

# **Low-cost und open-source Lösungen für die automatisierte Generierung von 3D-Punktwolken – ein kritischer Überblick**

Fabio Remondino  
3DOM - 3D Optical Metrology Unit  
FBK - Bruno Kessler Foundation  
via Sommarive 18  
38123 Povo-Trento, Italy  
remondino@fbk.eu  
<http://3dom.fbk.eu/en/home>

Thomas P. Kersten  
HafenCity Universität Hamburg  
Labor für Photogrammetrie & Laserscanning  
Hebebrandstrasse 1  
22297 Hamburg  
thomas.kersten@hcu-hamburg.de  
<http://www.hcu-hamburg.de/geomatik/kersten>

## **1 Einleitung**

Die Erstellung von 3D-Modellen von Gebäuden, Kulturobjekten, Monumenten oder der bebauten Umwelt ist heute Standard für viele Anwendungen wie Dokumentation, digitale Restaurierung, Visualisierung, Inspektion, Planung, AR/VR, Spiele, Unterhaltung und viele mehr. 3D-Modellierung umfasst dabei das gesamte Verfahren, um aus entsprechenden Vermessungsdaten (realitätsbasierte Methode) oder aus anderen Informationsquellen ein dreidimensionales Produkt (oder Objekt) zu erstellen. Daten können dafür mit Digitalkameras oder aktiven Sensoren aufgenommen werden, um sie mit den bekannten bildbasierten (Remondino & El-Hakim 2006) oder distanzbasierten (Vosselman & Maas 2010) Methoden zu verarbeiten. Die bildbasierte Methode gilt im Allgemeinen als preiswert (insbesondere für terrestrische Anwendungen), flexibel,

portable und als fähig, verlorene Szenarien einfach aus Archivbildern zu rekonstruieren (Grün et al. 2004). In den letzten Jahren sind verschiedene Lösungen für die automatisierte Verarbeitung von Bildern mit der entsprechenden Ableitung von 3D-Informationen und Modellen verfügbar geworden. Diese Verarbeitungsschritte umfassen hauptsächlich die Bildorientierung und die 3D-Rekonstruktion aus dichten Punktwolken mit einem hohen Automatisierungsgrad.

Dieser Beitrag untersucht das Leistungspotential und die Zuverlässigkeit einiger preiswerter kommerzieller und open-source Softwarepakete, die große Bildverbände automatisch verarbeiten und dabei unbekannte Kamerastandpunkte bestimmen können. Verschiedene Datensätze wurden verwendet, um die Ergebnisse der evaluierten Softwarepakete visuell und metrisch zu vergleichen und analysieren zu können.

## **2 Bisherige Arbeiten**

Mehrbildorientierung ist die wichtigste (grundlegende) Aufgabe in der Photogrammetrie und im Computer Vision. Sie stellt die wesentliche Voraussetzung für viele Anwendungen dar (Grün & Huang 2001) und sie schließt häufig die simultane Kamerakalibrierung (auch als Selbstkalibrierung bekannt) ein. Die Genauigkeit und die Zuverlässigkeit der Bildorientierung und der Kamerakalibrierung beeinflussen erheblich die Qualität aller folgenden Prozesse wie die Bestimmung des 3D-Punktes und die nachfolgende 3D-Modellierung. Bildorientierung und Kamerakalibrierung erfolgen bereits seit den frühen 1990ziger Jahren in der Nahbereichsphotogrammetrie mithilfe kodierter Zielmarken unter Laborbedingungen und für industrielle Anwendungen vollautomatisch (Ganci & Handley 1989). Jedoch in einigen anderen Anwendungsgebieten (im Feld) wie in der Architektur, Archäologie und Denkmalpflege können kodierte Zielzeichen nicht immer eingesetzt werden. In diesen Fällen ist die Identifikation von Verknüpfungspunkten für die Bestimmung der äußeren Orientierung komplexer und muss vorzugsweise unter Verwendung vollautomatischer Matchingverfahren gelöst werden. Andererseits ist eine genaue Passpunktmessung für die Skalierung und die Georeferenzierung von Bildverbänden (absolute Orientierung) noch eine manuelle (interaktive) Aufgabe.

Auf der einen Seite legt die Photogrammetrie mehr Wert auf *Genauigkeit* und *Zuverlässigkeit* bei der Bildorientierung und Kamerakalibrierung für Aufgaben

in den Bereichen Mapping, Kartographie, Dokumentation und Monitoring. Demgegenüber richtet die Computer Vision ihren Fokus auf einen hohen *Automationsgrad* dieser Prozesse in der Robotik und Inspektion, in dem häufig die geometrische Qualität der Ergebnisse vernachlässigt werden kann. Photogrammetrie benutzt ein Lochkameramodell im Allgemeinen mit zehn Parametern für die innere Orientierung (IO): Hauptpunkt (2 Parameter), Kamerakonstante (1), radial-symmetrische Verzeichnung (2-3), tangentielle Verzeichnung (2), Affinität (1) und Scherung (1). Computer Vision verwendet die projektive Geometrie, wobei die innere Orientierung der Kamera auf den Bildhauptpunkt (2), Brennweite (1) und radial-symmetrische Objektivverzeichnung (1-2) begrenzt sind. Als eine der Ersten hat Brown (1971) das photogrammetrische Kameramodell vorgestellt. In einem bemerkenswerten Artikel haben Clarke & Fryer (1998) die historische Entwicklung der gesamten Kamerakalibrierungstechniken einschließlich einer vollständigen Literaturübersicht zusammengefasst. Einen Überblick der aktuellen Methoden für die Kamerakalibrierung in der Nahbereichsphotogrammetrie und im Computer Vision mit operationellen Aspekten für die Selbstkalibrierung werden in Remondino & Fraser (2006) dargestellt, während durch Fraser (2012) der aktuelle Stand der automatisierten Kamerakalibrierung in der Nahbereichsphotogrammetrie erläutert wird. Eine Kamerakalibrierung sollte mit einem photogrammetrischen Bildverband (ideale Schnittgeometrie) berechnet werden, da ein typischer Bilddatensatz für eine 3D-Rekonstruktion oft keine signifikante Bestimmung aller Parameter der inneren Orientierung erlaubt. Dennoch favorisiert der Computer Vision Bereich die Selbstkalibrierung der Kamera im Feld bzw. am Objekt durch die so genannte *Structure-from-Motion* (SfM) Methode, die die simultane Bestimmung der inneren und äußeren Kameraparametern einschließlich Geometrie der 3D-Szene beinhaltet. Kamerakalibrierung wird im Allgemeinen mittels kodierter Zielmarken oder Schachbrettmuster durchgeführt, um höhere Genauigkeit in der Verknüpfungspunktidentifizierung und in der Kameraparameterschätzung zu erzielen. Eine Kamerakalibrierungsmethode ohne Zielmarken werden von Barazzetti et al. (2011) vorgestellt.

