

# Texturierung von 3D Gebäudemodellen unter Verwendung von Schrägsichtaufnahmen flugzeuggetragener TIR-Kameras

IWASZCZUK DOROTA & STILLA UWE

*Zusammenfassung: Thermische Infrarotaufnahmen von Gebäuden werden immer häufiger zur Untersuchung der Energieeffizienz von Gebäuden eingesetzt. Die Kombination der thermischen Daten mit 3D-Gebäudemodellen durch z.B. Texturierung hilft u.a. bei der Interpretation und Geolokalisierung der Aufnahmen. In diesem Beitrag wird ein Verfahren zur automatischen Texturierung von 3D-Gebäudemodellen mit thermischen Bildsequenzen präsentiert. Der automatische Vorgang umfasst die direkte Georeferenzierung von Aufnahmen, die automatische Verfeinerung der Korregistrierung und die automatische Auswahl und Extraktion von Gebäudetexturen inklusive einer Sichtbarkeitsprüfung. Die Neuerung der präsentierten Methode besteht hauptsächlich in der Berücksichtigung und maximalen Reduzierung der Ungenauigkeiten des 3D-Gebäudemodells. Die präsentierte Methode wird anhand einer von einem Hubschrauber aufgenommenen thermischen Bildsequenz untersucht.*

## 1 Einleitung

Thermische Infrarotaufnahmen von Gebäuden werden immer häufiger zur Untersuchung der Energieeffizienz von Gebäuden und zur Inspektion des Zustandes von Gebäudehüllen eingesetzt. Die Kombination der thermischen Daten mit der dreidimensionalen Geometrie der Gebäude hilft bei der Identifikation von Wärmeverlusten und thermischen Überwachung einzelner Gebäude, Gebäudekomplexe sowie ganzer Stadtteile. Dazu können die thermischen 2D-Aufnahmen mit 3D-Gebäudemodellen mittels Texturierung zusammengeführt werden.

Heutzutage werden 3D-Stadtmodelle oft mit sichtbaren (VIS) Texturen angereichert. Texturierungen von 3D-Gebäudemodellen finden Anwendung in vielen Gebieten, wie z.B. bei der Stadtplanung, beim Krisenmanagement und in der Werbung.

Es wurden einige Systeme zur semi-automatischen Texturierung, wie z.B. SAMATS - Semi-Automated Modeling and Texturing System (HEGARTY & CARSWELL, 2009) oder Toposcopy (GRONEMAN, 2004) entwickelt. Verfügbar ist auch kommerzielle Software die zur automatischen Texturierung entwickelt wurde (z.B. UVM SYSTEMS, 2014). Diese Software setzt aber meistens eine perfekte Koregistrierung zwischen den 3D-Modellen und denen zur Texturierung verwendeten Bildern voraus oder benötigt manuelle Vorarbeit. Die Koregistrierung kann durch die Berechnung der äußeren Orientierung erreicht werden, die typischerweise in der Aerotriangulation berechnet wird. Die äußere Orientierung kann auch durch die Verwendung eines GPS, häufig in Verbindung mit einer IMU (Eng. inertial measurement unit), bestimmt werden. Oft ist jedoch die mit dem GPS und der IMU aufgenommene Position und Orientierung nicht genau genug, um die aufgenommenen Daten genau georeferenzieren zu können und mit den Daten aus verschiedenen Quellen zu fusionieren. Außerdem werden in den kommerziellen

Systemen manche Aspekte der Texturierung, wie die Fremdverdeckung (z.B. durch Vegetation) und Auswahl der besten Textur nicht oder nur am Rande berücksichtigt.

In einigen wissenschaftlichen Beiträgen ist das Thema der automatischen Texturierung unter den Aspekten vieler auftretender Probleme behandelt worden. In vielen Arbeiten werden Strategien zur Sichtbarkeitsprüfung vorgestellt, die sich in zwei Gruppen unterscheiden lassen: bildbasierte Methoden (e.g. FRUEH et al., 2004; Karras et al., 2007) und vektorbasierte Methoden (e.g. HANUSCH, 2008). Die Auswahl der besten Texturen wird auf Basis von verschiedenen Kriterien durchgeführt: Sichtbarkeit (BÉNITEZ & BAILLARD 2009), Blickwinkel (WANG ET AL., 2008) oder einer Kombinationen aus Sichtbarkeit, Auflösung, Blickwinkel, Entfernung und Kohärenz mit dem benachbarten Dreieck des 3D-Modells (FRUEH et al., 2004; IWASZCZUK & STILLA, 2010). LORENZ (2011) präsentiert eine Methode zur automatischen Texturierung mit texelpräziser Berechnung des Qualitätskriteriums. Die Fremdverdeckung wird durch BÖHM, 2004; ABDELHAFIZ & NIEMEIER, 2009; BÉNITEZ et al., 2010 behandelt.

