

Kontinuierliche Selbstkalibrierung von Stereokameras

Thao Dang, Christoph Stiller, Universität Karlsruhe

Manuskripteingang: 13. Januar 2009; zur Veröffentlichung angenommen: 19. Januar 2009

Dieser Artikel beschreibt ein Verfahren zur kontinuierlichen Selbstkalibrierung von Stereokameras und gibt in [1] gewonnene Ergebnisse in verkürzter Form wieder. Die Bedeutung der einzelnen Kameraparameter für die Tiefenrekonstruktion wird untersucht und eine theoretisch fundierte Beschreibung der gesamten Verarbeitungskette einer stereoskopischen Selbstkalibrierung geliefert. Unterschiedliche geometrische Bedingungsgleichungen – die Epipolarbedingung zwischen korrespondierenden Punkten in zwei Bildern, die trilinearen Bedingungen zwischen Bildtripeln und der Bündelausgleich bei längeren Bildverbänden – werden einheitlich in einem Gauß–Helmert-Modell formuliert, welches eine Kalibrierung mit hoher Genauigkeit erlaubt. Die Bestimmung der Kameraparameter erfolgt mit einem robusten, rekursiven Schätzverfahren für Gauß–Helmert-Modelle, basierend auf einem Iterativen Erweiterten Kalman-Filter. Es konnte gezeigt werden, dass eine Kombination verschiedener Bedingungsgleichungen die Zuverlässigkeit der Parameterschätzung erhöht. Die Leistungsfähigkeit der kontinuierlichen Selbstkalibrierung wird anhand realer Bilddaten eines aktiven Stereosystems demonstriert.

Schlagwörter: Selbstkalibrierung, Stereosehen, aktives Sehen

Continuous Self-Calibration of Stereo Cameras

This article describes an algorithm for continuous self-calibration of stereo cameras and highlights results obtained in [1]. The contribution of this work is a detailed analysis of the influence of individual camera parameters on three-dimensional reconstruction and a theoretically consistent presentation of all processing steps required for stereo self-calibration. Different geometric constraint equations – such as the epipolar constraint between corresponding points in two camera images, the trilinear constraints between image triplets and bundle adjustment for long image sequences – are formulated in a common Gauss-Helmert type model, which allows calibration with high accuracy. Estimates of the camera parameters are obtained using a robust, recursive filter algorithm for Gauss-Helmert models. This algorithm is adapted from an iterative extended Kalman filter. It is shown that a combination of different constraint equations increases the reliability of parameter estimation. The performance of the continuous self-calibration is demonstrated on real imagery of an active stereo camera system.

Keywords: Self-calibration, stereo vision, active vision

1 Einleitung

Stereosehen gewinnt durch die steigende Kapazität moderner Rechner in vielen Bereichen zunehmend an Bedeutung. Stereokameras erlauben die direkte Erfassung dreidimensionaler (3D) Information und werden beispielsweise in modernen Fahrerassistenzsystemen, welche die Sicherheit und den Komfort im Straßenverkehr erhöhen, zur Detektion und Klassifikation anderer Verkehrsteilnehmer eingesetzt. Andere Anwendungen finden sich in der Robotik und im Multimediabereich.

Um mit Hilfe von Stereobildern 3D-Messwerte zu bestimmen, ist eine exakte Kalibrierung der Kameras unerlässlich. Beispielsweise ist eine Bestimmung der relativen Orientierung der Kameras in der Praxis häufig mit einer Genauigkeit in einer Größenordnung von 10^{-2} Grad erforderlich. Herkömmlicherweise werden Kameraparameter offline mit exakt vermessenen, häufig schachbrettartigen Kalibrierobjekten bestimmt [2; 3]. In der jüngeren Vergangenheit wurden dagegen einige Arbeiten vorgestellt, in denen Methoden der Selbstkalibrierung von Mono- und Stereokameras untersucht wurden [4–7].

Unter Selbstkalibrierung von Stereokameras verstehen wir die automatische Bestimmung von Kameraparametern – wie z. B. Brennweiten und Bildhauptpunkte, aber auch relative Orientierungen der Kameras zueinander – aus nahezu beliebigen Bildsequenzen. Grundlage für eine Selbstkalibrierung von Stereokameras sind lediglich geometrische Bedingungsgleichungen zwischen korrespondierenden Bildpunkten in zeitgleich aufgenommenen Bildern der linken und rechten Kamera (räumliche Korrespondenzen) sowie korrespondierende Punkte in zeitlich aufeinander folgenden Bildern einer einzelnen Kamera (zeitliche Korrespondenzen). Spezielle Kalibrierobjekte sind hier nicht erforderlich.

Eine verlässliche Stereoselbstkalibrierung könnte die aufwendige Offline-Kalibrierung ersetzen und somit zu einer Einsparung von Zeit und Kosten in der Produktion führen. Zum anderen kann in der praktischen Anwendung nur durch eine Selbstkalibrierung die Verfügbarkeit von Stereosehen über längere Zeiträume garantiert werden. Gerade in Fahrzeugen können durch mechanische Einflüsse oder große Temperaturschwankungen relevante Änderungen der Kameraparameter auftreten, die eine kontinuierliche Nachführung der Kalibrierung erforderlich machen. Ein weiteres Anwendungsfeld ist das an biologischen Systemen orientierte aktive Sehen (Bild 1), bei dem eine Selbstkalibrierung aufgrund der ständigen Bewegung der Kameras unverzichtbar ist.

