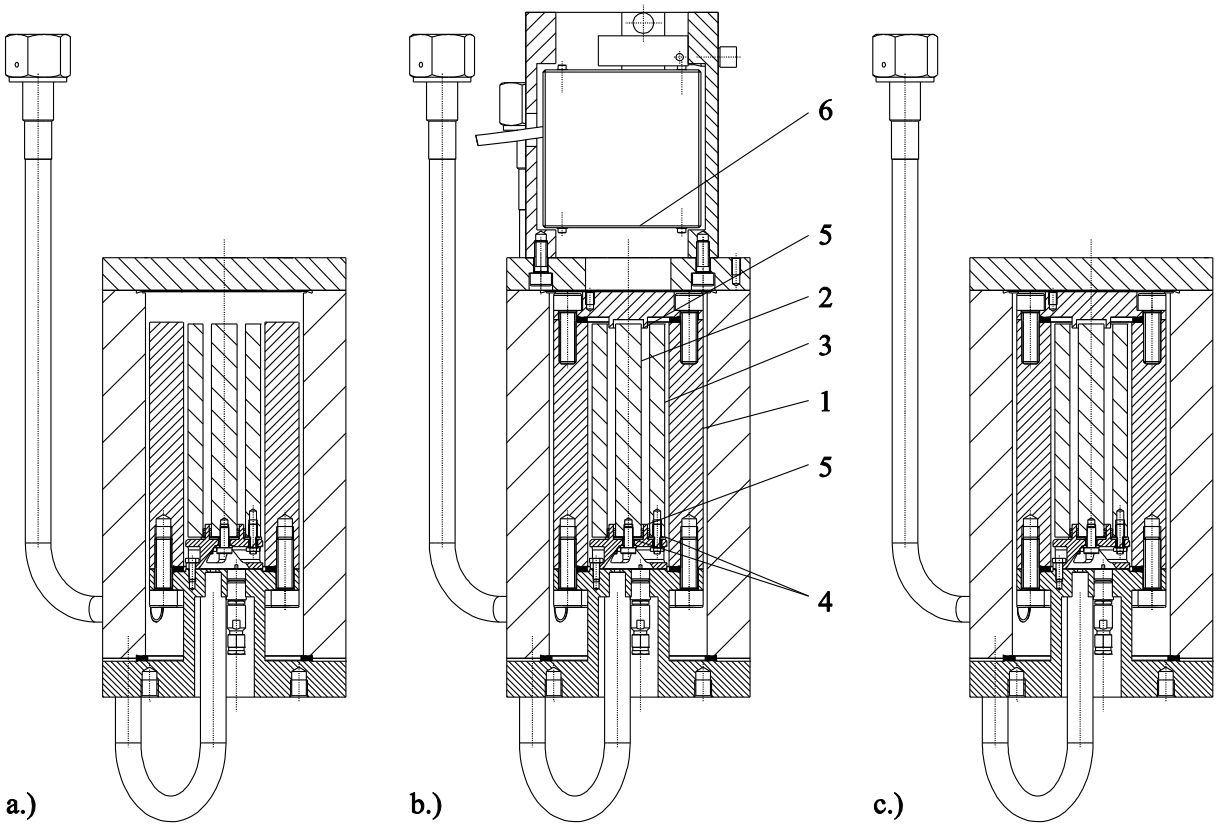


Bild 2: Die verschiedenen Betriebsmodi der neuen Zylinderkondensatoren: a.) Kondensator ohne oberen Masseschirm, b.) Messung der Verbiegung mit Interferometer, c.) Kondensator mit zusätzlichem Druckbehälter und oberem Masseschirm, 1 Druckbehälter (mit Masse verbunden), 2 innere Elektrode (Außendurchmesser 12 mm, Höhe 100 mm), 3 äußere Elektrode (Innendurchmesser 20 mm, Höhe 100 mm), 4 elektrische Isolierung; 5 unterer Masseschirm, 6 Halterung für einen Planspiegel-Interferometer.

In allen drei möglichen Betriebsmodi wird die Kompressibilität des verwendeten Kondensator-Materials (zurzeit wird der Edelstahl 1.4122 / X39CrMo17-1 eingesetzt) benötigt, welche mittels Resonanz-Ultraschallspektroskopie [18] bestimmt wird. Weitere Korrekturen werden mithilfe der Finite-Elemente-Methode durchgeführt [19].

Um die erzielten Ergebnisse mit den Zylinderkondensatoren abzusichern, wird als weiterer Kondensatortyp ein ringförmiger Kreuzkondensator mit 25 Ebenen verwendet [17]. Auch andere Kondensatormaterialien, wie z. B. Wolframkarbid, welches eine um den Faktor 2 bis 3 kleinere Kompressibilität hat, werden in Zukunft zum Einsatz kommen.

Die Bestimmung der Suszeptibilität χ des Messgases stellt extreme Anforderungen an die Messung von Kapazitätsänderungen. So hat Helium bei $p = 1$ MPa und $T = 273$ K eine Suszeptibilität von nur ca. 0.001, das heißt, eine relative Unsicherheit in der Bestimmung von χ und damit des Verhältnisses R/A_ε von 1×10^{-6} erfordert, die Kapazitätsänderung durch das Gas mit einer relativen Unsicherheit von 1×10^{-9} zu messen. Auch hier konnte der notwendige Fortschritt nur durch eine wesentliche Änderung und Optimierung der experimentellen Parameter im Vergleich zu den Tieftemperaturexperimenten erreicht werden.

Zum einen war der Druck zu erhöhen: Es werden Drücke bis 7 MPa verwendet. Um die notwendige Extrapolation auf $p = 0$ genügend genau vornehmen zu können, ist dabei die Isotherme hinunter bis zu Drücken auszumessen, die ca. 1/3 des Maximaldruckes betragen. Unter diesen Bedingungen hängt die Gesamt-Standardmessunsicherheit der Bestimmung der Boltzmann-Konstante vom Maximaldruck ab. Bei Maximaldrücken unter etwa 3 MPa steigt die Unsicherheit des linearen Terms in Gleichung (5) durch die Grenzen der Kapazitätsmessung stark an. Oberhalb dieses Wertes tragen dagegen unterschiedliche Komponenten gleichberechtigt zur Gesamtunsicherheit bei. Zum anderen kann die durch die Kapazitätsmessung bedingte Unsicherheitskomponente reduziert werden, indem Messgase eingesetzt werden, die eine höhere Polarisierbarkeit besitzen. Hier kommen vor allem Neon ($A_\varepsilon(\mathrm{Ne})/A_\varepsilon(\mathrm{He}) \approx 2$) und Argon ($A_\varepsilon(\mathrm{Ar})/A_\varepsilon(\mathrm{He}) \approx 8$) infrage. Da nur für Helium eine genügend genaue Ab-initio-Berechnung der Polarisierbarkeit möglich ist, muss allerdings die Polarisierbarkeit anderer Gase an der von Helium kalibriert werden. Hierzu sind mit diesen Gasen und mit Helium Isothermenmessungen bei der gleichen Temperatur im Tieftemperaturbereich durchzuführen [16].

Zur Messung der Kapazitätsänderungen durch das Messgas wurde eine spezielle Spartransformator-Kapazitätsmessbrücke entwickelt. Ihr Hauptbestandteil ist ein von der PTB entwickelter hochpräziser induktiver 1:1 Spannungsteiler, der eine Spartransformator-Konfiguration verwendet [20]. In [20] konnte gezeigt werden, dass Kapazitätsänderungen von wenigen 0,1 % mit einer relativen Standardmessunsicherheit unter 1 ppm gemessen werden können, was eine relative Standardmessunsicherheit von einigen ppb bezogen auf die Kapazität bedeutet.

Für die Messung von Drücken bis zu 7 MPa mit einer relativen Standardmessunsicherheit von 1 ppm werden Kolbenmanometer eingesetzt, deren Kalibrierung auf höchstem metrologischen Niveau in der Arbeitsgruppe 3.13 „Druck“ der PTB erfolgte. Dieses Ziel war eine große Herausforderung, weil es Kolbenmanometer mit bisher unerreichter Qualität erforderte und sogar das nationale Normal der PTB deutlich verbesserte. Für absolute Druckmessungen in Helium bis zu einem Druck von 7 MPa wurden deshalb wie in [21, 22] beschrieben spezielle Kolbenmanometer konstruiert, gebaut und getestet [23]. Das System umfasst zwei Kolbenmanometer-Plattformen, drei Kolben-Zylinder-Einheiten (PCUs) mit Querschnittsflächen von 20 cm² und drei 2 cm² PCUs. Jede Plattform ist mit einer automatisierten Massenauflage und einem Massensatz mit der Gesamtmasse von 150 kg ausgestattet. Die 11 Hauptmassen von jeweils 12,5 kg wurden bezüglich ihrer mechanischen und magnetischen Eigenschaften in Übereinstimmung mit den Anforderungen der Empfehlung OIML R 111-1, Massen der Klasse E1, hergestellt [24]. Das Design der PCUs und deren Aufbau wurde bezüglich der Reduzierung des Druckverformungskoeffizienten und den durch die Montage induzierten Verformungen optimiert. Die Spaltbreite der PCUs wurde angepasst, um einen Kompromiss zwischen Fallgeschwindigkeit und Empfindlichkeit zu erreichen.

Die Rückführbarkeit der Druckmessungen auf die SI-Basiseinheiten bis zu 7 MPa wurde in zwei Schritten realisiert. Zuerst wurde die wirksame Querschnittsfläche der 20 cm² PCUs für den Druck $p = 0$ aus dimensionellen Messungen bestimmt. In einem zweiten Schritt wurden dann die 2 cm² PCUs durch Vergleichsmessungen kalibriert. Der Grund für dieses Vorgehen liegt darin begründet, dass die dimensionellen Messungen auf dem erforderlichen Niveau von weniger als 1 ppm nur für die 20 cm² PCUs möglich sind. Für die Messung von Durchmesser, Geradheit und Rundheit konnte die erforderliche relative Unsicherheit von 1 ppm für die 20 cm² PCUs aber nur durch den Einsatz verbesserter dimensionellen Messtechnik [25, 26, 27] erreicht werden. Dabei konnten zum ersten Mal Unsicherheiten für die dreidimensionalen Daten der Kolben von 8 nm und für die Zylinder von 16 nm erreicht werden.

Die Berechnung der effektiven Flächen und die Abschätzung ihrer Unsicherheiten erforderten mehrere Versuche, welche ausführlich in [22] diskutiert werden. Die Berechnungen ergaben eine relative Unsicherheit der effektiven Querschnittsflächen der drei 20 cm² PCUs zwischen 0,5 ppm und 0,7 ppm. Die effektive Querschnittsfläche, für die in den DCGT-Messungen verwendete 2 cm²

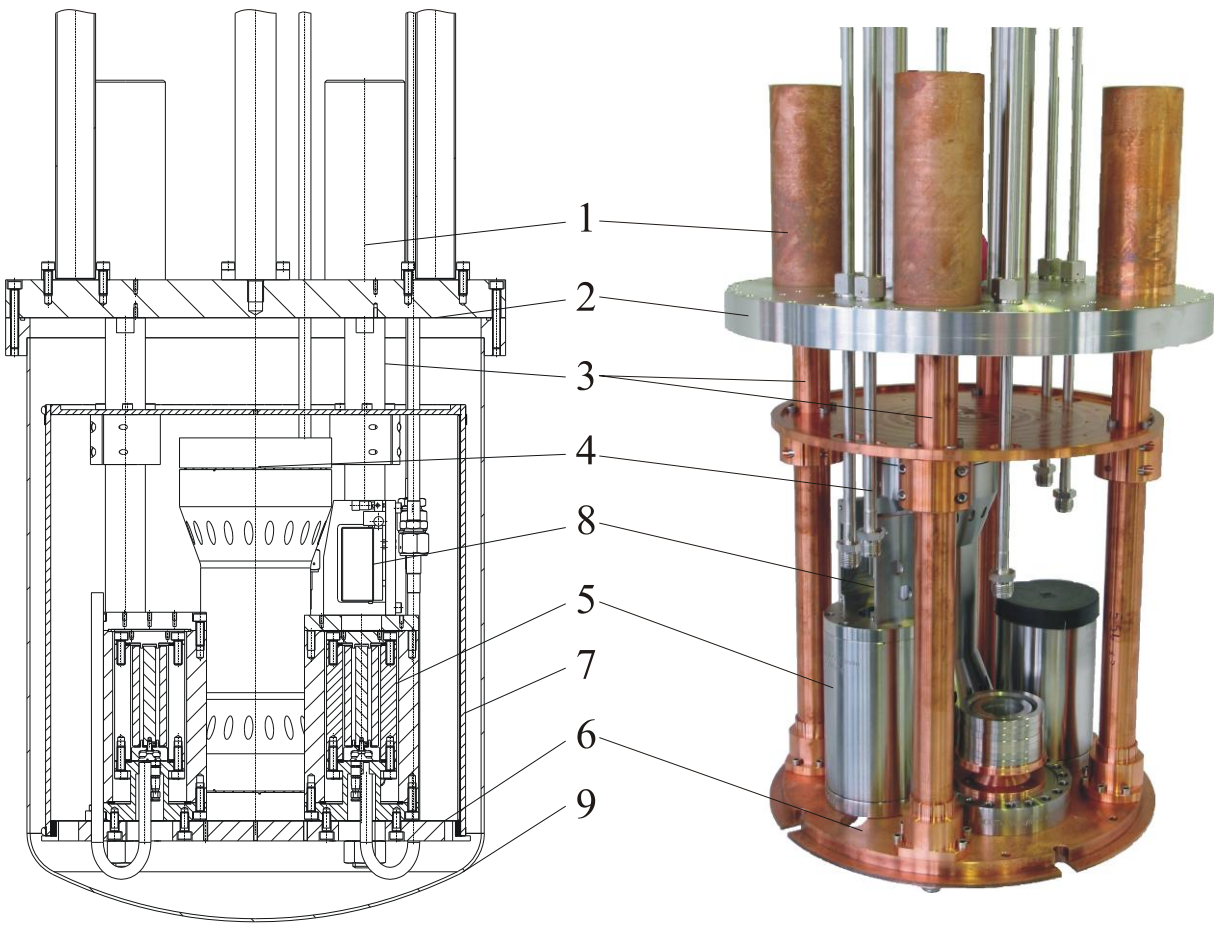


Bild 3: Konstruktionszeichnung und Foto des DCGT-Messsystems: 1 Zylinder aus Kupfer für die Verbesserung des thermischen Kontakts zum Bad, 2 oberer Flansch des Vakuumsystems, 3 Kupferstangen, 4 Druckbehälter für 25-Ebenen-Kreuzkondensator, 5 Druckbehälter für einen Zylinder-Kondensator; 6 zentrale Kupferplatte; 7 isotherme Abschirmung, 8 Halterung für das Planspiegel-Interferometer, 9 Vakuumkammer.

PCU, wurde durch Vergleiche mit den drei 20 cm² PCUs mit einer Unsicherheit von 1,5 ppm bestimmt.

Wie in der Einleitung beschrieben, müssen die Experimente für die Neubestimmung der Boltzmann-Konstante an der Temperatur des Wassertripelpunktes durchgeführt werden. Um diese Messungen nachvollziehbar auf die Definition der Basiseinheit Kelvin zurückzuführen, muss die Temperatur des Gases und damit der Kondensatorelektroden mit einer Unsicherheit von 0,1 mK gemessen werden. Zur Realisierung einer thermischen Umgebung von ausreichender Qualität wurde deshalb eine Dreiebenenanordnung in dem neuen DCGT-Versuchsaufbau realisiert [17, 28, 29].

Jeder Kondensator ist von einem Druckbehälter umgeben, der thermisch mit einer 2 cm dicken zentralen Kupferplatte des Messsystems verbunden ist (siehe Bild 2). Die Temperatur der zentralen Kupferplatte wird mit Hilfe von drei Präzisionsplatinwiderstandsthermometern (Kapseltyp) gemessen. Neben der Kupferplatte an der Unterseite des Messsystems besteht das Messsystem weiterhin noch aus einer oberen Platte, vier dicken Kupferstangen zwischen den beiden Platten und einer isothermen Abschirmung. Dieses Messsystem befindet sich zur thermischen Isolierung in einer Vakuumkammer, welche in einem Flüssigkeits-Thermostaten eingesetzt ist.

Der Thermostat hat ein Gesamtvolumen von ca. 800 l. Das zentrale Arbeitsvolumen, in dem sich die Vakuumkammer befindet, hat einen Durchmesser von 500 mm und eine Höhe von 650 mm. Die Temperaturstabilität und die Temperaturhomogenität an der Grenze des Arbeitsvolumens ohne und mit der Kammer wurden sorgfältig unter verschiedenen experimentellen Bedingungen untersucht. Es konnte gezeigt werden, dass unter optimalen Bedingungen, sowohl die Instabilität als auch die Inhomogenität der Temperatur im Arbeitsvolumen deutlich unter 1 mK sind [17, 28]. Weitere Experimente [17, 28] konnten zeigen, dass die Kupferplatte, auf der sich die Kondensatoren befinden, mit einer Temperaturinstabilität kleiner als 0,1 mK geregelt wird.

3. Erste Ergebnisse und Ausblick

Insgesamt elf Isothermen mit Helium und vier mit Neon konnten bereits mit dem neuen DCGT-Versuchsaufbau realisiert werden [19]. Dabei wurde die erste Variante der oben beschriebenen Zylinderkondensatoren verwendet (kein oberer Masseschirm). Die Isothermen bestehen aus bis zu vierzehn Tripel, die die Temperatur T, den Druck p und χ (Ergebnis der Kapazitätsmessung) enthalten. Die Gesamtzahl der Tripel beläuft sich auf 121 für Helium und 56 für Neon. Aus einer Ausgleichsrechnung (Gleichung (5)) wurde der lineare Koeffizient und mit dessen Hilfe dann nach Gleichung (7) die Boltzmann-Konstante bestimmt: $k_{\mathrm{TPW}} = 1.380654 \cdot 10^{-23}$ J/K. Eine detaillierte Analyse der Unsicherheitsquellen, einschliesslich von Monte-Carlo-Simulationen, ergaben eine Schätzung für die relative Standardmessunsicherheit für den k_{TPW}-Wert von 9,2 ppm.

Mit dem Tieftemperaturexperiment wurden im Temperaturbereich von 21 K bis 27 K weitere Messungen für die Bestimmung der Boltzmann-Konstante durch-

geführt. Die Auswertung dieser Daten ergaben für die Boltzmann-Konstante einen Wert von $k_{LT} = 1.380657 \cdot 10^{-23}$ J/K mit einer Standardmessunsicherheit von 15,9 ppm [19, 30]. Der gewichtete Mittelwert der Werte k_{LT} und k_{TPW} beträgt $k = 1.380655 \cdot 10^{-23}$ J/K. Dieser Wert hat eine relative Standardmessunsicherheit von 7,9 ppm und unterscheidet sich vom derzeitig international akzeptierten Wert von CODATA [10] von $1{,}3806504 \cdot 10^{-23}$ J/K um 3,2 ppm.

Mit Blick auf die Unsicherheiten der anderen möglichen Methoden der Primärthermometrie [AGT, RIGT, JNT, DBT], kann festgestellt werden, dass derzeit das DCGT die einzige Methode ist, mit der die Boltzmann-Konstante ohne Anwendung von akustischen und/oder Mikrowellen-Resonatoren, mit einer Unsicherheit unter 10 ppm bestimmt werden kann (siehe Bild 4). Solche Resonatoren werden für die akustische Gasthermometrie [4] und die Brechungsindex-

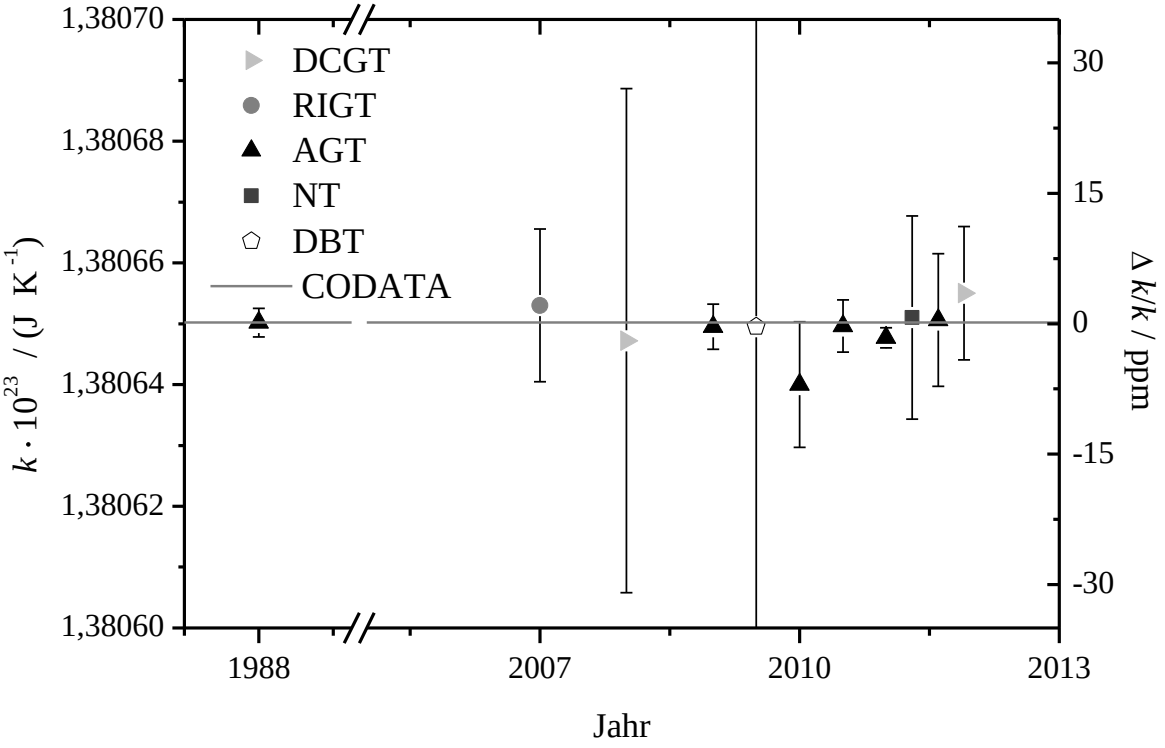


Bild 4: Übersicht über die Bestimmung des Boltzmann-Konstante seit 1988. Die Balken beschreiben die Standardmessunsicherheiten (DBT 38 ppm). Die durchgezogene Linie zeigt den numerischen Wert für die Boltzmann-Konstante von CODATA [10], der derzeitig international akzeptierten wird.

Gasthermometrie [5] verwendet. Diese Tatsache ist von besonderer Bedeutung, da die Dielektrizitätskonstanten-Gasthermometrie von den anderen Methoden unabhängig ist, weil sie ganz andere Unsicherheitsbeiträge hat.

Bei der hier gezeigten ersten Bestimmung der Boltzmann-Konstante mit dem neuen DCGT-Messaufbau am TPW tragen im Wesentlichen die Instabilität und die effektive Kompressibilität des Zylinderkondensators zur Unsicherheit der bestimmten Boltzmann-Konstante bei. Fortschritte sind hier durch die schon oben erwähnten anderen Kondensatorendesigns und alternative Elektrodenmaterialien zu erwarten. Erste Stabilitätsuntersuchungen mit ringförmigen Kreuzkondensatoren verliefen erfolgreich. Ein Einsatz dieser Kondensatoren würde den Unsicherheitsbeitrag der effektiven Kompressibilität verringern, weil die Berechnung der effektiven Kompressibilität hier genauer möglich ist.

Weiterhin wird zurzeit daran gearbeitet, die Unsicherheit der Druckmessung auf ein Niveau von 1 ppm und den Einfluss von Verunreinigungen im Messgas zu verringern. Angesicht der Erfahrungen während der ersten DCGT-Experimente am TPW scheint es realistisch zu sein, die relative Unsicherheit der Boltzmann-Konstante innerhalb der nächsten zwei Jahre auf ein Niveau von etwa 2 ppm zu senken.

Literatur

[1] H. Preston-Thomas, Metrologia 1990, 27, 3. Bezugsquelle: Bureau International des Poids et Mesures (BIPM), Pavillon de Breteuil, 92310 Sèvres Cedex, France, oder http://www.bipm.org/en/publications/its-90.html.

[2] Recommendation 1 (CI-2005) of the CIPM, www.bipm.org/cc/CIPM /Allowed/94/CIPM-Recom1CI-2005-EN.pdf)

[3] Fellmuth, B. u. A.: Determination of the Boltzmann constant—status and prospects, Meas. Sci. Technol. **17**, R145-R159, 2006

[4] Moldover, M. R. u. A.: Measurement of the universal gas constant R using a spherical acoustic resonator, Phys. Rev. Lett. **60**, 249-252, 1988

[5] Schmidt, J. W.: Polarizability of Helium and Gas Metrology, PRL **98** 254504-4, 2007

[6] Benz, S. P. u. A.: An electronic measurement of the Boltzmann constant Metrologia, **48**, 142-153, 2011

[7] Lemarchand, C u. A.: Determination of the Boltzmann Constant by Laser Spectroscopy as a Basis for Future Measurements of the Thermodynamic Temperature, Int. J. Thermophys., **31**, 1347-1359, 2010

[8] Gugan, D. und Michel, G. W.: Dielectric Constant Gas Thermometry from 4.2 to 27.1 K, Metrologia **16**, 149, 1980

[9] Luther, H. u. A.: Determination of thermodynamic temperature and 4He virial coefficients between 4.2 K and 27.0 K by dielectric constant gas thermometry, Metrologia **33**, 341-352, 1996

[10] Mohr, P. J. et al.: CODATA recommended values of the fundamental physical constants: 2006, Rev. Mod. Phys. **80**, 633, 2008.

[11] Łach G. u .A.: Radiative Corrections to the Polarizability of Helium, PRL **92** 233001-4, 2004

[12] Puchalski M. u. A.: Blackbody-radiation correction to the polarizability of helium, Phys. Rev. A **83**, 042508-7, 2011

[13] Fellmuth, B. u. A.: The international Boltzmann project – the contribution of the PTB, C. R. Physique **10** 828-834, 2009

[14] Gaiser, C. und Fellmuth, B.: Helium virial coefficients—a comparison between new highly accurate theoretical and experimental data, Metrologia **46** 525 533, 2009

[15] Gaiser, C. u. A.: Dielectric-Constant Gas-Thermometry Scale from 2.5 K to 36 K Applying 3He, 4He, and Neon in Different Temperature Ranges, Int. J. Thermophys. **31** 1428-1437, 2010

[16] Gaiser, C. und Fellmuth, B.: Experimental benchmark value for the molar polarizability of neon, Europhysics Letters **90** 63002-p1-63002-p5, 2010

[17] Zandt, T. u. A.: Dielectric-Constant Gas-Thermometry Measuring System for the Determination of the Boltzmann Constant at PTB, Int. J. Thermophys. **31**, 1371-1385, 2010

[18] Maynard, J. D.: The use of piezoelectric film and ultrasound resonance to determine the complete elastic tensor in one measurement, J. Acoust. Soc. Am. **91** 1754-1762, 1992

[19] Fellmuth, B. u. A.: Determination of the Boltzmann constant by dielectric-constant gas thermometry, Metrologia, angenommen, 2011

[20] Fellmuth, B. u. A.: High-Precision Capacitance Bridge for Dielectric-Constant Gas Thermometry, IEEE Trans. Instr. Meas**. 60**, 2522-6, 2011

[21] Sabuga, W.: Towards 7 MPa Pressure Standards with $1 \cdot 10^{-6}$ Uncertainty, MAPAN – J. Metrology Soc. India **22**, 3-11, 2007

[22] Sabuga, W.: Pressure measurements in gas media up to 7.5 MPa for the Boltzmann constant redetermination, PTB-Mitteilungen, eingereicht, 2011

[23] Sabuga W u. A.: 2011'Design and Evaluation of Pressure Balances with $1 \cdot 10^{-6}$ Uncertainty for the Boltzmann Constant Project, PTB-Mitteilungen, eingereicht, 2011

[24] OIML R 111 1, Weights of classes E1, E2, F1, F2, M1, M1–2, M2, M2–3 and M3 Part 1: Metrological and technical requirements, OIML, 2004

[25] Jusko, O. u. A.: Dimensional calibration techniques for pressure for pressure balances balances to be used in the new determination of the Boltzmann constant, CIMMEC – 1st International Congress on Mechanical Metrology, 2008

[26] Jusko, O. u. A.: Form Measurements by Optical and Tactile Scanning, Proc. ASPE Spring Topical Meeting Mechanical Metrology and Measurement Uncertainty, Albuquerque, NM, USA, p. 11-14 ISBN 978 1 887706 49 0, 2009

[27] Jusko, O. u. A.: Final Results of the Geometrical Calibration of the Pressure Balances to be Used for the new Determination of the Boltzmann Constant, Key Engineering Materials **437**, 150, 2010

[28] Zandt, T. u. A.: Dielectric-Constant Gas-Thermometry Measuring System for the Determination of the Boltzmann Constant at PTB, Int. J. Thermophys. DOI: 10.1007/s10765-011-0980-4, 2011

[29] Merlone A u. A.: Construction and Start-up of a Large-Volume Thermostat for Dielectric-Constant Gas Thermometry, Int. J. Thermophys. **31,** 1386-1395, 2010

[30] Gaiser, C. und Fellmuth, B., Metrologia, in Bearbeitung, 2011

Rückgeführte Größenmessung von Nanopartikeln im Transmissionsmodus eines Rasterelektronenmikroskops (TSEM)

T Klein, E Buhr, K-P Johnsen, C G Frase

Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig
E-mail: Tobias.Klein@ptb.de

Kurzfassung

Es wird ein neues elektronenmikroskopisches Verfahren zur rückgeführten Größenmessung von Nanopartikeln beschrieben, bei dem ein Transmissionsdetektor in einem Rasterelektronenmikroskop genutzt wird (TSEM, Transmission Scanning Electron Microscopy). Im Zuge der Bildauswertung werden mit Hilfe eines Simulations-Modells der TSEM-Abbildung der Rand und die Größe jedes einzelnen Partikels individuell bestimmt. Die Simulation basiert auf grundlegenden Streutheorien für Elektronen unter Beachtung der Eigenschaften des Partikels und des Mikroskops. Die Partikelgrößenverteilung wird durch die automatisierte Auswertung einer Serie von TSEM-Aufnahmen bestimmt, die insgesamt Tausende Partikel zeigt. Nanopartikel-Proben aus drei verschiedenen Materialklassen (Gold, Silika und Latex) im Größenbereich zwischen 5 nm und 60 nm werden exemplarisch untersucht. Eine Unsicherheitsbetrachtung zeigt erweiterte Unsicherheiten für die mittlere Partikelgröße im Bereich zwischen 1 nm und 3 nm.

1 Einleitung

Die größenabhängigen Eigenschaften von Nanopartikeln und die damit verbundenen neuen physikalischen Effekte können für verbesserte Produkte genutzt werden. Auf der anderen Seite bedingt die zunehmende Anwendung von Nanopartikeln in Alltagsprodukten Fragen und Sorgen bezüglich Gesundheit und Umwelt. Hochgenaue Größenmessungen von Nanopartikeln sind eine Voraussetzung, um diesen Fragen zu begegnen und Ursache-Wirk-Zusammenhänge zu

untersuchen. Gleichzeitig hilft eine rückführbare Messtechnik, die einzigartigen Eigenschaften von Nanopartikeln für technologische Anwendungen zu nutzen.

Im Hinblick auf rückführbare Messungen von Nanoobjekten scheint die Nutzung eines Transmissionsdetektors in einem Rasterelektronenmikroskop vielversprechend. Mit dieser Technik wurden bereits Gitterebenen abgebildet und rückgeführte Messungen an Membran-Masken durchgeführt [1,2]. Weiterhin konnte die Anwendbarkeit der Technik für Messungen von Nanopartikeln gezeigt werden [3]. Im Vergleich zur Abbildung mittels Sekundärelektronen, die üblicherweise im SEM angewandt wird, haben Transmissonsmessungen eine Reihe von Vorteilen wie z.B. verbesserte Auflösung und geringere Empfindlichkeit gegenüber elektrostatischen Probenaufladungen. Des Weiteren ist eine genaue Simulation der Bildentstehung von TSEM-Aufnahmen weniger komplex als vergleichbare Simulationen von Sekendärelektronen-Abbildungen [4]. Diese Vorteile können ausgenutzt werden, um die Bestimmung der Partikel-Ränder und damit die Größenmessung zu verbessern.

2 Durchführung

2.1 Elektronenmikroskop

Der experimentelle Aufbau ist ausführlich in [3] beschrieben. Partikelproben werden in einem herkömmlichen SEM (Zeiss Supra 35 VP) untersucht, das mit einem als Zubehör erhältlichen Transmissionsdetektor ausgestattet ist. Der Detektor besteht aus fünf Halbleiterdetektoren, von denen vier als Dunkelfelddetektoren genutzt werden. Der fünfte Detektor ist unterhalb einer kleinen Apertur positioniert und dient als Hellfelddetektor. Er registriert Elektronen, die ohne relevante Ablenkung durch die Probe hindurchtreten. Um den Messvorgang zu beschleunigen und den Einfluss des Nutzers zu minimieren, wurde eine Routine zur automatischen Bildaufnahme entwickelt, die auf der Abarbeitung einer Liste von nicht überlappenden Aufnahmefeldern auf der Probe basiert.

