

# **Kalibrierung mobiler Multikamerasysteme mit disjunkten Sichtfeldern**

DISSERTATION

KIT - KARLSRUHER INSTITUT FÜR TECHNOLOGIE  
FRAUNHOFER IOSB - FRAUNHOFER-INSTITUT FÜR Optronik,  
SYSTEMTECHNIK UND BILDAUSWERTUNG

**Frank Pagel**





# **Kalibrierung mobiler Multikamerasysteme mit disjunkten Sichtfeldern**

zur Erlangung des akademischen Grades eines  
Doktors der Ingenieurwissenschaften /  
Doktors der Naturwissenschaften

der Fakultät für Informatik  
des Karlsruher Instituts für Technologie (KIT)

**genehmigte**

## **Dissertation**

von

**Frank Pagel**

aus Bruchsal

Tag der mündlichen Prüfung: 8.11.2012

Hauptreferent: Prof. Dr.-Ing. Jürgen Beyerer

Korreferent: Prof. Dr.-Ing. Christoph Stiller



---

## **Kurzfassung**

---

Moderne Fahrzeuge, sowohl im zivilen als auch im militärischen Bereich, verfügen über immer komplexer werdende Sensorsysteme zur Erfassung der Umwelt. Kamerasysteme spielen hierbei eine wichtige Rolle, da sie vergleichsweise preiswert sind und einen hohen Informationsgehalt bieten. Die Kenntnis der Kameraanordnung, der d. h. extrinsischen Kalibrierparameter, liefert wertvolle räumliche Informationen. Z. B. um sich anhand der Bilddaten besser in der Szene zu orientieren, oder zur Bestimmung, wo sich bestimmte Objekte relativ zum Fahrzeug befinden. Bedingt durch Anforderungen an Design oder Energiemanagement kann es vorkommen, dass die Sichtfelder eines Multikamerasystems keine Schnittmenge aufweisen.

Die vorliegende Arbeit beschreibt und untersucht ein bildgestütztes Verfahren zur extrinsischen Kalibrierung eines Multikamerasystems mit nichtüberlappenden Sichten auf einer mobilen Plattform. Die Schwierigkeit besteht darin, dass durch die nichtüberlappenden Sichten zum Aufnahmepunkt keine korrespondierenden Referenzpunkte zwischen den einzelnen Kameras verwendet werden können, wie dies üblicherweise bei Mehrkamerasystemen mit überlappenden Sichtbereichen der Fall ist. Stattdessen wird die wohlbekannte sog. Hand-Augen-Kalibrierung für die gegebene Problemstellung adaptiert, einer Methodik aus dem Bereich der Robotik. Grundlage hierfür ist die Extraktion von Bildmerkmalen und deren zeitliche Verfolgung in den Bildfolgen jeder einzelnen Kamera. Damit können mit einem Bündelausgleichsverfahren die Kamerabe-

wegungen berechnet werden. Aus den geschätzten Bewegungen der Kameras lassen sich unter bestimmten Voraussetzungen Bedingungsgleichungen formulieren, die auf die Anordnung der Kameras rückschließen lassen.

Die Neuartigkeit des hier vorgestellten, bewegungsbasierten Ansatzes beruht zum einen auf der kontinuierlichen Schätzung der Kalibrierparameter. Das implementierte Verfahren kann daher als eine Art Selbstkalibrierung verstanden werden, die die extrinsischen Parameter fortlaufend, d. h. schritthal tend schätzt. Zu diesem Zweck erfolgt die Schätzung der Kalibrierparameter mittels eines iterativen erweiterten Kalmanfilters.

Des weiteren wird ein modulares und somit skalierbares Fusionskonzept vorgestellt, das eine parallelisierte Prozessierung der einzelnen Bilddatenströme erlaubt. Jede Kamera agiert als eigenständiges Rechenmodul, das lokal, d. h. auf das jeweilige Modul bezogen eine Parameterschätzung durchführen kann. Mittels Propagierung der Parameter und deren Unsicherheiten entstehen redundante Schätzergebnisse, die sich durch die Fusion zusammenführen lassen. Dies hat beispielsweise den Vorteil, dass einzelne Module mit schlecht verwertbaren Sensorinformationen nur gering gewichtet in die Fusion eingehen und somit durch sicherere Informationen anderer Module kompensiert werden können.

Weitere Aspekte, die in der vorliegenden Arbeit gezielt berücksichtigt wurden, sind die Eigenschaften planarer Bewegungen, wie sie meist bei Straßenfahrzeugen auftreten und die Handhabung von Bewegungstransformationen, deren absoluter Verschiebungsbetrag nicht bekannt ist, d. h. lediglich deren Verschiebungsrichtung. Dieser Umstand ist insbesondere bei monokularen Bewegungsschätzungen relevant. Für letzteres wurde die Methode des Bewegungsausgleichs entwickelt und implementiert.