Dieser Beitrag gibt einen kurzen Überblick über das in [1] entwickelte Rahmenwerk zur kontinuierlichen Stereoselbstkalibrierung. Basierend auf einer Analyse der Einflüsse von Fehlern in der Kamerakalibrierung auf die 3D-Rekonstruktion werden wichtige Parameter iden-

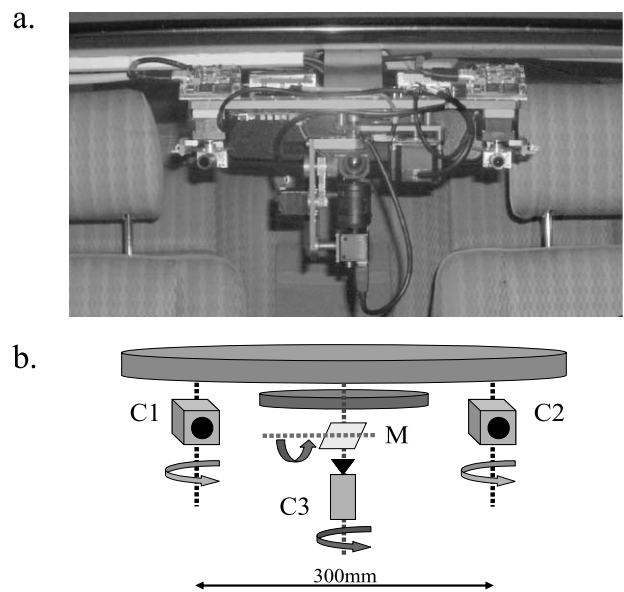


Bild 1: Aktive Kameraplattform des Instituts für Mess- und Regelungstechnik (MRT). Die Kameraplattform besteht aus zwei Stereokameras (C1 und C2), die unabhängig voneinander um ihre Hochachse gedreht werden können, und einer Telemekamera (C3), bei der über einen Spiegel (M) die vertikale und horizontale Blickrichtung verändert werden kann.

Figure 1: Active camera platform of the Institut für Mess- und Regelungstechnik (MRT). The platform consists of two stereo cameras (C1 and C2) which can be rotated independently about their yaw axes, and a tele camera capable of adjusting both its horizontal and vertical gaze direction via a mirror (M).

tifiziert, welche durch ein Stereokalibrierungsverfahren bestimmt werden müssen (Kap. 2). Verschiedene geometrische Bedingungsgleichungen werden auf ihre Eignung für eine kontinuierliche Selbstkalibrierung hin untersucht und in einem Gauß-Helmert-Modell einheitlich formuliert (Kap. 3). In [1] werden für die Ausgleichung in solchen Gauß-Helmert-Modellen robuste, rekursive Optimierungsverfahren entwickelt. Experimentell wurden die beschriebenen Verfahren an mehreren Beispielen aus dem Bereich der mobilen Wahrnehmung getestet (Kap. 4). Die Verfahren erlauben sowohl die Kalibrierung eines statischen Stereosystems als auch die Nachführung der Parameter einer aktiven Kameraplattform, welche im Rahmen dieser Arbeit zur Umfeldwahrnehmung in kognitiven Automobilen entwickelt wurde.

2 Kameraparameter und ihr Einfluss auf die Stereorekonstruktion

In dieser Arbeit wird zur Beschreibung von Stereokameras ein ideales Lochkameramodell verwendet. In homogenen Koordinaten kann hier die Abbildung eines 3D-Punktes $\mathbf{X} = [X, Y, Z, 1]^T$ auf einen Bildpunkt $\mathbf{x}_L = [x_L, y_L, 1]^T$ der linken Kamera mit Hilfe der einfachen Beziehung¹

$$\mathbf{x}_L = \lambda_L \mathbf{K}_L \mathbf{R}_L [\mathbf{I}, -\mathbf{C}_L] \mathbf{X} \quad (1)$$

beschrieben werden. λ_L bezeichnet dabei einen unbekannt skalarer Faktor. \mathbf{K}_L ist die Matrix der inneren Orientierung mit der Brennweite f_L und dem Bildhauptpunkt $\mathbf{c}_L = [c_{L,x}, c_{L,y}]^T$:

$$\mathbf{K}_L = \begin{bmatrix} f_L & 0 & c_{x,L} \\ 0 & f_L & c_{y,L} \\ 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

Analoge Beziehungen gelten für die rechte Stereokamera.

Die Rotationsmatrix \mathbf{R}_L und der Vektor \mathbf{C}_L definieren die Orientierung und die Position der linken Stereokamera bezüglich eines globalen Koordinatensystems (*extrinsische Parameter*). In dieser Arbeit gelte für die rechte und linke Kamera

$$\begin{aligned} \mathbf{R}_R &= \mathbf{R}(\Psi_R, \Phi_R = 0, \Theta_R), \quad \mathbf{C}_R = [b/2, 0, 0]^T, \\ \mathbf{R}_L &= \mathbf{R}(\Psi_L, \Phi_L, \Theta_L), \quad \mathbf{C}_L = [-b/2, 0, 0]^T, \end{aligned} \quad (3)$$

wobei b die Basisbreite des Stereosystems und $\Psi_{L/R}$, $\Phi_{L/R}$, $\Theta_{L/R}$ die Gier-, Nick- und Rollwinkel der Kameraorientierungen bezeichnen. Mit Gl. (3) wird festgelegt, dass der Ursprung unseres Weltkoordinatensystems im Mittelpunkt zwischen beiden Kameras liegt und die z -Achse des Koordinatensystems parallel zur Blickrichtung der rechten Kamera verläuft ($\Phi_R = 0$). Gl. (3) ist besonders für aktive Kamerasysteme von Vorteil, bei denen die Orientierung beider Kameras ständig verändert werden kann. Insgesamt werden Stereokameras also durch sechs intrinsische und sechs extrinsische Parameter beschrieben.

Herkömmlicherweise wird zur Vereinfachung der Stereorekonstruktion eine *Rektifizierung* der Kamerabilder vorgenommen. Diese Transformation überführt real aufgenommene Bilder mit Koordinaten $\mathbf{x}_L, \mathbf{x}_R$ in virtuelle Aufnahmen einer idealen Stereoaordnung mit Koordinaten $\tilde{\mathbf{x}}_L, \tilde{\mathbf{x}}_R$. Eine solche ideale Stereokamera besteht aus zwei gedachten identischen Kameras, deren

Bildebenen parallel zueinander ausgerichtet sind. Wichtigste Eigenschaft der Rektifizierung ist, dass korrespondierende Bildpunkte $\tilde{\mathbf{x}}_L$ und $\tilde{\mathbf{x}}_R$ eines 3D-Objektpunktes \mathbf{X} in den resultierenden Bildern stets in der gleichen Bildzeile liegen, d. h.

$$\begin{pmatrix} \tilde{x}_L - \tilde{x}_R \\ \tilde{y}_L - \tilde{y}_R \end{pmatrix} = \begin{pmatrix} \tilde{d} \\ 0 \end{pmatrix}. \quad (4)$$

Wegen $\tilde{y}_L - \tilde{y}_R = 0$ kann die Suche nach korrespondierenden Bildpunkten im linken und rechten Stereobild auf eine Bildzeile beschränkt werden. Dies ist insbesondere für den erforderlichen Rechenaufwand einer Stereorekonstruktion von großer Bedeutung.

