

---

# Untersuchung zur Genauigkeit der Kamerakalibrierung über Merkmalspunkte in SfM-Werkzeugen

Matthias SOOT, Marc SCHULZE und Christian MULSOW

## Zusammenfassung

Im Bereich der Photogrammetrie werden Kameras als Messinstrumente genutzt. Hierbei können bei den Messungen relative Genauigkeiten von besser als 1:100.000 der Objektgröße erreicht werden. Um diese Genauigkeiten zu erreichen, ist es, wie auch bei allen anderen Messinstrumenten notwendig, das Messinstrument zu kalibrieren.

Einer der aufwendigsten Vorbereitungsschritte für eine photogrammetrische Messung stellt heute die Vermarkung diskreter Punkte an einem Objekt dar. Hier kommen i.d.R. codierte und uncodierte Zielzeichen zum Einsatz, die an verschiedenen Stellen des Objektes aufgeklebt werden.

In der Computer Vision (CV) ist bereits ein hoher Automatisierungsgrad erreicht. Beispielsweise können Structure from Motion (SfM)-Werkzeuge ausschließlich aus Aufnahmen von einem 3D-Objekt ohne weitere Nutzerinteraktion und Vorinformation 3D-Koordinaten von Punkten des Objektes rekonstruieren. Bei Forschungen in diesem Bereich lag der Fokus bisher bei der Automatisierung und der Schnelligkeit der Prozesse. Die Genauigkeit der Rekonstruktion spielt jedoch zunehmend eine größere Rolle. Teil der SfM-Prozedur ist die Detektion von Merkmalspunkten in den Aufnahmen des Objektes sowie die Zuordnung dieser Punkte in mehreren Aufnahmen. Die in den SfM-Werkzeugen zugeordneten Punkte können dann analog zu zugeordneten codierten Zielmarken für eine Kamerakalibrierung genutzt werden.

Im Rahmen dieser Arbeit wird zunächst die Qualität der in den SfM-Werkzeugen ermittelten Merkmalspunkte untersucht und die Kalibrierung der Kamera mit diesen Merkmalspunkten durchgeführt. Die Kalibrierung wird parallel auch mit Zielmarken durchgeführt und mit der Kalibrierung mit Merkmalspunkten verglichen. Auch die SfM-Werkzeuge berechnen bereits eigene Parameter für die innere Orientierung der Kamera.

Die Qualität der Prozedur sowie der ermittelten Parameter und Modelle wird im Rahmen dieser Arbeit ebenfalls untersucht.

## 1 Einleitung

Ziel der Arbeit ist die Untersuchung der erreichbaren Genauigkeit einer Kamerakalibrierung mit Merkmalspunkten. Hierzu wurden die in verschiedenen SfM-Werkzeugen zugeordneten Merkmalspunkte in Auswertungen als Beobachtungen in ein Industriemesssystem (AICON 3D-Studio) eingeführt, um dort im Rahmen einer Bündelblockausgleichung die Parameter der inneren Orientierung zu ermitteln. Derselbe Bildverband wurde in einer weiteren Auswertung genutzt, um die Parameter der inneren Orientierung in einer Bündelblockausgleichung mit beobachteten codierten Zielmarken durchzuführen. Des Weiteren

wurden die Ergebnisse der Simultankalibrierung des Aufnahmeverbandes in den SfM-Werkzeugen analysiert.

Im Rahmen der Arbeit wurden verschiedene SfM-Werkzeuge mit unterschiedlichen Implementierungen des von LOWE (2004) entwickelten SIFT-Detektors genutzt. Die genutzten SfM-Werkzeuge sind:

- Bundler (vgl. SNAVELY et. al. (2007)) mit einer SIFT Implementierung aus der VLFeat-Bibliothek
- VisualSfM (vgl. WU (2013)) mit SiftGPU
- PhotoScan mit einer eigenen, an SIFT angelehnten, Implementierung eines Merkmalsdetektors

