

Präzise Berechnung von Kameraposen in Manhattan-Welten

Darko Vihar^{1,2}, Rico Nestler^{1,3}, Karl-Heinz Franke¹

¹Zentrum für Bild- und Signalverarbeitung e. V.,
Werner-von-Siemens-Straße 10, 98693 Ilmenau

²TU Ilmenau, Fakultät für Informatik und Automatisierung,

³TU Ilmenau, Fakultät für Maschinenbau, FG QBV,
Postfach 10 05 65, 98684 Ilmenau

eMail: darko.vihar@zbs-ilmenau.de, rico.nestler@tu-ilmenau.de,
karl-heinz.franke@zbs-ilmenau.de

Zusammenfassung. Im Beitrag wird das Prinzip der szenenbezogenen Kalibrierung in Manhattan-Welten vorgestellt und in exemplarischen Einsatzsituationen bewertet. Die Extraktion von extrinsischen Kameraparameter aus Fluchtpunkten stellt insbesondere für große urbane oder industrielle Arbeitsräume eine Alternative zu den klassischen targetbezogenen Kalibrierverfahren dar. Es werden Fehlereinflüsse bei der Berechnung von Kameraposen analysiert und allgemeine Einsatzbeschränkungen dieses Ansatzes abgeleitet. Darüber hinaus werden Methoden für die Berechnung der Lage und Orientierung mehrerer Kameras zum Welt-Koordinatensystem vorgestellt und die Wirkung unpräziser bzw. falsch berechneter Fluchtpunkte beschrieben. Das vorgestellte Verfahren wird mit realen Aufnahmen eines Anwendungsfalls der Mensch-Roboter-Kooperation (MRK) getestet und mit einer Perspective-n-Point (PnP)-Methode verglichen.

1 Einleitung und Motivation

Die Voraussetzung und eine gleichzeitig sehr wichtige Herausforderung für die 3D-Erfassung mit mehreren Kameras nach dem geometrischen Prinzip der Triangulation ist die präzise Bestimmung der inneren Kameraparameter sowie der räumlichen Bezüge der Kameras in einem gemeinsamen Koordinatensystem.

Gängige Praxis zur Bestimmung der (intrinsischen) Abbildungseigenschaften einer Kamera ist die Verwendung eines geometrisch bekannten Kalibrierkörpers, der aus verschiedenen Blickrichtungen aufgenommen wird. Aus bekannten Welt- zu ermittelten Bildpunktkorrespondenzen lassen sich die inneren Kameraparameter z.B. durch das Prinzip des Bündelausgleichs sehr genau ermitteln.

Analog kann auch für die Bestimmung der *extrinsischen* Parameter, der geometrischen Lage und Orientierung der Kameras in Weltkoordinaten, vorgegangen werden. Jedoch ist bei großen Arbeitsräumen, z.B. im urbanen oder industriellen Bereich, und den dort erforderlichen größeren Abständen der Kameras diese Form der Kalibrierung problematisch. Die wesentliche Ursache dafür ist, das gut handhabbare Kalibrierobjekte, die im gesamten Messraum positioniert werden müssen, in den Kamerabildern oft zu klein oder auch nur partiell abgebildet werden können. Damit wird die



Abb. 1: Demonstrator für eine MRK im ZBS e.V.. Die im linken Bild segmentierten Liniensegmente sind farblich entsprechend der zugeordneten Achsenrichtungen eines Manhattan-Weltmodells dargestellt. Das Bild rechts zeigt ein Weltkoordinatensystem, das aus Bildmerkmalen Fluchtpunkten abgeleitet werden kann.

Bestimmung der Passpunktpositionen auflösungsbedingt nur unvollständig und ungenau. Weiterhin stören in konkreten Szenen Objekte, wie z.B. Maschinen und Einrichtungen im industriellen Umfeld, die Positionierbarkeit und Sichtbarkeit der Kalibriertargets erheblich. Damit kann bei dieser Verfahrensweise im praktischen Einsatz oft nur einen Teil des Messraumvolumens in die Systemkalibrierung einbezogen werden.