2.2 Kalibrierung der Pixelgröße

Um die Pixelgröße zu bestimmen, wird das SEM mit Hilfe eines 2D-Gitters kalibriert, das aus kleinen, zylinderförmigen Aluminium-Strukturen auf Silizium besteht, die einen nominellen Abstand von 144 nm haben (150-2D von Advanced Surface Microscopy Inc.). Der tatsächliche Abstand wird in unserer Arbeitsgruppe mittels Laserbeugung im tiefen UV bestimmt, die rückgeführte Werte für den mittleren Abstand liefert [5,6].

Während der Kalibrierung werden dieselben Einstellungen des SEM gewählt, die später auch während der Messungen genutzt werden. Da mittels Elektronenmikroskopie individuelle Abstände zwischen benachbarten Strukturen bestimmt werden, muss eine große Zahl von Abständen über das gesamte Gitter gemessen werden, um mit hoher Genauigkeit Aussagen über den mittleren Abstand treffen zu können. Als Beispiel zeigt Bild 1 die Verteilung der gemessenen Abstände für 25000-fache Vergrößerung. Die Verteilung der Abstände lässt sich größtenteils auf die Variation der Positionen der einzelnen Aluminium-Strukturen auf dem Substrat zurückführen. Messungen wurden an verschiedenen Stellen des Substrates ausgeführt, um die genügende Homogenität des mittleren Abstands über das gesamte Gitter nachzuweisen.

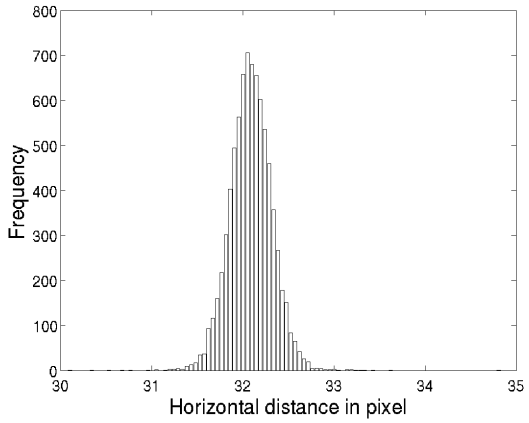


Bild 1: Verteilung der gemessenen horizontalen Abstände (in Pixeln) zwischen benachbarten Aluminium-Strukturen des Kalibrierstandards, hier gezeigt für 25000-fache Vergrößerung.

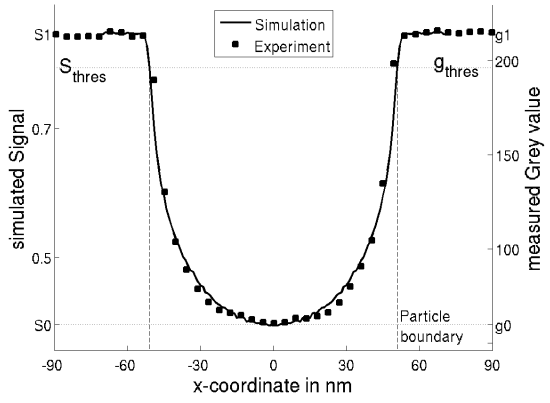


Bild 2: Simuliertes TSEM-Signalprofil über eine Latexkugel zusammen mit dem entsprechenden gemessenen Grauwert-Profil.

2.3 Probenpräparation

Es wurden Referenzpartikel aus Gold (NIST RM8011, nomineller Durchmesser 10 nm), Silika-Partikel für die Qualitätskontrolle (IRMM 304) und Latex-Kugeln (Duke Scientific, nomineller Durchmesser 50 nm) untersucht. Die Partikel liegen in Suspension vor. Sie werden auf herkömmlichen TEM-Netzchen präpariert, die aus einem dünnen Kohlenstofffilm bestehen, der von einem Metall-Netzchen gehalten wird. Die Probenpräparation wird unter Reinraumbedingungen durchgeführt. Die TEM-Netzchen werden in die Mulden einer Teflon-Platte gelegt, bevor jeweils ein Tropfen unverdünnter Suspension mit Hilfe einer Spritze aufgebracht wird. Die Proben werden für eine gewisse Zeit (Minuten bis Stunden) in einer gesättigten Wasserdampf-Atomsphäre aufbewahrt, die ein Austrocknen verhindert. Anschließend werden die Tröpfchen mit Reinraumpapier komplett aufgesaugt. Bis zu zwölf TEM-Netzchen finden auf einem Vielfachhalter Platz und können in Transmission untersucht werden.

3 Simulation und Bestimmung des Schwellwerts

Bei der Auswertung von TSEM-Aufnahmen von Nanopartikeln ist eine präzise Bestimmung des Partikelrands für hochgenaue Größenmessungen essentiell. Eine bequeme und oft genutzte Methode zur Bestimmung des Partikelrands in elektronenmikroskopischen Aufnahmen ist die Anwendung eines globalen Schwellwerts. Unterschiedliche Algorithmen zur Bestimmung des Schwellwerts stehen zur Verfügung und können genutzt werden, um Partikel vom Hintergrund zu trennen [7]. Allerdings weichen die resultierenden Größenverteilungen voneinander ab, wenn verschiedene Algorithmen angewandt werden [8]. Herkömmliche Schwellwertverfahren nutzen lediglich die Grauwertverteilung des zu analysierenden Bildes und ignorieren die zugrundeliegenden physikalischen Mechanismen der Bildentstehung. Die Bestimmung des Partikelrands und somit auch die Größenmessung kann verbessert werden, wenn der physikalische Prozess der Bildentstehung in die Bestimmung des Schwellwerts mit einbezogen wird. Die Bildentstehung im Rasterelektronenmikroskop wird üblicherweise mittels Monte Carlo Simulationen modelliert [9].

In dieser Arbeit wird das vielseitige Programmpaket „MCSEM“ zur Simulation genutzt (Monte Carlo Simulation of Scanning Electron Microscopy) [10]. Als

Beispiel zeigt Bild 2 das simulierte TSEM-Signalprofil für eine Latexkugel auf einem Kohlenstofffilm. Durch Streuprozesse im Kohlenstofffilm kommt es zu Ablenkungen der Elektronen, die für einige Elektronen so groß sind, dass der Hellfelddetektor nicht mehr getroffen wird. Die Signalstärke außerhalb des Partikels S_1 ist daher bei den hier verwendeten Bedingungen auf etwa 85% reduziert. Die minimale Signalstärke S_0 misst man im Zentrum des Partikels. Da der Latex-Partikel für Elektronen teiltransparent ist, ist S_0 größer als null. Die Signalstärke am Rand des Partikels wird als Schwellwertsignal S_{thres} bezeichnet.

Das Schwellwertsignal S_{thres} hängt sowohl von der Größe als auch vom Material des Partikels ab, siehe Bild 3. Die Materialabhängigkeit basiert auf unterschiedlichen Streueigenschaften. So streuen etwa Goldpartikel aufgrund der hohen Kernladungszahl stärker als z.B. Latexpartikel. Die Abnahme des Schwellwertsignals mit wachsender Partikelgröße resultiert aus der zunehmenden Streuung des Elektronenstrahls, der das Partikel trifft: Wenn die Mitte des Elektronenstrahls am Rand des Partikels liegt, wird die Hälfte des Strahls mit dem Partikel in Wechselwirkung treten und je nach durchstrahlter Dicke mehr oder weniger stark gestreut. Die zweite Hälfte der Elektronen fliegt ungestreut am Partikel vorbei in Richtung Detektor.

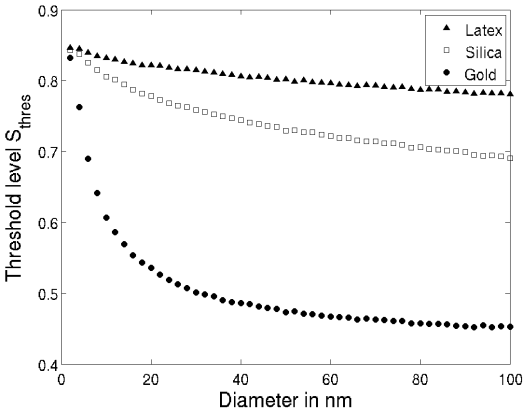


Bild 3: Signalstärke am Partikelrand in Abhängigkeit von der Partikelgröße für die drei hier untersuchten Materialien.

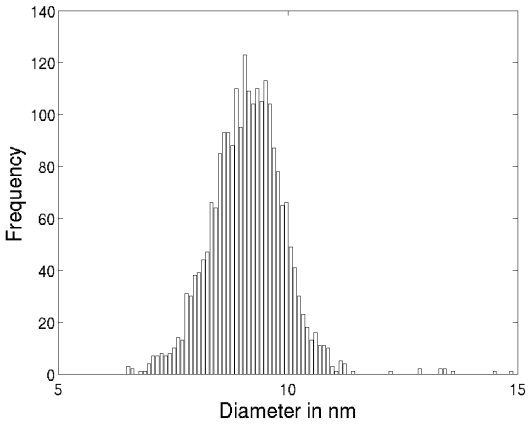


Bild 4: Größenverteilung der Gold-Nanopartikel.

Um das simulierte Schwellwertsignal auf ein aufgenommenes TSEM-Bild anwenden zu können, müssen die Simulationsergebnisse zu den gemessenen

Grauwerten der Aufnahme in Bezug gesetzt werden, wie in Bild 2 gezeigt. Für jeden einzelnen Partikel einer Aufnahme werden die benötigten Grauwerte aus einer so genannten „region of interest“ (ROI) bestimmt, die individuell um jeden Partikel herum definiert wird. Der Grauwert g_1 (der der Signalstärke S_1 entspricht) wird als Median der Grauwerte derjenigen Bildpixel bestimmt, die zum Hintergrund der ROI gehören. Der Grauwert g_0 (der der Signalstärke S_0 entspricht) wird als Mittelwert einiger weniger Grauwerte berechnet, die sich im Schwerpunkt des Partikels befinden. Mit diesen Größen lässt sich der gesuchte Schwellwert-Grauwert g_{thres} einfach mittels Dreisatz berechnen, vgl. Bild 2:

$$g_{thres}(d) = \frac{S_{thres}(d) - S_0(d)}{S_1 - S_0(d)} \cdot (g_1 - g_0) + g_0$$

In dieser Gleichung ist die Abhängigkeit der Signalstärken vom Partikeldurchmesser d explizit angegeben. Nicht nur die Schwellwert-Signalstärke S_{thres}, sondern auch die Signalstärke S_0 im Zentrum des Partikels hängt von der Partikelgröße ab; sie ist umso geringer je größer der Partikel ist.

4 Auswertung der TSEM-Bilder

Um die TSEM-Bilder auszuwerten, wurde ein Programm in Matlab [13] entwickelt, das Auswerte-Routinen der freien Bildauswertungs-Software ImageJ nutzt [14]. Der Rand jedes einzelnen Partikels wird individuell bestimmt. Dazu wird die Schwellwert-Signalstärke genutzt, die von der Simulation für einen Partikel entsprechender Größe und entsprechenden Materials vorhergesagt wird. Da dieser Schwellwert von der Partikelgröße abhängt, die anfangs nicht bekannt ist, wird ein iterativer Prozess angewandt. Zunächst wird ein globales Schwellwertverfahrens durchgeführt [15]. Es ermöglicht eine erste Schätzung der Partikelgröße, die für eine Schätzung des Schwellwerts genutzt werden kann, die wiederum zu einer verbesserten Schätzung der Größe führt, und so weiter. Die Iteration wird beendet, wenn Schwellwert und Größe konstant bleiben, wofür in der Regel weniger als zehn Iterationsschritte benötigt werden.

Der endgültige, durch die Iteration berechnete Schwellwert wird zur Bestimmung des Partikelrands genutzt. Um den Einfluss von Digitalisierungsfehlern

zu verringern, wird die Partikelgröße mit Subpixel-Genauigkeit bestimmt. Dazu wird die ROI, die den Partikel enthält, durch Interpolation sowohl in x- als auch in y-Richtung fünffach vergrößert; das gesamte Bild wird also 25 Mal größer. Nachdem der Partikel mit Hilfe des individuellen Schwellwerts vom Hintergrund getrennt wurde, wird die Anzahl der zum Partikel gehörenden Subpixel gezählt und mit der Fläche eines Subpixels (in nm^2) multipliziert. Die Partikelgröße ist schließlich definiert als der Durchmesser eines flächengleichen Kreises.

Um vereinzelte Nanopartikel von Agglomeraten und Artefakten wie etwa Verschmutzungen zu unterscheiden, werden Objekte nicht ausgewertet, die nicht innerhalb eines bestimmten Größenintervalls liegen oder die nicht rund genug sind. Dieses relativ simple Vorgehen beinhaltet das Risiko, Partikel mit außergewöhnlichen Größen und Formen zu ignorieren, während Artefakte, die Partikeln ähneln, in die Auswertung einbezogen werden. Daher müssen die drei Parameter - minimale und maximale Größe sowie Rundheit – präzise eingestellt werden. Um Artefakte zu vermeiden, sollte ebenso auf reine Proben und eine durchdachte Präparation unter Reinraumbedingungen geachtet werden.

5 Messunsicherheit

Die Messunsicherheit der mittleren Partikelgröße wurde nach den Regeln des „GUM“ bestimmt [17]. Im Folgenden werden die Beiträge zur Messunsicherheit diskutiert.

Die Kalibrierung der Längenskala des SEM ist in Abschnitt 2 beschrieben. Effekte, die die Pixelgröße beeinflussen, wie z.B. die so genannte „leading edge distortion“, müssen zusammen mit der Unsicherheit des Kalibrierstandards betrachtet werden, um die Unsicherheit der Pixelgröße zu bestimmten. Es stellt sich heraus, dass der resultierende Beitrag zur Gesamtunsicherheit im Vergleich zu den im Folgenden diskutierten Beiträgen nur eine geringe Rolle spielt.

Die Unsicherheit der Schwellwert-Signalstärke S_{thres} ist ein wichtiger Beitrag zur Gesamtunsicherheit, da sie unmittelbar die Position des Partikelrands und damit die gemessenen Partikelgröße beeinflusst. Der Grund für die mit dem

Schwellwert verbundene Unsicherheit liegt in den Eingangsparametern der Simulation, die nicht beliebig genau bestimmt werden können. Weiterhin kann jeweils nur mit einer beschränkten Anzahl von Elektronen gerechnet werden und die für die Simulation verwendeten Streuquerschnitte und Materialdichten sind nur innerhalb gewissen Grenzen bekannt. Ebenso sind Annahmen, wie beispielsweise die sphärischer Partikel, nur näherungsweise erfüllt. All diese Beiträge müssen berücksichtigt werden, um die Unsicherheit der Schwellwert-Signalstärke abzuschätzen.

Ein weiterer wichtiger Beitrag zur Gesamtunsicherheit resultiert aus der Auswertung der aufgenommenen TSEM-Bilder. So unterliegt die Bestimmung des Grauwerts g_1 des Hintergrunds sowie des Grauwerts g_0 im Zentrum des Partikels gewissen Unsicherheiten. Einflussgrößen sind unter anderem das Bildrauschen und lokale Schwankungen des Hintergrundsignals aufgrund von Unregelmäßigkeiten der Kohlenstoffschicht oder aufgrund von eingetrockneten Resten der Dispersionsflüssigkeit.

Bildrauschen in der Nähe des Partikelrands hat hingegen keinen großen Einfluss: Einige Pixel auf dem Rand oder in dessen Nähe können Grauwerte aufweisen, die zufällig unterhalb oder oberhalb des Schwellwerts liegen. Dies kann dazu führen, dass diese Pixel fälschlicherweise zum Partikel gezählt werden bzw. fälschlicherweise nicht zum Partikel gerechnet werden. Der Erwartungswert der resultierenden Abweichung der Partikelgröße ist null und eine detaillierte Auswertung zeigt, dass die damit verbundene Unsicherheit im Vergleich zu anderen Beiträgen vernachlässigt werden kann.

Da runde Objekte in einem Digitalbild als eine Menge von quadratischen Pixeln dargestellt werden, kommt es zu Digitalisierungsfehlern. Der daraus resultierende Beitrag zur Unsicherheit wird durch die Interpolation der Bilder vor der Größenbestimmung wesentlich verkleinert.

Eine angemessene Anzahl von Partikeln sollte ausgewertet werden, um statistisch aussagekräftige Resultate zu erhalten. Da keine Annahmen zur Größenverteilung gemacht werden, wird die statistische Unsicherheit als experimentelle Standardabweichung der mittleren Partikelgröße berechnet. (Unter der Annah-

me einer Lognormal-Verteilung kann die statistische Unsicherheit geringer abgeschätzt werden [18].) Da hier eine genügend große Anzahl an Partikeln ausgewertet wird (einige tausend) und die Größenverteilungen der Partikel vergleichsweise schmal sind, ist der Beitrag der statistischen Unsicherheit im Vergleich zu den systematischen Unsicherheitsbeiträgen gering.

Die Probenpräparation sollte sicherstellen, dass sich ein repräsentativer Teil der suspendierten Partikel auf dem Substrat befindet. Ein laufendes Forschungsvorhaben untersucht diese Fragestellung [19]. Zurzeit ist ein Unsicherheitsbeitrag aufgrund einer möglicherweise nicht perfekten Probenpräparation nicht in der Messunsicherheit enthalten. Die angegebenen Daten wie die Breite der Größenverteilung und ihr Mittelwert beschreiben die Nanopartikel, die tatsächlich auf dem Substrat vorliegen.

Die quadratische Summation aller Unsicherheitsbeiträge liefert die kombinierte Unsicherheit, von der die erweiterte Messunsicherheit der mittleren Partikelgröße abgeleitet wird. Hier wird als Erweiterungsfaktor k=2 gewählt, was einem Vertrauensintervall von 95% entspricht. Für die drei Nanopartikel-Proben, die hier untersucht wurden, liegt die erweiterte Messunsicherheit im Bereich zwischen 1 nm und 3 nm, Details siehe in Abschnitt 6. Es soll darauf hingewiesen werden, dass diese erweiterten Messunsicherheiten alle bekannten systematischen Beiträge beinhalten, inklusive des relativ kleinen Beitrags der statistischen Unsicherheit. Dies sollte beachtet werden, wenn diese Unsicherheiten mit denen aus anderen Veröffentlichungen verglichen werden, in denen teilweise nur der statistische Unsicherheitsbeitrag angegeben wird.

6 Ergebnisse und Diskussion

In diesem Abschnitt werden die Ergebnisse der Messungen an den drei Nanopartikel-Proben dargestellt und diskutiert. In Tabelle 1 sind die relevanten Ergebnisse zusammengestellt. Sie enthält die mittlere Partikelgröße und die dazugehörige erweitere Unsicherheit, Angaben zur Größenverteilung sowie die Anzahl der aufgenommenen Bilder und ausgewerteten Partikel.

Die Größenverteilung der Goldpartikel (NIST RM8011,) ist in Bild 4 gezeigt. Sie basiert auf der Auswertung von 2300 Partikeln, deren mittlere Größe 9,1 nm beträgt.

	Gold	Silika	Latex
Mittlere Partikelgröße in nm	9,1	27,5	44,5
Erweitere Unsicherheit der mittleren Größe in nm	1,2	1,6	2,6
Statistische Standardabweichung der mittleren Größe in nm	0,02	0,09	0,17
Breite der Größenverteilung in nm (Standardabweichung der Partikelgröße)	0,8	5,2	7,9
Median der Größe in nm	9,1	28,9	45,7
Modalwert der Größe in nm	9,1	29,6	46,7
Anzahl der aufgenommenen TSEM-Bilder	266	230	271
Anzahl der ausgewerteten Partikel	2318	3144	2113

Tabelle 1: Ergebnisse der Messungen der drei untersuchten Nanopartikel-Proben.

Die Messungen der Silika-Partikel (IRMM 304,) zeigen eine kleine Bimodalität der Größenverteilung auf, siehe Bild 5. Da mehr als 3100 Partikel analysiert wurden, ist die Bimodalität statistisch signifikant. Dieses Beispiel mach einen bedeutenden Vorteil von abbildenden Verfahren wie der Elektronenmikroskopie gegenüber Ensemble-Messungen wie DLS deutlich: Obwohl Ensemble-Techniken oft schneller sind als mikroskopische Einzelmessungen, ist es sehr schwierig, solche schwach ausgeprägten Bimodälitäten mittels DLS u.ä. zu quantifizieren [20].

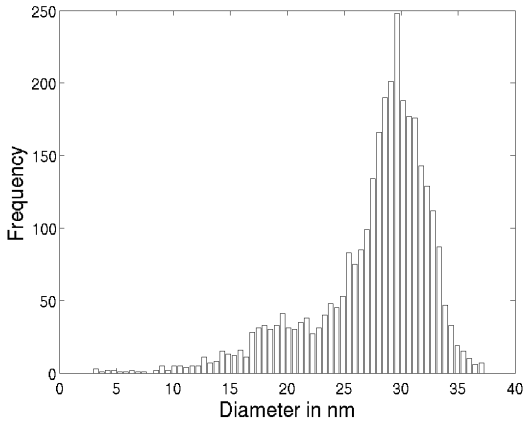


Bild 5: Größenverteilung der Silika-Nanopartikel.

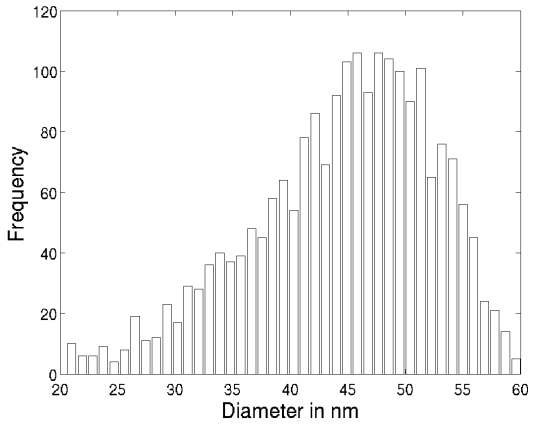


Bild 6: Größenverteilung der Latex-Partikel.

Bild 6 zeigt die Größenverteilung der gemessenen Latex Partikel von Duke Scientific. Das Größenhistogramm weist eine deutliche Abweichung von symmetrischen Verteilungen wie etwa einer Normalverteilung auf. Auch derartig ungewöhnliche Größenverteilungen sind mittels Ensemble-Techniken nur schwer nachzuweisen [20].

7 Zusammenfassung

Die Anwendung der TSEM-Technik für rückgeführte Größenmessungen von Nanopartikeln mit Größen bis hinunter zu etwa 5 nm wurde exemplarisch für drei verschiedene Materialklassen gezeigt. Hochgenaue Größenmessungen sind auf eine präzise Bestimmung des Partikelrandes angewiesen, die durch die Nutzung physikalischer Modelle zur Simulation der Bildentstehung verbessert werden kann. Diese Simulationen zeigen eine Abhängigkeit der TSEM-Signalstärke am Partikelrand von Partikelgröße und –material. Infolgedessen ist ein Iterationsprozess während der Bildauswertung nötig, wobei der Schwellwert individuell für jeden einzelnen Partikel bestimmt wird. Trotz dieses Aufwands kann eine vollständige Charakterisierung einer Partikelprobe durch den Einsatz automatisierter Routinen für Bildaufnahme und -auswertung erstaunlich schnell durchgeführt werden, wobei Tausende von Partikeln auf Hunderten von Bildern ausgewertet werden. Dadurch wird der Nachteil ungenügender Statistik überwunden, der oftmals mit Elektronenmikroskopie verbunden wird. Mit der Methode können Unsicherheiten zwischen 1 nm und 3 nm mit kommerziell erhält-

licher Messausrüstung erreicht werden. Die durch die Probenpräparation bedingte Unsicherheit ist in den genannten Werten allerdings noch nicht enthalten.

Literatur

[1] Van Ngo V, Hernandez M, Roth B and Joy D C 2007 Microscopy Today 15 12–16

[2] Farrow R C, Postek M T, Keery W J, Jones S N, Lowney J R, Blakey M et al. 1997 J. Vac. Sci. Technol. B. 15 2167-72

[3] Buhr E, Senftleben N, Klein T, Bergmann D, Gnieser D, Frase C G et al. 2009 Meas. Sci. Technol. 20 084025

[4] Postek M T, Lowney J R, Vladar A E, Keery W J, Marx E and Larrabee R D 1993 J. Res. Nat. Inst. Stand. Technol. 98 415-45

[5] Buhr E, Michaelis W, Diener A and Mirandé W 2007 Meas. Sci. Technol. 18 667-74

[6] Chernoff D A, Buhr E, Burkhead D L and Diener A 2008 Proc. SPIE 6922 69223J

[7] Sezgin M and Sankur B 2004 J. Electron. Imaging 13 146–68

[8] Sadowski T E, Broadbridge C C and Daponte J 2007 Mater. Res. Soc. Symp. Proc. 982 25–30

[9] Frase C G, Gnieser D and Bosse H 2009 J. Phys. D: Appl. Phys. 42 183001

[10] Johnsen K P, Frase C G, Bosse H and Gnieser D 2010 Proc. SPIE 7638 76381O

[11] Salvat F and Mayol R 1993 Comp. Phys. Comm. 74 358-74

[12] Joy D C and Luo S Scanning 11 176-80

[13] 2009 MATLAB R2009a (Natick, Massachusetts: The MathWorks Inc.)

[14] Abramoff M D, Magalhaes P J and Ram S J 2004 Biophotonics International 11 36–43

[15] Prewitt J M and Mendelsohn M L 1965 Ann. N.Y. Acad. Sci. 128 1035–53

[16] Vincent L and Soille P 1991 IEEE Trans. Pattern Anal. Mach. Intell. 13 583–98

[17] 1993 Guide to the Expression of Uncertainty in Measurement (Geneva: International Organization for Standardization)

[18] 2004 ISO 13322 - 1 Particle Size Analysis - Image Analysis Methods - Part 1. Static Image Analysis Methods (Geneva: International Organization for Standardization)

[19] DIN INS “Präparationsmethoden für Größenmessungen - Realisierung und Optimierung von Präparationsmethoden für zuverlässige Größenmessungen mit AFM und TSEM“

[20] Rasteiro M G, Lemos C C and Vasquez A 2008 Part. Sci. Technol. 26 413-37

Neuartiges Kalibrierstrahlersystem zur Präzisionskalibrierung von Strahlungsthermometern

Marc Schalles, Thomas Fröhlich

Technische Universität Ilmenau, 98693 Ilmenau, E-Mail:
marc.schalles@tu-ilmenau.de

Kurzfassung

Ein Dreifach-Fixpunktstrahler (DFS) zur Präzisionskalibrierung von Strahlungsthermometern wurde am Institut für Prozessmess- und Sensortechnik entwickelt, der erstmals eine Mehrpunkt-Kalibrierung mit direktem Bezug zur Internationalen Temperaturskala (ITS-90) ermöglicht. Er enthält die Fixpunktsubstanzen Zinn, Zink und Aluminium und kann daher zur Kalibrierung bei $T_1 = 231,928\,^{\circ}\mathrm{C}$, $T_2 = 419,527\,^{\circ}\mathrm{C}$ und $T_3 = 660,323\,^{\circ}\mathrm{C}$ mit Darstellungsunsicherheiten von $31\,\mathrm{mK} \ldots 86\,\mathrm{mK}$ ($k = 2$) eingesetzt werden.

1 Einleitung

Die Kalibrierung von Strahlungsthermometern wird in der Regel mit Hilfe von temperaturvariablen Hohlraumstrahlern durchgeführt (Bild 1, a). Dies sind Körper, meistens Wärmerohre, die mit Hilfe von präzise geregelten Öfen auf eine idealerweise konstante, beliebig einstellbare Temperatur T_n gebracht werden. Im Inneren der Körper befindet sich ein Hohlraum, an dessen Apertur ein sehr hoher Emissionsgrad ϵ nahe 1 vorliegt. Das zu kalibrierende Strahlungsthermometer misst die Strahldichte $L(T_\mathrm{n})$ der aus dem Inneren kommenden Wärmestrahlung und wird an dieser kalibriert. Das Problem dieser Kalibriermethode besteht darin, dass die Temperatur der Referenzstrahlung mit Hilfe eines zweiten Transferstrahlungsthermometers (TSRT) oder Standard-Platinwiderstandsthermometers (SPRT) bestimmt werden muss, um eine direkte Rückführbarkeit der Kalibrierung auf die Temperaturen der Fixpunk-

te $T_{1,\,2\,\ldots,\mathrm{n}}$ der Internationalen Temperaturskala ITS-90 zu gewährleisten. Die mit diesem Verfahren realisierte erweiterte Unsicherheit bei Kalibrierungen durch die Physikalisch Technische Bundesanstalt, PTB, liegt im Bereich von $\pm 40\,\mathrm{mK} \ldots \pm 280\,\mathrm{mK}$ bei $270\,^{\circ}\mathrm{C} \ldots 962\,^{\circ}\mathrm{C}$ [1].

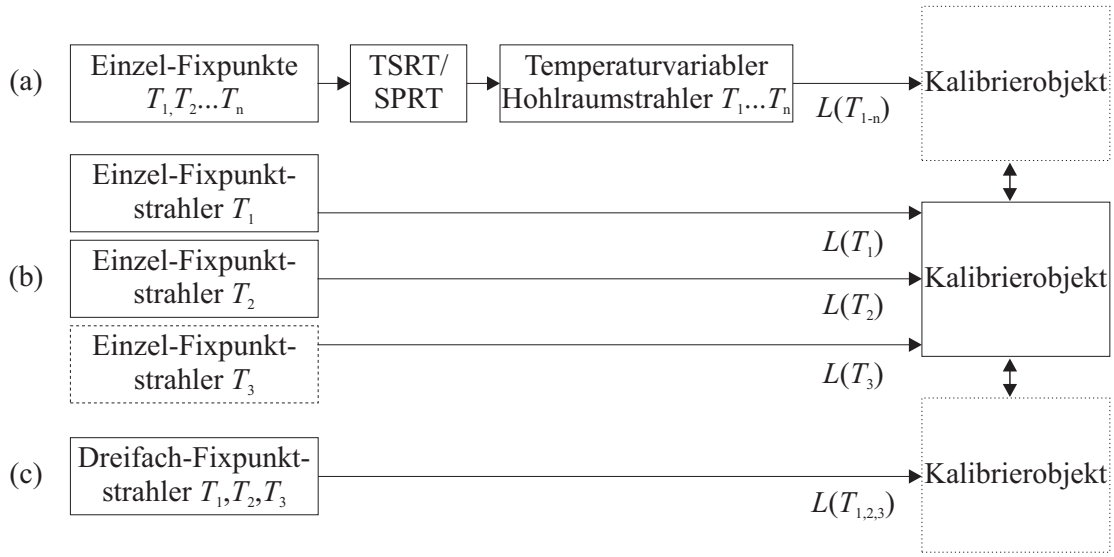


Bild 1: Vergleich von Methoden zur Kalibrierung von Strahlungsthermometern mit Rückführbarkeit zur ITS-90

Möchte man die Präzision der Kalibrierung erhöhen und gleichzeitig direkt rückführbar auf die ITS-90 sein, benutzt man anstelle von temperaturvariablen Hohlraumstrahlern sogenannte Fixpunktstrahler. Dies sind Hohlraumstrahler, in deren Innerem sich zusätzlich ein Reservoir gefüllt mit einer ITS-90 Fixpunktsubstanz, also meistens hochreine Metalle, befindet. Schmilzt man diese auf, bzw. lässt sie erstarren, bleibt die Temperatur im Strahlerhohlraum konstant auf der bekannten Phasenumwandlungstemperatur T_{90} und das Ausgangssignal des Strahlungsthermometers $Y(T_{90})$ kann direkt an der emittierten Strahldichte $L(T_{90})$ kalibriert werden. Die typische erweiterte Messunsicherheit dieser Kalibrierungen ist besser $\pm 100\,\mathrm{mK}$.

Da Strahlungsthermometer für gewöhnlich deutlich nichtlineare Kennlinien besitzen, müssen jedoch mehrere Fixpunktstrahler mit unterschiedlichen Fixpunktsubstanzen für die Kalibrierung genutzt werden (Bild 1, b). Durch die erforderliche Nutzung mehrerer Öfen, Wärmerohre und Regler entsteht dabei ein erheblicher gerätetechnischer und durch die erneute Justierung des Strahlungsthermometers auch zeitlicher Aufwand.

Diese Nachteile werden durch den vorgestellten Dreifach-Fixpunktstrahler (DFS) beseitigt. Durch seine verbesserte Bauform und ermöglicht er eine schnelle, automatisierbare Kalibrierung an 3 Schmelz- und 3 Erstarrungsvorgängen bei nur einmaliger Justage des Kalibrierobjekts (Bild 1, c). In einem Aufheiz- und Abkühlvorgang kann somit ein Kalibrierobjekt vollständig kalibriert (Bild 2) und auch die Kurzzeitdrift des Strahlungsthermometersignals $Y(T)$ ermittelt werden.

