

David Uhlig\* und Michael Heizmann

# Multi-Stereo-Deflektometrie mit einer Lichtfeldkamera

Multi-stereo deflectometry with a light-field camera

DOI 10.1515/teme-2018-0042

**Zusammenfassung:** Die optische dreidimensionale Form-erfassung spiegelnd reflektierender Objekte ist in der Messtechnik immer noch eine schwierige Aufgabe. Die Deflektometrie rekonstruiert die Oberfläche durch Beobachtung verzerrter Bilder einer reflektierten Referenzszenen. In diesem Beitrag wird erstmalig ein neuer Ansatz präsentiert, bei dem das deflektometrische Messverfahren effizient mit einer Lichtfeldkamera kombiniert wird. Eine Interpretation der Kamera als hochgradig multiples Kamera-Array erlaubt es, das deflektometrische Messproblem mit einem Multi-Stereo-Ansatz zu lösen. Die Leistungsfähigkeit der Oberflächenrekonstruktion wird in experimentellen Messungen bestätigt und es zeigt sich, dass durch die Kombination von Deflektometrie mit Lichtfeldkameras hohe Genauigkeiten erreicht werden können.

**Schlüsselwörter:** Deflektometrie, Lichtfeldkamera, 3D-Rekonstruktion.

**Abstract:** The optical three-dimensional shape reconstruction of specular objects is still a difficult task. Deflectometric methods reconstruct the surface by observing the distorted image of a reflected reference scene. In this article a new approach is presented in which the deflectometric measurement method is carried out efficiently with a light-field camera. An interpretation of the camera as a massive multiple camera array allows to solve the deflectometric measurement problem with a multi-stereo approach. The performance of the surface reconstruction is confirmed in experimental measurements and it is shown that high accuracies can be achieved with the combination of deflectometry and light-field cameras.

**Keywords:** Deflectometry, light-field camera, 3D reconstruction.

---

\*Korrespondenzautor: David Uhlig, Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik (IIT), Hertzstraße 16/06.35, 76187 Karlsruhe, E-Mail: david.uhlig@kit.edu

Michael Heizmann, Karlsruher Institut für Technologie, Institut für Industrielle Informationstechnik (IIT), Hertzstraße 16/06.35, 76187 Karlsruhe, E-Mail: michael.heizmann@kit.edu

## 1 Einleitung

Einem menschlichen Beobachter gelingt es oftmals, aus den Reflexionen eines spiegelnden Objektes auf die spiegelnde Oberfläche selbst zu schließen. Die automatische Sichtprüfung solcher Oberflächen ist in der Messtechnik jedoch immer noch eine große Herausforderung. Zur Vermessung spiegelnder Objekte haben sich die deflektometrischen Messmethoden durchgesetzt. Sie rekonstruieren die 3D-Gestalt, indem sie die durch die reflektierende Oberfläche verzerrten Bilder einer kalibrierten Szene beobachten. Während hierbei sehr hohe Auflösungen der Oberflächenneigung erreicht werden können, hat die Deflektometrie die inhärente Eigenschaft, dass die absolute Höhe mit einer einzelnen herkömmlichen Kamera nicht eindeutig rekonstruiert werden kann. Stattdessen erhält man eine unendliche Vielzahl möglicher Oberflächen [13].

Die Mehrdeutigkeit kann dennoch mittels Zusatzwissen aufgelöst werden. Zhang et al. [14] rekonstruieren spiegelnde Oberflächen durch Vorwissen über eine Referenzebene. Petz und Tutsch [8] lösen die Mehrdeutigkeit auf, indem sie eine Referenzstruktur mittels einer Linearpositioniereinheit in unterschiedliche Positionen fahren und eine sequentielle Erfassung der resultierenden Spiegelbilder auswerten. Balzer und Werling [1] beobachten die spiegelnde Oberfläche aus unterschiedlichen Kameraperspektiven und lösen die Aufgabe mit dem Prinzip der Stereo-Deflektometrie.

Alle bisherigen Verfahren haben gemeinsam, dass ihnen eine herkömmliche Kamera als Aufnahmegesetz dient. In diesem Beitrag soll ein neuer Ansatz präsentiert werden, welcher eine Lichtfeldkamera nutzt, um implizit eine hochgradig multiple Stereo-Deflektometrie zu realisieren. Im Gegensatz zur herkömmlichen Stereo-Deflektometrie mit zwei Kameras kann hier die Szene implizit aus vielen (> 100) Richtungen beobachtet werden. Damit wird eine robustere Messung erreicht, die Mehrdeutigkeit der deflektometrischen Messung aufgelöst und eine hoch genaue Oberflächenrekonstruktion ermöglicht.

Der Beitrag gliedert sich wie folgt: Zunächst werden in Abschnitt 2 die Grundlagen der Deflektometrie beschrieben, das Prinzip der Lichtfeldkamera erklärt und die Kalibrierung des gesamten Messsystems präsentiert. Abschnitt 3 befasst sich mit dem Rekonstruktionsalgorithmus. Schließlich werden in Abschnitt 4 experimentelle Messungen vorgestellt und die Ergebnisse diskutiert.