Für urbane oder industrielle Umgebungen ist typisch, dass eine Vielzahl von 3D-Objektkanten in der Szene paarweise parallel auftreten. Lassen sich in diesem Fall die Parallelenscharen drei orthogonalen Richtungen zuordnen, wie sie bei 3D-Objekten mit hauptsächlich rechten Winkeln auftreten, wird dies als „Manhattan-World“ beziehungsweise „Manhattan-Weltmodell“ bezeichnet. Bei der perspektivischen Abbildung bilden die Eigenschaften dieses Weltmodells die Grundlage für den in diesem Beitrag beschriebenen Ansatz der *szenenbezogenen Kamerakalibrierung*. Hierbei sollen aus verzeichnungsfrei korrigierten Aufnahmen einer Manhattan-Welt mit intrinsisch vorkalibrierten Kameras die Kameraposen bezüglich des Manhattan-Weltkoordinatensystems der Szene und anschließend die Lage und Orientierung der Kameras zueinander nur aus Szenenmerkmalen berechnet werden. Das hier beschriebene Verfahren ist Bestandteil eines Kalibrierungs-Toolkits des ZBS e.V., das eine weitestgehend automatische extrinsische Kalibrierung von Kameras ohne zusätzliches Expertenwissen erlaubt.

3 Theorie

3.1 Lochkameramodell und klassische targetbezogene Kalibrierung

Die Grundlage für den in diesem Beitrag vorgeschlagenen Ansatz bildet die perspektivische Abbildung von Manhattan-Welten modelliert durch ein ideales Lochkameramodell. Es definiert die Transformation des dreidimensionalen euklidischen Raums auf die zweidimensionale Bildebene nach dem Prinzip der Zentralprojektion. Wenn zunächst das 3D-Koordinatensystem der Welt gleich dem der Kamera mit dem Ursprung im Punkt C gesetzt wird, dann ist bei der Verwendung einer Notation gemäß

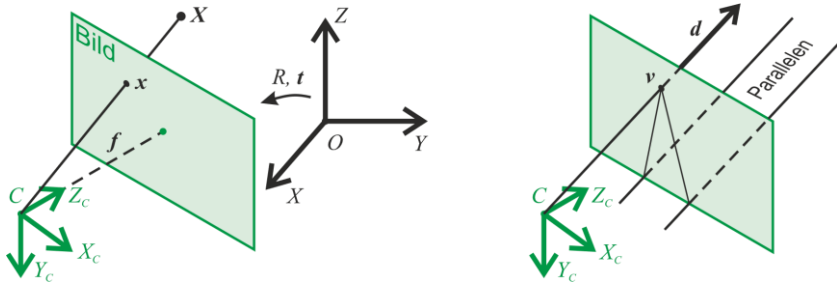


Abb. 2: Lochkameramodell (links), Fluchtpunkt aus Schnitt von 2D-Projektionen paralleler 3D-Geraden der Welt (rechts).

[1] diese Projektion kompakt mit der Gleichung

$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_C \quad (1)$$

beschrieben. Mit \mathbf{x} und \mathbf{X}_C werden die *homogenen* Koordinatenvektoren des Bildpunktes beziehungsweise des Raumpunktes bezeichnet. Diese sind bis auf einen Skalierungsfaktor eindeutig bestimmt. Der Index C verdeutlicht, dass der Raumpunkt \mathbf{X}_C im Kamerakoordinatensystem definiert ist. Die 3x3-Matrix K

$$K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

wird als Kameramatrix oder auch Kalibrierungsmatrix des Lochkameramodells bezeichnet. Sie besteht im einfachsten Fall aus der Bildweite (Brennweite) f , dem Abstand der Bildebene vom Kamerazentrum, und dem Bildhauptpunkt (p_x, p_y) .

Ist der Raumpunkt \mathbf{X} in einer 3D-Welt, z.B. einer Manhattan-Welt, definiert und die Orientierung sowie die Lage der Kamera in dieser Welt allgemein mit der Rotationsmatrix R und der Translation \mathbf{t} beschrieben, dann muss \mathbf{X} zunächst affin in das Kamera-Koordinatensystem transformiert werden $\mathbf{X}_C = \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{X}$ und wird erst im Anschluss nach Gleichung (1) auf die Bildebene projiziert. Zusammengefasst entsteht der bekannte algebraische Zusammenhang des Lochkameramodells

$$\mathbf{x} = K[R|\mathbf{t}]\mathbf{X} \quad (3)$$

Klassische Kalibrierungsansätze, wie zum Beispiel in [2], leiten die unbekanntenen Kamera-Parameter K, R, \mathbf{t} mit einem geometrischen Kalibriertarget ab, dessen bekannte Welt-Punkte \mathbf{X}_i und dazu korrespondierende Bild-Punkte \mathbf{x}_i für die Lösung von (3) verwendet werden können. Auch für die in diesem Beitrag präsentierte Vorgehensweise werden die Kameras intrinsisch (K) nach diesem Prinzip vorkalibriert.