Im Zuge dieser Arbeit wurde die Funktionsfähigkeit des vorgestellten Kalibrierverfahrens anhand einer prototypischen Implementierung und simulierten sowie reellen Daten untersucht. Eine rein kamerabasierte Kalibrierung eines Mehrkamerasystems

---

mit nichtüberlappenden Sichten ist demnach mit dem vorgestellten Ansatz prinzipiell durchführbar. Die geschätzten Parameter wurden mit einer Referenzkalibrierung verglichen und evaluiert. Zusätzlich wurde eine Applikation realisiert, die mit Hilfe der Kalibrierparameter die Kamerasichten in eine gemeinsame Vogelperspektive transformiert. Diese Applikation erlaubt zusätzlich eine qualitative Analyse der Ergebnisse. Es konnte gezeigt werden, dass eine Abhängigkeit zwischen der Art der Bewegung der Sensorplattform und der Güte des Kalibrierergebnisses besteht. Die Ergebnisse zeigen, dass weiterer Forschungsbedarf und die Ergänzung des Kamerasytems um weitere sensorische Informationsquellen sinnvoll ist.

*Meiner Familie*

---

# Inhaltsverzeichnis

---

<b>I Einleitung</b>	<b>1</b>
<b>1 Einführung in die Kalibrierung multiokularer, nichtüberlappender Kameras in Fahrzeugen</b>	<b>3</b>
1.1 Die Kalibrierung und ihre Bedeutung in der Bildfolgenauswertung – Zielsetzung und Herausforderungen . . . . .	3
1.2 Stand der Forschung . . . . .	13
1.3 Die Kalibrierung als »Henne-Ei-Problem« . . . . .	18
1.3.1 Prämisse . . . . .	18
1.3.2 Lösungsansatz und Beitrag der Arbeit . . . . .	19
1.4 Aufbau der Arbeit . . . . .	27
<b>2 Allgemeine Grundlagen</b>	<b>29</b>
2.1 Das Kameramodell . . . . .	29
2.2 Das Bewegungsmodell und die nichtholome Bewegungsbedingung . . . . .	33
2.3 Nichtlineare Levenberg-Marquardt-Optimierung . . . . .	36
2.4 Das iterative erweiterte Kalmanfilter . . . . .	37
2.5 Zustands- und Fehlerpropagierung . . . . .	40
2.6 Gauß'sche Fusion . . . . .	43

<b>3 Visuelle Odometrie</b>	<b>47</b>
3.1 Extraktion und Verfolgung von Bildpunkten . . . . .	47
3.1.1 Ausreißereliminierung . . . . .	50
3.2 Der Bündelausgleich . . . . .	51
<b>4 Das allgemeine Kalibriermodell</b>	<b>55</b>
4.1 Die Hand-Auge-Kalibrierung . . . . .	55
4.2 Generelle Einschränkungen bewegungsbasierter Kalibrierungen . . . . .	58
4.3 Zur Verwendung skalierter Translationsbeträge der Kamerabewegung . .	61
4.4 Notwendigkeit und Einführung einer Historie . . . . .	67
<b>II Bewegungsbasierte, kontinuierliche extrinsische Kalibrierung</b>	<b>69</b>
<b>5 Das erweiterte Kalibriermodell für Straßenfahrzeuge</b>	<b>71</b>
5.1 Erweiterte Annahmen . . . . .	73
5.2 Fallspezifische Auswertung der Bewegung . . . . .	74
5.2.1 Schätzung der Kameraorientierungen bei nichtrotatorischer Be- wegung . . . . .	75
5.3 Einbeziehung der Bodenebene . . . . .	76
5.3.1 Schätzung der Rotationsachse . . . . .	77
5.3.2 Selektion von Punkten der Bodenebene . . . . .	78
5.3.3 Parameterbestimmung mittels nichtlinearer Optimierung . . . . .	79
5.3.4 Anmerkungen zur Abhängigkeit des Höhenparameters von der Geschwindigkeit . . . . .	80
<b>6 Der Bewegungsausgleich</b>	<b>85</b>
6.1 Effiziente Berechnung des dünnen Bewegungsausgleichs . . . . .	91
6.2 SMA relativ zu einem zentralen Bezugspunkt . . . . .	94
6.3 Anmerkungen zur Schätzbarkeit der Kalibrierparameter . . . . .	97

