

Automatische Orientierung mit und ohne Messmarken – Das Mögliche und das Unmögliche

HELMUT MAYER¹

Zusammenfassung: Mit Hilfe des 5-Punkt Algorithmus von NISTÉR (2004) kann die relative Orientierung von zwei kalibrierten Kameras direkt, d.h., ohne Näherungen, bestimmt werden. Zusammen mit Random Sample Consensus (RANSAC) können Bildpaare mit natürlicher Textur auch bei großer Basis und z.B. 95% falschen Zuordnungen automatisch stabil orientiert werden. Durch Verknüpfung zu Tripeln und von Tripeln erhält man eine relative Orientierung für Sequenzen auf alleiniger Grundlage von Bildern und (u.U. genäherter) Kalibrierung. Dieser Möglichkeit steht das Unmögliche gegenüber, nämlich die Orientierung stark projektiv verzerrter Bilder mit unkodierten, angeblitzten, gleich aussehenden Messmarken, für die es keine Einschränkung der Zuordnung gibt. Als Lösungsmöglichkeit wird die Verknüpfung der beiden Vorgehensweisen vorgeschlagen: Für die Orientierung wird die in nicht durch Blitz völlig überstrahlten Bildern meist vorhandene natürliche Textur verwendet. Unkodierte Messmarken werden über die dann bekannte Epipolargeometrie und Trifokaltensoren zugeordnet, was neue Anwendungsmöglichkeiten durch die verstärkte Automatisierung eröffnen sollte. Mögliches und Unmögliches werden an Hand von Beispielen erläutert

1 Einleitung

Die automatische Orientierung von (evtl. unkalibrierten) Bildern hat in den letzten Jahren einen hohen Stand erreicht. Gute Beispiele sind (POLLEFEYS et al. 2004, LHULLIER & QUAN 2005). In der Photogrammetrie dokumentiert (REMONDINO & EL-HAKIM 2006) den Stand. Der vorgelegte Beitrag zeigt neue Entwicklungen für die Orientierung kalibrierter Bilder auf und stellt dar, wie diese für die Orientierung mit unkodierten Messmarken genutzt werden könnten, was durch die verstärkte Automatisierung neue Anwendungen eröffnen sollte.

Mit dem 5-Punkt Algorithmus (NISTÉR 2004) kann die relative Orientierung von zwei kalibrierten Kameras auf Grundlage von fünf korrespondierenden Punkten direkt, d.h., ohne Näherungen, bestimmt werden. In Kombination mit Random Sample Consensus – RANSAC (FISCHLER & BOLLES 1981) ermöglicht er die automatische Orientierung, auch wenn bei der Bildzuordnung bis zu 95% der Punkte falsch zugeordnet werden. Damit wird es möglich, Bildpaare mit natürlicher Textur ohne Messmarken auch bei großer Basis stabil zu orientieren. Die Grundlage für die Zuordnung auf Basis von Kreuzkorrelation und kleinsten Quadraten bilden markante Punkte. Um mehr als zwei kalibrierte Bilder relativ zu orientieren, werden Tripel gebildet. Tripel werden als Grundelemente des vorgestellten Ansatzes verwendet. Für sie können Punkte auf Grundlage geometrischer Schnittbedingungen eindeutig kontrolliert werden, was die Zuverlässigkeit stark erhöht. Für vier und mehr Bilder werden Tripel über ihre Projektionsmatrizen verknüpft. Zur Steigerung der Effizienz werden Bildpyramiden verwendet. Insgesamt erhält man eine vollauto-

¹ Helmut Mayer: Universität der Bundeswehr München, Institut für Photogrammetrie und Kartographie, 85577 Neubiberg, Tel.: +49-89-6004-3429, Fax: +49-89-6004-4090, e-mail: Helmut.Mayer@unibw.de

matische relative Orientierung von Sequenzen auf alleiniger Grundlage von Bildern und (u.U. genäherter) Kalibrierung.