## 2 Grundlagen

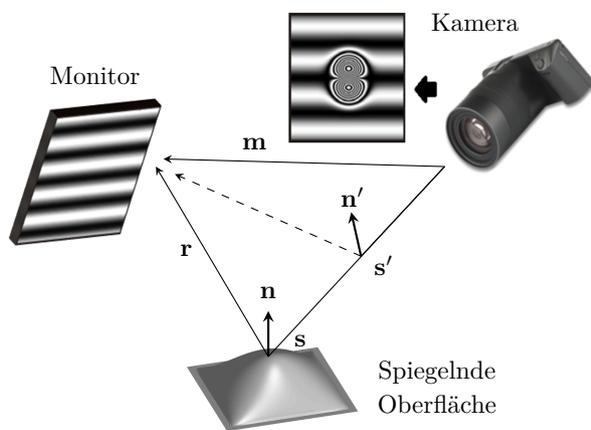
### 2.1 Messprinzip der Deflektometrie

Deflektometrische Messverfahren berechnen die Oberfläche durch Analyse der Oberflächennormalen. Der prinzipielle deflektometrische Messaufbau ist in Abb. 1 zu sehen. Ein Sichtstrahl<sup>1</sup>  $\hat{\mathbf{s}} = \frac{\mathbf{s}}{\|\mathbf{s}\|}$  der Kamera trifft auf die unbekannte spiegelnde Fläche im Punkt  $\mathbf{s} \in S^* \subset \mathbb{R}^3$ , wird reflektiert und beobachtet die Reflexion des Monitors im Punkt  $\mathbf{m}(\hat{\mathbf{s}}) \in M \subset \mathbb{R}^3$ . Die Oberflächennormale eines Spiegelpunktes  $\mathbf{s}$  ergibt sich dann zu:

$$\hat{\mathbf{n}}(\mathbf{s}) = \frac{\hat{\mathbf{s}} - \hat{\mathbf{r}}}{\|\hat{\mathbf{s}} - \hat{\mathbf{r}}\|}, \quad (1)$$

wobei für die Reflexion  $\mathbf{r} = \mathbf{m}(\hat{\mathbf{s}}) - \mathbf{s}$  gilt. Eine Integration über das zur Oberfläche gehörende Normalenfeld kann schließlich die Oberfläche rekonstruieren.

Im Allgemeinen ist in Gl. (1) die Länge des Vektors  $\mathbf{s}$  unbekannt, wodurch sich für jeden Kamerasichtstrahl eine einparametrische Schar hypothetischer Oberflächennorma-



**Abb. 1:** Messprinzip der Deflektometrie. Für jeden Punkt  $\mathbf{s}'$  kann eine hypothetische Normale  $\mathbf{n}'$  berechnet werden.

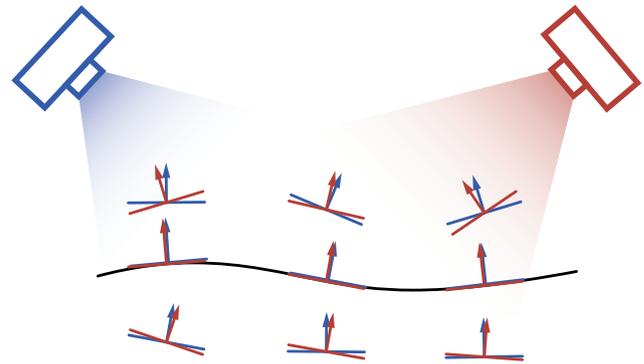
<sup>1</sup> Vektoren und Matrizen werden im Folgenden **fett** geschrieben, Einheitsvektoren mit Dach. Es gilt also  $\|\hat{\mathbf{s}}\| = 1$ , mit  $\hat{\mathbf{s}} := \frac{\mathbf{s}}{\|\mathbf{s}\|}$ .

len berechnen lässt, was wiederum auf eine Mehrdeutigkeit der Oberflächenschätzung führt [1].

Die Mehrdeutigkeit kann durch Zusatzwissen aufgelöst werden. In diesem Beitrag wird dies mit dem Prinzip der Multi-Stereo-Deflektometrie erreicht. Hierbei wird die Oberfläche von mindestens einer weiteren Kamera beobachtet. Im Gegensatz zur klassischen Stereo-Tiefenmessung auf diffusen Oberflächen können auf spiegelnden Oberflächen nicht so einfach korrespondierende Punkte in beiden Kameras gefunden werden. Sichtstrahlen der Kameras, die denselben Oberflächenpunkt beobachten, sehen unterschiedliche Punkte in der Monitorebene. Für beide Kameras kann jedoch eine einparametrische Schar hypothetischer Oberflächennormalen berechnet werden. Da die Normalen auf der wahren Oberfläche identisch sein müssen, kann die Oberfläche durch Vergleich der Normalenfelder gefunden werden. Abbildung 2 verdeutlicht das Prinzip. Eine genauere Beschreibung der Vorgehensweise folgt in Kap. 3.

Um eine deflektometrische Normalenmessung durchzuführen, muss jedem Kamerasichtstrahl (bzw. Pixel)  $\hat{\mathbf{s}}$  ein Punkt in der Monitorebene  $\mathbf{m}(\hat{\mathbf{s}})$  zugeordnet werden. Hierzu wird zunächst eine Registrierung von Kamerapixeln mit Monitorpixeln gesucht  $\hat{\mathbf{s}} \rightarrow \mathbf{l}(\hat{\mathbf{s}})$ . Anschließend müssen die Punkte  $\mathbf{l}$  aus dem lokalen Monitorkoordinatensystem in Punkte  $\mathbf{m}$  aus dem Kamerakoordinatensystem transformiert werden.