Andererseits wurde die automatische Orientierung terrestrischer Bilder ohne Zielmarken in der Photogrammetrie bereits in den letzten Jahren etabliert (Läbe & Förstner 2006, Remondino & Ressel 2006, Barazzetti et al. 2010, Pierrot-Deseilligny & Cléry 2011 und Del Pizzo & Troisi 2011). Doch die größten Entwicklungen und Innovationen kamen aus dem Bereich der Computer Vision seit dem Ende der 1990ziger Jahren mit der SfM-Methode (Fitzgibbon & Zisserman 1998, Pollefeys et al. 1999, Nister 2004, Vergauwen & Van Gool 2006, Snavely

et al. 2007, Agarwal et al. 2009, Strecha 2010, Wu 2011). Die durch das SfM abgeleiteten 3D-Rekonstruktionen sind im Allgemeinen hauptsächlich für Visualisierung, objekt-basierte Navigation oder andere ähnliche Zwecke, aber nicht für photogrammetrische Anwendungen und Kartierungsaufgaben nutzbar.

Aufgrund dieser Fortschritte der Algorithmen sowie durch die stetigen Verbesserungen in der Computertechnologie können heutzutage eine große Anzahl von Bildern in einem beliebig definierten Objektkoordinatensystem unter Verwendung einer Vielzahl von Algorithmen automatisch orientiert werden. Dafür werden häufig open-source Software (VisualSfM, Bundler, Apero, Insight3D, etc.) oder freie Webservices (Microsoft Photosynth, Autodesk 123D Catch Beta, My3DScanner, Hypr3D, Arc3D, etc.) verwendet. Verschiedene Softwarepakete wie PMVS2 (Furukawa & Ponce 2010) und MicMac (Pierrot-Deseilligny & Paparoditis 2006) liefern bereits nach Bestimmung der Kameraparameter dichte Punktwolken oder polygonale Modelle als 3D-Rekonstruktionen. Dieser Automatisierungsgrad beim Ablauf der 3D-Rekonstruktion aus Bildern hat auch vermehrt zur Zusammenarbeit in den Bereichen Mapping und Modellierung geführt, die sich z.B. in den Projekten PhotoCity (Tuite et al. 2011) oder Volunteer Geographic Information (VGI) (Uden & Zipf 2012) widerspiegeln. Einerseits hilft diese hohe Automation und die web-basierte Verarbeitung der Bilddaten Personal in der Datenverarbeitung zu sparen, aber andererseits bestehen keine Garantie auf Datensicherheit sowie ein Anspruch auf Qualität für das 3D-Modell. Im kommerziellen Bereich sind Softwarepakete wie PhotoModeler-Scanner und Agisoft PhotoScan vor kurzem auf dem Markt erschienen, die automatische Bildorientierung, Kamerakalibrierung und 3D-Objektrekonstruktion aus Bildsequenzen anbieten. Jedoch fehlt eine Bewertung der Genauigkeit, Zuverlässigkeit und Wiederholbarkeit dieser automatischen Verfahren. Daher könnte es u.a. für Nicht-Experten ziemlich nützlich sein, wenn bild-basierte Methoden der 3D-Rekonstruktion einander gegenüber gestellt werden, um so das Potential dieser Verfahren besser beurteilen zu können.

### **3 Bildtriangulation**

In der Photogrammetrie versteht man unter Bildtriangulation im Allgemeinen das Verfahren, um aus einem Bildverband die inneren und äußeren Orientierungsparameter abzuleiten. Das Verfahren erfordert eine zuverlässige Verknüpfung der Bilder (Bildkorrespondenzen) durch manuell oder automatisch extrahierte Verknüpfungspunkte, die die Grundlage für eine Bündelausgleichung nach

der Methode der kleinsten Quadrate bilden. Die wesentlichen Probleme, die die Softwareentwicklung von zuverlässiger kommerzieller Methoden erschweren, sind konvergente Bildaufnahmen, unvorhersehbare Basislinien zwischen den Bildern, Maßstabsänderungen und Beleuchtungsveränderungen, homogen strukturierte Aufnahmebereiche, sich wiederholende Muster im Objekt, sowie komplexe Objektstrukturen, etc. Diese Effekte stellen für die automatisierte Identifizierung von Verknüpfungspunkten in Bildern eine große Herausforderung dar.

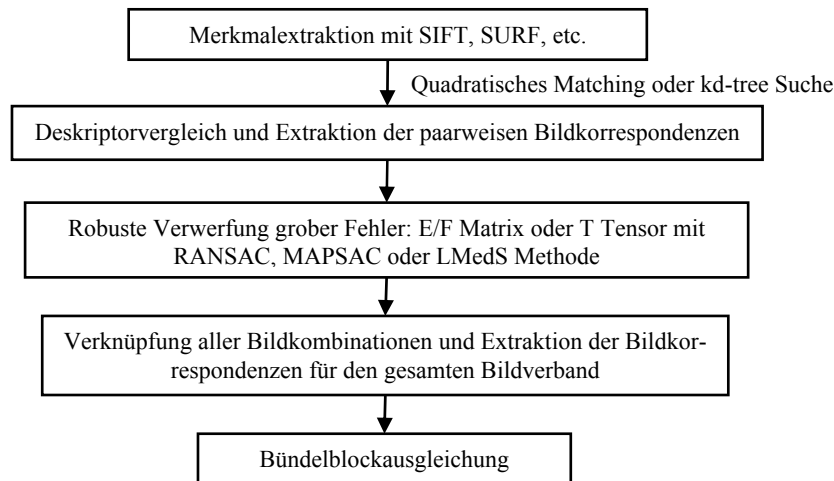


Abb. 1. Typischer Workflow für die automatisierte Extraktion von Bildkorrespondenzen.

