

Konzeption und Realisierung eines neuen portablen 360°-Stereokamerasystems

DAVID HOLDENER¹

Zusammenfassung: Die Nachfrage nach der 3D-Vermessung von Innenräumen nimmt mit dem Trend der Digitalisierung in der Baubranche stetig zu. Eine effiziente Lösung zur Aufnahme des Innenbereichs ist das Mobile Mapping. Bildbasierte Systeme mit 360°-Panoramaabdeckung erlauben eine schnelle Datenerfassung und können zu georeferenzierten 3D-Bilddaten prozessiert und in cloudbasierte 3D-Geoinformationsservices integriert werden.

Das in dieser Arbeit entwickelte Kamerasystem erreicht eine 360°-Stereoabdeckung durch die ringförmige Anordnung von fünf horizontalen Stereobildpaaren. Die kostengünstige Realisierung basiert auf einer Kamerahalterung, welche mit dem 3D-Drucker gefertigt wurde, und aus Action-Kameras mit Fischaugen-Objektiv. Das Stereokamerasystem wurde erfolgreich kalibriert. Mit dem aufgezeigten Auswerteprozess können automatisch Tiefenkarten und Punktwolken aus den Stereobildern gerechnet werden. Ein Vergleich von 3D-Messungen mit Referenzdaten zeigt maximale Abweichungen von 3 cm auf typische Distanzen für Innenaufnahmen von 2-8 m. Trotz der kostengünstigen Implementierung kann so das Potential eines bildbasierten Aufnahmesystems aufgezeigt werden.

1 Einleitung

Bildbasiertes Mobile Mapping ist eine etablierte Methode zur dreidimensionalen Erfassung von Infrastrukturbauwerken wie Straßen und Bahntrassen. Um eine kontinuierliche und vollständige Abdeckung solcher komplexen Umgebungen zu gewährleisten, können die Aufnahmesysteme mit stereoskopisch ausgerichteten Panoramakameras ausgerüstet werden. Mit dem fortschreitenden Trend der Digitalisierung in der Architektur und Bauwirtschaft steigt aber auch die Nachfrage nach einer effizienten und genauen dreidimensionalen Erfassung von Innenräumen. Sei dies zur Grundlagenerfassung für eine spätere Planung oder die Baufortschrittskontrolle während der Ausführung.

Die bestehenden Systeme zur messtechnischen Erfassung von Innenräumen wie beispielsweise von BERGSLI & SCHROTH (2017) präsentiert, basieren häufig auf einer hybriden Aufnahmesensorik. Dabei wird die 3D-Information von Laserscannern erfasst und anschließend mit Bildern ergänzt. Eine Ableitung von georeferenzierten 3D-Bildern und deren Einbindung in cloudbasierte 3D-Geoinformationsservices ermöglicht auch fachfremden Anwendern Auswertungen vorzunehmen. Mittels Dense Image Matching können Tiefenwerte direkt aus Stereobildern gerechnet werden. Dies garantiert, im Gegensatz zum Laserscanning, die räumliche und zeitliche Kohärenz zwischen den Farb- und Tiefendaten eines 3D-Bildes und erübrigt eine Co-Registrierung zwischen Scannern und Kameras (NEBIKER et al. 2015).

¹ Fachhochschule Nordwestschweiz, Institut Vermessung und Geoinformation, Gründenstrasse 40, CH-4132 Muttenz, E-Mail: david.holdener@hotmail.com

Entsprechend ist das Ziel dieser Arbeit die Entwicklung und Evaluierung eines 360°-Stereo-kamerasystems für flächenhafte 3D-Aufnahmen von Innenräumen und eine spätere Einbindung in ein Indoor Mobile Mapping System (IMMS). Dabei ist insbesondere den engen Platzverhältnissen und kurzen Aufnahmedistanzen bei Innenaufnahmen Rechnung zu tragen.

2 Konzeption des Aufnahmesystems

Für Außenaufnahmen existieren bereits fahrzeuggestützte Stereosysteme basierend auf mehrköpfigen Panoramakameras. Die Stereobasis wird dabei entweder vertikal (EARTHMINER 2014) oder horizontal (BLASER et al. 2017) angeordnet. Aus dem Bereich der Robotik kommt der Ansatz von omnidirektionalen Stereopanoramasytemen, bestehend aus zwei Kameras mit katadioptrischen Spiegeln (z.B. LUI & JARVIS 2010). Dabei können jedoch wie bei kommerziellen Systemen für Virtual Reality Anwendungen keine messtechnischen Genauigkeiten im cm-Bereich erreicht werden.