Zur Koregistrierung des 3D-Modells mit den zur Texturierung verwendeten Bildern werden ebenfalls verschiedene Techniken angewendet. FRUEH et al. (2004) projizieren das 3D-Gebäudemodell mit verschiedenen, zufälligen Parametern ins Bild und korreliert die Modelllinien mit den extrahierten Liniensegmenten. Aus mehreren Versuchen wird dazu die Kameraposition mit der besten Anpassung ausgewählt. In jedem Bild der Sequenz läuft der Vorgang gleich ab. Diese Methode nutzt keine Zusammenhänge zwischen den Frames. HU et al. (2006) verbinden die Frames mit markanten Punkten, die jedoch keinen Zusammenhang mit den Gebäudestrukturen aufweisen. Unter Annahme der sog. „Manhattan“ oder „Legoland scenes“ mit dominanten horizontalen und vertikalen Richtungen kann die Ausrichtung der Kamera mithilfe der Fluchtpunkte berechnet werden (DING & ZAKHOR, 2008). Dieses Verfahren hilft bei ungenauen Navigationsparametern, die Fehler der Drehwinkel zu reduzieren. Um die Texturen präzise zuordnen zu können, wird eine Methode benötigt, die für das Matching bei Gebäuden charakteristische Bildmerkmale verwendet und diese von Frame zu Frame verfolgt.

In diesem Beitrag wird eine Methode zur Texturierung von 3D Gebäudemodellen unter Verwendung von Schrägsichtaufnahmen luftgetragener TIR-Kameras präsentiert. Der Schwerpunkt der vorgestellten Methode liegt auf der Koregistrierung von Bildsequenzen mit 3D-Gebäudemodellen. Dafür werden die Linien des 3D-Gebäudemodells mit den in den Bildern detektierten Linien gematcht. Um die Verbindung zwischen den Frames herzustellen, werden die dem 3D-Modell zugeordneten Liniensegmente verfolgt. Der Rechenaufwand wird durch die Einführung von sog. Key-Frames reduziert.

## **2 Automatische Texturierung**

In diesem Beitrag wird die automatische Texturierung als direkte Georeferenzierung von Schrägsichtaufnahmen, die automatische Verfeinerung der Korregistrierung und die automatische Auswahl und Extraktion von Gebäudetexturen inklusive einer Sichtbarkeitsprüfung behandelt. Für jede Fläche des 3D-Modells wird eine Textur erzeugt, die ein vollständiges Rasterbild der jeweiligen Fläche darstellt. Wenn die Fläche kein Rechteck bildet, wird um die Fläche eine rechteckige Bounding-Box erstellt und diese als Textur verwendet. Die Pixel der

Bounding-Boxen, welche nicht zur Modellfläche gehören, werden ausgeblendet. Diese Strategie hat den Vorteil, dass man in den extrahierten Texturen weitere Objekte detektieren kann (im Gegensatz zu Texturierung von Dreiecken), wie zum Beispiel Wärmelecken. Für jede Fläche wird wie in Iwaszczuk & Stilla (2010) beschrieben die Textur mit der höchsten Qualität ausgewählt.

### 3 Koregistrierung von TIR-Bildsequenzen mit 3D-Gebäudemodellen

Die Ursachen für die Diskrepanzen zwischen dem projizierten 3D-Modell und dem zur Texturierung verwendeten Bild sind typischerweise die unbekannte oder ungenaue äußere und innere Orientierung des Bildes und/oder die Ungenauigkeit des 3D-Modells<sup>1</sup>. Um die exakte Texturextraktion zu ermöglichen, sollen die beschriebenen Fehler in einer Modell-zu-Bild-Koregistrierung minimiert werden. Die innere Orientierung wird als bekannt angenommen, da sie typischerweise mit einer Labor- oder Selbst-Kalibrierung bestimmt werden kann. Die approximierten äußere Orientierung kann aus den Navigationsdaten abgeleitet werden. Diese approximierten äußere Orientierung ist aber nicht genau genug um die Texturen präzise extrahieren zu können. Bei den Ungenauigkeiten der Gebäudemodelle handelt es sich häufig um nicht modellierte Dachüberstände, die insbesondere bei der Erfassung der 3D- Gebäudemodelle aus Luftbildaufnahmen entstehen. Dadurch können die Texturen für die Fassaden nicht korrekt ausgeschnitten werden. Um diese Probleme zu lösen, wird einerseits die optimale Kameraposition berechnet und andererseits werden die Texturen lokal im Bild angepasst. Bei der Verwendung von Bildsequenzen werden die dem 3D-Modell zugeordneten Linien von Frame zu Frame verfolgt.

