

Gig3D – Eine Milliarde 3D-Punkte in einer Sekunde

Ralph SCHMIDT und Bernd-Michael WOLF

Zusammenfassung

Die Nahbereichsgruppe des Instituts für Photogrammetrie und GeoInformation der Leibniz Universität Hannover und SOLVing3D entwickeln im Rahmen des Vorhabens Gig3D ein Messsystem, mit dem in einer Sekunde bis zu einer Milliarde 3D-Punkte aufgezeichnet werden können. Zum Einsatz kommen hierbei modernste Hochgeschwindigkeitskameras und aktuelle Verfahren aus dem Bereich der stereoskopischen Oberflächenrekonstruktion bzw. der photogrammetrischen Bildzuordnung.

In diesem Beitrag wird anhand eines Beispiels ein lokales Zuordnungsverfahren mit dem Verfahren Semi-Global Matching verglichen. Zur Unterstützung der Oberflächenrekonstruktion in texturarmen Bereichen des Testobjektes wird ein stochastisches Muster auf die Oberfläche projiziert. Die damit erzeugten Ergebnisse werden mit Daten verglichen, die ohne projizierte Muster erzielt worden sind, denn in verschiedenen Anwendungsfällen ist die Projektion von Mustern nicht möglich. Das Anwendungsszenario erlaubt es aber, die Auswertung nach der Aufnahme durchzuführen; hier kann also der Fokus auf Vollständigkeit und höchste Präzision gelegt werden. Die Antastgenauigkeiten des hierfür genutzten Kamerasystems in Normalfallanordnung werden in Anlehnung an die VDI/VDE 2634 geprüft und mit den Längenabweichungen eines Systems mit identischen Kameras und der in der optischen Messtechnik üblichen Konvergenzanordnung verglichen.

1 Einleitung

Die automatische dreidimensionale Objektrekonstruktion statischer Szenen ist seit über drei Jahrzehnten ein wichtiger Forschungszweig in der Photogrammetrie und im Computer Vision. Im industriellen Bereich kommen heutzutage hauptsächlich zwei unterschiedliche Methoden zur Anwendung: 1. Strukturierte Beleuchtung mit Laserlinien oder einem zweidimensionalen Projektor und ein oder zwei Kameras 2. Time-of-Flight-Sensoren (TOF) in Laserentfernungsmessern oder TOF-Kameras. Während die TOF-Methoden die Tiefenwerte bzw. Distanzen des Objektes zum Sensor direkt über Laufzeit des Signals messen, werden bei den Ansätzen mit strukturierter Beleuchtung die Tiefenwerte über Triangulation bestimmt. Bei zeitkodierte Mustern wie Gray Code oder zeitvariierten Mustern wie Laserlinien wird eine eindeutige Zuordnung korrespondierender Pixel angestrebt (BATLLE et al. 1998, PAGÈS et al. 2003). Solche Verfahren benötigen für eine komplette pixelgenaue Rekonstruktion zahlreiche aufeinanderfolgende Messungen. Daher eignen sich diese Methoden nicht zur Messung von dynamischen Szenen, bei denen sich Form und/oder Position des Objektes schnell im Verhältnis zur benötigten Messzeit ändert. In solchen Fällen können prinzipiell nur zweidimensionale Sensoren dreidimensionale Deformationen und Bewegungen von Objekten erfassen.

In diesem Beitrag wird das in der Entwicklung befindliche Messsystem Gig3D vorgestellt, welches mit einem klassischen photogrammetrischen Aufbau bestehend aus zwei Hochgeschwindigkeitskameras in Verbindung mit automatischen Verfahren der digitalen Bildzu-

ordnung die Vermessung hochdynamischer Vorgänge erlaubt. Im nächsten Abschnitt werden die Zuordnungsverfahren vorgestellt, die bisher untersucht worden sind. In Abschnitt drei werden erste Ergebnisse anhand eines Beispiels gezeigt. Hierbei werden die Resultate eines üblichen lokalen fensterbasierten Zuordnungsansatzes denen von Semi-Global Matching gegenübergestellt. Da bei Hochgeschwindigkeitsaufnahmen möglicherweise kein ausreichend lichtstarker Projektor zur Verfügung steht, oder man aus anderen Gründen auf künstliche Muster verzichten muss, wird auch das Potenzial von Messungen ohne Muster demonstriert. In Abschnitt vier wird die Antastgenauigkeit des vorgestellten Messsystems gemäß Blatt 1 VDI/VDE 2634 (VDI/VDE 2634 2002) geprüft.