Die Registrierung von Kamerapixeln und Monitorpixeln wird im folgenden Abschnitt erläutert. Die Schätzung der Koordinatentransformation folgt in Kap. 2.4.



**Abb. 2:** Regularisierung durch Stereo. Auf der wahren Oberfläche stimmen die deflektometrisch gemessenen Normalen beider Kameras überein.

## 2.2 Codierte Beleuchtung

Die Registrierung von Kamerasichtstrahlen  $\hat{s}$  zu Monitorpunkten  $I(\hat{s})$  wird über eine zeitliche Codierung der Monitorpixel gefunden. Phasenschubverfahren eignen sich hier besonders gut, da sie robust sind gegenüber Rauschen, einer Tiefpassfilterung durch Defokussierung der Kamera und gegenüber einer Variation des Umgebungslichtes [15].

Zur Codierung einer normalisierten Monitorkoordinaten  $x \in [0, 1]$  wird eine Signalfolge von  $K$  um  $\Psi_k = \frac{2\pi k}{K}$  gegeneinander verschobenen sinusförmigen Mustern mit Phase  $\varphi(x) = 2\pi f x$  gezeigt:

$$g_k(x) = A + B \cos(\varphi(x) + \Psi_k). \quad (2)$$

In der Kamera wird die Signalfolge  $\tilde{g}_k$  mit  $k = 0 \dots K - 1$  aufgenommen und daraus die Phase berechnet:

$$a = \sum_{k=0}^{K-1} \tilde{g}_k \sin(\Psi_k), \quad b = \sum_{k=0}^{K-1} \tilde{g}_k \cos(\Psi_k) \quad (3)$$

$$\varphi = \arctan(-a, b) \in [-\pi, \pi]. \quad (4)$$

Für symmetrische  $K$ -Schritt-Phasenschubverfahren kann die Unsicherheit der Phasenmessung  $\sigma_\varphi$  in Abhängigkeit des Sensorrauschens  $\sigma_I$  angegeben werden [4]:

$$\sigma_\varphi = \frac{1}{f} \sqrt{\frac{2}{K} \frac{\sigma_I}{\hat{B}}}, \quad (5)$$

wobei  $\hat{B} = \frac{2}{K} \sqrt{a^2 + b^2}$  eine Schätzung der Modulation  $B$  des Signals darstellt. Eine Maximierung der Frequenz minimiert daher die Unsicherheit. Bei zu hoher Frequenz verringert jedoch das Tiefpassverhalten des Systems den Kontrast des Signals,  $\hat{B}$  wird kleiner und die Unsicherheit  $\sigma_\varphi$  wieder größer.

Bei Frequenzen  $f > 1$  ist keine eindeutige Rekonstruktion der Phase möglich, da die Phase Werte  $\varphi > 2\pi$  annehmen kann, der Arkustangens jedoch nur in den Bereich  $[-\pi, \pi]$  abbildet. Gleichung (4) wird  $2\pi$ -Unstetigkeiten aufweisen. Um einen kontinuierlichen Phasenverlauf zu erhalten, werden Methoden der Phasenfaltung benötigt. Hierbei wird die Phase für verschiedene Frequenzen  $f_i$  berechnet, um anschließend für jede Frequenz eine kontinuierliche Phasenschätzung  $\varphi_i \in [0, 2\pi]$  zu erhalten. Der interessierte Leser sei hier auf die Literatur verwiesen [16].

Für mehr Robustheit können die Phasenschätzungen mit einem gewichteten Mittelwert fusioniert werden. Die Gewichtung der jeweiligen Schätzung wird hier jeweils durch die geschätzte Unsicherheit  $\sigma_{\varphi_i}$  angegeben. Da die

Phase periodisch im Intervall  $[0, 2\pi]$  ist, darf nicht das übliche arithmetische Mittel berechnet werden. Stattdessen wird ein zirkularer Mittelwert gebildet, indem die Phase auf den Einheitskreis abgebildet und ein gewichtetes Mittel der komplexen Zeiger berechnet wird:

$$z_i = e^{j\varphi_i}, \quad z = \frac{\sum_i \frac{1}{\sigma_{\varphi_i}^2} z_i}{\sum_i \frac{1}{\sigma_{\varphi_i}^2}}. \quad (6)$$

Die Phase kann nun aus dem Argument des komplexen Zeigers berechnet werden:

$$\varphi = \arctan(\text{Im}(z), \text{Re}(z)) \bmod 2\pi, \quad (7)$$

woraus sich schließlich die decodierte Monitorcoordinate berechnen lässt:

$$x = \frac{\varphi}{2\pi}. \quad (8)$$

## 2.3 Lichtfelder und Lichtfeldkameras

Das Licht einer Szene kann in der geometrischen Optik durch die plenoptische Funktion mit sechs Variablen beschrieben werden: drei Ortskoordinaten, zwei Winkelkoordinaten, ein Spektralwert. In einer herkömmlichen Kamera kann i.A. nur ein Unterraum dieser Funktion eingefangen werden: zwei Ortskoordinaten mit Farb-/Intensitätswert. Eine Lichtfeldkamera erlaubt es, daneben noch zwei zusätzliche Winkeldimensionen einzufangen.