Dem hiermit Möglichen steht das Unmögliche gegenüber: Es besteht aus Bildern mit unkodierten Messmarken, die so angeblitzt wurden, dass der Hintergrund schwarz wird. Da alle Messmarken gleich aussehen, kann jede Messmarke im einen jeder Messmarke im anderen Bild zugeordnet werden. Dies führt schon bei zwanzig Messmarken zu 95% Fehlzuordnungen, dem Limit für RANSAC-basierte Ansätze. Zwar kann die Anordnung der Punkte zur Einschränkung der Zuordnung verwendet werden, allerdings funktioniert dies nur für nicht zu stark auseinander liegende und damit unterschiedlich projektiv verzerrte Ansichten.

In diesem Beitrag wird als Lösung vorgeschlagen, die beiden oben vorgestellten Vorgehensweisen zu kombinieren: Für die grundlegende Orientierung wird die in nicht durch Blitz völlig überstrahlten Bildern meist vorhandene natürliche zufällige und damit mehr oder weniger eindeutig zuordnebare Textur verwendet. Unkodierte Messmarken können dann über die bekannte Epipolarometrie und Trifokaltensoren zugeordnet werden.

In Abschnitt 2 werden die grundlegenden Ideen des 5-Punkt Algorithmus und von RANSAC vorgestellt. Hierauf baut der Abschnitt 3 auf, der sich mit der Orientierung von Bildsequenzen beschäftigt. Abschnitt 4 illustriert das Unmögliche und Abschnitt 5 zeigt eine Lösungsmöglichkeit auf. Dies wird mit einer längeren, aus jeweils sechs Megapixeln großen Bildern bestehender, Sequenz illustriert und diskutiert. Das Papier schließt mit einem Ausblick.

2 5-Punkt Algorithmus und RANSAC

Der 5-Punkt Algorithmus (NISTÉR 2004) löst zum ersten Mal die Orientierung von zwei Bildern direkt, d.h., ohne Näherungen. Grundlage hierfür ist die Verwendung der Bedingungen, die für den kalibrierten Fall gelten und eine geeignete mathematische Formulierung.

Auch wenn in diesem Papier davon ausgegangen wird, dass Kalibrierungsinformation für die Aufnahmen vorhanden ist, impliziert dies nicht, dass vor der Aufnahme eine (Labor-) Kalibrierung durchgeführt werden muss. Die Erfahrung zeigt, dass z.T. relativ grobe Näherungen, z.B. von mehr als 5% Fehler, die man von älteren ausgewerteten Sequenzen oder z.T. auch aus Erfahrungen mit ähnlichen Kameras erhält, für eine Orientierung ausreichen können.

Sobald eine genäherte Orientierung vorliegt, kann diese mit Selbstkalibrierung (MCGLONE et al. 2004) verbessert werden. Weiterhin ist anzumerken, dass z.B. mit dem Ansatz von (POLLEFEYS et al. 2004) aus Aufnahmen mit genügend großer Basislänge und 3D Struktur ausreichende Näherungen für die Kalibrierung direkt aus Bildern bestimmt werden können.

Grundlage des 5-Punkt Algorithmus ist die Komplanaritätsbedingung. Hierfür transformiert man die beobachteten Bildpunkte mittels euklidischer Rotationsmatrix \mathbf{R} und Kalibriermatrix \mathbf{K} . Man erhält reduzierte (homogene) Bildkoordinaten ${}^k\mathbf{x}'$ und ${}^k\mathbf{x}''$ für erstes und zweites Bild. Die erste Kamera kann mittels der Projektionsmatrix ($\mathbf{I}0$) und die zweite mittels ($\mathbf{I}\mathbf{T}$) beschrieben werden. Dabei ist \mathbf{T} der Basisvektor. Die Bedingung, dass sich die Strahlen schneiden, impliziert die Komplanarität von ${}^k\mathbf{x}'$, ${}^k\mathbf{x}''$ und \mathbf{T} . Es ergibt sich (das Vektorprodukt \times steht senkrecht auf der aus ${}^k\mathbf{x}'$ und \mathbf{T} gebildeten Ebene, in der auch ${}^k\mathbf{x}''$ liegt, somit wird das Skalarprodukt 0):

$${}^k \mathbf{x}' (\mathbf{T} \times {}^k \mathbf{x}'') = {}^k \mathbf{x}'^T \mathbf{S}_T {}^k \mathbf{x}'' = 0 \quad \text{mit} \quad \mathbf{S}_T = \mathbf{S}(\mathbf{T}) = \begin{pmatrix} 0 & -T_3 & T_2 \\ T_3 & 0 & -T_1 \\ -T_2 & T_1 & 0 \end{pmatrix}.$$