2 Zuordnungsverfahren

Traditionelle Zuordnungsverfahren identifizieren korrespondierende Punkte über die Analyse der Ähnlichkeit von Bildausschnitten, wobei typischerweise ein Referenzfenster (Mustermatrix) um das zu messende Pixel mit einem Suchfenster (Suchmatrix) über einen gewissen Suchbereich verglichen wird. Dabei können künstliche Muster auf die Szene projiziert werden, um auch in texturarmen Bereichen eine eindeutige Zuordnung zu ermöglichen (KANG et al. 1995). Dieses Verfahren kann auch für dynamische Szenen angewandt werden, indem jedes Frame-Paar unabhängig voneinander rekonstruiert wird. Bei einem kalibrierten Messsystem bestehend aus zwei Kameras kann der Suchbereich auf die Epipolarlinie beschränkt werden, die in entzerrten Epipolarbildern mit korrespondierenden Bildzeilen zusammenfällt. Der Suchbereich kann weiter auf den zu erwarteten Tiefenbereich beschränkt werden, so dass nicht die komplette Bildzeile nach dem korrespondierenden Pixel abgesucht werden muss. Ein sehr einfaches Ähnlichkeitsmaß ist die Summe der Absoluten Differenzen (SAD):

$$SAD = \sum_{p \in N} |I_r(p) - I_l(p)| \quad (1)$$

wobei I_r das rechte Bild, I_l das linke Bild, p die Pixel des Bildes und N das Nachbarschaftsfenster ist, über das die Matching-Kosten SAD berechnet werden. Zahlreiche andere Ähnlichkeitsmaße wurden entwickelt, welche hauptsächlich radiometrische Unterschiede zwischen den beiden Bildern berücksichtigen (HIRSCHMÜLLER & SCHARSTEIN 2009). Obwohl die Bilder bei unseren Experimenten unter kontrollierten Beleuchtungsbedingungen aufgenommen wurden, haben sich auch der Normalisierte Kreuzkorrelationskoeffizient (NCC) und Mutual Information (MI) als vorteilhaft erwiesen. Beim NCC (MORAVEC 1980) können durch die Normalisierung radiometrische Unterschiede bezüglich Helligkeit und Kontrast kompensiert werden:

$$NCC = \frac{\sum_{p \in N} (I_l(p) - \bar{I}_l(p))(I_r(p) - \bar{I}_r(p))}{\sqrt{\sum_{p \in N} (I_l(p) - \bar{I}_l(p))^2 \sum_{p \in N} (I_r(p) - \bar{I}_r(p))^2}} \quad (2)$$

Hierbei ist mit dem Überstrich der jeweilige Mittelwert des Fensters gekennzeichnet. Mutual Information hingegen misst die statistische Abhängigkeit zweier Datensätze (VIOLA & WELLS 1997, MAES et al. 1997). Hierbei erfolgt die Zuordnung über die Bestimmung der gegenseitigen Wahrscheinlichkeitsdichte (Joint Probability) zwischen zwei Bildern, was der Berechnung von Wahrscheinlichkeiten korrespondierender Intensitätswerte in einem Histogramm entspricht. Zur Schätzung der Wahrscheinlichkeitsdichte und der eigentlichen Maximierung der MI gibt es verschiedene Varianten. Aufgrund der Invarianz bezüglich globaler radiometrischer Unterschiede eignet sich dieses Maß insbesondere für die Registrierung multimodaler Aufnahmen wie sie z. B. in der Medizin vorkommen.

