

Bündelausgleichung für Multikamerasysteme

JOHANNES SCHNEIDER, FALKO SCHINDLER & WOLFGANG FÖRSTNER¹

Zusammenfassung: Wir stellen einen Ansatz für eine strenge Bündelausgleichung für Multikamerasysteme vor. Hierzu verwenden wir eine minimale Repräsentation von homogenen Koordinatenvektoren für eine Maximum-Likelihood-Schätzung. Statt den Skalierungsfaktor von homogenen Vektoren durch Verwendung von euklidischen Größen zu eliminieren, werden die homogenen Koordinaten sphärisch normiert, so dass Bild- und Objektpunkte im Unendlichen repräsentierbar bleiben. Dies ermöglicht auch Bilder omnidirektionaler Kameras mit Einzelblickpunkt, wie Fisheyekameras, und weit entfernte bzw. unendlich ferne Punkte zu behandeln. Speziell Punkte am Horizont können über lange Zeiträume beobachtet werden und liefern somit eine stabile Richtungsinformation. Wir demonstrieren die praktische Umsetzung des Ansatzes anhand einer Bildfolge mit dem Multikamerasystem „Ladybug3“ von Point Grey, welches mit sechs Kameras 80 % der gesamten Sphäre abbildet.

1 Einleitung

Multikamerasysteme bestehen aus mehreren gegenseitig fest orientierten Kameras. Sie dienen der Erhöhung der Auflösung oder der Integration von Kameras unterschiedlicher spektraler Empfindlichkeiten (Z/I DMC, Vexcel Ultracam) oder der Vergrößerung des effektiven Öffnungswinkels (Blom Pictometry, Rolleiflex Panoscan Mark III). Die Orientierungsbestimmung stützt sich entweder auf

1. die als unabhängig betrachteten Bilder, was eine Näherung darstellt,
2. eine Fusion der Bilder zu einem perspektiven Bild, was bei, im Vergleich zu den Objektentfernungen, kleinen Basislängen hinreichend genau ist, oder
3. das Multikamerasystem unter Nutzung der Kenntnis der gegenseitigen Orientierung.

In diesem Artikel stellen wir für dieses letzte und auch strenge Modell einen Ansatz für eine Bündelausgleichung vor. Dazu erweitern wir in Abschnitt 2 die Kollinearitätsgleichung um die jeweilige Orientierung der Einzelkameras im Multikamerasystem, die sinnvollerweise als bekannt vorausgesetzt wird. Anschließend führen wir in Abschnitt 3 eine minimale Repräsentation für die homogenen Koordinatenvektoren der Kollinearitätsgleichung ein, so dass eine Maximum-Likelihood-Schätzung ohne Zusatzbedingungen für die Normierung der homogenen Beobachtungs- und Koordinatenvektoren gelingt. In Abschnitt 4 diskutieren wir Verfahren zur Näherungswertbestimmung bei Multikamerasystemen. Abschnitt 5 berichtet schließlich über Ergebnisse aus der strengen Bündelausgleichung einer Bildfolge aus einer Kurvenfahrt mit dem omnidirektionalen Kamerasystem Ladybug3. Wir zeigen, dass dank des großen Sichtfeldes die Bewegung des Kamerasystems relativ zur Umgebung mit wesentlich mehr Struktur über einen großen Raumwinkelbereich bestimmt werden kann.

¹ Universität Bonn, Institut für Geodäsie und Geoinformation, Professur für Photogrammetrie, Nußallee 15, 53115 Bonn; {josch,falko.schindler}@uni-bonn.de, wf@ipb.uni-bonn.de

2 Die erweiterte Kollinearitätsgleichung

Die Kollinearitätsgleichung modelliert den Abbildungsprozess von Objektpunkten und stellt die perspektive Relation zu den beobachteten Bildkoordinaten her. Da Multikamerasysteme aus Konstruktionsgründen i. a. mehrere Projektionszentren besitzen, erweitern wir die klassische Kollinearitätsgleichung um eine Transformation M_t , welche die inverse Bewegung des Multikamerasystems im Objektkoordinatensystem beschreibt:

$$\lambda_{ijt} \mathbf{x}_{ijt} = K_j R_j [I_3 - Z_j] M_t \mathbf{X}_i = P_j M_t \mathbf{X}_i \quad (1)$$

Die Projektionsmatrizen P_j beinhalten dann die Orientierungen der Einzelkamera im Multikamerasystem sowie die innere Orientierung jeder Einzelkamera. Das Modell verlangt, dass die homogenen Bildkoordinaten \mathbf{x}_{ijt} der Objektpunkte \mathbf{X}_i , die mit der Kamera j zum Zeitpunkt t beobachtet wurden, bis auf einen beobachtungsspezifischen Skalierungsfaktor λ_{ijt} gleich der rechten Seite sind.