\mathbf{S}_T ist die schiefsymmetrische Matrix zum Vektor \mathbf{T} mit Rang 2, die das Vektorprodukt $\mathbf{V} = \mathbf{T} \times \mathbf{U}$ als Matrix-Vektor-Multiplikation darstellt. Damit ergibt sich insgesamt wegen $(\mathbf{R}^{n-1})^T = \mathbf{R}'$ (Drehmatrix) und mit \mathbf{K} als projektiver, d.h., mit einem Skalar multiplizierter, Kalibriermatrix

$$\mathbf{x}'^T (\mathbf{K}^{n-1})^T \mathbf{R}' \mathbf{S}_T \mathbf{R}^{n-1} \mathbf{K}^{n-1} \mathbf{x}'' = 0.$$

Diese Beziehung ist linear in den Bildkoordinaten beider Bilder, d.h., insgesamt bilinear. Mit der so genannten 3×3 „Fundamentalmatrix“

$$\mathbf{F} = (\mathbf{K}^{n-1})^T \mathbf{R}' \mathbf{S}_T \mathbf{R}^{n-1} \mathbf{K}^{n-1}$$

kann die Komplanaritätsgleichung als Bedingung für die Homologie der Bildpunkte einfach dargestellt werden: $\mathbf{x}'^T \mathbf{F} \mathbf{x}'' = 0$. Daraus, dass der Kern von \mathbf{F} die Matrix \mathbf{S}_T mit Rang 2 ist, ergibt sich die kubische Bedingung

$$\det(\mathbf{F}) = 0 \tag{1}$$

die sich auch darin ausdrückt, dass ein Singulärwert von \mathbf{F} 0 ist. Für die essentielle Matrix \mathbf{E} , die man erhält, wenn die Kalibrierung bekannt ist,

$$\mathbf{x}'^T (\mathbf{K}^{n-1})^T \mathbf{E} \mathbf{K}^{n-1} \mathbf{x}'' = 0$$

gilt weiterhin, dass die beiden übrigen Singulärwerte gleich groß sind. Dies kann mittels der folgenden neun kubischen Bedingungen beschrieben werden:

$$\mathbf{E} \mathbf{E}^T \mathbf{E} - \frac{1}{2} \text{Spur}(\mathbf{E} \mathbf{E}^T) \mathbf{E} = 0 \tag{2}$$

Die Lösung des Systems von Gleichungen basiert auf Singulärwertzerlegung. Für fünf Punkte erhält man einen rechten Nullraum aus vier Vektoren. Die essentielle Matrix entspricht einer Linearkombination der vier Vektoren. Zur Bestimmung der Skalare der Linearkombination werden sie in die kubischen Bedingungen der Gleichungen (1) und (2) eingesetzt und es wird eine Gauß-Jordan Elimination durchgeführt. Weitere Umformungen führen zu einem Polynom zehnten Grades. NISTÉR (2004) verwendet zur schnellen Lösung Sturm-Sequenzen.

In (NISTÉR 2004) wird gezeigt, dass bei realen Beispielen, wenn auch sehr selten, bis zu zehn Lösungen auftreten können. In den allermeisten Fällen gibt es allerdings zwischen einer und vier Lösungen. Während die Bestimmung der Fundamentalmatrix für eine Vorwärtsbewegung etwas stabiler sein kann, hat die essentielle Matrix für Seitwärtsbewegungen große Vorteile, so dass insgesamt die essentielle Matrix im Allgemeinfall vorzuziehen ist.

