

# Ableitung von 3D-Verschiebungsvektoren aus fusionierten Laserscan- und Bilddaten

## Determination of 3D Displacement Vectors from Combined Laser-Scan and Image Data

Wolfgang Wiedemann, Andreas Wagner, Thomas Wunderlich

Flächenhafte Messmethoden wie Laserscanning und (Stereo-)Photogrammetrie gewinnen im Bereich geodätischer Überwachungsmessungen zunehmend an Bedeutung. Jedes dieser Verfahren birgt dabei individuelle Vor- und Nachteile bei der Erfassung von Objektbewegungen. In 3D-Punktwolken von Laserscannern sind Bewegungen in Blickrichtung einfach identifizierbar. Das Erkennen von Bewegungen senkrecht zur Blickrichtung des Scanners ist hingegen nur unter bestimmten Bedingungen möglich. Im Gegensatz dazu besitzen hochaufgelöste Bilder die größte Sensitivität für Objektbewegungen senkrecht zur Blickrichtung und zeigen Schwächen bei reinen Distanzänderungen.

Mit der hier vorgestellten Fusionierung von Bild- und Laserscandaten zu sogenannten RGB+D-Bildern lassen sich die Schwächen der Einzelsysteme für die Deformationsanalyse beheben und die Vorteile vereinen. Dazu wird die 3D-Punktwolke in einen Tiefenkanal (D) umgewandelt und mit dem Farbbild (RGB) verschmolzen. In den kombinierten RGB+D-Bildern kann jedes Pixel direkt in 3D-Koordinaten umgerechnet werden. Die notwendige relative Orientierung zwischen Scanner und Kamera ist sowohl im Voraus durch eine sorgfältige Kalibrierung der Sensoren als auch nachträglich aus den Messdaten selbst bestimmbar. Durch das Auffinden korrespondierender Punkte in den RGB+D-Bildern aufeinanderfolgender Messperioden ist die direkte Bestimmung von 3D-Verschiebungsvektoren möglich. Diese Ergebnisse können in eine strenge Deformationsanalyse mit Signifikanztest eingebunden werden. Korrespondierende Punkte zwischen den Epochen lassen sich über etablierte Bildverarbeitungsalgorithmen ermitteln. Bekannte Vertreter zur Extraktion von Bildmerkmalen sind z. B. der SIFT-Algorithmus oder der binäre Deskriptor BRISK. Diese Algorithmen beschreiben den zu untersuchenden Bildausschnitt über numerische Werte, welche epochenübergreifend verglichen und einander zugeordnet werden können.

Die kombinierte Auswertung wird an einem künstlichen Versuchsaufbau detailliert beschrieben und mit herkömmlichen Methoden zur Auswertung von Laserscandaten verglichen. Es zeigt sich, dass die Ergebnisse des vorgestellten Ansatzes leichter und klarer zu interpretieren sind und Scheindeformationen vermieden werden. Es lassen sich Daten unterschiedlichster Aufnahmesysteme – wie z. B. moderner scannender Totalstationen, Mobile-Mapping-Systemen, UAV oder Roboterplattformen – auswerten und ihr volles Potenzial für die Deformationsanalyse nutzen.

*Schlüsselwörter:* Deformationsanalyse, geodätische Überwachungsmessungen, Bildverarbeitung, Laserscanning, RGB+D, Sensorfusion

*Areal measurements, like laser scans or camera images, are more and more frequently used for deformation monitoring. Each single acquisition method has its advantages and disadvantages in detecting displacements. In dense point clouds of laser scans, distance changes in line of sight are easy to identify. In contrast, high-resolution image data is sensitive to displacements perpendicular to the viewing direction of the camera, where the analysis of scan data weakens.*

*A promising solution of combining the advantages and reducing the drawbacks is the fusion of laser-scan and image data in the form of RGB+D images. Point clouds are converted into a depth image, resulting in an additional D-channel to the color (RGB) image data. In the combined RGB+D image, each pixel can be directly converted into 3D coordinates. The necessary mutual orientation of both data sets can be solved via a-priori sensor calibration or a-posteriori data registration. By identifying corresponding points in subsequent measurement epochs (RGB+D images), it is further possible to directly determine 3D displacement vectors. The results of the RGB+D method can be integrated in a rigorous geodetic deformation analysis with tests of significance.*

*Corresponding points in the images can be found by identifying key points and by describing them with image features, for example. Prominent representatives of image features are the SIFT algorithm or the binary descriptor BRISK. These algorithms are used to match two images at an abstract numerical level, instead of in the original image domain.*

*The combined analyzing strategy is described in detail with the help of an artificial experimental setup and is compared with conventional methods for the evaluation of laser scan data. It appears that the results of the presented approach can be interpreted more easily and clearly, and that it will avoid apparent deformations. Sensor data of any kind of acquisition system can be used, such as modern total stations, mobile mapping systems, unmanned aerial vehicles or robot platforms. It will exploit the full potential of sensor data of such systems, which already provide both kind of data, for the first time.*

*Keywords: Deformation monitoring, geodetic monitoring, image matching, laser scanning, RGB+D, sensor fusion*

## 1 EINFÜHRUNG

Im Bereich der geodätischen Überwachungsmessungen ist eine fundamentale Änderung zu erkennen: Zunehmend ergänzen flächenhafte Messsysteme die traditionelle Diskretisierung der Netzmessung (manuelle Auswahl von diskreten Punkten) bzw. lösen diese gar ab. Oft kommen hierfür Laserscanner zum Einsatz, die es erlauben, Objekte schnell und flächenhaft abzutasten und die gewonnenen Informationen als 3D-Punktwolke zu speichern. Allerdings erweist es sich als sehr schwierig, aus diesen Daten verlässliche Aussagen über eine Objektbewegung oder -deformation zu treffen; insbesondere fehlen die dafür nötigen Methoden, um statistische Kennzahlen abzuleiten. Dies hat seine Ursache in dem grundlegenden Problem des Laserscannings, in den Epochen einer Zeitreihe keine identischen Punkte messen zu können. Das vom Scanner erfasste Punktraaster kommt – auch bei identischen Scanner-Einstellungen – wahllos auf dem abgetasteten Objekt zu liegen. Dadurch ist keine strenge Deformationsanalyse mit statistischen Signifikanztests möglich /Wunderlich et al. 2016/. Bekannte Methoden zur Berechnung von Deformationen zwischen zwei Punktwolken (wie z. B. der Punkt-zu-Punkt-Vergleich zwischen zwei Epochen) sind im Wesentlichen nur sensitiv für Bewegungen in Blickrichtung des Scanners. Dem gegenüber steht die Bildverarbeitung, bei der Objektbewegungen senkrecht zur Blickrichtung der Kamera leicht detektiert werden können, eine Distanzvariation aber nur schwer zu erkennen ist.