Die Distanz \tilde{d} der rektifizierten Koordinaten in Gl. (4) wird als Disparität bezeichnet. Sie ist umgekehrt proportional zur Entfernung Z des betrachteten 3D-Punktes. Im Folgenden werden wir aus Gründen der Lesbarkeit auf die Tilde bei rektifizierten Koordinaten verzichten.

Bei bekannten Kameraparametern können die entsprechenden Transformationsvorschriften für eine Rektifizierung der beiden Bilder leicht bestimmt werden (s. z. B. [8]). Würde eine solche Transformation allerdings aus fehlerhaften Kameraparametern errechnet, so ist Gl. (4) nicht mehr gültig. Stattdessen ergeben sich auf der rechten Seite der Gleichung folgende Fehler:

1. *Disparitätsfehler Δd* : Horizontale Fehler in den Bildkoordinaten resultieren in einer fehlerhaften Stereodisparität $d + \Delta d$ und damit gleichzeitig in einer fehlerbehafteten Entfernungsmessung. Aus der bekannten Stereorekonstruktionsformel $Z = bf/d$ ergibt sich für die Tiefenrekonstruktion eine Abweichung $Z = Z^2/(bf) \cdot \Delta d$ von der wahren Entfernung Z . Bei gegebenen Disparitätsfehlern nimmt also die Ungenauigkeit der Positionsschätzung quadratisch mit der Entfernung zu.
2. *Vertikale Ausrichtungsfehler Δy* : Nahezu alle in der Praxis relevanten Verfahren zur Bestimmung korrespondierender Punkte in einem Stereobildpaar basieren auf der Annahme rektifizierter Bilder und beschränken den Suchraum der Korrespondenzanalyse auf eine Bildzeile. Ein vertikaler Ausrichtungsfehler $\Delta y = y_L - y_R$ kann also dazu führen, dass die Suche nach korrespondierenden Punkten fehlerhafte oder in vielen Fällen keine gültigen Ergebnisse liefern kann. Um ein zuverlässiges und dichtes Tiefenbild zu erhalten, sollte Δy in der Praxis kleiner als ein Pixel sein.

Tabelle 1 fasst die Auswirkungen von Kalibrierfehlern auf Δd , Δy und die Abweichung ΔZ zusammen. Zum einen können mit Hilfe dieser Ergebnisse zulässige Toleranzgrenzen der Genauigkeit zu schätzender Kameraparameter abgeleitet werden. Zum anderen lässt sich an Tabelle 1 beispielsweise ablesen, dass Fehler des Bildhauptpunkts \mathbf{c} zumindest in ers-

¹ Die Indizes L und R kennzeichnen Größen der linken bzw. rechten Kamera. Vektoren und Matrizen werden in diesem Artikel fett gedruckt.

Fehlerquelle	Auswirkung auf Bildkoordinaten (Index „N“ bezeichnet normalisierte Koordinaten)	Disparitätsfehler Δd und vertikaler Ausrichtungsfehler Δy	Lineare Empfindlichkeit
Gierwinkelfehler $\Delta\Psi_L$	$\Delta\mathbf{x}_{L,N} \approx \begin{pmatrix} 1+x_{L,N}^2 \\ x_{L,N}y_{L,N} \end{pmatrix} \Delta\Psi_L$	$\Delta d \approx f(1+x_{L,N}^2)\Delta\Psi_L$ $\Delta y \approx fx_{L,N}y_{L,N}\Delta\Psi_L$	$\frac{\Delta Z}{\Delta\Psi_L} \approx -\frac{Z^2}{b}(1+x_{L,N}^2)$
Nickwinkelfehler $\Delta\Phi_L$	$\Delta\mathbf{x}_{L,N} \approx -\begin{pmatrix} x_{L,N}y_{L,N} \\ 1+y_{L,N}^2 \end{pmatrix} \Delta\Phi_L$	$\Delta d \approx -fx_{L,N}y_{L,N}\Delta\Phi_L$ $\Delta y \approx -f(1+y_{L,N}^2)\Delta\Phi_L$	$\frac{\Delta Z}{\Delta\Phi_L} \approx \frac{Z^2}{b}x_{L,N}y_{L,N}$
Wankwinkelfehler $\Delta\Theta_L$	$\Delta\mathbf{x}_{L,N} \approx \begin{pmatrix} -y_{L,N} \\ x_{L,N} \end{pmatrix} \Delta\Theta_L$	$\Delta d \approx (y_L - c_y)\Delta\Theta_L$ $\Delta y \approx (x_L - c_x)\Delta\Theta_L$	$\frac{\Delta Z}{\Delta\Theta_L} \approx \frac{Z^2}{b}y_{L,N}$
Basisbreitenfehler Δb	$\Delta\mathbf{x}_{L/R,N} \approx \pm \begin{pmatrix} \frac{d_N}{2b} \\ 0 \end{pmatrix} \Delta b$	$\Delta d \approx \frac{\Delta b}{b}d$ $\Delta y \approx 0$	$\frac{\Delta Z}{\Delta b} \approx -\frac{Z}{b}$
Hauptpunktverschiebung $\Delta\mathbf{c}_L$	$\Delta\mathbf{x}_L \approx \Delta\mathbf{c}_L$	$\Delta d \approx \Delta c_{L,x}$ $\Delta y \approx \Delta c_{L,y}$	$\frac{\Delta Z}{\Delta c_{L,x}} \approx -\frac{Z^2}{bf}$
Brennweitenfehler Δf_L (eine Kamera)	$\Delta\mathbf{x}_L \approx (\mathbf{x}_L - \mathbf{c}) \frac{\Delta f_L}{f}$	$\Delta d \approx (x_L - c_x) \frac{\Delta f_L}{f}$ $\Delta y \approx (y_L - c_y) \frac{\Delta f_L}{f}$	$\frac{\Delta Z}{\Delta f_L} \approx -\frac{Z^2}{bf^2}(x_L - c_x)$
Brennweitenfehler Δf (beide Kameras)	$\Delta\mathbf{x}_{L/R} \approx (\mathbf{x}_{L/R} - \mathbf{c}) \frac{\Delta f}{f}$	$\Delta d \approx \frac{\Delta f}{f}d$ $\Delta y \approx 0$	$\frac{\Delta Z}{\Delta f} \approx -\frac{Z}{f}$