Der Aufbau einer Lichtfeldkamera ähnelt dem der herkömmlichen Kamera, mit dem Unterschied, dass vor dem Sensor ein Array von Mikrolinsen platziert ist [7]. Abbildung 3 zeigt den Aufbau der Lichtfeldkamera.

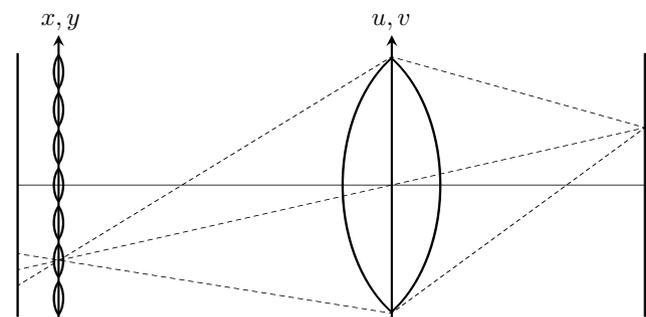


Abb. 3: Schematischer Aufbau der Lichtfeldkamera. Links: Sensor mit Mikrolinsenarray. Mitte: Hauptlinse. Rechts: Objektebene.

Durch das Hinzufügen des Mikrolinsenarrays wird es möglich, einen Ausschnitt des fünfdimensionalen Lichtfeldes  $L(u, v, x, y, c)$  einer Szene einzufangen. Hier bezeichnen  $x, y$  die Koordinaten der Mikrolinsen vor dem Sensor

2 Die Codierung erfolgt sowohl in horizontaler als auch in vertikaler Richtung.

und damit die örtliche Dimension des Lichtfeldes.  $u, v$  beschreiben die Koordinaten innerhalb der Mikrolinse relativ zum Mikrolinsenmittelpunkt und geben dadurch implizit Auskunft, wo ein Lichtstrahl durch die Hauptlinse gekommen ist. Sie repräsentieren die Winkelinformation des Lichtfeldes.  $c$  steht für die spektrale Dimension bzw. die Farbe.

Im Vergleich zur herkömmlichen Kamera erlauben es diese zusätzlichen Informationen, die Perspektive auf die Szene zu ändern, Tiefeninformation zu extrahieren [12] oder auch die Fokusebene im Nachhinein zu verschieben [7].

Hält man  $u = u^*$ ,  $v = v^*$  fest, wählt also aus jedem Mikrolinsenbild<sup>3</sup> die Pixel mit jeweils gleicher relativer Position zum Zentrum der Mikrolinse, so kann für jedes Paar ein Sub-Apertur-Bild (SAB) extrahiert werden [2]:

$$I_{u^*,v^*}(x, y) := L(u^*, v^*, x, y). \quad (9)$$

Diese SABs können näherungsweise als Lochkameras mit großer Tiefenschärfe interpretiert werden, da sie jeweils nur einen kleinen Teil der Hauptlinse beobachten. Insgesamt lässt sich die Lichtfeldkamera deswegen auch als ein massiv multiples Stereo-Kamera-Array von vielen leicht gegeneinander verschobenen, virtuellen Kameras interpretieren (Abb. 4). Bei bspw.  $14 \times 14$  Pixeln unter einer Mikrolinse erhält man demnach ein Array von 196 virtuellen Kameras<sup>4</sup>.

Dieses Kamera-Array kann genutzt werden, um die Mehrdeutigkeit der deflektometrischen Normalenmessung mithilfe eines multiplen Stereo-Ansatzes zu regularisieren.

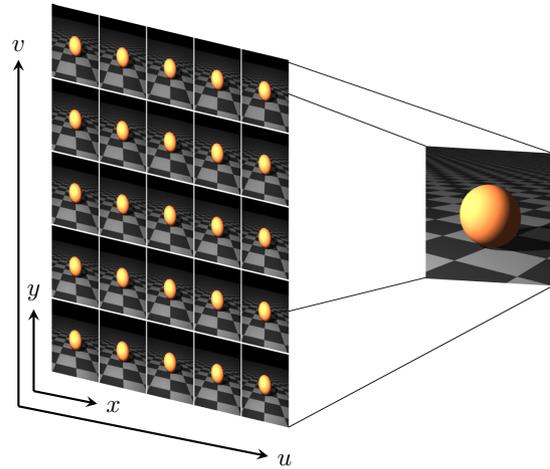
## 2.4 Kalibrierung des Messsystems

Damit die Oberflächennormalen aus Sichtstrahl  $\hat{s}$  und Monitorpunkt  $\mathbf{m}$  berechnet werden können, muss das gesamte Messsystem kalibriert werden.

Gängige modellbasiertes Verfahren [2] kalibrieren die Lichtfeldkamera und erzeugen aus den Lichtfelddaten rektifizierte SABs mit zugehörigen intrinsischen Kameraparametern sowie die extrinsischen Transformationen zwischen den SABs. Durch die Kalibrierung werden die SABs in einem virtuellen Kameraarray in äquidistanten Abständen zueinander angeordnet und für jedes Pixel kann ein Sichtstrahl  $\hat{s}(x, y)$  berechnet werden.

<sup>3</sup> Die Rohdaten des Sensors müssen zunächst decodiert werden, um ein nutzbares Lichtfeld  $L(u, v, x, y)$  zu erhalten, siehe [2].

