

# Verbesserung der Auffindwahrscheinlichkeit (POD) von Oberflächenfehlern in Metallerzeugnissen mittels optischer Inspektionsverfahren und Bildverarbeitung

Markus RAUHUT, Martin SPIES, Kai TAEUBNER  
Fraunhofer-Institut für Techno- und Wirtschaftsmathematik (ITWM)  
Abteilung Bildverarbeitung, 67663 Kaiserslautern

**Kurzfassung.** Berührungslos arbeitende optische Verfahren zur Oberflächenprüfung werden in einer Vielzahl von Produktionsanlagen zur Qualitätskontrolle eingesetzt. Für die quantitative Bewertung von optischen Inspektionssystemen bietet sich die POD-Analyse (englisch: Probability of Detection, POD) zur Bestimmung der Fehlerauffindwahrscheinlichkeit an. In diesem Beitrag konzentrieren wir uns auf die Analyse und Bewertung verschiedener Bildverarbeitungsalgorithmen mittels der POD-Analyse.

## 1. Einführung

Bei vielen metallischen Produkten wie beispielsweise Stanzteilen oder beschichteten Metallabdichtungen muss die Unversehrtheit der Oberfläche überprüft werden, um das bestimmungsgemäße Betriebsverhalten des jeweiligen Bauteils zu gewährleisten. Abhängig von der Güte und der Komplexität des Beleuchtungs- und (Video-) Kamerasystems sowie der nachgeschalteten Bildverarbeitungsalgorithmen werden unterschiedliche Fähigkeiten im Hinblick auf die Fehlerdetektion und auf die Fehlergrößenbestimmung erreicht. Um verschiedene, in der Praxis eingesetzte Systeme quantitativ bewerten und neue Systeme gezielt zusammenstellen zu können, wenden wir  $\alpha$ -Analysen zur Bestimmung der Fehlerauffindwahrscheinlichkeit (englisch: Probability of Detection, POD) an. Dem internationalen Standard MIL-HDBK-1823A in der aktuellen Version von 2009 folgend ermitteln wir POD-Kurven aus den an Testkörpern beispielsweise mit Bohrungen und Nuten unterschiedlicher Dimensionierung und Orientierung aufgenommenen, digitalisierten und algorithmisch verarbeiteten Daten. Am Beispiel von texturierten Oberflächen zeigen wir, dass applikationsspezifisch zusammengestellte Algorithmik zu erheblichen Verbesserungen der POD führen kann (siehe [7] für POD und Sensorik). Wir sehen darin einen wichtigen Beitrag im Hinblick auf den Einsatz optischer Oberflächeninspektionsverfahren zur quantitativen ZfP im Bereich der automatischen Qualitätskontrolle.

## 2. Fehlerauffindwahrscheinlichkeit (POD)

Das Konzept der Fehlerauffindwahrscheinlichkeit (POD) stellt einen wichtigen Teil der Untersuchung und Evaluierung der Integrität eines Bauteils dar [1-3]. Die POD bezeichnet die Wahrscheinlichkeit, einen Fehler im Bauteil zu finden; sie wird in diesem Beitrag als Funktion der Fehlergröße  $a$  bestimmt. Die resultierende POD-Kurve liefert zusammen mit den auferlegten Konfidenzintervallen die Fehlergröße, die mit einer ‚vernünftigen‘ Wahrscheinlichkeit detektiert werden kann. Diese Fehlergröße wird dann mit den Anforderungen an die Bauteilintegrität verglichen.

