

Komplementäre Bildfusion zur Inspektion technischer Oberflächen

Complementary Image Fusion for Inspection of Technical Surfaces

Fernando Puente León, DS2 Valencia*

Manuskripteingang: 8. Januar 2002; zur Veröffentlichung angenommen: 15. Februar 2002.

In diesem Beitrag wird die Fusion komplementärer Bildsensordaten am Beispiel metallischer Oberflächen demonstriert. Die Fusionsergebnisse selbst können Bilder, Merkmale oder Symbole sein. Das Fusionsproblem wird in Form eines Energiefunctionals formuliert. Dessen Summanden beschreiben gewünschte Eigenschaften, sinnvolle Forderungen und Vorwissen bezüglich der Bildserie, des Fusionsergebnisses und etwaiger problemrelevanter Zwischengrößen sowie ihre wechselseitigen Beziehungen. Durch Minimierung des Energiefunctionals erhält man das unter den getroffenen Annahmen optimale Fusionsresultat. In manchen Fällen erlaubt die Struktur der Energiefunktion, die oftmals aufwändige Optimierungsrechnung durch eine effizientere getrennte Minimierung der Einzelterme zu approximieren. Die vorgestellte Methodik wird mittels praxisrelevanter Beispiele erläutert.

This contribution addresses the fusion of complementary image sensor data using metallic surfaces as an example. The fusion results themselves can be images, features or symbols. The fusion problem is formulated in shape of an energy function. Its addends describe desirable characteristics, postulates, and prior knowledge regarding the image series, the fusion result, and eventual relevant intermediate magnitudes as well as their mutual relations. By minimizing the energy function, the optimal fusion result with respect to the assumptions met is obtained. In some cases, the structure of the energy function allows to perform the costly optimization by means of a more efficient separate minimization of the addends. The presented fusion strategy is demonstrated based on several examples from practice.

Schlagwörter: Automatische Sichtprüfung, Sensorfusion, Beleuchtung, Bildverarbeitung, Segmentierung, metallische Oberflächen, Bayes'sche Fusionstheorie

1 Einleitung

In der automatischen Sichtprüfung ist es oftmals nicht möglich, eine Oberfläche anhand einer einzigen Kameraaufnahme zuverlässig zu inspizieren. Beschränkungen optischer Systeme, Beleuchtungsprobleme sowie Schwierigkeiten, sämtliche interessierenden Objektbereiche gleichzeitig und mit der geforderten Auflösung zu erfassen, erfordern oftmals die Akquisition mehrerer Bilder bei verschiedenen Aufnahmesituationen. Meistens wird die Nutzinformation anschließend durch getrennte Weiterverarbeitung dieser Einzelbilder gewonnen.

Ein in vielen Fällen erfolgreicherer Ansatz basiert hingegen darauf, zur Extraktion der gewünschten Objekteigenschaften die gesamte Bildserie gleichzeitig auszuwerten. Hierbei können die Messdaten aus den verschiedenen Bildsensoren mit Methoden der Sensorfusion zu einem oder mehreren Ergebnissen verschmolzen werden, welche die Nutzinformation möglichst kompakt repräsentieren sollen [4]. Dabei ist es für die Formulierung der Fusionsaufgabe zunächst unbedeutend, ob die Ergebnisse selbst weiterhin Bilder, Merkmale oder Symbole darstellen.

In diesem Beitrag wird die Fusion von Bildsensoren anhand „nichtkooperativer“ metallischer Oberflächen demonstriert. Die hier verwendeten Sensoren sind in dem Sinne „virtuell“, dass mit einer einzigen Kamera eine Serie von Bildern

* Die Arbeit entstand am Institut für Mess- und Regelungstechnik (Prof. Dr.-Ing. C. Stiller) der Universität Karlsruhe (TH).

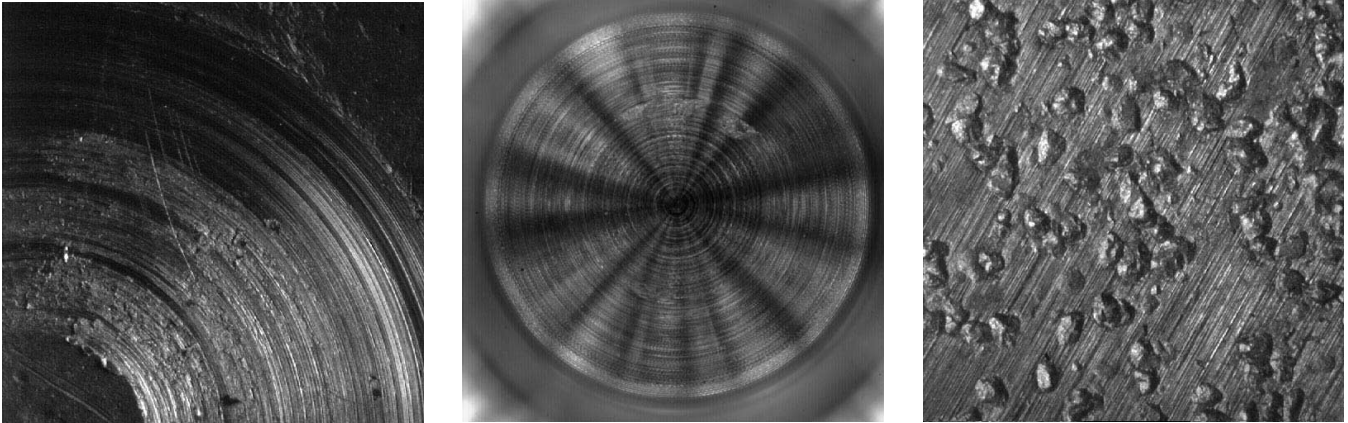


Bild 1: Makroskopische Aufnahmen metallischer Oberflächen: (links) Werkzeugspur; (Mitte) Membranoberfläche eines Drucksensors; (rechts) kugelgestrahlte Oberfläche.

mit unterschiedlichen Beleuchtungskonstellationen aufgenommen wird.

Anhand praxisrelevanter Beispiele – der Erfassung kriminaltechnisch relevanter Spuren, der Inspektion der Membranoberfläche von Drucksensoren sowie der Auswertung kugelgestrahlter Oberflächen (siehe Bild 1) – wird die vorgestellte Methodik erläutert. Durch die systematische Herstellung komplementärer Aufnahmesituationen und den Einsatz von Fusionsverfahren lässt sich oftmals ein erheblicher Qualitätsgewinn in der automatischen Sichtprüfung realisieren.