<sup>4</sup> Bei realen Lichtfeldkameras leiden im Allgemeinen die SABs am Rand der Mikrolinsen unter einer starken Vignettierung. Deshalb sind effektiv nicht alle SABs nutzbar.



**Abb. 4:** Interpretation des Lichtfeldes als Kamera-Array. Jedes SAB repräsentiert eine „virtuelle“ Kamera, die gegenüber den anderen Kameras leicht verschoben ist.

Der zweite Schritt in der deflektometrischen Kalibrierung ist die Schätzung der relativen Kamera-Monitor-Pose. Es wird die Transformation von Monitorkoordinaten  $\mathbf{l} = (l_x, l_y, 0)^\top$  zu Kamerakoordinaten  $\mathbf{m} = (m_x, m_y, m_z)^\top$  gesucht:

$$\mathbf{m} = \mathbf{R} \cdot \mathbf{l} + \mathbf{t} \quad (10)$$

mit Rotation  $\mathbf{R}$  und Translation  $\mathbf{t}$ . Die Kamera-Posen-schätzung benötigt Sichtkontakt mit dem Monitor. Im Allgemeinen wird dieser aber nicht im direkten Sichtfeld der Kamera liegen. Die Pose kann dennoch geschätzt werden, indem die Reflexion des Monitors in einem planaren Spiegel beobachtet wird. Die Pose von Kamera zum gespiegelten, virtuellen Monitor kann anschließend über Standard-PnP-Verfahren geschätzt werden [6]. Um die Anzahl der Punkt-zu-Punkt-Korrespondenzen zu erhöhen und um eine robuste Schätzung zu erreichen, wird hier jedem Kamerapixel mithilfe des Phasenschiebeverfahrens ein subpixelgenauer Punkt in der virtuellen Monitorebene zugeordnet.

Bei Änderung der Position des Spiegels sieht die Kamera jeweils eine neue Reflexion des Monitors. Bei gleichbleibender Kamera-Monitor-Pose kann nun die Transformation von Kamera- zu realen Monitorkoordinaten geschätzt werden [11].

Die Schätzung der Kamera-Monitor-Pose kann dabei für jedes SAB individuell durchgeführt werden. Die Kalibrierung der Lichtfeldkamera liefert aber bereits die Transformationen zwischen den SABs. Es wird deshalb insgesamt nur eine einzige Transformation von Monitorkoordinaten zu den Kamerakoordinaten eines Referenz-SAB benötigt. Daher können die einzelnen, individuell geschätz-

ten Transformationen zu einer Einzigem gemittelt werden, was die Robustheit der Schätzung erhöht.

Die mittlere Translation kann dabei über den herkömmlichen Mittelwert berechnet werden. Da Rotationsmatrizen Elemente der speziellen Orthogonalen Gruppe sind,  $\mathbf{R} \in \text{SO}(3)$ , kann die mittlere Rotation  $\mathbf{R}$  z. B. durch

$$\mathbf{R} = \tilde{\mathbf{R}} \left( \tilde{\mathbf{R}}^\top \tilde{\mathbf{R}} \right)^{-1/2} \quad (11)$$

aus den individuell geschätzten Rotationen  $\mathbf{R}_{u,v}$  der SABs berechnet werden, wobei  $\tilde{\mathbf{R}} = \frac{1}{N} \sum_{u,v} \mathbf{R}_{u,v}$  den arithmetischen Mittelwert der geschätzten Rotationsmatrizen darstellt [10].

### 3 Oberflächenrekonstruktion

Grundsätzlich wird für die Rekonstruktion der Oberfläche zunächst mithilfe des Phasenschubverfahrens die Registrierung zwischen Monitorpixeln und Kamerapixeln durchgeführt. Die Phasendecodierung erfolgt hier auf den Sensorrohdaten. Erst danach wird das rektifizierte Lichtfeld aus diesen decodierten Daten extrahiert, wodurch die Registrierung der SABs zum Monitor erhalten wird. Der Vorteil dieser Vorgehensweise liegt darin, dass die Qualität der Phasemessung nicht durch Interpolationen während des Sensor-Demosaicing leidet.

Mithilfe der intrinsischen Kalibrierung der Kamera kann für jedes SAB  $I_{u,v}(x, y)$  individuell eine hypothetische Oberflächennormale  $\hat{\mathbf{n}}_{u,v}(\mathbf{s})$  für jeden Punkt im Messraum  $\mathbf{s} \in S$  berechnet werden.

Auf der wahren Oberfläche müssen die Oberflächennormalen aller SAB übereinstimmen [1]. Die optimalen Punkte, auf denen die hypothetischen Normalen möglichst ähnlich sind, können durch Minimieren eines geeigneten Gütemaßes gefunden werden. Ein geeignetes Maß ist z. B. der mittlere Winkel zwischen den Normalen [5]:

$$J(\mathbf{s}) = \frac{1}{N} \sum_{u,v} \arccos \left( \left\langle \hat{\mathbf{n}}_{u,v}, \frac{1}{N} \sum_{u,v} \hat{\mathbf{n}}_{u,v} \right\rangle \right)^2, \quad (12)$$

wobei  $N$  die Anzahl der gültigen Oberflächennormalen im Punkt  $\mathbf{s}$  bezeichnet. Betrachtet man  $\mathbf{s}(x, y)$  entlang der Kamerasichtstrahlen, so kann das Gütemaß pixelweise mithilfe eines Linien-Suchverfahrens minimiert werden [3]. Dadurch wird die optimierte Oberfläche sowie die zugehörigen Oberflächennormalen erhalten:

$$\mathbf{s}^0 = \arg \min_{\mathbf{s} \in S} J(\mathbf{s}), \quad \hat{\mathbf{n}}^* = \hat{\mathbf{n}}(\mathbf{s}^0). \quad (13)$$