3.2 Fluchtpunkte in einer Manhattan-Welt

Der Schnittpunkt einer Menge paralleler Geraden der Richtung \mathbf{d} liegt im 3D-Kamerakoordinatensystem im Unendlichen bzw. im uneigentlichen Punkt $\mathbf{X}_\infty = (\mathbf{d}^T, 0)^T$ des zugehörigen projektiven Raums. Die perspektivische Abbildung dieses Fernpunktes auf die Bildebene heißt *Fluchtpunkt* und wird durch Einsetzen in die Gleichung (1) als $\mathbf{v} = K\mathbf{d}$ berechnet (vgl. [1]). Analog ist die Rückprojektion des Fluchtpunktes als Strahl durch das Kamerazentrum und dem Punkt \mathbf{v} mit der Richtung $\mathbf{d} = K^{-1}\mathbf{v}$

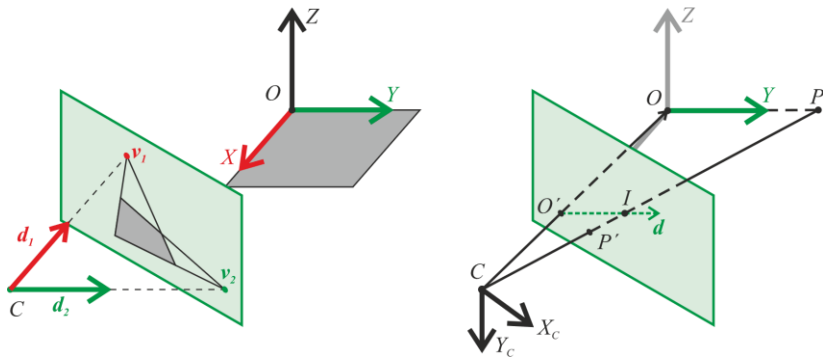


Abb. 3: Zur Herleitung der Rotation (links) und des Translationsvektors (rechts) aus Fluchtpunkten. 3D-Punkte O' und P' sind in Kamerakoordinatensystem definiert.

definiert. Folgerichtig gilt für die im Bild eines Manhattan-Objektes gefundenen Fluchtpunkte, die zu den drei senkrecht zueinander stehenden Raumrichtungen gehören, dass die Vektoren $\mathbf{d}_1 = K^{-1}\mathbf{v}_1$, $\mathbf{d}_2 = K^{-1}\mathbf{v}_2$ und $\mathbf{d}_3 = K^{-1}\mathbf{v}_3$ dazu parallel und ebenso orthogonal sind (siehe Abb. 3, links).

3.3 Ableitung der Rotationsmatrix R

Die nach den obigen Überlegungen abgeleiteten normierten Vektoren $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3$ bilden eine Orthonormalbasis eines Welt-Koordinatensystems

$$\mathbf{r}_x = \frac{\mathbf{d}_1}{\|\mathbf{d}_1\|}, \mathbf{r}_y = \frac{\mathbf{d}_2}{\|\mathbf{d}_2\|}, \mathbf{r}_z = \frac{\mathbf{d}_3}{\|\mathbf{d}_3\|}. \quad (4)$$

Die gesuchte Drehmatrix $R = [\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z]$ setzt sich spaltenweise aus diesen Basisvektoren zusammen. Hierbei ist die Reihenfolge der $\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z$ so zu wählen, dass ein rechtshändiges Koordinatensystem realisiert wird.

3.4 Ableitung der Translation t

Der Translationsvektor t kann aus der perspektivischen Abbildung ohne zusätzliches Wissen zur Szene nicht eindeutig berechnet werden. Ein Ansatz zur Bestimmung der Translation ist es, bekannte Distanzen der Szene zu verwenden. Dabei werden z.B. zwei bekannte Punkt-Paare $\{x_i, X_i\}$ zur Lösung von (3) mit der zuvor bestimmten Rotationsmatrix R verwendet. Fehler aus der Rotationsberechnung setzen sich daher in der Translation fort.

Eine Alternative dazu wurde in [4] vorgestellt. Die Grundlage bilden eine bekannte Richtung \mathbf{d} (hier: eine der Raumachsen der Manhattan-Welt), die reale Länge eines Liniensegments auf der Richtung \mathbf{d} und die Koordinaten des Start- und des Endpunktes des abgebildeten Liniensegmentes. Das Prinzip dieses Ansatzes ist in Abb. 3., rechts verdeutlicht.