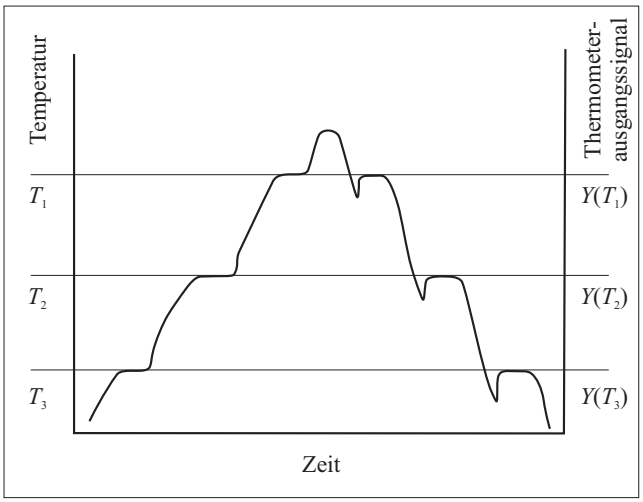


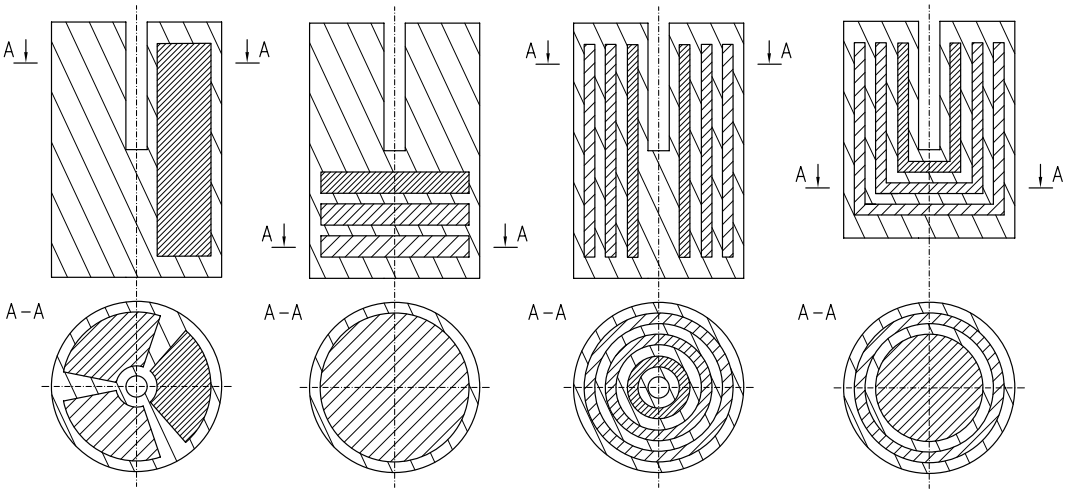
Bild 2: Schematische Darstellung der gemessenen Hohlraumstrahlertemperatur im DFS mit Phasenumwandlungsplateaus

2 Aufbau

Mit der Entwicklung des Dreifach-Fixpunktstrahlers wurde erstmalig die Integration von drei unterschiedlichen Fixpunktsubstanzen in nur eine Fixpunktstrahlerbauform vorgenommen. Als Fixpunktsubstanzen wurden Zinn (Erstarrungstemperatur $T_{\mathrm{sn}} = 231,928\,^{\circ}\mathrm{C}$), Zink ($T_{\mathrm{zn}} = 419,527\,^{\circ}\mathrm{C}$) und Aluminium und ($T_{\mathrm{al}} = 660,323\,^{\circ}\mathrm{C}$) ausgewählt. Sie sind in der ITS-90 aufgelistet, besitzen günstige thermische Materialeigenschaften, z.B. eine hohe Phasenumwandlungsenthalpie, und überdecken einen für die Kalibrierung von Strahlungsthermometern, z.B. mit InGaAs-Detektor, wichtigen Temperaturbereich.
Ebenfalls neuartig ist, dass der Grundkörper des DFS aus hochreiner Aluminiumoxidkeramik gefertigt ist, welche im Gegensatz zu herkömmlichen Grafit-Fixpunktstrahlern eine höhere mechanische Belastbarkeit der Bauform ergibt.

Thermische Spannungen die durch unterschiedliche Ausdehnungskoeffizienten der verwendeten Materialien und durch das Ausdehnen und Zusammenziehen der Fixpunktsubstanzen während der Phasenumwandlung können dadurch beherrscht werden.

Zu Beginn der Entwicklung wurde zunächst untersucht, welche Anordnung der Kammern mit den Fixpunktsubstanzen zu wählen ist, um eine gut messbare, lang anhaltende Phasenumwandlung und geringe statische und dynamische Temperaturgradienten im Strahlerhohlraum zu erreichen. Grundsätzlich sind dabei vier wesentliche Varianten eines Dreifach-Fixpunktstrahlers denkbar:



Kammern auf Umfang verteilt	Kammern axial im Strahlerboden	Hohlzylindrische Kammern	Becherförmige Kammern
(1)	(2)	(3)	(4)

Bild 3: Mögliche Anordnungen der Kammern im Dreifach-Fixpunktstrahler

Alle vier Varianten wurden mit Hilfe von FEM-Temperaturfeldberechnungen verglichen und auf ihre Verwendbarkeit überprüft. Exemplarisch sind in der folgenden Abbildung die Temperaturverteilungen in verschiedenen Fixpunktstrahlervarianten während der Schmelze von Aluminium dargestellt, da es die höchste Schmelztemperatur besitzt und sich somit Temperaturgradienten und Wärmeableitungsfehler am stärksten ausbilden. Das Aluminium befindet sich in der dem Strahlerhohlraum nächsten, Zink und Zinn in den weiteren Kammern. Als Temperatur der Strahleraußenfläche (Ofentemperatur während der Phasenumwandlung) wurden $665\,^{\circ}\mathrm{C}$ angenommen.

Bei den Varianten 1 und 2 sind die Kammern ungünstig positioniert, da sich

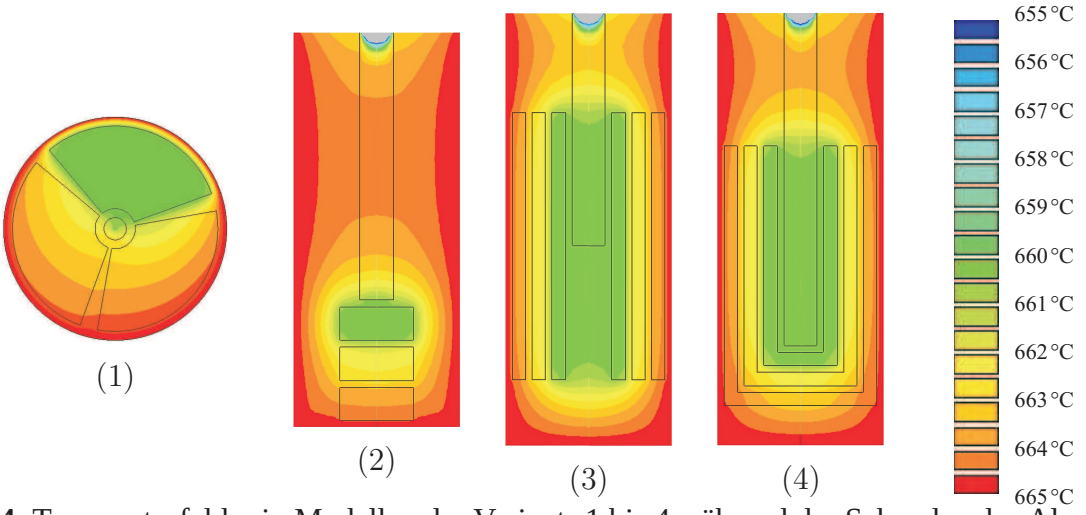


Bild 4: Temperaturfelder in Modellen der Variante 1 bis 4 während der Schmelze des Aluminiums

während der Phasenumwandlung des Aluminiums (grüne Bereiche) kein homogenes Temperaturfeld im Strahlerhohlraum ausbildet. Deutlich bessere Ergebnisse können mit der hohlzylindrischen Kammeranordnung der Variante 3 erreicht werden, da hier der Hohlraumstrahler weitgehend von den Fixpunktsubstanzen umschlossen ist. Jedoch kann auch hier eine Drift der Bodentemperatur des Hohlraumstrahlers aufgrund von Wärmeeinleitung vom offenen Kammerende am Boden nicht ausgeschlossen werden. In Variante 4 hingegen umschließen die Kammern den Hohlraum koaxial und am Boden. Die Isothermenfronten breiten sich während der Phasenumwandlung nahezu in radialer Richtung aus, weshalb die Temperatur des Strahlerbodens lange konstant bleibt und sehr ausgeprägte Temperaturplateaus entstehen. Diese Anordnung der Kammern wurde deshalb im DFS realisiert. Mit Hilfe weiterführender Optimierungsrechnungen wurden die Geometrie der Bauform sowie die Kammervolumina optimiert, um folgende metrologische Parameter zu erreichen:

- Durchmesser $50\,\mathrm{mm}$, maximale Länge $150\,\mathrm{mm}$
- Apertur des Hohlraums $1010\,\mathrm{mm}$
- effektiver Hohlraumemissionsgrad $\epsilon_{\mathrm{eff}} > 0,9995$
- Gesamtmasse der Fixpunktmaterialien $< 0,5\,\mathrm{kg}$
- Reproduzierbarkeit der Fixpunkttemperaturen $< 10\,\mathrm{mK}$
- Erweiterte Unsicherheit der Strahlertemperatur $U(T_{\mathrm{ph}}) < 100\,\mathrm{mK}$

Als Ergebnis dieser Arbeiten entstand die in Bild 5 dargestellte Bauform. Sie

besteht aus vier Zylindern 99,7%-iger Al_2O_3-Keramik welche beim Befüllungsprozess der Fixpunktsubstanzen (99,99%-ige Reinheit) zusammengeschoben und mit einer Deckplatte arretiert werden. Die Gesamtlänge des Strahlerhohlraumes beträgt $118\,\mathrm{mm}$ bei einer Apertur von $10\,\mathrm{mm}$ und ist damit vergleichbar mit bekannten Einzel-Fixpunktstrahlern [2], [3].
Es ist bekannt dass dielektrische Materialien stark wellenlängenabhängige, teils niedrige Emissionsgrade ϵ_λ haben [4]. Deshalb wurde am Institut für angewandte Materialforschung der Universität Duisburg-Essen der spektrale Emissionsgrad einer Al_2O_3-Probe überprüft. Die erwartete deutliche Wellenlängenselektivität konnte dabei bestätigt werden. Zur Erhöhung und Homogenisierung des Wandemissionsgrads und zur Verbesserung des effektiven Hohlraumemissionsgrads $\epsilon_{\mathrm{eff},\lambda}$, der sich nach Mehrfachreflexion der Strahlung im Hohlraum ergibt, wurde deshalb zusätzlich der Hochtemperatur-Emissionsgradlack HE23 der Firma Rolls Royce aufgetragen, sowie der Hohlraumboden konisch abgeschrägt. Mit diesen Maßnahmen konnte $\epsilon_{\mathrm{eff},\lambda}$ im Wellenlängenbereich um $1,5\,\mu\mathrm{m}$ auf hohe Werte von $0,99986 \ldots 0,99988$ gehoben werden [5].

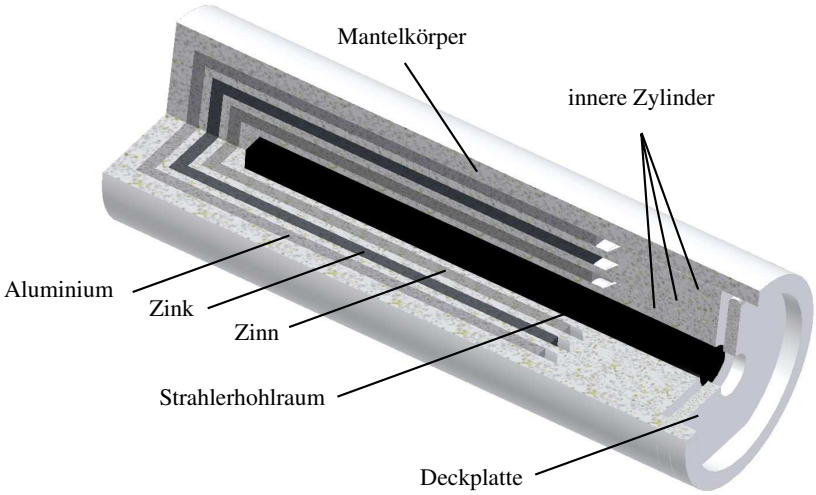


Bild 5: Schnittdarstellung eines Dreifach-Fixpunktstrahlers

3 Messungen

Zur Temperierung des DFS wurde ein Rohrofen mit integriertem Na-Wärmerohr verwendet und zur Verringerung der Störstrahlung während der Messungen

eine Wasser gekühlte Blende vor die Apertur des DFS montiert. Die Ofentemperatur wurde wenige Kelvin über oder unter die Phasenumwandlungstemperatur eingestellt, um die Schmelze oder Erstarrung der jeweiligen Fixpunktsubstanz auszulösen. Die Hohlraumstrahlertemperatur während der Phasenumwandlungen wurde mit einem Transferstrahlungsthermometer LP5 [6] mit InGaAs-Detektor, Messbereich $200\,°\mathrm{C}$ bis $2500\,°\mathrm{C}$ und $1\,\mathrm{mK}$ Auflösung vermessen. Es ergaben sich folgende Plateauverläufe:

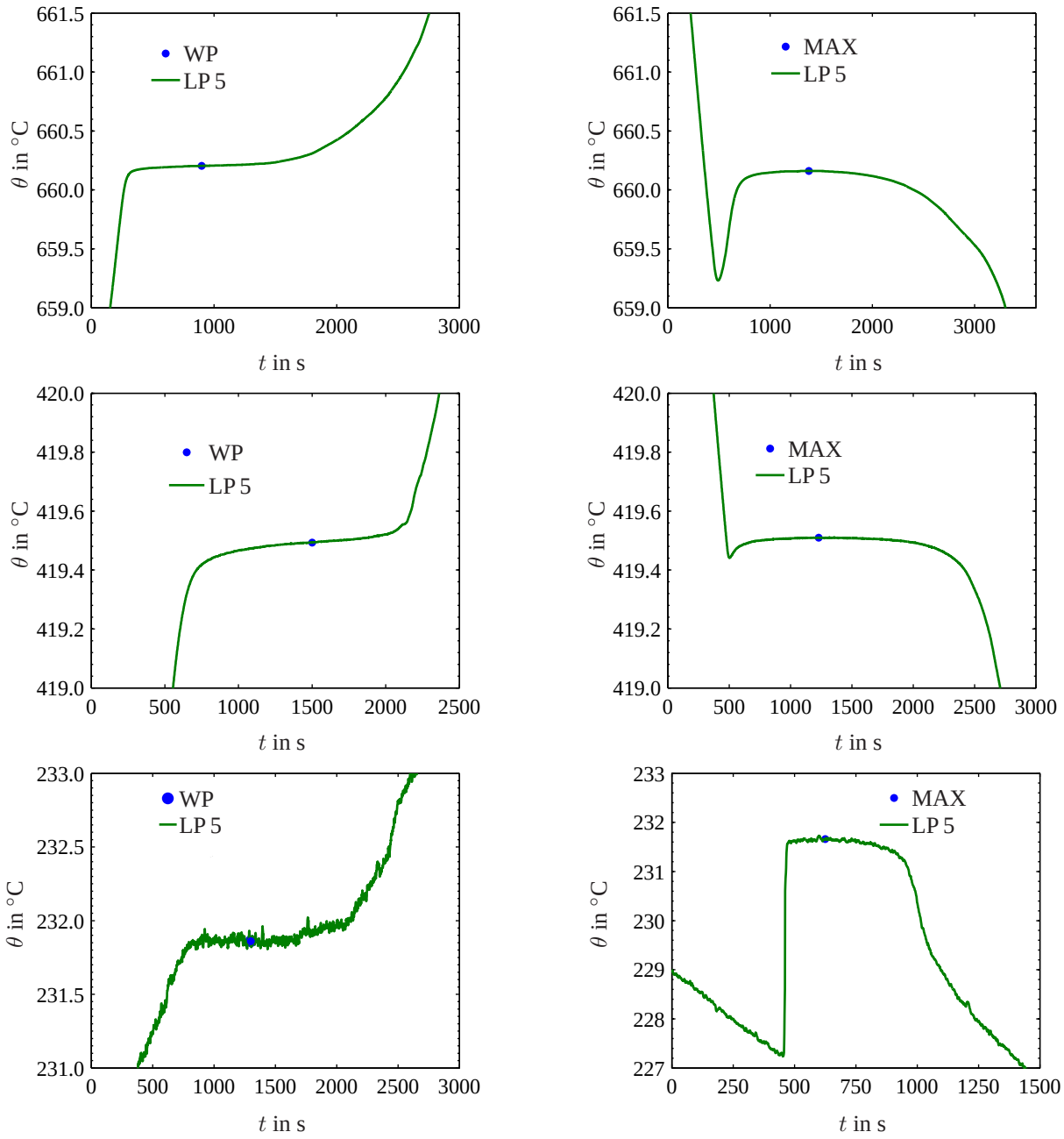


Bild 6: Al, Zn, Sn-Liquidustemperaturverlauf (links) und Solidustemperaturverlauf (rechts) bei Messung mit einem LP 5

In den Messungen gut zu erkennen sind die unterschiedlichen Materialeigenschaften der Fixpunktsubstanzen. Während Zink während der Erstarrung nahezu keine und Aluminium nur eine geringe Unterkühlung aufweist, unterkühlt Zinn sehr stark bevor es erstarrt. Dadurch fällt auch die Dauer der Zinn-Erstarrung mit ca. $750\,\mathrm{s}$ kürzer aus als bei Al und Zn (ca. $2000\,\mathrm{s}$).
Bei einer aufeinander folgenden Messung von mehr als zehn Phasenumwandlungen an jedem Fixpunkt wurde die Reproduzierbarkeit der Phasenumwandlungstemperaturen überprüft. Dazu wurden die Temperaturen der Wendepunkte der Liquidusplateaus T_l und der Maxima der Solidusplateaus T_s ermittelt. Sie sind in den Temperaturverläufen blau markiert. In Tabelle 1 sind die Standardabweichungen der Temperaturen aufgelistet.

Tabelle 1: Mittlere Fixpunkttemperaturen der LP 5-Messungen

	Schmelze	Erstarrung
Fixpunkt	$s(T_\mathrm{l})$ in mK	$s(T_\mathrm{s})$ in mK
Al	3	5
Zn	9	4
Sn	13	20

Es ergaben sich am Al- und Zn-Fixpunkt sehr geringe Werte von $3\,\mathrm{mK}$ bis $9\,\mathrm{mK}$, welche die gestellten Anforderungen bestens erfüllen. Die höheren Werte am Sn-Fixpunkt folgen einerseits aus den bereits beschriebenen Materialeigenschaften sowie dem stärkeren Rauschen des Strahlungsthermometers bei tieferen Temperaturen. Daher sind die erreichten Werte von $13\,\mathrm{mK}$ und $20\,\mathrm{mK}$ auch als sehr gut einzuschätzen.

Eine Aussage zur erreichbaren Unsicherheit der Darstellung der DFS-Temperatur lässt sich aus diesen Messungen nicht ableiten. Daher wurde eine Messunsicherheitsbetrachtung vorgenommen. Die Eingangsgröße des aufgestellten Modells ist dabei die Temperatur T_{90}, welche ein idealer Fixpunktstrahler lt. ITS-90 zur Kalibrierung bereitstellen sollte, die Ausgangsgröße ist die Temperatur T_{ph} welche tatsächlich an der Strahlerapertur emittierten Strahldichte $L(T_{90})$ während der Phasenumwandlung entspricht.

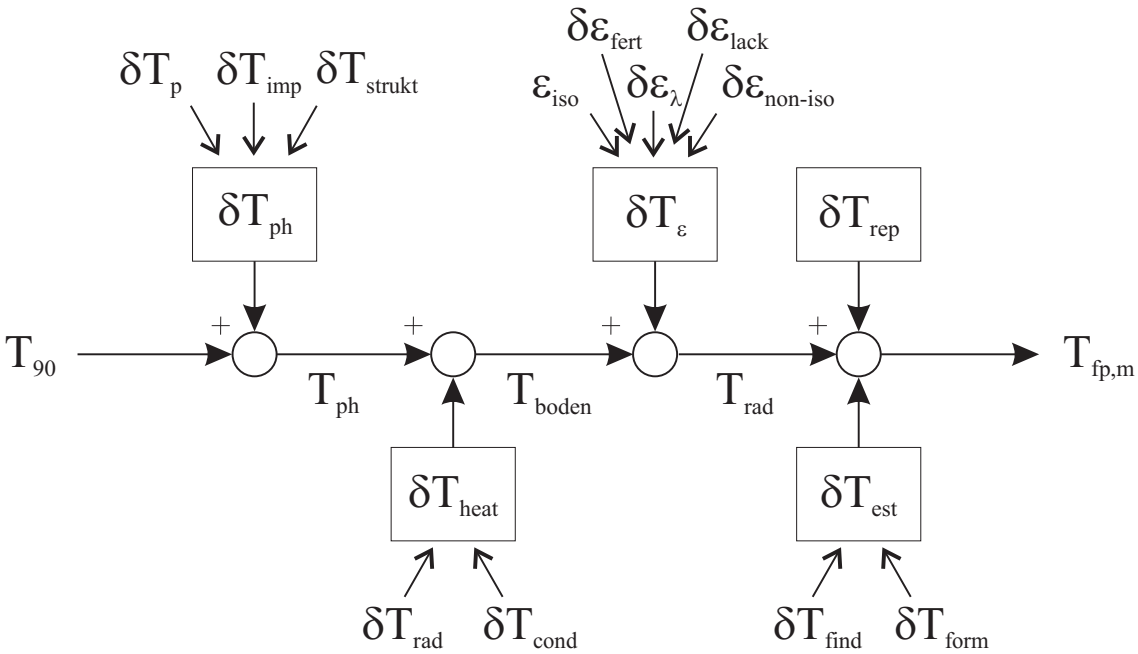


Bild 7: Modell des Messunsicherheitsbudgets

Die 13 verschiedenen Einflussgrößen sind entsprechend ihrer Ursache in 5 Untermodelle zusammengefasst, deren Ergebnisgrößen Temperaturabweichungen δT sind, die aufsummiert die unsicherheitsbehaftete Gesamtabweichung von T_{ph} zu T_{90} ergeben:

- $\boldsymbol{\delta T_{\mathrm{ph}}}$- Differenz der realen Phasenumwandlungstemperatur $\boldsymbol{T_{\mathrm{ph}}}$ der DFS-Fixpunktsubstanz zur Temperatur $\boldsymbol{T_{90}}$ der reinen Fixpunktsubstanz unter Normaldruck ($\boldsymbol{\delta T_{\mathrm{p}}}$- Korrektur hydrostatischer Druck, $\boldsymbol{\delta T_{\mathrm{imp}}}$ - Verunreinigung Fixpunktsubstanz, $\boldsymbol{\delta T_{\mathrm{strukt}}}$- Strukturveränderungen Fixpunktsubstanz)
- $\boldsymbol{\delta T_{\mathrm{heat}}}$ - Veränderung der Hohlraumstrahlertemperatur $\boldsymbol{T_{\mathrm{boden}}}$ durch Wärmeaustausch mit der Umgebung und dem Ofen ($\boldsymbol{\delta T_{\mathrm{cond}}}$-Wärmeleitung, $\boldsymbol{\delta T_{\mathrm{rad}}}$-Wärmestrahlung)
- $\boldsymbol{\delta T_{\epsilon}}$ - Temperaturabweichung der zu messenden Strahlungstemperatur $\boldsymbol{T_{\mathrm{rad}}}$ durch Emissionsgradabweichungen ($\boldsymbol{\delta \epsilon_{\mathrm{iso}}}$ - Bestimmung Emissionsgrad isothermer Hohlraum, $\boldsymbol{\delta \epsilon_{\mathrm{non-iso}}}$ - Nichtisothermie Hohlraum, $\boldsymbol{\delta \epsilon_{\mathrm{fert}}}$ - Fertigungstoleranzen des Hohlraumstrahlers, $\boldsymbol{\delta \epsilon_{\mathrm{lack}}}$-Lackierungsfehler, $\boldsymbol{\delta \epsilon_{\lambda}}$ - Messung Wandemissionsgrad)
- $\boldsymbol{\delta T_{\mathrm{rep}}}$ - Abweichungen durch die endliche Reproduzierbarkeit der Fixpunkttemperaturen

- $\boldsymbol{\delta T_{\mathrm{est}}}$ - bei der rechnerischen Ermittlung der Fixpunkttemperatur aus dem Plateauverlauf entstehende Temperaturabweichungen ($\boldsymbol{\delta T_{\mathrm{find}}}$ - Auffinden des Fixpunkts, $\boldsymbol{\delta T_{\mathrm{form}}}$ - Form des Temperaturplateaus).

Berücksichtigt man alle Einflussgrößen, ergeben sich aus dem Messunsicherheitsbudget folgende Fixpunkttemperaturen des DFS und ihre erweiterten Messunsicherheiten:

Tabelle 2: Fixpunkttemperaturen des DFS als Ergebnis des Unsicherheitsbudgets

	ITS-90 Fixpunkttemperatur T_{90}	DFS Fixpunkttemperatur T_{fp} ($k = 2$)	Differenz $T_{\mathrm{fp}} - T_{90}$
Al (Liqu.)	$660,323\,^{\circ}\mathrm{C}$	$660,285\,^{\circ}\mathrm{C} \pm 31\,\mathrm{mK}$	$-38\,\mathrm{mK}$
Al (Sol)		$660,263\,^{\circ}\mathrm{C} \pm 33\,\mathrm{mK}$	$-60\,\mathrm{mK}$
Zn (Liqu.)	$419,527\,^{\circ}\mathrm{C}$	$419,377\,^{\circ}\mathrm{C} \pm 39\,\mathrm{mK}$	$-150\,\mathrm{mK}$
Zn (Sol.)		$419,407\,^{\circ}\mathrm{C} \pm 36\,\mathrm{mK}$	$-120\,\mathrm{mK}$
Sn (Liqu.)	$231,928\,^{\circ}\mathrm{C}$	$231.919\,^{\circ}\mathrm{C} \pm 61\,\mathrm{mK}$	$-9\,\mathrm{mK}$
Sn (Sol.)		$231,914\,^{\circ}\mathrm{C} \pm 86\,\mathrm{mK}$	$-14\,\mathrm{mK}$

Diese Werte weichen von den Temperaturen lt. ITS-90 ab, weil keine hoch reinen Fixpunktsubstanzen verwendet wurden und die darin enthaltenen Verunreinigungen die Phasenumwandlungstemperatur verschieben. Zusätzlich dazu verursachen bekannte stationäre Temperaturgradienten im Strahlerhohlraum Emissionsgradveränderungen. Auch die bereits beschriebenen Materialeigenschaften (Zinn) bedingen Plateauverformungen, welche berücksichtigt werden müssen. Da diese Abweichungen jedoch mit geringer Unsicherheit bestimmbar sind, erreicht der DFS trotzdem geringe Messunsicherheiten in der Temperaturdarstellung von $31 \cdots 86\,\mathrm{mK}$ die dem international erreichten Stand der Technik mit Einzel-Fixpunktstrahlern entsprechen.
Eine ausführliche Darstellung des Messunsicherheitsbudgets (s. [7]) würde an dieser Stelle zu weit führen, deshalb gibt Tabelle 3 einen Überblick der relativen Beiträge in % zur jeweils ermittelten Messunsicherheit.

Tabelle 3: Zusammenstellung der Beiträge zur Gesamtunsicherheit der Fixpunkttemperaturen. Die größten Beiträge sind markiert. T_{fp}

Größe	Al-Liqu.		Al-Sol.		Zn-Liqu.		Zn-Sol.		Sn-Liqu.		Sn-Sol.	
	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$	$u(x_{\mathrm{i}})$	$\frac{u(x_{\mathrm{i}})}{u(y_{\mathrm{i}})}$
	mK	%	mK	%	mK	%	mK	%	mK	%	mK	%
δT_p	2,4	2,4	2,4	2,2	1,2	0,4	1,2	0,5	0,8	0,1	0,8	0,0
δT_{imp}	1,3	0,7	1,3	0,6	13,5	**47,4**	13,5	**57,0**	0,4	0,0	0,4	0,0
$\delta T_{\mathrm{strukt}}$	2,9	3,5	2,9	3,2	2,9	2,2	2,9	2,6	2,9	0,9	2,9	0,5
δT_{rad}	6,1	**15,4**	6,1	**14,0**	0,8	0,2	0,8	0,2	0,1	0,0	0,1	0,0
δT_{cond}	10,0	**41,4**	10,0	**37,6**	9,0	**21,1**	9,0	**25,3**	27,0	**79,5**	27,0	**39,4**
$\delta T_{\epsilon-\mathrm{iso}}$	0,9	0,0	0,9	0,0	0,5	0,0	0,5	0,0	0,3	0,0	0,3	0,0
$\delta T_{\epsilon-\lambda}$	8,7	**31,2**	8,7	**28,4**	5,1	6,8	5,1	8,1	2,8	0,9	2,8	0,4
$\delta T_{\epsilon-\mathrm{lack}}$	0,1	0,0	0,1	0,0	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0
$\delta T_{\epsilon-\mathrm{fert}}$	0,3	0,0	0,3	0,0	0,2	0,0	0,2	0,0	0,1	0,0	0,1	0,0
$\delta T_{\epsilon-\mathrm{noniso}}$	0,5	0,1	0,5	0,1	0,3	0,0	0,6	0,1	0,5	0,0	0,2	0,0
δT_{rep}	3,0	3,7	5,0	9,4	9,0	**21,1**	4,0	5,0	13,0	**18,4**	20,0	**21,6**
δT_{find}	0,3	0,0	0,4	0,1	1,2	0,4	0,3	0,0	0,4	0,0	12,0	7,8
δT_{form}	1,6	1,1	3,3	4,1	1,3	0,4	1,8	1	1,4	0,2	23,7	**30,3**
T_{fp}	15,5	100	16,3	100	19,6	100	17,9	100	30,5	100	43,0	100

4 Zusammenfassung

Im Artikel wurde ein Dreifach-Fixpunktstrahler vorgestellt, der erstmalig die Fixpunktsubstanzen Zinn, Zink und Aluminium in einer Bauform enthält. Mit dieser Kombination ist es nun möglich eine Kalibrierung eines Strahlungsthermometers an drei direkt auf die ITS-90 rückführbaren Temperaturen, in einem kontinuierlichen Kalibriervorgang, an nur einem Referenzstrahler durchzuführen. Ausgehend von numerischen Berechnungen zur Optimierung der Bauform, über Messungen mit einem Transferstrahlungsthermometer LP5 mit InGaAs-Detektor konnte abschließend in einem Messunsicherheitsbudget gezeigt werden, das die Messunsicherheit der Fixpunkttemperaturen bei Strahlungstemperaturmessungen bei Wellenlängen um $1,5\,\mu\mathrm{m}$ im Bereich von $31\,\mathrm{mK}$ bis $86\,\mathrm{mK}$ liegt (s. Tabelle 4) und damit die Werte von herkömmlichen Einzel-Fixpunktstrahlern hoher metrologischer Präzision erreicht. Der Dreifach-Fixpunktstrahler stellt somit eine einfachere, kostengünstigere und

metrologisch gleichwertige Alternative zu diesen dar.

Liquidustemperaturen ($k = 2$)	Solidustemperaturen ($k = 2$)
$T_{\mathrm{fp}}(\mathrm{Sn}) = 231,919\,^{\circ}\mathrm{C} \pm 61\,\mathrm{mK}$	$T_{\mathrm{fp}}(\mathrm{Sn}) = 231,914\,^{\circ}\mathrm{C} \pm 86\,\mathrm{mK}$
$T_{\mathrm{fp}}(\mathrm{Zn}) = 419,377\,^{\circ}\mathrm{C} \pm 39\,\mathrm{mK}$	$T_{\mathrm{fp}}(\mathrm{Zn}) = 419,407\,^{\circ}\mathrm{C} \pm 36\,\mathrm{mK}$
$T_{\mathrm{fp}}(\mathrm{Al}) = 660,285\,^{\circ}\mathrm{C} \pm 31\,\mathrm{mK}$	$T_{\mathrm{fp}}(\mathrm{Al}) = 660,263\,^{\circ}\mathrm{C} \pm 33\,\mathrm{mK}$

Literatur

[1] J. Hollandt, B. Friedrich, B. Gutschwager, D. Taubert und J. Hartmann, "High accuracy radiation thermometry at the national metrology institute of germany, the ptb", *High Temperatures High Pressures* **35/36**, S. 379–415, 2003/2004.

[2] B. Chu, H. McEvoy und G. Machin, "Use of an ingaas radiation thermometer to verify the accuracy of the npl blackbody reference sources from 156°C to 600°C", in *Temperature: Its Measurement and Control in Science and Industry*, D. C. Ripple (Hrsg.), **7**, S. 571–576, AIP, American Institute of Physics, (Chicago), 2003.

[3] B. Khlevnoi, S. Ogarev, V. Sapritskii, M. Morozova, S. P. Samoilov, M. Sakharov, V. Khromchenko und V. Shapoval, "Reference blackbody sources in the 100-3500 K range for precision measurements in radiometry, photometry, and optical thermometry", *Measurement Techniques* **48**(11), S. 7, 2005.

[4] W. Bauer und A. Moldenhauer, "Emissivities of ceramics for temperature measurements", in *Thermosense XXVI*, D. Burleigh, E. Cramer und G. Peacock (Hrsg.), *Proceedings of SPIE* **5405**, S. 340, SPIE, Society of Photographic Instrumentation Engineers, (Orlando), 2004.

[5] M. Schalles und Blumröder, G., *Calculation of the Effective Emissivity of Blackbodies Made of Alumina*. , in *Ilmenau Scientific Colloquium*, (Ilmenau), to be published 2011.

[6] E. Schreiber, U.-P. Arlt und G. Neuer, "Optimierung eines Strahlungsthermometers für Temperaturen von 230°C bis 1500°C", in *Temperatur 2003, VdI Berichte 1784*, S. 143–154, VDI Verlag GmbH, (Berlin), 2003.

[7] M. Schalles, *Dreifach-Fixpunktstrahler zur Kalibrierung von Strahlungsthermometern*. dissertation thesis, Technische Universität Ilmenau, 2010.