Der prinzipielle Verlauf der POD-Kurve zeigt, dass mit zunehmender Fehlergröße auch die Detektionswahrscheinlichkeit ansteigt. An der Größe  $a_{90/95}$  schneidet die untere 95 % Konfidenzgrenze das 90 % POD-Niveau. Diese Größe wird üblicherweise als die Fehlergröße betrachtet, die sicher zu detektieren ist. Um experimentell POD-Kurven zu ermitteln, müssen wohl definierte Inspektionen an geeigneten Testkörpern durchgeführt werden. Die Vorgehensweise, eine POD mittels einer  $\hat{a}$  versus  $a$ -Analyse zu bestimmen, ist wie folgt: ein Fehler der Größe  $a$  erzeugt ein Signal der Amplitude (in unserem Fall einer Metrik)  $\hat{a}$ . Üblicherweise wird ein Schwellwert  $\hat{a}_{th}$  definiert, der die kleinste vom Prüfsystem aufgezeichnete Amplitude darstellt, d.h. unterhalb dieses Wertes ist das Signal nicht mehr vom Rauschen zu unterscheiden. Die zweite Schwelle ist der Entscheidungsschwellwert  $\hat{a}_{dec}$ , oberhalb dessen das Signal als ‚Treffer‘ interpretiert wird. Die Schwellwert  $\hat{a}_{th}$  ist immer kleiner oder gleich dem Schwellwert  $\hat{a}_{dec}$ . Unter der Annahme, dass die Signalamplituden/Metriken statistisch normalverteilt sind, kann das  $\hat{a}$  versus  $a$ -Diagramm in eine POD-Kurve überführt werden. Eine detaillierte mathematische Beschreibung dieser Prozedur ist in [1, Appendix G] zu finden.

## 3. Optische Inspektionsverfahren

### 3.1 Hintergrund und Problemstellung

Bei optischen Inspektionssystemen, vor allem bei Oberflächeninspektionssystemen, entscheidet hauptsächlich das Erreichen der vom Endanwender vorgegebenen Detektionsrate und Auffindwahrscheinlichkeit darüber, ob ein solches System in der Praxis einsetzbar ist. Typischerweise werden diese Parameter in Form einer Vorstudie empirisch festgestellt, d.h. anhand von Musterteilen wird mehr oder weniger subjektiv entschieden, ob die Kundenanforderungen erfüllbar sind [3].

Die Detektionsqualität wird in der Bildverarbeitung hauptsächlich als Eigenschaft von Bildsensor (Kamera) und Beleuchtung gesehen. Durch die geeignete Wahl der Beleuchtung und dem Aufnahmewinkel zwischen Kamera und Beleuchtung können bestimmte Arten von Defekten auf einer Oberfläche stärker sichtbar und somit detektierbar gemacht werden.

Die Konfiguration von Beleuchtung und Kameratechnik ist in der Bildverarbeitung ausreichend bekannt und wird dementsprechend systematisch konzipiert (siehe dazu auch [4]). Es ist sicherlich auch wahr, dass die eigentliche Algorithmik die Detektionsqualität nur bedingt verbessern kann. Geht man aber davon aus, dass die Konfiguration von Sensorik und Beleuchtung optimal gewählt ist, ist die Leistungsfähigkeit der Bildverarbeitungsalgorithmen ein wichtiger Aspekt bei der Betrachtung der Gesamtqualität. In den letzten Jahren wurde eine Vielzahl von Algorithmen entwickelt, die wesentlich effektiver und effizienter arbeiten, als "einfache" Verfahren wie z.B. Kantendetektoren. Hinzu kommt, dass durch die Steigerung der Leistung heutiger Rechner komplexere Verfahren zum Einsatz kommen können. Ein Problem bei diesen "neuen" komplexen

Algorithmen ist jedoch, dass Bilddaten die für das menschliche Auge optimal sind, nicht unbedingt optimal für die Algorithmik sind.

Deshalb ist es für den Entwickler einer Bildverarbeitungssoftware wichtig, ein Werkzeug zu haben, das eine quantitative Beurteilung der Algorithmen ermöglicht. Hier fehlt es jedoch an Verfahren, die quantitative Werte liefern. Dementsprechend hoch ist für den Bildverarbeiter das Risiko, ein praxistaugliches Inspektionssystem realisieren zu können.