Aus der Ähnlichkeit der Dreiecke $CO'I$ und COP ist die Entfernung des Punktes O vom Kamerazentrum $\|\vec{CO}\| = \|\vec{CO'}\| \|\vec{OP}\| / \|\vec{O'I}\|$ und damit der Translationsvektor

$$\mathbf{t} = \vec{CO} = \|\vec{CO}\| \frac{\vec{CO'}}{\|\vec{CO'}\|} \quad (5)$$

bestimmbar.

4 Problem der Liniensegment-Extraktion

Vorstehend wurde gezeigt, wie aus Fluchtpunkten Informationen zur extrinsischen Kamerakalibrierung abgeleitet werden können. Dabei ist die Güte und Genauigkeit der gefundenen Fluchtpunktpositionen als eingegebene primäre Information entscheidend für die Güte aller davon abgeleiteten Größen. Die Voraussetzung für eine präzise Fluchtpunkt-Berechnung ist die präzise und robuste Erkennung von Liniensegmenten (Bildkanten). Die etablierte und weitverbreitete Vorgehensweise mit einem Canny-Kantendetektor gefolgt von einer parametrischen Geradensuche im Hough-Raum ist insbesondere in Regionen mit feinen Texturen sehr störungsanfällig (falsch-positive Kanten).

Für das hier beschriebene Verfahren wurden zur Liniendetektion zwei alternative Verfahren, der „Fast Line Detector“ (FLD) nach [5] und der „Line Segment Detector“ (LSD) nach [6], untersucht. Beide geben Liniensegmente mit Subpixelgenauigkeit aus und erzeugen subjektiv korrekte Ergebnisse. Der FLD wertet das Ergebnis des Canny-Operators durch heuristische Verfolgung (Kolinearitäts-Kriterium) entlang von Konturen aus. Der LSD teilt das Grauwertbild in sogenannte Linien-Support-Regionen, d.h. Bereiche ähnlicher lokaler Gradientenorientierung, auf. Diese Regionen werden per Region-Growing als Liniensegmente zusammengeführt.

5 Fluchtpunktberechnung

Ein Fluchtpunkt entsteht als Schnittpunkt sämtlicher in ein Bild perspektivisch projizierter Raumparallelen (vgl. Abb.2 rechts). Es ist daher erforderlich, Geraden im Bild, die zu einem Fluchtpunkt gehören, zu detektieren und daraus diesen Schnittpunkt zu berechnen.

Seit dem ersten erfolgreichen Versuch von Barnard [7] im Jahr 1983 wurden viele Verfahren zur genauen Erkennung der Fluchtpunkte vorgeschlagen. Die meisten Ansätze gehen in drei Schritten vor: Erkennung der Liniensegmente im Bild (Liniensegment-Extraktion), Kanten-Clustering zu Fluchtpunkten und anschließende Fluchtpunkt-Berechnung. Insbesondere die Zuordnung der erkannten Liniensegmente zu Fluchtpunkten unterscheidet die Verfahren, da diese entweder im Ausgangsbildraum oder in einem transformierten Raum, z.B. Houghraum und Varianten oder Gauss'schen Sphere-Raum, erfolgen kann. Die bekannten Randbedingungen der Manhattan-Welt werden dabei sowohl beim Clustering als auch bei der Fluchtpunktberechnung eingesetzt.

Entsprechend eigener Untersuchungen [8] verspricht das agglomerative Clusterverfahren J-Linkage [9] [10] die besten Ergebnisse. Ähnlich einem RANSAC-Schema werden bei diesem Ansatz Fluchtpunkthypothesen über sogenannte Konsensusmengen von Kanten bewertet. Die Trennung der Cluster wird anhand der sogenannten Jaccardi-Distanz realisiert.

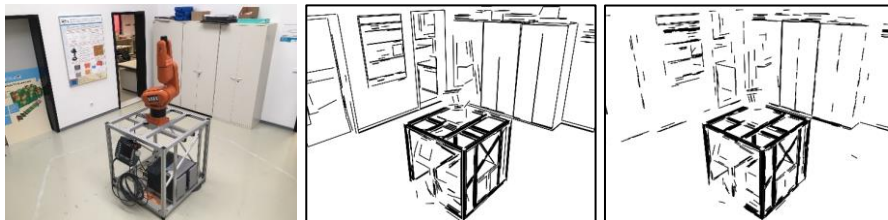


Abb. 4: Eingabebild (links), extrahierte Liniensegmente mit FLD (Mitte) und LSD (rechts).