Aus der Präanalyse zur Erreichung einer 360°-Abdeckung mit cm-Messgenauigkeit entstanden die Anforderungen von mindestens fünf Stereobasen für eine optimale Überlappung und eine Basislänge von zirka 50 cm. Am besten kann eine horizontale ringförmige Anordnung der Stereobasen diese Anforderungen erfüllen (siehe Abbildung 1). Wird das System als IMMS über oder um den Kopf getragen, kann eine komplette 360°-Stereoabdeckung mit hoher Genauigkeit erreicht werden.

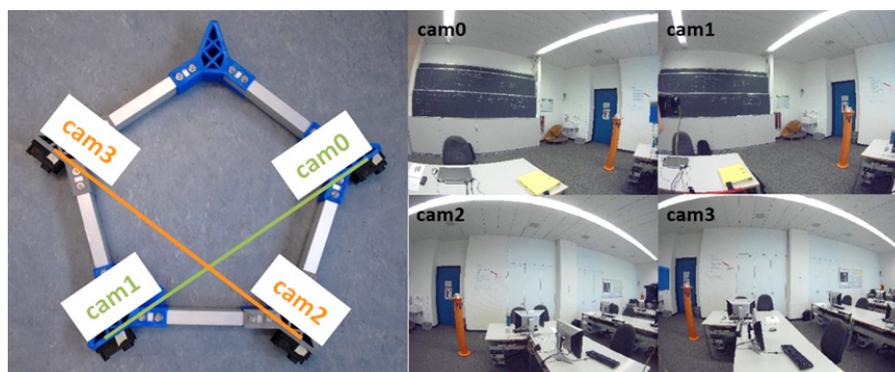


Abb. 1: Umgesetztes Kamera-Rig (links) mit den synchronen Aufnahmen von zwei der fünf Stereopaaren (rechts)

3 Implementierung und Kalibrierung

Die Implementierung des Systems erfolgte mit den kostengünstigen Action-Kameras Git2 von GitUp. Die Kameras haben eine Auflösung von 4608 x 3456 Pixel mit Kammerkonstanten von 3 mm bei einer Sensorpixelgröße von 1.34 μm . Dank der Fischaugen-Objektive wird ein Öffnungswinkel von 120° x 90° ermöglicht. Die Kamera-Halterungen und das Gestell wurden mittels additiver 3D-Drucktechnik hergestellt. Das Aufnahmesystem in der Ausführung, wie in Abbildung 1 ersichtlich, hat einen Gesamtdurchmesser von 70 cm und eine Basislänge von 60 cm. Die synchrone Auslösung der Kameras erfolgt mit dem Einplatinencomputer Raspberry Pi 3 mit zugehörigem Touchscreen. Der Auslösemechanismus mit Pulsweitenmodulation (PWM) ermög-

licht eine mittlere Differenz der Synchronität zwischen den Kameras von $20 \mu\text{s}$ und maximale Abweichungen von $40 \mu\text{s}$.

3.1 Kalibrierung

Die Kalibrierung erfolgte in einem Passpunktfeld mit Bündelblockausgleichung nach ELLUM & EL-SHEIMY (2017) und dem äquidistanten Kameramodell für Fischaugen-Objektive (ABRAHAM & FÖRSTNER, 2015). Die Implementierung konnte dabei vom Mobile Mapping System von BLASER et al. (2017) übernommen werden. Die resultierenden Parameter der Kalibrierung sind

- die inneren Orientierungsparameter (IOP) für jede Kamera;
- die relativen Orientierungsparameter (ROP) zwischen rechter und linker Kamera jedes Stereosystems;
- Bore-sight-Kalibrierung mit Verschiebungsvektorkomponenten (Lever Arm) und Verdrehungswinkel (Misalignment) zwischen den linken Kameras jedes Stereosystems und der Master-Kamera 0
- Äußere Orientierungsparameter (EOP) der Master-Kamera pro Epoche

Alle Parameter wurden in einer einzelnen Kalibrierung in einem Passpunktfeld (siehe Abb. 2) bestimmt. Dazu wurden in acht Epochen (32 Bilder) 3952 Bildmessung exklusive 61 Ausreißern durchgeführt. Die runden Passpunktmarken wurden mit einem Ellipsenfit nach dem Prinzip der kleinsten Quadrate angemessen.



