

---

# Automatische Bestimmung der Kameraorientierung eines LiDAR Mobile Mapping Systems

Janetta WODNIOK, Sabine HOFMANN, Claus BRENNER und Thomas LUHMANN

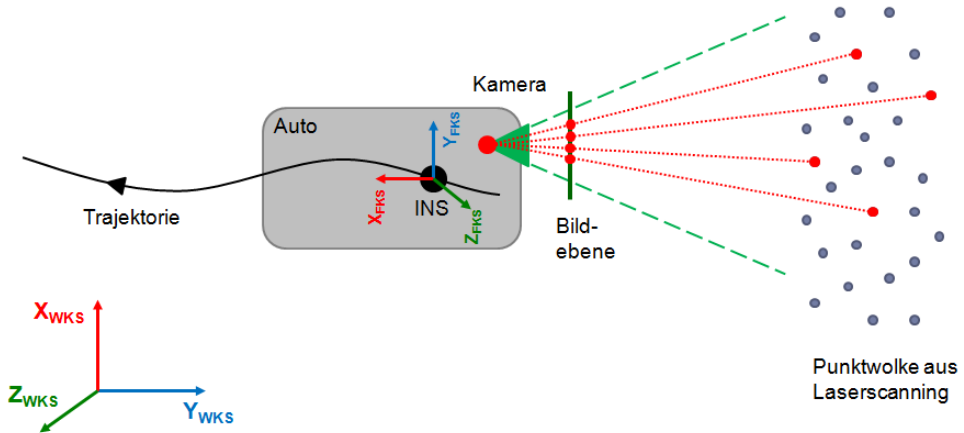
## Zusammenfassung

In diesem Beitrag wird die automatische relative Orientierung zwischen den Kameras und den Laserscannern eines Mobile Mapping Systems untersucht. Dadurch entfällt einerseits der Zeitaufwand einer manuellen Kalibrierung. Zugleich wird damit, aufgrund der Vielzahl von Bildern, erreicht, dass die Kalibrierparameter trotz Fehler in der GPS/IMU Trajektorie mit ausreichender Genauigkeit geschätzt werden können. Dazu wird die Zuordnung von Intensitätsbildern (aus der Punktwolke) zu optischen Bildern (der Kameras) mittels Interestpunkten vorgenommen. Als Interestpunkte dienen SURF Features, welche sowohl in den Scandaten anhand der Intensitätsbilder als auch in den optischen Bildern extrahiert werden. Anhand der bekannten 3D-Koordinaten der Interestpunkte wird die Orientierung der Kamera mittels Rückwärtsschnitt bestimmt. Hieraus ergibt sich in Verbindung mit der bekannten Orientierung des Systems zum Aufnahmezeitpunkt die relative Orientierung zwischen Kamera und Laserscanner. Diese wird über mehrere Standpunkte berechnet, um Fehler in der GPS/IMU Trajektorie auszugleichen. Der Beitrag erläutert die Vorgehensweise und zeigt die Ergebnisse, welche mit einem Riegl VMX-250-CS6 Mobile Mapping Datensatz der Innenstadt von Hannover erzielt wurden.

## 1 Einführung

### 1.1 Motivation

Mobile Mapping Systeme werden in der Praxis immer häufiger eingesetzt. Die dreidimensionale Erfassung der Umgebungsinformation kann etwa zur Visualisierung, zur Erfassung von Straßeninfrastrukturanlagen oder zur Erstellung hochgenauer Karten für Fahrerassistenzsysteme genutzt werden. Werden in einem solchen Mobile Mapping System verschiedene Sensoren zur Umgebungserfassung eingesetzt, z.B. Laserscanner und Kameras, ist es erforderlich, die relative Orientierung zwischen den Sensoren zu bestimmen. Hierfür existieren typischerweise manuelle Kalibrierprozeduren, für welche die Identifikation und Zuordnung korrespondierender Punkte im Scan und im Bild durchgeführt werden muss. Dieses Vorgehen hat jedoch den Nachteil, dass es nicht nur zeitaufwändig ist, sondern sich zudem die Kalibrierung auf eine einzige Aufnahme stützt, also von der dort vorhandenen Verteilung der Punkte abhängig ist. Da das verwendete Bild und die verwendeten Scanpunkte außerdem zu unterschiedlichen Zeitpunkten aufgenommen wurden, beeinflussen GPS/IMU Fehler in der Bestimmung der Trajektorie direkt das Ergebnis.