Achtung: Die Bestimmung einer „kalibrierten“ Fundamentalmatrix ist nicht gleichzusetzen mit der Bestimmung der essentiellen Matrix, da hierbei die Bedingungen aus Gleichung (2) nicht eingehalten werden. Praktisch führt dies zum Problem der Chiralität (HARTLEY & ZISSERMAN 2003), das darin besteht, dass Teile der Punkte vor und andere Teile hinter der Kamera liegen.

Dies ist kein Fehler in der Berechnung, sondern entspricht einer korrekten projektiven Rekonstruktion, wie sie implizit mittels der Fundamentalmatrix beschrieben wird.

Mittels des 5-Punkt Algorithmus kann direkt eine Lösung für die relative Orientierung bestimmt werden. Praktisch tritt aber das Problem auf, dass bei der Zuordnung von Punkten viele Fehler auftreten. In schwierigeren Fällen mit sich wiederholender Textur und großen Basen können z.B. 95% der Punkte inkorrekt sein. Um trotzdem eine richtige Lösung zu erhalten, eignet sich besonders das Random Sample Consensus – RANSAC (FISCHLER & BOLLES 1981) Verfahren.

Dessen grundlegende Idee ist konträr zu vielen anderen robusten Verfahren: Statt aus allen Beobachtungen eine Lösung zu berechnen und dann schlecht passende Beobachtungen zu eliminieren, wird zufällig ein Minimalsatz von Beobachtungen (z.B. kann aus drei Punkten ein Kreis bestimmt werden) zur Lösung des Problems ausgewählt und dann bewertet, wie gut die entsprechende Lösung zu den Beobachtungen passt (d.h., für obiges Beispiel, ob die Punkte nah genug am Kreis liegen). Dies geschieht in einer mehr oder weniger großen Zahl von Iterationen und am Ende wird die am besten bewertete Lösung als Ergebnis verwendet.

Der Vorteil ist, dass man für Probleme mit kleinen Minimalsätzen von Beobachtungen, d.h., für die die Kombinatorik gut ist, z.T. auch bei nur 5% oder weniger korrekten Daten eine stabile Lösung erhalten kann. Neben der Bedingung, dass der Minimalsatz klein sein sollte, ist ein Kriterium für die Evaluierung der Lösung notwendig. So muss bei dem obigen Beispiel bekannt sein, bis zu welchem Abstand von einem Kreisring ein Punkt auf dem Kreis liegt. Diese Kriterien basieren z.T. auf aus Gleichung abschätzbaren Messgenauigkeiten, sind aber manchmal auch anwendungsabhängig. So hat z.B. die Frage, ob Punkte einer Wand auf einer Ebene liegen, weniger mit der Mess- als vielmehr mit der Baugenauigkeit zu tun, die nicht direkt aus den Messungen ableitbar ist.

3 Orientierung von Bildsequenzen

Die oben beschriebenen 5-Punkt Algorithmus und RANSAC suchen aus Mengen von fünf Punktpaaren möglicher Zuordnungen die Menge aus, die bei der vorgegebenen Kalibrierung zur besten – d.h., möglichst viele Punkte korrespondieren möglichst genau – Orientierung führt.

Die korrespondierenden Punkte werden zuerst einzeln in den Bildern mit dem Förstner Operator (FÖRSTNER & GÜLCH 1987) bestimmt. Für die Punkte wird Rotationsinvarianz über die Nutzung von Richtungsinformation in Form des Eigenvektors des Förstner Operators erreicht. Die entsprechend des Eigenvektors rotierten Punkte werden durch farbige Kreuzkorrelation verglichen. Man erhält ein Histogramm der relativen Winkel der zugeordneten Punkte. Der zum Maximum des Histogramms korrespondierende Winkel wird als maximaler Konsens und damit als Schätzung der relativen Rotation der beiden Bilder gegeneinander betrachtet. Hiermit werden im Weiteren alle Bildpunkte des zweiten Bildes gedreht und es wird noch einmal die Kreuzkorrelation durchgeführt. Liegt der Korrelationskoeffizient oberhalb eines empirisch bestimmten Schwellwertes von 0.5, so erfolgt eine kleinste Quadrate Zuordnung, nach der Punktpaare auf Grundlage von Genauigkeit und Korrelation > 0.9 aussortiert werden.