2 Sensorfusion

Die Aufgabe, Messdaten aus unterschiedlichen Sensoren zu kombinieren, wird in der Literatur als *Sensorfusion* bezeichnet [4;9]. Hier soll unter Sensorfusion die Kombination entsprechender Bildinhalte aus einer Serie $\mathcal{D} := \{d(\mathbf{x}, \omega_i), i = 0, \dots, B-1\}$ von B Bildern zu C Ergebnissen verstanden werden, wobei

$$\omega = (\phi, \theta, \dots)^T \quad (1)$$

den Parametervektor der Aufnahmesituation bezeichnet. Dabei sind ϕ und θ jeweils Azimut und Elevationswinkel der Beleuchtungseinrichtung.¹ Die Fusionsergebnisse können Bilder, Merkmale oder symbolische Bilddeskriptoren sein.

Abhängig vom Abstraktionsniveau, auf dem die Fusion erfolgt, werden drei Fusionsarten unterschieden:

- Bei der *Datenfusion* werden Bildsignale kombiniert, um dadurch neue Bilder zu erhalten. Wenn bildgebende Sensoren als Datenquellen dienen, wird hierfür manchmal auch der Begriff *Bildfusion* verwendet. Im Rahmen dieses Beitrags wird hierunter jedoch allgemein die Verschmelzung von Bildsensoren verstanden – unabhängig

von der Abstraktionsebene, auf der die Fusion selbst erfolgt.

- *Merkmalsfusion* kann durchgeführt werden, um verbesserte numerische Schätzwerte bestimmter Bildeigenschaften – wie z. B. die Ausprägung von Oberflächendefekten – zu erhalten.
- Auf dem höchsten Abstraktionsniveau werden symbolische Bilddeskriptoren kombiniert – wie z. B. Detektions- oder Klassifikationsergebnisse bezüglich Einzelbildern der Serie. Hierfür haben sich die Begriffe *Entscheidungsfusion* bzw. *Symbolfusion* etabliert.

Sowohl bei der Datenfusion als auch bei der Merkmalsfusion kann es günstig sein, Ergebnisse aus höheren Abstraktionsebenen zu berücksichtigen, wie Bild 2 zeigt. Beispielsweise könnte es vorteilhaft sein, Segmentierungsergebnisse (d. h. eine symbolische Bildbeschreibung) von Einzelbildern der Serie zu verwenden, um die Fusion auf Datenebene zu unterstützen, da der Segmentierungsprozess Bereiche mit unterschiedlichen Dateneigenschaften liefert [10].

Eine weitere Unterscheidung bei der Fusion betrifft die Wahl der verwendeten Sensoren:

- *Gleichartige* bzw. *homogene Sensoren* werden oftmals eingesetzt, um eine Reduktion der Messzeit, eine Erhöhung der Zuverlässigkeit eines Systems oder eine Erhöhung des Signal-zu-Rausch-Verhältnisses zu erzielen. Wegen der Gleichartigkeit der Sensordaten kann

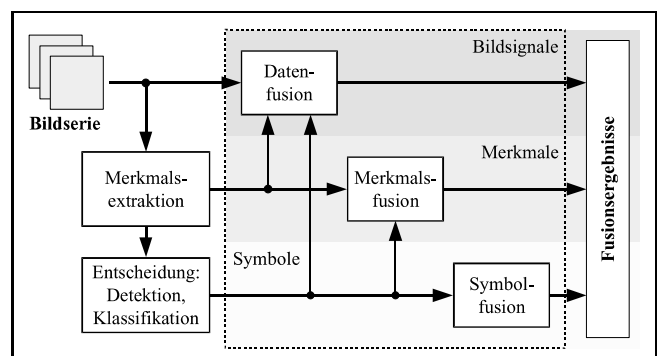


Bild 2: Sensorfusion von Bildsignalen.

¹ Bei Bedarf können zusätzliche Parameter berücksichtigt werden, wie z. B. die Integrationszeit der Videokamera, die Wellenlänge des verwendeten Lichtes oder die Objektpose relativ zur Aufnahmeoptik.

die Fusion mit einem Minimum an Vorverarbeitung erfolgen.

- Anordnungen aus *Sensoren unterschiedlicher Art* können verwendet werden, um Beschränkungen und Schwächen der Einzelsensoren zu kompensieren. Allgemein ist hier vor der Fusion eine Vorverarbeitung erforderlich, damit die Daten im gleichen Format vorliegen.

Die in diesem Aufsatz vorgestellten Bildsignale stammen aus Sensoren, die in dem Sinne virtuell sind, dass mit nur einem realen Sensor – einer Videokamera – eine Serie von Bildern in B unterschiedlichen Aufnahmesituationen $\{\omega_i, i = 0, \dots, B-1\}$ gewonnen wird. Wegen der Gleichartigkeit der „Einzelsensoren“ handelt es sich hierbei um eine *homogene Sensorfusion*.

Schließlich kann die Verschmelzung der Sensordaten *konkurrierend* oder *komplementär* erfolgen, je nachdem, ob für einen Ort \mathbf{x} die Nutzinformation über mehrere Bilder verteilt vorliegt, oder ob sie in einem oder in wenigen Bildern konzentriert enthalten ist. Ein Beispiel für eine konkurrierende Fusion ist die Akkumulation von Bildern, die unter gleichen Aufnahmebedingungen erfasst wurden, zum Zwecke der Rauschunterdrückung. Durch die systematische Variation der Aufnahmeparameter liegt in diesem Beitrag die Nutzinformation zwar verteilt auf die Bilder der Serie vor, für einen bestimmten Ort jedoch auf wenige Bilder konzentriert, weshalb hier eine *komplementäre Bildfusion* angestrebt wird.