**Abb. 1:** Anhand der erzeugten Punktwolke soll die relative Orientierung zwischen Kamera und der INS-Laserscanner-Einheit automatisch bestimmt werden.

## 1.2 Verwandte Arbeiten

Zur Bestimmung der relativen Orientierung von Punktwolke und Kamera wurden bereits mehrere Methoden entwickelt. SWART u. a. (2011) beschreiben eine Methode, Panoramabilder mit Punktwolken über eine dynamische Version des Iterative Closest Point (ICP) Algorithmus zu registrieren. Der dynamische ICP besteht aus dem starren ICP-Algorithmus und einer Bündelausgleichung. Es handelt sich um eine 3D zu 3D Registrierung der rekonstruierten dünnen Punktwolke aus den Bildern mit der Punktwolke aus Laserscanning.

LI-CHEE-MING & ARMENAKIS (2010) haben das Problem der Registrierung auf das zweidimensionale beschränkt. Das Verfahren basiert auf Punktverknüpfungen zwischen optischem Bild und Intensitätsbild. Zur Erzeugung des Intensitätsbildes wird die unregelmäßige Punktwolke in ein Raster umgerechnet. Zur Registrierung wird ein automatisierter Kantenfinder verwendet. Im Anschluss wird ein flächenbasiertes Bildmatching durchgeführt, um Korrespondenzen zwischen Intensitätsbild und optischem Bild zu finden. Die 3D-Punkte der korrespondierenden Kantenpunkte des Intensitätsbildes werden aus der Punktwolke entnommen. Anhand dieser Punkte wird die innere und äußere Orientierung in einer Bündelausgleichung mit Simultankalibrierung berechnet.

BÖHM & BECKER (2007) lösen die Registrierung ebenfalls im zweidimensionalen Raum. Zunächst wird ein Intensitätsbild der Punktwolke für den Standpunkt des Laserscanners erzeugt. Hierfür wird die rasterweise Abtastung des Objektes als Bild repräsentiert. Im Anschluss werden mit dem SIFT-Algorithmus Interessenspunkte zwischen Intensitätsbild und optischem Bild gefunden und anschließend nach Korrespondenzen gesucht. Mit den gefundenen Punktpaaren werden die Transformationsparameter berechnet.

## 2 Daten

Die Daten wurden mit einem Mobile Mapping System Riegl VMX-250-CS6 mit vier Kameras (Abbildung 2) erzeugt. Das System liefert die Trajektorie mit Position und Orientierung des Messsystems, eine Punktwolke aus Laserscanning sowie optische Bilder. Die Punktwolke wurde mittels der Software RiPROCESS berechnet und als ASCII-Datei weiterverarbeitet. Zu jedem Punkt liegen der Zeitpunkt der Erfassung, die 3D-Koordinaten sowie ein Intensitätswert vor. Die Dichte der Punktwolke ist in Fahrtrichtung abhängig von Scanfrequenz und Fahrtgeschwindigkeit, quer dazu von Pulsrate und Distanz.

Die optischen Bilder wurden mit zwei unterschiedlichen Kamertypen erzeugt, je zwei Spiegelreflexkameras (12,1 MPixel) und zwei Industriekameras (5 MPixel). Die Bildfrequenz lag bei 1 Bild/s. Die Orientierung der Kameras relativ zum Gesamtsystem wurde mittels der von Riegl vorgesehenen Prozedur bestimmt. Diese sieht die manuelle Identifikation korrespondierender Punkte in den Scan- und Bilddaten vor. Anhand der Orientierung des Gesamtsystems in einem übergeordneten Koordinatensystem kann mittels des Zeitstempels die äußere Orientierung für jedes Bild berechnet werden und dient als Näherungswert für das entwickelte Verfahren. Zusätzlich liegen die inneren Orientierungen der Kameras vor, die von Riegl kalibriert wurden. Die hier verwendeten Daten stammen aus einer Fahrt in der Innenstadt Hannover, welche im April 2012 durchgeführt wurde.



**Abb. 2:** Mobile Mapping System Riegl VMX-250-CS6 mit zwei seitlich ausgerichteten Spiegelreflexkameras und zwei rückwärtig gerichteten Industriekameras.