Tabelle 1: Auswirkungen von Fehlern der Kamerakalibrierung auf die Stereorekonstruktion. Normalisierte Koordinaten ergeben sich aus $\mathbf{x}_{L,N} = \mathbf{K}_L^{-1}\mathbf{x}_L$ und $\mathbf{x}_{R,N} = \mathbf{K}_R^{-1}\mathbf{x}_R$.

Table 1: Sensitivity of 3D reconstruction to erroneous camera calibration. Normalized coordinates are computed as $\mathbf{x}_{L,N} = \mathbf{K}_L^{-1}\mathbf{x}_L$ and $\mathbf{x}_{R,N} = \mathbf{K}_R^{-1}\mathbf{x}_R$.

ter Näherung durch eine Superposition von Gier- und Nickwinkelfehlern kompensiert werden können. Dies bestätigt die Beobachtung, dass die Kalibrierung des Bildhauptpunktes $\mathbf{c}_{L/R}$ nur geringe Auswirkungen auf das Rekonstruktionsergebnis hat [6]. Aus diesem Grund wird in dieser Arbeit auf die Bestimmung von $\mathbf{c}_{L/R}$ verzichtet. Zielgrößen der Selbstkalibrierung sind damit die Brennweiten f_L , f_R sowie die fünf Winkel der äußeren Orientierung aus Gl. (3). Die Basisbreite b ist ein Skalierungsfaktor, der allein aus Bilddaten ohne zusätzliches Wissen nicht beobachtbar ist.

3 Kontinuierliche Schätzung von Kameraparametern

Bild 2 zeigt die Gesamtstruktur des vorgeschlagenen Algorithmus zur kontinuierlichen Selbstkalibrierung von Stereokameras. An dieser Stelle wird ein Abriss der wesentlichen Komponenten des entwickelten Schätzver-

fahrens gegeben. Ausführlichere Beschreibungen sind in [1] zu finden.

Ausgangspunkt für Selbstkalibrierung sind korrespondierende Punkte in zwei oder mehr Bildern, welche einem gemeinsamen 3D-Objektpunkt entsprechen. Diese Messdaten können beispielsweise mit Hilfe der Scale Invariant Feature Transform (SIFT) [9] in hoher Qualität extrahiert werden. Wir unterscheiden dabei räumliche Korrespondenzen zwischen gleichzeitigen Aufnahmen der linken und rechten Stereokamera sowie zeitliche Korrespondenzen in aufeinander folgenden Bildern einer Kamera. Zwischen diesen extrahierten Punktkorrespondenzen gelten verschiedene, von den Kameraparametern abhängige, geometrische Bedingungsgleichungen:

1. *Epipolarbedingung:* Die Epipolarbedingung (z. B. [10]) oder Komplanaritätsbedingung stellt einen Zusammenhang zwischen korrespondierenden Punkten in zwei Bildern – in unserem Fall dem linken und rechten Stereobild – her. Sie besagt anschaulich, dass Lichtstrahlen, welche durch korrespondierende Bildpunkte \mathbf{x}_L und \mathbf{x}_R sowie die Kamerabrennpunkte \mathbf{C}_L und \mathbf{C}_R definiert werden, stets in derselben Ebene liegen. Mathematisch

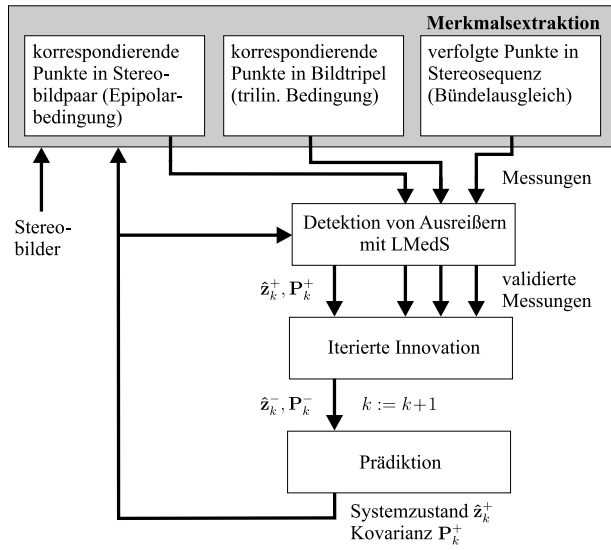


Bild 2: Gesamtstruktur des rekursiven Selbstkalibrierungsverfahrens für Stereokameras.
Figure 2: Structure of the recursive self-calibration algorithm for stereo cameras.

kompakt lässt sich dieser Zusammenhang durch die Fundamentalmatrix \mathbf{F} beschreiben:

$$h_1(\mathbf{x}_L, \mathbf{x}_R, \mathbf{F}) = \mathbf{x}_L^T \mathbf{F} \mathbf{x}_R \stackrel{!}{=} 0, \quad (5)$$

mit

$$\mathbf{F} = \mathbf{K}_L^{-T} \mathbf{R}_L \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -b \\ 0 & b & 0 \end{bmatrix} \mathbf{R}_R^T \mathbf{K}_R^{-1}. \quad (6)$$