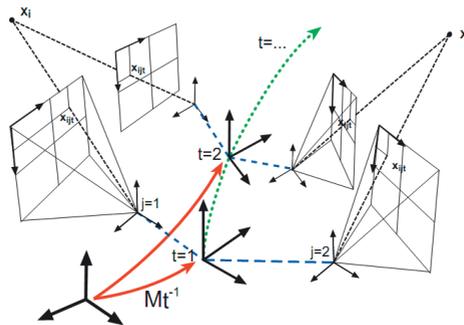


Abb. 1: Bei der skizzierten Bewegung eines Multikamerasystems werden zwei Objektpunkte X_i zu zwei Zeitpunkten t auf die Bildebenen der Kameras j abgebildet. Die Position und Orientierung des Multikamerasystems relativ zu einem übergeordneten Koordinatensystem wird durch die Bewegungsmatrix M_t beschrieben.

Die Kollinearitätsgleichung kann zur folgenden homogenen Bedingung umformuliert werden

$$\mathbf{0} = \mathbf{x}_{ijt} \times P_j M_t \mathbf{X}_i = S(\mathbf{x}_{ijt}) P_j M_t \mathbf{X}_i \quad (2)$$

wobei $S(\mathbf{x}_{ijt})$ der schiefsymmetrischen 3×3 Matrix mit den Elementen des homogenen 3×1 Vektors \mathbf{x}_{ijt} entspricht. Die Schätzung der unbekannt Parameter erfolgt aufgrund des impliziten funktionalen Zusammenhangs in Gleichung (2) zwischen den Beobachtungen und den unbekannt Parameter im Gauß-Helmert-Modell. Um lineare Abhängigkeiten zu vermeiden werden nur die zwei Zeilen verwendet, die das betragsmäßig größte Element in $S(\mathbf{x}_{ijt})$ beinhalten.

Da der unbekannte Skalierungsfaktor λ_{ijt} nicht durch die Verwendung von euklidischen Koordinaten eliminiert wird, können weiterhin Bild- und Objektpunkte im Unendlichen repräsentiert werden. Das sind Bildpunkte, die nicht auf einer Bildebene abbildbar sind, z. B. bei der Verwendung omnidirektionaler Kameras und Objektpunkte, die im Vergleich zu einer kleinen Bewegung des Multikamerasystems sehr weit entfernt liegen, wie z. B. Objektpunkte am Horizont.

3 Minimale Repräsentation

Da homogene Koordinatenvektoren i. a. singuläre Kovarianzmatrizen haben, verwenden wir eine minimale Repräsentation homogener Koordinaten, mit denen eine Maximum-Likelihood-Schätzung durchführbar ist, ohne auf euklidische Koordinaten überzugehen (FÖRSTNER, 2010). Die minimale Repräsentation eines sphärisch normierten homogenen Punktvektors erfolgt durch eine Transformation in den Tangentialraum der Einheitskugel an der Stelle des Punktvektors. Der Nullraum der Kovarianzmatrix des sphärisch normierten Koordinatenvektors entspricht der Richtung des homogenen Koordinatenvektors, liegt demnach also genau senkrecht zu dem genannten Tangentialraum (vgl. Abbildung 2). Daher wird durch diese Transformation die Kovarianzmatrix des reduzierten Koordinatenvektors regulär. Der Tangentialraum $J = [s, t]$ an den geschätzten Näherungen für die Punkte $\hat{\mathbf{x}}_{ijt}$ und entsprechend für \mathbf{X}_i ergibt sich durch

$$J_x(\hat{\mathbf{X}}_{ijt}^a) = \text{null}(\hat{\mathbf{x}}_{ijt}^{aT}) \qquad J_X(\hat{\mathbf{X}}_i^a) = \text{null}(\hat{\mathbf{X}}_i^{aT}) \qquad (3)$$

wobei J eine orthonormale Matrix darstellt. Die Koordinatenkorrekturen können in dem jeweiligen Tangentialraum der Näherungskordinaten in einem linearem Ersatzmodell geschätzt werden und anschließend wieder in den projektiven Raum durch