Die POD-Analyse ist eine Möglichkeit, die Qualität eines oder mehrerer Algorithmen zu beurteilen. Bevor diese Analyse eingesetzt werden kann sind einige Vorarbeiten notwendig:

### 3.2 Objektivierungsansatz

Die POD-Analyse stellt eine Möglichkeit dar, die minimalen Fehlergrößen zu einer hohen Detektionswahrscheinlichkeit quantitativ zu berechnen. Als Eingabeparameter für die POD-Analyse musste eine Metrik entwickelt werden, die die Detektionswahrscheinlichkeit eines Defektes erfasst. Diese Metrik sollte dabei folgende Eigenschaften haben:

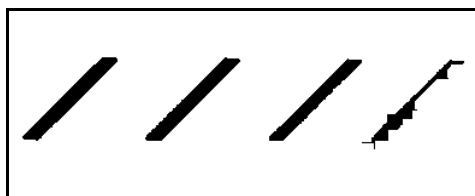
- Die Metrik soll einfach zu berechnen sein, damit schon in der Vorstudie eine quantitative Aussage zur Detektionswahrscheinlichkeit ermittelt werden kann.
- Die Metrik möglichst unabhängig von den verwendeten Algorithmen sein, aber dennoch die Eigenschaften typischer Analyseverfahren berücksichtigen.

Für diese Arbeit haben wir uns auf zwei Defekttypen konzentriert: Risse und Löcher, die durch Nuten und Bohrungen simuliert werden.

### 3.3 Eine Metrik zur Berechnung der Detektionswahrscheinlichkeit

Basis für die POD-Analyse ist eine Metrik, die beschreibt wie sicher ein Defekt erkannt wurde. Vereinfacht kann man für Bildverarbeitungsverfahren sagen, dass wenn sich die Pixelwerte der Defekte lokal stark von der Oberfläche (dem Gut-Bereich) unterscheiden, dann ist das Verfahren stabil. Der Unterschied von komplexen Verfahren zu den einfachen Algorithmen ist dabei, dass komplexe Verfahren lokal adaptiv rechnen, d.h. der Unterschied der Pixelwerte muss nicht global groß sein, sondern nur in der Umgebung des Defekts. Einfache Verfahren arbeiten sehr effektiv, wenn der globale Pixelabstand zwischen Defekt und Gut-Region hoch ist.

Heutzutage werden effiziente Klassifikatoren eingesetzt, die anhand von geometrischen Eigenschaften des vorher detektierten Defekts wie z.B. der Form eine Validierung der potentiellen Defekte durchführen. Dies bedeutet aber, dass Detektionsalgorithmen nicht nur Defekte erkennen müssen, sondern auch deren Form korrekt bestimmen müssen. Gerade hier scheitern einfache Verfahren wie z.B. der Laplace-Operator häufig. Abbildung 1 ist ein Beispiel für eine fehlerhafte Formerkennung.

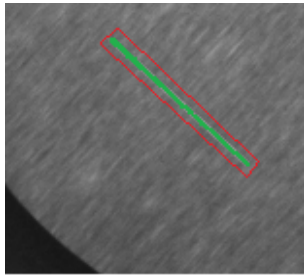


**Abbildung 1:** Beispiel für eine fehlerhafte Formerkennung. Der rechte Riss wurde erkannt, aber dessen Form nicht korrekt detektiert. Dies wird bei einer Klassifikation evtl. zu einem Fehler führen.

Letztendlich ist aber der lokale Kontrast eine gute Approximation für die Detektionswahrscheinlichkeit eines Algorithmus. Würde eine komplexere Metrik verwendet, könnte diese implizit die Verfahrensweise bestimmter Algorithmen bevorzugen.

Würde z.B. die 2. Ableitung der Pixelwerte zwischen Defekt und Gut-Region als Maß für die Detektionswahrscheinlichkeit verwendet, würden alle Verfahren welche die 2. Ableitung verwenden später in der POD-Analyse bevorzugt.

Aus diesen Beobachtungen lässt sich eine einfache Metrik zumindest für die Detektionswahrscheinlichkeit von Kratzern, Lunkern, Beulen, Dellen und Löchern ableiten, komplexere Fehler wie Texturabweichungen lassen sich nicht so einfach quantifizieren [6]. Wir definieren den lokalen Kontrast als  $K = Abs(Avg_D - Avg_E)$ , der Absolutdifferenz des Mittelwertes der Pixel in der Defektregion  $Avg_D$  und des Mittelwertes der Pixel in der näheren Umgebung  $Avg_E$ , z.B. mit einer Breite von vier Pixeln um den Defekt (siehe Abbildung 2).