Als endgültiges korrespondierendes Pixel wird dasjenige mit den geringsten aggregierten Kosten bzw. dem höchsten NCC gewählt; ein Vorgehen, das auch als „Winner-Take-All-Strategie“ bezeichnet wird. Allerdings weisen diese lokalen Zuordnungsverfahren den Nachteil auf, dass steile Disparitätssprünge nicht exakt modelliert und Kanten verschmiert werden. Dieser Effekt kann durch die Wahl kleiner Pixelfenster reduziert werden. Texturarme Bereiche erfordern dagegen möglichst große Pixelfenster, so dass bei der Matching-Strategie meistens ein Kompromiss in der Wahl der Fenstergröße erforderlich ist. Globale Matching-Verfahren wie Dynamic Programming (BIRCHFIELD & TOMASI 1999), Semi-Global Matching (HIRSCHMÜLLER 2008) oder Graph Cuts (BOYKOV et al. 2001) vermeiden dieses Problem, indem sie die Matching-Kosten pixelweise berechnen und das Disparitätsfeld über eine zu minimierende Kostenfunktion bestimmen (TERZOPOULOS 1986):

$$E(p) = E_{Data}(p) + \lambda E_{Smooth}(p) \quad (3)$$

wobei $E_{Data}(p)$ die globale Summe der Kosten angibt, $E_{Smooth}(p)$ die Glattheitsannahmen zwischen benachbarten Pixeln formuliert und λ den Einfluss des Glattheitsterms reguliert. Eine aktuelle Evaluation verschiedener Stereoverfahren findet sich in (SCHARSTEIN & SZELISKI 2002). Die einzelnen globalen Verfahren unterscheiden sich im Wesentlichen durch die Energiefunktion (3) und das Verfahren ihrer Minimierung. Beim Dynamic Programming (DP) wird die Energiefunktion über jede Bildzeile separat minimiert. Durch die isolierte Betrachtung einzelner Zeilenpaare werden allerdings die Beziehungen zu den umliegenden Zeilen vernachlässigt, und es ergeben sich streifige Artefakte im Disparitätsbild. Diesen Nachteil vermeiden Algorithmen wie Graph Cuts (GC), indem sie die Kostenfunktion über das gesamte Bild minimieren. Der Rechenaufwand steigt wegen der komplexen Möglichkeiten hingegen enorm an, so dass insbesondere die Auswertung von Bildsequenzen dynamischer Szenen sehr lange Rechenzeiten bedeuten würden. Als Kompromiss kann das Semi-Global Matching (SGM) angesehen werden, bei dem die Kosten nicht nur entlang der Bildzeile, sondern auch zu 8 oder 16 Richtungen quer dazu minimiert werden. Dadurch wird der Rechenaufwand stark reduziert und die Streifenartefakte vermieden. In dieser Untersuchung betrachten wir einen lokalen fensterbasierten Ansatz und SGM.

3 Ergebnisse

In diesem Abschnitt wird anhand eines Beispieldatensatzes die lokale Methode mit dem SGM verglichen. Dabei wird auch untersucht, ob ohne projiziertes Muster vollständige und

korrekte Rekonstruktionen der statischen Szene erreicht werden können. In Abb. 1 ist das rechte Bild des untersuchten Datensatzes mit und ohne projiziertes Muster abgebildet.