## 3 Algorithmus

Zur Bestimmung der relativen Orientierung zwischen einer einzelnen Kamera und der Punktwolke sollen korrespondierende Punkte genutzt werden. Der im Folgenden vorgestellte Algorithmus beruht auf der Annahme, dass die Projektion einer Punktwolke in ein Bild geometrische und radiometrische Ähnlichkeit mit einem optischen Bild mit der gleichen inneren und äußeren Orientierung aufweist.

Für die Berechnung der relativen Orientierung sind mehrere Schritte notwendig. Zunächst wird die Punktwolke in ein Intensitätsbild projiziert. Anhand der näherungsweise bekannten äußeren Orientierung der Kamera kann ein Intensitätsbild mit der gleichen Orientierung

erzeugt werden. Die Projektion der Scanpunkte führt aufgrund der geringen Dichte der Punktwolke zu Lücken im erzeugten Bild. Daher wird das Ergebnis durch einen nachfolgenden Bildverarbeitungsschritt verdichtet. Anschließend wird der SURF-Algorithmus verwendet, um sowohl im Intensitätsbild als auch im optischen Bild Interessenspunkte, d. h. markante Punkte, zu finden. Diese Interessenspunkte werden anhand ihrer Eigenschaften, wie Position, Größe und Deskriptoren, miteinander verglichen. Dadurch können Korrespondenzen zwischen dem optischen Bild und dem Intensitätsbild gefunden werden. Im Anschluss wird die äußere Orientierung des Kamerastandpunktes bestimmt. Zu diesem Zweck wird ein Rückwärtsschnitt mittels iterativer Ausgleichung berechnet. Des Weiteren wird ein Ausreißertest durchgeführt, um vorhandene grobe Fehler aufzudecken und zu eliminieren. Als Ergebnis des Rückwärtsschnitts werden die äußere Orientierung des Bildes sowie die dazugehörige Genauigkeit ausgegeben. Aus dieser äußeren Orientierung wird mithilfe der bekannten Orientierung des Systems die relative Orientierung berechnet. Die relative Orientierung zwischen Kamera und System wird schließlich aus einer Vielzahl von äußeren Bildorientierungen durch Mittelwertbildung bestimmt.

### 3.1 Erzeugen eines Intensitätsbildes

Das Intensitätsbild der Punktwolke wird erzeugt, indem die Punkte der Punktwolke aus dem übergeordneten System in das Kamerasystem transformiert werden und anschließend über  $u = x/z \cdot f_x + c_x$  und  $v = y/z \cdot f_y + c_y$  in die Bildebene projiziert werden. Dazu werden die näherungsweise bekannten Parameter der äußeren Orientierung einer Kamera sowie die Parameter der inneren Orientierung (Brennweite  $f_x, f_y$  und Bildhauptpunkt  $c_x, c_y$ ) genutzt. Um eine größtmögliche Ähnlichkeit zum optischen Bild zu erzeugen, werden zusätzlich die Verzeichnungsparameter angebracht (vgl. RIEGL LASER MEASUREMENT SYSTEMS GMBH 2011). Die Intensitätswerte der Punktwolke liegen in einem Bereich von -16 dB bis -5 dB. Für eine gute Ähnlichkeit mit dem optischen Bild wird diese auf den Bereich von 0...255 skaliert.

Aufgrund der geringen Dichte der Punktwolke entsteht ein Intensitätsbild, das Lücken zwischen den projizierten Punkten aufweist. Um Korrespondenzen zum optischen Bild zu finden, wird dieses Bild verdichtet. Mit einer Filtermaske von 5 x 5 Pixel werden im Fall der genutzten rückwärtsgerichteten Industriekameras gute Ergebnisse erzielt (Abbildung 3).



**Abb. 3:** Optisches Bild einer rückwärtsgerichteten Industriekamera (links) und verdichtetes Intensitätsbild aus der Punktwolke (rechts).

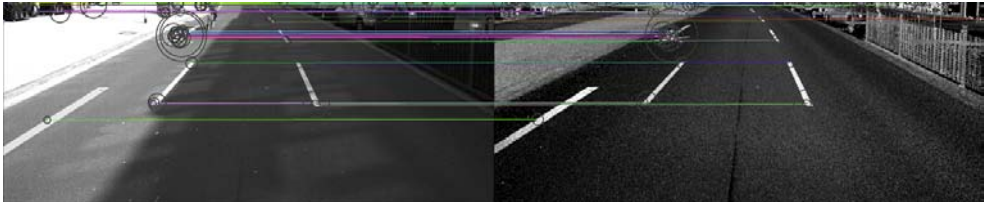
### 3.2 Auswahl korrespondierender 2D-Features

Die Verknüpfung zwischen optischem Bild und Intensitätsbild wird mittels einer Zuordnung von Interestpunkten hergestellt. Zunächst werden in den Bildern SURF-Features (Bay u. a. 2006) extrahiert. Auf Grund der Unterschiede zwischen Intensitätsbild und optischem Bild werden neben der Ähnlichkeit der Deskriptoren, die anhand der euklidischen Distanz geprüft wird, weitere Kriterien berücksichtigt.