Für drei Bilder, die wegen der im Vergleich zu Paaren deutlich erhöhten Zuverlässigkeit der Orientierungsbestimmung das Grundelement des vorgestellten Ansatzes bilden, da hier die Punkte über geometrische Schnittbedingungen eindeutig kontrolliert werden, können Trifokal-

tensoren verwendet werden (HARTLEY & ZISSERMAN 2003). Allerdings gibt es (bisher) keine direkte Lösung für den kalibrierten Trifokaltensor. Deswegen wird hier zweimal der 5-Punkt Algorithmus verwendet. Hierbei darf allerdings nicht wie bei (relativen) Bildpaaren zweimal die Basis auf eins gesetzt werden, sondern es muss die zweite Basis an die erste angepasst werden. Die verwendete Lösung besteht darin, dass für die durch die Zuordnung vorliegenden fünf Punkttupel jeweils in den zwei Bildpaaren die Punkte 3D trianguliert werden und für das gemeinsame Bild die Abstände bestimmt werden. Das Verhältnis der Abstände entspricht dem Verhältnis der Längen der Basen. Zur Verbesserung der Robustheit wird der Median verwendet. Für vier und mehr Bilder werden Tripel über die Projektionsmatrizen mit Hilfe von 3D Homographien verknüpft und durch kleinste Quadrate Zuordnung Mehrfachpunkte bestimmt. Die Kombination von Tripeln zu längeren Sequenzen erfolgt hierarchisch mit einer Überlappung von zwei Bildern. D.h., aus Tripeln werden vier Bilder, aus vier werden sechs, aus sechs zehn, achtzehn, vierunddreißig Bilder, etc.

Zur Beschleunigung der Berechnung werden Bildpyramiden verwendet. Bildpaare werden auf einer Bildgröße von ca. 100×100 Pixel orientiert, Bildtripel auf der doppelten und bei größeren Bildern auch auf der vierfachen Größe. Nach der Orientierung der gesamten Sequenz auf dieser Stufe werden am Ende die Punkte durch die Pyramide in die originale Auflösung verfolgt.

Zentraler Bestandteil des vorgestellten Ansatzes ist neben der hoch genauen, auf kleinsten Quadraten basierenden Bildzuordnung die robuste Bündelausgleichung (MCGLONE et al. 2004). Diese wird für jeden Schritt der Orientierung durchgeführt.

Weitere Details zu dem Ansatz und den erzielbaren Genauigkeiten sind in (MAYER 2005) dargestellt. Neuere Ergebnisse (MAYER 2007) zeigen, dass bei Schleifen von fünfzig bis hundert Bildern der Größe fünf Megapixel Genauigkeiten für die Translation im Bereich von unter 1 % des Durchmessers der Schleife und für die Rotation von wenigen hundertstel Grad pro Bild erzielt werden können.

4 Unmögliches

Abbildung 1 stellt eine typische geblitzte Aufnahme des photogrammetrischen Nahbereiches dar. Werden in diesem Fall die Codes der Punkte nicht gelesen, so ist jeder Punkt zu jedem anderen zuordnbar. Dies führt z.B. dazu, dass bei zwanzig Punkten, für die nur eine Zuordnung richtig sein kann, nur maximal 5% der Punkte korrekt sein können. Da meist deutlich mehr Punkte im Bild vorhanden sind, erhält man oft unter einem Prozent korrekte Punkte, was bei RANSAC zu vielen Iterationen und einer hohen Unzuverlässigkeit führt.