**Abbildung 2:** Defektregion (grün) und lokale Umgebung (rot)

Da mit steigender Größe von Defekten im Bild auch die Detektionswahrscheinlichkeit steigt, scheint es sinnvoll die Fläche der Defektregion mit in die Gesamtmetriek eingehen zu lassen. Betrachtet man die folgenden Möglichkeiten bei der Detektion durch einen Algorithmus, wird klar, dass dies zu einer fehlerhaften Metrik führen würde:

1. Der Algorithmus markiert eine Region die **größer** als der wirkliche Defekt ist: Die Form des Defekts wurde nicht korrekt erkannt, aber durch die Multiplikation mit der Fläche ergibt sich evtl. eine höherer Wert als bei der korrekten Formerkennung .
2. Der Algorithmus markiert eine Region die **kleiner** als der wirkliche Defekt ist: Die Form des Defekts wurde nicht korrekt erkannt, aber der Mittelwert der Umgebung (roter Bereich in Abbildung 2) enthält nun Teile des Defektes, d.h. der lokale Kontrast wird evtl. kleiner. D.h. der lokale Kontrast gibt schon die schlechte Erkennung wieder.
3. Der Algorithmus markiert die Region korrekt. Dementsprechend sollte der lokale Kontrast die Metrik optimal wieder spiegeln.

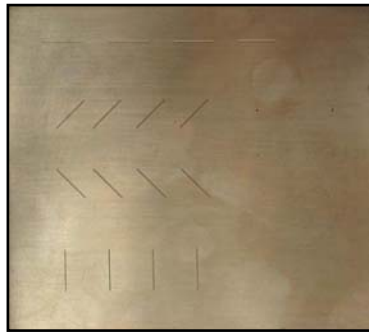
Aufgrund dieser Beobachtung wurde der lokale Kontrast  $K$  multipliziert mit einem Faktor als Eingangsgröße für die POD-Analyse verwendet. Da mit sinkender Auflösung die Detektionswahrscheinlichkeit fällt, wird mit diesem Faktor auf niedrige Auflösungen ein Malus gegeben. Die höchste Auflösung hat dabei den Wert 1.

#### 4. Inspektions-Setup und Durchführung der Messungen

Zur Detektion von Rissen und Bohrungen wurden folgende Komponenten verwendet:

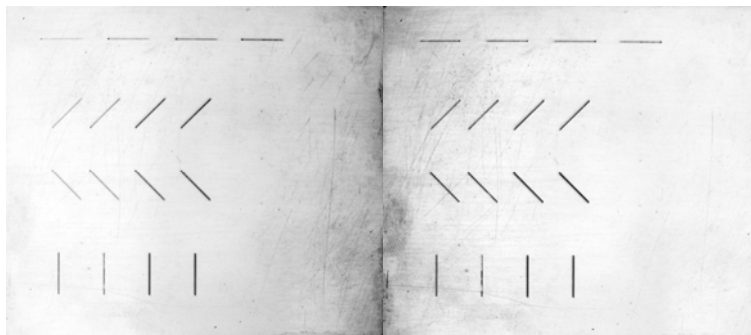
- Kamera: Prosilica GC1600M (1620 x 1220 Pixel), Pixelgröße 4.4  $\mu\text{m}$ ;
- Objektiv: Schneider Bad Kreuznach, Cingeon 1.4/12mm
- Beleuchtung: a) Seitenlicht mit Planistar Lichthaube (rote LEDs), b) koaxiales Auflicht mit Volpi ILP ACIS (weiße LEDs).
- Pixelauflösungen: 0,07 mm; 0,1 mm; 0,12mm. Die Auflösungen wurden durch eine Höhenverstellung der Kamera eingerichtet
- Aufnahmen: Für die Rissdetektion wurden zu jeder Pixelauflösung zwei Seitenlicht-Aufnahmen durchgeführt (Dunkelfeld), für die Bohrungsdetektion jeweils eine Aufnahme im Auflicht

Für die Detektion der Nuten und Bohrungen kam ein Messing-Prüfkörper (Abbildung 3) zum Einsatz.