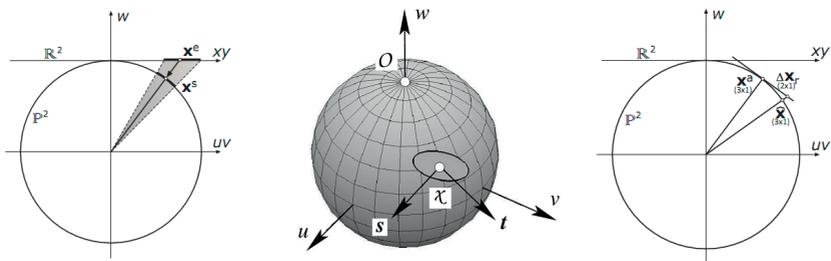


Abb. 2: Links: Durch sphärisches Normieren wird die Unsicherheit eines euklidisch normierten homogenen 2D-Punktes \mathbf{x}^e im \mathbb{R}^2 kegelförmig auf die Einheitskugel S^2 projiziert, wobei die zu \mathbf{x}^s gehörige Kovarianzmatrix auch im Falle eines unendlich fernen Punktes singulär ist. Mitte: Die Konfidenz-Ellipse eines sphärisch normierten unsicheren 2D-Punktes liegt in der Tangentialebene, die hier durch die Basisvektoren s und t aufgespannt wird. Die zugehörige 2×2 -Kovarianzmatrix ist regulär (aus FÖRSTNER, 2010). Rechts: In der Tangentialebene werden die Korrekturen $\Delta \mathbf{x}_i$ geschätzt und im Ausgleichsprozess iterativ auf die homogenen Näherungswerte \mathbf{x}^a angebracht.

$$\widehat{\mathbf{x}}_{ijt} = N(\widehat{\mathbf{x}}_{ijt}^a + J_x(\widehat{\mathbf{x}}_{ijt}^a) \widehat{\Delta \mathbf{x}}_{r,ijt}) \quad \widehat{\mathbf{X}}_i = N(\widehat{\mathbf{X}}_i^a + J_X(\widehat{\mathbf{X}}_i^a) \widehat{\Delta \mathbf{X}}_{r,i}) \quad (4)$$

zurück transformiert werden, worin $N(\cdot)$ das Argument auf die Länge 1 normiert. Das Verfahren wurde in MATLAB realisiert (SCHNEIDER, 2010) und bildet die Grundlage für die in Kapitel 5 dargestellten Ergebnisse.

4 Näherungswertbestimmung

Die Bündelausgleichung erfordert Näherungswerte für die Orientierungsdaten und Objektpunktkoordinaten. Falls eine dieser beiden Gruppen von Parametern bekannt ist, können die anderen durch Vorwärts- oder Rückwärtsschnitt bestimmt werden.

Bei klassischen Kameras erfolgt die Bestimmung von Näherungswerten, ausgehend von einem Bildpaar i. a. durch sukzessives Anhängen weiterer Bilder mit Rückwärtsschnitten oder skalierte relative Orientierungen. Bei großen Bildverbänden führt dies wegen der Fehlerakkumulation meist zu systematischen Fehlern. Dies kann man durch eine simultane Behandlung aller Bilder bzw. Bildpaare vermeiden.

MARTINEC (2007) stellt ein zweistufiges Verfahren vor, mit welchem eine simultane Mehrbildrekonstruktion ausgehend von relativ orientierten Bildpaaren gelingt. Zunächst werden konsistente Rotationsmatrizen \mathbf{R}_i der Kameras im Objektkoordinatensystem durch eine lineare kleinste Quadrate Ausgleichung auf der Basis der berechneten relativen Rotationen \mathbf{R}_{ij} aus allen linearen Bedingungen $\mathbf{R}_i = \mathbf{R}_{ij} \mathbf{R}_j$ bestimmt. Anschließend können die Kameratranslationen der äußeren Orientierung und die Objektpunktkoordinaten durch Einführung unbekannter Skalierungsfaktoren für die Raumstrahlen auf der Basis aller Raumrichtungen \mathbf{r}_{ij} in einer linearen Ausgleichung analog aus allen linearen Bedingungen $\mathbf{X}_i - \lambda_{ij} \mathbf{r}_{ij} = \mathbf{X}_j$ bestimmt werden.