2. *Trilineare Bedingungen:* Die Trilinearitäten [11] sind eine Generalisierung der Epipolarbedingung für drei Bilder. Bei der hier vorgestellten Selbstkalibrierung besteht ein Bildtripel aus dem Stereobildpaar zum Zeitpunkt k und dem nächsten Bild der rechten Kamera zum Zeitpunkt $k+1$. Die trilinearen Bedingungen verwenden damit sowohl räumliche ($\mathbf{x}_L(k) \leftrightarrow \mathbf{x}_R(k)$) als auch zeitliche Punktkorrespondenzen ($\mathbf{x}_R(k) \leftrightarrow \mathbf{x}_R(k+1)$) und erfordern deshalb die Kenntnis der Kamerabewegung relativ zu einer betrachteten starren Szene. Die Selbstkalibrierung muss also zusätzlich zu den Kameraparametern drei rotatorische und drei translatorische Bewegungsparameter bestimmen. Die mathematische Beschreibung ist von ähnlicher Form wie bei der Epipolarbedingung,

$$\mathbf{h}_2(\mathbf{x}_L(k), \mathbf{x}_R(k), \mathbf{x}_R(k+1), \mathbf{T}) \stackrel{!}{=} \mathbf{0}, \quad (7)$$

wobei \mathbf{T} den von den Kameraparametern sowie der Kamerabewegung abhängigen Trifokaltensor bezeichnet (s. z. B. [1]).

3. *Bündelausgleich mit reduzierter Strukturrepräsentation:* Grundlage des Bündelausgleichs ist die sogenannte Kollinearitätsgleichung, welche besagt, dass ein 3D-Objektpunkt, sein zugehöriger Bildpunkt und die Brennpunkte der Kamera stets auf einer Geraden liegen. Betrachtet man eine Menge von Objektpunkten $\mathbf{X}_i, i = 1..N$, welche sich relativ zur Kamera starr bewegen, so lässt sich mit Hilfe der Kollinearitätsgleichung ein Zusammenhang zwischen den 3D-Koordinaten \mathbf{X}_i , ihren Bildpunkten in der Stereosequenz, der 3D-Objektbewegung und den Kameraparametern angeben. Nachteilig bei der Formulierung der Kollinearitätsgleichung ist allerdings, dass die 3D-Positionen \mathbf{X}_i explizit bekannt sein müssen. Im Gegensatz zu den beiden vorangehenden Bedingungsgleichungen muss deshalb eine Selbstkalibrierung hier neben den Kameraparametern stets die 3D-Struktur der betrachteten Szene ermitteln. In [1] wird eine kompakte Bedingungsgleichung für die Selbstkalibrierung aufgestellt, welche pro betrachtetem Objektpunkt nur dessen Entfernung benötigt und somit statt $3N$ nur N Parameter zur Repräsentation der 3D-Struktur erfordert: Ausgangspunkt sind wie bei den trilinearen Bedingungen räumliche und zeitliche Punktkorrespondenzen $\mathbf{x}_{i,L}(k) \leftrightarrow \mathbf{x}_{i,R}(k) \leftrightarrow \mathbf{x}_{i,R}(k+1)$. Zusätzlich nehmen wir für die Formulierung der Bedingungsgleichung an, dass die Entfernungen der Objektpunkte im Zeitschritt k durch $\rho_i(k)$ gegeben sind. Aus $\rho_i(k)$ und $\mathbf{x}_{i,R}(k)$ können die Koordinaten des Objektpunktes $\mathbf{X}_i(k)$ eindeutig bestimmt werden,

$$\mathbf{X}_i(k) = \mathbf{\Pi}_R^{-1}(\mathbf{x}_{i,R}(k), \rho_i(k)), \quad (8)$$

wobei $\mathbf{\Pi}_R^{-1}(\cdot)$ die Inversion der Projektionsgleichung (1) für die rechte Kamera bezeichnet. Gehen wir weiter davon aus, dass sich alle betrachteten Objektpunkte relativ zur Kamera starr bewegen und ihre 3D-Bewegung durch eine Rotationsmatrix $\mathbf{R}_M = \mathbf{R}(\Psi_M, \Phi_M, \Theta_M)$ und einen Translationsvektor \mathbf{T}_M vollständig beschrieben werden kann, so lässt sich aus Gln. (1) und (8) folgende Bedingungsgleichung für die extrahierten Merkmale aufstellen:

$$\mathbf{h}_3(\mathbf{x}_{L,i}(k), \mathbf{x}_{R,i}(k), \mathbf{x}_{R,i}(k+1), \mathbf{z}(k)) = \begin{bmatrix} \pi_R\{\mathbf{R}_M \mathbf{\Pi}_R^{-1}(\mathbf{x}_{i,R}(k), \rho_i(k)) + \mathbf{T}_M\} - \mathbf{x}_{i,R}(k+1) \\ \pi_L\{\mathbf{\Pi}_R^{-1}(\mathbf{x}_{i,R}(k), \rho_i(k))\} - \mathbf{x}_{i,L}(k) \end{bmatrix} \stackrel{!}{=} \mathbf{0} \quad (9)$$

$\pi_R(\cdot)$ und $\pi_L(\cdot)$ repräsentieren dabei die Projektion nach Gl. (1) in die linke und rechte Kamera. Der Vektor $\mathbf{z}(k)$ umfasst die Parameter der beiden Kameras, der relativen Objektbewegung sowie der 3D-Struktur.