Die hier vorgestellte kombinierte Auswertung von Bilddaten und Laserscandaten vereint die Vorteile der beiden Verfahren und behebt den jeweiligen Nachteil der Einzelauswertung. Die Messdaten des Laserscanners werden dazu in einen Tiefenkanal umgewandelt und mit den Kameradaten zu RGB+D-Bildern ver-

schmolzen /Wagner 2016/. Um die Funktionsweise dieser neuen Mess- und Auswertemethode zu demonstrieren, wurde ein künstlicher Versuchsaufbau gewählt (Abschnitt 2). Die Erfassung der Messdaten und deren Umwandlung in RGB+D-Bilder wird in Abschnitt 3 beschrieben. Mit den gewonnenen Daten mehrerer Epochen wird anschließend eine Deformationsanalyse auf Basis der RGB+D-Bilder durchgeführt und die Ergebnisse werden mit herkömmlichen Methoden zur Auswertung von Laserscandaten verglichen (Abschnitt 4). Eine Zusammenfassung und ein Ausblick (Abschnitt 5) schließen den Beitrag ab.

## 2 VERSUCHSAUFBAU

*Abb. 1* zeigt die erfassten Daten eines künstlichen Versuchsaufbaus, anhand dessen die Auswertemethode und Ergebnisse demonstriert werden. Das hochauflösende Farbbild (RGB) (*Abb. 1a*) zeigt drei unterschiedliche Objekte vor einem texturierten Hintergrund. In *Abb. 1b* ist die erfasste Punktwolke (XYZ) – eingefärbt nach den Intensitätswerten (*I*) – abgebildet. Beide Datensätze werden mit einer bildgebenden scannenden Totalstation erfasst; die äußere Orientierung zwischen Scan- und Kameraeinheit ist aus einer vorangegangenen Kalibrierung bekannt.

Die Szene hat eine Abmessung von ca. 75 cm × 30 cm und eine Tiefe von ca. 80 cm. 15 überlappende Einzelbilder der Teleskopkamera einer Leica MS60 (2560 Pixel × 1960 Pixel, 0,61 mgon/Pixel) werden zu einem hochauflösenden Panorama (10600 Pixel × 4821 Pixel) zusammengefügt und überblendet, vgl. /Wagner 2016/. Der mittlere Abstand zwischen dem

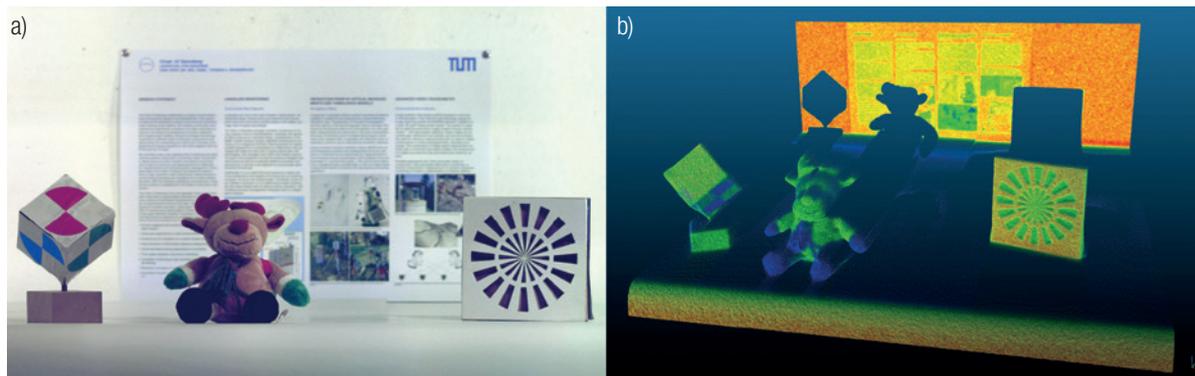


Abb. 1 | Versuchsaufbau: a) hochauflösendes Kamerabild (RGB), b) Laserscandaten (XYZ+I)

Projektionszentrum der Kamera und dem Ursprung des Laserscans zu den Objekten beträgt 8,85 m. Gescannt wird die Szene mit einem mittleren Punktabstand von 0,5 mm, was einer Winkelauflösung von ca. 3,2 mgon entspricht. Das Messrauschen des Laserscans wird vom Hersteller in der gemessenen Distanz mit 0,6 mm, die absolute Positionsgenauigkeit mit 2,0 mm angegeben /Leica Geosystems 2016/.

Das mittlere Objekt (Stofftier) im Vordergrund wird während des Experiments auf unterschiedlichen Positionen platziert, wie in *Abb. 2* dargestellt. Es werden insgesamt drei Epochen aufgenommen. Nach der ersten Messung wird das Objekt um ca. 20 Zentimeter nach links verschoben und die Daten für Epoche 2 erfasst. Für die Aufnahme der 3. Epoche wird das Objekt erneut ca. 20 cm, diesmal in Richtung des Aufnahmestandpunkts, verschoben. Damit wird eine Deformation senkrecht zur Aufnahme- und anschließend eine reine Distanzänderung simuliert.