Zunächst wird untersucht, ob der Wert des Laplace-Operators für die Kandidaten identisch ist. Damit kommen nur noch Interestpunkte in Frage, die beide hell und auf dunklerem Hintergrund bzw. umgekehrt liegen.

Im Anschluss wird die Epipolargeometrie (Luhmann 2010) genutzt, um die Suche einzuschränken. Ein 3D-Punkt und die zugehörigen Bildpunkte bilden die Epipolarebene. Im vorliegenden Fall haben beide Bilder die gleiche äußere Orientierung, d.h. korrespondierende Bildpunkte und der 3D-Punkt liegen auf einem gemeinsamen Raumstrahl. Die Interestpunkte sollten folglich dieselbe Bildkoordinate aufweisen. Da die Orientierung nur näherungsweise bekannt ist, sind die Raumstrahlen nicht identisch und die Bildkoordinaten weichen voneinander ab. Die Information kann dennoch genutzt werden, um den Suchradius auf wenige Pixel einzuschränken.

Die verbleibenden Kandidaten werden anschließend auf ihre Größe geprüft, sodass alle Punkte ausgeschlossen werden, die in unterschiedlichen Skalierungsstufen gefunden wurden. Zum Schluss wird die Distanz der Deskriptoren berechnet. Ein korrespondierendes Punktpaar wird gefunden, wenn die Distanz 0,7-mal kleiner ist als die Distanz zum zweitnächsten Nachbarn (Bay u. a. 2006). Außerdem darf die minimale Distanz einen vorgegebenen Schwellwert nicht überschreiten. Dadurch wird sichergestellt, dass eine Ähnlichkeit vorhanden ist (Ausschnitt in Abbildung 4).



**Abb. 4:** Korrespondenzen zwischen optischem Bild (links) und Intensitätsbild (rechts).

### 3.3 Zuordnung von 3D-Punkten zu 2D-Features

Die Zuordnung der 3D-Punkte zu den Interestpunkten ist notwendig, da nicht zu jedem Bildpunkt ein 3D-Punkt vorliegt. Der Grund hierfür ist die geringe Punktdichte der Punktwolke. Durch das Verdichten des Intensitätsbildes werden zusätzliche Grauwerte im Bild erzeugt, wodurch eine größere Ähnlichkeit mit dem optischen Bild erreicht wird. Allerdings werden keine neuen 3D-Punkte generiert. Aus diesem Grunde liegt nur in Ausnahmefällen an der Position des Interestpunktes ein 3D-Punkt vor. Wird an der Position des Features kein 3D-Punkt gefunden, muss dieser in der Nachbarschaft ermittelt werden. Die Suche nach dem besten 3D-Punkt basiert auf der Annahme, dass Punkte, die in einer Ebene liegen, in der Zuordnung weniger fehleranfällig sind.

Dazu werden zunächst alle 3D Punkte ausgewählt, die im 2D-Bild Nachbarn des entsprechenden Features sind. Wird dabei kein oder nur ein 3D Punkt in der Nachbarschaft detektiert, wird das Feature verworfen, da keine Kontrolle möglich ist. Gibt es mindestens zwei Punkte, werden die Distanzen der 3D-Punkte untereinander bestimmt. Überschreitet mindestens eine Distanz einen definierten Schwellwert, wird der Interestpunkt verworfen. Da keine neuen 3D-Punkte erzeugt werden sollen, wird auf Basis der vorhandenen 3D-Punkte, der nächstgelegene Punkt bestimmt. Dazu wird der Abstand der in Frage kommenden 3D-Punkte zum Raumstrahl des Feature-Punktes bestimmt und der Punkt mit dem geringsten Abstand dem Interestpunkt zugeordnet.