Die geometrischen Beziehungen (6), (7) und (9) sind also durch formal ähnliche implizite Gleichungen zwischen den idealen Beobachtungen und den wahren Parametern gegeben. Aufgabe der Selbstkalibrierung ist nun, aus gegebenen fehlerbehafteten Messungen optimale Kameraparameter zu bestimmen. Dies geschieht über eine Ausgleichung mit Nebenbedingungen nach dem Gauß–Helmert-Modell: Das zu minimierende quadratische Gütekriterium lautet

$$J = \|\hat{\mathbf{z}}_0 - \mathbf{z}_0\|_{\Sigma_{zz,0}}^2 + \sum_{k=1}^K \|\mathbf{x}_k - \hat{\mathbf{x}}_k\|_{\Sigma_{ee,k}}^2 + \|\hat{\mathbf{z}}_k - \mathbf{f}(\hat{\mathbf{z}}_{k-1})\|_{\Sigma_{zz,k}}^2, \quad (10)$$

wobei die ausgeglichenen Beobachtungen und der gesuchte Schätzwert des Parametervektors die Nebenbedingungen

$$\mathbf{h}(\mathbf{x}_k, \hat{\mathbf{z}}_k) = \mathbf{0} \quad (11)$$

erfüllen müssen. In Gln. (10)–(11) bezeichnen:

- \mathbf{K} die Anzahl der verwendeten Stereobildpaare,
- $\hat{\mathbf{x}}_k$ bzw. \mathbf{x}_k die Vektoren aus den gemessenen, fehlerbehafteten Punktkorrespondenzen bzw. den ausgeglichenen (d. h. korrigierten) Beobachtungen im Zeitschritt k ,
- $\hat{\mathbf{z}}_k$ die beste Schätzung der Kamera-, Bewegungs- und 3D-Strukturparameter zum Zeitpunkt k ,
- $\mathbf{f}(\cdot)$ die zeitliche Dynamik des gesuchten Parametervektors,
- $\mathbf{h}(\cdot)$ das aus verschiedenen Bedingungsgleichungen h_1, h_2 und h_3 zusammengesetzte implizite Beobachtungsmodell und
- $\|\mathbf{a} - \mathbf{b}\|_{\Sigma}^2$ die mit der Kovarianzmatrix Σ normierte Mahalanobisdistanz zwischen zwei Vektoren \mathbf{a} und \mathbf{b} .

Mit Hilfe der in Gln. (10)–(11) vorgestellten Ausgleichung nach dem Gauß–Helmert-Modell können beliebige Kombinationen der oben angeführten geometrischen Bedingungsgleichungen zur stereoskopischen Selbstkalibrierung verwendet werden. Wichtig ist, dass mit dem Gütekriterium J gemäß Gl. (10) ein Fehler in der Bildebene und damit ein geometrisches Fehlermaß minimiert wird. Dadurch ergibt sich zwar eine Optimierungsaufgabe mit nichtlinearen Nebenbedingungen. Es kann aber eine höhere Genauigkeit als bei einer einfacheren Ausgleichung mit algebraischem Fehlerkriterium erreicht werden.

In [1] wird ein rekursives Verfahren zur Lösung des Optimierungsproblems nach Gln. (10)–(11) auf Basis eines Iterativen Erweiterten Kalman Filters (IEKF) vorgeschlagen (Tabelle 2). Durch wiederholte Korrektur der Systemparameter und der erfassten Messdaten wird eine verbesserte Linearisierung der Nebenbedingungen und damit eine höhere Genauigkeit erreicht. Mit die-

Systemmodell: $\mathbf{z}_k = \mathbf{f}(\mathbf{z}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1})$, $\mathbf{w}_k \sim N(\mathbf{0}, \Sigma_{ww})$.

Beobachtungsmodell:

$\hat{\mathbf{x}}_k = \mathbf{x}_k + \mathbf{e}_k$, $\mathbf{h}(\mathbf{x}_k, \mathbf{z}_k) = \mathbf{0}$, $\mathbf{e}_k \sim N(\mathbf{0}, \Sigma_{ee})$.

Initialzustand: $\hat{\mathbf{z}}_0 = \mathbf{z}_0$, $\mathbf{P}_0^- = \mathbf{P}_0$, $k = 0$.

1. Innovation

$\check{\mathbf{x}}_{k,0} = \hat{\mathbf{x}}_k$, $\check{\mathbf{z}}_{k,0} = \hat{\mathbf{z}}_k^+$.

For $l = 0 \dots L - 1$:

$\mathbf{A}_l = \partial \mathbf{h} / \partial \mathbf{z} |_{\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}}$, $\mathbf{B}_l = \partial \mathbf{h} / \partial \mathbf{x} |_{\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}}$.

$\mathbf{K}_{k,l} = \mathbf{P}_k^- \mathbf{A}_l^T [\mathbf{A}_l \mathbf{P}_k^- \mathbf{A}_l^T + \mathbf{B}_l \Sigma_{ee} \mathbf{B}_l^T]^{-1}$.

$\mathbf{r}_{k,l} = \mathbf{h}(\check{\mathbf{x}}_{k,l}, \check{\mathbf{z}}_{k,l}) + \mathbf{B}_l \cdot (\hat{\mathbf{x}} - \check{\mathbf{x}}_{k,l}) + \mathbf{A}_l \cdot (\hat{\mathbf{z}}_k^- - \check{\mathbf{z}}_{k,l})$.

$\check{\mathbf{z}}_{k,l+1} = \hat{\mathbf{z}}_k^- - \mathbf{K}_{k,l} \mathbf{r}_{k,l}$.

$\check{\mathbf{x}}_{k,l+1} = \hat{\mathbf{x}}_k - \Sigma_{ee} \mathbf{B}_l^T [\mathbf{B}_l \Sigma_{ee} \mathbf{B}_l^T]^{-1} \mathbf{r}_{k,l}$.

$\hat{\mathbf{z}}_k^+ = \check{\mathbf{z}}_{k,L}$.

$\mathbf{P}_k^+ = [\mathbf{I} - \mathbf{K}_{k,L-1} \mathbf{A}_{L-1}] \mathbf{P}_k^-$.

2. Prädiktion:

$\mathbf{F}_z = \partial \mathbf{f} / \partial \mathbf{z} |_{\hat{\mathbf{z}}_k^+, \mathbf{u}_k, \mathbf{0}}$, $\mathbf{F}_w = \partial \mathbf{f} / \partial \mathbf{w} |_{\hat{\mathbf{z}}_k^+, \mathbf{u}_k, \mathbf{0}}$.

$\hat{\mathbf{z}}_{k+1}^- = \mathbf{f}(\hat{\mathbf{z}}_k^+, \mathbf{u}_k, \mathbf{0})$.