Unsicherheit bei der Datenfusion dimensioneller Messungen

Maryna Galovska, Marcus Petz, Rainer Tutsch

Institut für Produktionsmesstechnik, TU Braunschweig
Tel. 0531-3917043, e-mail: maryna.galovska@tu-bs.de

Kurzfassung

Bei der Auswertung von Messdaten wird eine detaillierte Analyse der Messunsicherheit gefordert. Die Fusion dimensioneller Daten mehrerer Sensoren umfasst Transformationen und Ausgleichsrechnungen. Die resultierende Messunsicherheit ist häufig nicht direkt bestimmbar, kann aber mit der Monte-Carlo Methode abgeschätzt werden. Anhand mehrerer Simulationsrechnungen zur Datenfusion für die Formmessung eines Zylinders werden Effekte aufgezeigt, die die resultierende Unsicherheit entweder vergrößern oder verkleinern.

1 Aufgaben und Realisierungsarten der Datenfusion in der dimensionellen Messtechnik

Datenfusion, definiert als das Zusammenführen der Daten unterschiedlicher Quellen zur Maximierung des Informationsgehalts [1], kommt heute in einer Vielzahl von Anwendungen in der dimensionellen Messtechnik zum Einsatz. Damit lassen sich die Effizienz und der Anwendungsbereich der Methoden sowie die Wiederholpräzision und die Messabweichung verbessern. Abhängig von der zu lösenden Aufgabe und den gewählten Quellen kann Datenfusion in verschiedenen Phasen der Messdatenverarbeitung eingesetzt werden. Hinsichtlich der Ausgangsgrößen unterscheidet man drei Ebenen der Datenfusion (Bild 1): die Rohdatenebene, die Merkmalsebene und die Entscheidungsebene [1, 2]. Eine Klassifizierung ist auch anhand der Eingangs-/Ausgangspaarungen möglich [3]. Bei der Konfiguration der Quellen unterscheidet man zwischen kompetitiven, komplentären und kooperativen Quellen [2].

Die Ziele der Datenfusion sind einerseits die Datensammlung, andererseits die Datenverbesserung um *zusätzliche* oder *genauere* Kenntnisse über physikalische Größen zu erhalten [4]. Datenfusion zur *Datensammlung* kann als ein Weg zur

Kombination von Daten unterschiedlichen Inhalts angesehen werden. Die Notwendigkeit für die Datenfusion ergibt sich oft aus den Eigenschaften des Messobjekts wie Größe, Form, Oberflächenbeschaffenheit usw. Typische Anwendungen sind Messungen auf unterschiedlichen Größenskalen, Messungen unterschiedlicher Merkmale oder die Erfassung großer Messbereiche.

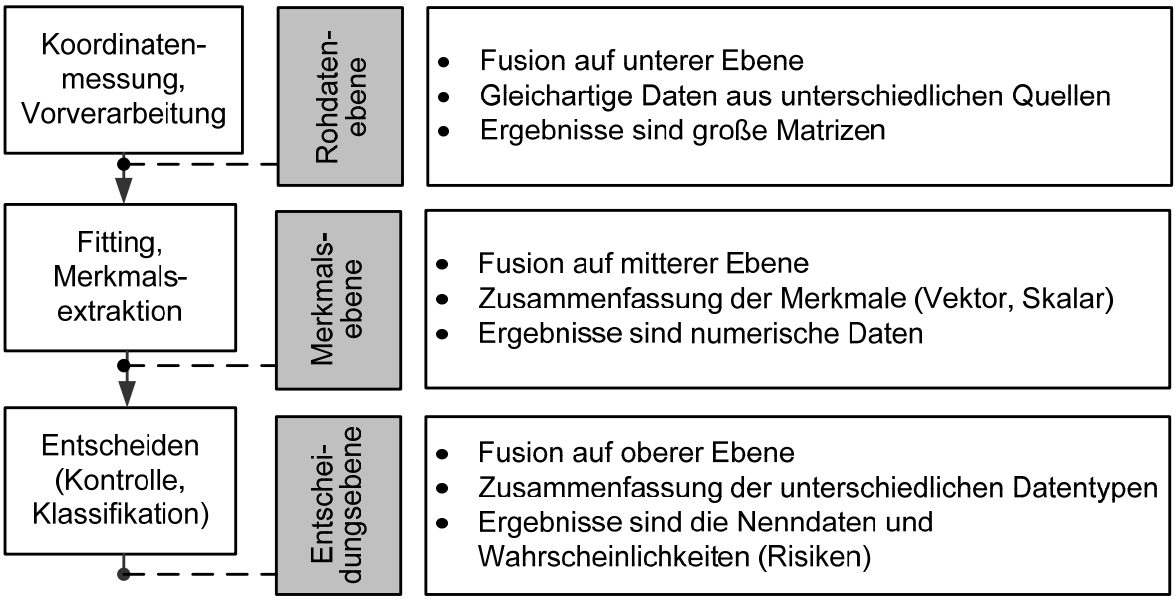


Bild 1: Ebenen der Datenfusion

Als Beispiel wird in [5] ein Ansatz zur mehrskaligen Messung beschrieben, bei dem optische Sensoren mit unterschiedlicher Auflösung und unterschiedlichem Messbereich zur Erfassung von Mikro- und Makrogestalt verwendet werden. Die Multisensor-Datenfusion zur Datensammlung wird meist unter Verwendung kooperativer Sensoren auf der Rohdatenebene durchgeführt, wobei der Prozess der *Registrierung* von entscheidender Bedeutung ist.

Die grundlegende Idee zur *Datenverbesserung* durch Datenfusion ist der Einsatz *redundanter Messungen* zur Reduzierung der Messunsicherheit. Entsprechend dem Modell des Messergebnisses gemäß ISO 5725 ($y=m+B+e$, mit m – Schätzwert für den Messwert, B – systematische Abweichung, e – zufällige Abweichung), verringert die Mittelung über mehrere Messungen die zufällige Komponente der Abweichung. Zur Verringerung der unbekannten (nicht korrigierten) systematischen Komponente müssen Vergleichspräzisionsmessungen durchgeführt werden. Definitionsgemäß bedeutet das die Durchführung der Messungen an demselben oder einem ähnlichen Objekt bei Variation verschiedener Bedingungen wie Ort, Bediener oder Messgerät. Dieser Typ der Datenfusion erhöht ebenfalls die Robustheit der Ergebnisse. Bei der Multisensordaten-

fusion zur Datenverbesserung werden üblicherweise kompetitive und komplementäre Sensoren eingesetzt.

2 Bestimmung der Messunsicherheit

Die im grundlegenden Normungsdokument GUM (ISO Guide to the Expression of Uncertainty in Measurement) vorgeschlagene Modellierung einer Messung basiert auf eindimensionalen und explizit formulierbaren Ausgangsfunktionen. In der Koordinatenmesstechnik ist die Ausgangsgröße oft mehrdimensional und das Modell ist nicht als Funktion sondern als Berechnungsvorschrift gegeben. Oft werden Ausgleichsrechnungen oder Optimierungsoperationen durchgeführt. Die Messunsicherheit mehrdimensionaler Messwerte wird oft durch Ellipsoide dargestellt. Numerisch beschrieben wird sie durch die Kovarianzmatrix, die die Parameter der mehrdimensionalen Wahrscheinlichkeitsverteilung charakterisiert. Die Monte-Carlo Methode (ISO/IEC Guide 98-3/Suppl. 1) hat sich bei der Anwendung auf beliebige Modelle auch im mehrdimensionalen Fall bewährt.

Bei der Verarbeitung von Messdaten muss die damit verbundene Transformation der Messunsicherheit detailliert analysiert werden. Die Messunsicherheit enthält Beiträge des Messinstruments, aber auch der gewählten Messstrategie und der Datenverarbeitungsalgorithmen einschließlich der Datenfusion. Das Problem der Messunsicherheitsfortpflanzung wird in [2] angesprochen, erfordert aber detaillierte Forschungsarbeit. In dieser Studie werden die verschiedenen Stufen der Datenverarbeitung untersucht und die Komponenten der Messunsicherheit werden analysiert. Zuerst wird eine Ausgangsunsicherheit einzelner Messpunkte aus der Unsicherheit der einzelnen Koordinaten abgeleitet. Die 3D-Unsicherheit von Messpunkten für optische Messverfahren wird in [6] charakterisiert. Nachfolgend werden die Unsicherheiten der Koordinaten jedes Einzelpunkts und die Unsicherheiten u_i der Punkte innerhalb der Messpunktwolke vereinfachend als gleich und völlig unkorreliert angenommen.

In der Datenvorverarbeitung sollen die Daten “bereinigt”, d.h. von Ausreißern befreit, gefiltert und gegebenenfalls in ein anderes Datenformat transformiert werden. Das Problem der Unsicherheitsfortpflanzung im Falle einer Filterung wird in [7] untersucht. Die nächsten Verarbeitungsschritte enthalten meist Ausgleichsrechnungen: Ausrichten, Registrieren, Merkmalsextraktion usw. Das Problem der Bestimmung von Messgrößen aus Punktwolken unter Berücksichtigung der Unsicherheit der Einzelpunkte wird in [6, 8] angesprochen.

3 Koordinatentransformationen

3.1 Ansätze

Die Datenverarbeitung von Punktwolken bei Multisensormesssystemen erfordert Koordinatentransformationen: einerseits für die Ausrichtung, andererseits für die Transformation der Punktwolken in ein gemeinsames Koordinatensystem (Registrierung). Dafür muss folgendes festgelegt werden: (I) die Anpassungsmethode (matching) und die Optimierungskriterien, (II) die Art der Transformation (z.B. Starrkörper-, affine oder polynomiale Transformation) und eine Anzahl von Transformationsparametern, (III) ein Referenzkoordinatensystem, das eine einfache Ausrichtung erlaubt und die durch die Transformation verursachte Unsicherheit reduziert. Für die Anpassung können die Sensorpositionen bei der Datenaufnahme verwendet oder Messpunkte in überlappenden Bereichen der Punktwolken angeglichen bzw. Referenzmarken zur Deckung gebracht werden.

Eine Starrkörpertransformation in einem kartesisches Koordinatensystem ist durch einen Verschiebungsvektor $\vec{x}_0$ und eine 3x3 Rotationsmatrix R darstellbar: $\vec{x}_{tr} = R \cdot \vec{x} + \vec{x}_0$. Zur Berechnung der Parameter können verschiedene Optimierungskriterien angewandt werden, z.B. die Minimierung der quadratischen Abstandssumme oder die Maximierung der Kreuzkorrelationsfunktion im Überlappbereich. In jedem Fall sind die so bestimmten Transformationsparameter nicht exakt, sondern mit einer Unsicherheit behaftet (Bild 2).

3.2 Simulation

Eine Monte-Carlo Simulation der Messung von drei punktförmigen Objektmerkmalen mit zwei Sensoren liefert jeweils drei Punktwolken. Die Transformation der beiden Datensätze aufeinander zeigt eine signifikante Vergrößerung der Messunsicherheit der transformierten Punkte (Bild 2). Dabei wurde vereinfachend die Unsicherheit in Form unkorrelierter zufälliger Ereignisse modelliert.

Im nächsten Schritt wurde die Auswirkung der Unsicherheit der Transformationsparameter auf die Bestimmung von Geometriemerkmalen untersucht. Als elementares ebenes Modell wurde ein Kreis mit dem Radius R und den Mittelpunktskoordinaten (x_0, y_0) verwendet. Zwei unabhängige Messungen sollen jeweils eine Punktwolke liefern. Die erste Punktwolke soll so transformiert werden, dass sie mit der zweiten möglichst zur Deckung kommt. Dazu wird ein

least-squares-fit anhand von 5 unsicherheitsbehafteten Referenzpunkten der zweiten Punktwolke durchgeführt, was unsicherheitsbehaftete Transformationsparameter ergibt. Nach der Transformation ist die Unsicherheit der Punkte aus Punktwolke 1 aufgrund der Unsicherheit der Transformationsparameter erhöht. In Bild 3 ist erkennbar, dass das Ergebnis empfindlich von der Anordnung der 5 Referenzpunkte abhängt. Ebenfalls dargestellt ist die entsprechende 3D-Simulation zur Transformation einer Kugelfläche.

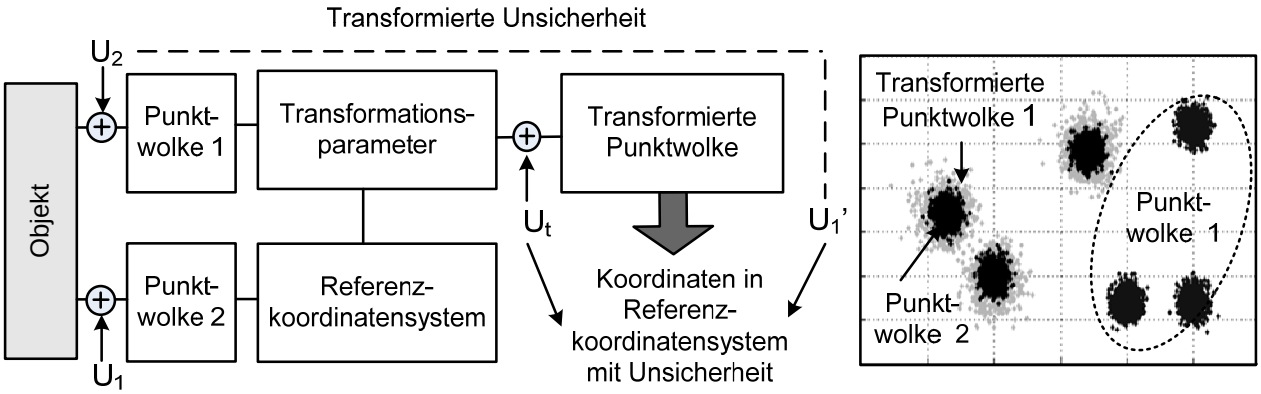


Bild 2: Berücksichtigung der Unsicherheit der Koordinatentransformation

Offensichtlich hängt die Unsicherheit der Transformationsparameter und damit die der transformierten Koordinaten von der Anzahl und der Anordnung der Referenzpunkte ab, die zur Berechnung der Transformationsparameter verwendet werden. Bei gleichmäßiger Anordnung der Referenzpunkte auf dem Kreisumfang reduziert die Erhöhung der Zahl der Referenzpunkte von 5 auf 6 die Unsicherheit der Radiusberechnung um ca. 30%.

Es ist wichtig, zu beachten, dass obwohl die beiden Punktwolken als unkorreliert angenommen wurden, nach der Koordinatentransformation aufgrund der Kopplung über die Unsicherheitsfortpflanzung eine Korrelation vorliegt. Simulationsrechnungen für unterschiedliche Fälle wurden durchgeführt und einige Ergebnisse sind in Bild 4 dargestellt.

Die Diagramme zeigen die Korrelation zwischen den Koordinaten der transformierten Punkte. Im idealen *Fall 1,* bei dem die Referenzpunkte und folglich die Transformationsparameter ohne Unsicherheit bekannt sind, ist keine Korrelation erkennbar. Eine signifikannte Unsicherheit der Referenzpunkte führt zu einer ausgeprägten Korrelation, die von der Anordnung der Referenzpunkte abhängt. Bei gleichmäßiger Anordnung ist die Korrelationsfunktion symmetrisch

(*Fall 2*). Ungleichmäßige Anordnung der Referenzpunkte führt zu ungleicher Unsicherheit der einzelnen transformierten Messpunkte (Bild 3) und damit zu einer asymmetrischen Korrelationsfunktion (*Fall 3*).

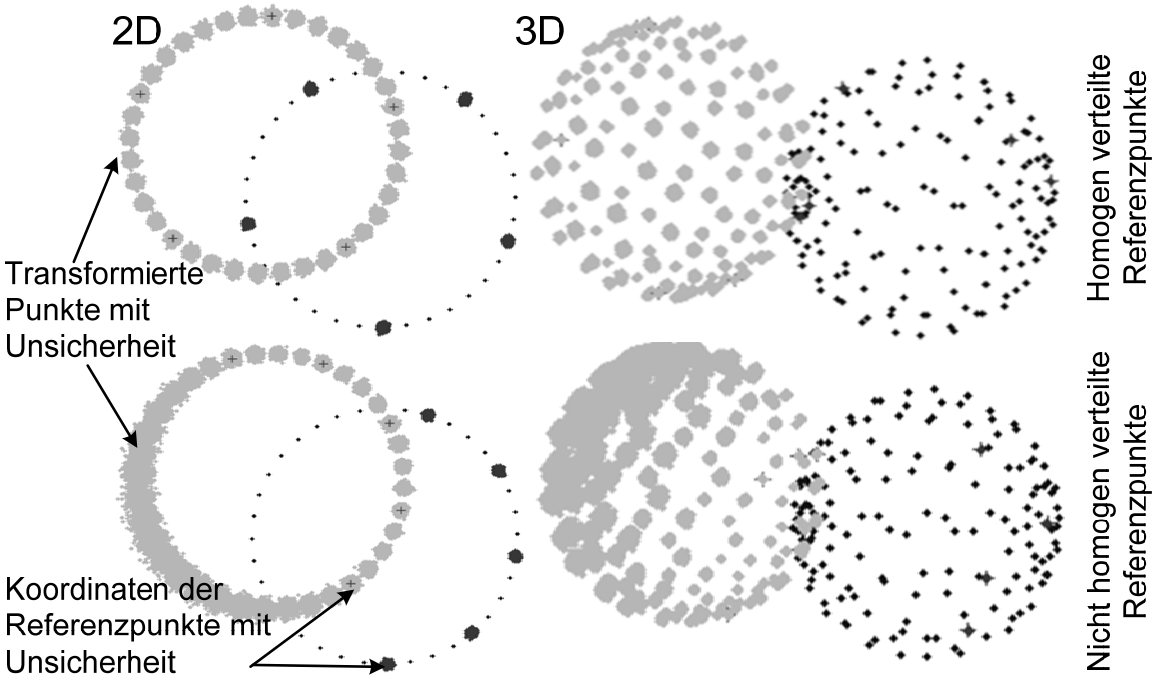


Bild 3: Vergrößerung der Unsicherheit durch die Koordinatentransformation für ein 2D und ein 3D Modell

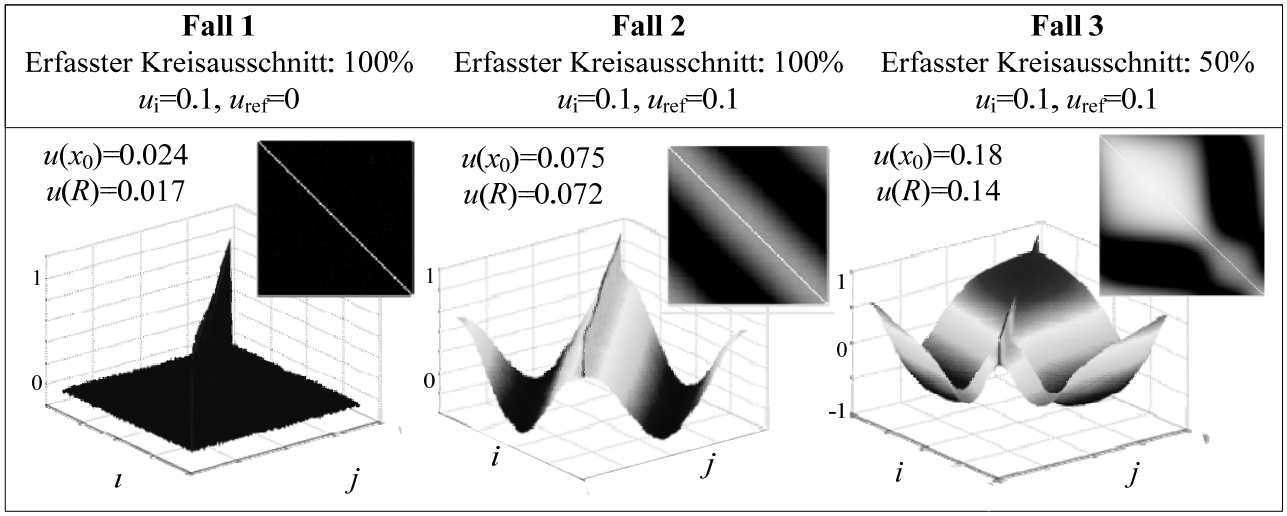


Bild 4: Kreuzkorrelation infolge der Koordinatentransformation

3.3 Anwendungsbeispiel: Stitching von Punktwolken

Das Zusammensetzen eines Bildes (2D oder 3D) aus einzelnen Teilbildern wird in der Fotografie und der Messtechnik mit dem Begriff „Stitching“ bezeichnet. In der Nomenklatur der Datenfusion handelt es sich um die Anwendung „Datensammlung“.

Für unser Anwendungsbeispiel wurde ein Verfahren zur kostengünstigen und gleichzeitig genauen Messung der Geometrie eines Zylinders entwickelt. Ein optisches Streifenprojektionsmessgerät nach dem Prinzip der passiven Triangulation erfasst einen Ausschnitt (Patch) der Oberfläche des Zylinders. Dieser liegt auf einer Vorrichtung aus vier Präzisionskugeln, deren Anordnung mit einem Koordinatenmessgerät gemessen worden ist. Der Zylinder wird manuell schrittweise gedreht, bis die Oberfläche vollständig mit einander überlappenden optisch gemessenen Ausschnitten abgedeckt ist. Die Aufgabe besteht darin, alle gemessenen Patches zu einer vollständigen Abbildung der Zylinderoberfläche zusammenzuführen. Dabei kann die mit der Unsicherheit der KMG-Messung behaftete bekannte Geometrie der Vorrichtung genutzt werden; die relativen Drehlagen sind jedoch unbekannt.

Für den ersten Lösungsansatz wird die Stirnseite des Zylinders mit einem stochastischen Muster versehen und mit einem zusätzlichen Kamerasystem aufgenommen (Bild 5). Mittels Bildkorrelation wird die Drehlage des Zylinders berechnet [9].

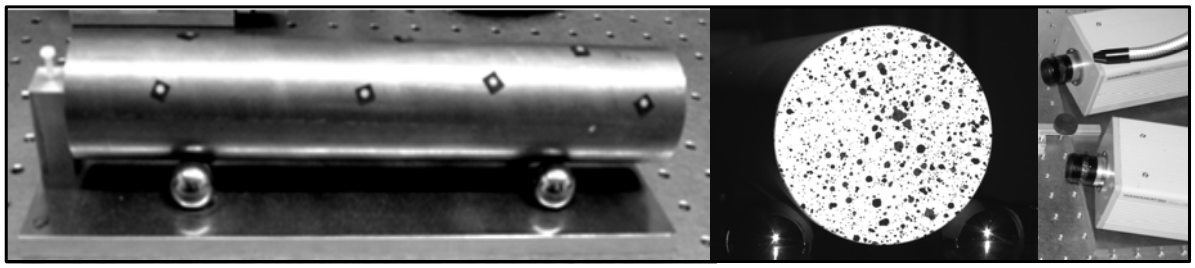


Bild 5: Messung der Form eines Zylinders

Für N Patches und n_i Punkte ist folgende Optimierungsaufgabe zu lösen:

$$\min\left(\sum_{i=1}^{N}\sum_{j=1}^{n_i}\left(x_{ij}-\left(k_{x_i}\cdot z_{ij}+a_i\right)\right)^2+\left(y_{ij}-\left(k_{y_i}\cdot z_{ij}+b_i\right)\right)^2-R^2\right)$$

Mit den gefundenen Optimierungsparametern werden die Punktwolken der Patches unter Anwendung einer Fusion mit den Daten der Winkelsensorik transformiert (ϕ_i: gemessener Winkel): Für N Patches und n_i Punkte ergibt sich für jeden Patch die Optimierungsaufgabe:

$$x_{ij}^{tr} = (x_{ij} - k_{xi} \cdot z_{ij} - a_i)) \cdot \cos(\phi_i) - (y_{ij} - k_{yi} \cdot z_{ij} - b_i)) \cdot \sin(\phi_i)$$
$$y_{ij}^{tr} = (x_{ij} - k_{xi} \cdot z_{ij} - a_i)) \cdot \sin(\phi_i) + (y_{ij} - k_{yi} \cdot z_{ij} - b_i)) \cdot \cos(\phi_i)$$

Nach der Durchführung der Transformation können die Messpunkte für ein 3D-Modell der Zylinderoberfläche verwendet werden.

Für einen zweiten Lösungsansatz werden Kreismarken auf den Zylindermatel geklebt und für ein merkmalsbasiertes Stitching ausgewertet. Sequentielles Stitching führt allerdings zur Kumulation der Unsicherheitsbeiträge. Dies kann durch eine gleichzeitige Optimierung für alle Patches vermieden werden, wie sie in [10] für die Messung einer Präzisionskugel verwendet wurde. Als Alternative können zusätzliche ortsfeste Marken für die Definition eines allen Patches gemeinsamen Referenzkoordinatensystems verwendet werden.

4 Unsicherheit der fusionierten Daten

Nach der Registrierung werden die Daten gemäß der jeweiligen Messaufgabe ausgewertet. Verschiedene Analysetechniken für die Datenfusion unsicherheitsbehafteter Daten sind verfügbar [11]: Wahrscheinlichkeitstheorie (Bayes-Ansatz), Möglichkeitstheorie (Fuzzy Sets), Dempster-Shafer's Evidenztheorie, Rough Sets Theorie nach Pawlak. Am häufigsten wird ein probabilistischer Ansatz mit einer Gewichtung einzelner Beobachtungen angewandt. Die Merkmalsextraktion bei der Fusion auf Rohdatenebene wird durch eine gewichtete Ausgleichsrechnung repräsentiert und für die Fusion auf Merkmalsebene wird meist eine gewichtete Mittelung angewandt (Fig. 6a):

$$\tilde{x}_F = \sum_{i=1}^{m} \tilde{x}_i w_i, \quad \sum_{i=1}^{m} w_i = 1, \quad u_f = \left(\sum_{i=1}^{m} w_i \right)^{-\frac{1}{2}}, \qquad (1)$$

mit $\tilde{x}_i$: Schätzung für die *i–te* Quelle mit Standardunsicherheit u_i,
w_j: Gewichtungen, u_f: Standardunsicherheit des Schätzwertes

Richtig gewählte Gewichtungen der Quellen sind entscheidend für den Nutzen der Datenfusion. Die Gewichtung kann global oder lokal erfolgen.

4.1 Simulation

In [8] werden Ergebnisse der Monte-Carlo Simulation der Messung eines Kreises vorgestellt, wobei sowohl zufällige als auch nicht erkennbare systematische Abweichungen untersucht wurden (Bild 6b). Zufällige Effekte wurden mittels eines Zufallsgenerators für jeden Messpunkt generiert. Dagegen wird der unbekannte systematische Effekt durch eine additive Abweichung modelliert, die für jeden Punkt einer Iteration identisch ist, aber von Iteration zu Iteration stochastisch variiert. Durch die daraus resultierende Korrelation der Unsicherheiten der Messpunkte wird das Modell realitätsnäher.

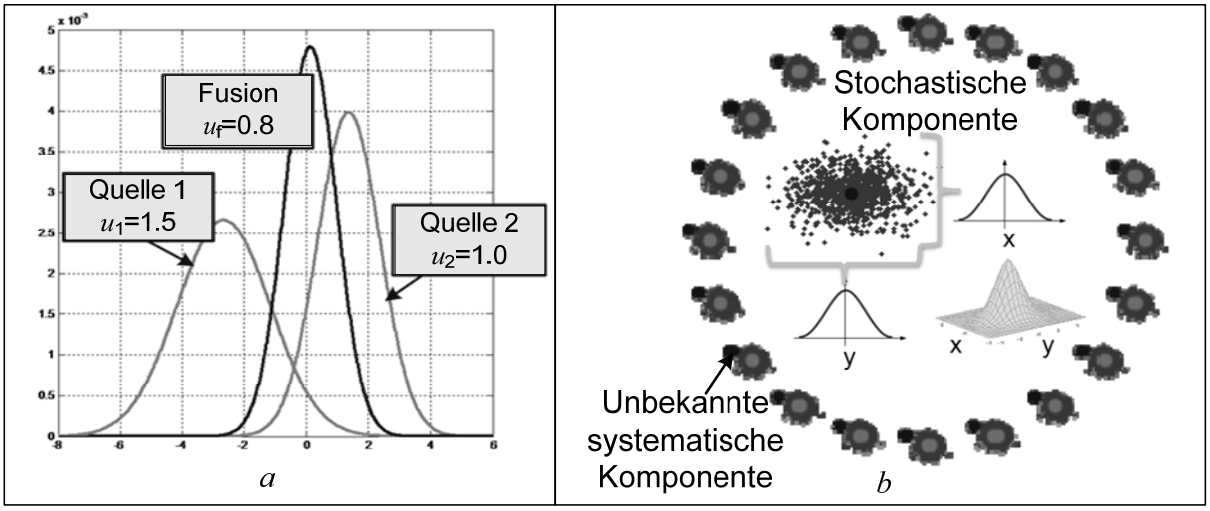


Bild 6: Grundlagen der Unsicherheitsanalyse: (*a*) probabilistiche Fusion, (*b*) Monte-Carlo Simulation für die Messung von Kreismerkmalen (R, x_0, y_0)

Bild 7 zeigt das Ergebnis der Simulation einer Messung von Kreismerkmalen mit zwei Sensoren unterschiedlicher Charakteristik. Der erste Sensor liefert eine große Anzahl unsicherer Messpunkte, während der zweite Sensor wenige Messpunkte liefert, diese jedoch mit hoher Genauigkeit. Die Datenfusion lässt sich grafisch als Schnitt der beiden Unsicherheitsellipsoide darstellen und führt zu einer signifikanten Verbesserung des Messergebnisses gegenüber beiden Einzelmessungen.

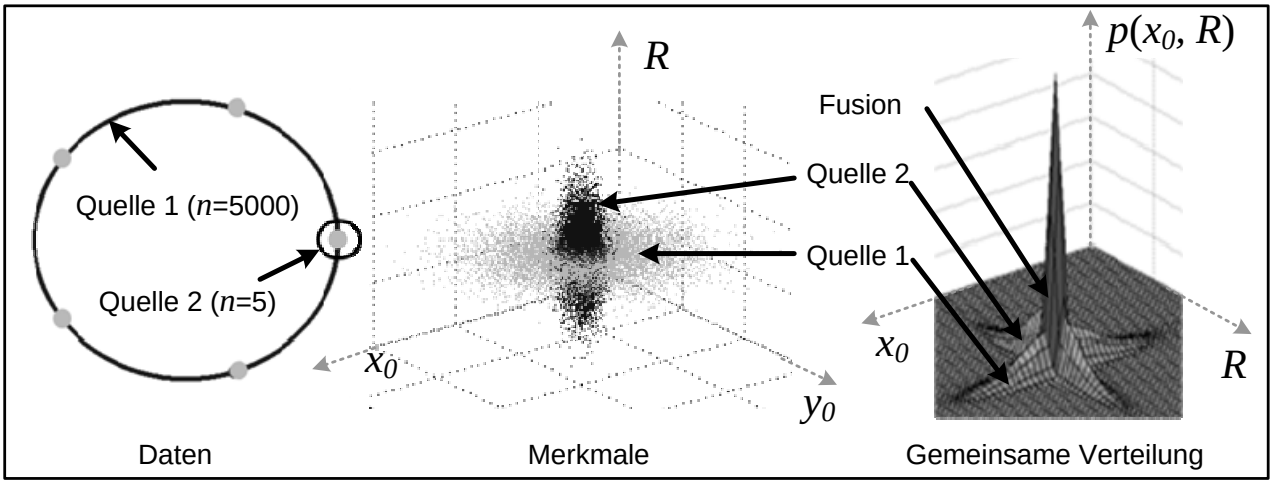


Bild 7: Simulation der Datenfusion bei der Messung von Kreismerkmalen.

4.2 Beispiele

Um den Einfluss der nicht erkennbaren systematischen Abweichungen zu reduzieren, werden bei Wiederholmessungen verschiedene Parameter des Messsystems variiert. Exemplarisch werden in [12] Aufnahmen einer Szene mit unterschiedlichen Beleuchtungsparametern, Kamerapositionen, Fokuseinstellungen, spektralen Filterungen usw. fusioniert.