<b>7 Kontinuierliche Zustandsfilterung</b>	<b>107</b>
7.1 Bestimmung der Kovarianzmatrizen der Teiltransformationen . . . . .	109
7.2 Zusammenführen der Teilergebnisse und Filterung der extrinsischen Parameter . . . . .	110
<b>8 Multimodulare Fusion</b>	<b>115</b>
8.1 Globales Schema zur Bewegungsschätzung . . . . .	117
8.1.1 Systeminitialisierung durch initiale Bewegungsschätzung . . . . .	118
8.1.2 Lokale Bewegungsschätzung . . . . .	118
8.1.3 Propagierung und Fusion der Bewegungsparameter . . . . .	120
8.1.4 Ergebnisse zur fusionierten Bewegungsschätzung . . . . .	123
8.1.5 Prüfen der Bewegungskonsistenz . . . . .	123
8.2 Globales Kalibrierschema . . . . .	129
8.2.1 Lokale Kalibrierung . . . . .	129
8.2.2 Propagierung und Fusion der Kalibrierparameter . . . . .	133
8.2.3 Zusammenfassung: Bewegungsphasen und Systementwurf . . . . .	135
<b>III Experimentelle Ergebnisse und Bewertung</b>	<b>141</b>
<b>9 Experimentelle Ergebnisse</b>	<b>143</b>
9.1 Ergebnisse mit simulierten Daten . . . . .	143
9.2 Versuche mit realen Messdaten . . . . .	153
9.2.1 Referenzkalibrierung . . . . .	153
9.2.2 Ergebnisse . . . . .	155
<b>10 Abschließende Bemerkungen</b>	<b>169</b>
10.1 Zusammenfassung . . . . .	169
10.2 Ausblick . . . . .	171
<b>A Anhang</b>	<b>175</b>
A.1 Darstellung von Rotationen im dreidimensionalen Raum . . . . .	175

A.2 Beweis der Kommutativität der paarweisen Gauß'schen Fusion . . . . .	179
A.3 Die Epipolargeometrie . . . . .	180
A.4 Initialisierung bei rein translatorischer Bewegung . . . . .	182
A.5 Überführung der Bewegungstransformationen in die Grundebene . . . . .	185
A.6 Berechnung der Vogelperspektive . . . . .	187
A.7 Zur musterbasierten extrinsischen Kalibrierung . . . . .	189
<b>Literaturverzeichnis</b>	<b>191</b>
<b>Bisherige Veröffentlichungen des Autors</b>	<b>205</b>

---

## Symbolverzeichnis

---

### Abkürzungen

BA, SBA	Bündelausgleich bzw. dünner Bündelausgleich (engl. Sparse Bundle Adjustment)
EKF, IEKF	Erweitertes bzw. iteratives erweitertes Kalmanfilter
GF	Gaußsche Fusion
GFP	Gaußsche Fehlerpropagierung
LGS	Lineares Gleichungssystem
LM	Levenberg-Marquardt (-Optimierung)
MMD	Multiokulare Kamerasysteme auf mobiler Plattform mit disjunkten Sichtfeldern
RANSAC	Random Sample Consensus
SFM	Structure from Motion
SLAM	Simultaneous Localization and Mapping
SMA	Dünner Bewegungsausgleich (engl. <i>Sparse Motion Adjustment</i> )
TCP	Tool Center Point
UAV	Unbemannte Luftfahrzeuge (engl. <i>Unmanned Aerial Vehicle</i> )

### Indizes

$\xi$	<i>Random Sampling-Index</i>
-------	------------------------------

$i, j, h$	Modulindizes
$k$	Globaler Zeitindex der Zustandsfilterung
$L$	Anzahl der Iterationen beim IEKF
$l$	Iterationszähler beim IEKF
$m, n$	Allgemeine Indizes
$p$	Bild- und Szenenpunktindex
$r$	Index innerhalb eines Bewegungshistorienblocks
$t$	Index innerhalb der Bewegungshistorie

## Symbole

$\alpha, \gamma$	Eulerwinkel der Transformation in die Grundebene
$\boldsymbol{\alpha}_R$	3D-Rotationsachse $[\alpha_x, \alpha_y, \alpha_z]^T$
$(\mathbf{a} \times \mathbf{b})$	Kreuzprodukt der Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$
$(\mathbf{q}_1 \star \mathbf{q}_2)$	Multiplikation zweier Quaternionen
$\varepsilon_C$	Kalibrierfehler
$\Sigma$	Allgemeine Kovarianz- bzw. Rauschmatrix
$\bar{\mathbf{q}}$	Konjugiertes Quaternion zu $\mathbf{q}$
$C$	Zentraler Bezugspunkt der Fahrzeugbewegung beim Einspurmodell
$\text{Diag}\left(a_1, \dots, a_n\right)$	Diagonalmatrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ mit den Diagonaleinträgen $a_1, \dots, a_n$
$\text{diag}(\mathbf{M})$	Vektor $\mathbf{d} \in \mathbb{R}^N$ der Diagonaleinträge von $\mathbf{M} \in \mathbb{R}^{N \times N}$
$\mathbf{g}$	Ebene $\mathbf{g} = (g_1, g_2, g_3, g_4)^T$
$\underline{C}$	Homogene, Euklidische, extrinsische $\in \mathbb{R}^{4 \times 4}$ Transformationsmatrix in der Bodenebene (2D)
$\mathbf{E}, \mathbf{F}$	Essentielle und Fundamentalmatrix $\in \mathbb{R}^{3 \times 3}$
$\varepsilon, \boldsymbol{\varepsilon}$	Restfehler, Residuenvektor
$C$	Homogene, Euklidische, extrinsische $\in \mathbb{R}^{4 \times 4}$ Transformationsmatrix
$f$	Brennweite