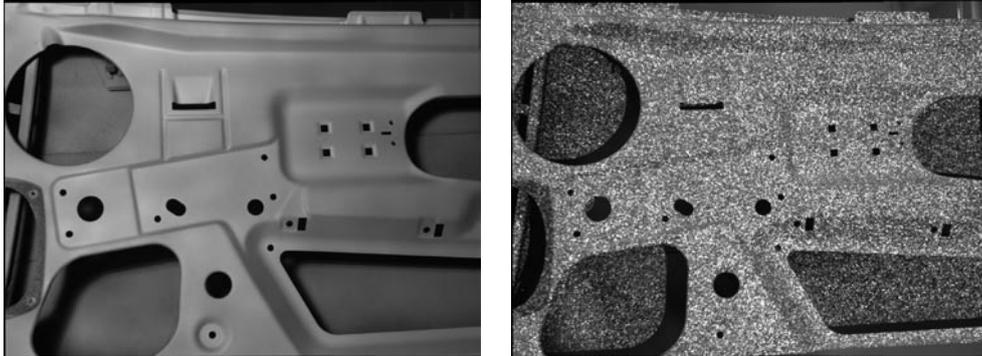


Abb. 1: Datensatz „Autotür“ links ohne und rechts mit projizierten Muster

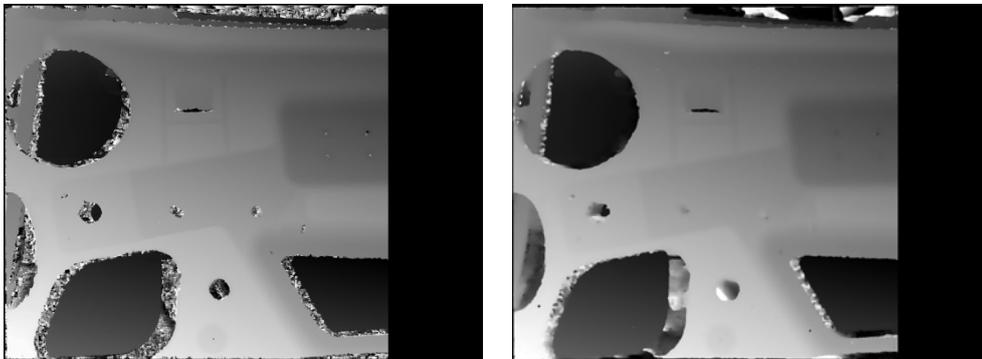


Abb. 2: Ergebnisse mit Textur: NCC (links) und SGM (rechts)

In einem ersten Test werden die Ergebnisse NCC und SGM mit projizierten Muster miteinander verglichen. In Abb. 2 sind die entsprechenden Disparitätsbilder dargestellt, wobei zur besseren Visualisierung die Grauwerte auf das Intervall $[0, 255] = \{d \in \mathbb{N} | 0 \leq d \leq 255\}$ skaliert wurden. Für den NCC wurden verschiedene Fenstergrößen getestet, wobei das Ergebnis mit der Größe 9×9 anhand visueller Kriterien als bestes erachtet wurde (s. Abb. 2 links). Auch beim SGM wurden zahlreiche Parametereinstellungen getestet, und das visuell beste Ergebnis zum Vergleich ausgewählt (s. Abb. 2 rechts). Eine qualitative Analyse ergibt, dass das vom NCC und SGM erzeugte Disparitätsfeld die glatte Oberfläche sehr gut modelliert. Es fällt jedoch auf, dass insbesondere beim SGM kleine Öffnungen nicht korrekt oder überhaupt nicht erfasst werden.

In einem zweiten Test werden nun die Ergebnisse ohne projiziertes Muster miteinander verglichen (s. Abb. 3). Mit dem lokalen Ansatz NCC bei einer Fenstergröße von 9×9 Pixel kann ohne das Muster auf dieser sehr texturarmen Oberfläche kein zufriedenstellendes

Ergebnis erzielt werden (s. Abb. 3 l. o.). Der Test mit einem beträchtlich größeren Fenster von 49×49 Pixel ergibt zumindest eine komplette Abdeckung des Objektes (s. Abb. r. o.). Die Glattheit der Oberfläche wird allerdings nicht korrekt wiedergegeben. Der Effekt des „Foreground Fattening“ (SCHARSTEIN & SZELISKI 2002) bei fensterbasierten Ansätzen wird hier besonders deutlich, d. h. an Tiefendiskontinuitäten werden die Disparitäten des hinteren Objektes fälschlicherweise dem vorderen zugeschlagen. Beim SGM ist bei der Verwendung von AD als Ähnlichkeitsmaß zunächst zu beobachten, dass aufgrund geringer Helligkeitsunterschiede in den beiden Bildern weite Teile der Tür nicht rekonstruiert werden können (s. Abb. 3 l. u.). Mit MI kann die gesamte Tür dennoch rekonstruiert werden (s. Abb. r. u.). Im Vergleich zum fensterbasierten Ansatz wird die Oberfläche besser dargestellt, und der „Foreground Fattening“-Effekt tritt nur sehr gering auf.