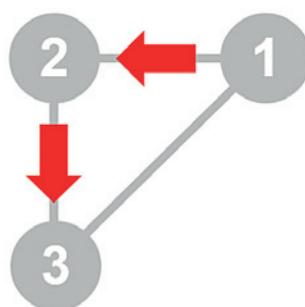


Abb. 2 | Positionen des bewegten Objekts in den einzelnen Epochen (Draufsicht)

Die Verwendung solcher Systeme bietet den Vorteil eines einheitlichen Koordinatensystems bei der Datenerfassung – entsprechende Kalibrierung vorausgesetzt /Wagner 2016/. Da das Scannerzentrum und das Projektionszentrum der Kamera nahezu identisch sind, werden zusätzlich unerwünschte Abschattungen vermieden. Alternativ ist es aber auch möglich, die Daten mit getrennten Geräten zu erfassen. Das Zusammensetzen der Panoramabilder kann, bei entsprechend großen Überlappungsbereichen, z. B. über Feature-Matching erfolgen /Paar et al. 2005/. Die nötige Registrierung

zwischen Bild und Scandaten lässt sich auch über Passpunkte lösen oder kann auf Basis der Intensitätswerte und Geometrien, die sowohl aus den Laserscandaten als auch aus den Bildern abgeleitet werden können, bestimmt werden /Taylor & Nieto 2012/.

### 3 RGB+D-BILDER

#### 3.1 Datenerfassung

Um RGB+D-Bilder zu erstellen, ist es nötig, Kamerabilder und Laserscandaten einer Szene zu erfassen. Dies kann bei der Verwendung von modernen bildgebenden Totalstationen mit einem einzigen Gerät erfolgen. Diese Instrumente sind in der Lage, hochauflösende Bilder durch das Instrumententeleskop zu erfassen. Da das Sichtfeld durch die ca. 30-fache Vergrößerung stark begrenzt ist, müssen meist mehrere Bilder zu einem Panorama zusammengefügt werden. Die dafür nötige innere und äußere Orientierung der Kamera muss vorab in Bezug auf das Tachymetersystem kalibriert werden. Da die Kamera mit dem Instrument rotiert, bleibt die äußere Orientierung während der Messung gegenüber der Zielachse konstant und kann als Winkelablesung des Instruments ausgedrückt werden. Zusammen mit den Stationierungswerten des Instruments sind die Bilder direkt georeferenziert und orientiert. Die neuesten Instrumentengenerationen verfügen zusätzlich über eine Scanfunktion, die – wenn auch langsamer – mit denen eines Laserscanners vergleichbar ist.

#### 3.2 Tiefenbild

Die Laserscandaten werden nach der Aufnahme so weiterverarbeitet, dass sie mit den Bilddaten verschmolzen werden können. Da die Kamera bei der Aufnahme rotiert, wird für die Panoramagenerierung eine sphärische Abbildungsfläche gewählt, in der die Raumrichtungen erhalten bleiben. Demzufolge müssen auch die Laserscandaten in ein sphärisches Tiefenbild umgewandelt werden. Dazu wird jeder Einzelpunkt der Punktwolke über eine perspektivische Projektion auf das Panoramabild abgebildet. Im Falle eines getrennten Aufnahmezentrums von Laserscanner und Kamera müssen zusätzlich aus Kameransicht verdeckte Punkte herausgefiltert werden. Für diesen Vorgang, der auch als Z-Buffer-Problem bekannt ist, existieren unterschiedliche Ansätze, siehe z. B. /Qin & Grün 2014/.

Das Ergebnis der Projektion der Punktwolke auf das sphärische Panorama ist ein unregelmäßiges Gitter aus 2D-Punkten und deren Abstand zum Projektionszentrum der Kamera. Das eigentliche Tiefenbild wird durch die Interpolation der Punkte auf ein regelmäßiges Punktgitter in der Auflösung des Kamerabilds erzeugt; die Pixelwerte codieren dabei die Objektdistanz. Das Tiefenbild bildet so den D-Kanal (engl. depth-channel) des RGB+D-Bilds. Die drei anderen Kanäle repräsentieren die Farbwerte der Kameraaufnahmen: Rot, Grün und Blau (*Abb. 3*).

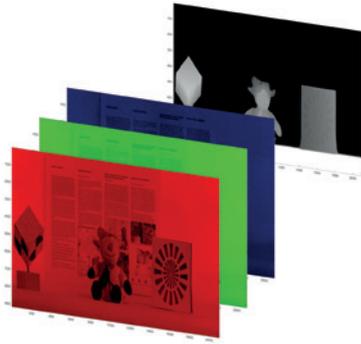


Abb. 3 | Die vier Kanäle eines RGB+D-Bilds

### 3.3 Eigenschaften des RGB+D-Bilds

Hauptvorteil dieser Darstellungsart ist, dass jedem Pixel eine Raumrichtung zugeordnet ist. Mit der zusätzlichen Tiefeninformation lassen sich dadurch aus dem Bild direkt sphärische Koordinaten abgreifen, welche einfach in ein kartesisches System zu überführen sind.

Allerdings muss hier angemerkt werden, dass die Genauigkeit der Tiefeninformation stark mit der Abtastauflösung auf der Objektoberfläche und der gewählten Interpolationsmethode korreliert ist. Bei der Aufnahme mit hochauflösenden Kameras oder zusammengesetzten Panoramas aus mehreren Einzelbildern ist die RGB-Bildauflösung meist deutlich höher als die Auflösung der Distanzinformation. Im angegebenen Versuchsaufbau ist beispielsweise der mittlere Punktabstand des Laserscans am bewegten Objekt mit 0,44 mm ca. 5-fach höher als der Abstand zweier Pixel (0,085 mm) im Bild. Unter der (pessimistischen) Annahme einer pixelgenauen Feature-Detektion kann die reine Translation quer zur Aufnahme- richtung besser als 1/10 mm aufgelöst und bestimmt werden. In Aufnahme- richtung ist das Ergebnis einem Messrauschen von 0,5 mm unterworfen und nur ca. jedes fünfte Pixel beinhaltet eine tatsächliche Distanzmessung. Dazwischen werden die Entfernungswerte interpoliert. Für die 3D-Verschiebungsvektoren im konkreten Fall kann somit eine Präzision von  $< 1$  mm angegeben werden. Die absolute Messgenauigkeit wird von der Distanzmessgenauigkeit des Laserscanners dominiert. Das Verhältnis der Genauigkeiten in Quer- und Längsrichtung dreht sich mit zunehmender Distanz um, wie dies auch bei klassischen Tachymetermessungen der Fall ist.