**Abbildung 3:** Messing-Prüfkörper mit Nuten und Bohrungen

Die Nuten sind auf dem Prüfkörper in  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  und  $135^\circ$  angeordnet. Dies soll eine unbekannte Rissrichtung für den Algorithmus wieder spiegeln. Bei einer bekannten Richtung der Nuten wäre es möglich speziell angepasste Algorithmen zu entwerfen. Dies ist jedoch in der Praxis meist nicht gegeben. Die Nuten haben von links nach rechts, die Breiten: 0,13 mm; 0,14 mm; 0,22 mm; 0,25 mm wobei die Breite von oben nach unten in jeder "Spalte" identisch ist. Die Bohrungen haben von oben nach unten die Durchmesser 0,4 mm; 0,3 mm; 0,27 und 0,13 mm, wobei jeweils zwei Bohrungen pro "Zeile" den gleichen Durchmesser haben. Abbildung 4 zeigt die Eingabebilder zur Erkennung der Nuten.



**Abbildung 4:** Seitenlichtaufnahmen von oben und unten. Die Bilder wurden zur besseren Darstellung gedreht, invertiert und kontrastverstärkt.

## 5. Algorithmen

In diesem Abschnitt sollen zwei Algorithmen zur Erkennung der Nuten erläutert werden. Der Algorithmus zur Erkennung von Bohrungen wurde aus Platzgründen ausgelassen. In der ersten Variante wird aus den beiden Eingabebildern ein Gesamtbild erstellt und dann die Analyse durchgeführt. In Variante 2 wird pro Eingabebild ein Algorithmus gestartet und die Ergebnisse beider Verfahren zusammen geführt.

### Variante 1:

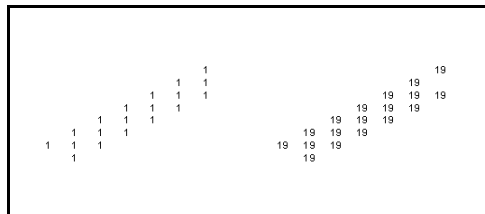
Diese Variante berechnet aus den beiden Seitenlichtaufnahmen ein Gesamtbild.

1. Der Maximum jedes Pixelpaars aus den beiden Eingabebildern ergibt das Gesamtbild.
2. Durch eine Schwellwertoperation wird das Bild segmentiert (siehe Abbildung 6)
3. Dieses Bild wird gelabelt, d.h. zusammengehörige Regionen erhalten die gleiche Zahl.
4. Das Labelbild wird so verändert, dass jede Region als Label ihre Größe (Anzahl Pixel) erhält.
5. Durch eine weitere Schwellwertoperation werden alle Regionen entfernt die unter einem gewissen Wert liegen (siehe Abbildung 7).

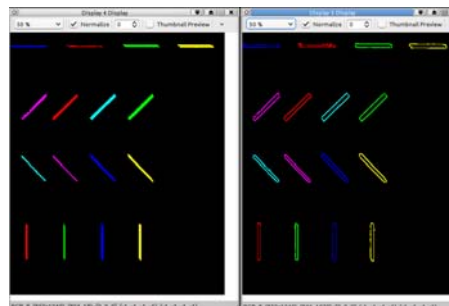
6. Das Bild wird erneut gelabelt.
7. Das Bild wird kopiert und dilatiert.
8. Das Ergebnis aus 7. wird mit dem Ergebnis aus 6. subtrahiert.
9. Nun sind beide Regionen für die Metrik vorhanden, d.h. es werden Regionenfeatures berechnet und die Metrik bestimmt (siehe Abbildung 8).