Abb. 2: Kalibrierfeld

Die resultierenden Parameter haben allesamt eine Standardabweichung von unter 0.7 mm respektive 30 mdeg . Die vollständige Liste der erreichten Genauigkeiten ist in Tabelle 1 aufgezeigt. Auf Grund der teilweise hohen Korrelation zwischen den Parametern ist eine unabhängige Analyse der einzelnen Werte jedoch nicht sinnvoll.

Weiter wurde eine durchschnittliche Größe der Residuen von 0.7 Pixel erreicht. Die Residuen zeigen auf manchen Bildern ein systematisches Verhalten auf. Dies lässt auf den Rolling-Shutter-Effekt der Kameras oder eine ungenügende Synchronität schließen. Zudem nehmen die Residuen mit zunehmendem Bildradius zu, da dort die Bildverzerrungen grösser sind und im Fall der günstigen Action-Kameras nicht mehr so gut mit dem mathematischen Modell übereinstimmen.

Tab. 1: Empirische Genauigkeiten der Kalibrierung

Parameter	Kamera	St. Abw. der Lage [mm]			St. Abw. der Orientierung [mdeg]		
		X	Y	Z	ω	φ	κ
ROP Basis 1	0 - 1	0.3	0.3	0.6	16	24	3
ROP Basis 2	2 - 3	0.4	0.3	0.7	12	19	2
Boresight	0 - 2	0.5	0.3	0.5	31	20	28
EOP (mittel 8 Epochen)		0.4	0.4	0.4	16	12	11

3.2 Auswerteprozess

Die Stereobildauswertung konnte ebenfalls von BLASER et al. (2017) adaptiert werden. Das Ziel der Prozessierung sind georeferenzierte 3D-Bilder im äquidistanten Modell und zugehörigen Tiefenwerten, im Optimalfall für jedes Pixel. In einem ersten Schritt werden dazu die Bilder nach den IOP und ROP korrigiert und mittels Epipolar-Rektifizierung in den Stereonormalfall gebracht. Diese Modellkonvertierung zur Normalisierung erfolgt nach dem epipolar äquidistanten Modell von ABRAHAM & FÖRSTNER (2005). Diese Konvertierung ist Voraussetzung für das anschliessende Dense Image Matching mit der Software SURE (ROTHERMEL et al., 2012). Die Disparitätskarte und deren Werte werden dann wieder in das äquidistante Projektionsmodell rückkonvertiert und stimmen so mit den Farbbildern überein. Aus diesen 3D-Bildern ist auch eine Ableitung der Punktwolke möglich.

4 Systemevaluierung

Zur Evaluierung der relativen Genauigkeit wurden manuelle Messungen im Passpunktfeld durchgeführt. Zudem wurden die abgeleiteten Punktwolken mit einem Referenzscan von einem terrestrischen Laserscanner (TLS) verglichen.

4.1 Manuelle Punktmessung

Mit Bildern des Passpunktfelds (Abb. 2) wurde die relative Genauigkeit innerhalb des Stereobildpaares untersucht. Die Bildmessungen fanden dabei in den nach IOP und ROP korrigierten Bildern mit einem Ellipsenfit nach den kleinsten Quadraten statt. Zwei Passpunkte formen zusammen eine 3D Referenzdistanz die als fehlerlos betrachtet werden kann. Bei einem Vergleich der photogrammetrisch gemessenen Distanz mit deren Referenz beeinflussen somit nur die Unsicherheiten von IOP, ROP und der Bildmessung sowie die Güte des mathematischen Modells die erreichbare Genauigkeit.

Auf typische Distanzen für Innenaufnahmen von 2 - 8 m wurden mit der 60 cm langen Basis in drei Epochen 78 3D-Distanzen gemessen. Die resultierende mittlere Differenz zwischen der photogrammetrischen bestimmten Distanz und Referenzdistanz beträgt 5.7 mm mit einer Standardabweichung der Differenzen von 5.6 mm (1σ). Davon ausgeschlossen sind zwei Ausreißer grösser als 70 mm. Die verbleibende maximale Abweichung beträgt 33 mm.