$\mathbf{P}_{k+1}^- = \mathbf{F}_z \mathbf{P}_k^+ \mathbf{F}_z^T + \mathbf{F}_w \Sigma_{ww} \mathbf{F}_w^T$.

3. Setze $k := k + 1$. Gehe zu 1.

Tabelle 2: Iteratives Schätzverfahren für Gauß–Helmert-Modelle.

Table 2: Iterative estimation algorithm for Gauss–Helmert models.

sem Ansatz kann eine kontinuierliche Selbstkalibrierung zur Nachführung und Verbesserung von Kameraparametern realisiert werden. In der Praxis ist zu beachten, dass eine Selbstkalibrierung robust sein muss gegenüber häufig auftretenden Modellverletzungen wie z. B. vereinzelt großen Fehlern in den extrahierten Punktkorrespondenzen (häufig bei periodischen Strukturen im Bild). Dies wird in [1] durch die Kombination des IEKF mit einem *Least-median-of-squares*-Schätzer erreicht.

Vergleicht man die Eignung der verschiedenen geometrischen Bedingungsgleichungen für die Selbstkalibrierung, so zeigt die Simulation, dass die vorgestellte Variante des Bündelausgleichs die weitaus höchste Genauigkeit liefert. Nachteilig beim Bündelausgleich sind aber der höhere Rechenaufwand und die in der Praxis häufig verletzte Annahme einer starren Szene. Die Epipolarbedingung ist auf eine solche Annahme nicht angewiesen. Aus diesem Grund ist es sinnvoll, eine Kombination von Bündelausgleich und Epipolarbedingung für verschiedene Messungen zu verwenden, um damit in der praktischen Anwendung eine zuverlässigere Selbstkalibrierung zu ermöglichen.

4 Beispiele

Das im vorangehenden Abschnitt beschriebene Verfahren zur Selbstkalibrierung wurde sowohl mit synthe-

tischen Bildern in der Simulation als auch mit realen Bildsequenzen in verschiedenen Umgebungen getestet. Die Tests wurden mit einem aktiven Kamerasystem wie auch mit starren Stereokameras durchgeführt und lieferten gute Ergebnisse. Messungen ergaben, dass die Selbstkalibrierung unter realen Bedingungen einen mittleren relativen 3D-Rekonstruktionsfehler von ca. 50% (vor Beginn der Kamerakalibrierung) in wenigen Sekunden auf unter 5% reduzieren kann.

Exemplarisch wird hier ein Beispiel mit dem aktiven Kamerasystem im Versuchsträger des MRT (Bild 1) vorgestellt. Wie in Kap. 3 erläutert, wurde für die Stereoselbstkalibrierung eine Kombination von Epipolarbedingungen und rekursivem Bündelausgleich verwendet. Bild 3 zeigt beispielhaft die für die Selbstkalibrierung verwendeten, automatisch extrahierten räumlichen und zeitlichen Punktkorrespondenzen. Die Anfangsschätzung, mit der die kontinuierliche Selbstkalibrierung initialisiert wurde, ging von parallel ausgerichteten Kameras und von Brennweiten aus, welche um 10% von den tatsächlichen Daten abwichen. Wie in Bild 4 zu sehen ist, war mit diesen Startwerten keine sinnvolle Stereorekonstruktion möglich (Bild 0). Bereits nach fünf Bildern konnte ein dichtes Disparitätsbild gewonnen



Bild 3: Zweites und drittes Stereobild der Beispielsequenz. Ebenfalls dargestellt sind die automatisch extrahierten Korrespondenzmerkmale (+: erfolgreich getrackte Merkmale, ◇: Stereokorrespondenzen zur alleinigen Verwendung in der Epipolarbedingung, *: ungültige Korrespondenzmerkmale). Die prädizierten Koordinaten der getrackten Punkte werden durch ihre Fehlerkovarianzellipsen veranschaulicht.

Figure 3: Second and third stereo frame of sample sequence. The automatically selected features are also shown: +: successfully tracked features, ◇: stereo features for epipolar constraint, *: invalid tracking features. Predicted positions of the tracked features are marked by their covariance ellipses.

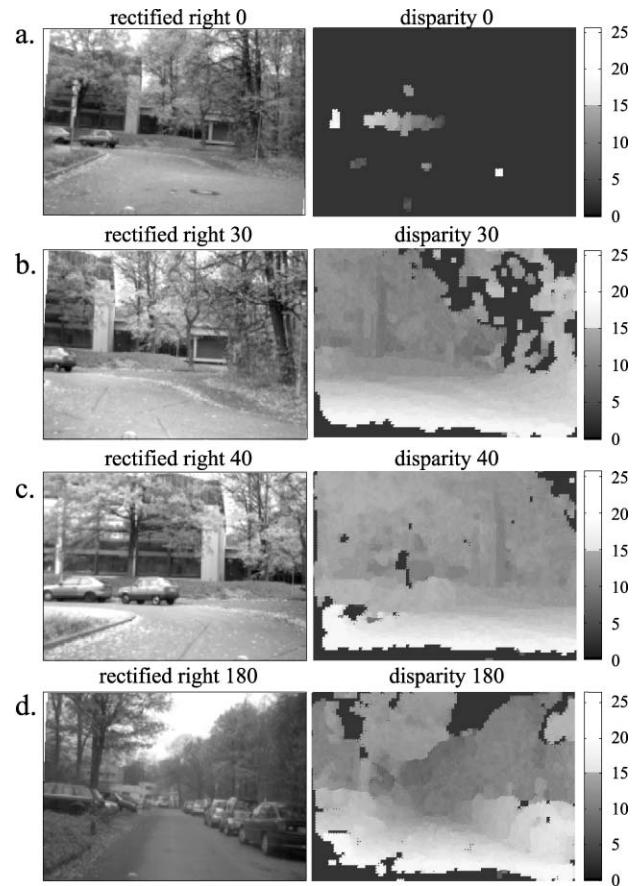


Bild 4: Rekonstruktion nach Selbstkalibrierung mit aktivem Kamerasystem. Gezeigt werden das (mit den Ergebnissen der Selbstkalibrierung rektifizierte) rechte Kamerabild sowie das farblich kodierte Tiefenbild für verschiedene Bilder.