## 4 DEFORMATIONSANALYSE

Der RGB+D-Ansatz kombiniert Laserscan- und Bilddaten. Vergleicht man die Sensitivitäten der beiden Datensätze für Objektbewegungen, zeigen diese ein komplementäres Verhalten. Distanzvariationen unter einer bestimmten Raumrichtung lassen sich in Laserscandaten einfach ermitteln, wohingegen Bewegungen senkrecht zur Aufnahme- richtung nur schwer (bis gar nicht) aufzudecken sind. Bei der Verfolgung von Objekten in (Mono-)Kamerabildern mittels Bildverarbeitungs- algorithmen zeigt sich ein genau gegensätzliches Verhalten. Hier können Bewegungen senkrecht zur Blickrichtung einfach erfasst werden. Geringe Distanzänderungen, die lediglich als kleine Änderungen in der Skalierung von bewegten Objekten wiedergegeben werden, können hingegen nur schwer detektiert

werden. Durch eine kombinierte Auswertung wird der Nachteil des einen Verfahrens durch den Vorteil des anderen Verfahrens ausgeglichen.

Mithilfe identischer Punkte in RGB+D-Bildern aufeinanderfolgender Epochen können 3D-Verschiebungsvektoren berechnet werden. Wie bereits beschrieben, repräsentieren die einzelnen Bildpunkte direkt sphärische Koordinaten, die in Welt-/Objekt-Koordinaten umgerechnet werden können. Zwei identische Punkte aus unterschiedlichen Epochen bilden somit den Start- und Endpunkt eines 3D-Verschiebungsvektors. Identische Bildpunkte können auch in RGB+D-Bildern unterschiedlicher Standpunkte einer Epoche gefunden werden, wodurch eine Überbestimmung der Einzelpunkte vorliegt. In Verbindung mit einer Netzmessung können über eine freie Netzausgleichung Genauigkeitsmaße für die Objektpunkte ermittelt werden. Liegen mehrere Messepochen vor und enthält das Netz zusätzlich stabile Stützpunkte, können die frei ausgeglichenen Netze unter den Voraussetzungen gleicher Lagerung, identischen Datumsdefekten und der Verwendung identischer Näherungskordinaten einer strengen Deformationsanalyse im Sinne eines Kongruenzmodells zugeführt werden. Durch einen zweistufigen Netzaufbau lassen sich stabile von bewegten Punkten trennen und die Signifikanz der ermittelten Verschiebungsvektoren statistisch belegen.

### 4.1 Merkmalsbasierte Beschreibung von Bildpunkten

Um korrespondierende Punkte oder Regionen in unterschiedlichen Bildern derselben Szene zu finden, können verschiedene Algorithmen der Bildverarbeitung oder der Computer Vision zum Einsatz kommen, die auf der Verteilung der Pixelwerte in den Bildern basieren. Verfahren, wie die normalisierte Kreuzkorrelation oder das Least Squares Matching, verwenden dabei direkt die Pixelwerte eines kleinen Bildausschnitts; andere Verfahren versuchen Merkmale der lokalen Umgebung eines Untersuchungspunkts numerisch zu beschreiben und fassen diese in sogenannten Feature- oder Merkmalsvektoren zusammen.

Einer der weitverbreitetsten Algorithmen zur merkmalsbasierten Beschreibung von Bildpunkten ist der SIFT-Algorithmus (Scale Invariant Feature Transform) /Lowe 2004/. Er beruht auf der Berechnung von dominierenden Gradientenrichtungen in der lokalen Umgebung eines Untersuchungspunkts in unterschiedlichen Auflösungsstufen. Diese Werte werden zu einem gewichteten Gradienten-Histogramm zusammengefasst und bilden einen 128 Werte langen Feature-Vektor. Hauptmerkmal des SIFT-Algorithmus ist seine Invarianz gegenüber Skalierungs- und Orientierungsänderungen des Untersuchungspunkts. Gegenüber anderen Verfahren zeichnet sich der SIFT-Algorithmus des Weiteren durch seine Robustheit gegenüber Beleuchtungsänderungen, Rauschen und affinen Verzerrungen der Abbildung aus /Lowe 2004/.

Neuere Methoden zur merkmalsbasierten Beschreibung von Bildpunkten versuchen verstärkt den Berechnungsaufwand zu optimieren. Dabei erreichen sie ähnliche Genauigkeiten und Detektierungs- raten wie die klassischen Ansätze. Bei den neuen Algorithmen wird der Feature-Vektor aus 1-bit-Werten gebildet, die wesentlich effizienter zu verarbeiten sind und so Rechenzeit einsparen. Ein