Gemeinsam mit Dr. Jusko (PTB) wurde dieser Ansatz auf die Messung eines Präzisionszylinders angewandt. Die Messung mit einem Formtester im „Helical-Scan-Mode“ (spiralförmige Antastung) liefert eine hohe Messpunktdichte und ergab scheinbar einen konischen Formfehler (Bild 8). Die Fusion mit einer geringen Anzahl (15) hochgenauer Durchmessermessungen führte zu einer Korrektur der systematischen Abweichung (wahrscheinliche Ursache: Zylinderachse nicht parallel zu Drehachse). In diesem Fall hatten die zu fusionierenden Datensätze sehr unterschiedliche Unsicherheiten, was sehr unterschiedliche Gewichtungen gemäß (1) zur Folge hat:

$$\tilde{x}_F = (w_1 x_1 + w_2 x_2), \quad u_1 << u_2 \Rightarrow w_1 >> w_2 \Rightarrow \tilde{x}_F \cong x_1.$$

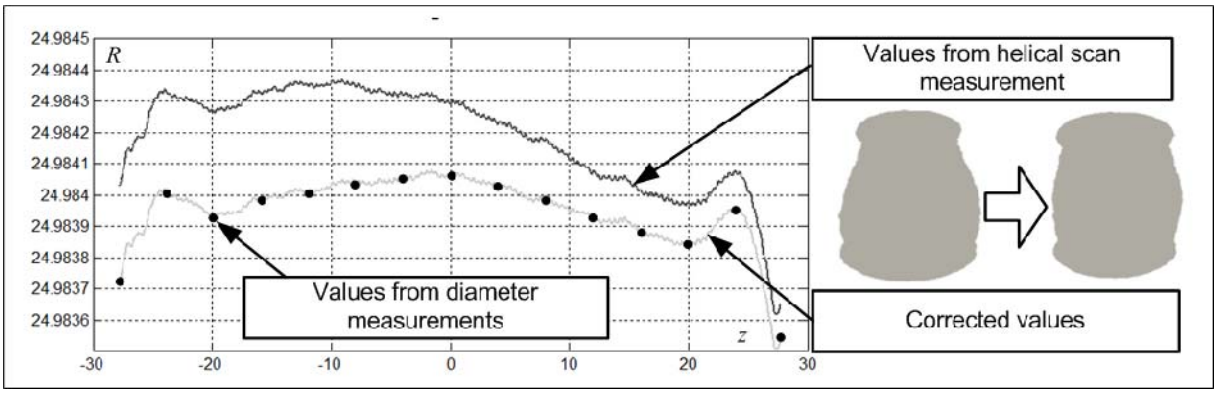


Bild 8: Fusion der Daten des Helical Scan und der Durchmessermessung

5 Zusammenfassung

Die Transformation der ursprünglichen Unsicherheit in den verschiedenen Phasen der Messdatenauswertung vergrößert oder verkleinert die resultierende Messunsicherheit. Die Kenntnis dieser Zusammenhänge hilft bei der gezielten Beeinflussung der Unsicherheitstransformation und ermöglicht eine Optimierung des Messprozesses ohne Reduzierung des Unsicherheitsbeitrags des Messgerätes. Die Bedeutung einer sorgfältigen Bestimmung der Transformationsparameter wurde gezeigt. Die Verbesserung der Qualität des Messergebnisses durch die Datenfusion ist auf die Redundanz und geeignete Gewichtung der Beobachtungen zurückzuführen.

Danksagung

Diese Arbeit wurde durch die Braunschweig International Graduate School of Metrology (IGSM) unterstützt.

Literatur

[1] Girão, P. M.; Dias Pereira, J. M.; Postolache, O.: Multisensor Data Fusion and Its Application to Decision Making - Chapter in Advanced Mathematical & Computational Tools in Metrology VII, P Ciarlini et al., World Scientific Publishing, New Jersey, 2006.

[2] Weckenmann, A.; Jiang, X.; Sommer, K.-D.; Neuschaefer-Rube, U.; Seewig, J.; Shaw, L.; Estler, T.: Multisensor data fusion in dimensional Metrology, p. 701-721. In: CIRP Annals - Manufacturing Technology 58, 2009.

[3] Nachtigall, L. ; Puente León, F. ; Pérez Grassi, A.: Konzepte zur Fusion von Bildserien mit variabler Beleuchtungsrichtung, Forum Bildverarbeitung., S.: 325 – 336 Hrsg.: F. Puente León, Verlag: KIT Scientific Publishing, Karlsruhe, 2010.

[4] Ruser, H. ; Puente León, F.: Informationsfusion – eine Übersicht, Technisches Messen: TM 74/3, S.: 93-102, 2007.

[5] Shaw, L.; Weckenmann, A.: Optical 3D-Characterization for multiscale workpieces. - 10th ISMQC (05.-09.09.2010, Osaka, Japan), 2010 - Proceedings on CD-ROM.

[6] Tutsch, R. ; Petz, M.: Ermittlung der Oberflächengeometrie aus Meßpunktewolken in der optischen 3D-Meßtechnik, Expertenforum „Informationsfusion in der Mess- und Sensortechnik“, S.: 127 – 155, Hrsg.: J. Beyerer, F. Puente León, K.-D. Sommer, Universitätsverlag Karlsruhe, 2006.

[7] Krystek, M.: Measurement uncertainty propagation in the case of filtering in roughness measurement, Meas. Sci. Technol. 12, p. 63-67, 2001.

[8] Tutsch, R.; Petz, M.; Galovska, M.: Messstrategie für zylindrische Werkstücke ohne Verwendung einer Rotationsachse. In: Tagungsband VDE-Wissensforum, Form- und Konturmessung, Leonberg, 15.-16.06.2010

[9] Tutsch, R.; Petz, M.; Galovska, M.: Strategy and uncertainty evaluation of cylindrical and spherical measurements without rotation axis, Manufacturing 2010, p. 194, Institute of Mechanical Technology, Poznan University of Technology, Poznan, Poland, 2010.

[10] Bartl, G.; Krystek, M.; Nicolaus, A.; Giardini, W.: Interferometric determination of the topographies of absolute sphere radii using the sphere interferometer of PTB, Meas. Sci. Technol. 21 (2010).

[11] Chong, C.Y., Mori, S.: Convex combination and covariance intersection algorithms in distributed fusion, Proceedings of the 4th Inter. Conf. on information fusion, pp. WeA2-11-18, vol. 1, Montreal, Canada, 2001.

[12] Heizmann, M.; Puente León Fusion von Bildsignalen, Technisches Messen: TM 74/3, S.: 130-137, 2007.

Ermittlung einer aufgabenspezifischen Messunsicherheit für Verzahnungsmessungen

Kerstin Rost[1], Frank Härtig[1], Gert Goch[2]

[1]Physikalisch-Technische Bundesanstalt, Braunschweig und Berlin
[2]Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft, BIMAQ, Universität Bremen
Tel.: 0531 5925322, email: Kerstin.Rost@ptb.de

Kurzfassung

Die Kenntnis der Messunsicherheit ist von entscheidender Bedeutung für die sichere Beurteilung eines Messergebnisses. Übereinstimmungsprüfungen mit Spezifikationen können nur durchgeführt werden, wenn die Eignung des eingesetzten Messverfahrens sichergestellt ist. Je höher die Ansprüche an die Qualität eines Werkstücks sind, desto größer ist der Bedarf an einer eindeutigen und rückgeführten Qualifizierung der eingesetzten Messtechnik.

Gerade in der dimensionellen Fertigungsmesstechnik ist die Ermittlung der Messunsicherheit über experimentelle oder analytische Methoden mit erheblichem Aufwand verbunden. Dies gilt insbesondere bei der Messung von Zahnrädern mit ihrer breiten Produktpalette. Um den erforderlichen Eignungsnachweises von Verzahnungsmessgeräten zu vereinfachen, wurde nun basierend auf dem in der PTB entwickelten „Virtuellen Koordinatenmessgerät“ ein Verfahren zur Ermittlung der Messunsicherheit durch Simulation entwickelt. Die hierbei durchgeführten Betrachtungen zur Berechnung der Messunsicherheit werden nachfolgend beschrieben.

1 Einleitung

In Deutschland werden jährlich mindestens 150 Millionen Zahnräder hauptsächlich für den Automobilbau produziert [1]. Durch die Umbrüche in der Energiepolitik und den damit verbundenen Ausbau von Windenergieanlagen (WEA) wächst darüber hinaus der Bedarf an Großverzahnungen. In beiden Gebieten ist die Qualität der Zahnräder von entscheidender Bedeutung. Es ist da-

her unumgänglich für die zur Messung der Zahnradgeometrie eingesetzten Koordinaten- und Verzahnungsmessgeräte ebenfalls einen Qualitätsnachweis zu führen, d. h. für jedes Prüfmerkmal eine aufgabenspezifische Messunsicherheit anzugeben [2, 3].

Zur Ermittlung der Messunsicherheit gibt es drei Ansätze, die sich nach dem „Guide to the Expression of Uncertainty in Measurement“ (GUM) [4] richten.

Aufstellen einer komplexen Messunsicherheitsbilanz

Die Aufstellung der analytischen Modellgleichung erfordert die genaue Kenntnis des Messprozesses unter Betrachtung sämtlicher wirksamer Einflussgrößen des Messgerätes, des Werkstückes und der Umgebungsbedingungen. Diese Methode ist sehr aufwändig und erfordert eine hohe messtechnische Expertise, so dass sie nur von wenigen Instituten oder Kalibrierlaboratorien angewandt wird.

Messunsicherheitsermittlung unter Verwendung eines kalibrierten Werkstücks (Substitutionsverfahren) [5]

Diese anwenderfreundliche Methode erfordert, wie der Name bereits sagt, ein kalibriertes Werkstück. Dies ist auch gleich der Nachteil des Verfahrens, da geeignete Normale und deren Kalibrierung kostspielig sind und daher oft nicht zur Verfügung stehen. Insbesondere für Groß- und Mikroverzahnungen existieren nur vereinzelt geeignete Normale, so dass eine aufgabenspezifische Messunsicherheit nicht angegeben werden kann. Dennoch ist dies heutzutage für Verzahnungsmessungen die gebräuchlichste Methode zur Unsicherheitsermittlung .

Abschätzung der Messunsicherheit durch Monte-Carlo-Simulation

Dieses Verfahren ist universell und eignet sich zur Ermittlung einer aufgabenspezifischen Messunsicherheit. Das Verfahren wurde in der PTB entwickelt und ist unter dem Namen „Virtuelles Koordinatenmessgerät“ (VCMM) [6] im Markt bekannt. Seit ca. einem Jahrzehnt ist es ein etabliertes Verfahren, dass die Ermittlung der Messunsicherheit für die Messung von Regelgeometrien in Einzelpunktantastung auf Koordinatenmessgeräten ohne kalibriertes Normal ermöglicht. Es wurde daher als Grundlage für die Entwicklung eines universellen Verfahrens zur Messunsicherheitsermittlung für Verzahnungsmessgeräte genutzt. Besondere Bedeutung hatte neben der mathematischen Modellierung der auf den Verzahnungsmessprozess wirkenden Unsicherheitseinflüsse auch der Nachweis der Konformität der Methode mit dem *GUM Supplement 1*.

2 Messunsicherheitsermittlung durch Simulation

2.1 Prinzip des Virtuellen Koordinatenmessgeräts (VCMM)

Der Grundgedanke des VCMM ist die Nachbildung des Messprozesses durch numerische Simulationen. Das Modell des Messprozesses mit allen auf diesen wirkenden Einflussgrößen bildet wie bei einer klassischen analytischen Unsicherheitsbilanz die Basis für die Ermittlung der Messunsicherheit. Die Berechnung erfolgt jedoch nicht über die Lösung der partiellen Ableitungen der Modellgleichung und die gewichtete quadratische Addition der einzelnen Unsicherheitsbeiträge, sondern durch ein virtuelles Experiment, das die reale Messung unter sehr vielen verschiedenen Bedingungen simuliert.

Eine Besonderheit des VCMM ist, dass das mathematische Modell des Messprozesses die Einflüsse auf *jeden einzelnen* Messpunkt beschreibt. Damit kann es zur Ermittlung einer aufgabenspezifischen Messunsicherheit für beliebige Messaufgaben auf Koordinatenmessgeräten eingesetzt werden. Die Eingangsgrößen des mathematischen Modells werden hierbei individuell für jedes Messgerät erfasst, so dass die Berechnung der Messunsicherheit für die Maschinentypen durchgeführt werden kann, deren kinematisches Fehlermodell mit den hinterlegten Modellen übereinstimmt.

2.2 Modell-Erweiterung für Verzahnungsmessgeräte

Zur Übertragung des Verfahrens auf Verzahnungsmessgeräte wurde ein mathematisch-physikalisches Modell des Verzahnungsmessprozesses aufgestellt. Berücksichtigt wurden die Einflüsse der translatorischen und rotatorischen Maschinenachsen, der Umgebungsbedingungen auf Messgerät und Werkstück, des Tastsystems sowie der Werkstückaufspannung und des Scannings.

Für die Modellierung der kartesischen Achsen von Verzahnungsmessgeräten und des Umgebungseinflusses konnte auf Arbeiten zurückgegriffen werden, die im Rahmen der Entwicklung des VCMM für prismatische Werkstücke auf Koordinatenmessgeräten durchgeführt wurden [7]. Aufgrund der Ähnlichkeit von Koordinaten- und Verzahnungsmessgeräten war hier nur eine Anpassung von Einzeleinflüssen notwendig; beispielsweise der Einfluss transversaler Temperaturgradienten auf die Säule des Verzahnungsmessgerätes.

Für die bei Verzahnungsmessungen zusätzlich wirksamen Unsicherheitseinflüsse durch die Nutzung eines Drehtisches, das scannende Antasten sowie durch die Werkstückaufspannung mussten dagegen neue mathematische Modelle entwickelt und umgesetzt werden. Für jedes dieser Modelle wurden Methoden zur individuellen und auf die SI-Einheiten rückführbaren Erfassung der Einflussgrößen entwickelt [8].

2.3 Softwaretechnische Umsetzung

Die Modellerweiterungen für Verzahnungsmessgeräte wurden im Softwarepaket VCMM-Gear zusammengefasst. Zur Auswertung der Simulationsergebnisse wurde das Konzept des Virtuellen Koordinatenmessgerätes um ein statistisches Auswertemodul erweitert. Es erlaubt die Steuerung und Kontrolle der Messunsicherheitsberechnung über klar definierte und GUM-konforme Methoden. Das Softwarepaket kann über definierte Schnittstellen prinzipiell in jede Messsoftware eingebunden werden. Hierdurch ist die Nutzung aller Informationen und Funktionen möglich, über die die Messgerätesoftware verfügt. Dies betrifft insbesondere die Auswertung der Prüfmerkmale (Kurvenanpassung, Filterung). Damit ist das VCMM-Gear genauso universell einsetzbar wie die entsprechende Messsoftware.

Das Zusammenspiel von Messgerätesoftware, VCMM-Gear und integriertem Auswertemodul wird in Bild 1 veranschaulicht. Der messgeräteunabhängige interne Rechenkern, verwaltet das kinematische Fehlermodell und kommuniziert über eine Schnittstelle mit der Messgerätesoftware. Die über die Schnittstelle übergebenen Messpunkte werden mehrfach entsprechend den implementierten Fehlermodellen verrauscht. Anschließend erfolgt die Auswertung der Prüfmerkmale in der Messgerätesoftware. Sowohl der aus den tatsächlich erfassten Messpunkten ermittelte Messwert (gelb) als auch die n aus den verrauschten Messpunktsätzen berechneten simulierten Messwerte (lila) werden an das PTB-Auswertemodul übergeben und liefern nach der statistischen Auswertung die Messunsicherheit U. Das vollständige Messergebnis wird dann vom Auswertemodul an die Messgerätesoftware zurückgeliefert.

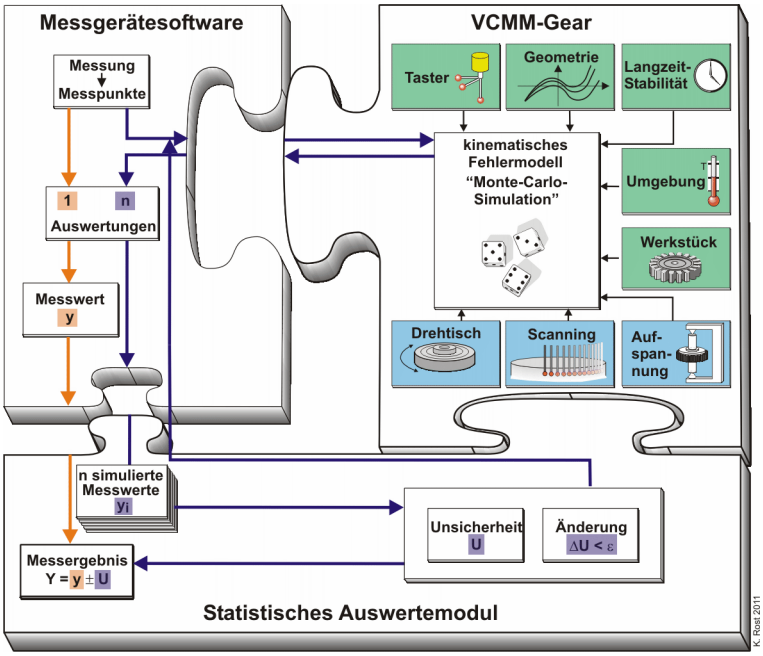


Bild 1: Prinzipskizze der Messunsicherheitsermittlung durch Simulation

3 Berechnung der Messunsicherheit im VCMM-Gear

Mit der Herausgabe des *GUM Supplement 1* (GUM-S1) existiert erstmals eine normative Grundlage für die Messunsicherheitsermittlung durch Monte-Carlo-Simulationen. Damit wurde eine Untersuchung und Anpassung der im ursprünglichen VCMM eingesetzten Berechnungsmethoden im Hinblick auf ihre GUM-Konformität erforderlich.

Im GUM-S1 werden zwei Methoden zur Berechnung der Messunsicherheit aufgeführt. Einerseits kann aus der Standardunsicherheit u der virtuellen Messwerte unter Berücksichtigung eines im Vorfeld festgelegten Erweiterungsfaktors k die erweitere Messunsicherheit U bestimmt werden. Andererseits kann aus der Wahrscheinlichkeitsverteilung der Messgröße Y direkt ein Überdeckungsintervall einer definierten Überdeckungswahrscheinlichkeit p ermittelt werden. Beide aufgeführten Methoden berücksichtigen jedoch nur die Streuung der simulierten Messwerte. Es wird angenommen, dass systematische Unsicherheitsbeiträge entsprechend im Vorfeld korrigiert wurden [9].

Sowohl die Untersuchungen der einzelnen Einflussgrößen auf Verzahnungsmessgeräten [10] als auch frühere Untersuchungen an Koordinatenmessgeräten [6] haben gezeigt, dass bei 3D-Koordinatenmessungen systematische Unsicher-

heitsbeiträge wirken. Ihre Korrektur im Vorfeld ist nicht immer möglich, da die Auswirkung auf die simulierten Messwerte vom Prüfmerkmal abhängt. Sind vergleichbare kalibrierte Werkstücke vorhanden, kann eine nachträgliche Korrektur durch Substitutionsmessungen erfolgen. Erfolgt keine Korrektur, müssen die restlichen systematischen Unsicherheitsbeiträge für eine GUM-konforme Berechnung ebenso behandelt werden wie zufällige Beiträge [11].

Da das VCMM-Gear als universelle Methode zur Berechnung einer aufgabenspezifischen Messunsicherheit für Verzahnungsmessaufgaben entwickelt wurde, müssen *sämtliche vorhandenen* Einflüsse berücksichtigt werden, auch die wirksamen systematischen. Hierzu werden alle Beiträge quadratisch zur kombinierten Standardunsicherheit u_c addiert. Werden die zufälligen Unsicherheitsbeiträge mit u_i bezeichnet und der systematische Beitrag mit b, gilt:

$$u_c = \sqrt{\sum_{i=1}^{n} u_i^2 + b^2} \tag{1}$$

Entscheidend für die praktische Anwendung dieser Formel ist die Bestimmung der systematischen Unsicherheitsbeiträge. Der im VCMM-Gear umgesetzte Ansatz wird am Beispiel einer Kreismessung verdeutlicht. Bild 2 stellt die Wirkung eines konstanten systematischen Messpunktversatzes einem zufälligen Fehler gegenüber. Auf der linken Seite oben ist die Wirkung zufälliger Unsicherheitsbeiträge auf die Messpunkte dargestellt. Die graue Scheibe kennzeichnet den Bereich, in dem die Kreismittelpunkte x_i im Vergleich zum ‚unverrauschten' Referenzmittelpunkt x_{ref} liegen. Darunter ist die sich ergebende Häufigkeitsverteilung der Verschiebung des Mittelpunktes durch eine Gaußsche Glockenkurve angedeutet. Die rechte Seite zeigt die Position der Messpunkte sowie die sich ergebende Häufigkeitsverteilung, wenn zusätzlich zu den zufälligen Unsicherheitsbeiträgen ein unkorrigierter systematischer Fehler b wirkt. Erkennbar ist sowohl die Verschiebung der möglichen Lagen des Kreismittelpunktes als auch die Verschiebung der Häufigkeitsverteilung gegenüber der unverrauschten Position. Die graue Scheibe und der Mittelwert der simulierten Positionen $\bar{x}$ liegen nun um b zum Referenzmittelpunkt verschoben. Die Differenz zwischen unverrauschtem Messwert x_{ref} und dem Mittelwert der simulierten Werte $\bar{x}$ liefert damit den systematischen Fehler b. Es gilt $b = \bar{x} - x_{ref}$.

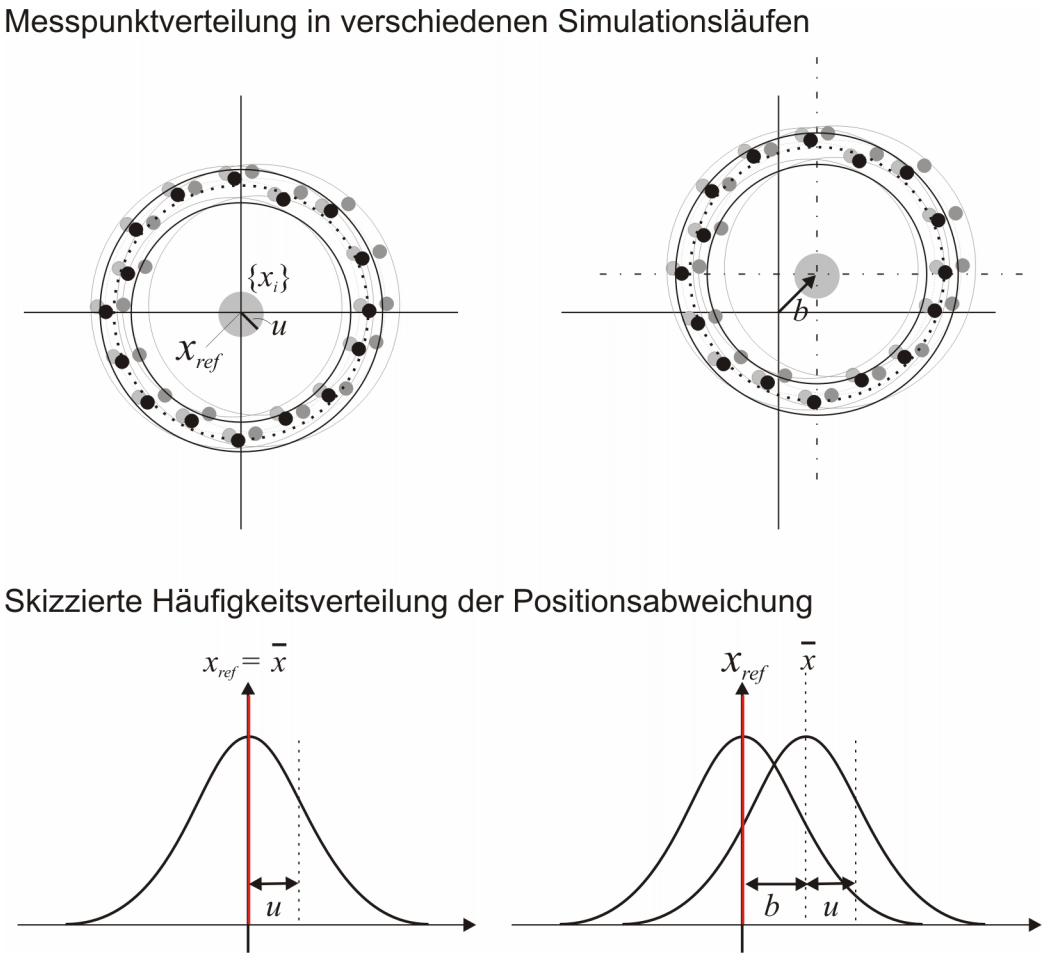


Bild 2: Einfluss systematischer Restfehler auf die Messunsicherheit am Beispiel der Lage eines Kreismittelpunkts

Analog zu Gleichung 1 kann der systematische Beitrag auch in jedem einzelnen Simulationsschritt berücksichtigt werden, indem die Varianz der simulierten Messwerte nicht auf den Mittelwert $\bar{x}$, sondern auf den Referenzwert x_{ref} bezogen wird. Dann gilt:

$$u_c = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - x_{ref}\right)^2} \qquad (2)$$

Beide Formeln liefern ein mit den Vorgaben des GUM konformes Messunsicherheitsergebnis. Damit stehen sie im Gegensatz zu der im ursprünglichen VCMM vertretenen Berechnungsformel, die den systematischen Fehler nur linear addiert und damit schon formal nicht GUM-konform ist.

Ergänzend zu der Diskussion um die Berücksichtigung systematischer Fehler haben die Untersuchungen zur Messunsicherheit von Verzahnungsmessungen gezeigt, dass die Prüfmerkmale nicht immer normalverteilt sind. Bild 3 zeigt zwei Beispiele für typische Häufigkeitsdichteverteilungen. Während die dargestellte Profilwinkelabweichung *fHa* in guter Näherung eine Normalverteilung aufweist, gilt dies nicht für die Profilformabweichung *ffa*.

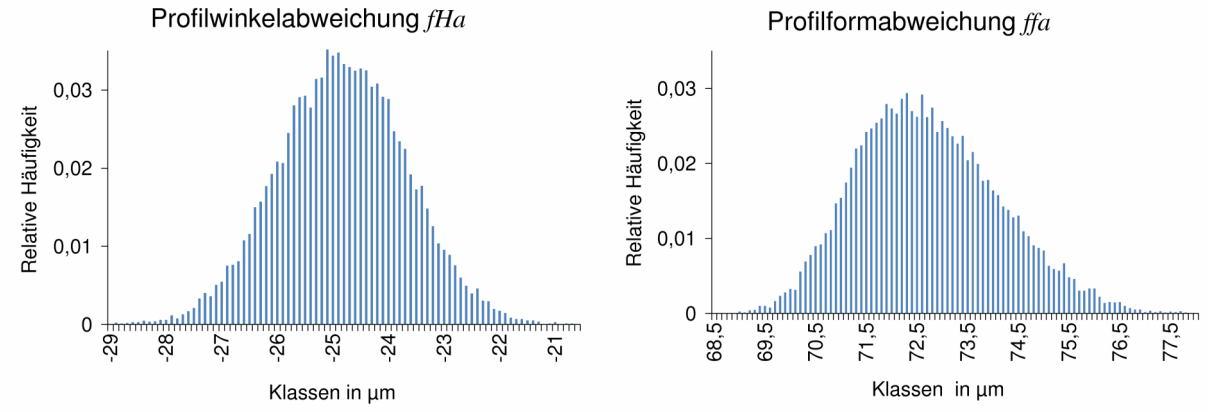


Bild 3: Verteilung der simulierten Messwerte für die Profilwinkelabweichung *fHa* und die Profilformabweichung *ffa*

Damit ist die Berechnung eines Überdeckungsintervalls einer definierten Überdeckungswahrscheinlichkeit p ohne weitere Kenntnis der Wahrscheinlichkeitsdichteverteilung der Messgröße nicht möglich. Die Angabe eines 95%-Überdeckungsintervalls mit Hilfe des Erweiterungsfaktors $k = 2$ aus der Standardmessunsicherheit u_c gilt nur für Normalverteilungen. Für andere Verteilungen ergeben sich andere Erweiterungsfaktoren k. Unabhängig von der Bestimmung der erweiterten Messunsicherheit kann ein symmetrisches Überdeckungsintervall einer definierten Überdeckungswahrscheinlichkeit p (vgl. [9, Kap. 5.3]) direkt aus den Quantilen der Häufigkeitsverteilung berechnet werden. Zur Berücksichtigung der systematischen Unsicherheitsbeiträge wird die Verteilung auf den Referenzwert bezogen, indem die Absolutwerte der Differenzen der Messwerte x_i zum Referenzwert aufsteigend sortiert werden und durch Abzählen das Unsicherheitsintervall bestimmt wird. Wenn n der Anzahl der vorhandenen Messwerte und p der gesuchten Überdeckungswahrscheinlichkeit entspricht, dann gilt für das Überdeckungsintervall U_p:

$$U_p = [x]_{\frac{n(1+p)}{2}} - [x]_{\frac{n(1-p)}{2}} \tag{3}$$

Tabelle 1 listet die sich mit den verschiedenen Formeln ergebenden Messunsicherheiten auf, wobei für die erweiterte Messunsicherheit nach Gleichung (2) die Annahme einer Normalverteilung getroffen wurde. Dies ist das übliche Vorgehen beispielsweise bei der Berücksichtigung von Kalibrierunsicherheiten, wenn keine weiteren Angaben zur Verteilung der Messwerte gemacht werden. Wie die Tabelle zeigt, wird in beiden Fällen durch die quadratische Addition die erweiterte Messunsicherheit überschätzt. Insbesondere wenn große systematische Unsicherheitsanteile vorhanden sind, kann dieser Effekt dominieren. Dies ist ein deutlicher Hinweis, den Messprozess gegebenenfalls zu verbessern.

	fHa	***ffa***
$U = 2\,u_c = 2\sqrt{s^2 + b^2}$	2,42	13,6
U_{95}	2,38	8,78
Erweiterungsfaktor $k = U_{95}/u_c$	1,96	1,29

Tabelle 1: Vergleich der Berechnungsmethoden für die Messunsicherheit der in Bild 3 dargestellten Häufigkeitsverteilungen der Profilparameter

4 Stabilität der Messunsicherheitsberechnung

Damit die statistische Aussagekraft der simulierten Messunsicherheit sichergestellt ist, muss eine ausreichend große Anzahl von Simulationen durchgeführt werden. Im Hinblick auf einen zeiteffizienten Einsatz des VCMM-Gear in der Industrie wurde ein adaptives Monte-Carlo-Verfahren umgesetzt, bei dem die Anzahl der durchzuführenden Simulationen von der Stabilität der ermittelten Messunsicherheit bestimmt wird.

Aus den Erfahrungen des VCMM ergibt sich die Forderung, dass unter gleichen Bedingungen die bei wiederholten Simulationsläufen ermittelten Messunsicherheiten maximal 10 % voneinander abweichen sollten. Dies entspricht zwei signifikanten Stellen bei der Angabe der Messunsicherheit. Höhere Anforderungen an die Genauigkeit sind nur dann sinnvoll, wenn auch die Eingangsgrößen mit größerer Genauigkeit bekannt sind. Eine Erhöhung der signifikanten Stellen würde ansonsten nur die Simulationszeit verlängern, ohne eine bessere Kenntnis des Messprozesses zu liefern. Zur Überwachung der Stabilität der Messunsicherheit eignen sich sowohl die Änderung der Messunsicherheit zwischen zwei Simulationsläufen als auch die Varianz der berechneten Messunsicherheit, da

sie analog zur Standardmessunsicherheit ein Maß für die Streuung des jeweils simulierten Wertes ist.

Zur Festlegung eines geeigneten Abbruchkriteriums wurden anhand einer typischen Profilmessung Simulationen mit verschiedenen Abbruchkriterien durchgeführt und die Streuung der Messunsicherheiten der einzelnen Prüfmerkmale betrachtet. Zusätzlich wurden die Abweichungen zu einer Simulation mit 500.000 Einzelwerten ermittelt. Tabelle 2 listet für ausgewählte Abbruchkriterien die mittleren simulierten Messunsicherheiten sowie die relative Spanne der Wiederholungssimulationen auf.

	Vergleichs-unsicherheit	**relative Änderung $\Delta s / s$ zwischen zwei Simulationsläufen**			**absolute Änderung Δs zwischen zwei Simulationsläufen**		
		$1e^{-1}$	$1e^{-2}$	$1e^{-3}$	$1e^{-2}$	$1e^{-7}$	$1e^{-9}$
Fa	1,86 µm	1,85 µm 32 %	1,87 µm 16 %	1,85 µm 5 %	1,8 µm 44 %	1,94 µm 31 %	1,84 µm 5 %
fHa	2,42 µm	2,36 µm 42 %	2,39 µm 17 %	2,42 µm 5 %	2,51 µm 24 %	2,40 µm 50 %	2,39 µm 4 %
ffa	13,60 µm	13,9 µm 9 %	13,6 µm 4 %	13,6 µm 1 %	13,8 µm 15 %	13,8 µm 8 %	13,6 µm 2 %

Tabelle 2: Messunsicherheiten und zugehörige Variation in % für drei Profilparameter bei Erreichen der Stabilität für verschiedene Abbruchkriterien

Die Ergebnisse zeigen, dass zwei signifikante Stellen der Messunsicherheit entweder durch eine maximale relative Änderung zwischen zwei Durchläufen von weniger als $1e^{-3}$ oder eine maximale absolute Änderung von weniger als $1e^{-9}$ erreicht werden können. Im VCMM-Gear wurde das relative Abbruchkriterium implementiert, damit die Anzahl der bis zur Stabilitätsgrenze durchgeführten Simulationen unabhängig von der Größenordnung der ermittelten Messunsicherheit ist. Zum Erreichen des relativen Stabilitätskriteriums $1e^{-3}$ wurden im Mittel für alle Prüfmerkmale etwa 2000 Simulationen benötigt.

5 Zusammenfassung und Ausblick

Mit den Untersuchungen zur Messunsicherheit und der Entwicklung des geräteunabhängigen statistischen Auswertemoduls konnte gezeigt werden, dass eine GUM-konforme Berücksichtigung systematischer Unsicherheitsbeiträge bei ei-

ner Monte-Carlo-Simulation möglich ist. Die Betrachtungen zur Stabilität der ermittelten Messunsicherheit zeigen darüber hinaus, dass die Messunsicherheitsermittlung durch Monte-Carlo-Simulation auch in der Industrie eingesetzt werden kann, da der Zeitaufwand für die Simulation geringer ist als für reale Wiederholungsmessungen zur Ermittlung der Messunsicherheit nach [4].