6 Vorgeschlagenes Verfahren und Evaluierung

Das Verfahren zur Bestimmung der Kamerapose aus der Abbildung einer Manhattan-Welt wird durch folgende Schritte realisiert: Extraktion der Liniensegmente, Klassifikation der Linien und Fluchtpunktberechnung, Bestimmung der Drehmatrix und anschließende Berechnung des Translationsvektors.

Der Ablauf wird hier exemplarisch mit einer Beispielszene durchgerechnet. Die intrinsischen Parameter der verwendeten Kamera sind: $f = 1721.11\text{px}$, Hauptpunkt = $(1001.15, 753.91)$ bei einer Bildauflösung von 2016×1512 Pixeln.

6.1 Liniensegment-Extraktion

Beide Algorithmen geben bei Eingabe eines Bildes die Liniensegmente in Form subpixelgenau berechneter Start- und Endpunkte aus. Der Parameter der minimalen Liniennlänge ist kritisch und daher in der Abhängigkeit von der Auflösung des Eingabebildes und den Szeneneigenschaften entsprechend zu wählen.

Für das Beispielbild (Abb.4 links) lieferte FLD 563 und LSD 611 Liniensegmente. Entsprechend der subjektiven Bewertung ist die Konsistenz der Ergebnisse ähnlich zu bewerten. Aus der statistischen Bewertung der Liniensegmente lässt sich ableiten, dass der LSD im Vergleich zu FLD im Schnitt kürzere Liniensegmente mit geringerer Längenstreuung ausgibt. Da sich auf dieser Grundlage noch keine Präferenz ergibt, erfolgt die Verfahrensauswahl erst im Kontext mit dem nachfolgenden Verarbeitungsschritt.

6.2 Kanten-Clustering und Fluchtpunktberechnung

Bei der Eingabe von berechneten Liniensegmenten clustert der J-Linkage-Algorithmus die Liniensegmente zu dazugehörigen Fluchtpunkten. Da eine Manhattan-Welt vorausgesetzt wird, entsprechen in der Regel die drei Cluster mit der größten Anzahl zugeordneter Liniensegmente den gesuchten Raumachsen (Fluchtpunkten).

Der Schnittpunkt der gleich klassifizierten Geraden wurde mit der Methode der minimalen Fehlerquadrate bestimmt. Zusätzlich zum berechneten Fluchtpunkt wird der RMS-Fehler, über den kleinsten Abstand der Geraden zum Schnittpunkt, ausgegeben. Dieser repräsentiert ein Maß für die Genauigkeit der Kantendetektoren und gibt ein Indiz für die Lagegüte von Fluchtpunkten nahezu paralleler Liniensegmente (Abschnitt 6.3). Die Ergebnisse sind jeweils für eine Liniensegmentierung nach dem Prinzip des FLD sowie des LSD in der folgenden Tabelle dargestellt.

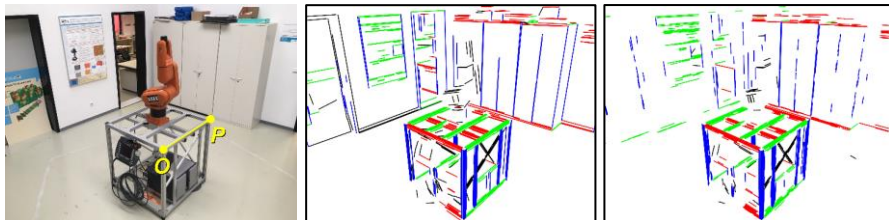


Abb. 5: Eingabebild (links), FLD-basiertes Ergebnis (Mitte) und LSD-basiertes Ergebnis (rechts). Die Linien sind farblich entsprechend den zugeordneten Fluchtpunkten dargestellt. Schwarze Linien konnten keinem Fluchtpunkt zugeordnet werden.