Im Rahmen dieser Arbeit wurde zur Modellierung der Bildfehler beschreibenden Funktionen das auf BROWN (1971) basierende Modell, welches heute in fast allen gängigen Photogrammetrie-Softwares implementiert ist, verwendet. Es berücksichtigt die wesentlichen physikalisch begründbaren Fehlereinflüsse wie die radialsymmetrische Verzeichnung ( $\Delta r_{\text{sym}}$ ), radialasymmetrische Verzeichnung ( $\Delta r_{\text{asy}}$ ) sowie Affinität (Scherung und Maßstab). Im Rahmen der Arbeit wurde die DSLR Kamera Nikon D700 mit einem 28 mm Nikkor Festbrennweiten-Objektiv kalibriert, welche mit einem Vollformatsensor (24 x 36 mm) ausgestattet ist.

## 2 Voruntersuchungen

### 2.1 Geeignete Texturen für Merkmalspunktedektoren

Im Rahmen der Arbeit wurden verschiedene Texturen unter gleichen Aufnahmebedingungen analysiert und die Anzahl und Qualität (Zuordenbarkeit) der Merkmalspunkte untersucht. Ziel war es, eine Textur zu finden, die eine möglichst große Zahl an Merkmalspunkten guter Qualität und eine gleichmäßige Verteilung dieser Merkmalspunkte ermöglicht. Schachbrettmuster und Millimeterpapier liefern bspw. eine große Zahl an Merkmalspunkten, die jedoch auf Grund der sehr ähnlichen Deskriptoren sehr schlecht zugeordnet werden können. Texturen mit geringem Kontrast oder sehr grober Textur liefern eine geringe Anzahl an Merkmalspunkten.

Als Fazit kann festgehalten werden, dass unregelmäßige Texturen mit starken Variationen verwendet werden sollten, um viele SIFT-Punkte zuordnen zu können. Des Weiteren sollte auf eine gleichmäßige Intensität der Textur geachtet werden, um eine regelmäßige Verteilung der Merkmalspunkte innerhalb der Aufnahmen zu erreichen.

#### 2.1.1 Qualität der Deskriptoren

Im Rahmen der Arbeit wurden verschiedene Implementierungen des Merkmalsdetektors SIFT genutzt. Um die Qualität der Merkmalsdeskriptoren einzuordnen, wurden die Verhältnisse der detektierten Punkte je Bild zu den zugeordneten Punkten vor einer robusten Analyse der Zuordnung und nach dieser Analyse untersucht. Die durchschnittliche Anzahl an Merkmalspunkten pro Bild liegt bei rd. 8.000 Punkten bei Bundler und VisualSfM sowie 40.000 bei PhotoScan (höchste Genauigkeit). Von diesen Punkten werden rd. 1/5 (PhotoScan) bis 1/7 (Bundler) einem Merkmalspunkt in einem weiteren Bild zugeordnet. Die

robuste Prüfung der Zuordnung durch RANSAC (Random Sample Consensus) zeigt, dass die vorherige Zuordnung der Punkte bei Bundler am zuverlässigsten durchgeführt wurde (rd. 90 % der Zuordnungen bestätigt). Die schlechtesten Ergebnisse liefert hier PhotoScan mit lediglich 70 % bestätigter Zuordnung.

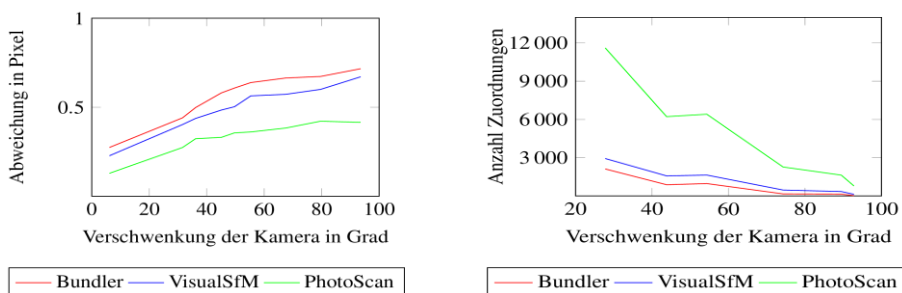
### 2.1.2 Lagequalität der Merkmalspunkte