### 3.1 Identifikation von Bildkorrespondenzen

Die Identifizierung von Bildkorrespondenzen beginnt mit der Extraktion von Punkten oder Flächen unter Verwendung von flächen- oder merkmals-basierten Algorithmen (Abb. 1). Heutzutage liefern SIFT (Lowe 2004) und SURF (Bay et al. 2008) Algorithmen sehr markante Merkmale, die mit einem entsprechenden Deskriptor (64- oder 128-dimensionaler Vektor) zu jedem extrahierten Bildmerkmal invariant gegenüber Bildmaßstab und -rotation sind. Entsprechende Punkte werden dann gefunden, indem man die Deskriptoren in allen möglichen Bildkombinationen in einer vollständigen Analyse vergleicht. Strenge, aber langsame Verfahren (z.B. quadratisches Matching) oder aber schnelle Methoden als Näherung (z.B. kd-tree Suche) können hierbei angewendet werden. Um mögliche Ausreißer zu eliminieren, werden robuste Schätzer (RANSAC, MAPSAC, LMedS, etc.) zur Überprüfung der Epipolar-Bedingung eingesetzt, die in der fundamentalen Matrix  $F$ , in der essentiellen Matrix  $E$  (falls innere Orientierungsparameter bekannt sind) oder in dem trifokalen Tensor  $T$  (im Falle von

Bildtriplets) verkapselt sind. Sobald entsprechende Punkte für jedes Bildpaar (oder Bildtripel) verfügbar sind, werden sie in ihren Bahnen (Tracks) abgelegt und der Vergleich der Zahlenwerte (Pixel) aller Bildpunkte ergibt den Satz von Bildkorrespondenzen für den gesamten Block oder für die Bildsequenz.

### **3.2 Bündelblockausgleichung**

Sobald die Bildkorrespondenzen gefunden sind, werden ihre 3D-Objektkoordinaten mittels einer Bündelausgleichungsmethode berechnet. Die Methode der Bündelausgleichung wurde bereits in den 1950ziger Jahren in der Photogrammetrie konzipiert (Brown 1976) und sie ist seit den 1980ziger Jahren die Standardlösung für die 3D-Objektrekonstruktion. Während der 1990ziger Jahre ist sie in zunehmendem Maße auch im Bereich der Computer Vision eingesetzt worden (Triggs et al. 1999, Lourakis & Argyros 2009, Agarwal et al. 2010). Eine Bündelausgleichung ist ein Optimierungsproblem für die gesuchten Parametern des 3D-Objektes und der Kamera (Translation, Rotation und eventuell die inneren Orientierungsparameter) zwecks Bestimmung einer 3D-Rekonstruktion, die unter bestimmten Annahmen optimal ist: wenn die Fehler in den Bildpunktmesungen der Gauß'schen Normalverteilung entsprechen, dann ist die Bündelausgleichung eine Schätzung mit der größten Wahrscheinlichkeit. Die optimale Schätzung wird gefunden, indem man eine Kostenfunktion herabsetzt, die den Fehler der Modellanpassung quantitativ bestimmt und gleichzeitig die extrahierten 2D-Merkmale in 3D und die Kameraparameter berechnet. Im Allgemeinen wird eine Bündelausgleichung als nicht-lineares Kleinste-Quadrate-Problem mit einer Kostenfunktion formuliert, dessen Fehler für die Merkmals-Rückprojektion quadratisch ist (d.h. die Summe der quadratischen Differenzen zwischen gemessenen und zurück-projizierten Bildpunkten). Die numerische Lösung zum Problem der Funktionsminimierung erfolgt im Allgemeinen mit Methoden wie Levenberg-Marquardt, Gauß-Newton oder Gauß-Markov.

## **4 Verwendete Datensätze und untersuchte Software**

Die verwendeten Bilddatensätze (siehe Tab. 1 und Abb. 2) repräsentieren sowohl echte Objektaufnahmen als auch Labordatensätze, die für Qualitäts- und Genauigkeitsanalysen bestens geeignet sind. Dabei wurden verschiedene Digitalkameras mit unterschiedlicher Bildauflösung für die Objektaufnahme eingesetzt, während als Referenzdaten je nach Objekt kalibrierte Maßstäbe, Passpunkte, kalib-

rierte Parameter der inneren Orientierung und bekannte Objektformen zur Verfügung standen. Schwierigere Objekte mit wiederholenden Mustern („Cerere“) und eher texturlosen Szenen („Alter Bahnhof Bergedorf“) wurden auch untersucht. Es handelt sich bei den verwendeten Datensätzen um terrestrische Daten; nur beim „Leuchtturm Neuwerk“ wurden zusätzlich auch Helikopteraufnahmen verwendet.

Tab. 1: Charakteristik der verwendeten Datensätze für die Evaluierung der Softwarepakete zur automatisierten Generierung von 3D-Punktwolken.

Datensatz	# Fotos	Auflösung (pixel)	Brennweite (mm)	Dim. L x B x H (m)	Referenzdaten
Würfel	24	6048x4032	50	0.1x0.1x0.1	Maßstab, kalib. Kamera, bek. Form
Kugel	67	3008x2000	35	0.2	Maßstab, bekannte Form
Leuchtturm	119	3008x2000, 4256x2832	18, 85	14x14x40	Passpunkte
Cerere	212	6048x4032	35, 50	15x30x10	-
Bahnhof	147	4256x2832, 4288x2848	20, 28	15x15x10	Passpunkte
Navona	92	4000x3000	6	50x250x15	kalib. Kamera

Folgende Softwarepakete für die automatisierte Generierung von 3D-Punktwolken wurden für die hier vorgestellten Untersuchungen verwendet:

- *Agisoft Photoscan* wurde als kommerzielle und preiswerte (low-cost) Software von Agisoft LLC in St. Petersburg (Russland) entwickelt. Es können damit großen Bilddatensätze orientiert werden, um daraus dichte 3D-Punktwolken zu generieren. Aus kommerziellen Gründen gibt es kaum Informationen über die verwendeten Algorithmen.

- *Apero* wurde als open-source Software durch das Matis-Labor des französischen IGN (Institut Géographique National) in Paris entwickelt, mit der eine Näherungslösung durch Methoden aus der Computer Vision und eine strenge photogrammetrische Lösung kombiniert werden, um optimale Ergebnisse mit der klassischen Gauß-Newton Ausgleichung zu erhalten (Pierrot-Deseilligny & Cléry 2011). Diese Software benutzt einen geänderten SIFT++-Merkmalsextraktor (Vedaldi & Fulkerson 2010) und sie erlaubt dem Anwender zwischen einigen Kameramodellen (Parameter von Brown, Fisheye, etc.) zu wählen. *Apero* wird im Allgemeinen mit *MicMac* verwendet, eine open-source Software für Dense Image Matching (Pierrot-Deseilligny & Paparoditis 2006).

- *Bundler* wurde von der University of Washington in Seattle und Microsoft mit dem Ziel entwickelt, eine 3D-Szene unter Verwendung einer großen Anzahl von

Bildern, die vom Internet mit unbekanntem Kameraparametern heruntergeladen werden, zu rekonstruieren (Snavely et al. 2008). Die Software verwendet eine typische SfM-Methode zusammen mit RANSAC, mit der die F-Matrix der extrahierten SIFT-Merkmale geschätzt wird und mögliche Ausreißer für jedes Bildpaar eliminiert werden. In der anschließenden Ausgleichung werden neun Kameraparameter (jeweils drei Translationen und Rotationen, Brennweite und zwei radial-symmetrische Verzeichnungsparameter) für jedes Bild bestimmt.