Verf. ($\#v_i$)	v_1	v_2	v_3
FLD (28)	(-1054.88, 43.10)	(2672.13, -20.03)	(1087.19, 4668.74)
RMS	40.57	21.86	59.54
LSD (11)	(-1087.41, 30.96)	(2653.04, -5.65116)	(1086.31, 4711.57)
RMS	27.47	21.44	36.4

Der RMS-Fehler des LSD ist stets geringfügig kleiner als der des FLD-basierten Ergebnisses. Das J-Linkage liefert nahezu doppelt so viele Cluster bei FLD (28) als bei LSD (11). Beide Ergebnisse hängen von der Präzision der Liniensegmente ab. Die Auswertung zusätzlichen Testbilder lieferte ähnliche Unterschiede stets zu Gunsten des LSD-Verfahrens, daher wird es zur Liniensegmentierung favorisiert.

6.3 Bestimmung der Rotationsmatrix R

Für die Berechnung der Drehmatrix gemäß (4) wurden die Fluchtpunkte im Uhrzeigersinn bezüglich des Bildhauptpunktes geordnet. Für die aus den Abb.5 (links) berechneten Richtungsvektoren lauten die paarweisen Winkelbeziehungen zwischen den Basisvektoren: $\angle(\mathbf{r}_1, \mathbf{r}_2) = 89.49^\circ$, $\angle(\mathbf{r}_2, \mathbf{r}_3) = 89.5^\circ$ und $\angle(\mathbf{r}_1, \mathbf{r}_3) = 90.36^\circ$. Die resultierende Drehmatrix wurde mit einer Rodriguez-Transformation orthogonalisiert.

Je nach Orientierung der Kamera in der Manhattan-Welt, kann eine Drehmatrix nach (4) nicht immer berechnet werden. Dieser Fall tritt ein, wenn eine oder alle drei Achsen des Kamerakoordinatensystems parallel zu Achsen des Welt-Koordinatensystems sind. Damit sind die Projektionen von Raumparallelen auf der Bildebene ebenfalls parallel und zugehörige Fluchtpunkte liegen auch in der Bildebene im Unendlichen.

Im Fall nur eines unendlich fernen Fluchtpunktes kann der zugehörige Richtungsvektor als Kreuzprodukt aus den anderen Richtungen berechnet werden. Die übrigen Fälle sollten möglichst vermieden werden, z.B. durch Veränderung der Kamerapose, da die weitere Auswertung andernfalls nicht möglich ist.

6.4 Bestimmung der Translation t

Die Auswertung von Gleichung (5) zur Ableitung der Translation ist trivial. Lediglich der Punkt I kann möglicherweise falsch interpretiert werden. Er entsteht als Schnittpunkt der Strahlen $C + \lambda_1 \overline{CP}^i$ und $O' + \lambda_2 \mathbf{d}$ im 3D-Raum.

Für die Berechnung des Translationsvektors t nach (5) wurden zwei Punkte auf der oberen Seite des Robotergestells festgelegt (siehe Abb. 5, links) und deren Koordinaten ermittelt. Die Seitenlänge im Raum beträgt 880mm. Damit beträgt die Entfernung der Kamera zum Ursprung der Manhattan-Welt $\|t\| = 2411.37\text{mm}$ und die Orientierung, hier zur leichteren Interpretation als Euler-Winkel (z, y', x"-Konvention), Gier = 138.98°, Nick = -0.68°, Roll = 66.72°.

6.5 Bestimmung von R und t zwischen zwei Kameras

Wird eine Manhattan-Welt aus zwei Ansichten aufgenommen und die zugehörigen extrinsischen Parameter der Kameras als $X_{C_1} = \begin{bmatrix} R_1 & t_1 \\ \mathbf{0}^T & 1 \end{bmatrix} X$ und $X_{C_2} = \begin{bmatrix} R_2 & t_2 \\ \mathbf{0}^T & 1 \end{bmatrix} X$ ermittelt, dann ist die Orientierung und Lage der Kamera C_2 bezüglich C_1

$$R_{21} = R_2 R_1^T, \quad t_{21} = t_2 - R t_1 \tag{6}$$

gegeben.

Hierbei ist zu beachten, dass die R_1, R_2, t_1 und t_2 im selben Koordinatensystem bestimmt worden sind. Im hier beschriebenen Verfahren werden die in 6.3. berechneten Drehmatrizen durch manuelle Zuweisung der Welt-Achsen zugeordnet. Zukünftig soll dieser Schritt bildmerkmalbezogen automatisch umgesetzt werden.