Die herkömmlich zur Kamerakalibrierung genutzten codierten Zielmarken können mit einer Lagegenauigkeit von  $1/20$  bis  $1/50$  Pixel bestimmt werden. Zur Beurteilung der Lagegenauigkeit der SIFT-Merkmalspunkte wurde eine flächenbasierte Merkmalszuordnung in Form eines 6 Parameter Least-Squares-Matching (LSM) durchgeführt, bei dem die Genauigkeit zwischen  $1/10$  und  $1/50$  Pixel angegeben werden kann. Die Abweichungen der Lage der Merkmalspunkte zu den beim LSM ermittelten korrespondierenden Punkten kann als Lagegenauigkeit des SIFT-Merkmalspunktes angenommen werden.

Bei der Untersuchung der Lagegenauigkeit wurde festgestellt, dass die SIFT Implementierungen aus der VLFeat-Bibliothek und SiftGPU die genutzten Bilder verkleinern, bevor der SIFT-Algorithmus durchgeführt wird. Dies führt zu einer deutlichen Verschlechterung der Punktlagegenauigkeit bezogen auf die Ausgangsauflösung der Bilder. Bei Bildern mit geringer Auflösung ( $1596 \times 1096$  Pixel) konnten mit allen SIFT-Implementierungen Genauigkeiten von  $1/10 - 1/5$  Pixel erreicht werden. Bei Bildern in großer Auflösung ( $4256 \times 2832$  Pixel) konnten bezogen auf die Originalauflösungen mit Bundler und VisualSfM nur Genauigkeiten von  $> 1/2$  Pixel erreicht werden. Mit PhotoScan konnten auch in voller Auflösung Genauigkeiten von  $< 1/5$  Pixel erreicht werden.

### 2.1.3 Affinvarianz des SIFT Algorithmus

In der einschlägigen Literatur zur Kamerakalibrierung in der Nahbereichsphotogrammetrie (vgl. u.a. MASON (1994)) wird herausgestellt, dass Aufnahmen zur Erreichung hoher Genauigkeit möglichst konvergent erfolgen sollten, um hohe Punktlagegenauigkeiten zu erhalten. Hierzu muss der Merkmalsdetektor affinvariant sein. Der entwickelte skaleninvariante Merkmalsdetektor SIFT wurde im Rahmen dieser Arbeit auf seine Affinvarianz untersucht. Hierbei wurden Aufnahmen mit einer jeweiligen Verschwenkung von rd. 10 Bogengrad durchgeführt.

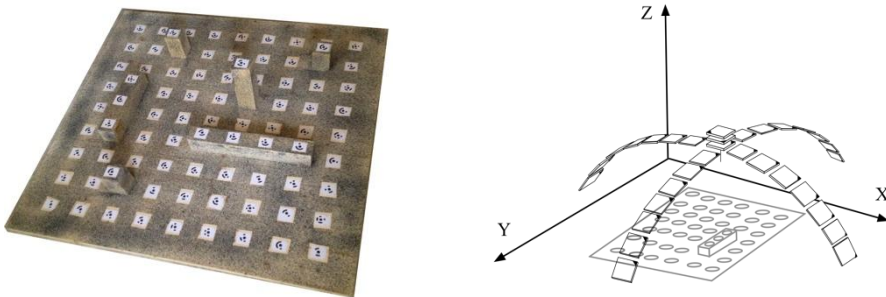


**Abb. 1:** Darstellung der Abhängigkeit der Lagegenauigkeit der SIFT-Punkte von der Verschwenkung der Kamera (Auflösung von  $1596 \times 1096$  Pixel) sowie die Abhängigkeit der Anzahl zuordenbarer Punkte mit der Kameraverschwenkung.