### 3.4 Berechnung der äußeren Orientierung

Anhand der extrahierten Interestpunkte und ihrer 3D-Punkte wird die äußere Orientierung der Bilder in einer Ausgleichung nach der Methode der kleinsten Quadrate bestimmt. Als funktionales Modell dient der räumliche Rückwärtsschnitt.

Als Beobachtungen gehen die Bildkoordinaten der Interestpunkte des optischen Bildes ein. Für die Näherungswerte der Beobachtungen werden die zugeordneten 3D-Punkte ins Bild projiziert.

Die Beobachtungen sowie die Näherungswerte der Beobachtungen gehen als unverzerrte Bildkoordinaten ein, da die Verzeichnung innerhalb der Ausgleichung nicht berücksichtigt wird. Für das stochastische Modell wird angenommen, dass die Beobachtungen alle gleichgenau vorliegen.

Die Berechnung der Unbekannten erfolgt iterativ, bis die Quadratsumme der Zuschläge der Unbekannten unter einem Schwellwert liegt.

Um vorhandene grobe Fehler aufzudecken und zu eliminieren, wird ein Ausreißertest durchgeführt. Damit sollen vor allem fehlerhafte Zuordnungen der 3D-Punkte zu den Interestpunkten erkannt werden. Hierfür wird die Kofaktormatrix  $Q_{vv} = Q_{ll} - A \cdot Q \cdot A^T$  der Verbesserungen aufgestellt. Anhand der normierten Verbesserung  $w_i$ , die sich daraus nach

$$w_i = v_i / \hat{s}_{v_i} \quad \text{mit} \quad \hat{s}_{v_i} = \hat{s}_{l_i} \cdot \sqrt{(Q_{vv})_{ii}}$$

berechnen lässt, wird mittels eines Schwellwerts  $k$  geprüft, ob ein grober Fehler vorhanden ist. Bei  $w_i > k$  liegt mit einer Irrtumswahrscheinlichkeit von  $\alpha = 1\%$  ein grober Fehler vor ( $k = 2,56$ ). Im Rahmen der Ausgleichung werden iterativ alle aufdeckbaren groben Fehler eliminiert, indem der jeweilige betroffene Interestpunkt aus der Berechnung entfernt wird. Anschließend wird die Ausgleichung ohne diesen Punkt erneut durchgeführt (LUHMANN 2010).

### 3.5 Berechnung der relativen Orientierung

Aus der geschätzten äußeren Orientierung der Bilder erfolgt im nächsten Schritt die Berechnung der relativen Orientierung als gewichtetes Mittel aus allen Bildern, deren Orientierung unter Verwendung von mehr als 20 zugeordneten Interestpunkten bestimmt wurde. Als Gewicht werden die Standardabweichungen der Parameter der äußeren Orientierung verwendet.

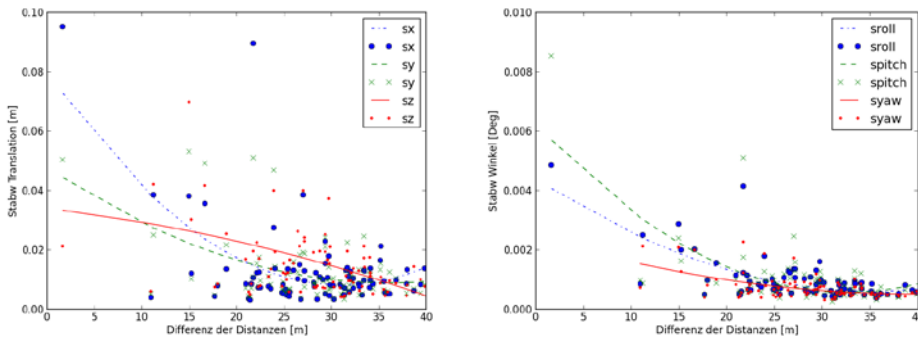
## 4 Ergebnisse

Zur Bestimmung der relativen Orientierung wurden insgesamt 90 Bilder ausgewertet. Diese stammen aus einer Messfahrt und wurden unter ähnlichen äußeren Bedingungen aufgenommen. Davon wurden 76 Bilder zur Berechnung der relativen Orientierung verwendet. Die a posteriori Standardabweichung der Gewichtseinheit lag bei  $s_0 = 0,36$  Pixel. Die Standardabweichungen der einzelnen Parameter liegen bei den Translationen bei 6 - 8 mm und bei den Rotationen bei 10 - 12 mGrad.