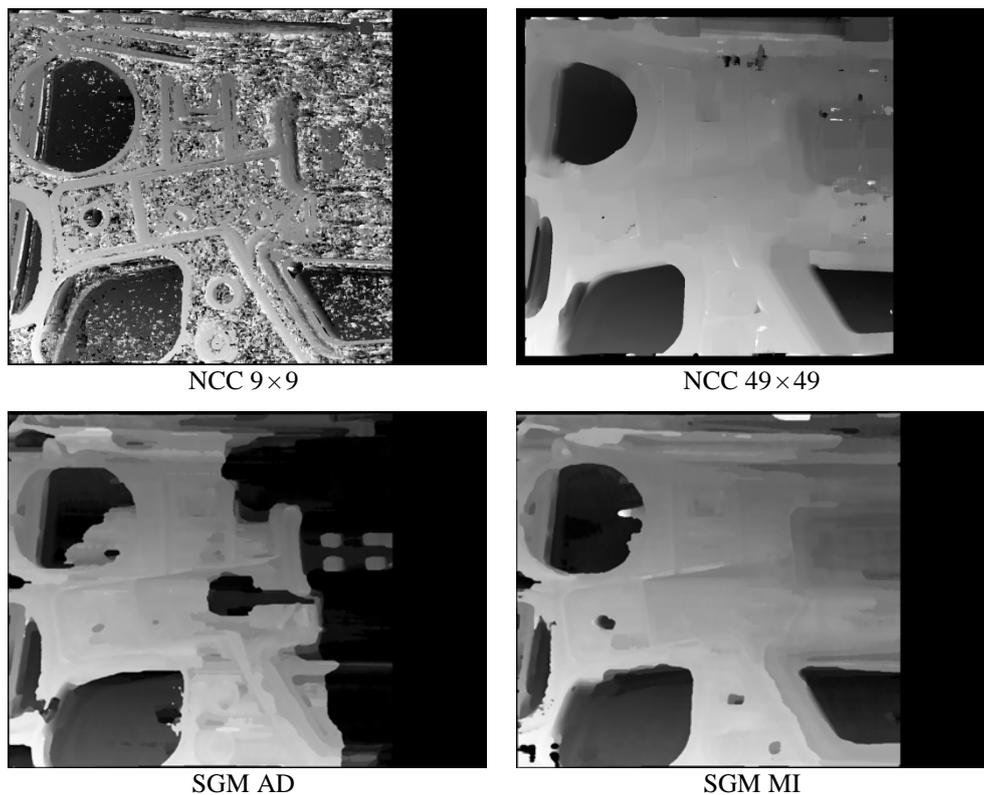


Abb. 3: Ergebnisse ohne projizierte Textur

Zusammenfassend lässt sich schließen, dass unter Verwendung eines projizierten Musters das SGM keine Vorteile gegenüber dem klassischen fensterbasierten Ansatz erzielen kann. Wenn dagegen texturarme Bereiche eines Objektes rekonstruiert werden sollen, ermöglicht das SGM eine vollständige und annähernd korrekte Rekonstruktion der Oberfläche.

4 Genauigkeitsuntersuchungen

Eine dichte Oberflächenrekonstruktion mit Matching-Verfahren erfordert eine sehr schmale Basis. Optimale Voraussetzungen für die Oberflächenrekonstruktion in Form einer möglichst dichten 3D-Punktwolke ergeben sich bei nahezu parallelen Aufnahmerichtungen und kleiner Basis. Demgegenüber steht für hohe Genauigkeiten bei der photogrammetrischen Punktbestimmung die Forderung nach einem Höhen-Basis-Verhältnis von möglichst 1, das sich bei einer Schwenkung der Kameras von $26,5^\circ$ ($= \arctan(B/2H)$) aus der Senkrechten zur Basis B_2 einstellt (s. Abb. 4 rechts).