Bei Verschwenkungen  $<10^\circ$  können Genauigkeiten bis  $1/10$  Pixel in der Lagegenauigkeit erreicht werden. Die Lagegenauigkeit nimmt nahezu linear zur Verschwenkung der Kamera ab und erreicht bei den Implementierungen des SIFT aus der VLFeat-Bibliothek und SiftGPU Lagegenauigkeiten schlechter als  $1/2$  Pixel (vgl. Abbildung 1). Neben der Lagegenauigkeit nimmt ebenfalls die Anzahl der zuordenbaren Punkte deutlich ab (vgl. Abbildung 1). Der SIFT-Algorithmus ist also erkennbar nicht affinvariant.

### 3 Auswertung

Zur Durchführung der Analysen mit Textur und Zielmarken wurde ein Testfeld mit geeigneter, künstlicher Textur und 81 codierten 12-Bit AICON Zielmarken erstellt (siehe Abbildung 2). Im Rahmen der Arbeit wurde ein Bildverband mit 31 Aufnahmen des Testfeldes ausgewertet. Die Aufnahme erfolgte dabei nach einer Erweiterung des Modells von GODDING (1993) (vgl. Abbildung 2). Die deutlich höhere Anzahl an Bildern wurde verwendet, um Merkmalspunkte auch über größere Verschwenkungen der Aufnahmen zuordnen zu können.



**Abb. 2:** Bild des Testfeldes mit codierten Zielmarken und künstlicher Textur sowie Darstellung der erweiterten Aufnahmekonfiguration nach GODDING (1993)

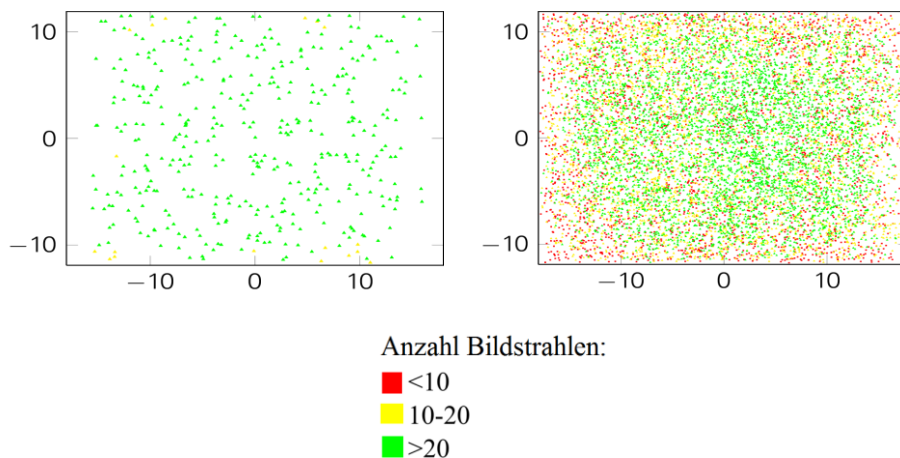
Die Auswertung der Aufnahmen wurde dann zunächst in den 3 SfM-Werkzeugen durchgeführt. Die robust zugeordneten Merkmalspunkte aus den SfM-Werkzeugen wurden anschließend in das Industriemesssystem AICON 3D-Studio als Beobachtungen eingeführt und die Bündelblockausgleichung durchgeführt. Hierzu wurden lediglich die Punkte genutzt, die mindestens fünf Mal beobachtet wurden. Dies entspricht ca.  $1/3$  der rekonstruierten Merkmalspunkte.

Bei den Bündelblockausgleichungen mit den Merkmalspunkten wurden sehr hohe Gesamtredundanzen erreicht (vgl. Tabelle 1). Die Standardabweichungen der Gewichtseinheit sind für die Auswertungen mit Merkmalspunkten um den Faktor 5-15 mal größer. Dies kann auf die schlechtere Punktlagegenauigkeit der Merkmalspunkte zurückgeführt werden. Auf Grund der hohen Anzahl von Punkten in der Bündelblockausgleichung ist die benötigte Zeit hier deutlich größer (vgl. Tabelle 1).