Der Einfluss der Distanz zwischen Kamera und Messobjekt sowie des Bildradius auf die resultierende Messgenauigkeit wurde vertieft analysiert. Auf Grund der schlechter werdenden Schnittgeometrie mit zunehmender Messdistanz wird auch eine Zunahme der Messunsicherheit erwartet. Das durchgeführte Experiment zeigt jedoch lediglich eine leichte nicht signifikante Abnahme der Messgenauigkeit auf (Abb. 3, links). Eine Abnahme der Messgenauigkeit wird im Falle von

Fischaugen-Objektiven auch mit zunehmendem Bildradius erwartet. Dies gilt insbesondere für Action-Kameras mit kostengünstiger Optik auf Grund der zunehmenden Diskrepanz zwischen mathematischem Modell und effektiven Eigenschaften des Objektivs. Die beobachteten Differenzen innerhalb dieses Experiments zeigen eine Zunahme der Messunsicherheit mit zunehmendem Bildradius auf (Abb. 3, rechts).

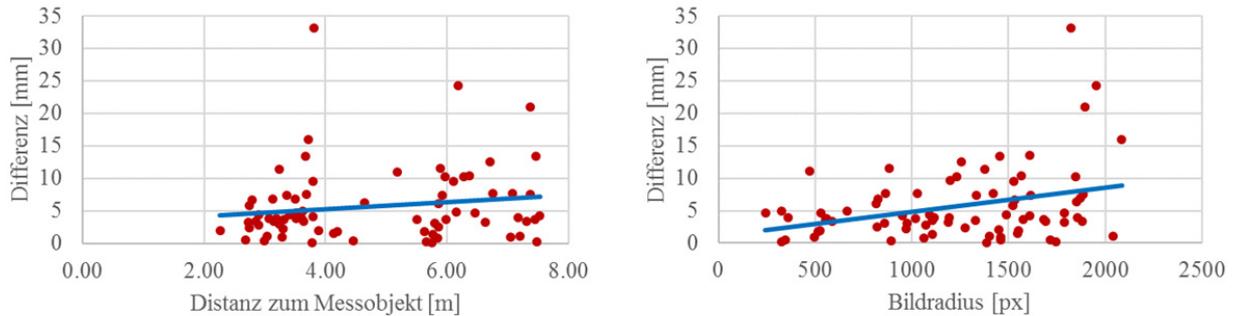


Abb. 3: Differenzen zwischen photogrammetrisch bestimmten 3D-Distanzen und Referenzdistanzen bezüglich der Distanz zum Messobjekt (links) und dem Bildradius (rechts)

4.2 Vergleich der Punktwolken

Zur Evaluierung des gesamten Auswerteprozesses wurde die Punktwolke des Stereosystems mit einer Referenzpunktwolke von einem TLS verglichen. Der Referenzscan wurde mit einer Leica P20 vom gleichen Standort wie die Kameraaufnahmen erfasst. Die Koregistrierung beider Punktwolken erfolgte mit dem ICP-Algorithmus. Die Testszene repräsentiert einen typischen Innenraum mit texturierten und untexturierten Flächen, Möblierung und Aufnahmedistanzen von 1.5 - 7 m. Der Vergleich zeigt erhebliche Differenzen im Bereich von ebenen und untexturierten Flächen wie Wänden und Decken (Abb. 4). Auf Grund der fehlenden Merkmale für eine saubere Zuordnung von Verknüpfungspunkten ist dies ein typisches Problem des Dense Image Matchings.

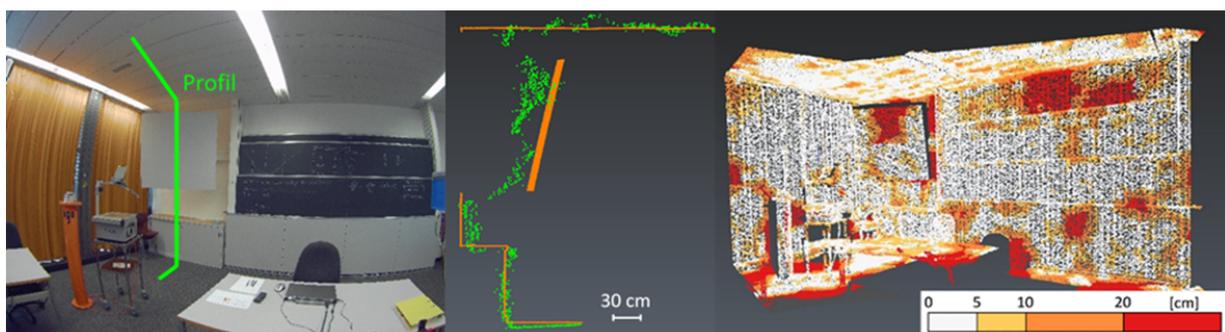


Abb. 4: Testszene der Punktwolkenevaluierung mit einem Profil und einem flächenhaften Vergleich