- *Photosynth* ist eine als Webservice optimierte Version von Bundler, die durch Microsoft Live Labs freigegeben wurde. Sie funktioniert wie eine Blackbox, in dem der Anwender die Bilder auf eine Internet-Website laden und die fertige Punktwolke anschließend herunterladen kann.

- *VisualSfM* wurde durch die Universität von Washington & Google Inc. entwickelt. Es ist eine optimierte SfM-Software mit graphischer Benutzeroberfläche, die mit Multicore-Architektur und auf der Graphics Processing Unit (GPU) arbeitet. Die Software benutzt mit dem SiftGPU seinen eigenen Merkmalsextraktor (Wu 2007) und eine Bündelausgleichung für Multicore-Verarbeitung (Wu et al. 2011), die den klassischen Levenberg-Marquardt-Algorithmus mit den Preconditioned Conjugate Gradients (Wu 2011) kombiniert. In einer Kleinste-Quadrate-Ausgleichung werden acht Kameraparameter (drei Translationen, drei Rotationen, die Brennweite und ein radial-symmetrischer Verzeichnungsparameter) für jedes Bild berechnet.



Würfel



Kugel



Leuchtturm Neuwerk



Cerere



Alter Bahnhof Bergedorf



Navona

Abb. 2: Abbildungen der untersuchten Datensätze.



## 5 Resultate, Vergleiche und Analysen

### 5.1 Visuelle Analyse der Orientierungsergebnisse

Eine schnelle aber oft akzeptable Analyse der Ausgleichungsergebnisse ist eine visuelle Betrachtung der berechneten Kamerastandpunkte und der 3D-Objektpunktkoordinaten. In den Datensätzen „Cerere“ und „Navona“ (Abb. 3 und 4) werden durch die visuelle Analyse der Ausgleichungsergebnisse einige Probleme der Softwarepakete mit SfM bei langen und komplexen Bildsequenzen aufgedeckt. Das Fehlverhalten und die Divergenz der Bündelausgleichung können durch verschiedene Gründe wie (i) Schätzung nur eines Parametersatzes der inneren Orientierung für jedes Bild, (ii) sequentielle Bündelausgleichung und keine globale Ausgleichung, (iii) falsche Initialisierung der Parameter und (iv) unkontrollierte Fehlerfortpflanzung, etc. verursacht werden.

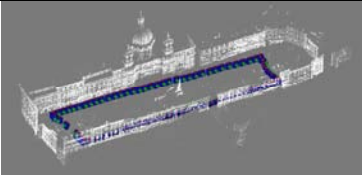
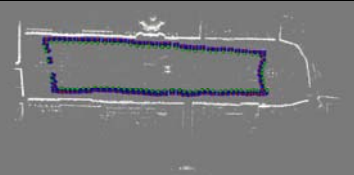
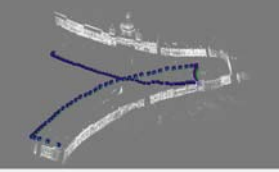
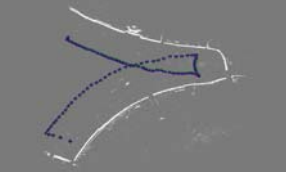
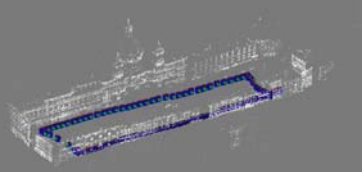
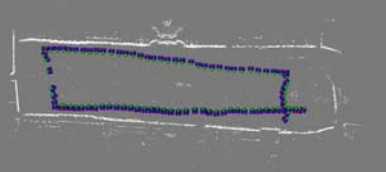
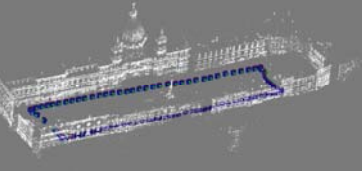
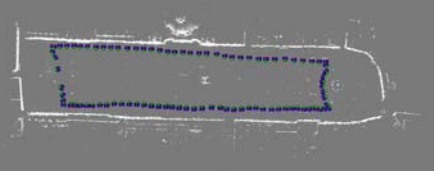
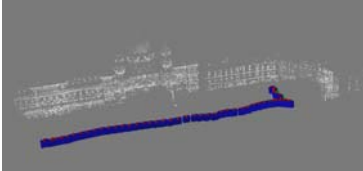
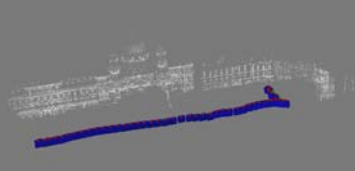
NAVONA		
		<b>AGISOFT:</b> Orientierte Bilder: 92/92 # 3D-Punkte: 193451
		<b>PHOTOSYNTH:</b> Orientierte Bilder: 92/92 # 3D-Punkte: 71068
		<b>BUNDLER:</b> Orientierte Bilder: 92/92 # 3D-Punkte: 73020
		<b>APERO:</b> Orientierte Bilder: 92/92 # 3D-Punkte: 334908
		<b>VISUALSFM:</b> Orientierte Bilder: 56/92 # 3D-Punkte: 21334

Abb. 3: Visuelle Analyse der berechneten Orientierungsdaten des Bildverbandes Navona (92 Bilder, 12 Mpx, Brennweite 6 mm): die korrekte Schließung des Bildverbandes vs. Divergenz und Fehlverhalten der Bündellösung.