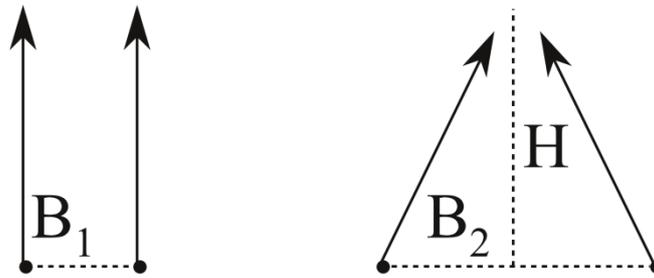


Abb. 4: Links: Optimale Anordnung für dichte Oberflächenrekonstruktion mit paralleler Basis; Rechts: konvergente Anordnung für ausgeglichene Genauigkeit bei der photogrammetrischen Punktbestimmung mit Messmarken

Die Ergebnisse der Oberflächenrekonstruktion zeigen, dass sich bereits mit einfachen Verfahren Ergebnisse mit geringem Rauschen in der Tiefe erzielen lassen, allerdings nur für den Normalfall mit einer vergleichsweise schmalen Basis von etwa 100 mm bei einem Messvolumen von 300 mm Kantenlänge und einem Aufnahmeabstand H von etwa 500 mm. In Anlehnung an die VDI/VDE 2634 werden Maßstäbe mit Messmarken an verschiedenen Stellen im Messvolumen platziert, und die sich aus der 3D-Punktbestimmung ergebenden Strecken werden mit den Sollstrecken auf den Maßstäben verglichen.

Das Stereosystem mit konvergenter Anordnung der beiden Kameras hat eine Basis von etwa 280 mm und eine Schwenkung von etwa 28° , also nahe am theoretischen Wert von $26,5^\circ$ für ein Höhen-Basis-Verhältnis von 1; das Normalfallsystem besitzt eine Basis von 100 mm.

In dieser Genauigkeitsuntersuchung werden photogrammetrische Streckenmessungen miteinander verglichen, die mit einem Stereokamerasystem einmal in Normalfall- und einmal in Konvergenzanordnung aufgenommen wurden. In beiden Fällen werden dieselben 1,3 Megapixel-Kameras mit hochauflösenden Objektiven eingesetzt, deren innere Orientierungen mit dem Parametersatz nach (BROWN 1971) modelliert wurden. Die Parameter der relativen Orientierung sind in beiden Fällen im Rahmen der Kalibrierung als Unbekannte mitbestimmt worden.

Die Anordnung der Strecken ist zusammen mit den jeweiligen Längenabweichungen in Abb. 5 für den Normalfall dargestellt. Die (blauen) Striche innerhalb des Messvolumens

markieren die räumliche der Lage der Strecken, die Längenabweichungen sind als (rote) Kreise jeweils an einem Ende der Strecke aufgetragen. Zur Verdeutlichung sind die Längenabweichungen stark vergrößert dargestellt.

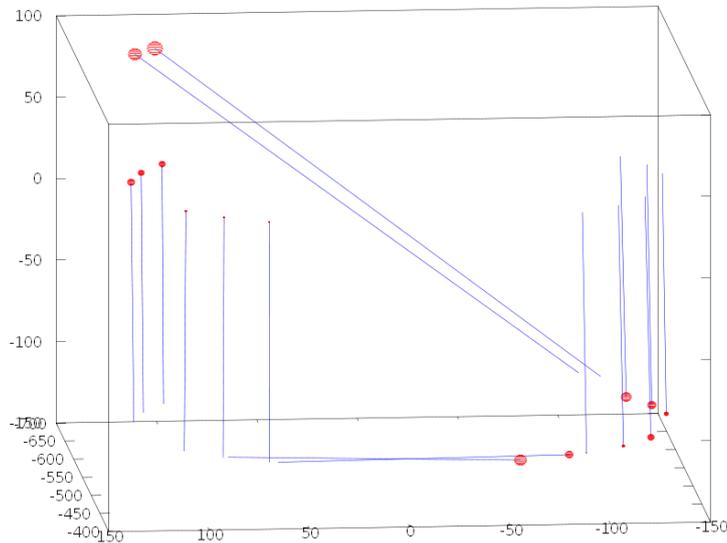


Abb. 5: Anordnung der Probestrecken im Messvolumen (Normalfall)



Abb. 6: Soll-Ist-Vergleich Streckenmessung im Normalfall als orange, gestrichelte Linie und für den Konvergentfall als blaue, durchgezogene Linie. Auf der Abszisse sind Streckenbezeichnungen (Punktnummeranfang-Punktnummerende) aufgetragen, an der Ordinate die entsprechenden Längenabweichungen in mm.