Weiter konnten auch in der Punktwolke systematische Abweichungen mit zunehmendem Bildradius festgestellt werden. Dank dem weiten Öffnungswinkel der Kameras sind die Überlappungen zwischen den Stereobereichen jedoch groß genug um die signifikant abweichenden Bereiche an den Bildrändern jeweils zu ersetzen.

5 Fazit & Ausblick

Die präsentierte Realisierung eines bildbasierten 3D-Aufnahmesystems bestehend aus einem Ring mit fünf horizontal angeordneten Stereobasen kann Innenräume mit einer 360°-Abdeckung dreidimensional erfassen. Die kostengünstige Realisierung erfolgte mit einem Rig aus dem 3D-Drucker, Action-Kameras (GitUp Git2) mit Fischaugen-Objektiven und einem Einplatinencomputer (Raspberry Pi 3) mit dazugehörigem Touchscreen. Die Kalibrierung des Systems mit zwei der fünf Stereobasen erzielte Genauigkeiten der ROP von unter 0.7 mm respektive 30 mdeg. Der Auswerteprozess für Fischaugen-Stereobildpaare nach ABRAHAM & FÖRSTNER (2005) wurde erfolgreich auf die Aufnahmen des Prototyps angewandt.

Die Nachteile der kostengünstigen Umsetzung mit Action-Kameras sind der Rolling-Shutter der Kameras, die ungenaue Synchronisierung und die teilweise großen Differenzen der Optik zum mathematischen Modell. Nichtsdestotrotz zeigte die Evaluierung mit Genauigkeiten im cm-Bereich das Potential des Systems für Innenaufnahmen auf. Die Tiefenkarten und Punktwolken zeigen typische Abweichungen für Bereiche mit großem Bildradius. Dank einer genügenden Überlappung können solche Bereiche jedoch ignoriert werden.

Durch die Integration von Positionierungssensorik könnte das präsentierte Aufnahmesystem für weiterführende Analysen als ganzheitliches IMMS evaluiert werden. Eine Lösung für die vielen untexturierten Flächen ist ebenfalls Teil der ausstehenden Untersuchungen.

6 Literaturverzeichnis

- ABRAHAM, S. & FÖRSTNER, W., 2005: Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **59**(5), 278-288.
- BERGSLI, M. & SCHROTH, G., 2017: NavVis - Enabling Digital Value Creation Indoors. FIG Working Week. Helsinki, Finnland.
- BLASER, S., NEBIKER, S. & CAVEGN, S., 2017: System Design, Calibration and Performance Analysis of a Novel 360° Stereo Panoramic Mobile Mapping System. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. ISPRS Workshop. Hannover, Deutschland, 207-213.
- EARTHMINE, 2014: earthmine Mars Collection System. http://www.earthmine.com/html/products_mobile.html, [Zugriff vom 09.11.2017].
- ELLM, C. & EL-SHEIMY, N., 2002: The calibration of image-based mobile mapping systems. *Proceedings of the 2nd Symposium on Geodesy for Geotechnical and Structural Engineering*. The International Association of Geodesy (IAG). Berlin, Deutschland.
- LUI, W.L.D. & JARVIS, R., 2010: Eye-Full Tower: A GPU-based variable multibaseline omnidirectional stereovision system with automatic baseline selection for outdoor mobile robot navigation. *Robotics and Autonomous Systems*, **58**(6), 747-761.
- NEBIKER, S., CAVEGN, S. & LOESCH, B., 2015: Cloud-Based Geospatial 3D Image Spaces – A Powerful Urban Model for the Smart City. *ISPRS International Journal of Geo-Information*, **4**(4), 2267-2291.
- ROTHERMEL, M., WENZEL, K., FRITSCH, D. & HAALA, N., 2012: SURE: Photogrammetric Surface Reconstruction from Imagery. *Proceedings LC3D Workshop*. Berlin, Deutschland.