2.1 Allgemeine Fusionsstrategie

Ein universeller Ansatz zur Formalisierung von Fusionsproblemen besteht darin, sowohl das gesamte verfügbare *A-priori*-Wissen, gewünschte Eigenschaften und sinnvolle Forderungen bezüglich der zu fusionierenden Daten \mathcal{D} , der Zwischengrößen \mathcal{Z} und der Fusionsergebnisse \mathcal{R} als auch ihre wechselseitigen Beziehungen in Form von verallgemeinerten Energietermen $E_k(\mathcal{D}, \mathcal{R}, \mathcal{Z})$ zu beschreiben [3]. Diese „Energien“ E_k müssen derart gewählt werden, dass sie das verfügbare Wissen und die gegebenen Forderungen monoton widerspiegeln in dem Sinne, dass das Resultat um so wünschenswerter ausfällt, je geringer die Energie ist. Die einzelnen Terme E_k werden sodann durch gewichtete Summation zu einem so genannten Energiefunktional E zusammengefasst:

$$E = \sum_k \lambda_k E_k(\mathcal{D}, \mathcal{R}, \mathcal{Z}), \quad \lambda_k > 0. \quad (2)$$

E stellt eine implizite, kompakte Darstellung der Fusionsaufgabe dar. Aufgrund der Monotonieeigenschaften von E kann die Fusion insgesamt durch eine simultane Minimierung des Energiefunctionals E bezüglich der Fusionsresultate \mathcal{R} und der Zwischengrößen \mathcal{Z} durchgeführt werden:

$$\{\mathcal{R}^*, \mathcal{Z}^*\} = \arg \min_{\mathcal{R}, \mathcal{Z}} \{E\}, \quad (3)$$

wobei \mathcal{R}^* und \mathcal{Z}^* die infolge der Optimierung resultierenden Größen bezeichnen.

Vorteilhaft an diesem Ansatz sind dessen große Allgemeinheit und die Möglichkeit, Zusatzinformation anhand weiterer Energieterme zu berücksichtigen. Nachteilig ist vor allem die Schwierigkeit, das Energiefunktional zu optimieren. Anhand der in Bild 1 dargestellten Beispiele soll nun die Methodik gemäß Gl. (2) für die unterschiedlichen Abstraktionsebenen anschaulich erläutert werden.

2.2 Fusion auf Datenebene

Das erste Beispiel beschäftigt sich mit dem Einsatz von Datenfusion zur Erzeugung qualitativ hochwertiger Bilder feiner Oberflächenspuren, wie sie etwa in der Kriminaltechnik zur Erkennung von Schusswaffen oder Werkzeugen benötigt werden [6; 10]. Dafür wird die Spur mit einer gerichteten Lichtquelle beleuchtet und dabei eine Bildserie \mathcal{D} des Spurenträgers durch Variation des Beleuchtungswinkels ϕ erfasst, sodass die interessierenden Oberflächenbereiche vollständig und mit hohem Kontrast durch die gesamte Serie abgedeckt werden:

$$\mathcal{D} := \{d(\mathbf{x}, \phi_i), i = 0, \dots, B-1\}, \quad (4)$$

$$\phi_i = \phi_0 + i\Delta\phi, \quad 0 \leq \phi_i < 2\pi. \quad (5)$$

Daraus sollen die besten Bildbereiche „herausgeschnitten“ und zu einem hochwertigen Resultat $\mathcal{R} = \{r(\mathbf{x})\}$ zusammengefügt werden, welches die Vorteile der Einzelbilder vereint.

Ein allgemeines Energiefunktional zur Durchführung der Datenfusion im Sinne von Gl. (2) lautet:

$$E = E_D(\mathcal{D}, r) + \lambda E_K(r), \quad \lambda > 0. \quad (6)$$

Dabei verkörpert $E_D(\mathcal{D}, r)$ den Zusammenhang zwischen den Bilddaten \mathcal{D} und dem Fusionsresultat $r(\mathbf{x})$. $E_K(r)$ modelliert wünschenswerte oder *a priori* bekannte Eigenschaften von $r(\mathbf{x})$, und λ fungiert als ein Regularisierungsparameter, der beide Energiekomponenten gewichtet. Das Resultat $r(\mathbf{x})$ erfüllt die *A-priori*-Erwartungen desto besser, je geringer $E_D(\mathcal{D}, r)$ und $E_K(r)$ sind.

Für die vorliegende Aufgabe muss Gl. (6) jedoch entsprechend erweitert und konkretisiert werden:

$$E = \sum_i \sum_{\mathbf{x}} (r(\mathbf{x}) - d(\mathbf{x}, \phi_i))^2 \gamma(\phi_i - \varphi(\mathbf{x})) \quad (7)$$

$$+ \lambda_1 \sum_{\mathbf{x}} (h(\mathbf{x}) * \varphi(\mathbf{x}))^2 - \lambda_2 \sum_{\mathbf{x}} K\{r(\mathbf{x})\}$$

$$= E_D(\mathcal{D}, r, \varphi) + \lambda_1 E_G(\varphi) + \lambda_2 E_K(r) \quad (8)$$

mit $0 \leq \varphi(\mathbf{x}) < 2\pi$, $\lambda_1, \lambda_2 > 0$, $h(\mathbf{x})$: Hochpass.

Bei Gl. (7) handelt es sich um eine kompakte Beschreibung des Fusionsproblems, welche die bekannten und gewünschten Eigenschaften der an der Fusion beteiligten Größen sowie deren Zusammenhänge widerspiegelt. Das optimale Fusionsresultat $r^*(\mathbf{x})$ erhält man durch simultane

Minimierung von E bezüglich $r(\mathbf{x})$ und der benötigten Hilfsgröße $\varphi(\mathbf{x}) \in \mathcal{Z}$:

$$E(r^*, \varphi^*, \mathcal{D}) := \min_{r, \varphi} \{E\}. \quad (9)$$

Formal taucht in Gl. (7) im Vergleich zu Gl. (6) das bei der Fusion benötigte Beleuchtungsazimut $\varphi(\mathbf{x})$ sowie der zugehörige Energieterm $E_G(\varphi)$ auf. An dieser Stelle wird postuliert, dass sich das optimale Beleuchtungsazimut $\varphi^*(\mathbf{x})$ örtlich langsamer ändert als die sichtbar zu machende Oberflächentextur $t(\mathbf{x})$. Typischerweise handelt es sich bei der Textur $t(\mathbf{x})$ um ein Bandpass-Signal [2; 10]. Um schädliche Wechselwirkungen zwischen $\varphi(\mathbf{x})$ und der Textur $t(\mathbf{x})$ infolge der örtlichen Variation der Beleuchtungsrichtung zu vermeiden, wird $\varphi(\mathbf{x})$ als Tiefpass-Signal angesetzt. Im Idealfall sollten die Träger der Fourierpektren $T(\mathbf{f})$ bzw. $\Phi(\mathbf{f})$ der Signale $t(\mathbf{x})$ bzw. $\varphi(\mathbf{x})$ disjunkt sein. Der zweite Summand $E_G(\varphi)$ in Gl. (7) „bestraft“ daher unzulässig hochfrequente Signalanteile von $\varphi(\mathbf{x})$, indem er die Energie des hochpassgefilterten Signals $h(\mathbf{x}) * \varphi(\mathbf{x})$ wiedergibt. Dieser Term stellt damit eine örtliche Glattheitsforderung für den optimalen Beleuchtungswinkel $\varphi^*(\mathbf{x})$ dar.