### CERERE

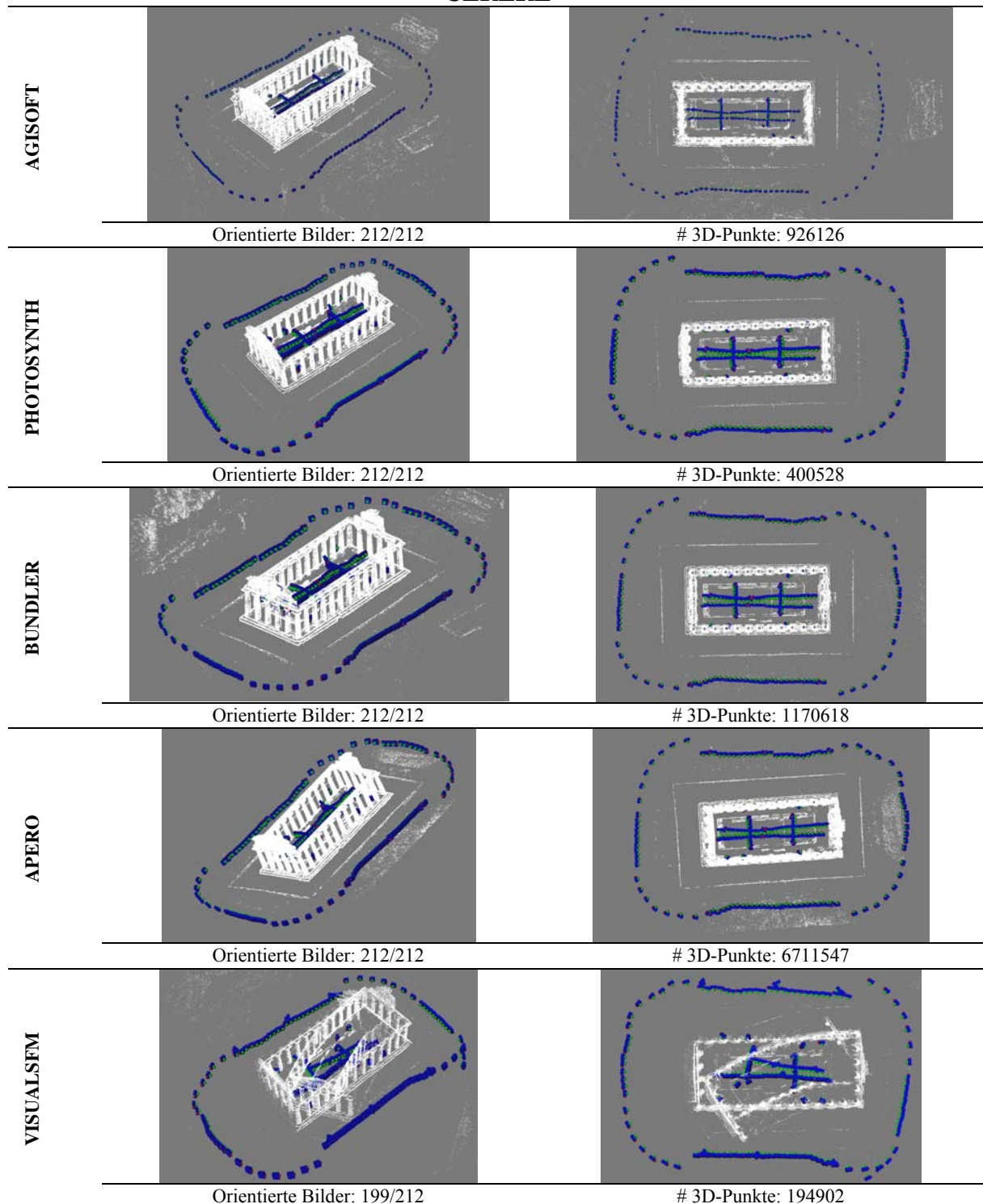


Abb. 4: Visuelle Analyse der berechneten Orientierungsdaten des Bildverbandes "Cerere" (212 Bilder, 24 Mpx, Brennweite 35 mm/50 mm): korrekte Verbindung internes/externes Netzwerk vs. Divergenz und Fehlverhalten der Bündellösung.

## 5.2 Vergleich der kalibrierten Parameter der inneren Orientierung

Die Datensätze „Würfel“ und „Navona“ beinhalten kalibrierte Parameter der inneren Orientierung, die von einem räumlich gut angeordneten Bildverband und einer photogrammetrischen Bündelausgleichung abgeleitet wurden. Die kalibrierten Werte sind mit denen aus den verschiedenen eingesetzten Softwarepaketen verglichen worden, um deren Leistungspotential und Zuverlässigkeit bei der Selbstkalibrierung zu überprüfen.

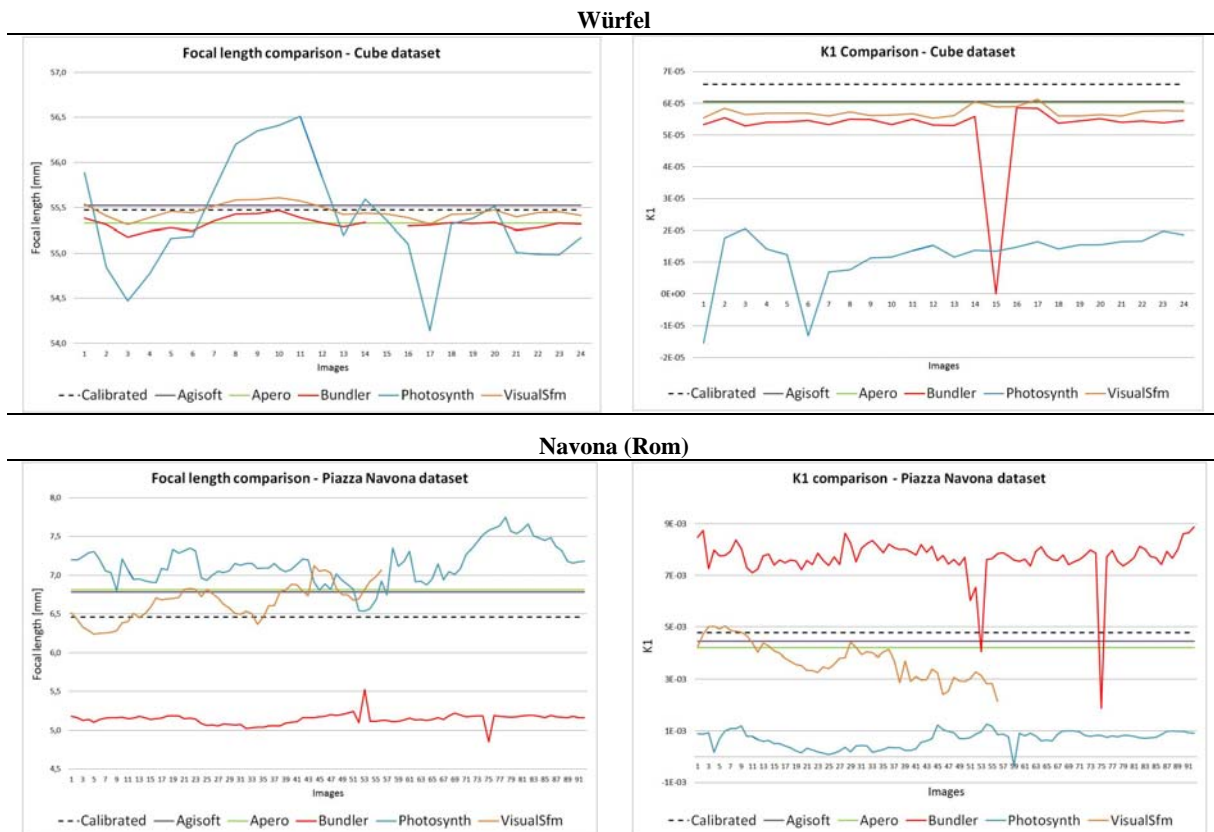


Abb. 5: Vergleich zwischen den kalibrierten Parametern der Inneren Orientierung (IO) und denen aus der Selbstkalibrierung abgeleitet. Die Instabilität der Inneren Orientierung ist eindeutig sichtbar für die Software mit SfM.