**Tabelle 1:** Gegenüberstellung der Informationen zur Bündelblockausgleichung in AICON 3D-Studio.

	Zielmarken	Merkmalspunkte aus		
	Bildmessung	Bundler	VisualSfM	PhotoScan
Anzahl genutzter Punkte	81	5.559	8.936	31.594
Gesamtredundanz ( $r=n-u$ )	3.657	149.129	159.316	623.800
$\sigma_0$	0,00045	0,00583	0,0048	0,002
benötigte Zeit	3 min	30 min	1h	2h

Bei der Auswertung wurde festgestellt, dass insbesondere in den Ecken der Bildsensoren weniger Merkmalspunkte beobachtet wurden und die hier beobachteten Punkte schlechte Strahlenschnittgeometrien aufweisen (geringe Anzahl an Strahlenschnitten (vgl. Abbildung 3) und schlechte Schnittwinkel).



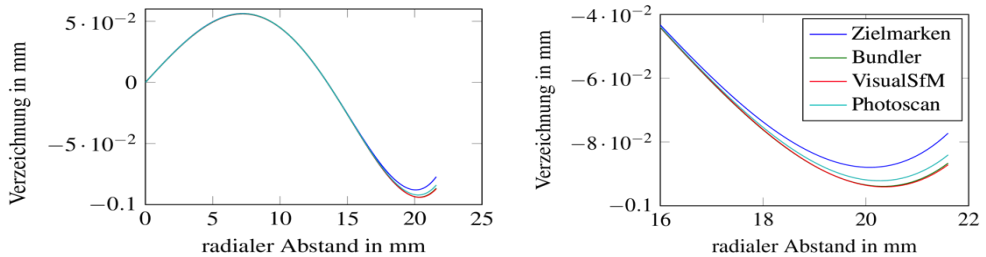
**Abb. 3:** Darstellung der geometrischen Stabilität der auf dem Sensor beobachteten Punkte anhand der Anzahl der Beobachtungen pro Punkt für Auswertungen mit Zielmarken (links) und Auswertungen mit Merkmalspunkten in Bundler (rechts).

## 4 Ergebnisse

### 4.1 Vergleich der Auswertungen in AICON 3D-Studio

Die Parameter der Kamerakonstanten ( $c_k$ ) und des Bildhauptpunktes ( $x_h, y_h$ ) weichen nicht signifikant voneinander ab. Bei den Parametern der bildfehlerbeschreibenden Funktionen (A1, A2, B1, B2 und C1) sind signifikante Unterschiede festzustellen. Der Parameter C2 (Scherung) wurde in allen Auswertungen nicht signifikant bestimmt und wurde daher von der Ausgleichung ausgeschlossen. Die Abweichungen der Parameter der radialsymmetrischen Verzeichnung (A1 und A2) führen bis 80% der halben Sensordiagonalen zu Abwei-

chungen von  $< 1/10$  Pixel. Darüber hinaus steigen die Abweichungen auf maximal 0,8 Pixel in der äußersten Sensorecke (vgl. Abbildung 4). Diese Differenzen sind auf die fehlenden bzw. schlechten Merkmalspunktbeobachtungen in den Ecken des Sensors zurückzuführen. Die Abweichungen der Parameter der radialsymmetrischen und tangentialen Verzeichnung (B1 und B2) führen bei der Auswertung mit den Merkmalspunkten aus VisualSfM zu Unterschieden von bis zu  $1/5$  Pixel in den äußersten Ecken. Bei den sonstigen Auswertungen bleiben die Diskrepanzen im Bereich von  $< 1/10$  Pixel. Die Abweichungen des Parameters C1 (Skalierung) führt zu Verzeichnungen von  $\approx 1/20$  Pixel. Die Auswirkungen der Einflüsse aller Parameter in Summe ist in Abbildung 6. dargestellt. Hierbei dominieren die Effekte der radialsymmetrischen Verzeichnung.



**Abb. 4:** Darstellung der Funktionsverläufe der radialsymmetrischen Verzeichnung für den ausgewerteten Bildverband (Pixelgröße  $8\mu\text{m}$ ) sowie Detailansicht des Funktionsverlaufs am Sensorrand.