Der erste Summand $E_D(\mathcal{D}, r, \varphi)$ sorgt für Datennähe vom Ergebnisbild $r(\mathbf{x})$. Lokal soll $r(\mathbf{x})$ möglichst demjenigen Bild $d(\mathbf{x}, \phi_i)$ der Serie gleichen, das am Ort \mathbf{x} optimale Beleuchtungseigenschaften zeigt. Die Funktion $\gamma(\phi_i - \varphi(\mathbf{x}))$ lässt im ersten Energieterm nur das bestgeeignete Bild $d(\mathbf{x}, \phi_i)$ und dessen Nachbarn im Beleuchtungsraum eingehen.

Der dritte Summand $E_K(r)$ bewertet, ob das lokale, mit der Bildqualität monoton wachsende Gütekriterium $K\{\cdot\}$ – z. B. ein lokales Kontrastmaß – im Fusionsresultat $r(\mathbf{x})$ global möglichst hohe Werte erzielt.

Implementationsstrategie: Anstelle der sehr aufwändigen Minimierung des Energiefunktionals E gemäß Gl. (9) wird

eine suboptimale Vorgehensweise eingeschlagen, die auf der Selektion der bestbeleuchteten Bildausschnitte der Serie für jeden Ort \mathbf{x} mit Hilfe des lokalen Kontrastes K basiert; siehe Bild 3. Dafür wird im ersten Schritt die so genannte *Beleuchtungs- oder Fusionskarte*

$$\tilde{\phi}(\mathbf{x}) = \arg \max_{\phi_i} \{K\{d(\mathbf{x}, \phi_i)\}\} \quad (10)$$

berechnet, die für jeden Ort \mathbf{x} Auskunft über denjenigen Parameterwert ϕ_i gibt, welcher den lokalen Kontrast K maximiert.² Dies führt im Wesentlichen zu einer Minimierung des dritten Summanden $\lambda_2 E_K(r)$. Danach wird die Beleuchtungskarte $\tilde{\phi}(\mathbf{x})$ mit einem Tiefpassfilter geglättet:

$$\hat{\varphi}^*(\mathbf{x}) = \mathcal{LTP}\{e^{j\tilde{\phi}(\mathbf{x})}\}. \quad (11)$$

Dadurch wird im Wesentlichen der zweite Summand minimiert. Bei der Filterung von $\tilde{\phi}(\mathbf{x})$ muss die Zyklizität von ϕ beachtet werden, da $\phi = \phi + 2\pi k$, $k \in \mathbb{Z}$ gilt.

Die eigentliche Fusion erfolgt durch gewichtete Überlagerung maximal zweier, bezüglich ϕ benachbarter Bilder $d(\mathbf{x}, \phi_i)$ mit Hilfe eines linearen Interpolators γ . Dies führt zu einer Minimierung vor allem des ersten Summanden. Das Fusionsresultat ergibt sich zu:

$$\hat{r}^*(\mathbf{x}) = \frac{(\hat{\varphi}^*(\mathbf{x}) - \phi_k) \bmod 2\pi}{(\phi_l - \phi_k) \bmod 2\pi} d(\mathbf{x}, \phi_k) + \frac{(\phi_l - \hat{\varphi}^*(\mathbf{x})) \bmod 2\pi}{(\phi_l - \phi_k) \bmod 2\pi} d(\mathbf{x}, \phi_l), \quad (12)$$

$$l := (k + 1) \bmod B,$$

$$(\phi_l - \phi_k) \bmod 2\pi \leq (\hat{\varphi}^*(\mathbf{x}) - \phi_k) \bmod 2\pi.$$

Die Interpolation sorgt für glattes Überblenden zwischen benachbarten Bildern. Bei hinreichend dichter Abtastung

² Man beachte, dass in Bild 3 die Winkelwerte von $\tilde{\phi}(\mathbf{x})$ durch Grauwerte codiert sind.

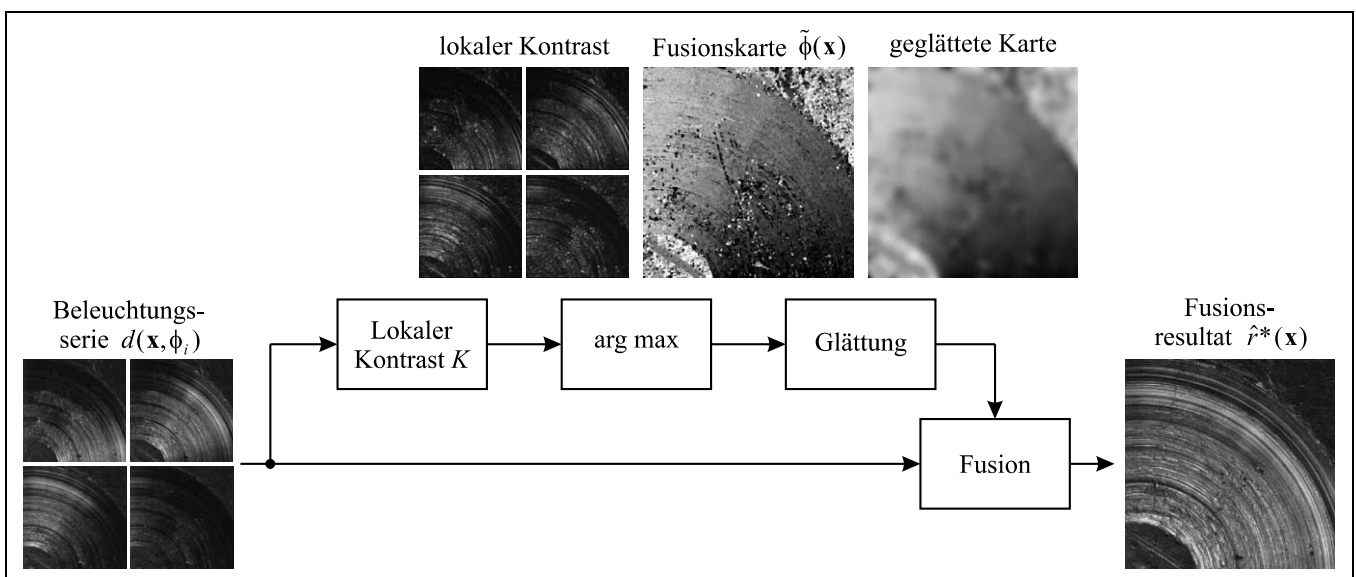


Bild 3: Fusion auf Datenebene: Struktur der Strategie zur effizienten Minimierung von E .