Bei realen Messungen werden aufgrund von Rauschen die Oberflächennormalen der berechneten Fläche  $\mathbf{s}^0$  und

die deflektometrisch gemessenen Normalen voneinander abweichen. Da diese aber möglichst gut übereinstimmen sollten, ist eine weitere Optimierung möglich, indem die Differenz dieser Normalen im Sinne der in [9] vorgeschlagenen Normalenfeldintegration mit Nebenbedingung weiter minimiert wird.

Die Oberfläche  $\mathbf{s}^0$  kann dazu mithilfe einer Zentralprojektion auf eine Tiefenkarte  $z^0(x, y)$  mit regulärem Gitter  $(x, y) \in \Omega$  abgebildet werden. Der Gradient der gesuchten Oberfläche  $z(x, y)$  lässt sich aus den gemessenen Normalen  $\hat{\mathbf{n}}^* = (n_1, n_2, n_3)^\top$  berechnen:

$$\nabla z(x, y) = \mathbf{g}(x, y) := \left( -\frac{n_1}{n_3}, -\frac{n_2}{n_3} \right)^\top. \quad (14)$$

Integration über den Gradienten würde die Oberfläche mit einer einparametrischen Mehrdeutigkeit rekonstruieren. Die durch Minimierung von Gl. (12) gefundenen Punkte können das Problem jedoch regularisieren und die Mehrdeutigkeit auflösen. Minimieren des in [9] vorgeschlagenen Funktionals

$$z^* = \min_z \int \int_{(x,y) \in \Omega} \|\nabla z(x, y) - \mathbf{g}(x, y)\|^2 + \lambda(x, y) [z(x, y) - z^0(x, y)]^2 dx dy \quad (15)$$

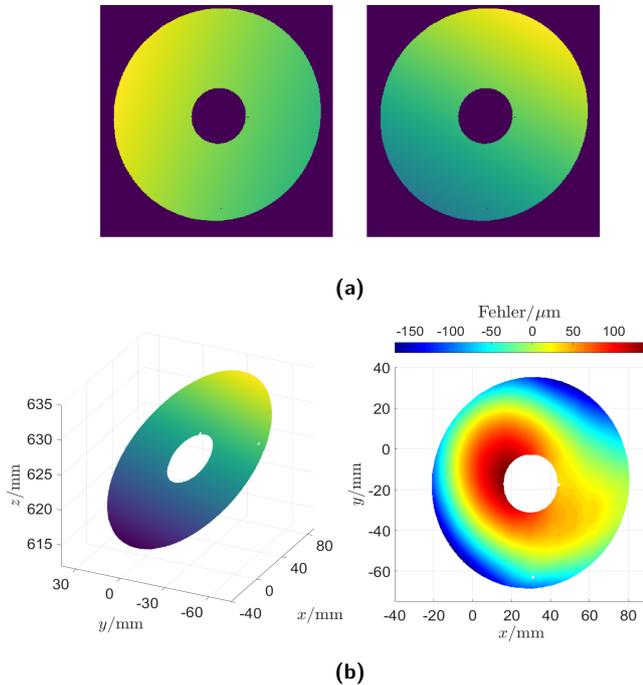
liefert die rekonstruierte Oberfläche. Der Faktor  $\lambda(x, y)$  legt hierbei fest, wie sehr die durch die Stereo-Deflektometrie gefundenen Regularisierungspunkte in der Normalenintegration beachtet werden sollen. Prinzipiell würde ein einzelner Punkt ausreichen, um die komplette Oberfläche eindeutig aus dem Normalenfeld zu integrieren [1]. Mehr Punkte erhöhen jedoch die Robustheit des Verfahrens.

Nach der Optimierung besitzt die Oberfläche minimalen Abstand zu den Regularisierungspunkten und gleichzeitig wird der Unterschied von berechneten zu deflektometrisch gemessenen Normalen minimiert.

## 4 Ergebnisse

### 4.1 Evaluation

Zur Evaluation und Verifikation des vorgestellten Ansatzes wurden experimentelle Untersuchungen an Testobjekten durchgeführt. Das verwendete Aufnahmegerät ist eine Lytro Illum. Nach der Kamerakalibrierung werden  $13 \times 13$  SABs mit einer Auflösung von  $338 \times 552$  Pixeln erhalten. Von den 169 SABs werden effektiv 100 für die Multi-Stereo-Deflektometrie genutzt. Ein Monitor (Sam-