Abb. 5 zeigt für beide Datensätze das Verhalten der Brennweite und des Wertes  $K_1$  (radial-symmetrische Verzeichnung) in den verschiedenen Softwarepaketen im Vergleich. Die Software mit SfM berechnet für jedes Bild andere Parameter der inneren Orientierung, was zu sehr merkwürdigen Oszillationen der Parameter führt. Dieses ist in der Photogrammetrie ziemlich ungewöhnlich, da beim Einsatz einer einzigen Kamera für die Objektaufnahme auch nur ein Satz für die

Parameter der inneren Orientierung in der Bündelausgleichung berechnet wird. Diese Instabilität ist insbesondere für den Datensatz „Würfel“ seltsam, da die verwendete Geometrie des Bildverbandes eigentlich sehr geeignet für eine Selbstkalibrierung ist (Abb. 5 oben). Außerdem führte möglicherweise die Tatsache, dass nur ein Satz der inneren Orientierungsparameter für jedes Bild berechnet wurde, zu falschen um die Verzeichnung korrigierten Bildern.

Tab. 2: Der Datensatz „Würfel“ – Vergleich der Resultate von der untersuchten Software und der Referenzorientierung.

Würfel						
	Referenz	Agisoft	Photosynth	Bundler	Apero	VisualSfM
$\Delta$ kalib. Maßstäbe	-	0.301	-0.547	0.220	-0.379	0.258
RMSE [pix]	0.325	2.212	7.149	3.961	3.440	0.597
$\sigma_{XYZ}$ [mm]	0.021	0.143	0.611	0.365	0.141	0.047

Tab. 3: Der Datensatz „Kugel“ – Vergleich der Resultate von der untersuchten Software und der Referenzorientierung.

Kugel						
	Referenz	Agisoft	Photosynth	Bundler	Apero	VisualSfM
$\Delta$ kalib. Maßstäbe	-	-0.070	-0.547	0.220	-0.127	0.156
RMSE [pix]	0.019	0.372	7.149	3.961	3.440	0.860
$\sigma_{XYZ}$ [mm]	0.021	0.031	0.611	0.365	0.141	0.071

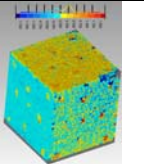
### 5.3 Quantitative Analyse durch Objektformen und Strecken

Die Datensätze „Würfel“ und „Kugel“ enthalten (i) kalibrierte Maßstäbe (78,2 mm und 577,8 mm), (ii) unterschiedlich kodierte Zielmarken und (iii) Objekte mit bekannten geometrischen Formen. Diese Informationen werden verwendet, um das Leistungspotential der verwendeten Softwarepakete für die Bildorientierung geometrisch zu beurteilen. Beide Datensätze (Bildsequenzen) werden zuerst mit einer photogrammetrischen Software bearbeitet, um daraus ein Referenzkoordinatensystem und Referenzdaten (3D-Koordinaten der Zielmarken, Eckpunkte des Würfels, Radius und Zentrum der Kugel) abzuleiten. Anschließend wurden die Ergebnisse der verschiedenen Softwarepakete in das Referenzkoordinatensystem transformiert und so skaliert, um (i) den Längenunterschied des Maßstabes ( $\Delta$ ), (ii) Projektionsfehler der gemessenen Bildkoordinaten (RMSE) und (iii) die Standardabweichung der berechneten Objektkoordinaten ( $\sigma_{XYZ}$ ) zu berechnen (Tabelle 2 und 3). Schließlich wurde mit den bekannten Objektkörpern (ebene Flächen des Würfels und runde Form der Kugel) eine

Abweichungsanalyse durchgeführt, um daraus statistische Aussagen für einen geometrischen 3D-Vergleich (Tabelle 4 und 5) ableiten zu können.

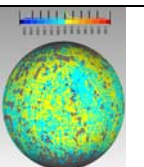
Tab. 4: Der "Würfel" Datensatz – Abweichungen zum idealen Körper.

Würfel [mm]					
	Agisoft	Photosynth	Bundler	Apero	VisualSfM
Ø Abw.	0.090	0.081	0.126	0.117	0.115
Std. abw.	0.401	0.557	0.517	0.425	0.381



Tab. 5: Der "Kugel" Datensatz – Durchschnittliche Abweichung als durchschnittliche Distanz zwischen 3D-Punkten und der idealen Kugel. Die Standardabweichung stellt die Streuung der generierten 3D-Punkte dar.

Kugel [mm]					
	Agisoft	Photosynth	Bundler	Apero	VisualSfM
Ø Abw.	-0.054	0.024	0.033	0.022	0.004
Std. abw.	1.247	1.110	0.789	0.754	0.485



Tab. 6: Ergebnisse für den Datensatz "Alter Bahnhof Bergedorf" – Vergleich der durchschnittlichen Abweichung und Standardabweichung der Passpunkte.

Alter Bahnhof Bergedorf (Hamburg)					
	Agisoft	Photosynth	Bundler	Apero	VisualSfM
Orientierte Bilder	147/147	146/147	138/147	147/147	144/147
# 3D Punkte	157307	33119	54177	552480	22539
Ø Abw. (m)	1.101	0.083	5.671	0.013	0.102
Std. abw. (m)	0.573	0.066	3.648	0.006	0.101

## 5.4 Quantitative Analyse über Passpunkte

Die Datensätze „Alter Bahnhof Bergedorf“ und „Leuchtturm Neuwerk“ verfügen jeweils über ein 3D-Netz von Passpunkten, die benutzt wurden, um die Genauigkeit der Bündelausgleichung und der abgeleiteten 3D-Punktwolken zu validieren. Für jeden Datensatz wurden die gleichen vier Passpunkte verwendet, um durch eine Ähnlichkeitstransformation jedes Projekt in das gleiche Referenzkoordinatensystem zu transformieren. Die verbliebenen ca. 20 Passpunkte wurden als Kontrollpunkte verwendet, um daraus das RMSEs der berechneten Objektpunkte zu berechnen. Tabelle 6 und Abb. 6 fassen die durchschnittlichen Abweichungen und die Standardabweichungen der Kontrollpunkte für beide Projekte zusammen, während in Abb. 6 zusätzlich die berechneten Kamerapositionen für den Datensatz „Leuchtturm Neuwerk“ dargestellt sind.