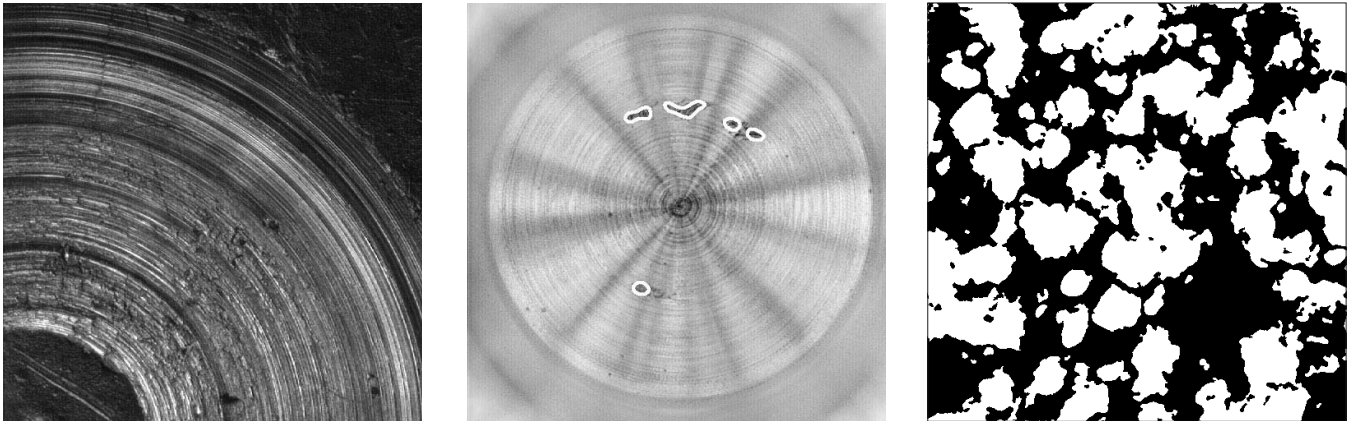


Bild 4: Fusionsergebnisse: (links) Werkzeugspur; (Mitte) Merkmalsbild der Membranoberfläche eines Drucksensors mit erkannten Defekten; (rechts) detektierte Kugeleinschläge.

des Beleuchtungsraumes kann die Ausdehnung von γ so schmal gewählt werden, dass für einen Ort \mathbf{x} keine Grauwerte von sehr unterschiedlich beleuchteten Bildern in das Ergebnisbild eingehen. Bild 4 (links) zeigt das für die Werkzeugspur aus Bild 1 (links) resultierende Fusionsergebnis, welches sich als Grundlage für einen rechnergestützten Vergleich zum Zwecke der Spurenidentifikation hervorragend eignet [7].

Schließlich sei noch angemerkt, dass anstelle der Änderung des Azimuts ϕ auch eine Variation des Elevationswinkels θ möglich ist. In diesem Fall vereinfachen sich die Gleichungen (11) und (12) erheblich, weil θ keine zyklische Größe ist [10].

Für den allgemeineren Fall, dass sowohl das Azimut ϕ als auch der Elevationswinkel θ variiert werden, muss die Bildfusion zweidimensional erfolgen. Bei geeigneter Abtastung des Beleuchtungsraumes kann diese jedoch auch durch eine zweistufige Implementation eindimensionaler Fusionsverfahren gemäß Bild 3 realisiert werden, womit ebenfalls eine Vereinfachung erzielt wird [10].

2.3 Fusion auf Merkmalsebene

Im zweiten Beispiel wird die Nutzung komplementärer Aufnahmen zur Quantifizierung und zuverlässigen Detektion von Oberflächenfehlern am Beispiel der Membranoberfläche von Drucksensoren demonstriert; siehe Bild 1 (Mitte). Dafür wird durch systematische Variation der Beleuchtung eine Bildserie \mathcal{D} gemäß Gl. (4) aufgenommen. Eine punktweise Analyse der Grauwertintensitäten im Beleuchtungsraum erlaubt in diesem Fall die Ermittlung eines Merkmalsbildes als Fusionsresultat, welches das Vorhandensein selbst kleinster Oberflächendefekte sicher indiziert.

Um das Fusionsprinzip zu erläutern, zeigt Bild 5 acht der 16 Bilder der Beleuchtungsserie. Wird nun für einen festen Oberflächenpunkt \mathbf{x} der Verlauf der Grauwerte $d(\mathbf{x}, \phi_i)$ in Abhängigkeit des Azimuts ϕ untersucht, so lässt sich ein klares Intensitätsmuster für defektfreie Bereiche erkennen: Bei senkrechter Beleuchtung erscheint die intakte

Drehtextur hell; anderenfalls erhält man dunkle Grauwerte. Abweichungen von diesem Verhalten weisen auf Oberflächenfehler hin.