---

$\Gamma$	Größe eines Bewegungshistorienblocks
$\boldsymbol{l}$	Gerade $\boldsymbol{l} = (l_1, l_2, l_3)^T$
$\mathbf{G}$	Homogene, Euklidische $\in \mathbb{R}^{4 \times 4}$ Transformationsmatrix zur Transformation einer Kamera in die Bodenebene
$\mathbf{h}(\mathbf{x}, \mathbf{z}), \mathbf{h}(\mathbf{x}, \mathbf{z}; \mathbf{K})$	Allgemeine Bedingungsgleichung in Abhängigkeit des Parametervektors $\mathbf{x}$ und Messung $\mathbf{z}$
$\mathbf{h}_{\mathbf{C}}$	Bedingungsgleichung zur Filterung der extrinsischen Parameter
$\mathbf{h}_{\mathbf{G}}$	Bedingungsgleichung zur Bestimmung der Bodenebene
$\mathbf{h}_{\mathbf{M}}$	Bedingungsgleichung des Bündelausgleichs
$\mathbf{h}_{SMA}$	Bedingungsgleichung des Bewegungsausgleichs
$\hat{\mathbf{x}}, \hat{\mathbf{z}}$	Schätzwerte
$\mathbf{H}$	Homographie $\in \mathbb{R}^{3 \times 3}$
$\mathbf{H}_{\mathbf{x}}, \mathbf{H}_{\mathbf{z}}$	Jacobimatrizen zur Bedingungsgleichung $\mathbf{h}(\mathbf{x}, \mathbf{z})$ mit $\mathbf{H}_{\mathbf{x}} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big _{\mathbf{x}, \mathbf{z}} \in \mathbb{R}^{\dim(\mathbf{h}) \times \dim(\mathbf{x})}$ und $\mathbf{H}_{\mathbf{z}} = \frac{\partial \mathbf{h}}{\partial \mathbf{z}} \Big _{\mathbf{x}, \mathbf{z}} \in \mathbb{R}^{\dim(\mathbf{h}) \times \dim(\mathbf{z})}$
$\mathbf{I}$	Einheitsmatrix $\text{Diag}(1, \dots, 1)$
$\mathbf{C}$	Bezeichner für die Extrinsischen Kalibrierparameter
$\underline{\mathbf{C}}$	Bezeichner für die Extrinsischen Parameter in der Bodenebene
$\mathbf{G}$	Bezeichner für die Bodenebene (engl. <i>Ground</i> )
$\mathcal{I}_k$	Bild zum Zeitpunkt $k$
$\mathbf{M}$	Bezeichner für Bewegung (engl. <i>Motion</i> )
$\underline{\mathbf{M}}$	Bezeichner für Bewegung in der Bodenebene
$\mathbf{J}$	Jacobi-Matrix
$\kappa, \boldsymbol{\kappa}$	Skalierungsfaktor bzw. Vektor der Skalierungsfaktoren beim Bewegungsausgleich
$\mathbf{K}$	Intrinsische Kameramatrix $\in \mathbb{R}^{3 \times 3}$
$\lambda$	Abstand entlang des Sichtstrahls zwischen $\mathbf{X}$ und Kamerazentrum
$\varepsilon_M$	Mahanalobis-Distanz