**Abbildung 6:** Das segmentierte Maximumbild



**Abbildung 7:** Durch einen Labeling-Algorithmus erhält jede Region eine Zahl. Links ist eine Region mit dem Label 1 zu sehen. Später wird dieses Label durch die Größe der Region ersetzt (hier 19). Durch eine Schwellwertoperation können nun alle größeren oder kleineren Region aus dem Bild entfernt werden.



**Abbildung 8:** Links zu sehen die detektierten Defekte, rechts die Umgebung die für die Metrik-Berechnung verwendet wird.

### Variante 2:

Diese Variante rechnet auf jeweils einem Eingabebild.

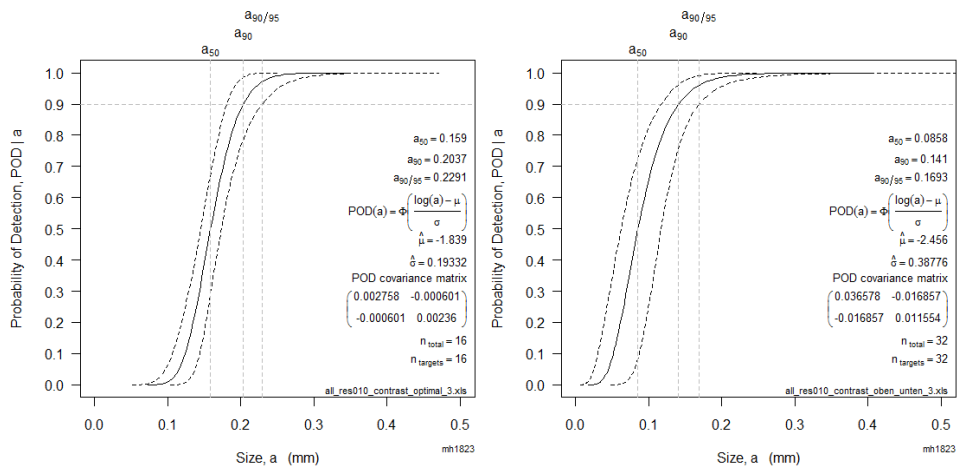
1. Auf dem Eingabebild wird eine Shadingkorrektur durchgeführt. Durch einen sehr großen Mittelwertfilter (Maske 100x100) wird eine Kopie des Eingabebilds geglättet und dann vom Eingabebild dieses Bild abgezogen.
2. Danach führt der Algorithmus die Schritte 2.-9. von Variante 1 durch (mit anderen Parametern).

### 5. POD Berechnung

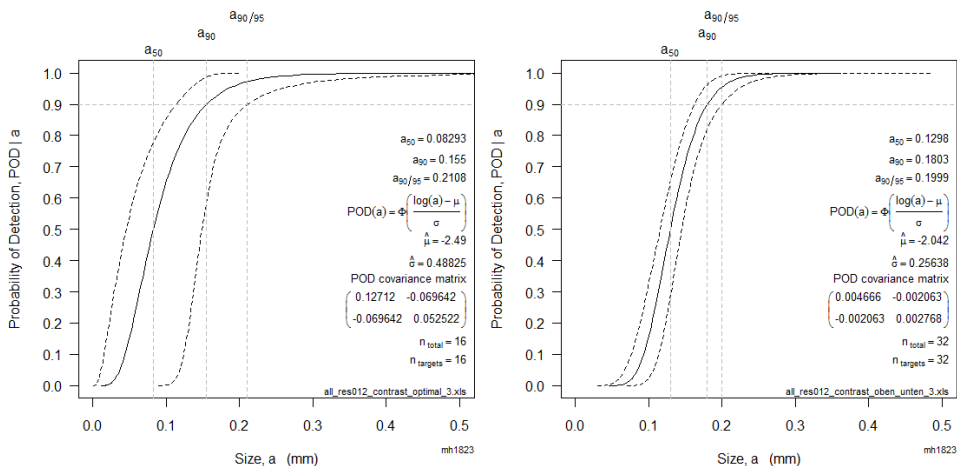
Die derart verarbeiteten Datensätze haben wir mit der mh1823 POD-Software (Version 2.5.2, 2009 Update des MIL-HDBK-1823 unter <http://StatisticalEngineering.com>) analysiert. Bei der Anwendung der Software ist folgendes zu beachten: (1) das eingesetzte ZfP-Verfahren/-System muss ein quantitatives Signal  $\hat{a}$  liefern; (2) die Testkörper müssen ‚Ziele‘ mit messbaren Charakteristiken haben, in unserem Fall ist dies die Metrik; (3) die