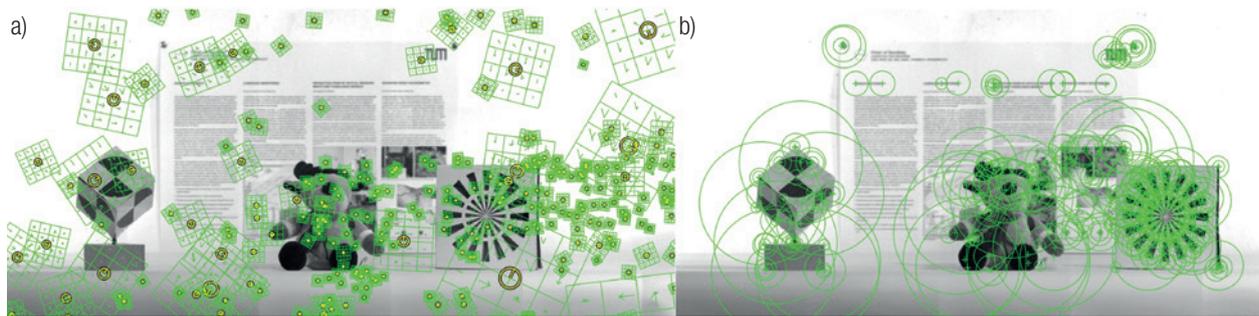


Abb. 4 | Teilmenge der berechneten Features in der Testszene: a) SIFT, b) BRISK

Repräsentant aus dieser Klasse ist der BRISK-Algorithmus (Binary Robust Invariant Scalable Keypoints) /Leutenegger et al. 2011/, der aus drei verschiedenen Teilen besteht: dem Abtastmuster, der Bestimmung der Orientierung sowie der Paarung der verwendeten Abtastpunkte. Das Abtastmuster enthält Untersuchungspunkte auf unterschiedlich skalierten konzentrischen Kreisen, zwischen denen paarweise die lokalen Gradienten berechnet werden. Gradienten weit auseinanderliegender Abtastpunkte werden verwendet, um die Orientierung des Features zu bestimmen. Die Beschreibung des eigentlichen Feature-Vektors erfolgt über die Gradienten nahe beieinanderliegender Abtastpunkte.

Abb. 4 zeigt die berechneten SIFT- und BRISK-Features für den Versuchsaufbau. Zur besseren Visualisierung wird nur eine (zufällige) Teilmenge der detektierten und beschriebenen Merkmale dargestellt. Das Zentrum der grünen Quadrate bzw. der Mittelpunkt der Kreise markieren die Lage des untersuchten Bildpunkts; die Ausdehnung gibt den Bildbereich an, über den der Beschreibungsvektor berechnet wurde. Die vorgestellten Algorithmen beschreiben

nicht jeden Bildpunkt mit einem Merkmalsvektor, sondern ermitteln besondere Punkte/Regionen, die in ihrer lokalen Nachbarschaft unverwechselbar sind. Wie die zwei Beispiele in Abb. 4 verdeutlichen, liefern die verschiedenen Feature-Algorithmen unterschiedliche Bildpunkte (Zentren der Quadrate/Kreise), deren Verteilung sich stark unterscheidet. Je nach Art der zu untersuchenden Szene muss der geeignetste Algorithmus ausgewählt und eventuell dessen Parameter auf die aktuellen Gegebenheiten angepasst bzw. müssen verschiedene Algorithmen kombiniert werden.

## 4.2 Zuordnung

Um identische Punkte in zwei Messepochen zu finden, werden die oben beschriebenen Feature-Vektoren genutzt. Die Suche kann so auf einer abstrakten numerischen Ebene durchgeführt werden, auf der lediglich die Ähnlichkeit der Feature-Vektoren verglichen wird. Im Detail bedeutet dies, dass für jedes Bildmerkmal der ersten



Abb. 5 | Ergebnis des Feature-Matchings zwischen zwei Epochen

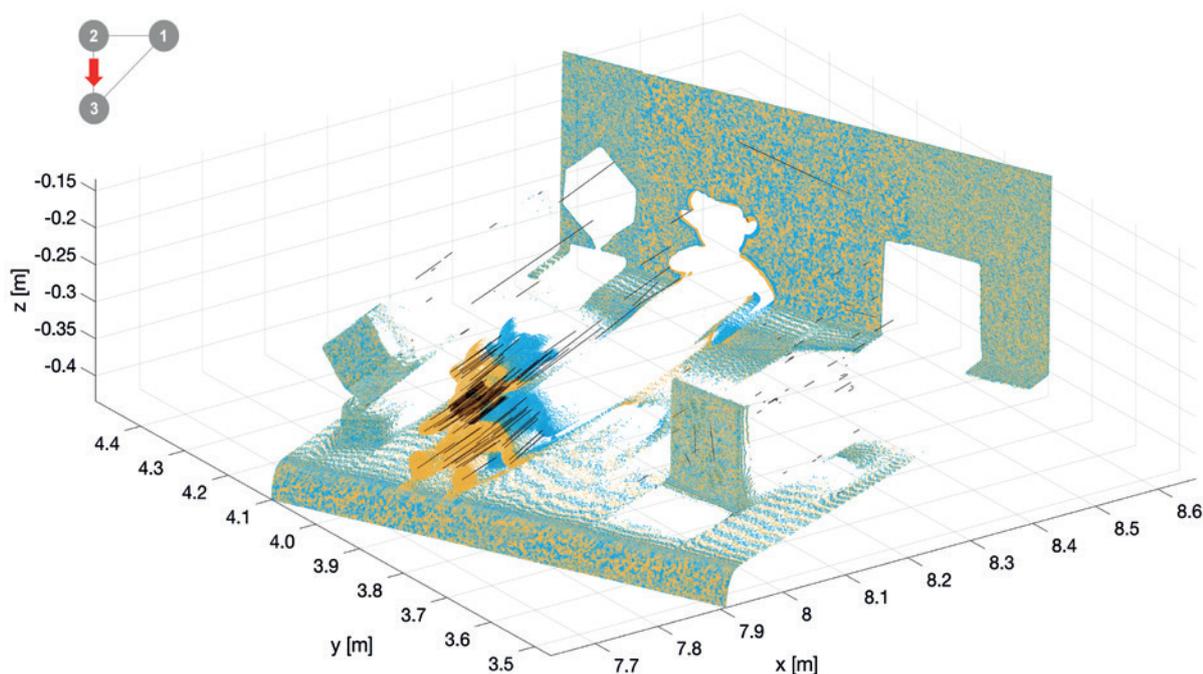


Abb. 6 | 3D-Ansicht des Feature-Matchings aus Epoche 2 (blau) und Epoche 3 (gelb), 3D-Verschiebungsvektoren in Schwarz

Epoche das passendste Feature aus der zweiten Epoche gefunden werden muss. Diese Zuordnung (Feature-Matching) kann bei SIFT-Merkmalen z.B. über die Summe der quadrierten Differenzen zwischen den Feature-Vektoren erfolgen. Beim Vergleich der Binär-Vektoren bietet sich die Berechnung der Hamming-Distanz /Hamming 1950/ an, die über einen XOR-Operator (Exklusiv-ODER) sehr effizient zu berechnen ist.