$\mathcal{S}, \underline{\mathcal{S}}$	Allgemeine Punktmenge bzw. Menge der 3D-Punkte der Boden-ebene
$\mathbb{E}$	Erwartungswert
$\mathbf{A}, \mathbf{B}$	Allgemeine Matrizen, $\mathbf{A} = (a_{mn}) \in \mathbb{R}^{M \times N}$
$\mathbf{D}$	Diagonalmatrix, $d_{mn} = 0, m \neq n$
$\mathbf{D}_a$	Affine 2D-Transformationsmatrix
$\mathbf{G}$	Kalmanverstärkungsfaktor beim IEKF
$\mathbf{V}$	Ladungsmatrix
$\text{vec}(\mathbf{M})$	Der vec-Operator stapelt die Spalten von $\mathbf{M}_{m \times n}$ in einem $m \cdot n$ -Vektor
$\det(\mathbf{M})$	Determinante $\rho \in \mathbb{R}$ von $\mathbf{M} \in \mathbb{R}^{N \times N}$
$\dim(\mathbf{M})$	Dimension der quadratischen Matrix $\mathbf{M} \in \mathbb{R}^{N \times N}$
$\mu$	Mittelwert
$\underline{\mathbf{M}}$	Homogene, Euklidische $\in \mathbb{R}^{4 \times 4}$ Transformationsmatrix der 2D-Bewegung einer Kamera in der Bodenebene
$\mathbf{A}^{-1}$	Inverse Matrix von $\mathbf{A}$
$\mathbf{P}_{\mathbf{M}}$	Projektionsmatrix für Bewegungen
$\mathcal{M}$	Kameramodul
$\mathbf{M}$	Homogene, Euklidische $\in \mathbb{R}^{4 \times 4}$ Transformationsmatrix einer 3D-Bewegung
$\mathbf{A}^{-T}$	Transponierte, invertierte Matrix von $\mathbf{A}$
$\mathbf{A}^T$	Transponierte Matrix von $\mathbf{A}$
$\mathfrak{N}$	Anzahl der Module bzw. Kameras
$\mathfrak{H}$	Anzahl der Bewegungsblöcke in einer Bewegungshistorie
$\mathbf{n}$	Normalenvektor
$\ \mathbf{a}\ $	Euklidische Norm $\sqrt{a_1^2 + \dots + a_d^2}$ von $\mathbf{a} = [a_1, \dots, a_d]^T$
$\ \mathbf{A}\ _F$	Frobeniusnorm von $\mathbf{A}$
$\mathfrak{P}$	Anzahl der Bildkorrespondenz- bzw. Szenenpunkte
$\mathfrak{P}_{\mathbf{G}}$	Anzahl der Punkte für die Bodenschätzung

$\mathfrak{P}_{\mathcal{S}}$	Anzahl der Punkte innerhalb einer Untermenge beim <i>Random Sampling</i>
$\mathfrak{M}_{\mathcal{S}}$	Anzahl der Untermengen beim <i>Random Sampling</i>
$v$	Geschwindigkeitsbetrag
$\mathcal{N}(\mu, \Sigma)$	Normalverteilung mit Mittelwert $\mu \in \mathbb{R}^N$ und Kovarianz $\Sigma \in \mathbb{R}^{N \times N}$
$f(\hat{x}, \hat{z})$	Prädiktionsfunktion des Kalmanfilters
$\mathbf{g}_l^0, \mathbf{g}_n^0$	Allgemeine lineare bzw. nichtlineare Funktionen
$r$	Residuenvektor beim IEKF
$\Pi$	Kameramatrix
$\pi(\mathbf{X}; \Pi)$	Projektion $\pi: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ eines 3D Punktes $\mathbf{X}$ auf den Bildpunkt $\mathbf{x}$ in Abh. der Projektionsparameter $\Pi$ , bestehend aus den intrinsischen und extrinsischen Parametern
$\mathbf{o}$	Kamera- bzw. Projektionszentrum
$\mathbf{x}, \mathbf{X}$	2D-Bildpunkt $[x, y]^T$ bzw. 3D-Szenenpunkt $[X, Y, Z]^T$
$i, j, k$	Imaginäre Einheiten eines Quaternions
$\mathbf{q}$	Quaternion $[q_w, q_x, q_y, q_z] = [q_w, v]$ mit Realteil $q_w \in \mathbb{R}$ und Imaginärteil $v \in \mathbb{R}^3$
$\mathbf{Q}_{\mathbf{q}}^>, \mathbf{Q}_{\mathbf{q}}^<$	$4 \times 4$ -Matrix zum Quaternion $\mathbf{q}$ zur links- bzw. rechtsseitigen Quaternionenmultiplikation
$\mathbf{Q}_{\mathbf{v}}^>, \mathbf{Q}_{\mathbf{v}}^<$	$4 \times 4$ -Matrix zum imaginären Quaternion $\mathbf{v}^q$ zur links- bzw. rechtsseitigen Quaternionenmultiplikation
$\rho$	Skalar der erweiterten Normalengleichung
$\mathbf{R}$	Extrinsische $3 \times 3$ -Rotationsmatrix
$\mathbf{r}$	Vektor der extrinsischen Eulerwinkel $[r_x, r_y, r_z]^T$
$\mathbf{R}(r_x, r_y, r_z)$	$3 \times 3$ -Rotationsmatrix für extrinsische Transformationen zwischen Kamerakoordinatensystemen i.Abh. der Eulerwinkel
$\underline{\mathbf{R}}$	Extrinsische $2 \times 2$ -Rotationsmatrix in der Bodenebene
$\underline{\mathbf{R}}(\beta), \underline{\Omega}(\psi)$	$2 \times 2$ -Rotationsmatrizen i.Abh. der Eulerwinkel
$\Omega$	$3 \times 3$ -Rotationsmatrix der Bewegungstransformation