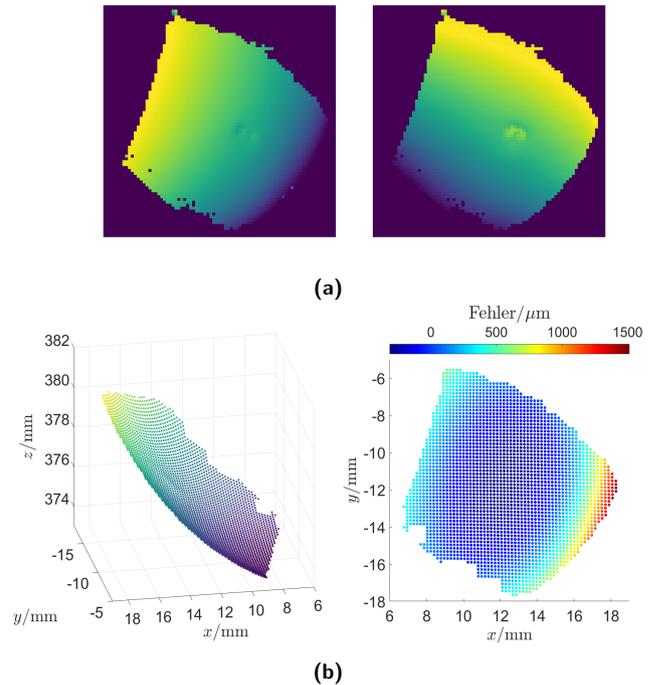


**Abb. 5:** Festplattendisk: a) Messung der Phase in horizontaler und vertikaler Richtung für das zentrale SAB. b) Links: Rekonstruierte Oberfläche. Rechts: Abweichung der Oberfläche zu idealer Ebene (mittlerer Fehler: RMSE  $\approx 62 \mu m$ ).

sung SyncMaster 2493HM, Auflösung  $1920 \times 1200$ , Pixelgröße  $0,27 \text{ mm}$ ) erzeugt die Phasenschubmustersequenzen. Die Phasenfaltung erfolgt mit einem Multi-Frequenz-Verfahren [16]. Die relativen Frequenzen der Sinusmuster werden zu  $f_1 = 1$ ,  $f_2 = 1/4$ ,  $f_3 = 1/16$ ,  $f_4 = 1/50$  gewählt.

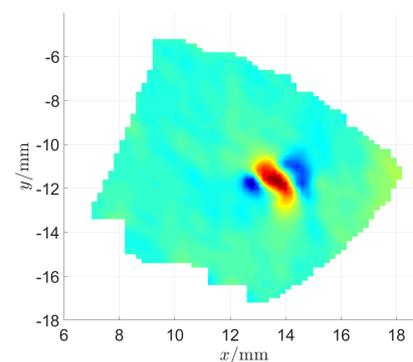
Als Testobjekte dienen eine Disk aus einer Festplatte und eine Metallkugel. Zu beiden liegt keine absolute Ground Truth Information vor. Festplattendisks weisen jedoch eine sehr hohe Planarität und geringe Rauheit auf, weswegen hier eine ideale Ebene als Referenz angenommen wird. Die Metallkugel hat einen Radius von  $2 \text{ cm}$ . Eine ideale Kugel mit selbem Radius dient daher als Referenz.

Die Vermessung der Festplattendisk und das Ergebnis der 3D-Rekonstruktion zeigt Abb. 5. Zur Berechnung eines Rekonstruktionsfehlers wurde eine ideale Ebene in die rekonstruierte Oberfläche gelegt. Die Abweichungen der rekonstruierten Oberfläche von dieser idealen Ebene bewegen sich im Bereich  $\pm 150 \mu m$ , wobei die Wurzel der summierten Fehlerquadrate zu  $RMSE = 62,7 \mu m$  berechnet wird. Es sei anzumerken, dass eine Vermessungen der Festplattendisk aus unterschiedlichen Positionen und mit unterschiedlichen Kalibrierungen ähnliche Rekonstruktionsfehler lieferte. Eine leichte Krümmung der Disk ist deshalb nicht auszuschließen.



**Abb. 6:** Metallkugel: a) Messung der Phase in horizontaler und vertikaler Richtung für das zentrale SAB. b) Links: Rekonstruierte Oberfläche. Rechts: Abweichung der Oberfläche zu idealer Kugel mit Radius  $2 \text{ cm}$  (mittlerer Fehler: RMSE  $\approx 113 \mu m$ ).

Die Vermessung der Metallkugel und das Ergebnis der 3D-Rekonstruktion zeigt Abb. 6. Auch hier wurde zur Berechnung des Rekonstruktionsfehlers eine ideale Kugeloberfläche in das berechnete Ergebnis gelegt. Die Abweichungen zur idealen Oberfläche bewegen sich hier im Bereich von  $-300 \mu m$  bis  $1,5 \text{ mm}$  mit einem mittleren Fehler  $RMSE = 112,6 \mu m$ . Während die globale Oberflächenstruktur gut rekonstruiert werden kann, so lassen sich auch lokale Defekte erkennen. Eine Analyse der lokalen Krümmung [5] macht dies leicht sichtbar, siehe Abb. 7.



**Abb. 7:** Metallkugel: Analyse der lokalen Krümmung der in Abbildung 6 rekonstruierten Oberfläche macht Defekte sichtbar.