Die absoluten Werte der Längenabweichungen sind in Abb. 6 dargestellt, für den Normalfall als orange, gestrichelte Linie und für den Konvergentfall als blaue, durchgezogene Linie. Auf der Abszisse sind die Bezeichnungen der einzelnen im Messvolumen verteilten Maßstäbe mit Punktnummeranfang und Punktnummerende aufgetragen, an der Ordinate die entsprechenden Längenabweichungen in Millimetern. Das ungünstige Höhen-Basis-Verhältnis von 5 beim Einsatz des Normalfallsystems führt zu größeren Längenmessabweichungen, für den Normalfall ergeben sich maximale Streckenmessabweichungen $\pm 0,1$ mm und für den Konvergentfall Werte zwischen $+0,01$ und $-0,05$ mm (Abb. 6).

5 Fazit und Ausblick

Ziel des Gig3D-Vorhabens ist die Entwicklung eines high-speed tauglichen, flächenhaft antastenden optischen Messsystems, mit dem bis zu einer Milliarde Oberflächenpunkte in einer Sekunde aufgenommen werden können. Entsprechend leistungsfähige Kameras stehen seit Anfang des Jahres zur Verfügung, zum Beispiel die Pco.dimax mit Bildraten von bis zu 1279 Bildern pro Sekunde (fps) bei voller Auflösung von 2016×2016 Pixeln. Unter Verwendung eines 3D-CamSplitters sind damit erstmals 1000×1000 3D-Punkte pro Stereobildpaar möglich; in einer Sekunde bei 1000 fps also 10^9 (= ein Giga) 3D-Punkte. Entsprechend leistungsfähige Projektoren sind derzeit nicht verfügbar. Aus diesem Grund fokussiert sich die Arbeit im Rahmen des Vorhabens auf die Ableitung von Oberflächen mit Zuordnungsverfahren ohne projizierte Muster. Lediglich zur Bestimmung von Referenzoberflächen wird hier ein stochastisches Muster auf die Oberflächen projiziert.

Von den untersuchten Zuordnungsverfahren hat sich das einfach und schnell zu berechnende Ähnlichkeitsmaß SAD als ausreichend leistungsfähig für Oberflächen mit Muster herausgestellt. Von den komplexeren Verfahren hat sich in den meisten Versuchen SGM als das leistungsfähigste herauskristallisiert. Die langen Rechenzeiten werden den Einsatz von SGM allerdings auf diejenigen Anwendungsfälle einschränken, bei denen die Bildsequenzen zunächst aufgezeichnet und dann offline ausgewertet werden können. Für die meisten Hochgeschwindigkeitsanwendungen stellt das jedoch kein Problem dar, da die Filme in der Regel dauerhaft gespeichert werden und somit auch für eine Offline-Auswertung zur Verfügung stehen. Vor diesem Hintergrund sind auch die Ergebnisse des Vergleiches zwischen Normalfallanordnung mit schmaler Basis und Konvergentanordnung aus Anwendersicht von großem Interesse. Die Normalfallanordnung der Kameras vermeidet nämlich die aus Sicht des Anwenders häufig unerwünschten Schrägansichten, die bei konvergenten Anordnungen unvermeidbar sind und den Wert der Filmaufnahmen für den Anwender oft reduzieren. Aus messtechnischer Sicht ist in der Theorie die Konvergentanordnung aufgrund des günstigeren Höhen-Basis-Verhältnisses dem Normalfall vorzuziehen. Die praktische Untersuchung mit zwei identischen Kamerapaaren in unterschiedlicher Anordnung hat hier gezeigt, dass zumindest für kleinere Messvolumen und bei sorgfältiger Kalibrierung auch mit einem ungünstigen Höhen-Basis-Verhältnis von 5 auch im Normalfall eine hohe Messgenauigkeit erreicht werden kann.