---

$\omega$	Vektor der Eulerwinkel $[\omega_x, \omega_y, \omega_z]^T$ der Bewegung
$\Omega(\omega_x, \omega_y, \omega_z)$	$3 \times 3$ -Rotationsmatrizen der Bewegungstransformationen der Kamerakoordinatensysteme i. Abh. der Eulerwinkel
$\underline{\Omega}$	$2 \times 2$ -Rotationsmatrix der Bewegungstransformation in der Bodenebene
$r_x, r_y, r_z$	Eulerwinkel der extrinsischen Transformationen
$\beta$	Eulerwinkel der Orientierung zwischen zwei Kameramodulen innerhalb der Grundebene
$\hat{\mathbf{x}}_k^+, \mathbf{P}_k^+$	Aktualisierte Zustands- und Kovarianzmatrizen zum Zeitpunkt $k$
$\hat{\mathbf{x}}_k^-, \mathbf{P}_k^-$	Prädizierte Zustands- und Kovarianzmatrizen zum Zeitpunkt $k$
$\hat{\mathbf{x}}_k$	Schätzung bzw. Zustand von $\mathbf{x}$ zum Zeitpunkt $k$
$\hat{\mathbf{x}}_k, \mathbf{P}_k$	Zustand und Kovarianzmatrix zum Zeitpunkt $k$
$\zeta_{\mathbf{C}}(\mathbf{C})$	Abbildung der Transformation $\mathbf{C}$ auf den Parametervektor $\mathbf{x}_{\mathbf{C}} = [t_x, t_y, t_z, r_x, r_y, r_z]^T$
$\zeta_{\mathbf{M}}(\mathbf{M})$	Abbildung der Transformation $\mathbf{M}$ auf den Parametervektor $\mathbf{x}_{\mathbf{M}} = [v_x, v_y, v_z, \omega_x, \omega_y, \omega_z]^T$
$\sigma, \sigma^2$	Standardabweichung bzw. Varianz
$\mathbf{T}^G_i, \mathbf{x}_i^G, \mathbf{P}_i^G$	Globale (fusionierte) Schätzung der Parameter von Modul $\mathcal{M}_i$
$\theta$	Funktion zur Prädiktion der Messung beim Bündelausgleich
$\Lambda_{\mathbf{M}}$	Schwellwert zur Klassifikation des Bewegungstyps
$\Lambda_{\mathbf{F}}$	Schwellwert zur Bewertung des Abstands eines Punktes zur Epipolargeraden
$\Lambda_{\varphi}$	Schwellwert zur Bewertung des Winkels zwischen zwei Vektoren
$\check{\mathbf{T}}_i^{(j)}, \check{\mathbf{x}}_i^{(j)}, \check{\mathbf{P}}_i^{(j)}$	(Lokale) Schätzung der Parameter von Modul $\mathcal{M}_i$ anhand der lokalen Informationen von Modul $\mathcal{M}_j$
$\mathbf{T}, \mathbf{T}(\mathbf{R}, \mathbf{t})$	Allgemeine, homogene, Euklidische Transformationsmatrix $\in \mathbb{R}^{4 \times 4}$
$\mathbf{t}$	3D-Translationsvektor $[t_x, t_y, t_z]^T$ der extrinsischen Transformationen