Das Verfahren kann auf Multikamerasysteme übertragen werden. Wir gehen davon aus, dass zum einen die Orientierungen zwischen den Einzelkameras des Multikamerasystems bekannt und die Einzelkameras kalibriert sind. Zur Extraktion und Zuordnung von markanten, korrespondierenden Bildpunkten wird der SIFT-Operator von D. LOWE (2004) verwendet. Relative Orientierungen zwischen den Bildpaaren werden mittels des direkten 5-Punkte-Algorithmus nach D. NISTÉR (2004) in Kombination mit dem RANSAC-Verfahren nach FISCHLER&BOLLES (1981) robust bestimmt. Während der Bestimmung der relativen Orientierungen können die Punktzuordnungen gleichzeitig von groben Fehlern bereinigt werden.

Die relativen Rotationen erhält man entweder aus der Kalibrierung des Kamerasystems oder durch relative Orientierungen überlappter Bilder in gleicher Weise wie bei MARTINEC. Die Positionen der Multikamerasysteme können neben den unbekanntem Objektkoordinaten mittels der so bestimmten Rotationen der Multikamerasysteme und den bereinigten Punktzuordnungen in einem linearen Gleichungssystem bestimmt werden. Hierzu müssen lediglich die Skalierungen λ_{ijt} bzw. $\lambda_{j(ij)'}'$ der Kamerastrahlen \mathbf{r}_{ijt} und der Basislinien $\mathbf{b}_{j(ij)'}'$ als unbekannte Parameter mit eingeführt werden:

$$- \mathbf{q}_{jt} = \mathbf{T}_t + \lambda_{ijt} \mathbf{r}_{ijt} - \mathbf{X}_i \quad (5)$$

$$\mathbf{q}_{(jt)'} - \mathbf{q}_{jt} = \mathbf{T}_t + \lambda_{j(ij)'} \mathbf{b}_{j(ij)'} - \mathbf{T}_{t'} \quad (6)$$

wobei q_{jt} die Translationsvektoren vom Multikamerasystem t zum Projektionszentrum der Kamera j , T_t die Ortsvektoren der Multikamerasysteme und X_j die Objektpunkte sind. Alle Größen sind im Objektkoordinatensystem orientiert, Abbildung 3 verdeutlicht die Geometrie.

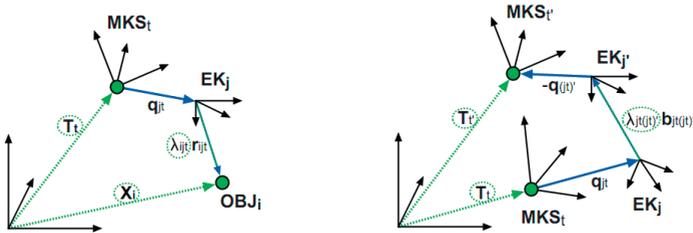


Abb. 3: Links: Der Schleifenschluss aus Translation T_t des Multikamerasystems MKS_t , der relativen Position q_{jt} der Einzelkamera EK_j und der Bewegung des an der Position X_j befindlichen Objektpunktes OBJ_j auf dem Sichtstrahl r_{jt} kann als Bedingung in einem Ausgleichungsmodell formuliert werden. Rechts: Durch die relative Orientierung zwischen Einzelkameras EK_j und EK_j' sind Basislinien $b_{j(j')}$ bis auf einen unbekanntem Skalierungsfaktor $\lambda_{j(j')}$ bekannt. Die Schleife von Translationen T_t , relativen Positionen q und der Basis b muss sich ebenfalls schließen.