## 4.2 Vergleich der Ergebnisse aus AICON 3D-Studio mit der Kalibrierung in den SfM-Werkzeugen

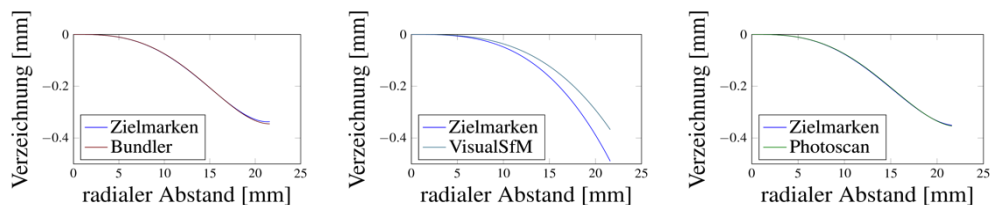
In einer weiteren Untersuchung wurde die Qualität der Simultankalibrierung der Kamera in den SfM-Werkzeugen untersucht. Die genutzten SfM-Werkzeuge nutzen unterschiedliche Ausprägung des Modells nach Brown (vgl. Tabelle 2).

**Tabelle 2:** Gegenüberstellung der Anzahl der ableitbaren Parameter in AICON 3D-Studio und den genutzten SfM-Werkzeugen

	AICON 3D	Bundler	VisualSfM	PhotoScan
Radialsym. Verzeichnung	1 bis 3	2	1	3 oder 4
Radialasym. Verzeichnung	1 oder 2	n.v.	n.v.	2
Affinität	1 oder 2	n.v.	n.v.	1 oder 2

Die mathematische Beschreibung der Parameter ist nicht identisch, lässt sich jedoch durch eine Normierung ineinander überführen. In AICON 3D-Studio wurden die Parameter, die in den jeweiligen SfM-Werkzeugen mitbestimmt werden, durch Auswertungen mit Zielmarken in übergeordneter Genauigkeit ermittelt. Die Parameter der inneren Orientierung aus den SfM-Werkzeugen und AICON 3D-Studio wurden anschließend miteinander verglichen. Die Abweichungen der Parameter der radialsymmetrischen Verzeichnung bei

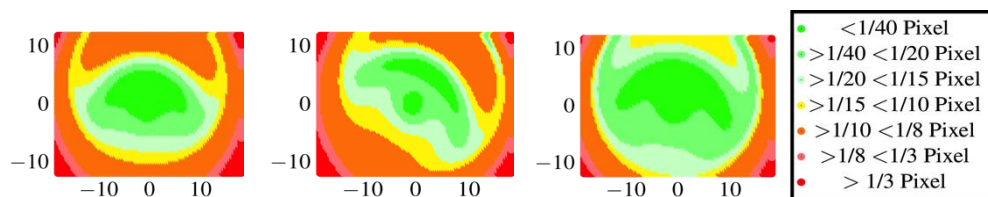
Bundler und PhotoScan bleiben im geringen Bereich ( $< 1/10$  Pixel für weite Teile des Sensors) und erreichen erst zum äußeren Sensorrand eine Größenordnung von  $\approx 1$  Pixel. Lediglich die Auswertung der radialsymmetrischen Verzeichnung mit VisualSfM weicht stark (einige Pixel) von der Auswertung in AICON 3D-Studio ab (vgl. Abbildung 5). Dies ist vermutlich auf unterschiedliche Nutzung von Bildradien im mathematischen Modell zurückzuführen.



**Abb. 5:** Gegenüberstellung der Funktionsverläufe der radialsymmetrischen Verzeichnung für den ausgewerteten Bildverband (Pixelgröße  $8,4 \mu\text{m}$ ) für Bundler (2 Parameter) VisualSfM (1 Parameter) und PhotoScan (3 Parameter).