Software geht davon aus, dass die Input-Daten korrekt sind, also dass die angegebene Größe die korrekte Größe ist und die zugehörige Response die korrekte Response – ist dies nicht der Fall, liefert die Software ‚nur‘ Näherungsergebnisse.

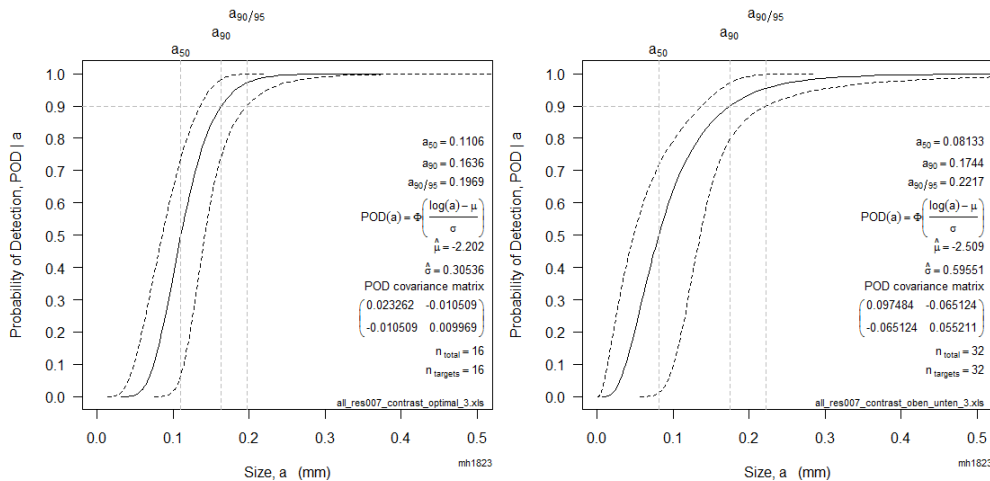
Der erste Schritt in der POD-Analyse ist die Darstellung der Daten in der Form von  $\hat{a}$  vs  $a$ -,  $\hat{a}$  vs  $\log(a)$ -,  $\log(\hat{a})$  vs  $a$ - und  $\log(\hat{a})$  vs  $\log(a)$ -Plots und der Auswahl des einen linearen Zusammenhang am Besten annähernden Modells. Die mh1823 POD-Software stellt die linearen Approximationen der Daten zur Auswahl dar. Die entsprechenden POD-Kurven werden mit den wichtigen Informationen des angewandten Modells dargestellt: die Modellparameter und ihre Kovarianzmatrix; die Fehlergröße  $a_{50}$ , die Größe mit 50 % POD;  $a_{90}$ , die Größe mit 90 % POD;  $a_{90/95}$ , die 95 % Konfidenzgrenze für die  $a_{90}$ -Schätzung; die Gleichung für das POD-Modell. Wir haben hier das  $\log(\hat{a})$  vs  $\log(a)$ -Modell ausgewählt. Die Abbildungen 9 bis 11 zeigen repräsentative Ergebnisse unserer Auswertung. In den Abbildungen 9 und 10 ist zu sehen, dass jeweils die Variante 2 des Algorithmus bessere Ergebnisse liefert. Obwohl die Auflösung in Abbildung 11 höher ist als in Abbildung 9 und 10 sind hier schlechtere Ergebnisse zu sehen. Dies lag an einer nicht-optimalen Parametrisierung der Verfahren. Durch Optimierung der Parameter konnte die Algorithmik für die Auflösung 0,07mm jedoch so verbessert werden, dass die Werte unter den  $a_{90/95}$ -Werten der Auflösungen 0,1mm und 0,12mm liegen. Dieses Resultat zeigt, dass die POD-Analyse ein gutes Werkzeug zur Beurteilung der Qualität von Algorithmen ist. Die POD-Kurven für die Bohrungsdetektion werden aus Platzgründen nicht aufgeführt.