Um zuverlässige Eingangsgrößen für die Ausgleichungsaufgabe zu erhalten, sind noch zusätzliche Prüfungen erforderlich, die unaufgedeckte Fehlzuordnungen und grobe Abweichungen beseitigen. Etwa sind Bildtripel zu prüfen, um eine zuverlässigere Suche nach Ausreißern zu gewährleisten und Punkte, die nicht vor den jeweiligen Kameras liegen, müssen eliminiert werden. Abbildung 4 zeigt, dass dieser Ansatz eine direkte Lösung für die Objektpunkte und die Kameraorientierungen aus Bildfolgen liefern kann. Bei diesem Datensatz handelt es sich um eine Feldbefliegung mit einer Drohne mit einer einzelnen Kamera, bei welchem es zu wenigen Fehlzuordnungen kommt. Die o. g. Prüfungen sind bei Kameraanordnungen mit kleinen Basen, wie in unserer Anwendung, nicht ausreichend, hier sind etwa die Vorzeichenbedingungen für die relative Lage der Punkte zu den Kameras in die gemeinsame Ausgleichung zu integrieren.

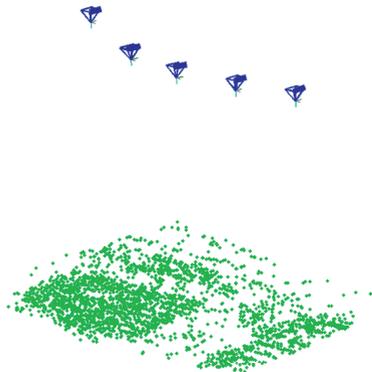


Abb. 4: Räumliche Darstellung der automatisch bestimmten Objektpunkte und der äußeren Orientierungen der Kamerapositionen zu 5 Zeitpunkten einer Befliegung mit einer Drohne.

5 Die freie Bündelausgleichung

Mit den an den Näherungswerten linearisierten Bedingungsgleichungen (siehe Formel 2) werden die im Sinne der Statistik optimalen Schätzwerte für die homogenen Größen der Kollinearitätsgleichung iterativ bestimmt. Für das stochastische Modell der Beobachtungen nehmen wir an, dass alle Bildpunkte gegenseitig unkorreliert sind, jedoch die Koordinaten eines Bildpunktes korreliert sein können (MCGLONE, 2004, Kapitel 11). Die Kovarianzmatrix der Beobachtungen sei demnach gegeben durch eine Blockdiagonalmatrix, bei der sich die jeweiligen Blockmatrizen durch Varianzfortpflanzung bei der sphärischen Normierung und der Überführung in die minimale Repräsentation der jeweiligen Bildpunkte ergeben.

Bei einer freien Bündelausgleichung werden keine Kontrollinformationen eingeführt oder sonstige Zwänge auferlegt. Das Netz wird auf die Näherungskoodinaten der Objektpunkte aufgefördert, um den Datumsdefekt zu beheben. Der Maßstab ist zwar durch die Basen zwischen den Einzelkameras gegeben, jedoch kann die räumliche Ausdehnung des Netzes verglichen zu den Basen sehr groß sein, was zu einem schlecht gestellten Normalgleichungssystem führt. Daher wird die Bedingung auf den Maßstab entsprechend schwach gewichtet.

5.1 Testdaten

Wir haben eine 15 Aufnahmen umfassende Bildfolge mit dem Multikamerasystem „Ladybug3“ der Firma Point Grey aufgenommen, welches mit sechs Kameras 80 % der gesamten Sphäre abbildet (siehe Abbildung 5). Das Kamerasystem wurde auf einen Handwagen montiert und um eine Kurve gefahren. Zwei der jeweils sechs Aufnahmen, die vom Operateur und vom Zenit, wurden von der Auswertung ausgeschlossen.

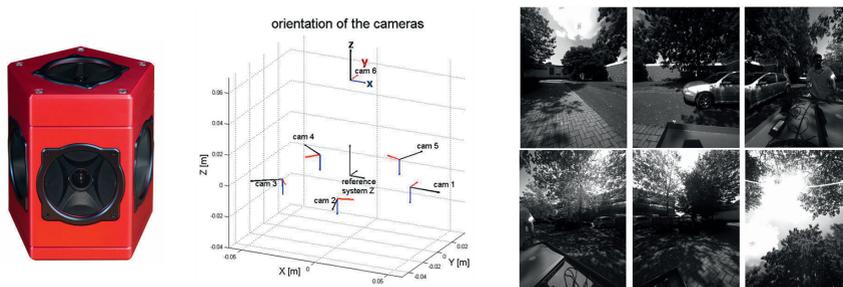


Abb. 5: Links: Das Multikamerasystem Ladybug3. Fünf Kameras sind horizontal angeordnet, eine Kamera blickt nach oben. Mit dieser Konfiguration ist die komplette obere Hemisphäre abgedeckt. Mitte: Orientierungen der lokalen Koordinatensysteme der sechs Einzelkameras und das Referenzsystem. Rechts: Bilder der sechs Kameras einer Aufnahme des Systems.