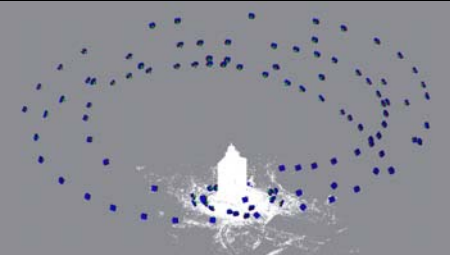
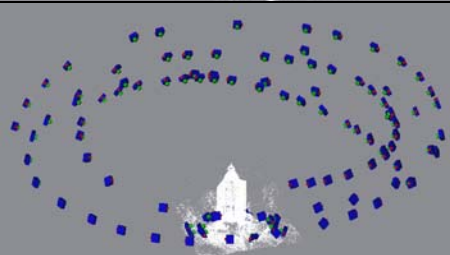
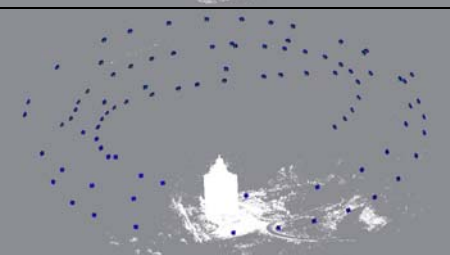
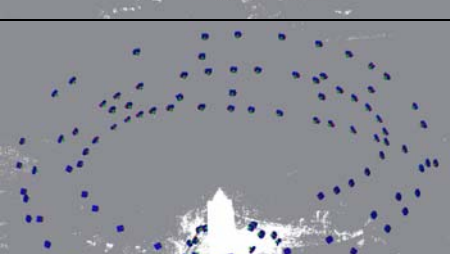
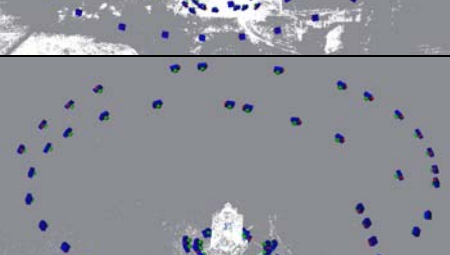
<b>Leuchtturm Neuwerk (Hamburg)</b>		
<b>AGISOFT</b>		Resultat der Orientierung:
		Orientierte Bilder: 118/118
		# 3D-Punkte: 3.374.993
		Kontrollpunktanalyse:
		Mean (m) 0.060 Std. abw. (m) 0.035
<b>PHOTOSYNTH</b>		Resultat der Orientierung:
		Orientierte Bilder: 107/118
		# 3D-Punkte: 86.746
		Kontrollpunktanalyse:
		Mean (m) 0.086 Std. abw. (m) 0.055
<b>BUNDLER</b>		Resultat der Orientierung:
		Orientierte Bilder: 88/118
		# 3D-Punkte : 881.340
		Kontrollpunktanalyse:
		Mean (m) 0.168 Std. abw. (m) 0.128
<b>APERO</b>		Resultat der Orientierung:
		Orientierte Bilder: 118/118
		# 3D-Punkte: 5.742.688
		Kontrollpunktanalyse:
		Mean (m) 0.056 Std. abw. (m) 0.020
<b>VISUALSFM</b>		Resultat der Orientierung:
		Orientierte Bilder: 66/118
		# 3D-Punkte: 30.703
		Kontrollpunktanalyse:
		Mean (m) 0.050 Std. abw. (m) 0.031

Abb. 6: Berechnete Kamerastandpunkte und statistische Angaben zu den Kontrollpunkten für den Bilddatensatz vom „Leuchtturm Neuwerk“.

## **6 Fazit und Ausblick**

Der Beitrag stellt einen kritischen Überblick und eine geometrische Bewertung von einigen Softwarepaketen für die automatisierte Generierung von 3D-Punktwolken aus Bilddaten dar. Dafür sind verschiedene Datensätze mit bekannter Geometrie großer und komplexer 3D-Szenen, mit genauen Passpunkten, kalibrierten Kameras und Maßstäben verwendet worden. Die Ergebnisse zeigen, dass die SfM-Verfahren bei komplexen und langen Bildsequenzen Probleme hinsichtlich Zuverlässigkeit und Wiederholbarkeit aufweisen. Aus photogrammetrischer Sicht kann die Fehlerquelle der Bündellösung auch darin begründet sein, dass SfM oft nur einen Parametersatz der inneren Orientierung für jedes Bild berechnet.

Wenn man sich jedoch die berechneten Objektpunktkoordinaten trotz der ungewöhnlichen Oszillationen für die Parameter der inneren Orientierung (in Abb. 5) anschaut, dann fallen die Ergebnisse überraschenderweise oft positiv aus (z.B. Tabelle 6 und siehe auch Kersten et al. 2012). Vermutlich werden die falsch geschätzten Parameter der inneren Orientierung durch die berechneten Parameter der äußeren Orientierung kompensiert. Weitere Untersuchungen über die Bestimmung der Kameraparameter sind daher notwendig. Wenn jedoch ein robuster Bildverband mit guter Geometrie vorliegt, dann liefern alle Softwarepakete ähnliche Ergebnisse im Hinblick auf die theoretische Genauigkeit der berechneten Objektpunktkoordinaten (RMSEs) und der abgeleiteten Parameter der inneren Orientierung. Da aber ein robuster Bildverband im Feld nicht immer aufgenommen werden kann, muss das Ergebnis kritisch analysiert werden, wenn solche Softwarepakete als Blackbox verwendet werden. Dennoch sind all diese Softwarepakete, insbesondere die SfM verwenden, in der Lage, die Kameraparameter (innere und äußere Orientierung) und die Geometrie der aufgenommenen 3D-Szene mit entsprechendem Skalierungsfaktor zu rekonstruieren. Allerdings kann die nachfolgende Ähnlichkeitstransformation, die im Allgemeinen zur Erzeugung geometrischer Resultate angewendet wird, nur korrekte Ergebnisse erzielen, wenn sich der aufgenommene Bildverband durch eine gute Netzgeometrie auszeichnet. Andernfalls können mögliche geometrische Modelldeformationen nicht kompensiert werden.