7 Verfahrensvergleich mit einem Referenz-Target

Um die hier beschriebene Vorgehensweise der szenenbezogenen Kalibrierung mit einem targetbasierten Verfahren zu vergleichen, wurde ein Kalibriertarget in eine Szene gelegt und bezüglich der Manhattan-Welt ausgerichtet. Die Kamera-Pose wurde zuerst mit dem vorgeschlagenen Verfahren und anschließend mit einem iterativen PnP-Algorithmus aus [3] ermittelt. Die Passpunkte des Targets wurden subpixelgenau berechnet. Die Manhattan-Welt ist in diesem Fall durch das Koordinatensystem des Kalibriertargets festgelegt. Das für die Translationsberechnung nach (5) benötigte Liniensegment wurde aus dem Target-Detektor entnommen. Die intrinsischen Parameter der Kameras entsprechen dem Beispiel im vorherigen Abschnitt.

Die Differenz der Translationsvektoren wird durch eine euklidische Norm der Vektor-Differenz $\|t_{soll} - t_{ist}\|$ bewertet. Die Differenz der Drehmatrizen wird als Winkelfehler

$$\theta = \cos^{-1} \left(\frac{\text{trace}(R_{soll} R_{ist}^T) - 1}{2} \right) \tag{7}$$

bestimmt, der eine Distanz-Metrik für Drehmatrizen umsetzt.

Die Fluchtpunkte mit dazu gehörigen RMS-Fehlern sind in der folgenden Tabelle dargestellt.

	v_1	v_2	v_3
Abb.6 oben	(818.41, 4805.54)	(-1551.14, -137.14)	(2435.64, 53.72)
RMS	43.81	40.54	30.03
Abb.6 unten	(639.56, 2826.98)	(-9053.44, -2453.98)	(1696.53, -559.244)
RMS	33.06	113.97	23.95

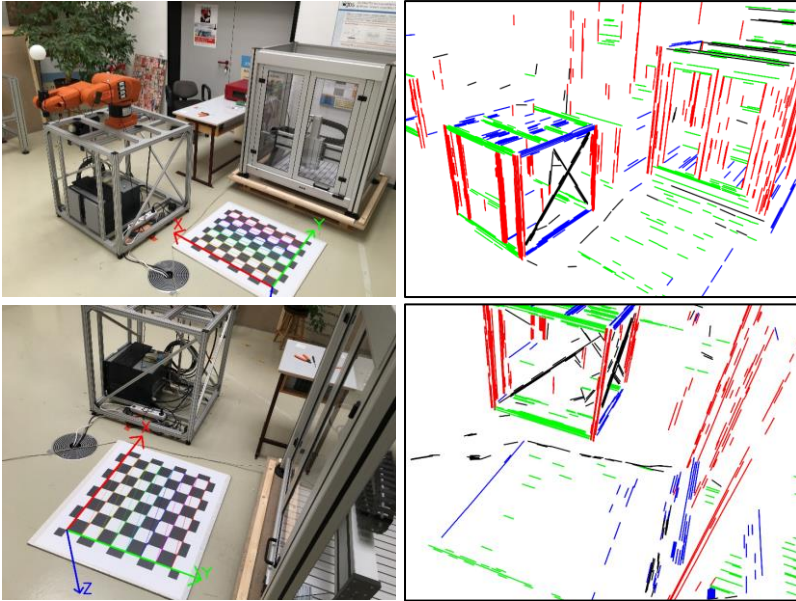


Abb. 6: Eingangsbilder zweier unterschiedlicher Kameraansichten mit dem eingezeichneten Koordinatensystem des Targets (linke Spalte) und mit den nach dem LSD-Verfahren und dem J-Linkage-Algorithmus extrahierten bzw. klassifizierten Liniensegmenten (rechte Spalte).

Für die Situation in Abb.6, unten, (fast parallele grüne Liniensegmente) liegt der berechnete Fluchtpunkt v_2 deutlich außerhalb des Bildes. Der dazugehöriger RMS-Fehler ist erwartungsgemäß am größten. Nichtsdestotrotz stehen die Basisvektoren $\angle(r_1, r_2) = 90.1^\circ$, $\angle(r_2, r_3) = 90.1^\circ$, $\angle(r_1, r_3) = 89.57^\circ$ nahezu senkrecht aufeinander, so dass an dieser Stelle keine Sonderbehandlung notwendig ist.

Die Ergebnisse der Kameraposen sowie die daraus resultierende relative Orientierung und Lage der Kameras zueinander wurden wie in den nachfolgenden Tabellen dargestellt bestimmt.