Vergleichsmessungen auf verschiedenen Verzahnungsmessgeräten konnten die grundsätzliche Eignung des Verfahrens zur Bestimmung der aufgabenspezifischen Messunsicherheit von Verzahnungsmessungen nachweisen. Mit der Integration in kommerzielle Verzahnungsmesssoftware erlaubt das entwickelte Verfahren erstmals eine Bestimmung der aufgabenspezifischen Messunsicherheit unabhängig von produktähnlichen Prüfkörpern für eine breite Palette von Verzahnungsmessaufgaben. Dies gilt insbesondere auch für Groß- und Mikroverzahnungen, für die nur wenige geeignete Prüfkörper existieren. Damit erweitert sich das Spektrum von Verzahnungen, die rückgeführt gemessen werden können.

Die entwickelten Modelle bieten eine Grundlage für die Erweiterung des Einsatzgebietes des ursprünglichen VCMM auf 3D-Koordinatenmessgeräten (KMG). Diese besitzen vielfach ebenfalls Drehtische und können die Werkstückoberfläche scannend erfassen. Damit decken die im ursprünglichen VCMM berücksichtigten Modelle nur noch einen Teil der Aufgaben ab, für die KMG heutzutage eingesetzt werden. Zur Übertragung der für Verzahnungsmessgeräte entwickelten Modelle, insbesondere des Scannings, sind jedoch weitere Untersuchungen notwendig, da aufgrund der vielfältigen Einsatzgebiete für KMG ein komplexeres Abweichungsverhalten zu erwarten ist.

Ein weiteres Einsatzgebiet der Monte-Carlo-Simulation von Verzahnungsmessungen sind F&E-Arbeiten zur Verbesserung des Messprozesses. Das VCMM-Gear ermöglicht die Identifizierung signifikanter Unsicherheitsbeiträge und erlaubt damit eine Konzentration der Entwicklungsarbeit der Messgerätehersteller auf die Aspekte, die besonders großen Einfluss auf die Messunsicherheit haben, wie das Tastsystem und die Geometriefehler der translatorischen Achsen. Insbesondere die Bestimmung und Verringerung systematischer Fehler kann den Messprozess verbessern.

6 Literatur

[1] Goch G: Verzahnungsmesstechnik – unverzichtbarer Bestandteil einer Schlüsseltechnologie. VDI-Berichte 1880, pp 1-23, 2005

[2] EN ISO14253-1: Prüfung von Werkstücken und Messgeräten durch Messen, Teil 1: Entscheidungsregeln für die Feststellung von Übereinstimmung. CEN, 1999.

[3] DIN EN ISO 9001, 2000: Qualitätsmanagementsysteme – Anforderungen.

[4] JCGM 100:2008: Guide to the Expression of Uncertainty in Measurement (GUM), 2008

[5] ISO/TS 15530-3: 2004: Geometric Product Specifications (GPS) – Coordinate measuring machines (CMM): Techniques for determining the uncertainty of measurement – Part 3: Use of calibrated workpieces or standards.

[6] Trapet E., Busch K., Franke M., Schwenke H., Wiegand U.: Rückführung von Koordinatenmeßgeräten durch Abschätzung der zu erwartenden Messabweichungen durch Simulation. Abschlussbericht zum DFG-Forschungsvorhaben TR 350 /1, PTB Braunschweig, 1997

[7] Trapet E., Franke M., Härtig F., Schwenke H., Wäldele F., Cox M., Forbes A., Delbressine F., Schellekens P., Trenk M., Meyer H., Moritz G., Guth Th., Wanner N.: Traceability of Coordinate Measurements According to the Method of the Virtual Measuring Machine. Physikalisch-Technische Bundesanstalt (Hrsg.), PTB-F-35, Braunschweig, März 1999, ISBN 3-89701-330-4

[8] Kniel, K; Härtig, F; Rost, K: Automatic Determination of the Measurement Uncertainty for Involute Gear Measurands. Proceedings of MPT2009-Sendai, JSME International Conference on Motion and Power Transmission, Mai 13- 15, 2009, Matsushima Isles Resort, Japan, 2009

[9] JCGM 101:2008: Evaluation of measurement data - Supplement 1 to the Guide to the expression of uncertainty in measurement - Propagation of distributions using a Monte Carlo method. 2008

[10] Schulze K; Kniel K, Härtig F: Abschlussbericht FVA 491: “Ermittlung einer aufgabenspezifischen Messunsicherheit von 3D-Verzahnungsmessungen“. Nr. 858, 2008

[11] Härtig, F.; Krystek, M.: Berücksichtigung systematischer Fehler im Messunsicherheitsbudget. 4. Fachtagung „Messunsicherheit praxisgerecht bestimmen“, Erfurt, 12-13. November 2008, 2008

Design und Fertigung hochminiaturisierter Fasersonden für interferometrische Messungen

Robert Schmitt, Niels König

Fraunhofer-Institut für Produktionstechnologie IPT, Steinbachstr. 17, 52074 Aachen,
E-Mail: niels.koenig@ipt.fraunhofer.de

Kurzfassung

Die kurzkohärente Interferometrie ist eine etablierte Messtechnik, welche die Möglichkeit bietet, schnelle und hochgenaue dimensionelle Messungen zuzulassen. Um diese Vorteile auch für die Messung in kleinen Bauräumen wie Bohrungen oder Röhrchen zu nutzen, sind miniaturisierte Messsonden notwendig. Auf Basis von optischen Glasfasern lassen sich entsprechende Sonden auslegen und fertigen.

Dieser Beitrag beschreibt die Entwicklung von faseroptischen Messsonden für die Anwendung in der kurzkohärenten Interferometrie. Die Strahlformung und -umlenkung der sog. All-Fiber-Sonde wird mittels Gradientenindex-Fasern und Schleifbearbeitung der Faserenden erzielt. Auf diese Weise kann auf den Einsatz refraktiver Optiken verzichtet und ein Sondendurchmesser von 80 µm und weniger erreicht werden.

1 Einleitung

Moderne Fertigungsverfahren ermöglichen die Herstellung von Bauteilen mit immer kleineren und filigraneren Strukturen. Ein Beispiel hierfür sind die Spritzlöcher von Einspritzsystemen, welche nach derzeitigem Stand der Technik mit Durchmessern von 50 µm hergestellt werden können. In diesen Größenordnungen gewinnen Formabweichungen und Ablagerungen von Verbrennungsrückständen vor dem Hintergrund der Verbrennungseffizienz und Schadstoffemission stärker an Bedeutung. Um die Qualität solcher Bauteile zu

gewährleisten, können optische Messmittel nur eingesetzt werden, sofern sie für den Einsatz in Bohrungen entsprechend stark miniaturisiert werden können. Um eine geringe Baugröße zu erzielen, eignen sich Messsonden auf Basis von Lichtwellenleitern (LWL). Dieser Beitrag beschreibt daher die Auslegung und Fertigung von LWL-Sonden für interferometrische Abstandsmessungen.

2 Kurzkohärente Interferometrie für die Messung in kleinen Bauräumen

Auf dem Gebiet der hochminiaturisierten Abstandsmesstechnik existieren bisher nur wenige kommerziell erhältliche Lösungen.

Von der Fa. Tropel wurde speziell für die Messung von Einspritzdüsen das Thetaform-System entwickelt, welches nach dem Prinzip der Zwei-Wellenlängen-Interferometrie arbeitet. Die punktuell messende Sonde besteht aus refraktiven Mikrooptiken zur Strahlformung und -lenkung und hat einen Durchmesser im einstelligen Millimeterbereich. Die Sonde kann mit bis zu 600 upm rotiert und dabei die komplette Mantelfläche des inneren Düsenkörpers gemessen werden. Über einen speziellen Strahlteiler mit den Abstrahlrichtungen 90° und 60° kann die Messung zusätzlich auf der Dichtfläche des Kegelsitzes für die Düsennadel durchgeführt werden [1]. Das Produkt wurde mittlerweile vom Markt genommen.

Die Firma Bosch hat mit dem WhitePoint-System ein kurzkohärentes Zwei-Wellenlängen-Heterodyninterferometer realisiert, welches kommerziell in Formtestern und Spezial-Koordinatenmessgeräten angeboten wird [2]. Die Messsonden werden grundsätzlich in zwei Varianten gebaut: Sonden, die ausschließlich auf Basis von Lichtwellenleitern (LWL) gebaut werden und Sonden, die zur Strahlformung und -umlenkung Gradientenindex-Linsen sowie Mikroprismen einsetzen. Letztere haben durch die Verwendung der refraktiven Optiken Durchmesser von mindestens 0,7 mm. Die faseroptischen Sonden nutzen zum Teil ebenfalls Gradientenindex (GRIN) Technologie. Es gibt darüber hinaus Ausführungen, die eine am Faserende aufgeschmolzene Kugel aufweisen, welche nochmals unter 45° schräg abgeschliffen wurde. Der Strahl wird da-

durch an der Sondenspitze an der geschliffenen Fläche durch Totalreflexion um 90° umgelenkt, wo er dann an der Kugelfläche gebrochen und fokussiert wird. Eine Sonde dieser Bauform, welche im Fraunhofer IPT hergestellt wurde, zeigt die Abb. 1. Fasersonden mit einem Kugeldurchmesser von 30 µm wurden bereits bei Bosch realisiert.

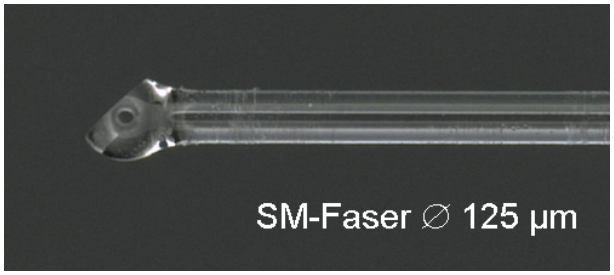


Bild 1: Messsonde aus Singlemode-LWL mit 125 µm Durchmesser. Das Kugelende wurde mit einem Faserschmelzkoppler aufgeschmozen. Die Kugel wurde mit Diamant-Schleifffolien unter einerm Winkel von 45° abgeschliffen.

Durch das Fraunhofer IPT wurde ein faserbasiertes Abstandsinterferometer entwickelt, dessen Messprinzip auf der kurzkohärenten Interferometrie basiert [3]. Das System besteht aus einem Doppelinterferometer, welches aus einem Fizeau-Interferometer aus Faserendfläche und Messobjekt sowie einem Michelson-Interferometer zum Ausgleich des optischen Gangunterschieds der Referenz- und Objektwelle aus der Messsonde aufgebaut ist. Der Ausgleich erfolgt ohne mechanisch bewegliche Bauelemente. Im Gegensatz zum WhitePoint-System wurde hier bei der Entwicklung von Beginn an auf ein rein faseroptisches Design (sog. All-Fiber-Ausführung) abgezielt, um ein möglichst hohes Miniaturisierungspotenzial zu erschließen.

Für die beiden letztgenannten Messsysteme werden faserbasierte Messsonden benötigt, welche durch ihr Potenzial zur starken Miniaturisierung Alleinstellungsmerkmale für beide Systeme darstellen und deren Einsatz für Messungen in kleinen Bauräumen wie z.B. Bohrungen ermöglicht. Grundlegende Unterschiede in den Funktionsweisen der Sonden sind in der Referenzbildung zu sehen. Während beim System des IPT die Referenzwelle am Lichtaustrittsfenster durch Fresnelreflexion erzeugt wird, verfügt das WhitePoint-System über eine separat erzeugt Referenz, die entweder durch ensprechende parallel verlegte LWL-Referenzstrecken oder durch Teilerbeschichtungen gebildet wird. Ein ent-

scheidender Vorteil, die Referenz in der selben Faser abzugreifen, die die Objektwelle führt, ist die Unempfindlichkeit gegenüber Einflüssen auf die Faser wie z.B. Temperaturschwankungen, Polarisationseffekte durch mechanischen Stress oder optische Weglängenverschiebungen durch Faserdehnung.

3 Optisches Design faseroptischer Sonden

Miniaturisierte optische Messsonden auf Basis klassischer refraktiver Linsen oder Prismen lassen sich nur mit Durchmessern von mindestens 0,5 mm realisieren, da solche Optiken mit Durchmesser bzw. Kantenlänge erst ab 0,3 - 0,5 mm erhältlich sind und diese vorzugsweise mit Röhrchen geschützt werden müssen. Dies bedeutet, dass zur Strahlformung und -lenkung bei noch kleineren Bauformen alternative Ansätze nötig sind.

3.1 Fokussierung durch Gradientenindex-LWL

Über die Strahlformung mit Hilfe von GRIN-LWL als sog. Fiber-lenses wurde bereits im Jahr 1987 berichtet [4]. Im Gegensatz zu Stufenindex-LWL haben GRIN-LWL einen parabellförmigen Verlauf des Brechzahlprofils im Kern (siehe Abb. 2).

Gemäß dem Fermat'schen Prinzip breitet sich Licht in einem inhomogenen Medium derart aus, dass der optische Weg AB entlang eines geometrischen Pfads S_0 gegenüber benachbarten Pfaden einen Extremalwert besitzt:

$$\left(\frac{\delta\overline{AB}}{\delta S}\right)_{S_0} = \delta \int_A^B n\,(s)\,ds = 0. \tag{1}$$

Dies hat zur Folge, dass schräg in den GRIN-LWL eingekoppelte Strahlen immer zum Kern hin gebrochen werden, was über alle innerhalb der numerischen Apertur (NA) der Faser eingekoppelten Strahlen eine charakteristische, sinusförmige Propagation zur Folge hat, wie die Abb. 3 zeigt.

Dieser Effekt kann für die Fokussierung ausgenutzt werden, indem der GRIN-LWL bspw. durch Cleaven[1] auf eine exakte Länge gekürzt wird. Durch Spleißen

[1] Das sog. Cleaven ist ein Trennvorgang, bei dem der LWL von zwei Seiten eingespannt und durch eine Klinge (i.d.R. diamantbesetzt) eine Läsion in der Mantelfläche zugefügt wird. Bei gleichzeitiger Dehnungs- oder Torsionsbeanspruchung reißt der LWL und es entsteht eine nahezu ideal glatte Bruchfläche.

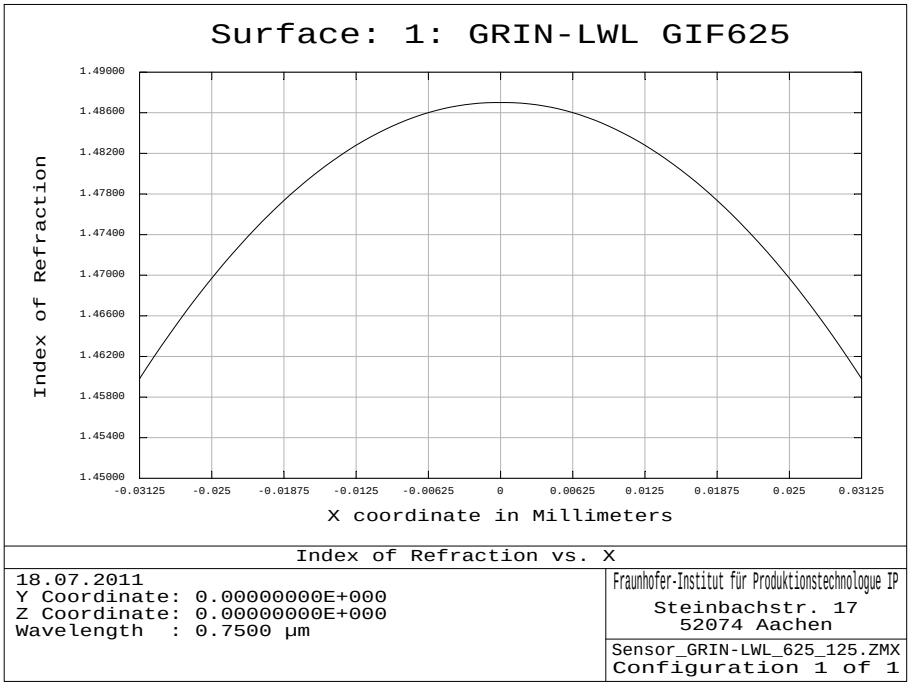


Bild 2: Parabellförmiges Brechzahlprofil des GRIN-LWL Nufern GIF 625 ($n_0 = 1,487$, $n_1 = 1,46$, $\oslash_{Kern}$=62,5 µm, $\oslash_{Cladding}$=125 µm)

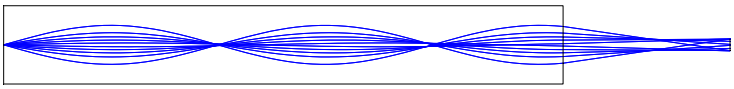


Bild 3: Sinusförmige Propagation durch GRIN-Faser (GIF 625, λ=750 nm).

des GRIN-LWL an die lichtführende Singlemode-Faser kann auf diese Weise wahlweise eine Kollimation oder Fokussierung des Messstrahls bewirkt werden. Die Brennweite f einer so erzeugten Faserlinse beträgt in Abhängigkeit von der Länge L des GRIN-LWL

$$f(L) = \frac{n\left(1 - \frac{\omega^4}{\omega_0^4}\right)\cos(g \cdot L)\sin(g \cdot L)}{g \cdot n_0\left[\sin^2(g \cdot L) + \frac{\omega^4}{\omega_0^4}\cos^2(g \cdot L)\right]} \tag{2}$$

mit ω_0 als Taillendurchmesser des Strahls aus der Singlemode-Faser und n dem Brechungsindex des Mediums, in das der Strahl aus der Faserlinse propagiert (z.B. Luft oder Mikroprisma). Die Gradienten-Konstante g und n_0 beziehen sich auf das Profil des Brechungsindex der verwendeten GRIN-Faser. g beträgt in der Regel 2 für die meisten GRIN-LWL. Der Faktor ω hängt zusätzlich von der Wellenlänge λ ab mit $\omega = \lambda/(\pi g n_0)$ [5].

Abb. 4 zeigt, dass die Abhängigkeit der effektive Brennweite f von der Länge

L des GRIN-LWL ein periodisches Verhalten aufweist. An den Polstellen der Funktion erhält man die Länge L, mit der sich ein kollimierender Strahlengang der Fasersonde realisieren lässt.

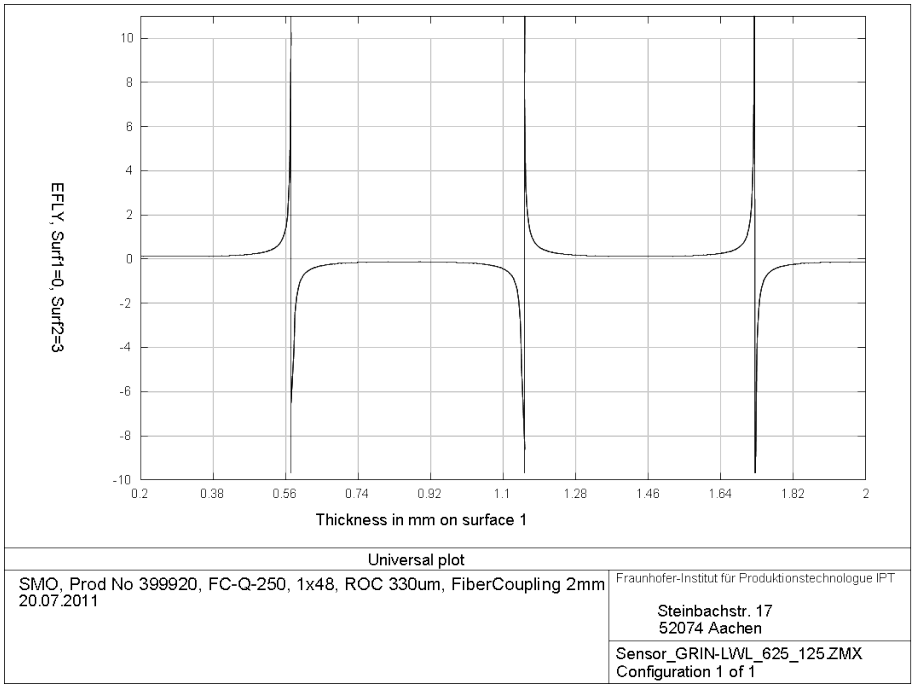


Bild 4: Effektive Fokuslänge in Abhängigkeit der Länge der GRIN-Faser (GIF 625, λ=750 nm).

Die Strahlformung von GRIN-LWL lässt sich im Optikdesign-Programm ZEMAX simulieren. Ebenso können damit die optischen Eigenschaften der Sonde bestimmt werden. Dazu gehört beispielsweise die Größe des Messspots. Für den betrachteten Fall mit einer GRIN-Faser GIF 625 und der SM-Faser HI780 (Modenfelddurchmesser 4,6 µm bei 780 nm bzw. 5,0 µm bei 850 nm) als lichtführender Faser liegt der kleinstmögliche beugungsbegrenzte Spotdurchmesser bei 0,7 µm. Dieser Wert basiert auf dem simulierten RMS-Radius. Die Beugungsbegrenzung wird durch das Airy-Scheibchen angegeben, welches für den betrachteten Fall (Länge der GRIN-Faser: 451 µm, Arbeitsabstand: 47 µm) einen Durchmesser von 7,832 µm hat.

3.2 Strahlumlenkung durch Bare-Fiber Polishing

Zur Strahlumlenkung kann die Faserendfläche schräg geschliffen werden, um so über Totalreflexion den Messstrahls z.B. in Richtung einer Bohrungswandung umzulenken. Strahlformung und -umlenkung können also rein faserbasiert rea-

lisiert werden. Länge und Schrägschliff der Faserspitze sind damit Designfeatures der Messsonde.

Zur Schleifbearbeitung des Faserendes wurde im IPT ein sog. Bare-Fiber Polishing Prozess eingesetzt. Der wichtigste Unterschied zur weit verbreiteten Schleif- und Polierbearbeitung von LWL bei der Konfektionierung mit Steckern ist, dass beim Bare-Fiber Polishing die Faser nicht durch Kleben eingebettet sondern lediglich lösbar eingespannt wird. Ein weiteres Merkmal ist die Faserzuführung, die mit einem frei einstellbaren Winkel die Faser auf den Schleifteller führt.

Das im IPT eingesetzte Schleifgerät zeigt die Abb. 5. Die Faserspitze wird in einer Keramikferrule geführt, welche mit dem Zielwinkel bereits vorgeschliffen ist. Der Schleifprozess ist parametrisiert durch die Spindeldrehzahl, die Vorschubgeschwindigkeit der Faser, die Körnung und Korndichte der Schleiffolien, den Anpressdruck in Form von Kontergewichten sowie die Wahl von Kühlschmierstoff und Oszillation der Faserführung während des Schleifprozesses.

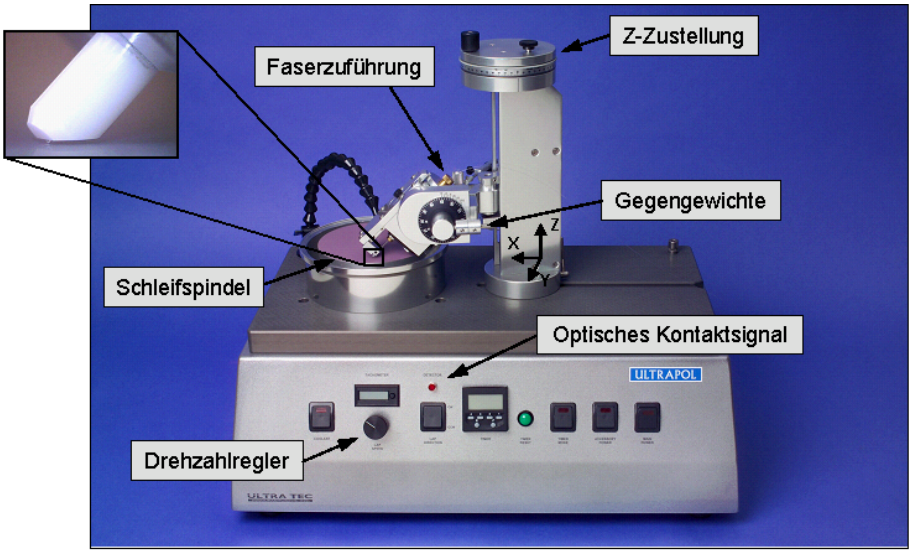


Bild 5: Schleifmaschine für das Bare-Fiber Polishing von LWL mit frei einstellbarem Winkel (Hersteller: Ultrapol)

Zur Erzielung einer möglichst hohen Strahlqualität nach Totalreflexion musste zunächst ein Schleifprozess zur Erzielung einer möglichst glatten Schliffffläche entwickelt werden. Hierfür wurde ein vollfaktorieller Versuchsplan aufgestellt,

die geschliffene Faseroberfläche mit einem Weißlichtinterferenzmikroskop gemessen und die 3D-Rauheit S_a nach DIN EN ISO 25178-2 ausgewertet. Das Ziel der Herstellung einer optischen Funktionfläche mit einer Rauheit im einstelligen Nanometerbereich, konnte erreicht werden, wie die Abb. 6 zeigt.

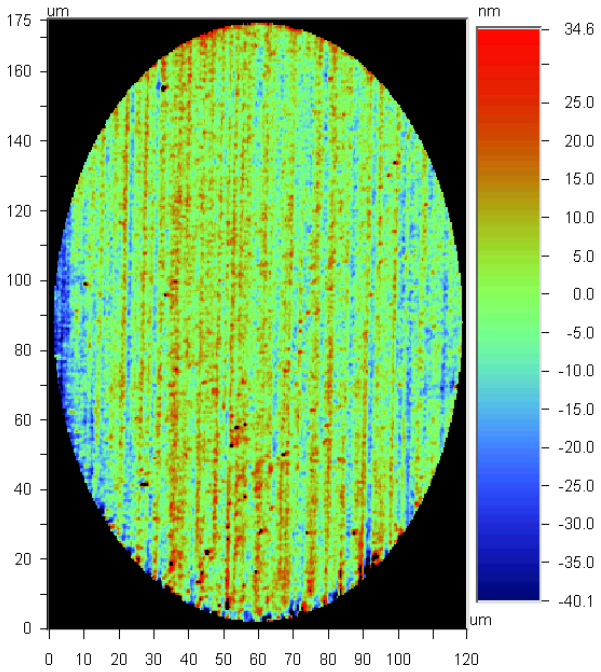


Bild 6: Mittels Weißlichtinterferenzmikroskop gemessene Faserendfläche (Rauheit S_a=7,29 nm). Die Messung zeigt die charakteristischen Riefen aus dem Schleifprozess mit Diamantschleiffolie.

Die mit dem Schleifprozess herstellbaren Geometrien der Sondenspitzen zeigt die Abb. 7. Die einfach geschliffene Fasersonde mit 90°-Strahlumlenkung (Abb. 7 links) eignet sich zur Messung in Bohrungen. Für die Messung des Kegelsitzes in Einspritzdüsen wurde im IPT ein weiterer Sondentyp mit einem zweifachen 30°-Schliff entwickelt (Abb. 7 rechts). Die beiden Flächen bilden ein verschobenes Equilateralprisma. Da der Strahl an der ersten Fläche totalreflektiert wird, tritt er an der gegenüberliegenden Fläche orthogonal aus, so dass diese als Referenzfläche für die interferometrische Messung genutzt werden kann.

Um Strahlformungs- und Strahllenkungseigenschaften in miniaturisierter Bauform zu vereinen, wird die Schleifbearbeitung vorzugsweise auf der GRIN-Faserlinse durchgeführt, nachdem diese mit dem lichtführenden SM-LWL gespliced wurde. Um die Länge des GRIN-LWL zu berechnen, wurde die

Fasersonde zunächst in ZEMAX im sog. Non-Sequential Mode modelliert (siehe Abb. 8). Durch Totalreflexion werden die Strahlen um 90° umgelenkt. Die Grenzflächen zwischen Faserkern- und Mantel bzw. Mantel und Luft wirken bei Strahlaustritt wie Zylinderlinsen. Dies resultiert in einem elliptischen Messspot (Abb. 9 links). Je nach Anwendung und Probengeometrie lässt sich der Messspot durch Variation der Länge des GRIN-LWL entweder in der Längs- oder der Quer-Achse minimieren (Abb. 9 links).

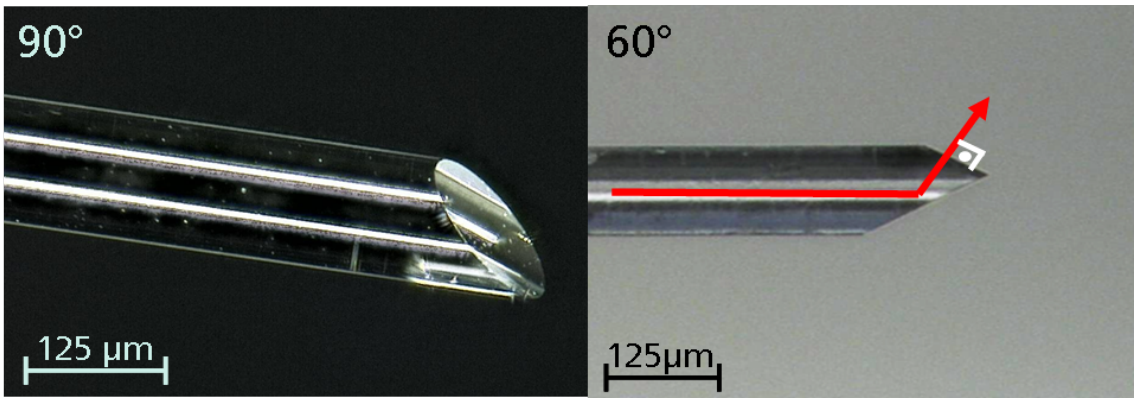


Bild 7: Links: Fasersonde mit 45°-Schliff für die 90°-Strahlumlenkung (SM-LWL Corning HI780), Rechts: Fasersonde mit zwei 30°-Schliffen. Der Strahl wird um 60° aus der optischen Achse abgelenkt und tritt nach Totalreflexion senkrecht zum Lichtaustrittsfenster aus.

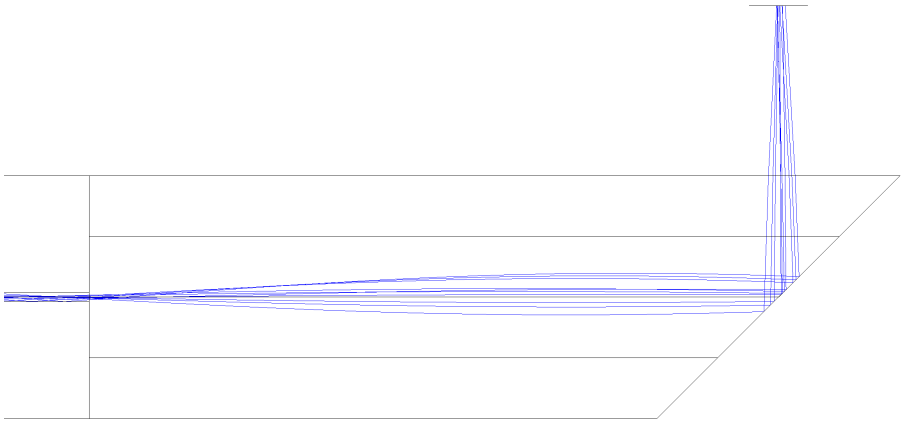


Bild 8: Fasersonde aus SM-LWL und GRIN-LWL mit 45°-Facette. Durch Totalreflexion werden die Strahlen durch den Mantel aus der Sonde ausgekoppelt.

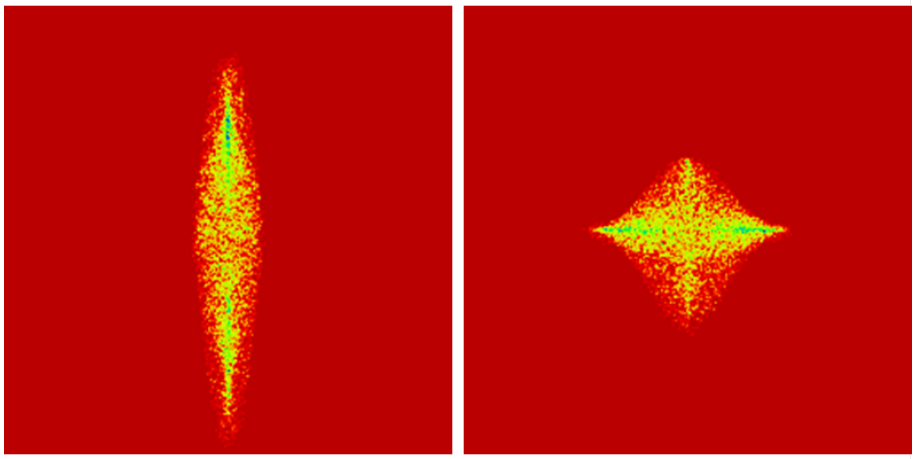


Bild 9: Simulierte Messspots der Fasersonde mit 90°-Strahlumlenkung. Links: Elliptischer Spot mit minimalen RMS-Radius von 0,88 µm in Quer-Richtung (Radius in Längs-Richtung 5,56 µm), Rechts: Spot mit minimalem RMS-Radius 2,61 µm

4 Zusammenfassung und Ausblick

In diesem Beitrag wurde gezeigt, wie sich faseroptische Sonden für interferometrische Messungen aufbauen und fertigen lassen. Das Sonden-Setup mit GRIN-Faserlinsen ermöglicht eine Strahlformung bei gleichzeitig kompakter Bauform. Die Herausforderung bei der Umsetzung der Designvorgaben bestehen in der exakten Ablängung des GRIN-LWL. Hierbei beträgt die Längenabweichung von der Zielgröße durch die von uns verwendete Cleave-Technik etwa 5 µm. Zur Strahlumlenkung eignet sich das Schleifen der LWL-Sondenspitze, wodurch eine Montage von Mikroprismen umgangen wird und zusätzlich kleinste Sonden etwa für die Bohrungsmessung realisiert werden können.