#### 3.1 Bildsequenzen und Key-Frame-Konzept

Bei der Verwendung von Bildsequenzen mit einer hohen Frequenzrate ist die geometrische Veränderung zwischen den Frames sehr gering. Daher sind das rechenintensive Matching und die Berechnung der äußeren Orientierung nicht für jedes Frame zwingend notwendig. Deswegen werden sog. Key-Frames eingeführt, in deren Koregistrierung vollständig durchgeführt wird. In den anderen Frames wird ein vereinfachtes Matching verwendet. Die Liniensegmente werden dafür zwischen den Frames verfolgt.

In der Abb. 1 wird der Prozess schematisch dargestellt. Er besteht aus fünf Hauptteilen: Vorprozessieren, Entscheidung, Matching, Tracking, und Verifizierung der Zuordnungen. Alle Ergebnisse werden protokolliert und in einer Chronik gespeichert.

#### 3.2 Entscheidung

Beim Entscheidungsprozess (in Abb. 1 decision) soll festgestellt werden, ob das aktuelle Frame ein Key-Frame oder nicht sein soll. Zur Auswahl eines Key-Frames entscheidend sind folgende Eigenschaften:

- die vordefinierten Abstände zwischen den Key-Frames (z.B. jedes zehnte Frame)

---

<sup>1</sup> Es können auch weitere Effekte Fehler verursachen (z.B. Rolling-Shutter-Effekt), die zu dieser Diskrepanz beitragen. Diese werden in diesem Artikel nicht weiter behandelt.

- zu kleine Überlappung mit dem vorherigen Frame
- nicht ausreichende Zuverlässigkeit des Matchings im vorherigen Frame
- Ergebnisse der Qualitätsberechnung und Texturauswahl (Frames bei denen gute Texturen extrahiert werden können, sollen Key-Frames werden)