**Abbildung 9:** Auflösung 0,1 mm/Pixel; links zu sehen die POD-Kurve von Variante 1 des Algorithmus, rechts Variante 2. Der wichtigste Wert  $a_{90/95}$  zeigt, dass Variante 2 mit einer POD von 0,17 bessere Ergebnisse liefert als Variante 1 (0,23).



**Abbildung 10:** Auflösung 0,12 mm/Pixel; links zu sehen die POD-Kurve von Variante 1 des Algorithmus, rechts Variante 2. Der wichtigste Wert  $a_{90/95}$  zeigt, dass Variante 2 mit einer POD von 0,2 bessere Ergebnisse liefert als Variante 1 (0,21).



**Abbildung 11:** Auflösung 0,07 mm/Pixel; links zu sehen die POD-Kurve von Variante 1 des Algorithmus, rechts Variante 2. Der wichtigste Wert  $a_{90/95}$  zeigt, dass Variante 1 mit einer POD von 0,2 bessere Ergebnisse liefert als Variante 2 (0,22).

## 6. Zusammenfassung und Ausblick

Die durchgeführten Versuche zeigen, dass die POD-Analyse ein hilfreiches Verfahren zum Entwurf und zur Validierung von Bildverarbeitungsalgorithmen ist. Neben der wichtigen Aussage, ob eine bestimmte Fehlergröße sicher detektiert werden kann, liefert die POD-Analyse vor allem quantitative Werte für die sicher detektierbaren Fehlergrößen zu einem gegebenen Algorithmus sowie Beleuchtungs- und Kamera-Setup. Allerdings müssen Fehldetektionen (Falsch-Positive) durch eine weitere Analyse erkannt und minimiert werden. Weiterhin zeigt unsere Analyse, dass die Qualität eines optischen Inspektionssystem sehr wohl von der Qualität der Algorithmik abhängt und es sich durchaus lohnt eine optimale Parametrisierung durchzuführen.

Für die Zukunft scheint es sinnvoll durch physikalisch korrekte Simulationen (z.B. Raytracing) Fehlerbilder zu erzeugen und dann mit einer größeren Datenbasis sehr viel genauere Ergebnisse zu erhalten.

## Referenzen

- [1] Department of Defense Handbook Draft 2009. Nondestructive Evaluation System Reliability Assessment, MIL-HDBK-1823
- [2] W.D. Rummel. NDE Procedure Validation and Use in NDE System Calibration for NDE Applications. In: Review of Progress in Quantitative NDE (2004), Vol. 24, Melville, New York, American Institute of Physics (AIP Conference Proceedings CP760), 1982-1986 (2005)
- [3] M. Rauhut. Typischer Aufbau eines Online-Oberflächeninspektionssystems. Materialien zum Fraunhofer Vision Seminar ‚Inspektion und Charakterisierung von Oberflächen mit Bildverarbeitung‘, Erlangen (2009)
- [4] J. Beyerer, Th. Längle. Bildgewinnung bei der Oberflächenprüfung. Materialien zum Praktikum der Allianz Vision zur Oberflächeninspektion. Erlangen (2009)
- [5] R. Gonzalez, R. Woods. Digital Image Processing. 2nd Edition, Prentice Hall (2002)
- [6] R. Rösch. Fehlerdetektion in texturierten Oberflächen im praktischen Einsatz. Vortrag auf dem Fraunhofer Vision Technologietag, Kaiserslautern (2009)
- [7] M. Rauhut, M. Spies, K. Taebner. Detektion und Auffindwahrscheinlichkeit (POD) von Oberflächenfehlern in Metallen mittels optischer Inspektionsverfahren, DGZfP-Jahrestagung 2010, Erfurt, Mai 2010