Abb. 5 zeigt die Ergebnisse des Feature-Matchings (SIFT) für die Messepochen 2 und 3 (Bewegung in Blickrichtung des Sensors). Die Mittelpunkte der roten Kreise markieren die Position der Features in Epoche 2, die grünen Kreuze die Position in Epoche 3. Korrespondierende Punkte sind mit einer gelben Linie verbunden, welche die Verschiebung der Features in der Bildebene repräsentiert. Da die erfassten Feature-Punkte im RGB+D-Bild direkt mit 3D-Koordinaten verknüpft sind, können die Deformationen als 3D-Raumvektoren aus den Pixelkoordinaten abgeleitet werden. In Abb. 6 sind dieselben Epochen 2 und 3 des Versuchs als 3D-Ansicht dargestellt. Die Punktwolke der Epoche 2 ist blau, die Punktwolke der Epoche 3 gelb eingefärbt; die 3D-Verschiebungsvektoren werden als schwarze Linien gezeichnet. Die sichtbaren Ausreißer – inkl. der falschen Zuordnungen an Kanten – können leicht detektiert und entfernt werden, indem die Orientierung und Länge der Vektoren mit benachbarten Elementen verglichen werden. Dabei liegt die Annahme zugrunde, dass Veränderungen in lokal begrenzten Bereichen eine annähernd gleiche Größenordnung und Richtung besitzen.

### 4.3 Vergleich mit herkömmlichen Auswertemethoden

Die berechneten 3D-Verschiebungsvektoren der RGB+D-Auswertung sind für Nutzer einfach zu interpretieren und spiegeln die tat-

sächlichen Objektbewegungen wider. Zusätzlich können sie einfach auf Signifikanz getestet werden. Durch die Kombination von Bild und Scandaten können so Deformationen abgeleitet werden, die bei Verwendung von nur einem Datentyp (Bild oder Scan) nicht eindeutig bestimmbar sind.

Im dargestellten Fall ist aus den reinen Bilddaten (Abb. 5) keine starke Bewegung zu erkennen, obwohl das mittlere Objekt (Stofftier) deutlich auf den Betrachter zubewegt wurde. Die meisten korrespondierenden Feature-Punkte fallen in den Bildern zusammen und könnten auch als ein reines Kippen des Objekts interpretiert werden. Die tatsächliche Objektbewegung/Deformation kann nicht rein aus den Bilddaten abgeleitet werden.

Die Deformationsanalyse aus reinen Scandaten zeigt in der hier dargestellten Szene ebenfalls nur unzureichende Informationen über die tatsächliche Verschiebung und lässt großen Spielraum für Interpretationen. Verwendet man beispielsweise nicht die Bilddaten, sondern nur die Tiefenkanäle der RGB+D-Bilder, so lassen sich Distanzänderungen zwischen zwei Messepochen einfach ermitteln. Hierfür müssen lediglich die Tiefenkanäle pixelweiße voneinander abgezogen werden. Abb. 7 zeigt die Distanzänderungen zwischen den unterschiedlichen Objektpositionen. Bei einer Bewegung senkrecht zur Blickrichtung des Scanners (Vergleich Epoche 1 und 2, Abb. 7a) ergeben sich zwei große Änderungsbereiche. An der Stelle, an der das Objekt verschwindet, erhöhen sich die Distanzen. An der Position, in der das Objekt in der Folgepoche zu liegen kommt, verkürzen sich die Distanzen. Einzig über die Objektkonturen können Rückschlüsse auf die tatsächliche Objektbewegung gezogen werden. Sind die Objektränder – anders als in diesem Beispiel – nicht scharf vom Hintergrund abgehoben, kann die Szene nicht eindeutig interpretiert werden. An Stellen, an denen die Objektpositionen zwischen den Epochen überlappen, entstehen zusätzlich Bereiche, in denen keine Aussage über eine Deformation getroffen

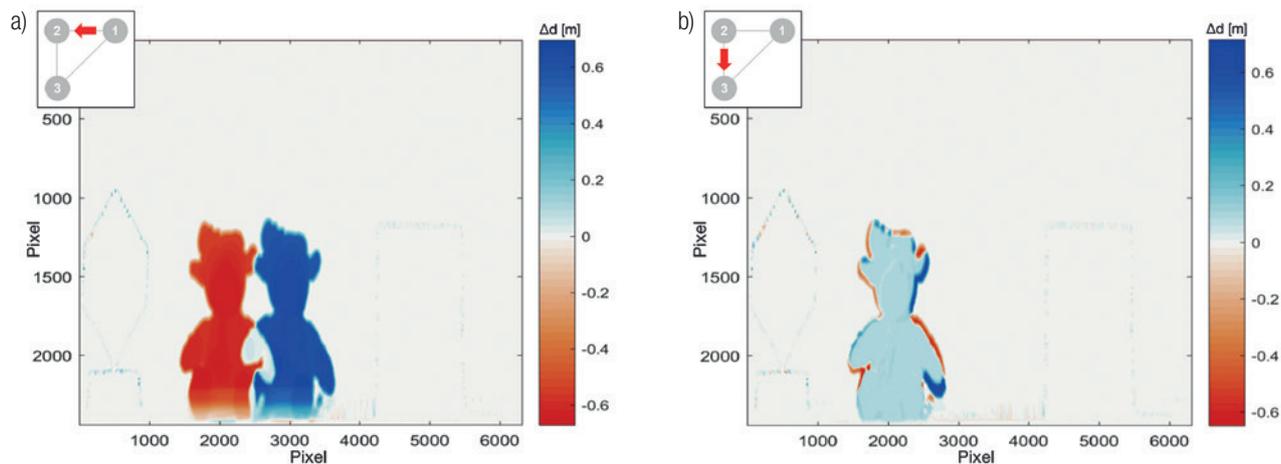


Abb. 7 | Differenzen [m] aus den Tiefenbildern: a) Objektposition 1 & 2, b) Objektposition 2 & 3