Durch harmonische Analyse der Signale $d(\mathbf{x}, \phi_i)$ lässt sich ein geeignetes Merkmal definieren, welches eine robuste Quantifizierung dieser Defekte erlaubt:

$$m(\mathbf{x}) = \frac{|D(\mathbf{x}, f_\phi = 1)|}{|D(\mathbf{x}, f_\phi = 1)| + |D(\mathbf{x}, f_\phi = 0)|}, \quad (13)$$

wobei

$$D(\mathbf{x}, f_\phi) := \mathcal{F}_\phi\{d(\mathbf{x}, \phi)\} \\ = \sum_{b=0}^{B-1} d(\mathbf{x}, \phi_b) \cdot \exp\left(-j 2\pi \frac{b f_\phi}{B}\right) \quad (14)$$

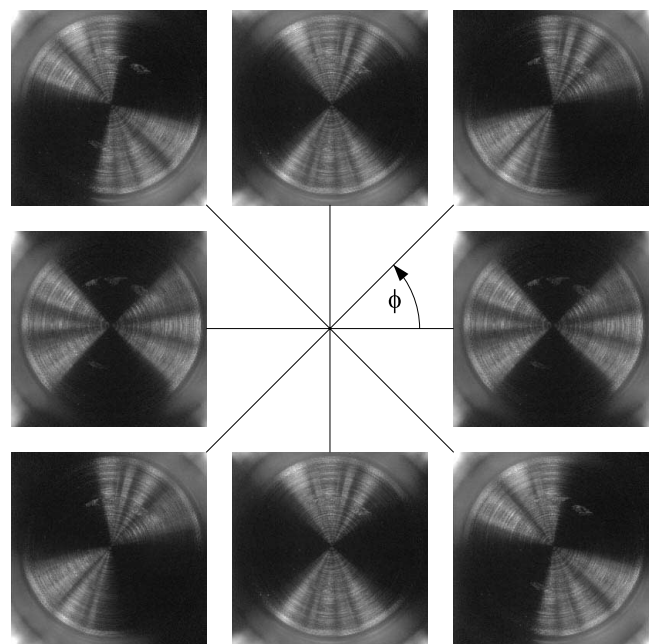


Bild 5: Fusion auf Merkmalsebene: Bildserie der Membranoberfläche eines Drucksensors.

die eindimensionale diskrete Fouriertransformation (DFT) der Serie bzgl. des Beleuchtungswinkels ϕ bezeichnet. Gleichung (13) berechnet ein Merkmal, basierend auf der Grundschwingung und dem Gleichanteil der Bildintensitäten am Ort \mathbf{x} , und bildet diese auf die Werte $[0; 1]$ ab. Ein Wert größer als 0,5 bedeutet, dass die Grundwelle überwiegt – d.h. die Textur ist an der Stelle \mathbf{x} defektfrei. Anderenfalls wird \mathbf{x} den fehlerhaften Bereichen zugeordnet. Das resultierende Energiefunktional ist trivial und besteht aus lediglich einem Summanden:

$$E = \sum_{\mathbf{x}} [r(\mathbf{x}) - m(\mathbf{x})]^2 = E_{\text{Merkmal}}(m(\mathbf{x}), r(\mathbf{x})), \quad (15)$$

wodurch in diesem Fall die aufwändige Optimierung entfällt: $r^*(\mathbf{x}) = m(\mathbf{x})$. Ein Beispiel für die Fusion auf Merkmalsebene ist in Bild 4 (Mitte) zusammen mit den Ergebnissen der anschließenden Defektdetektion zu sehen.

2.4 Fusion auf Symbolebene

Im letzten Beispiel geht es um die automatische Bestimmung des Überdeckungsgrades kugelgestrahlter Oberflächen, wie sie wegen ihrer hohen Ermüdungsfestigkeit u.a. im Flugzeugbau und in der Raumfahrt eingesetzt werden; siehe Bild 1 (rechts) [11]. Um eine wohldefinierte Überdeckung der Oberfläche mit Kugeleinschlägen zu erhalten, muss der Kugelstrahlprozess durch Ermittlung des Überdeckungsgrades als Funktion der Zeit kalibriert werden. Dies erfolgt bisher durch visuellen Vergleich von Testoberflächen mit einem Katalog von Referenzmustern. Aufgrund der komplexen Struktur solcher Oberflächen gelingt oftmals selbst erfahrenen Experten keine eindeutige Unterscheidung zwischen Einschlägen und Hintergrundtextur, was diese Vorgehensweise äußerst subjektiv werden lässt. Im Gegensatz dazu wird im vorgestellten Ansatz eine Bildserie \mathcal{D} gemäß Gl. (4) bei systematischer Variation des Beleuchtungswinkels ϕ aufgenommen und zu einem symbolischen Ergebnis $r(\mathbf{x})$ verschmolzen, wodurch eine wesentlich robustere Detektion der interessierenden Bereiche erreicht wird.

In diesem Fall kann das Fusionsproblem vollständig anhand eines Energiefunktionals E mit drei Termen formuliert werden:

$$E = E_{\text{Merkmal}}(m(\mathbf{x}), r(\mathbf{x})) + \lambda_1 \cdot E_{\text{Größe}}(r(\mathbf{x})) + \lambda_2 \cdot E_{\text{Form}}(r(\mathbf{x})). \quad (16)$$

Der erste Summand, $E_{\text{Merkmal}}(m(\mathbf{x}), r(\mathbf{x}))$, basiert auf einem Merkmal $m(\mathbf{x})$, das eine Entscheidung ermöglicht, ob ein Bildpunkt eher zur Hintergrundtextur oder zu einem Kugeleinschlag gehört. Entsprechend sorgt dieser Term für Merkmalsnähe und führt somit zu einer beleuchtungs-gestützten Klassifikation des Hintergrundes und der Einschläge.

Der zweite Energieterm, $E_{\text{Größe}}(r(\mathbf{x}))$, misst die Größe sämtlicher Regionen, die im Fusionsresultat $r(\mathbf{x})$ als Kugeleinschläge gekennzeichnet sind und beurteilt, ob die

unterschiedlichen Größen wahrscheinlich sind. Dabei werden die Durchmesser typischer Einschläge berücksichtigt. Ist das nicht der Fall, so liefert dieser Term hohe Werte.

Analog dazu beurteilt der dritte Term, $E_{\text{Form}}(r(\mathbf{x}))$, die Kompaktheit der Regionen im Ergebnis $r(\mathbf{x})$ und führt zu hohen Werten, falls die Form eines Einschlags von der Form einer Kreisscheibe abweicht.

Die folgende Gleichung zeigt das Energiefunktional E im Detail:

$$E = \sum_{\mathbf{x}} [r(\mathbf{x}) - m(\mathbf{x})]^2 + \frac{\lambda_1}{R} \sum_i \Psi \left\{ \sum_{\mathbf{x}} \Lambda_i \{r(\mathbf{x})\} \right\} + \frac{\lambda_2}{R} \sum_i \Gamma [\Lambda_i \{r(\mathbf{x})\}], \quad (17)$$

wobei R die Bereichszahl im binären Ergebnisbild $r(\mathbf{x})$ bezeichnet. Der Operator $\Lambda_i \{ \}$ führt eine Zusammenhangsanalyse eines binären Bildes durch und wählt die i -te Region aus, $\Psi \{ \}$ beurteilt die Bereichsgröße, indem kleine Werte bestraft werden, und $\Gamma [\]$ berechnet die Kompaktheit³ eines Bereiches.