Zur Beurteilung der Bilder ist die Ähnlichkeit zwischen optischem Bild und Intensitätsbild wichtig, um ausreichend viele korrespondierende Punkte zu finden, da die Anzahl der zur Ausgleichung verwendbaren Punkte bei der Zuordnung von 3D-Punkten zu Interesspunkten deutlich verringert wird. Im optischen Bild werden bei guter visueller Übereinstimmung durchschnittlich 3500 Interesspunkte, im Intensitätsbild ca. 5000 Punkte gefunden. Davon wurden pro Bildpaar nur für ca. 200 Punkte Korrespondenzen hergestellt. Nach der Zuordnung der 3D-Punkte blieben durchschnittlich 32 Punkte pro Standpunkt übrig. Bei geringer Ähnlichkeit lag die Anzahl der Punkte deutlich geringer, so finden sich in den Bildern ca. 1000 Interesspunkte, davon verbleiben weniger als 10 Punkte für die Ausgleichung.

Es wurde untersucht, ob die Größe der Verbesserungen an den Bildkoordinaten der Interesspunkte von der Distanz der zugehörigen 3D-Punkte zur Kamera abhängig ist. Für die getesteten Daten ist allerdings keine Korrelation erkennbar.

Weiterhin wurde der Einfluss der räumlichen Verteilung der genutzten 3D-Punkte auf das Ergebnis der Ausgleichung untersucht. Dabei zeigt sich, dass die Verwendung der räumlichen Tiefe im 3D-Raum bedeutender ist, als die komplette Ausnutzung des Bildformats. Liegen die 3D-Punkte mit großer unterschiedlicher Entfernung zum Projektionszentrum, verbessert sich die Standardabweichung der einzelnen Parameter.



**Abb. 5:** Standardabweichung der geschätzten Parameter in Abhängigkeit von der räumlichen Tiefe der 3D-Punkte für die Translationen (links) und Rotationen (rechts).

## 5 Zusammenfassung und Ausblick

Das vorgestellte Verfahren bietet eine Möglichkeit, die relative Orientierung von Kameras und Laserscannern eines Mobile Mapping Systems vollautomatisch aus den erfassten Daten selbst zu bestimmen. Die Voraussetzung hierfür sind allerdings gute Näherungswerte. Im nächsten Schritt muss die Suche nach korrespondierenden Interestpunkten so erweitert werden, dass eine Zuordnung auch bei schlechteren Näherungswerten möglich ist.

Weitere Arbeiten sind auch erforderlich, um das Verfahren auf andere Kameratypen anzupassen. Es zeigte sich bei der Auswertung der seitlich ausgerichteten Spiegelreflexkameras, dass trotz verdichten der Intensitätsbilder auf Grund der höheren Auflösung keine ausreichende Ähnlichkeit mit den optischen Bildern erreicht werden konnte und Artefakte, die aus der Struktur der Punktwolke entstehen, die Suche nach korrespondierenden Interestpunkten stark beeinflusst.

Außerdem sollte die Information zwischen den Aufnahmen mit genutzt werden, da die relative Orientierung der Kameras während der Messfahrt stabil ist. Damit könnte im Rahmen einer Bündelausgleichung ein Bildverband berechnet und anschließend mit der Punktwolke verknüpft werden.

## Literatur

- Bay, H., Tuytelaars, T., Gool, L.V. (2006): SURF: Speeded Up Robust Features. ECCV, S. 404-417.
- Böhm, J. & Becker, S. (2007): *Automatic Marker-Free Registration of Terrestrial Laser Scans Using Reflectance Features*. 8th Conference on Optical 3D Measurement Techniques, S. 338 – 344.
- Li-Chee-Ming, J. & Armenakis, C. (2010): *Fusion of optical and terrestrial laser scanner data*. The 2010 Canadian Geomatics Conference and Symposium of Commission I, ISPRS Convergence in Geomatics - Shaping Canada's Competitive Landscape.
- Luhmann, T. (2010): *Nahbereichsphotogrammetrie*. Wichmann Verlag, 2010
- RiegI Laser Measurement Systems GmbH (2011): *RiPROCESS - Software Description and User's Instructions*.
- Swart, A., Broere, J., Veltkamp, R. & Tan, R. (2011): *Refined non-rigid registration of a panoramic image sequence to a LiDAR point cloud*. Photogrammetric Image Analysis, Springer, S. 73-84.