## 4.2 Fazit

Die Experimente zeigen, dass der vorgestellte Ansatz funktioniert und sowohl global die Oberfläche, als auch lokal Defekte rekonstruieren kann. Trotz der sehr kleinen Stereo-Basis zwischen den Sub-Apertur-Bildern (i.A.  $< 1$  mm), aber aufgrund ihrer immensen Anzahl kann eine hohe Genauigkeit des Messverfahrens erreicht werden. Die Lichtfeldkamera kombiniert in kompakter Bauform die Vorteile eines Stereo-Kamera-Systems. Ein weiterer Vorteil liegt in der sehr hohen Tiefenschärfe der Sub-Apertur-Bilder, wodurch der Kamerafokus auf die Oberfläche gelegt werden kann, um eine detailreiche Vermessung zuzulassen. Gleichzeitig sind aber auch die Codemustersequenzen auf dem Monitor noch beobachtbar. Das im Gegenzug schlechtere Signal-zu-Rausch-Verhältnis durch die geringere Lichtintensität in den SABs kann bspw. durch ein höherwertiges Phasenschubverfahren wieder ausgeglichen werden.

Dennoch bleibt Optimierungsbedarf. Während die Integration des Normalenfeldes algorithmisch sehr effizient ist, ist die pixelweise Minimierung des mittleren Normalenwinkels ein sehr aufwändiges Verfahren. Im Weiteren weist die Lichtfeldkamera für die äußeren SABs immer größer werdende Kalibrierungsfehler auf. Eine Optimierung der Kamerakalibrierung für messtechnische Anwendungen könnte hier die Genauigkeit des Verfahrens noch erhöhen.

## 5 Zusammenfassung

In der vorgestellten Arbeit wird erstmalig ein Ansatz präsentiert, welcher die speziellen optischen Eigenschaften einer Lichtfeldkamera nutzt, um spiegelnde Oberflächen deflektometrisch zu vermessen. Die Lichtfeldkamera wird hierbei als ein hochgradig multiples Kamera-Array interpretiert, womit sich die Möglichkeit eröffnet, die Mehrdeutigkeit der deflektometrischen Normalenschätzung mittels eines Multi-Stereo-Ansatzes zu regularisieren. Durch Vergleich der mehrfachen Normalenschätzungen in einem Punkt wird eine initiale Schätzung der Oberfläche gefunden. Eine weitere Optimierung wird möglich, indem diese initiale Schätzung als Regularisierungspunkte einer Flächenberechnung durch Integration über das Normalenfeld dienen. Letztlich zeigen auch experimentelle Untersuchungen, dass sich die multiple Stereo-Deflektometrie effizient und in kompakter Bauform mithilfe einer Lichtfeldkamera realisieren lässt und eine hoch genaue Oberflächenrekonstruktion ermöglicht.

## Literatur

- [1] J. Balzer und S. Werling. Principles of shape from specular reflection. *Measurement*, 43(10):1305–1317, 2010.
- [2] Y. Bok, H.-G. Jeon und I. S. Kweon. Geometric calibration of micro-lens-based light field cameras using line features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2):287–300, 2017.
- [3] R. P. Brent. *Algorithms for minimization without derivatives*. Courier Corporation, 2013.
- [4] M. Fischer, M. Petz und R. Tutsch. Modellbasierte Rauschvorhersage für Streifenprojektionssysteme – Ein Werkzeug zur statistischen Analyse von Auswertalgorithmen. *tm - Technisches Messen*, 84(2), 2017.
- [5] S. Kammel. *Deflektometrische Untersuchung spiegelnder reflektierender Freiformflächen*. PhD thesis, Universitätsverlag Karlsruhe, Karlsruhe, 2005.
- [6] V. Lepetit, F. Moreno-Noguer und P. Fua. EPnP: An accurate  $O(n)$  solution to the PnP problem. *International Journal Computer Vision*, 81(2), 2009.
- [7] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz und P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Computer Science Technical Report CSTR*, 2(11):1–11, 2005.
- [8] M. Petz und R. Tutsch. Rasterreflexions-Photogrammetrie zur Messung spiegelnder Oberflächen (reflection grating photogrammetry for the measurement of specular surfaces). *tm - Technisches Messen*, 71:389–397, 2004.
- [9] Y. Quéau, J.-D. Durou und J.-F. Aujol. Normal integration: a survey. *Journal of Mathematical Imaging and Vision*, 60(4): 576–593, 2018.
- [10] I. Sharf, A. Wolf und M. B. Rubin. Arithmetic and geometric solutions for average rigid-body rotation. *Mechanism and Machine Theory*, 45(9):1239–1251, 2010.
- [11] K. Takahashi, S. Nobuhara und T. Matsuyama. A new mirror-based extrinsic camera calibration using an orthogonality constraint. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, S. 1051–1058, 2012.
- [12] A. Vianello, G. Manfredi, M. Diebold und B. Jähne. 3d reconstruction by a combined structure tensor and hough transform light field approach. *tm - Technisches Messen*, 2017.
- [13] S. Werling, M. Mai, M. Heizmann und J. Beyerer. Inspection of specular and partially specular surfaces. *Metrology and Measurement Systems*, 16(3):415–431, 2009.
- [14] H. Zhang, S. Han, S. Liu, S. Li, L. Ji und X. Zhang. 3d shape reconstruction of large specular surface. *Applied Optics*, 51(31):7616–7625, 2012.
- [15] S. Zhang. Phase unwrapping error reduction framework for a multiple-wavelength phase-shifting algorithm. *Optical Engineering*, 48(10):105601–105601–8, 2009.
- [16] C. Zuo, L. Huang, M. Zhang, Q. Chen und A. Asundi. Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review. *Optics and Lasers in Engineering*, 85:84–103, 2016.