$\underline{t}$	2D-Translationsvektor $[\underline{t}_x, \underline{t}_z]^T$ der extrinsischen Transformationen in der Bodenebene
$\underline{v}$	3D-Translationsvektor $[\underline{v}_x, \underline{v}_y, \underline{v}_z]^T$ der Bewegungstransformationen
$\mathbf{v}^q$	Rein imaginäres Quaternion $[0, \mathbf{x}]$ zum Vektor $\mathbf{v}$
$\underline{\mathbf{v}}$	2D-Translationsvektor $[\underline{v}_x, \underline{v}_z]^T$ der Bewegungstransformationen in der Bodenebene
$\underline{t}_x, \underline{t}_z$	Extrinsische Translationskomponenten innerhalb der $X - Z$ -Bodenebene
$\dim(\mathbf{a})$	Dimension des Vektors $\mathbf{a} \in \mathbb{R}^N$
$\langle \mathbf{a}, \mathbf{b} \rangle$	Skalarprodukt $\mathbf{a}^T \mathbf{b}$
$\mathbf{v}^s$	Skalierter Vektor $\mathbf{v} / \ v\ $
$\delta, \delta_x, \delta_y$	Differenz- bzw. Aktualisierungsvektoren
$\mathbf{a}, \mathbf{b}$	Allgemeine Vektoren, $\mathbf{a} = [a_1, a_2, \dots, a_N]^T \in \mathbb{R}^N$
$\mathbf{c}$	Durchstoßpunkt oder Bildhauptpunkt $[c_x, c_y]^T$
$\mathbf{e}_n$	Kanonischer Einheitsvektor im $\mathbb{R}^3$ , insbesondere $\mathbf{e}_1 = [1, 0, 0]^T, \mathbf{e}_2 = [0, 1, 0]^T, \mathbf{e}_3 = [0, 0, 1]^T$
$ \mathcal{S} $	Kardinalität, d. h. Anzahl der Elemente einer Menge $\mathcal{S}$
$ \mathbf{a} $	Betrag von $\mathbf{a}$
$\tilde{\mathbf{a}}$	Homogene Koordinaten $\tilde{\mathbf{a}} = [\mathbf{a}^T, 1]^T$
$[\mathbf{v}]_\times$	Schiefsymmetrische $3 \times 3$ -Matrix von $\mathbf{v} \in \mathbb{R}^3$
$\underline{v}_x, \underline{v}_z$	Translatorische Bewegungskomponenten innerhalb der $X - Z$ -Bodenebene
$\omega_x, \omega_y, \omega_z$	Eulerwinkel der Bewegungstransformationen
$\psi$	Eulerwinkel der Bewegung eines Kameramoduls innerhalb der Grundebene
$\mathbf{P}$	Unsicherheit eines Zusstands- bzw. Parametervektors
$\mathbf{x}$	Allgemeiner Parametervektor
$\mathbf{z}$	Messung
$c(\mathbf{x})$	Gewicht des Punktes $\mathbf{x}$

$d$	Abstand
$h_y$	Translatorische Komponente zur Transformation in die Bodenebene (Höhe)
$M, N, R$	Dimensionsbezeichner
$p(\mathbf{a})$	<i>a priori</i> Wahrscheinlichkeit
$p(\mathbf{a} \mathbf{b})$	<i>a posteriori</i> oder bedingte Wahrscheinlichkeit
$\cup$	Symbolischer Fusionsoperator
$A \iff B$	Äquivalenzrelation

## **Teil I**

### **Einleitung**



# 1

---

## **Einführung in die Kalibrierung multiokularer, nichtüberlappender Kameras in Fahrzeugen**

---

### **1.1 Die Kalibrierung und ihre Bedeutung in der Bildfolggenauswertung – Zielsetzung und Herausforderungen**

Die Auswertung von bildgebenden Sensoren auf bewegter Plattform ist ein wichtiges und spannendes, zeitgenössisches Forschungsthema, das immer mehr Einzug in unseren Alltag hält. So gehören Kameras in Straßenfahrzeugen mittlerweile bei vielen Autoherstellern zur Serienausstattung. Im Bereich der Fahrerassistenz sind Kameras weit verbreitet [Dic02]. Kamerabasierte Anwendungen in PKWs unterstützen den Fahrer bei der Umfelderfassung, beispielsweise bei Einparkmanövern [Par10], bei Überhol- und Einschervorgängen [Bae05], in Form von Fahrspurassistenten [Sch06] oder als markierendes Licht [MT11].

Aber auch im Luftverkehr sind Kamerasensoren weitläufig im Einsatz. Fliegende Drohnen übermitteln Videoaufzeichnungen zu Aufklärungszwecken [Kru01] [Mue07]. Ob zu Lande oder in der Luft, die Detektion, Verfolgung und Klassifikation von Objekten sind

hierbei wesentliche Kernaufgaben [Dan02], [Kol92], [Kan05], [Mue08], [Bac06].

Vor allem die Aktivitäten und Anstrengungen im Bereich des autonomen Fahrens trieb in den vergangenen Jahren den Stand der Forschung massiv voran [Hei00], [Hei02], [Sti08], [DAR]. Im gleichen Atemzug gewinnt die semantische Auswertung von Multisensorsystemen im Straßenverkehr immer mehr an Bedeutung [Wac97], [Kan80], [Vac09], [Hum08].