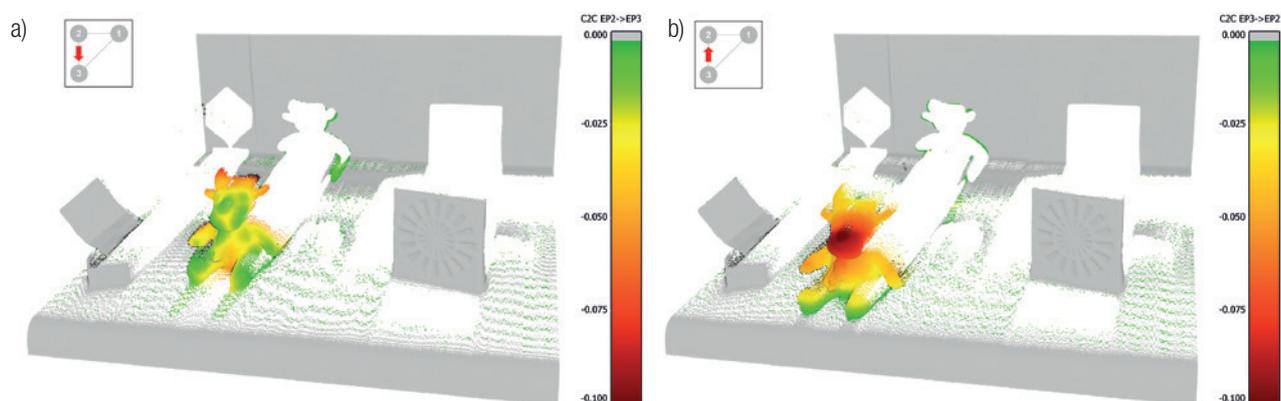


Abb. 8 | Cloud-to-Cloud-Distanzen [m] zwischen Epoche 2 & 3: a) Referenz-Epoche 2, b) Referenz-Epoche 3

werden kann. Bewegt sich das Objekt auf den Betrachter zu (Vergleich Epoche 2 und 3, *Abb. 7b*), entstehen deutlich sichtbare Artefakte an den Rändern des Objekts, aus denen ebenfalls nicht auf die Objektbewegung rückgeschlossen werden kann.

Eine weitere Auswertestrategie bei ausschließlicher Verwendung der Laserscandaten ist der reine Distanzvergleich zwischen Punktwolken (Cloud-to-Cloud) unterschiedlicher Epochen, wie in *Abb. 8* dargestellt. Dabei wird für jeden Punkt aus der Punktwolke einer Referenzeпоche der nächste Nachbarpunkt aus einer Vergleichsepoche gesucht und die Raumdistanz zwischen beiden berechnet. Die Abstände werden farblich codiert in der Referenzpunktwolke dargestellt. In *Abb. 8a* wurde Epoche 2 als Referenz gewählt, in *Abb. 8b* Epoche 3. Es ist deutlich zu erkennen, dass bereits die Berechnungsrichtung (Epoche 2 zu Epoche 3 bzw. Epoche 3 zu Epoche 2) deutliche Unterschiede zeigt. Diese ergeben sich aus Zuordnungsproblemen zwischen den beiden Punktwolken, da lediglich nach dem geometrisch nächsten Nachbarn gesucht werden kann; eine Suche nach identischen Punkten findet nicht statt. Es weichen sowohl die ermittelten Beträge der Deformation als auch die scheinbar deformierten Bereiche stark voneinander ab. Eine sinnvolle Interpretation der Objektbewegung ist nur schwer bis gar nicht möglich.

Fortgeschrittenere Methoden, wie z.B. der M3C2-Algorithmus /Lague et al. 2013/, die zur Bestimmung der Abweichungen in Punkt-

wolken zusätzlich die Normalenrichtung für die Nachbarschaftssuche und Abstandsberechnung heranziehen, liefern ebenfalls keine zufriedenstellenden Ergebnisse. Aussagekräftige Ergebnisse können nur durch starke manuelle Eingriffe in den Auswerteprozess erlangt werden. Weitere Auswertemethoden verwenden das ICP-Verfahren, um Deformationen zwischen Punktwolken abzuleiten. Dabei wird die Punktwolke in kleine Ausschnitte (Patches) zerlegt, die anschließend so transformiert werden, dass der Abstand zur Punktwolke der Folgeepoche minimal wird. Die mit dem ICP-Verfahren geschätzten Transformationsparameter liefern Informationen für eine Starrkörperbewegung. Das Verfahren eignet sich nicht, um Verformungen aufzudecken, und ist stark abhängig von der Einteilung der Patches /Chmelina et al. 2012/, /Wujanz et al. 2014/.

## 5 ZUSAMMENFASSUNG UND AUSBLICK

Der vorgestellte Ansatz zur Deformationsanalyse über fusionierte Bild- und Laserscandaten vereint die Vorteile beider Methoden. Mit dichten Laserscanpunktwolken werden Distanzänderungen in Sichrichtung erfasst; hochauflösende Bilder zeigen dagegen hohe Sensitivität für Verschiebungen senkrecht zur Aufnahmeichtung.