## 5 Fazit und Ausblick

Die Untersuchungen zu den Merkmalspunkten haben gezeigt, dass die Anzahl zuordenbarer Punkte bei großen Verschwenkungen sehr gering werden. Die erreichbare Lagegenauigkeit der SIFT-Punkte liegt bei  $\approx 1/10$  Pixel und nimmt mit der Verschwenkung der Aufnahmen näherungsweise linear ab. Trotz der deutlich schlechteren Lagegenauigkeit der Merkmalspunkte im Vergleich zur Verwendung von Zielmarken können in AICON 3D-Studio gute Kalibrierergebnisse erzielt werden. Dies ist insbesondere auf die deutlich größere Gesamtredundanz in der Bündelblockausgleichung zurückzuführen. Die Abweichungen der Parameter der inneren Orientierung führen in weiten Teilen des Sensors zu Abweichungen  $< 1/10$  Pixel. Lediglich in den Sensorecken kommt es zu größeren Abweichungen (vgl. Abbildung 6). Auch die in den SfM-Werkzeugen durchgeführte Simultankalibrierung liefert bereits gute Ergebnisse.



**Abb. 6:** Darstellung der Wirkung in Pixel auf dem Bildsensor die durch die Abweichungen der Verzeichnungsparameter zwischen der Kalibrierung mit Zielmarken und Merkmalspunkten entstanden sind. Zum Bildrand hin wirken die Abweichungen stärker. Zum Bildzentrum bleiben die Abweichungen unter 1/40 Pixel. Für die Abweichungen bei der Kalibrierung mit Merkmalspunkten (v.l.n.r.) aus Bundler, VisualSfM und PhotoScan.

Zur Erhöhung der Zuordenbarkeit der Merkmalspunkte sollten in weiteren Untersuchungen alternative Merkmalsdetektoren wie bspw. der von MOREL & YU (2009) entwickelte ASIFT verwendet werden, die eine höhere Zahl an Zuordnungen auch bei großen Verschwenkungen zulassen.

Die geringe Punktdichte von stabil beobachteten Punkten am Rand und in den Ecken des Sensors ist auch auf die verwendete Aufnahmekonstellation zurückzuführen. Ein weiterer Grund für die schlechte Zuordnung kann in der Verzeichnung selbst gesehen werden. Die Merkmalsdeskriptoren weichen auf Grund der großen Verzeichnung am Bildrand stark von Deskriptoren im anderen Bereich des Bildes ab, sodass eine Zuordnung nicht möglich ist. Um diesen Mangel zu beseitigen, kann in weitergehenden Arbeiten eine iterative Vorgehensweise der Entzeichnung und erneuten Zuordnung von Merkmalspunkten durchgeführt werden. Des Weiteren sollten alternative Aufnahmekonfigurationen getestet werden.

## Literatur

- Brown, D. (1971): Close-range camera calibration. *Photogrammetric Engineering*, 37(8), S. 855-866.
- Godding, R. (1993): *Ein photogrammetrisches Verfahren zur Überprüfung und Kalibrierung digitaler Aufnahmesysteme*. *Zeitschrift für Photogrammetrie und Fernerkundung*, 2/1998, S. 82-90.
- Lowe, D. (2004): Distinctive image features from scale-invariant keypoints. *International journal of Computer Vision*, 60(2), S. 91-110.
- Mason, S. (1994): Expert system based design of photogrammetric networks. Dissertation an der ETH Zürich.
- Morel, J. & Yu, G. (2009): ASIFT: A new framework for fully affine invariant image comparison. *SIAM Journal on Image Sciences*, 2(2), S.438-469.
- Snavely, N., Seitz, S., & Szeliski, R. (2007): Modeling the World from Internet Photo Collections. *International Journal of Computer Vision*, 80(2), S. 189-210.
- Wu, C. (2013): Towards linear-time incremental structure from motion. *3DTV-Conference 2013*, S. 127-134.



**Bitte geben Sie auf einer separaten Seite zu allen genannten Autoren folgende Informationen:**

Soot, Matthias  
TU Dresden, Geodätisches Institut  
matthias.soot@t-online.de

Schulze, Marc  
TU Dresden, Institut für Photogrammetrie und Fernerkundung  
marc.schulze@tu-dresden.de

Mulsow, Christian  
TU Dresden, Institut für Photogrammetrie und Fernerkundung  
christian.mulsow@tu-dresden.de