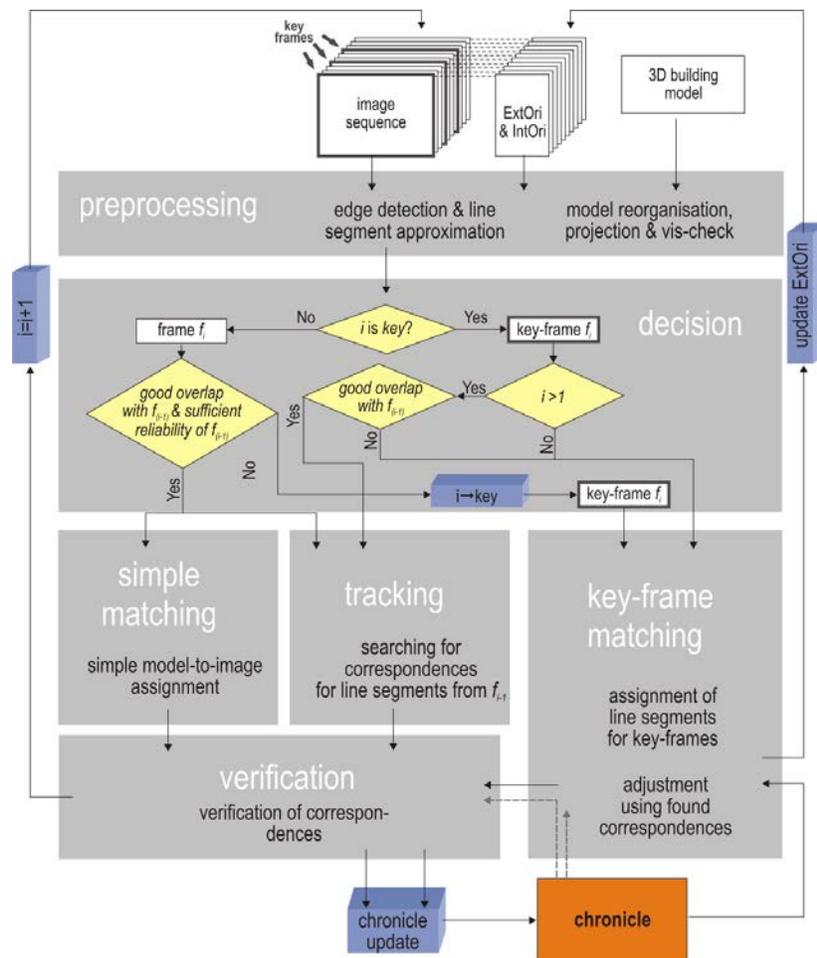


Abb. 1: Vereinfachtes Schema von Linienverfolgung

### 3.3 Matching

Nachdem die Entscheidung getroffen wird, ob das aktuelle Frame ein Key-Frame ist, wird das entsprechende Matching-Verfahren angewendet: Key-Frame Matching für die Key-Frames und vereinfachtes Matching für sonstige Frames.

#### 3.3.1 Key-Frame Matching

Für jede projizierte Modellkante wird in der lokalen Umgebung nach passenden Liniensegmenten im Bild gesucht. Die Größe des Suchbereiches wird durch die Fehlerfortpflanzung der Unsicherheit der Kameraposition und -Orientierung und der Unsicherheit der Modellkante berechnet. Die richtigen Korrespondenzen werden mit einem Ausreißer-Detektor selektiert. Dann werden die Position und Orientierung der Kamera neu

berechnet. Um die Unsicherheiten sowohl der äußeren Orientierung wie auch des Gebäudemodells als Konsequenz der ungenauen Erfassung und der Generalisierung zu berücksichtigen, wird ein stochastisches Modell in der Ausgleichung formuliert.

### 3.3.2 Vereinfachtes Matching

Beim vereinfachten Matching wird in der Umgebung der verfolgten Linie überprüft, ob sich dort eine passende Linie befindet, die im neuen Frame als dieselbe Linie erkannt werden kann. Dafür werden statistische Tests verwendet.

## 3.4 Linien-Tracking

Jede Linie, die in einem Key-Frame einer Modellkante zugeordnet wird, wird im nächsten Frame gesucht. Da bei einem Video mehrere Frames pro Sekunde aufgenommen werden, sind die Unterschiede in der Erscheinung der Objekte in benachbarten Frames sehr gering. Deswegen eignet sich die Kreuzkorrelation für die Linienverfolgung, da die geometrischen Verzerrungen von Frame zu Frame sehr gering sind und die Ergebnisse der Kreuzkorrelation kaum beeinflussen. Die neue Position der Linie im nächsten Frame kann aus den Navigationsdaten des Luftfahrzeuges geschätzt werden. Dafür wird um jede zugeordnete Linie ein Bildbereich gewählt und mit dem Bildbereich, bei dem dieselbe Linie zu erwarten ist, korreliert. Dadurch ergibt sich für jede Linie ein  $offset_i = [\Delta row_i, \Delta col_i]$ , aus dem eine Verschiebung zwischen den Frames als  $median(offset_j)$  berechnet wird.

## 3.5 Lokale Anpassung

Basierend auf den Korrespondenzen aus dem Matching, wird für jedes Frame die äußere Orientierung neu berechnet, um die beste Anpassung zwischen dem projizierten 3D-Modell und den Bildern zu erzielen. Wegen den vorher erwähnten Ungenauigkeiten im 3D-Modell ist jedoch die Anpassung für die einzelnen Texturen nicht ausreichend. Dafür wird noch zusätzlich in der lokalen Umgebung jeder projizierten Modellkante eine Gerade im Gradientenbild gesucht und für das Ausschneiden der Textur übernommen.

# 4 Experiment

## 4.1 Testdaten

Die IR-Daten wurden während einer Flugkampagne über einem dicht bebauten Bereich mit einer im Hubschrauber montierten IR-Kamera aufgenommen. Die Bilder wurden mit einer Frequenz von 25 pro Sekunde aufgenommen. Die Flughöhe betrug ca. 400 m über Grund und die Bilder wurden schräg nach vorne unter einem Pitch-Winkel von ca. 45° aufgenommen. Der Hubschrauber verfügte über ein GPS/INS Navigationssystem. Das Projektionszentrum der Kamera war im Verhältnis zur Antenne des Navigationsgerätes versetzt. Die innere Orientierung der Kamera war nur ungefähr bekannt. Daher wurde eine erweiterte Systemkalibrierung mit der Berechnung von boresight & leverarm Parametern (KOLECKI et al., 2010) durchgeführt. Die 3D-Gebäudemodelle wurden im Detailierungsgrad LOD2 aus Luftbildern semiautomatisch erstellt und im *CityGML*-Format gespeichert.