Mit Hilfe von rein faserbasierten Messsonden können sämtliche interferometrische Längenmessverfahren Messaufgaben in kleinen Bauräumen wie Düsenspritzlöchern oder Mikrobohrungen erschließen. Dazu zählen nicht nur die Laserinterferometrie, Mehrwellenlängen- und kurzkohärente Interferometrie. Entsprechende Applikationen mit Sondendurchmessern von bis zu 80 µm zur Messung in Einspritzdüsen wurden von uns bereits in [6] gezeigt. Kleinere Durchmesser lassen sich mit zusätzlicher Taperbearbeitung des LWL erzielen. Sonden für die optische Kohärenztomographie (OCT) z.B. für endoskopische Anwendungen wie die kardiovaskuläre Diagnostik lassen mit den gezeigten Verfahren ebenfalls realisieren. Entsprechende Eingriffe sind dank der kompakten Bau-

form minimal-invasiv und damit patientenschonend.

Zur weiteren Charakterisierung und Verbesserung des Sondenbaus ist es wichtig, das optische Verhalten an den geschliffenen Flächen besser zu verstehen. Fragestellungen wie der Einfluss der Rauheit (Schleifriefen) auf das Totalreflexionsvermögen bzw. das Streulichtverhalten können dabei simulativ betrachtet werden. Für das Design von Sonden, die das Lichtaustrittsfenster als Referenzfläche für die interferometrische Messung nutzen, wäre der Interferenzkontrast eine zusätzliche Optimierungsgröße. Die Erfahrung hat gezeigt, dass eine starke Fokussierung, d.h. geringer Arbeitsabstand und kleine Messspots nicht förderlich für einen hohen Interferenzkontrast sind, da durch die daraus resultierenden größeren Lichtaustrittswinkel die Fresnelreflexion von der Grenzfläche abnimmt. Im Gegensatz dazu ist für Anwendungen, bei denen das Referenzsignal anderweitig erzeugt wird, eine Fresnelreflexion hinderlich, da sie den Signal-Rausch-Abstand verringert. Für diesen Fall kann das Lichtaustrittsfenster entspiegelt und/oder schräg abgeschliffen werden, so dass die Reflexion nicht mehr in der Faser weitergeleitet wird.

Literatur

[1] T. Dunn, "Optical Metrology Enables New Generation of Fuel Injectors", *Optics & Photonics News* **14**(6), S. 28–33, 2003.

[2] O. Jusko, M. Neugebauer, H. Reimann, P. Drabarek, M. Fleischer und T. Gnausch, "Form measurements by optical and tactile scanning", in *Proc. ISIST*, 2008.

[3] F. Depiereux, P. Lehmann, T. Pfeifer und R. Schmitt, "Fiber-optical sensor with miniaturized probe head and nanometer accuracy based on spatially modulated low-coherence interferogram analysis", *Appl. Opt.* **46**, S. 3425–3431, 2007.

[4] W. Emkey und C. Jack, "Analysis and evaluation of graded-index fiber lenses", *J. Lightwave Technol.* **5**(9), S. 1156–1164, 1987.

[5] W. A. Reed, M. F. Yan und M. J. Schnitzer, "Gradient–index fiber–optic microprobes for minimally invasive in vivo low–coherence interferometry", *Opt. Lett.* **27**, S. 1794–1796, 2002.

[6] T. Pfeifer, R. Schmitt, N. König und G. F. Mallmann, "Interferometric measurement of injection nozzles using ultra-small fiber-optical probes", *Chinese Optics Letters* **9**, S. 071202, 2011.

Berührungslose Messung des Drehwinkels und der Axialbewegung an Rotorwellen von Windenergieanlagen

Jan F. Westerkamp, Michael Sorg, Carsten Bredemeier, Gert Goch

Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft (BIMAQ)
Prof. Dr.-Ing Gert Goch, Bremen
Tel.: 0421-218-64609, e-mail: wes@bimaq.de

Kurzfassung

Schwingungen der Komponenten einer Windenergieanlange führen oftmals zu Schäden und Ausfällen der Getriebe und Lager. Diese Defekte treten häufig wesentlich früher als prognostiziert ein. Umgekehrt führen Schäden in den oft schwer zugänglichen Bauteilen wiederum zu Schwingungen. Die frühzeitige Detektion von Initialschäden sowie die Prognose der zu erwartenden Restlebensdauer auf Grundlage beobachteter Betriebszustände sind eine wesentliche Voraussetzung zur Erhöhung der Anlagenverfügbarkeit. Eine aussagekräftige, messtechnische Analyse der Schwingungen und Beschleunigungen ist zurzeit zumeist noch nicht möglich und erfordert weitere prozessnahe oder In-Prozess-Messtechnik, um Schäden schon im laufenden Betrieb und in einem frühen Stadium zu detektieren.

In diesem Beitrag wird ein optischer Sensor zur berührungslosen Messung des Drehwinkels und der Axialbewegung der Rotorwellen von Windenergieanlagen vorgestellt. Es wird gezeigt, welchen Einfluss der Sensorabstand sowie die Oberflächengeschwindigkeit auf die Messunsicherheit haben.

Aufgrund des geringen Montageaufwands bietet sich der Sensor für die Nachrüstung und den mobilen Einsatz im Rahmen von Wartungsarbeiten an.

1 Einleitung

Schwingungen sind eine häufige Ursache für Schäden an Windenergieanlagen (WEA). Durch Massenunwuchten und aerodynamische Unwuchten der Rotorblätter werden beispielsweise Schwingungen erzeugt, die in Rotorwelle, Lager und Getriebe eingeleitet werden und dort zu Schäden führen können. Zudem verursachen häufig schwer zugängliche Bauteile in WEA Schwingungen, deren messtechnische Analyse im laufenden Betrieb zumeist noch nicht möglich ist. Die frühzeitige Erkennung von Initialschäden sowie eine Prognose der zu erwartenden Restlebensdauer lassen sich jedoch nur auf der Basis prozessnah erfasster Messdaten realisieren. Lösungen für die erweiterte Erfassung von Anlagenkenngrößen bietet hier die In-Prozess-Messtechnik.

Das Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft (BIMAQ) baut zurzeit in einem vom Bundesumwelt-ministerium geförderten Projekt ein „Inside-Sensoring Labor“ [1] auf. Hier werden neue Messmethoden entwickelt, um die lokale Verteilung von Kräften, Momenten, Vibrationen und Temperaturen im Triebstrang zu erfassen und um Verschleiß- und Ermüdungszustände von Anlagenbauteilen zu detektieren. Bereits während der Realisierung des Inside-Sensoring Labors entwickelt das BIMAQ Lösungen für neue Sensoren und Messsysteme und erprobt diese am WEA-Triebstrang des Labors. Mit dem Inside-Sensoring Labor ergänzt das BIMAQ die Forschungsinfrastruktur des gemeinsamen Zentrums für Windenergieforschung der Universitäten Oldenburg, Hannover und Bremen (ForWind).

Ein Beispiel aktueller Untersuchungen im Inside-Sensoring Labor ist ein Streulicht-Sensor, der sich für die berührungslose Messung des Drehwinkels und der Axialbewegung von Antriebswellen während des Betriebes eignet. Bild 1 zeigt axiale und radiale Schwingungen, die sich der Bewegung der rotierenden WEA-Rotorwelle überlagern und zu vorzeitigen Lagerschäden führen können.

Der Streulicht-Sensor erfasst die auf die Lager einwirkenden axialen und radialen Bewegungen der Welle. Die hochauflösend erfassten Drehwinkel-Messdaten stehen für die drehwinkelsynchrone Auswertung von Frequenz-spektren und für die Detektion von Torsionsschwingungen zur Verfügung.

Mit diesem Streulicht-Messverfahren [2] ist es möglich, laterale Bewegungen mit Oberflächengeschwindigkeiten bis zu 5.000 mm/s und mit einer Auflösung

bis zu 0,5 µm zu messen. Der Messabstand variiert dabei je nach Sensor zwischen wenigen Millimetern bis zu mehreren Dezimetern.

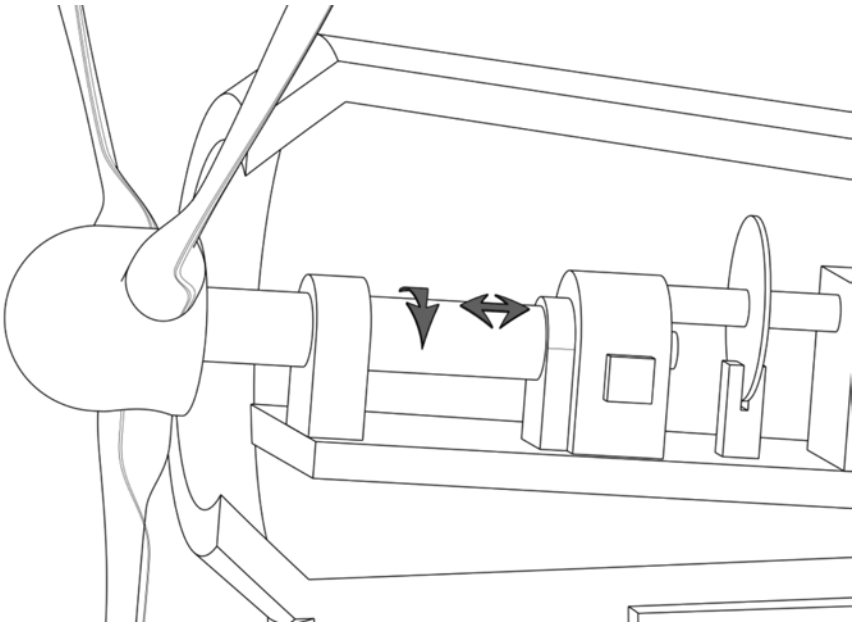


Bild 1: Überlagerte axiale und radiale Schwingungen an einer WEA-Rotorwelle

2 Messaufbau / Sensor

Für Voruntersuchungen im Inside-Sensoring Labor kam ein kostengünstiger, optischer Laser-Navigationssensor von Avago Technologies zum Einsatz. Bild 2 zeigt die Unterseite von zwei Labormustern des optischen Streulichtsensors, die gemeinsam auf einem Piezo-Mikrostelltisch befestigt sind.

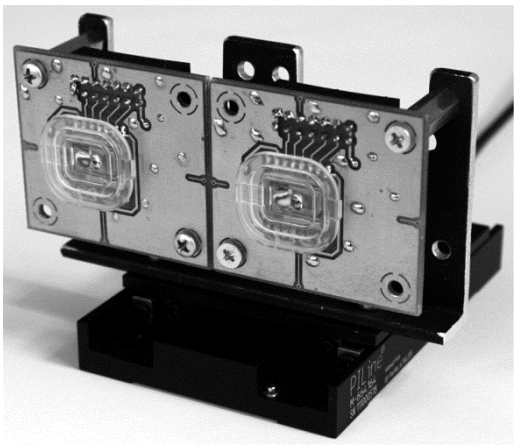


Bild 2: Labormuster zweier optischer Sensoren auf einem Piezo-Mikrostelltisch

Für die berührungslose Erfassung von Bewegungen beleuchtet eine Laserdiode einen kleinen Bereich der Oberfläche des Messobjekts. Eine Linse leitet das von der Oberfläche reflektierte Licht auf einen CMOS-Sensor (Bild 3), dessen Bildsignale ein digitaler Signalprozessor (DSP) weiter verarbeitet. Der im Sensor integrierte DSP vergleicht das aktuelle Bild mit dem jeweils vorherigen Bild und berechnet daraus den Vektor der Bildverschiebung.

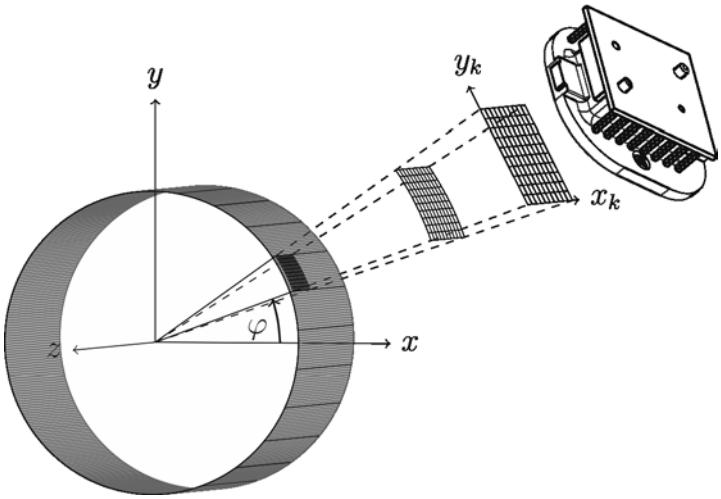


Bild 3: Messprinzip: ein gekrümmtes Oberflächensegment der Rotorwelle wird auf die ebene Fläche des CMOS-Sensors abgebildet.

Aus der erfassten Verschiebung des Sensorbildes in x_k- und y_k-Richtung lassen sich die Bewegungsrichtung und -geschwindigkeit des Messobjektes relativ zur Position des Sensors berechnen.

Für die Untersuchung der Sensoreigenschaften (z.B. Auflösung und Messunsicherheit) beim Einsatz an technischen Oberflächen wurde im Inside-Sensoring Labor eine Versuchsanordnung aufgebaut, welche reproduzierbare Relativbewegungen zwischen optischem Sensor und der Oberfläche eines Messobjektes ermöglicht. Die verwendete Versuchsanordnung ist in Bild 4 dargestellt. Ein Linearstelltisch bewegt den Sensor in x-Richtung am Messobjekt entlang. Der Abstand des Sensors zum Messobjekt wird mit einem Piezo-Mikrostelltisch eingestellt. Als Messobjekt dient ein Aluminiumprofil mit gewalzter Oberfläche.

Der Sensor übermittelt die gemessenen Daten an einen Messrechner, der auch die Linearstelltische ansteuert, um die Position des Sensors gegenüber der Oberfläche des Messobjektes zu variieren.

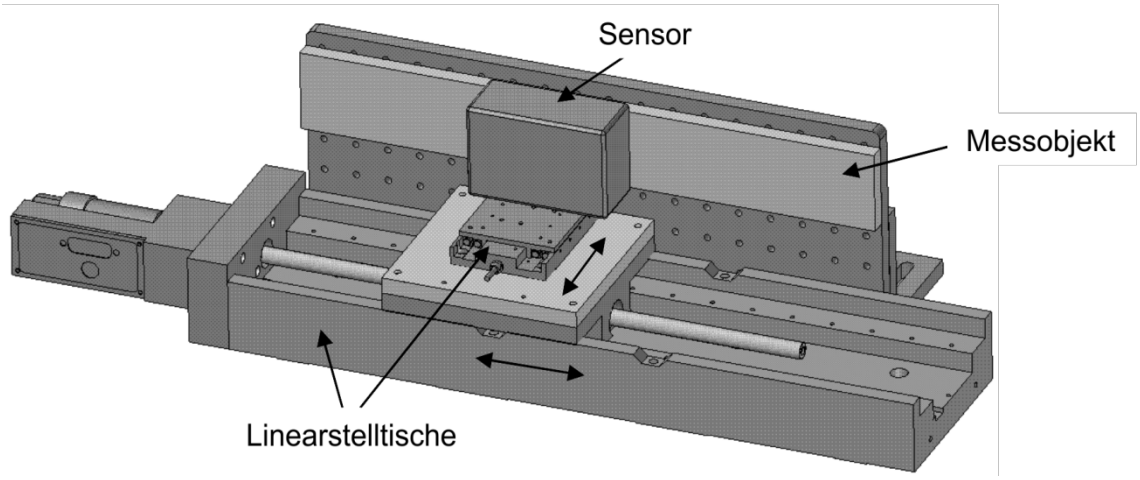


Bild 4: Versuchsaufbau mit zwei Linearstelltischen zur Bewegung des Sensors entlang einer Oberfläche sowie zur Variation des Messabstands

3 Messungen

Zur Charakterisierung des Sensors sind detaillierte Untersuchungen erforderlich, da für den beschriebenen Anwendungsfall keine Angaben zur Messunsicherheit der Wegmessung oder zum Einfluss des Sensorabstands auf das Messergebnis verfügbar sind. Untersuchungen zum Einsatz eines ähnlichen Sensors als Absolutdrehgeber auf einer Scheibe [3], zur Messung von Wegstrecken [4, 5, 6] oder auch zur Wegmessung mit selbst entwickelter Optik [7] lassen sich auf diesen Sensor nicht direkt übertragen. Die wesentlichen Unterschiede des hier untersuchten Sensors sind die nominelle Auflösung von bis zu 5040 cpi (counts per inch) sowie die hohe Bildwiederholrate von bis zu 11.750 Bildern pro Sekunde.

3.1 Einfluss des Sensorabstands auf das Messergebnis

Der Sensor soll in der Anwendung (z.B. für den mobilen Einsatz) mit einfachen Hilfsmitteln über dem Messobjekt befestigt werden. Daher ist die Abhängigkeit der Messwerte vom Abstand des Sensors zur Messoberfläche von großer Bedeutung, da eine aufwändige Justierung vermieden werden soll. Wie im Bild 5 zu sehen, steigt die empirisch ermittelte Standardunsicherheit des Mittelwertes der Sensorimpulse pro zurückgelegter Strecke an der oberen und unteren Grenze des Abstands signifikant an. Im Bereich von 2 mm bis 4 mm verändert sie sich hingegen nur geringfügig.

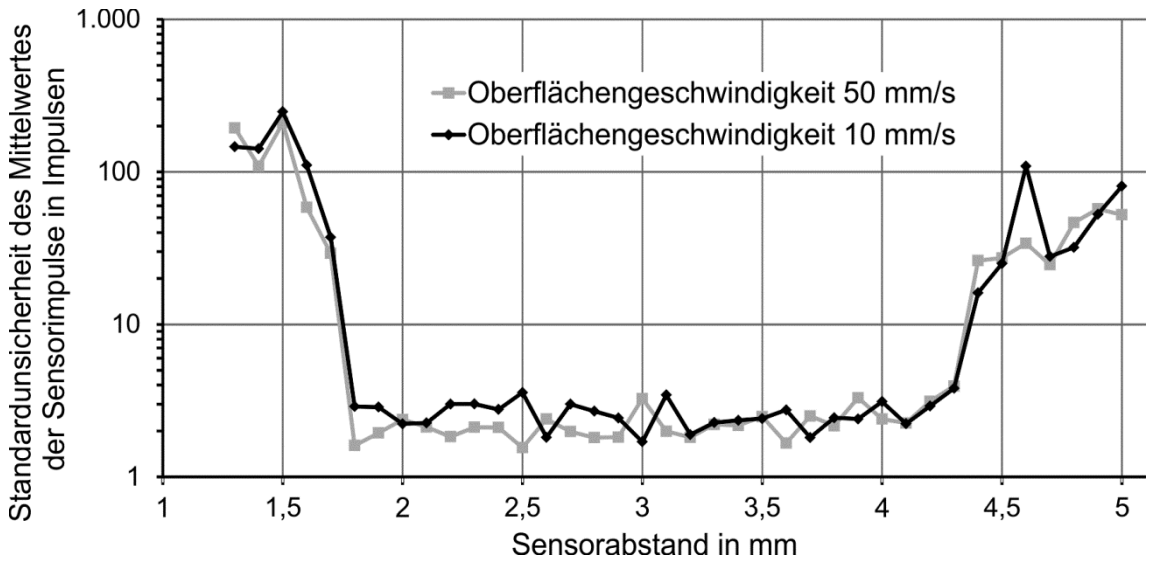


Bild 5: Empirisch ermittelte Standardunsicherheit des Mittelwertes der Sensorimpulse pro mm bei einem Sensorabstand von 1,3 mm bis 5 mm.

Die ermittelte Auflösung in Impulsen pro Millimeter zeigt im untersuchten Bereich hingegen eine deutliche Abhängigkeit vom Sensorabstand (Bild 6). Es sind daher Kalibrierungen für mehrere Abstände erforderlich und die Distanz zwischen Sensor und Messoberfläche muss in der Applikation mit einer geringen Unsicherheit bekannt sein.

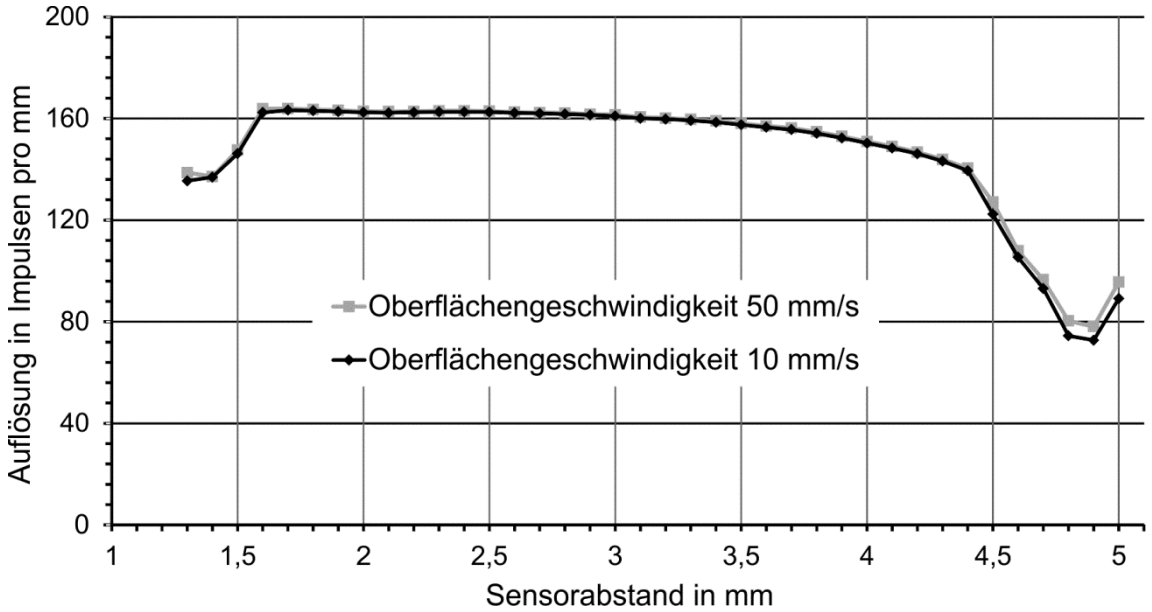


Bild 6: Anzahl der Sensor-Impulse pro mm bei einem Sensorabstand von 1,3 mm bis 5 mm in Abständen von 0,1 mm.

Für die Messung des Drehwinkels von Antriebswellen kann ein optischer Marker, der eine vollständige Umdrehung kennzeichnet, die ermittelte Abstandsabhängigkeit der Anzahl der Impulse kompensieren. Mit diesem Index

ist der Sensor in der Applikation kalibrierbar und zusätzlich wird das Aufsummieren von Messabweichungen auf jeweils eine volle Umdrehung der Welle beschränkt. Die Erkennung der Marker ist mit demselben Sensor möglich, da dieser zusätzlich zu den wegabhängigen Impulsen weitere Informationen zu jeder Messung ausgibt, wie beispielsweise die Anzahl der markanten Punkte im Bild oder die Belichtungszeit [8, 9].

3.2 Einfluss der Oberflächengeschwindigkeit auf das Messergebnis

Im Bild 6 ist bereits zu erkennen, dass die Anzahl der Sensorimpulse nur geringfügig von der Geschwindigkeit abhängt. Messungen bei einem festem Abstand von 2,4 mm, einer Geschwindigkeit von 1 mm/s bis 50 mm/s in Schritten von 5 mm/s, einer Messstrecke von 200 mm und je Geschwindigkeit 25 Wiederholungen ergeben den in Bild 7 gezeigten Verlauf. Der Unterschied beträgt weniger als 0,7 % vom Endwert.

Mit den im Labor untersuchten Geschwindigkeiten können in WEA z.B. Rotorwellen mit einem Durchmesser von 0,5 m bis zu einer Drehzahl von ca. 1,9 1/min untersucht werden. Allerdings sind die untersuchten Geschwindigkeiten zurzeit noch weit vom Maximalwert des Sensors von 5.000 mm/s entfernt.

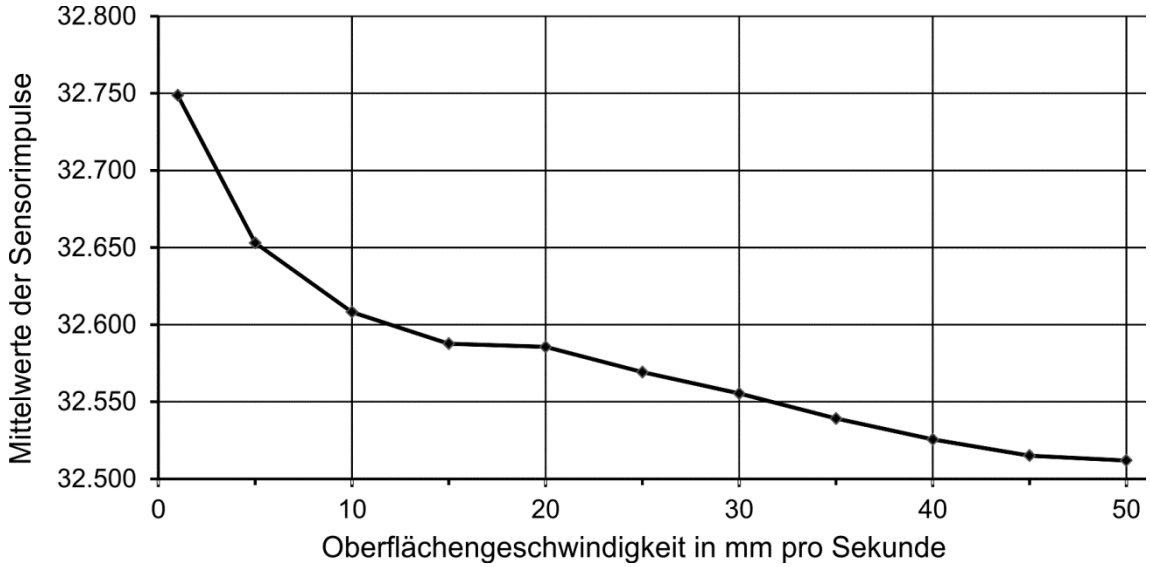


Bild 7: Anzahl der Sensorimpulse bei verschiedenen Geschwindigkeiten und einer konstanten Strecke von 200 mm.

3.3 Linearität der Längenmessung

In den Versuchen bewegte der Linearstelltisch den Sensor jeweils 25 Mal mit konstanter Geschwindigkeit (bis auf Beschleunigungs- und Verzögerungsrampen) einer vorgegebenen Strecke parallel zum Messobjekt. Bild 8 zeigt die Mittelwerte der vom Sensor ausgegebenen Impulse sowie die daraus ermittelte Standardunsicherheit des Mittelwertes für eine Geschwindigkeit von 25 mm/s. Die ermittelte Auflösung des Sensors bei einem Abstand von 2,4 mm von der Oberfläche beträgt 162,8 Impulse pro mm, entsprechend 6,1 µm pro Impuls.

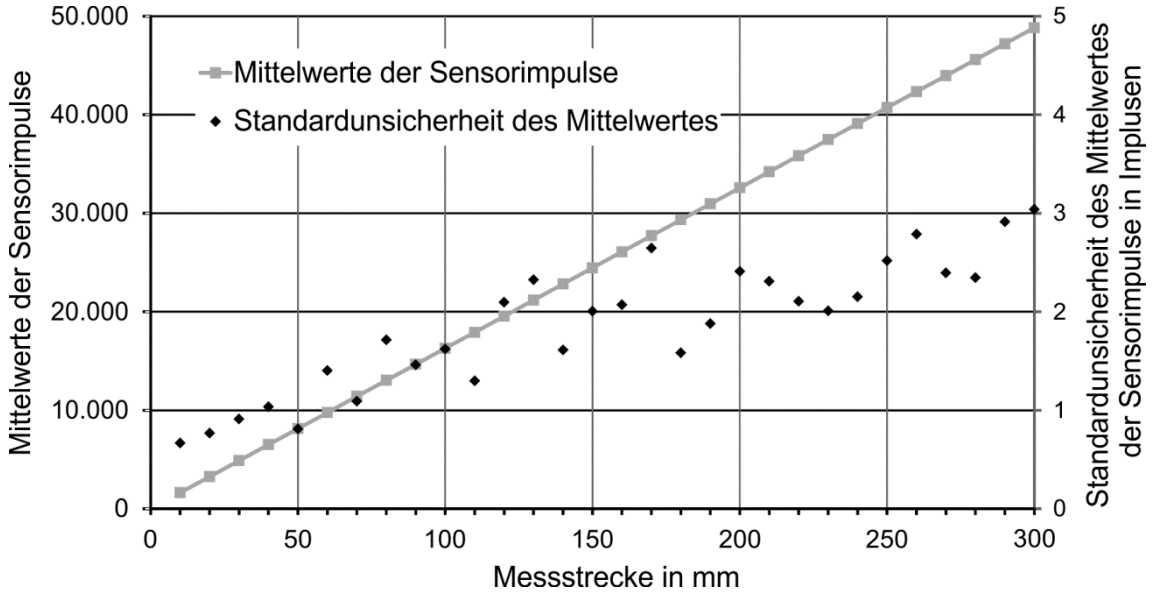


Bild 8: Ergebnisse der Weglängenmessung auf einem Linearstelltisch bei einer Geschwindigkeit von 25 mm/s.

Das Ergebnis in Bild 8 zeigt, dass die Standardunsicherheit des Mittelwertes der Sensorimpulse bei konstanten Messbedingungen (Messabstand und Geschwindigkeit konstant) sehr gering ist. Desweiteren zeigen die Messergebnisse, dass die Anzahl der Sensorimpulse praktisch unabhängig von der gemessenen Weglänge ist. Die Längenmessung mit dem Sensor zeigt ein sehr lineares Verhalten und erfüllt damit eine grundlegende Voraussetzung für den weiteren Einsatz.

4 Anwendungen / Diskussion

Für die Drehwinkelmessung an Rotorwellen ist der optische Laser-Navigationssensor aufgrund der geringen Messunsicherheit sowie der einfachen

Montage sehr gut geeignet. Zur Kalibrierung und um das Aufsummieren von Fehlern bei zeitlich langen Messungen zu reduzieren, ist ein Marker als Referenzpunkt auf der Welle erforderlich. Verschiedene Möglichkeiten, einen Marker zur erkennen sind in [8] beschrieben. Hier wird allerdings der Winkel auf einer drehenden Scheibe gemessen und die Marker wurden auf ein Papier gedruckt, dessen Oberfläche ein ähnlicher Sensor abtastet. Weitere Untersuchungen sollen geeignete Markermaterialien und Markerformen ermitteln, die für den Einsatz in Windenergieanlagen geeignet sind.

Weitere Einflussfaktoren auf die Messunsicherheit wie z.B. Beschleunigung der Oberfläche, Radius der Rotorwelle und damit die Krümmung der Oberfläche, Schwingungen der betrachteten Oberfläche sowie die Eigenschwingungen des Sensors und die Temperaturabhängigkeit sind Gegenstand noch laufender Untersuchungen.

5 Danksagung

Die Autoren danken dem Bundesministerium für Umwelt, Naturschutz und Reaktorsicherheit für die Förderung dieser Arbeiten im Rahmen des Projektes „Inside-Sensoring Labor für Großverzahnungen“ (FKZ 0325153).

Literatur

[1] Westerkamp, J. F.; Sorg, M.; Goch, G.: Inside-Sensoring Labor für Großverzahnungen. Forschungsjahrbuch Erneuerbare Energien 2009. FIZ Karlsruhe, 2010, S. 11.

[2] Ringe, M.; Patzelt, S.; Kruse, D.; Prekel, H.; Goch, G.: Measuring system for non-contact oscillation measurement. In: AMA Service GmbH (Hrsg.): Sensor+Test Conference 2009, Band 2, Wunstorf. 2009. - ISBN 978-3-9810993-5-5, S. 235-240.

[3] Tresanchez, M.; Palleja, T.; Teixido, M.; Palacin, J.: Using the image acquisition capabilities of the optical mouse sensor to build an absolute rotary encoder, Sensors and Actuators A: Physical 157 (2010) 1, S. 161-167.

[4] Minoni, U.; Signorini, A.: Low-cost optical motion sensors: An experimental characterization, Sensors and Actuators A: Physical 128 (2006) 2, S. 402-408.

[5] Ng, T. W.: The optical mouse as a two-dimensional displacement sensor, Sensors and Actuators A: Physical 107 (2003) 1, S. 21-25.

[6] Palacin, J.; Valganon, I.; Pernia, R.: The optical mouse for indoor mobile robot odometry measurement, Sensors and Actuators A: Physical 126 (2006) 1, S. 141-147.

[7] Popov, P.; Pulov, S.; Pulov, V.: A laser speckle pattern technique for designing an optical computer mouse, Optics and Lasers in Engineering 42 (2004) 1, S. 21-26.

[8] Tresanchez, M.; Palacin, J.; Teixido, M.; Palacin, J.: The optical mouse sensor as an incremental rotary encoder, Sensors and Actuators A: Physical 155 (2009) 1, S. 73-81.

[9] Bredemeier, C.: Entwicklung eines optischen Messsystems zur Erfassung von Drehwinkeln im Antriebsstrang von Windenergieanlagen. Hochschule Bremen, Fachbereich Elektrotechnik und Universität Bremen, Fachbereich Produktionstechnik, Diplomarbeit, 2010.