Für die Näherungswertbestimmung verwenden wir für dieses Experiment das institutseigene Programm „Aurelo“, welches einen Bildverband vollautomatisch orientiert (LÄBE, 2005). Dieses wurde vor kurzem weiterentwickelt, so dass auch omnidirektionale Kameras mit einem Projektionszentrum verwendet werden können. Da die Einzelkameras der Ladybug3 kein

einheitliches Projektionszentrum haben, werden auf die Strahlenbündel entfernungsabhängige Korrekturen angebracht (SCHMEING, 2011).

5.2 Ergebnisse

Im iterativen Ausgleichsprozess wird die Abweichungsquadratsumme minimiert und an den verbesserten Näherungswerten linearisiert. Häufig bleiben trotz der mittels RANSAC bereinigten

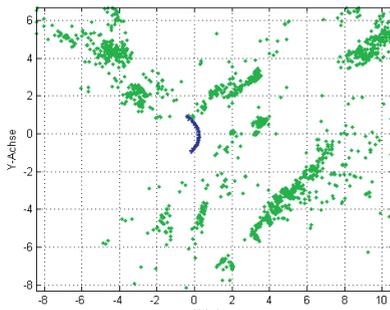


Abb. 6: 2D-Visualisierung der geschätzten Multikamerasystemstandpunkte (blau) und Objektpunkte (grün)

Punktzuordnungen fehlerhafte Zuordnungen vorhanden. Daher werden Beobachtungen, welche auf der Bildebene Verbesserungen größer $3\sigma_0$ Pixel erhalten, entfernt. Die restlichen Abweichungen werden nach der L2-Norm minimiert. Der a priori Varianzfaktor wird mit $\sigma_0 = 1$ Pixel angesetzt. Die Bündelausgleichung konvergiert nach sechs Iterationen. Der geschätzte Varianzfaktor ergibt sich zu $\sigma_0 = 0.62$ Pixel, was einen realistischen Schätzwert darstellt. Ein Ausschnitt aus den geschätzten Lagekoordinaten der Objektpunkte und Multikamerasysteme ist in Abbildung 6 dargestellt. In

Abbildung 7 sind die Schätzgrößen dreidimensional visualisiert. Die Kurvenfahrt des Multikamerasystems ist deutlich zu erkennen. Aufgrund des omnidirektionalen Sichtfeldes des Kamerasystems konnten Objektpunkte aus allen Richtungen aufgenommen werden.

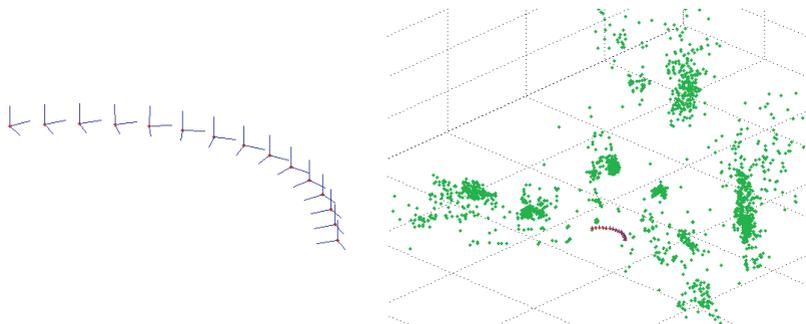


Abb. 7: Links: Orientierung des Multikamerasystems im globalen System während einer Kurvenfahrt. Rechts: Visualisierung der Trajektorie des Systems und der rekonstruierten Objektpunkte.

Die Verbesserungen der Bildkoordinaten sind nicht normalverteilt, wie es in dem Histogramm und dem Wahrscheinlichkeitsnetz in Abbildung 8 zu erkennen ist. Hier haben wird die

normierten Verbesserungen v/σ_v für alle Punkte mit drei oder mehr Strahlen verwendet. In der Mitte des Wahrscheinlichkeitsnetzes zeigen die Verbesserungen ein nahezu lineares Verhalten, sind also näherungsweise normalverteilt, während an den Flanken klare Hinweise auf nicht erkannte Ausreißer zu erkennen sind.