In den kombinierten RGB+D-Bildern werden durch die Extraktion von Bildmerkmalen identische Punkte ermittelt, die sich an

Hand ihres Beschreibungsvektors über mehrere Epochen verfolgen lassen. Da die Pixelkoordinaten aus den Farbbildern in Kombination mit dem Tiefenkanal direkt 3D-Koordinaten repräsentieren, können aus identischen Bildpunkten zwischen den Epochen 3D-Verschiebungsvektoren abgeleitet werden. Diese lassen sich auf Signifikanz testen und sind für den Nutzer einfach zu interpretieren. Mit der vorgestellten Methode können die Daten unterschiedlicher Sensoren – wie z.B. moderner Totalstationen, Mobile-Mapping-Systemen, UAV und Roboter-Plattformen – ausgewertet werden, sofern Bild- und Scandaten erhoben werden. Durch die gemeinsame Auswertung in RGB+D-Bildern lässt sich das volle Potenzial dieser Daten für die Deformationsanalyse nutzen. Mögliche Anwendungen sind Deformationsbereiche, die eine dichte, flächenhafte und berührungslose Erfassung erfordern, wie beispielsweise die Erfassung von Verformungen einer Tunnelschale oder Ortsbrust, überlagerte Setzungen (z.B. im Straßenbereich), die Überwachung von (Ingenieur-)Bauwerken, wie Häuserfassaden, Staudämme oder Schleusentore. Auch ein Einsatz in der Industrievermessung ist denkbar, solange die Oberfläche des Objekts genug Struktur für die Bildauswertung aufweist. Im Vergleich zu herkömmlichen Methoden zur Ableitung von Deformationen in Punktwolken bieten die gewonnenen 3D-Verschiebungsvektoren eine einfach interpretierbare Darstellung der Objektbewegung/Deformation. Scheindeformationen durch fehlerhafte Punktzuordnung bei der Abstandsberechnung werden vermieden.

Für die weitere Verbesserung des vorgestellten Ansatzes wird versucht, zusätzliche Feature-Typen zur Bestimmung identischer Punkte zu entwickeln. Hier können 3D-Features, die aus der Punktwolke gewonnen werden, die Auswertung stützen. Diese Features beschreiben – ähnlich wie bei der Extraktion der Bildmerkmale – die räumliche Verteilung der Nachbarkpunkte eines Untersuchungspunkts. Erste Ansätze hierfür sind mit 3D-SIFT /Scovanner et al. 2007/ und den Point-Feature-Histogrammen /Ruso 2009/ bereits in der Literatur beschrieben. Idealerweise resultieren daraus kombinierte Features, in die sowohl Bild- als auch Laserscandaten einfließen. Zusätzlich können die Intensitätswerte des Laserscanners als weiterer Kanal für die Analyse herangezogen werden.

## LITERATUR

- Chmelina, K.; Jansa, J.; Hesina, G.; Traxler, C. (2012): A 3-d laser scanning system and scan data processing method for the monitoring of tunnel deformations. In: *Journal of Applied Geodesy*, 6(2012)3-4, 177–185.
- Hamming, R. W. (1950): Error Detecting and Error Correcting Codes. In: *Bell System Technical Journal*, 29(1950)2, 147–160.
- Lague, D.; Brodu, N.; Leroux, J. (2013): Accurate 3D comparison of complex topography with terrestrial laser scanner. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 82(2013), 10–26.
- Leica Geosystems (2016): *Leica MS60/TS60 Gebrauchsanweisung*, Version 2.0. Heerbrugg, Schweiz.
- Leutenegger, S.; Chli, M.; Siegwart, R. Y. (2011): BRISK: Binary Robust Invariant Scalable Keypoints. In: *IEEE International Conference on Computer Vision (ICCV)*, November 6–13, 2011, Barcelona, Spanien, 2548–2555.
- Lowe, D. G. (2004): Distinctive image features from scale-invariant keypoints. In: *International Journal of Computer Vision*, 60(2004)2, 91–110.

Paar, G.; Bauer, A.; Kontrus, H. (2005): Texture-Based Fusion Between Laser Scanner and Camera for Tunnel Surface Documentation. In: *Proceedings of 7th International Conference on Optical 3-D Measurement Techniques*, Oktober 3–5, 2005, Wien, Österreich, 74–80.

Qin, R.; Grün, A. (2014): 3D change detection at street level using mobile laser scanning point clouds and terrestrial images. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, 90(2014), 23–35.

Ruso, R. B. (2009): *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. Dissertation, Fakultät für Informatik, Technische Universität München.

Scovanner, P.; Ali, S.; Shah, M. (2007): A 3-dimensional sift descriptor and its application to action recognition. In: *Proceedings of the 15th ACM International Conference on Multimedia*, September 23–28, 2007, Augsburg, 357–360.

Taylor, Z.; Nieto, J. (2012): A mutual information approach to automatic calibration of camera and lidar in natural environments. In: *Proceedings of Australasian Conference on Robotics and Automation*, Dezember 3–5, 2012, Wellington, Neuseeland, 3–5.

Wagner, A. (2016): A new approach for geo-monitoring using modern total stations and RGB+D images. In: *Measurement*, 82(2016), 64–74.

Wujanz, D.; Neitzel, F.; Krueger, D.; Manthe, C. (2014): DefoScan++: Oberflächenbasierte Registrierung terrestrischer Laserscans für die Deformationsanalyse. In: *allgemeine vermessungs-nachrichten (avn)*, 121(2014)1, 13–21.

Wunderlich, T.; Niemeier, W.; Wujanz, D.; Holst, C.; Neitzel, F.; Kuhlmann, H. (2016): Areal Deformation Analysis from TLS Point Clouds – the Challenge. In: *allgemeine vermessungs-nachrichten (avn)*, 123(2016)11-12, 340–351.

**M. Sc. Wolfgang Wiedemann**

TECHNISCHE UNIVERSITÄT MÜNCHEN  
LEHRSTUHL FÜR GEODÄSIE

Arcisstraße 21 | 80333 München  
w.wiedemann@tum.de



**Dr.-Ing. Andreas Wagner**

TECHNISCHE UNIVERSITÄT MÜNCHEN  
LEHRSTUHL FÜR GEODÄSIE

Arcisstraße 21 | 80333 München  
a.wagner@tum.de



**Prof. Dr.-Ing. habil. Thomas Wunderlich**

TECHNISCHE UNIVERSITÄT MÜNCHEN  
LEHRSTUHL FÜR GEODÄSIE

Arcisstraße 21 | 80333 München  
th.wunderlich@tum.de



Manuskript eingereicht: 11.07.2017 | Im Peer-Review-Verfahren begutachtet