Insgesamt ist feststellbar, dass bei der im Nahbereich oft gewählten Vorgehensweise von geblitzten völlig überstrahlten Bildern, bei denen praktisch keine Hintergrundtextur vorhanden ist, die oben vorgestellten direkten Verfahren in Kombination mit RANSAC oft eine niedrige Erfolgswahrscheinlichkeit besitzen

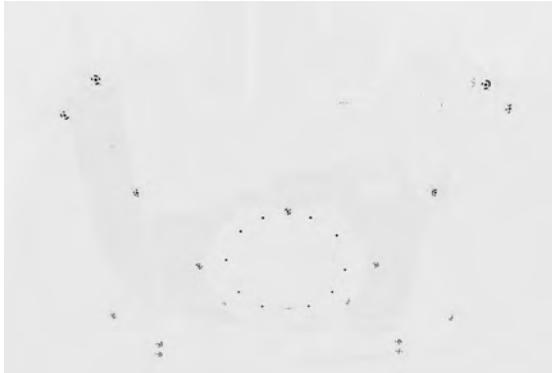


Abbildung 1: Typische geblitzte Nahbereichsaufnahme in inverser Darstellung: Ohne Lesen der Kodierung kann jeder Punkt zu jedem anderen zugeordnet werden. Das Bild wurde von der Firma SOLVing3D dankenswerterweise zur Verfügung gestellt.

5 Lösungsmöglichkeit

Im Weiteren werden Ergebnisse für eine Szene vorgestellt, die sowohl Bildtextur als auch Messmarken enthält und ein typisches Objekt des photogrammetrischen Nahbereichs, nämlich eine Autotür zeigt. In Abbildung 2 sind die Bilder 12 und 13 der Sequenz dargestellt. Diese wurden, obwohl sie stark gegeneinander verdreht sind, auf Grundlage der Information der Eigenvektoren des Förstner Operators korrekt zugeordnet.



Abbildung 2: Bilder 12 und 13 der Bildsequenz Autotür. Die Bilder wurden von der Firma SOLVing3D dankenswerterweise zur Verfügung gestellt.

Das Ergebnis für insgesamt 23 Aufnahmen ist in Abbildung 3 dargestellt. Es wurde ein σ_0 von 0.13 Pixel erzielt. Insgesamt ergaben sich 1580 3-fach, 951 4-fach, 372 5-fach, 252 6-fach, 245 7-fach, 210 8-fach, 141 9-fach, 179 10-fach, 94 11-fach und 14 12-fach Punkte. Hierbei ist zu beachten, dass Punkte nur so lange betrachtet werden, wie sie im nächsten Bild wiedergefunden werden. Punkte, die verloren gehen, werden neu initialisiert. Dieses Vorgehen vermeidet, dass in späteren Bildern verdeckte Punkte in Textur-reichen Bildern halluziniert werden.



Abbildung 3: Ergebnis für die Bildsequenz Autotür, für die die Bilder 12 und 13 in Abbildung 2 dargestellt sind – Ansicht von oben (links) und von der Seite (rechts)

Das Resultat zeigt, dass für die vorgegebenen Bilder, auch bei größeren Rotationen und projektiven Verzerrungen eine Zuordnung mit hoher Genauigkeit möglich ist. Der nächste Schritt, an dem im Moment gearbeitet wird, besteht darin, diese Ergebnisse dazu zu verwenden, um die (unkodierten) Messmarken zwischen den Bildern auf Grund der bekannten Epipolargeometrie sowie von aus den kalibrierten Kameramatrizen ableitbaren Trifokalensoren zuzuordnen. Dies würde einen Vergleich der erzielbaren Genauigkeiten ermöglichen. Eine Hypothese besteht hierbei darin, dass ähnlich wie bei der automatischen Aerotriangulation die Orientierungen bei genügend Textur auch ohne Messmarken sehr genau bestimmt werden kann, auch wenn die einzelnen Punkte (natürlich) weniger genau bestimmt werden.

6 Ausblick

Dieser Beitrag stellt dar, dass es für Bilder mit ausreichend natürlicher Textur möglich ist, ohne Näherungen bei gegebener (genäherter) Kalibrierung ein genaues (relatives) 3D Modell auch bei großen Basen abzuleiten. Als Grundlagen hierfür wurden der 5-Punkt Algorithmus von NISTÉR (2004) und RANSAC vorgestellt. Es wurde gezeigt, wie diese beiden zusammen mit hoch genauer Bildzuordnung von Punkten und der Verknüpfung zu Tripeln und n-fach Punkten zur genauen und verlässlichen Orientierung von Bildsequenzen führen. Darauf aufbauend wurde vorgeschlagen, dies mit (unkodierten) Messmarken zu kombinieren, was den Grad der Automatisierung erhöhen und damit neue Anwendungen eröffnen sollte.