Bei der Hintergrundtextur aus dem Beispiel Bild 1 (rechts) handelt es sich um eine Schleiftextur. In diesem Fall kann die Größe $m(\mathbf{x})$ wie folgt berechnet werden:

$$m_S(\mathbf{x}) = \frac{|D(\mathbf{x}, f_\phi = 1)|}{|D(\mathbf{x}, f_\phi = 1)| + |D(\mathbf{x}, f_\phi = 2)|}. \quad (18)$$

Gleichung (18) definiert ein Merkmal, basierend auf der Grundschwingung und der ersten Harmonischen der Bildintensitäten am Ort \mathbf{x} , und bildet diese auf die Werte $[0; 1]$ ab. Ein Wert größer als 0,5 bedeutet, dass die Grundwelle überwiegt – d.h. in diesem Fall gehört der Ort \mathbf{x} zu einem Einschlag. Anderenfalls wird \mathbf{x} der Schleiftextur zugeordnet. Für andere Hintergrundtexturen lassen sich entsprechende Unterscheidungsmerkmale analog hierzu definieren [11].

Implementationsstrategie: Auch bei der Fusion auf Symbolebene kann eine effizientere suboptimale Strategie angewendet werden, falls die Anforderungen an das Fusionsresultat sequentiell erfüllt werden. Bild 6 zeigt das Verarbeitungsschema für den Fall einer geschliffenen Hintergrundtextur. Dabei bezeichnet $\text{BP}\{ \}$ einen Bandpass-Operator. Die Riefenunterdrückung eliminiert helle Linien im Merkmalsbild $m_S(\mathbf{x})$, die aus breiten Riefen resultieren und welche zu ähnlichen Intensitäten wie die Kugeleinschläge führen, mit Hilfe eines grauwertmorphologischen Openings [5]. Bild 4 (rechts) zeigt das mit dieser Fusionsstrategie erzielte Ergebnisbild. Schließlich sei noch erwähnt, dass die mittels dieser Vorgehensweise ermittelte Überdeckung einen Fehler von weniger als 2%, bezogen auf die Gesamtfläche, aufweist [12].

³Zum Beispiel kann die Kompaktheit $\Gamma [\]$ eines Bereiches durch Vermessung seiner Konturlänge und Division durch seine Fläche bestimmt werden.

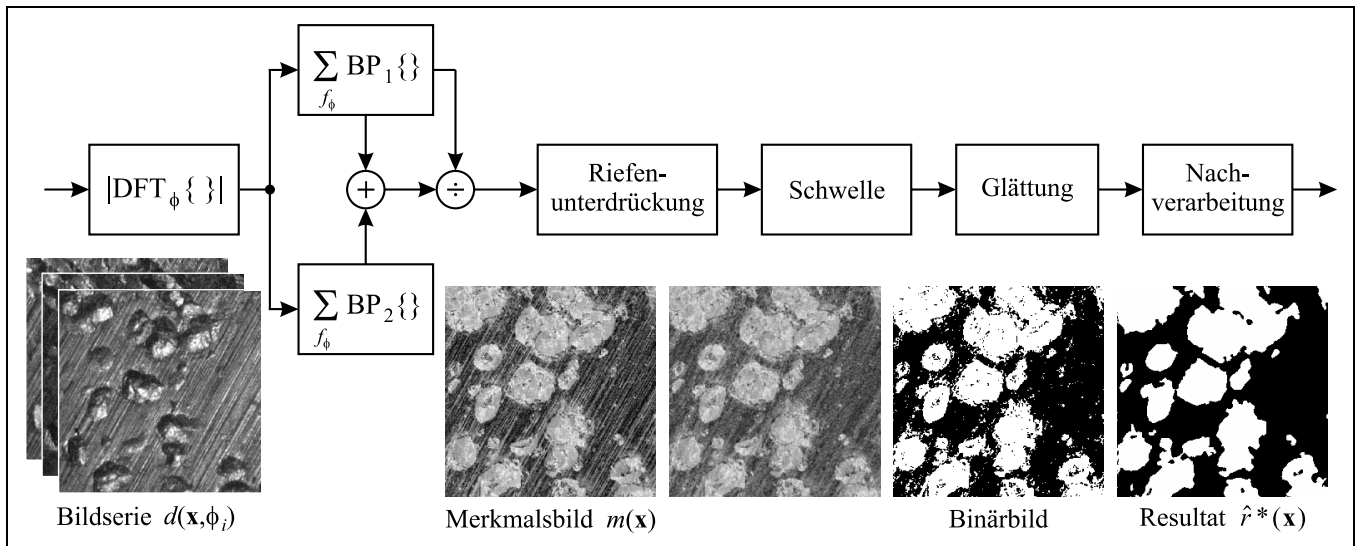


Bild 6: Fusion auf Symbolebene: Fusionsschema für kugelgestrahlte geschliffene Oberflächen.

2.5 Zusammenhang mit der Bayes'schen Statistik

Durch Einführung von Gibbs'schen Dichten lässt sich ein interessanter Zusammenhang zwischen der vorgestellten Fusionsmethodik und der Bayes'schen Statistik herstellen [3]. Für das Energiefunktional Gl. (2) kann eine Gibbs'sche Wahrscheinlichkeitsdichtefunktion (WDF) wie folgt definiert werden:

$$\text{WDF} \propto \exp\left(-\frac{E}{T}\right) = \prod_k \exp\left(-\frac{\lambda_k E_k}{T}\right). \quad (19)$$

Diese Gleichung entspricht der von der Physik her bekannten Boltzmann-Verteilung, wobei T als eine verallgemeinerte Temperatur aufgefasst werden kann. Da das Energiefunktional formal eine Summe von Termen ist, kann die WDF in Faktoren zerlegt werden. Wählt man darüber hinaus quadratische Energiefunktionen, so ergeben sich Gauß'sche Dichten.