## 7 Danksagung

Die Autoren danken Silvio Del Pizzo vom Department der angewandten Wissenschaften der Parthenope Universität Neapel für die Unterstützung bei den Berechnungen der Beispieldaten und für die interessanten Diskussionen über die Ergebnisse.

### Literatur

- Agarwal, S., Snavely, N., Simon, I., Seitz, S.M., Szeliski, R. (2009): Building Rome in a day. Proceedings of International Conference on Computer Vision, Kyoto, Japan.
- Agarwal, S., Snavely, N., Seitz, S.M., Szeliski, R. (2010): Bundle adjustment in the large. Proc. ECCV, Crete, Greece.
- Barazzetti, L., Scaioni, M., Remondino, F. (2010): Orientation and 3D modeling from markerless terrestrial images: combining accuracy with automation. *The Photogrammetric Record*, 25(132), 356–381.
- Barazzetti, L., Mussio, L., Remondino, F., Scaioni, M. (2011): Targetless camera calibration. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 38(5/W16), on CD-ROM.
- Bay, H., Ess, A., Tuytelaars, T., Van Gool, L. (2008): Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3): 346–359.
- Brown, D.C. (1971): Close-Range Camera Calibration. *Photogr. Eng.*, 37(8), 855-866.
- Brown, D.C. (1976): The bundle adjustment - progress and prospects. *Int. Archives of Photogrammetry*, 21(3).
- Clarke, T.A., Fryer, J.G. (1998): The Development of Camera Calibration Methods and Models. *Photogrammetric Record*, 16(91), 51-66.
- Del Pizzo, S., Troisi, S. (2011): Automatic orientation of image sequences in Cultural Heritage. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W16).
- Fitzgibbon, A., Zisserman, A. (1998): Automatic 3D model acquisition and generation of new images from video sequence. Proc. ESP Conf., 1261-1269.
- Fraser, C. (1996): Network design. In *Close Range Photogrammetry and Machine Vision* (K.B. Atkinson Ed.), Cap.9, Whittles Publishing, Caithness, Scotland, U.K.
- Fraser, C. (2012): Automatic camera calibration in close-range photogrammetry. Proc. ASPRS 2012, Sacramento, CA, USA.



- Furukawa, Y., Ponce, J. (2010): Accurate, dense, and robust multi-view stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(8), 1362-1376.
- Gruen, A., Huang, T.S. (2001): Calibration and Orientation of Cameras in Computer Vision. Springer, Berlin/Heidelberg
- Grün, A., Remondino, F., Zhang, L. (2004): Photogrammetric Reconstruction of the Great Buddha of Bamiyan, Afghanistan. *The Photogrammetric Record*, 19(107), 177-199.
- Ganci, G., Handley, H. (1989): Automation in Videogrammetry. *Int. Arch. Photogrammetry and Remote Sensing*, Vol. 32(5), Hakodate, Japan.
- Kersten, Th., Lindstaedt, M., Mechelke, K., Zobel, K. (2012): Automatische 3D-Objektrekonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie. *Publikationen der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation e.V., Band 21, Hrsg. E. Seyfert, CD-ROM, 137-148.*
- Laebe, T., Foerstner, W. (2006): Automatic relative orientation of images. *Proceedings of the 5th Turkish-German Joint Geodetic Days, Berlin*, 6 pages.
- Lourakis, M.I.A., Argyros, A.A. (2009): SBA: A software package for generic sparse bundle adjustment. *ACM Trans. Math. Software*, 36(1).
- Lowe, D.G. (2004): Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- Nister, D. (2004): Automatic passive recovery of 3D from images and video. *IEEE Proc. 2nd Int. Symp. on 3D Data Processing, Visualization and Transmission*, 438-445.
- Pierrot-Deseilligny, M., Paparoditis, N. (2006): A multi-resolution and optimization-based image matching approach: an application to surface reconstruction from SPOT5-HRS stereo imagery. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 36(1/W41).
- Pierrot-Deseilligny, M., Cléry, I. (2011): APERO, an open source bundle adjustment software for automatic calibration and orientation of a set of images. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38(5/W16).
- Pollefeys, M., Koch, R., Van Gool, L. (1999): Self-calibration and metric reconstruction inspite of varying and unknown internal camera parameters. *IJCV*, 32(1), 7-25.
- Remondino, F., Ressel, C. (2006): Overview and experiences in automated markerless image orientation. *Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(3), 248-254.

- Remondino, F., El-Hakim, S. (2006): Image-based 3D modelling: a review. *The Photogrammetric Record*, 21(115), 269-291.
- Remondino, F., Fraser, C. (2006): Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(5), 266-272.
- Snavely, N., Seitz, S.M., Szeliski, R. (2008): Modeling the World from Internet Photo Collections. *International Journal of Computer Vision*, 80(2), 189-210.
- Strecha, C., Pylvanainen, T., Fua, P. (2010): Dynamic and Scalable Large Scale Image Reconstruction. *Computer Vision and Pattern Recognition*, 406-413.
- Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A. (1999): Bundle Adjustment - A Modern Synthesis: Proc. Workshop on Vision Algorithms, Springer-Verlag, 298-372.
- Tuite, K., Tabing, N., Hsiao, D., Snavely, N., Popović, Z. (2011): PhotoCity: training experts at large scale image acquisition through a competitive game. Proc. ACM CHI Conference on Human Factors in Computing Systems.
- Uden, M., Zipf, A. (2012): OpenBuildingModels - Towards a platform for crowdsourcing virtual 3D cities. 7th 3D GeoInfo Conference, Quebec City, QC, Canada.
- Vedaldi, A., Fulkerson, B. (2010): VLFeat - An open and portable library of computer vision algorithms. Proc. 18th annual ACM Intern. Conf. on Multimedia.
- Vergauwen, M., Van Gool, L. (2006): Web-based 3D reconstruction service. *Mach. Vis. Appl.*, 17(6), 411-426.
- Vosselman, G., Maas, H.-G. (2010): *Airborne and Terrestrial Laser Scanning*. Whittles Publishing.
- Wu, C. (2007): SiftGPU: A GPU implementation of Scale Invariant Feature Transform (SIFT). <http://cs.unc.edu/~ccwu/siftgpu> (Zugriff am 30.04.12).
- Wu, C. (2011): VisualSFM: A Visual Structure from Motion System. [www.cs.washington.edu/homes/ccwu/vsfm/](http://www.cs.washington.edu/homes/ccwu/vsfm/) (Zugriff am 30.04.12)
- Wu, C., Agarwal, S., Curless, B., Seitz, S.M. (2011): Multicore Bundle Adjustment. The 24th IEEE Conference on Computer Vision and Pattern Recognition, 3057-3064.