Defizite des Ansatzes bestehen zuerst darin, dass die Ergebnisse nicht Maßstabs-invariant sind. Erfahrungsgemäß sind Unterschiede im Maßstab oft nur bis ca. 30% tolerierbar. Eine Möglichkeit besteht im SIFT Operator (LOWE 2004). Da die Erfahrungen mit der Erweiterung des Förstner Operators in Richtung Rotationsinvarianz gut sind, gibt es Überlegungen, diesen durch Maßstabsraum-Normierung Maßstabs-invariant zu machen. Hierbei soll nicht jeder einzelne Punkt auf verschiedenen Maßstäben zugeordnet, sondern zuerst wie bei der Rotation mittels maximalen Konsenses robust ein Maßstabsunterschied für das Bildpaar bestimmt werden.

Weiterhin ist der Ansatz dahingehend zu erweitern, dass nicht mehr nur Sequenzen, sondern auch flächenhaft verknüpfte Bildverbände orientiert werden können. Dies führt vor allem in Kombination mit Maßstabs- und Rotationsinvarianz zu einer sehr schlechten Kombinatorik, die in ähnlicher Weise angegangen werden soll wie in (SCHAFFALITZKY & ZISSERMAN 2002). Zuletzt ist als zentraler Punkt dieses Papiers die Kombination des vorgestellten Ansatzes mit Messmarken zu leisten.

7 Dank

Wir danken Bernd-Michael Wolf von der Firma SOLVing3D für die Überlassung von Beispielen und für interessante Diskussionen.

8 Literaturverzeichnis

- FISCHLER, M. & BOLLES, R., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. – *Communications of the ACM* **24** (6): 381–395.
- FÖRSTNER, W. & GÜLCH, E., 1987: A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. – *ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, Schweiz, 281–305.
- HARTLEY, R. & ZISSERMAN, A., 2003: *Multiple View Geometry in Computer Vision – Second Edition*. – Cambridge University Press, Cambridge, Großbritannien.
- LHULLIER, M. & QUAN, L., 2005: A Quasi-Dense Approach to Surface Reconstruction from Uncalibrated Images. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27** (3): 418–433.
- LOWE, D., 2004: Distinctive Image Features from Scale-Invariant Keypoints. – *International Journal of Computer Vision* **60** (2): 91–110.
- MAYER, H., 2005: Robust Least-Squares Adjustment Based Orientation and Auto-Calibration of Wide-Baseline Image Sequences. – *ISPRS Workshop in conjunction with ICCV 2005 “Towards Benchmarking Automated Calibration, Orientation and Surface Reconstruction from Images” (BenCos)*, Beijing, China, 1–6.
- MAYER, H., 2007: Least-Squares Adjustment Based Markerless Orientation of Wide-Baseline Image Sequences. – Submitted to *Photogrammetric Record* **26**.
- MCGLONE, J., BETHEL, J. & MIKHAIL, E. (EDITOR), 2004: *Manual of Photogrammetry*. – American Society of Photogrammetry and Remote Sensing, Bethesda, USA.
- NISTÉR, D., 2004: An Efficient Solution to the Five-Point Relative Pose Problem. – *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26** (6): 756–770.
- POLLEFEYS, M., VAN GOOL, L., VERGAUWEN, M., VERBIEST, F., CORNELIS, K. & TOPS, J., 2004: Visual Modeling with a Hand-Held Camera. – *International Journal of Computer Vision* **59** (3): 207–232.
- REMONDINO, F. & EL-HAKIM, S., 2006: Image-Based 3D Modelling: A Review. – *Photogrammetric Record* **21** (115): 269–291.
- SCHAFFALITZKY, F. & ZISSERMAN, A., 2002: Multi-view Matching for Unordered Images Sets, or “How Do I Organize My Holiday Snaps?”. – *Seventh European Conference on Computer Vision*, Volume I, 414–431.