## 4.2 Ergebnisse

In Abb. 2 werden Ergebnisse der Koregistrierung dargestellt. In Rot wird die ursprüngliche Projektion der Gebäudemodelle, in Grün die Projektion nach der Neuberechnung der Parameter der äußere Orientierung dargestellt.

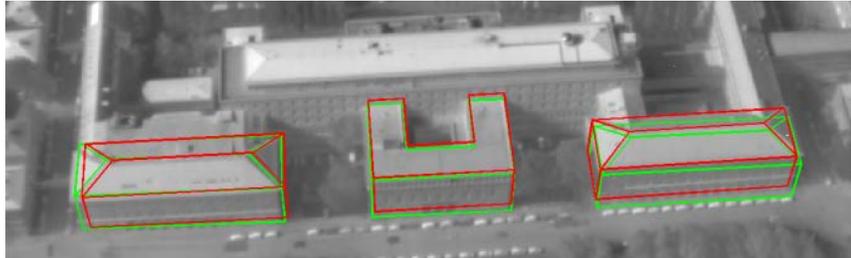


Abb. 2: Ergebnisse der Koregistrierung: Rot - ursprüngliche Projektion der Gebäudemodelle; Grün - die Projektion nach der Neuberechnung der Parameter der äußere Orientierung

In Abb. 3 wird das Ergebnis der lokalen Anpassung am Beispiel eines Gebäudes präsentiert. In Abb. 3a werden die ins Bild projizierten Modelkanten nach der Neuberechnung der Parameter der äußere Orientierung gezeigt, während in Abb. 3b die neu ermittelten Kanten in der Umgebung der projizierten Kante dargestellt werden. Die in diesem Beispiel erreichte Verbesserung der Position der Kante im Bild beträgt ca. 2-4 Pixel. Bei einer lokalen Pixelgröße auf der Fassade von ca. 25cm wird eine Verbesserung von bis zu 1m erreicht.

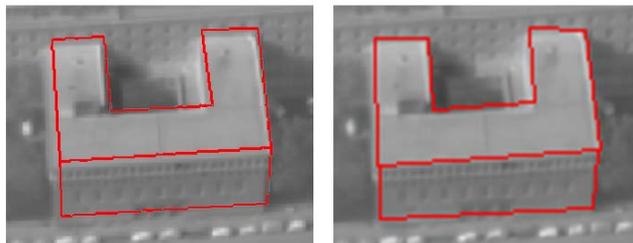


Abb. 3: Ergebnisse der lokalen Anpassung der Textur: a) ins Bild projiziertes 3D- Gebäudemodell; b) neugefundene Kanten in der Umgebung der projizierten Kante

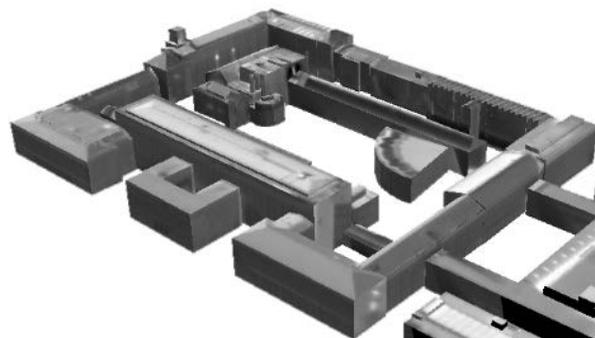


Abb. 4: Texturierte 3D-Gebäudemodelle

## 5 Diskussion und Ausblick

In diesem Beitrag wurde eine Methode zur automatischen Texturierung von 3D-Gebäudemodellen mit dem Schwerpunkt auf Korregistrierung von 3D-Gebäudemodellen mit Bildsequenzen vorgestellt. Um den Einfluss der Ungenauigkeiten des 3D-Modells zu reduzieren, wurden die projizierten Modellkanten noch lokal im Bild gefittet und die am besten passenden Texturen wurden extrahiert. Die ersten Ergebnisse haben die Eignung dieser Methode zur Extraktion von visuell passenden Texturen gezeigt. Die so erzeugten Texturen sind jedoch maßlich nicht korrekt. Um die geometrische Richtigkeit zu gewährleisten, muss die fehlerhafte Geometrie der 3D-Gebäudemodelle verbessert werden.