	$\ t_{soll}\ $ [mm]	$\ t_{ist}\ $ [mm]	$\ t_{soll} - t_{ist}\ $ [mm]	θ [°]
Abb.6 oben	2712.89	2710.76	2.13	0.60
Abb.6 unten	2092.83	2091.54	1.34	0.28

Die festgestellten Abweichungen der Translationen betragen nur wenige Millimeter. Der Winkelfehler der Drehmatrizen ist kleiner als ein Winkelgrad. Die Genauigkeit des vorgeschlagenen Verfahrens ist nur geringfügig schlechter als die einer PnP-Methode. Die Auswertung weiterer Aufnahmen bestätigt diese Ergebnisse. Im Falle der Stereoanordnung akkumulieren sich die Berechnungsfehler der Kameraposen.

Stereo	soll $\ t_{21}\ $ [mm]	ist $\ t_{21}\ $ [mm]	$\ \Delta t_{21}\ $ [mm]	θ_{21} [°]
Abb.6 oben und unten	2750.18	2746.08	7.57	0.33

8 Fazit und Ausblick

In Einsatzumgebungen, die eine targetbasierte extrinsische Kalibrierung nicht zulassen und die durch eine Manhattan-Welt repräsentiert sind, ist das in diesem Beitrag vorgestellte Verfahren einer szenenbezogenen Kalibrierung eine gute Einsatzalternative. Der Vergleich mit einem klassischen Verfahren zeigt, dass diese Vorgehensweise Güten erreicht, die für vielen Anwendungen geeignet sind.

In diesem Beitrag wurden stets ideale Manhattan-Welten mit drei aus der Szene ableitbaren, orthogonalen Raumachsen vorausgesetzt. Auch in Fällen, in denen aus Szenenmerkmalen nur zwei orthogonale Richtungen extrahierbar sind, können fehlende Informationen über entsprechende Heuristiken ergänzt oder fehlerhafte Informationen verbessert werden. Mit den im Beitrag entwickelten Fehlermaßen lassen sich solche Situationen sicher erkennen und nachfolgend automatisiert lösen. Diese Fragen sind Gegenstand der weiteren Forschungsarbeit auf diesem Gebiet.

9 Danksagung

Die in diesem Beitrag dargelegten Ergebnisse sind im Rahmen des Förderprogramms „zwanzig20 – Partnerschaft für Innovationen“ im Verbundvorhaben „Ergonomics Assistance Systems for Contactless Human-Machine-Operation (EASY COHMO)“ gefördert vom Bundesministerium für Bildung und Forschung entstanden.

Literatur

- [1] R. Hartley und A. Zisserman, *Multiple View Geometry in Computer Vision*, Second Edition, Cambridge University Press, 2004.
- [2] Z. Zhengyou, „A Flexible New Technique for Camera Calibration,“ *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. 22, Nr. 11, pp. 1330-1334, 2000.
- [3] G. Bradski, „The OpenCV Library,“ *Dr. Dobb's Journal of Software Tools*, 2000.
- [4] E. Guillou, D. Meneveaux, E. Maisel und K. Bouatouch, „Using vanishing points for camera calibration and coarse 3D reconstruction from a single image,“ *The Visual Computer*, Bd. 16, Nr. 7, pp. 396-410, 2000.
- [5] L. Jin Han, L. Sehyung, Z. Guoxuan, L. Jongwoo, C. Wan Kyun und S. Il Hong, „Outdoor place recognition in urban environments using straight lines,“ *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5550-5557, 2014.
- [6] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel und G. Randall, „LSD: a Line Segment Detector,“ *Image Processing On Line*, Bd. 2, pp. 35-55, 2012.
- [7] S. T. Barnard, „Interpreting Perspective Images,“ *Artificial Intelligence*, Bd. 21, Nr. 4, pp. 435-462, 1983.
- [8] L. Rehawi, „Geometrical Estimation of Multiple Cameras in a Manhattan-World,“ Bachelorarbeit TU Ilmenau - Fak. IA, FG SECSY, ZBS e.V., 2018.
- [9] R. Toldo und A. Fusiello, „Robust Multiple Structures Estimation with J-Linkage,“ in *Proceedings of the 10th European Conference on Computer Vision: Part I*, Marseille, Springer-Verlag, 2008, pp. 537-547.
- [10] J. P. Tardif, „Non-Iterative Approach for Fast and Accurate Vanishing Point Detection,“ in *IEEE 12th International Conference on Computer Vision*, Kyoto, 2009.