Auch wenn die hier vorgestellten Ergebnisse insbesondere in den texturarmen Regionen noch nicht perfekt sind, zeichnet sich deutlich das Potenzial der verfolgten Strategie ab. Bei den bisherigen Untersuchungen ist der hohe Dynamikumfang der Pco.dimax von 12 Bit noch nicht genutzt worden. Durch die bessere Ausdifferenzierung in den Grauwerten sind bei Nutzung des vollen Dynamikumfangs eine Verbesserung der Ergebnisse zu erwarten.

Mit Muster steht schon jetzt ein für Hochgeschwindigkeitsaufnahmen taugliches Verfahren für die präzise Oberflächenableitung zur Verfügung. Ohne Muster können schon jetzt mit den hier vorgestellten Verfahren – bei reduzierten Anforderungen an die Genauigkeit – 3D-Modelle von hochdynamisch ablaufenden Prozessen auf der Basis von Stereo-Hochgeschwindigkeitsfilmen in Normalfallanordnung berechnet werden.

6 Danksagung

Gig3D wird im Rahmen des Zentralen Innovationsprogramms Mittelstand (ZIM) als KF Kooperationsprojekt des Instituts für Photogrammetrie und GeoInformation der Leibniz Universität Hannover und der SOLVing3D GmbH vom Bundesministerium für Wirtschaft und Technologie gefördert. Unser Dank gilt ebenfalls den Mitarbeitern des Projektträgers AiF für ihre kompetente und engagierte Beratung und die ausgesprochen schnelle Bearbeitung der Anträge.

7 Literatur

- Battle, J., E. Mouaddib, J. Salvi (1998): *Recent Progress in Coded Structured Light as a Technique to Solve the Correspondence Problem: A Survey*. Pattern Recognition, 31(7), S. 963–982.
- Birchfield, S. & C. Tomasi (1999): *Depth Discontinuities by Pixel-to-Pixel Stereo*. Int. J. of Computer Vision, 35(3), S. 269–293.
- Boykov, Y., O. Veksler, R. Zabih (2001): *Fast Approximate Energy Minimization via Graph Cuts*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 23(11), S. 1222–1239.
- Brown, D. C. (1971): *Close-range camera calibration*. Photogrammetric Engineering, 37(8), S. 855–866.
- Hirschmüller, H. (2008): *Stereo Processing by Semi-Global Matching and Mutual Information*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 30(2), S. 328–341.
- Hirschmüller, H., D. Scharstein (2009): *Evaluation of Stereo Matching Costs on Images with Radiometric Differences*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 31(9), S. 1582–1599.
- Kang, S. B., J. Webb, L. Zitnick, T. Kanade (1995): *A Multibaseline Stereo System with Active Illumination and Real-Time Image Acquisition*. Proc. Fifth Int. Conf. on Computer Vision, S. 88–93.
- Maes, F., A. Collignon, D. Vandermeulen, G. Marchal, P. Suetens (1997): *Multimodality Image Registration by Maximization of Mutual Information*. IEEE Trans. on Medical Imaging, 16(2), 187–198.
- Moravec, H. P. (1980). *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*. Ph.D. Thesis, Stanford University, Report STAN-CS-80-813.

- Pagès, J., J. Salvi, R. García, C. Matabosch (2003): *Overview of Coded Light Projection Techniques for Automatic 3D Profiling*. Proc. IEEE Int. Conf. on Robotics and Automation, 1, S. 133–138.
- Scharstein, D. & R. Szeliski (2002): *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms*. Int. J. of Computer Vision, 47(1), S. 7–42.
- Terzopoulos, D. (1986): *Regularization of Inverse Visual Problems Involving Discontinuities*. IEEE Trans. on Pattern Analysis and Machine Intelligence, 8(4), S. 413–424.
- VDI/VDE 2634 (2002): *Blatt 1: Optische 3D-Messsysteme - Bildgebende Systeme mit punktförmiger Antastung*. VDI/VDE-Handbuch Messtechnik II, Beuth Verlag GmbH, Berlin, 10 Seiten
- Viola, P. & W. M. Wells III (1997): *Alignment by Maximization of Mutual Information*. Int. J. of Computer Vision, 24(2), S. 137–154.