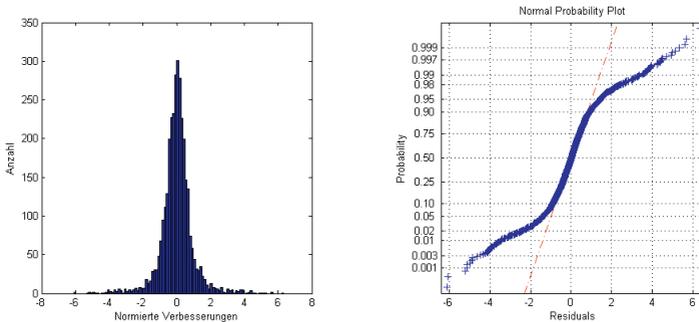


Abb. 8: Links: Histogramm über die geschätzten normierten Verbesserungen. Rechts: Darstellung der geschätzten normierten Verbesserungen in einem Wahrscheinlichkeitsnetz der Normalverteilung mit $\mu = 0$ und $\sigma = 0.62$. Wäre die Modellannahme normalverteilter Residuen vollständig erfüllt, lägen alle Datenpunkte auf einer Geraden. Dies ist nur im mittleren Bereich der Kurve der Fall.

6 Zusammenfassung und Ausblick

In diesem Aufsatz wurde ein Ansatz zur Formulierung einer Bündelausgleichung und erste Ergebnisse zur vollautomatischen Orientierung eines Multikamerasystems vorgestellt. Die Bündelausgleichung ist im statistischen Sinne optimal, da sie alle verfügbaren Informationen mit einbezieht und die stochastischen Eigenschaften der Beobachtungen berücksichtigt.

Die Bedingungsgleichungen für die Ausgleichung wurden mit homogenen Größen formuliert. Die Singularitäten der Kovarianzmatrizen der homogenen Koordinatenvektoren, welche keine Maximum-Likelihood-Schätzung zulassen, konnten durch eine minimale Repräsentation beseitigt werden. Somit sind wir in der Lage auch unendlich ferne Punkte in der Bündelausgleichung zu verwenden. Die Beobachtung unendlich ferner Punkte eignet sich besonders aufgrund der hohen Richtungsgenauigkeit und der Möglichkeit, diese über einen langen Zeitraum zu verfolgen, im besonderen Maße bei omnidirektionalen Kameras wie der Ladybug3.

Unser Ziel ist es, die Implementation der Bündelausgleichung in ein Programmsystem zu integrieren, welches autonom beliebig konstruierte Multikamerasysteme vollautomatisch orientieren kann.

7 Literaturverzeichnis

- FISCHLER, R. & BOLLES, R. C., 1981: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, Vol. 24, S. 381-395.
- FÖRSTNER, W., 2010: Minimal Representations for Uncertainty and Estimation in Projective Spaces. *Proceedings of Asian Conference on Computer Vision*.
- LÄBE, T. & FÖRSTNER, W., 2005: Erfahrungen mit einem neuen vollautomatischen Verfahren zur Orientierung digitaler Bilder. *DGPF Tagungsband 2005*, Rostock.
- LOWE, D., 2004: Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60** (2), S. 91-110.
- MARTINEC, D. & PAJDLA, T., 2007: Robust Rotation and Translation Estimation in Multiview Reconstruction. In *Proceedings of the Computer Vision and Pattern Recognition Conference 2007*.
- MCGLONE, J. C., editor, 2004: *Manual of Photogrammetry*. Fifth edition, American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland, USA.
- NISTÉR, D., 2004: An Efficient Solution to the Five-Point Relative Pose Problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **26** (6), S. 756-770.
- SCHMEING, B., LAEBE, T. & FÖRSTNER, W., 2011: Trajectory Reconstruction Using Long Sequences Of Digital Images From An Omnidirectional Camera. *DGPF Tagungsband 2011*, Mainz.
- SCHNEIDER, J., 2010: Entwicklung eines Modells zur Schätzung der Trajektorie eines Multikamerasystems mit Hilfe einer Bündelausgleichung. Hausarbeit am Institut für Geodäsie und Geoinformation, Universität Bonn.