Auswertung flächenhafter Messungen an Großverzahnungen

Jan F. Westerkamp, Dirk Stöbener, Axel von Freyberg,
Martina Fuhrmann, Gert Goch

Bremer Institut für Messtechnik, Automatisierung und Qualitätswissenschaft,
Prof. Dr.-Ing. Gert Goch, Bremen
Tel.: 0421-218-64609, e-mail: wes@bimaq.de

Kurzfassung

Die Lager und Getriebe von Windenergieanlagen sind extremen Belastungen ausgesetzt, was spezielle Ansprüche an Material, Fertigungsprozesse und Messtechnik stellt. Für Aussagen zu den Ursachen von Geräuschemissionen und Schwingungen, ihren schädigenden Auswirkungen auf die Komponenten sowie den wiederum durch die Schäden verursachten Schwingungen sind detaillierte Kenntnisse über die Zahnradgeometrie erforderlich.

Das Ziel der Forschungen im Labor für Großverzahnungs-Messungen ist es, die Zuverlässigkeit von Großgetrieben in Windenergieanlagen entscheidend zu verbessern. Im Zuge dieser Arbeiten ist am BIMAQ eine Software-Toolbox für die Auswertung von Verzahnungsmessungen entstanden. Sie ermöglicht die flächenhafte Betrachtung von Zahnradabweichungen und die Bestimmung von neu definierten, flächenhaften Abweichungsparametern. Damit bietet sie im Vergleich zum aktuellen Mess- und Auswerte-Standard detailliertere Erkenntnisse zur Zahnradgeometrie bzw. ihrer Verformung.

1 Einleitung

Bei der Herstellung und beim Betrieb von Antriebstrang-Komponenten, zum Beispiel für Kraftwerke oder Schiffe, bestehen langjährige Erfahrungen, welche sich nur bedingt auf Windenergieanlagen (WEA) übertragen lassen. Besonders die WEA-Lager und -Getriebe sind extremen Belastungen ausgesetzt, was spezielle Ansprüche an Material, Fertigungsprozesse und Messtechnik stellt. Con-

dition Monitoring Systeme erfassen zwar schon heute eine Vielzahl von Messgrößen wie Schwingungen und Beschleunigungen der WEA-Komponenten. Eine aussagekräftige, messtechnische Analyse und zuverlässige Diagnose auf Basis dieser Messdaten ist jedoch zurzeit nicht immer möglich.

Für Aussagen zu den Ursachen von Geräuschemissionen und Schwingungen, ihren schädigenden Auswirkungen auf die Komponenten sowie den wiederum durch die Schäden verursachten Schwingungen sind detaillierte Kenntnisse über die Zahnradgeometrie erforderlich.

Das BIMAQ baut in einem vom Bundesumweltministerium geförderten Projekt zurzeit ein Labor für Großverzahnungs-Messungen auf [1]. Mit dieser Infrastruktur lassen sich Zahnräder bis zu einem Durchmesser von 3,0 Meter hochgenau messen. Weitere Sensoren, z. B. zur hochauflösenden Messung des Drehwinkels, entwickelt das Inside-Sensoring Projekt des BIMAQ [2]. Daraus ergeben sich neue Erkenntnisse zu den Zusammenhängen zwischen Auslegung, Fertigung, Produktqualität und Funktionseigenschaften sowie zur Restlebensdauer von Großgetrieben in Windenergieanlagen. Ziel der Forschungen in dem Labor ist es, die Zuverlässigkeit von Großgetrieben in Windenergieanlagen entscheidend zu verbessern.

Im Zuge dieser Arbeiten ist eine Software-Toolbox für die Auswertung von Verzahnungsmessungen entwickelt worden. Sie ermöglicht die flächenhafte Betrachtung von Zahnradabweichungen und die Bestimmung von neu definierten, flächenhaften Abweichungsparametern. Damit bietet sie im Vergleich zum aktuellen Mess- und Auswerte-Standard detailliertere Erkenntnisse zur Zahnradgeometrie bzw. ihrer Verformung.

2 Mess-Strategie

Für die flächenhafte Auswertung von Zahnradgeometrien sind entsprechende Messdaten erforderlich. In der aktuellen industriellen Anwendung erfolgt die Qualitätsprüfung von Zahnrädern üblicherweise vergleichsweise lückenhaft durch Messung einer Stichprobe von vier Zähnen je Zahnrad. An jedem der Zähne werden je Flanke eine Profil- und eine Flankenlinie aufgenommen. Zusätzliche Sondermessungen erlauben beispielsweise die Auswertung von Teilungsabweichungen und Zahndicken. Diese standardisierte Messung ist jedoch nicht geeignet, um weitergehende Informationen zu Lageabweichungen aller

Zahnflanken sowie zu Verschränkungen und Verzugseffekten zu gewinnen. Für den flächenhaften Auswerteansatz der entwickelten Toolbox ist eine Erweiterung des Messumfanges erforderlich:

- Anzahl gemessener Zähne: erhöht von 4 auf alle Zähne, Messung auf beiden Zahnflanken
- Erhöhung der Anzahl der gemessenen Profillinien pro Flanke auf beispielsweise 9, verteilt über die gesamte Zahnbreite
- Erhöhung der Anzahl der gemessenen Flankenlinien pro Flanke auf beispielsweise 3, verteilt über die gesamte Zahnhöhe

Zusätzlich sieht die Mess-Strategie auch Messungen auf dem Zahnradgrundkörper vor, um den Einfluss von Grundkörper-Verformungen auf die Verzahnung erfassen und bei der Zahnradbewertung berücksichtigen zu können:

- Messung mehrerer Kreise auf beiden Stirnseiten
- Kreismessungen in der Zahnradbohrung bzw. auf dem Schaft
- Messung eines Punktes auf jedem Zahnkopf (äußere Zahnradmantelfläche, gebildet durch die Zahnköpfe)
- Messung je einer radialen Linie in der Mitte jedes Zahns auf beiden Stirnseiten

Der sich aus diesen Messungen ergebende Datensatz stellt ein enges Netz aus bis zu 100.000 Messpunkten dar (Bild 1). Die Erfassung dieser Datenmenge ist im Vergleich zum industriellen Standard jedoch deutlich zeitaufwändiger, bietet aber den entscheidenden Vorteil, eine flächenhafte Zahnradbewertung zu ermöglichen, die z.B. auch Zahnverschränkungen umfasst.

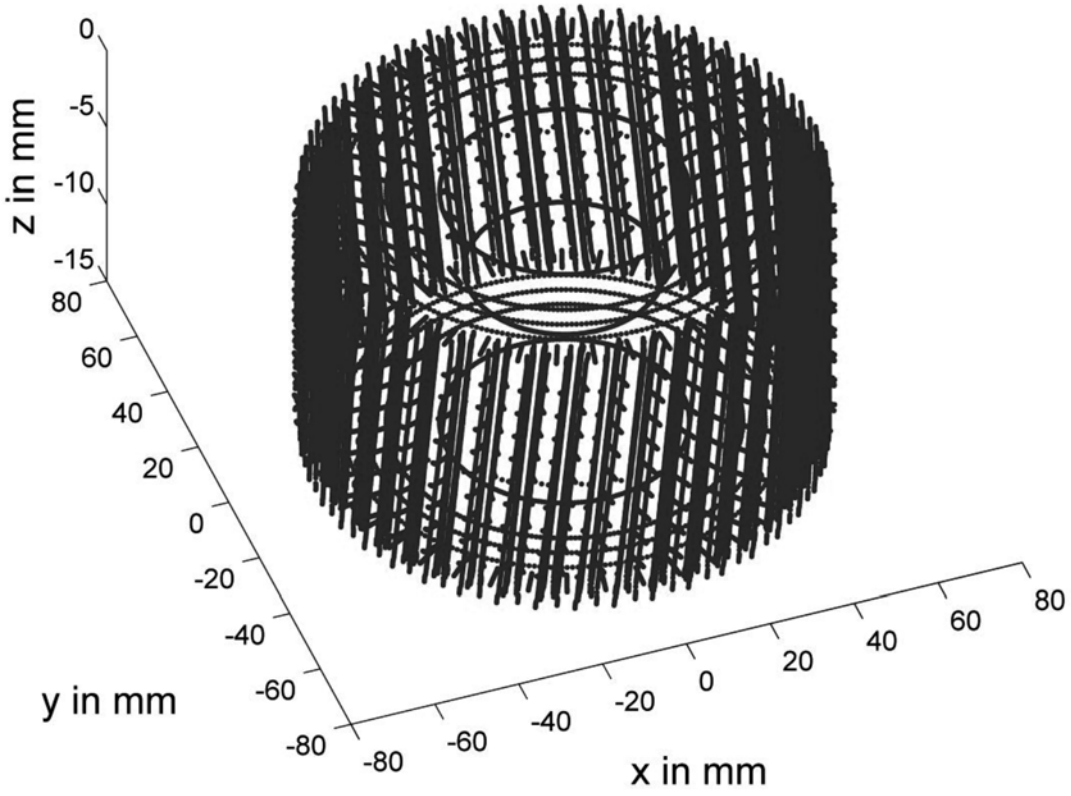


Bild 1: Messpunkte-Netz gemäß der für die Toolbox verwendeten Mess-Strategie.

3 Software-Toolbox

Die zur detaillierten Analyse von Geometrieabweichungen an evolventischen Zahnrädern im BIMAQ entwickelte Software-Toolbox zur flächigen Verzahnungsauswertung besteht zurzeit aus zwei Komponenten mit diversen Unterfunktionen:

- Algorithmen zur Berechnung von Abweichungen zwischen Messdaten und idealen, nicht-modifizierten Verzahnungsgeometrien,
- Simulation von verzogenen bzw. verformten Zahnrädern zur Analyse der Abweichungsursachen.

Diese Komponenten werden in den folgenden Unterkapiteln näher beschrieben.

Berechnung von Geometrieabweichungen

Für nicht-modifizierte Evolventenverzahnungen können die orthogonalen Abweichungen zu den Zahnflanken analytisch berechnet werden [3, 4, 5, 6, 7]. Die Visualisierung dieser Abweichungen, codiert als Graustufen, ist in Bild 2 und Bild 3 dargestellt.

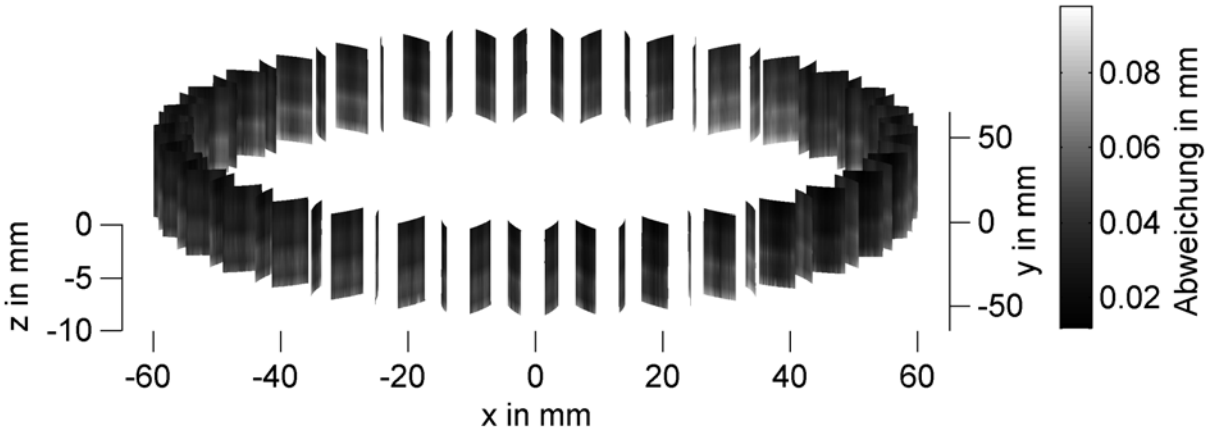


Bild 2: Auswertung von Flankenabweichungen auf Basis einer dreidimensionalen Approximation. Die geometrischen Abweichungen (in mm) sind in Graustufen dargestellt.

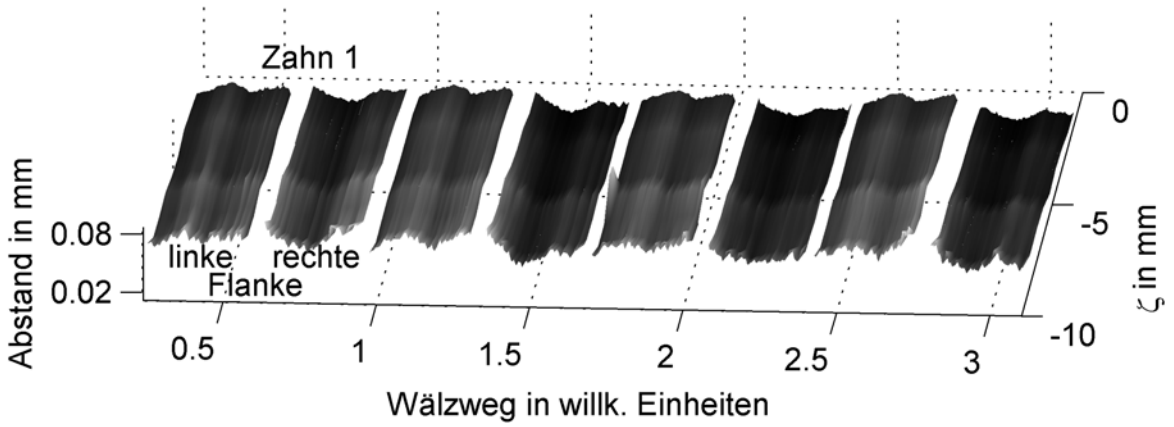


Bild 3: Darstellung der Flankenabweichungen für die ersten 4 Zähne (jeweils linke und rechte Flanke) des Zahnrades aus Bild 2 als Graustufen.

Für die Auswertung flächiger Abweichungs-Kenngrößen werden flankenweise die Auswertebereiche L_α und L_β (in Profillinien- (ξ-) und Flankenlinien- (ζ-)

Richtung) analog zu [8] festgelegt und Ebenen in die Abweichungen approximiert. Das Ergebnis ist an einem Beispiel in Bild 4 gezeigt.

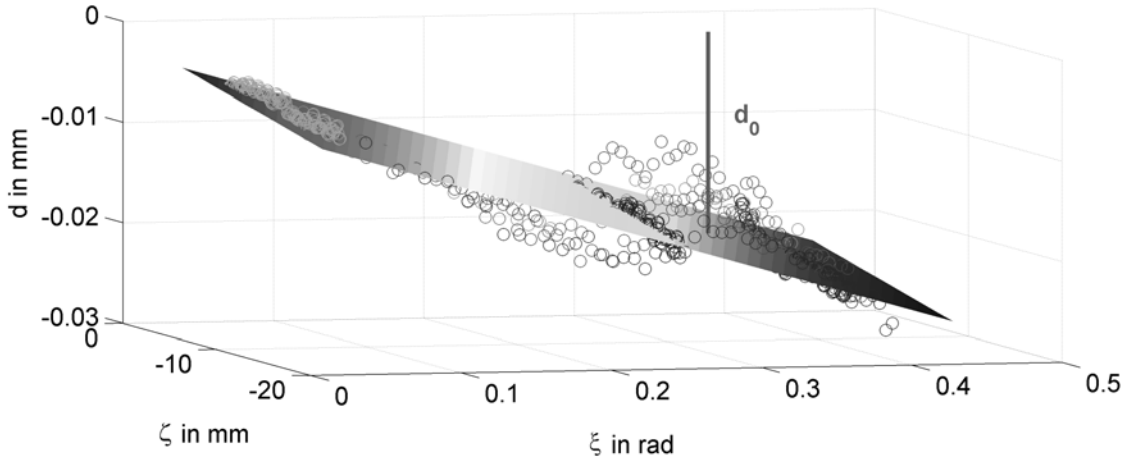


Bild 4: Approximation einer Ebene in die orthogonalen Abweichungen *d* einer Flanke.

Der Abstand d_0 der approximierten Ebene (in Höhe der Zahnmitte auf dem Radius des Teilkreises, siehe Bild 4) zur ξ-ζ-Ebene kann als ein Maß für Teilungsabweichungen interpretiert werden. Der Wert gibt nicht nur einen Anhaltspunkt auf die Flankenlageabweichungen der Zähne relativ zueinander, sondern beinhaltet auch Informationen über den Einzelzahn. Darüber hinaus werden die normgemäßen Parameter Flankenlinien- und Profillinien-Gesamtabweichung kombiniert zu einer flächigen Flanken-Gesamtabweichung F_a. Dieser Wert beziffert die Differenz zwischen der größten und der geringsten Abweichung innerhalb der Gesamtfläche des Auswertebereiches auf der Flanke. Die weitere Charakterisierung der Zahnradabweichungen erfolgt anhand der Parameter flächige Profilwinkelabweichung $f_{H\alpha a}$ als Steigung der approximierten Ebene in Profillinienrichtung und flächige Flankenwinkelabweichung $f_{H\beta a}$ als Steigung der Ebene in Flankenlinienrichtung sowie Flankenformabweichung f_{fa} als Kombination der genormten Profillinien- und Flankenlinien-Formabweichungen $f_{f\alpha}$ bzw. $f_{f\beta}$. Die Flankenformabweichung berechnet sich als Differenz der Abstände zweier maximaler Abweichungs-Datenpunkte von der Ebene innerhalb des gesamten Auswertebereiches.

3.2 Simulation verformter Zahnradgeometrien

Abweichungen der Verzahnungsgeometrie sind nicht allein dem Verzahnungsbereich eines Bauteils anzulasten, sondern entstehen vielmehr auch aus Verzügen und Abweichungen im Grundkörper [9, 10, 11]. Simulationen berechnen Datenpunkte definiert verzogener Zahnräder. Deren Auswertung ermöglicht Rückschlüsse auf die Auswirkungen der bekannten Grundkörperverzüge auf die im vorigen Kapitel beschriebenen Parameter. Entsprechend den bisherigen Erkenntnissen [9, 10] werden zunächst Datenwolken für oval verformte und skalierte Zahnräder sowie für zu einer Schalenform gewölbten Werkstücken (Dishing) erstellt.

Die folgenden Gleichungen (1) sowie Bild 5 verdeutlichen die Grundlagen zur Simulation von Ovalität (definiert durch die Amplitude a_0 und die Winkellage $\Delta\varphi$) und Maßänderung (definiert durch die Radiusänderung Δr) am Zahnrad. Die Transformationen werden auf die Geometrie eines idealen Zahnrades angewendet.

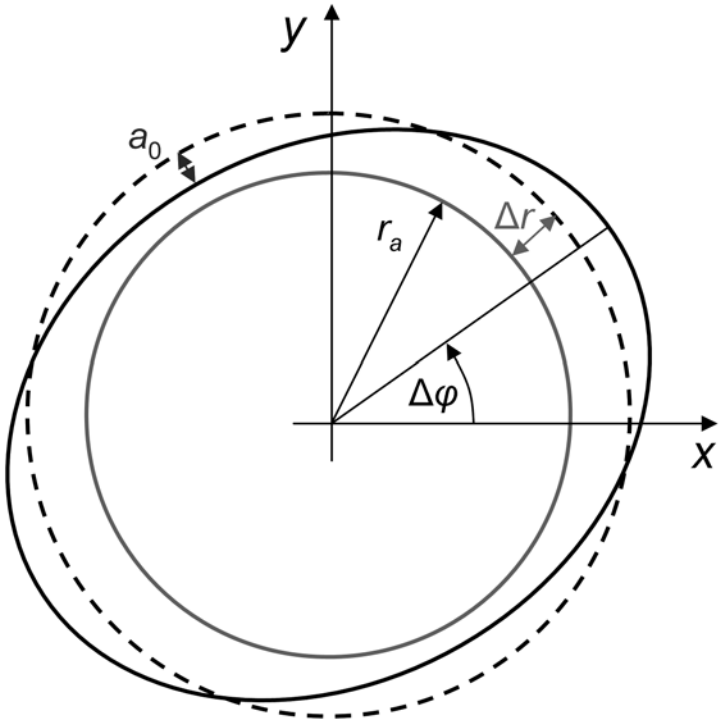


Bild 5: Modellierung von Maßabweichung und Ovalität eines Zahnrades.

$$r_v = r \cdot \left(1 + s_r + s_0 \cdot cos\left(2 \cdot \left(\varphi - \Delta\varphi\right)\right)\right)$$
$$s_r = \frac{\Delta r}{r_a}; \quad s_0 = \frac{a_0}{r_a} \tag{1}$$

Das Dishing wird anhand eines Zahnrades modelliert, das sich an eine Kugel mit sehr großem Durchmesser anschmiegt. Die folgenden Transformationsgleichungen (2) überführen die Datenpunkte eines idealen Zahnrades in ein „gedishtes“ Zahnrad entsprechend Bild 6. Dargestellt sind die Schnittbilder eines idealen (gestrichelt) und des verformten (durchgezogen) Zahnrades. Parametriert wird das Dishing durch den Kugeldurchmesser z_0.

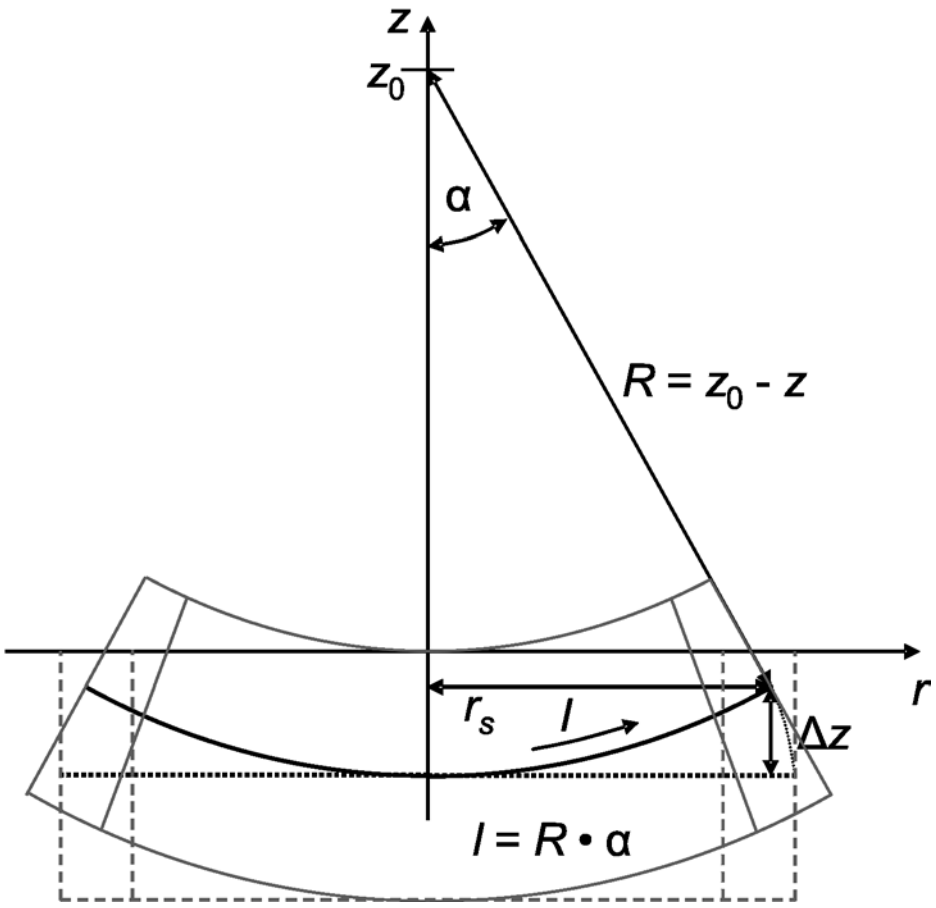


Bild 6: Modellierung des Dishings eines Zahnrades.

$$\begin{bmatrix} x_d \\ y_d \\ z_d \end{bmatrix} = \begin{bmatrix} |z_0 - z| \cdot sin\left(\frac{\sqrt{x^2 + y^2}}{|z_0 - z|} \right) \cdot \frac{x}{\sqrt{x^2 + y^2}} \\ |z_0 - z| \cdot sin\left(\frac{\sqrt{x^2 + y^2}}{|z_0 - z|} \right) \cdot \frac{y}{\sqrt{x^2 + y^2}} \\ z_0 - (z_0 - z) \cdot cos\left(\frac{\sqrt{x^2 + y^2}}{|z_0 - z|} \right) \end{bmatrix} \quad (2)$$

4 Anwendungen

Im Rahmen der flächigen Auswertung von Verzahnungsabweichungen wird die flächige Teilungsabweichung d_0 berechnet. Wie in Bild 7 dargestellt, korreliert dieser Parameter mit der normgemäßen Summenteilungsabweichung (nach Abzug des Mittelwertes und einer Vorzeichenkorrektur). Der Mittelwert der flächigen Teilungsabweichung enthält aber zusätzliche Informationen (z. B. ein Maß für die mittlere Zahndickenabweichung am Teilkreis), welche für die Analyse von Verzugsmechanismen hilfreich sind [12].

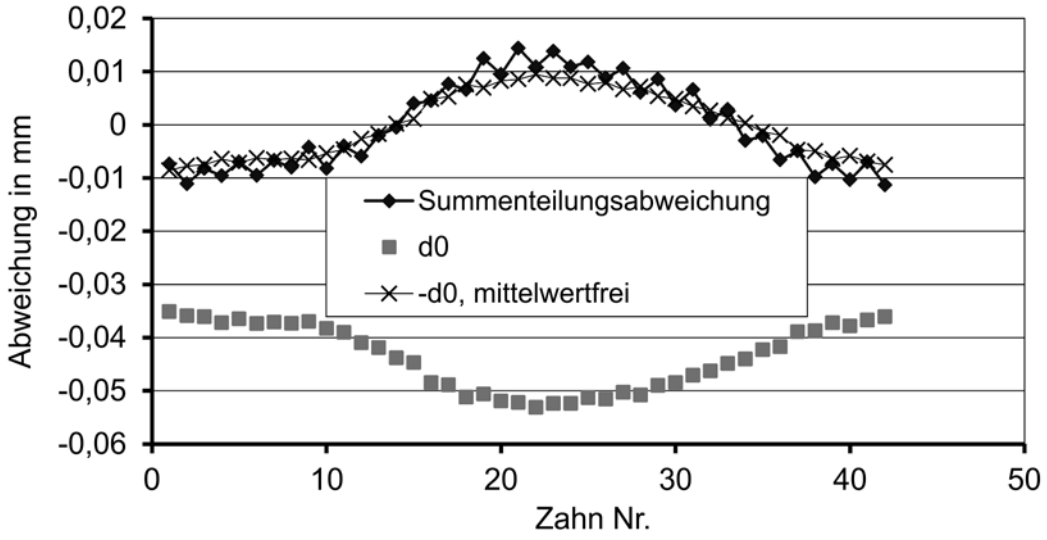


Bild 7: Vergleich der Summenteilungsabweichung mit der flächigen Teilungsabweichung d_0 der linken Flanken eines Zahnrades (Nr. 31266 vor der Wärmebehandlung).

Erste Auswertungen an simulierten Zahnraddaten sind sehr viel versprechend. Bild 8 zeigt beispielhaft das Ergebnis der flächig ausgewerteten Flankenlinien-Winkelabweichung an einem gradverzahnten Zahnrad auf Grundlage gemessener (Experiment) und simulierter Daten mit Dishing. Die Ergebnisse der Simulation bilden die Realität in diesem Fall sehr gut ab.

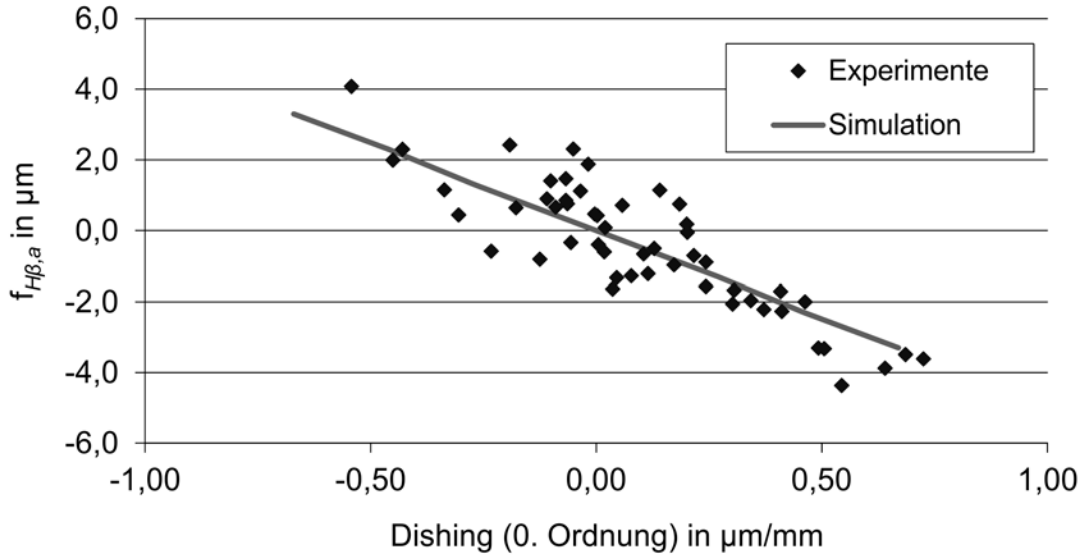


Bild 8: Vergleich der simulierten flächigen Flankenlinien-Winkelabweichung $f_{H\beta,a}$ gedishter, geradverzahnter Zahnräder mit experimentellen Messergebnissen.

Zukünftig steht die Toolbox mit ihrer umfangreichen Funktionsbibliothek zur Auswertung flächig gemessener Verzahnungsdaten zur Verfügung. Durch die flächige Messung und die Auswertung unter Berücksichtigung aller gültigen Messpunkte innerhalb des Auswertebereiches können gegenüber dem heutigen normgemäßen Ansatz deutliche Mehrinformationen ohne Sondermessungen (siehe Zahndickenabweichung) generiert werden. Darüber hinaus sind Rückschlüsse auf Abweichungsursachen möglich. Dies gibt beispielsweise bei der Produktionsoptimierung einen Hinweis auf die genauer zu untersuchenden Prozessschritte.

5 Danksagungen

Die Autoren danken der Deutschen Forschungsgemeinschaft für die Förderung dieser Arbeiten im Rahmen des SFB 570 „Distortion Engineering“, dem Bun-

desministerium für Umwelt, Naturschutz und Reaktorsicherheit für die Förderung im Rahmen des Projektes „Labor für Großverzahnungsmessungen“ (FKZ 0325140) sowie der Machine Tool Technologies Research Foundation (MTTRF).

Literatur

[1] Henske, G.; Lübke, K.; Freyberg, A.; Goch, G.: Labor für Großverzahnungsmessungen. Forschungsjahrbuch Erneuerbare Energien 2009. FIZ Karlsruhe, 2010, S. 12.

[2] Westerkamp, J.; Sorg, M.; Goch, G.: Inside-Sensoring Labor für Großverzahnungen. Forschungsjahrbuch Erneuerbare Energien 2009. FIZ Karlsruhe, 2010, S. 11.

[3] Stöbener, D.; von Freyberg, A.; Fuhrmann, M.; Goch, G.: Characterisation of Gear Distortions with Areal Parameters. In: Proc. 3rd Intern. Conference on Distortion Engineering, 14.-16. Sept. 2011, Bremen, akzeptiert.

[4] Goch, G.: Gear metrology, Annals of the CIRP 52 (2003) 2, S. 659-695.

[5] Günther, A.; Peters, J.; Goch, G.: Flächenhafte numerische Beschreibung, Ausrichtung und Auswertung von Zylinderrädern, tm - Technisches Messen 68 (2001) 4, S. 160-165.

[6] Günther, A.: Flächenhafte Beschreibung und Ausrichtung von Zylinderrädern mit Evolventenprofil. Universität Ulm, Abteilung Mess-, Regel- und Mikrotechnik, Diplomarbeit, 1996.

[7] Lotze, W.: Zahnradmessung mit Koordinatenmessgeräten. 3. Auflage. Eigenverlag, Dresden, 2006. - ISBN 3-00-015962-2.

[8] NORM VDI/VDE-Richtlinie 2607, 2000. Rechnergestützte Auswertung von Profil und Flankenlinienmessungen an Zylinderrädern mit Evolventenprofil.

[9] Clausen, B.; Frerichs, F.; Kohlhoff, T.; Lübben, T.; Prinz, C.; Rentsch, R.; Sölter, J.; Surm, H.; Stöbener, D.; Klein, D.: Identification of process parameters affecting distortion of disks for gear manufacture - Part I: casting, forming and machining, Materialwissenschaft und Werkstofftechnik 40 (2009) 5-6, S. 354-360.

[10] Clausen, B.; Frerichs, F.; Kohlhoff, T.; Lübben, T.; Prinz, C.; Rentsch, R.; Sölter, J.; Surm, H.; Stöbener, D.; Klein, D.: Identification of process

parameters affecting distortion of disks for gear manufacture - Part II: heating, carburizing, quenching, Materialwissenschaft und Werkstofftechnik 40 (2009) 5-6, S. 361-367.

[11] Steinbacher, M.; Clausen, B.; Hoffmann, F.; Zoch, H.-W.: Influencing parameters on the distortion of disk-shaped gear wheel base body derived from carburization. In: Proceedings of the IFHTSE Conference “New Challenges in Heat Treatment and Surface Engineering”, 9.-12. Juni 2009, Dubrovnik-Cavtat, Croatia. S. 349 - 349.

[12] Steinbacher, M.; Surm, H.; Clausen, B.; Lübben, T.; Hoffmann, F.: Methodical investigation of distortion biasing parameters during case hardening of spur wheels. In: Proc. 3rd Intern. Conference on Distortion Engineering, 14.-16. Sept. 2011, Bremen, akzeptiert.