Diese Faktoren können entweder als eine Likelihoodfunktion oder als A-priori-WDFs interpretiert werden. Setzt man beispielsweise das für die Datenfusion aufgestellte Energiefunktional Gl. (8) in Gl. (19) ein, so erhält man den folgenden Ausdruck:

$$p(D, r, \varphi) \propto e^{-\frac{E_D}{T}} \cdot \underbrace{e^{-\frac{\lambda_1 E_G}{T}} \cdot e^{-\frac{\lambda_2 E_K}{T}}}_{p(r, \varphi)}$$

$$p(D, r, \varphi) = \underbrace{p(D|r, \varphi)}_{\text{Likelihood}} \cdot \underbrace{p(r, \varphi)}_{\text{A-priori-WDF}} \quad (20)$$

$$\propto \underbrace{p(r, \varphi|D)}_{\text{A-posteriori-WDF}}$$

Nachdem das Energiefunktional Gl. (8) aus drei Energietermen besteht, weist die WDF drei Faktoren auf. Der erste hängt sowohl von den beobachteten Daten \mathcal{D} als auch von dem interessierenden Signal $r(\mathbf{x})$ und der Beleuchtungsfunktion $\varphi(\mathbf{x})$ ab. Von einem statistischen Standpunkt aus

kann er daher als die Likelihoodfunktion $p(\mathcal{D}|r, \varphi)$ aufgefasst werden, d.h. die bedingte WDF der Daten \mathcal{D} bei gegebenen $r(\mathbf{x})$ und $\varphi(\mathbf{x})$. Da der zweite und der dritte Term von den Daten \mathcal{D} unabhängig sind, spielen diese konsequenterweise die Rolle einer A-priori-WDF $p(r, \varphi)$, welche die Eigenschaften von $r(\mathbf{x})$ und $\varphi(\mathbf{x})$ beschreibt. Die Tatsache, dass die WDF $p(r, \varphi)$ in $p(r) \cdot p(\varphi)$ separiert werden kann, spiegelt wider, dass bereits bei der Formulierung der Energieterme das A-priori-Wissen über $r(\mathbf{x})$ und $\varphi(\mathbf{x})$ als voneinander unabhängig angenommen wurde.

Die Monotonie der Exponentialfunktion gewährleistet dabei, dass die Minimierung von E und die Maximierung der WDF äquivalent sind. Da die multiplikative Verknüpfung von Likelihoodfunktion $p(\mathcal{D}|r, \varphi)$ und A-priori-WDF $p(r, \varphi)$ bis auf einen Normierungsfaktor identisch mit der A-posteriori-WDF ist, führt die Minimierung des Energiefunktionals auf die Maximum-a-posteriori (MAP)-Lösung der Fusionsaufgabe [1].

Die Verallgemeinerung des Fusionsproblems gemäß den Gleichungen (2) und (19) hat den wesentlichen Vorteil, dass dadurch Methoden zur Lösung inverser Probleme erschlossen werden, die aus der statistischen Physik stammen, wie z.B. *Simulated Annealing* oder *Mean-Field-Methoden* [3; 8].

3 Zusammenfassung

Es wurden mehrere Verfahren zur Fusion von Bildsensoren am Beispiel metallischer Oberflächen vorgestellt. Die Ergebnisse selbst können Bilder, Merkmale oder Symbole sein.

Der Beitrag verdeutlicht insbesondere, dass sich durch die systematische Herstellung komplementärer Aufnahmesituationen und den Einsatz von Fusionsverfahren oftmals ein erheblicher Qualitätsgewinn in der automatischen Sichtprüfung realisieren lässt. Dem größeren Auf-



wand bei der Datenakquisition stehen oftmals eine einfachere bzw. robustere Signalverarbeitung und nicht zuletzt die Möglichkeit der Lösung neuer Sichtprüfungsaufgaben gegenüber.

Literatur

- [1] J. Beyerer, *Verfahren zur quantitativen statistischen Bewertung von Zusatzwissen in der Meßtechnik*, VDI-Verlag, Düsseldorf, 1999.
- [2] J. Beyerer und F. Puente León, „Suppression of inhomogeneities in images of textured surfaces“, *Optical Engineering* **36** (1), 85–93, 1997.
- [3] J. J. Clark und A. L. Yuille, *Data Fusion for Sensory Information Processing Systems*, Kluwer Academic Publishers, Boston, 1990.
- [4] X. E. Gros: *NDT Data Fusion*, Arnold, London, 1997.
- [5] R. M. Haralick und L. G. Shapiro, *Computer and robot vision*, Addison-Wesley, Reading, 1992.
- [6] M. Heizmann und F. Puente León, „Model-based analysis of striation patterns in forensic science“, In: *Enabling Technologies for Law Enforcement and Security*, S. K. Bramble, E. M. Carapezza und L. I. Rudin (Hrsg.), Proceedings of SPIE, Vol. 4234, S. 533–544, 2001.
- [7] M. Heizmann und F. Puente León, „Automated analysis and comparison of striated toolmarks“, In: *Proceedings of the European Meeting for Shoeprint/Toolmark Examiners (SPTM 2001)*, Berlin, Germany, 15–18 May 2001.
- [8] S. Kirkpatrick, C. D. Gelatt, Jr. und M. P. Vecchi, „Optimization by simulated annealing“, *Science* **220**, 671–680, 1983.
- [9] R. C. Luo und M. G. Kay, „Multisensor Integration and Fusion in Intelligent Systems“, *IEEE Transactions on Systems, Man, and Cybernetics* **19** (5), 901–931, 1989.
- [10] F. Puente León, *Automatische Identifikation von Schußwaffen*, VDI Verlag, Düsseldorf, 1999.
- [11] F. Puente León, „Model-based inspection of shot peened surfaces using fusion techniques“, In: *Machine Vision and Three-Dimensional Imaging Systems for Inspection and Metrology*, Proceedings of SPIE, Vol. 4189, S. 41–52, 2001.
- [12] F. Puente León, „Determination of the coverage of shot peened surfaces“, In: *Surface Treatment V: Computer Methods and Experimental Measurements*, S. 323–332, WIT Press, Southampton, 2001.

Dr.-Ing. F. Puente León DS2 (Design of Systems on Silicon), Parque Tecnológico de Valencia, C./ Charles Robert Darwin 2, E-46980 Paterna (Valencia), Spanien, Tel.: +34-96-136 6004, E-Mail: fernando.puente@ds2.es