Figure 4: Reconstruction after self-calibration of the active camera system. Shown are the rectified image of the right camera (obtained from the calibration results) and the color-coded depth image for several frames.

werden; nach ca. 40 Bildern erreichte das Filter einen eingeschwungenen Zustand. Bemerkenswert ist, dass die Selbstkalibrierung auch nach zwei gesteuerten Drehungen beider Kameras (vor Beginn der Linkskurve in Bild 38 um 15° nach links und in Bild 168 um 15° nach rechts) zuverlässige Ergebnisse liefert.

5 Zusammenfassung

In [1] wurde ein leistungsfähiges Verfahren zur kontinuierlichen Selbstkalibrierung von Stereokameras vorgestellt. Basierend auf einer Analyse der Empfindlichkeit der 3D-Rekonstruktion bezüglich Kalibrierfehlern

können die Bedeutung der einzelnen Kameraparameter für die Selbstkalibrierung bewertet und zulässige Toleranzgrenzen abgeleitet werden. Es wurde ein rekursives Schätzverfahren für Gauß-Helmert-Modelle entwickelt, mit Hilfe dessen eine Stereoselbstkalibrierung mit hoher Genauigkeit realisiert werden kann. Das Gauß-Helmert-Modell erlaubt die einheitliche Formulierung verschiedener geometrischer Bedingungsgleichungen – die Epipolarbedingung für korrespondierende Punkte in Stereobildpaaren, die trilinearen Bedingungen zwischen Bildtripeln und der Bündelausgleich mit reduzierter Strukturrepräsentation für Bildsequenzen. Wichtig für die Genauigkeit der Parameterschätzung ist dabei vor allem die Verwendung eines geometrischen, physikalisch relevanten Fehlermaßes. Robustheit gegenüber Modellverletzungen wird durch ein *Least-median-of-squares*-Optimierungskriterium erreicht.

Ein Vergleich der verschiedenen Bedingungsgleichungen für die Selbstkalibrierung zeigt, dass gerade die Epipolarbedingung und der Bündelausgleich in der Praxis komplementäre Eigenschaften haben und eine Kombination dieser Bedingungen in der Anwendung zu sehr guten Ergebnissen führt. Dies konnte in [1] an verschiedenen Beispielen demonstriert werden.

Die Autoren glauben, dass mit dem vorgestellten Ansatz eine kontinuierliche Nachführung wichtiger Kameraparameter möglich ist. Dies könnte den erforderlichen Wartungsaufwand für einen Stereosensor im praktischen Einsatz deutlich reduzieren. Eine wichtige offene Fragestellung, die in dieser Arbeit noch nicht untersucht wurde, ist die automatische Bestimmung von Parametern der Linsenverzeichnung. Diese hat vor allem bei günstigen und weitwinkligen Objektiven einen nicht zu vernachlässigenden Einfluss auf die Stereorekonstruktion.

Danksagung

Teile der berichteten Arbeiten werden durch die Deutsche Forschungsgemeinschaft, DFG, unterstützt. Unser Dank gilt der DFG, sowie den beteiligten Partnern des Sonderforschungsbereichs Transregio 28 Kognitive Automobile für die fruchtbare Zusammenarbeit.

Literatur

- [1] Dang, T.: *Kontinuierliche Selbstkalibrierung von Stereokameras*, Dissertation, Universitätsverlag Karlsruhe, 2007.
- [2] Tsai, R. Y.: A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. In *IEEE Journal of Robotics and Automation RA-3*, Nr. 4, 1987, S. 323–344.
- [3] Zhang, Z.: A flexible new technique for camera calibration. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11), 2000, S. 1330–1334.
- [4] Pollefeys, M.: *Self-Calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences*, Dissertation, Katholieke Universiteit Leuven, Belgien, 1999.
- [5] Horaud, R.; Csurka, G.; Demirdjian, D.: Stereo Calibration from Rigid Motions. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 2000, S. 1446–1452.
- [6] Zhang, Z.; Luong, Q.-T.; Faugeras, O.: Motion of an Uncalibrated Stereo Rig: Self-Calibration and Metric Reconstruction. In *IEEE Transactions on Robotics and Automation*, Nr. 12, 1996, S. 103–113.
- [7] Zisserman, A.; Beardsley, P.; Reid, I.: Metric calibration of a stereo rig. In *IEEE Workshop on Representations of Visual Scenes*. Boston, 1995, S. 93–100.
- [8] Fusiello, A.; Trucco, E.; Verri, A.: A Compact Algorithm for Rectification of Stereo Pairs. In *Machine Vision and Applications*, 12(1), 2000, S. 16–22.
- [9] Lowe, D. G.: Distinctive image features from scale-invariant keypoints. In *International Journal of Computer Vision*, 60(2), 2004, S. 91–110
- [10] Hartley, R. I.; Zisserman, A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2002.
- [11] Shashua, A.: Algebraic Functions for Recognition. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8), 1995, S. 779–789.

Dr.-Ing. Thao Dang promovierte am Institut für Mess- und Regelungstechnik der Universität Karlsruhe und arbeitet derzeit bei der Daimler AG in Böblingen im Bereich Aktive Sicherheitssysteme.

Forschungsschwerpunkte: Stereosehen, Selbstkalibrierung, Sensorfusion, Situationsanalyse für Fahrerassistenzsysteme.

Adresse: Institut für Mess- und Regelungstechnik, Engler-Bunte-Ring 21, 76131 Karlsruhe, E-Mail: thao.dang@daimler.com

Prof. Dr.-Ing. Christoph Stiller ist Ordinarius und Leiter des Instituts für Mess- und Regelungstechnik an der Fakultät Maschinenbau der Universität Karlsruhe.

Forschungsschwerpunkte: Maschinelles Sehen, Kognitive technische Systeme, Fahrerassistenzsysteme, Bahnortungssysteme, Sichtprüfungssysteme.

Adresse: Institut für Mess- und Regelungstechnik, Engler-Bunte-Ring 21, 76131 Karlsruhe, E-Mail: stiller@mrt.uka.de