Abb. 5: Beispiel für die Fremdverdeckung: die im 3D- Modell nicht erhaltenen Kamine wurden als Element der Fassade in der Textur dargestellt

Durch die Aufnahme aus verschiedenen Blickrichtungen kann nicht nur die äußere Orientierung der Kamera für jede Aufnahme berechnet werden, sondern auch die bestehende Geometrie der 3D-Gebäudemodelle verbessert werden. Dazu wird ein Ausgleichsmodell mit den Parametern der Kamera und den Parametern des Gebäudemodells benötigt. Dies soll der Gegenstand von zukünftigen Untersuchungen werden.

Ein weiteres Problem der automatischen Texturierung, der in diesem Artikel nicht behandelt wurde, stellt die Fremdverdeckung dar. Ein Beispiel dafür wurde in Abb. 5 präsentiert. Hier wurden die im 3D- Modell nicht enthaltenen Kamine als Elemente der Fassade in der Textur dargestellt.

## 6 Literaturverzeichnis

- ABDELHAFIZ, A. & NIEMEIER, W., 2009: Integration of Digital Images and Laser scanning Point Clouds – Texture Mapping with 3DImage. Journal of DVW, 60: 113–125. 91. DVW-Seminar Terrestrial laser scanning TLS2009 Fulda, Germany, November 19-2.
- BÉNITEZ, S. & BAILLARD, C., 2009: Automated selection of terrestrial images from sequences for the texture mapping of 3d city models. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38 (3/W4): 97–102. Proceedings of CMRT09.

- BÉNITEZ, S., DENIS, E. & BAILLARD, C., 2010: Automatic production of occlusion-free rectified facade textures using vehicle-based imagery. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38 (3A): 275–280. Proceedings of PCV 2010 - Photogrammetric Computer Vision and Image Analysis.
- BÖHM, J., 2004: Multi-image fusion for occlusion-free facade texturing. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences
- DING, M. & ZAKHOR, A., 2008: Automatic registration of aerial imagery with untextured 3D LiDAR models. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
- FRUEH C.; SAMMON R. & ZAKHOR A., 2004: Automated Texture Mapping of 3D City Models With Oblique Aerial Imagery. Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'04)
- GRONEMAN, A., 2004: TOPOSCOPY combines photogrammetric modelling with automatic texture mapping. The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences. Proceedings of the ISPRS workshop on Vision techniques applied to the Rehabilitation of City Centres, Lisbon, Portugal.
- HANUSCH, T., 2008: A new texture mapping algorithm for photorealistic reconstruction of 3D objects. The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 37 (5): 699–706
- HEGARTY, J. & CARSWELL, J., 2009: SAMATS - Texture extraction explained. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38 (5/W1).
- HU, J., YOU, S. & NEUMANN U (2006) Automatic Pose Recovery for High-Quality Textures Generation. Proceedings of 18th International Conference in Pattern Recognition: 561–565.
- IWASZCZUK, D. & STILLA, U., 2010: Quality measures for textures extracted from airborne IR image sequences. The International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences, 38 (3A): 79–84. PCV 2010 - Photogrammetric Computer Vision and Image Analysis.
- KARRAS, G., GRAMMATIKOPOULOS, L., KALISPERAKIS, I. & PETSAS, E., 2007: Generation of Orthoimages and Perspective Views with Automatic Visibility Checking and Texture Blending. Photogrammetric Engineering & Remote Sensing, 73 (4): 403–411.
- KOLECKI, J., IWASZCZUK, D. & STILLA, U., 2010: Calibration of an IR camera system for automatic texturing of 3D building models by direct geo-referenced images. Proceedings of Eurocow 2010, Castelldefels, Spain
- LORENZ, H., 2011: Texturierung und Visualisierung virtueller 3D-Stadtmodelle. Dissertation. Mathematisch-Naturwissenschaftliche Fakultät der Universität Potsdam
- UVM SYSTEMS, 2014: CityGRID. Onlineprospekt: <http://www.uvm-systems.com>, letzter Zugang am 05.02.2014
- WANG, Y., SCHULTZ, S. & GUIUFFRIDA, F., 2008: Pictometry's proprietary airborne digital imaging system and its application in 3D city modelling. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 37: 1065–1066.