Durch fallende Marktpreise für Videosensoren werden Mehrkamerasysteme in zunehmendem Maße attraktiv. Stereokameras berechnen durch Betrachten einer Szene aus zwei unterschiedlichen Standpunkten deren dreidimensionale Struktur. Voraussetzung hierfür ist die Kenntnis der relativen Lage der Kameras zueinander. Mehrkamerasysteme in Straßenfahrzeugen rekonstruieren mitunter ganze Städtemodelle [Mor06]. Aber auch monokulare Kamerasysteme sind in der Lage 3D-Informationen zu extrahieren [Pol99a] [New10]. Statt mehreren Kameras an verschiedenen Positionen werden verschiedenen Kamerapositionen derselben Kamera zu unterschiedlichen Zeitpunkten verwendet. D. h. zusätzlich zur Szenenstruktur wird die Kameraposition und damit die Bewegung der Kamera geschätzt. Derartige Schätzverfahren werden gemeinhin als *Structure from Motion*-Verfahren (SFM) bezeichnet. Die technischen Verfahren und Möglichkeiten der 3D-Rekonstruktion haben in den vergangenen Jahren einen rapiden Fortschritt erlebt. Abb. 1.1 und 1.2 zeigen Beispiele von aus monokularen Bildfolgen extrahierten 3D-Informationen.

Mehrakamerasysteme in Fahrzeugen ermöglichen i. A. eine erweiterte Perzeption der Umwelt, als dies mit einzelnen Kameras der Fall wäre. Aus technischen Kapazitäts-, Kosten- oder Designgründen ist die mögliche Anzahl an Kameras, die in ein Fahrzeug eingebaut werden kann, in der Praxis begrenzt. So kann es vorkommen, dass Kamerstellungen ohne überlappende Sichten entstehen.

Aber auch in Mehrkamerasystemen, bei denen die Kameras nicht dieselbe Szene sehen, ist die relative Lage der Kameras von Bedeutung. Die Kenntnis der Lage der Kameras ermöglicht es, die aus den unterschiedlichen Videosensoren extrahierte 3D-Information in einem gemeinsamen Koordinatensystem zu fusionieren. Die 3D-Daten können hier-

bei von vielfältiger Art sein. Beispielsweise können von den einzelnen Videosensoren generierte 3D-Punktwolken zusammengeführt werden. Hierbei spielt es keine Rolle, ob die 3D-Daten aus mehreren Stereopaaren stammen, oder bewegungsbasiert mittels SFM-Techniken erzeugt wurden [Pol99a],[Pol99b]. Eine sehr zentrale Bedeutung kommt der Bodenebene, also der befahrbaren Fläche zu. Beispielsweise ermöglicht die Kenntnis der Bodenebene die Generierung einer Vogelperspektive [Liu08] und Abb. 1.5. Auch in zweidimensionalen Bilddaten detektierten Objekten kann ein 3D-Bezug zugewiesen werden. Selbst ohne explizite Kenntnis der Entfernung des Objekts lässt sich aus der Verbindungsgeraden zwischen Kamerazentrum und Bildpunkt die dreidimensionale Richtung im Raum bestimmen, in der sich das Objekt befindet. D. h. mehrere Objekte können geometrisch zueinander in Relation gesetzt werden. Andere Applikationen greifen direkt auf 3D-Umgebungsmodelle oder Landmarken zurück, so dass das Fahrzeug in der Lage ist, sich relativ zur Szene zu orientieren [Emt10]; eine für Navigationsaufgaben oftmals eingesetzte Technik.

Der Prozess der Ermittlung der geometrischen Anordnung von Mehrkamerasystemen, die sog. *extrinsische Kalibrierung*, ist gemeinhin ein recht aufwändiger Prozess. Hierbei bezeichnen die extrinsischen Parameter die relative<sup>1</sup> Lage und Position der Kamera im dreidimensionalen Raum. Meist werden für derartige Aufgabenstellungen wohlbekannte Muster verwendet, die in verschiedenen Perspektiven von der Kamera erfasst werden. Durch die Kenntnis der Proportionen des Musters ist man in der Lage, sowohl die intrinsischen Parameter wie Brennweite und Linsenverzerrung, als auch die relative Lage der Kamera zum Muster zu bestimmen. Musterbasierte Verfahren finden auch Anwendung bei der Kalibrierung von Mehrkamerasystemen, zur Berechnung der relativen Lage der Kameras zueinander.

Kalibrierparameter werden meist als fix und unveränderlich angenommen. In Fahrzeugen und mobilen Anwendungen kann i. d. R. jedoch nicht davon ausgegangen

---

1 Die extrinsischen Parameter transformieren das Kamerakoordinatensystem mittels einer Euklidischen Transformation in ein anderes Koordinatensystem, wie z. B. das einer Straßenkarte, einer anderen Kamera, einem Gebäudemodell, einer Plattform (wie z. B. einem Roboterarm oder einem Fahrzeug), o. ä.

werden, dass die Optiken eines Mehrkamerasystems unverändert bleiben. Selbst die intrinsischen Parameter – wenngleich nicht Gegenstand der folgenden Untersuchungen – unterliegen im Laufe der Zeit Umwelteinflüssen und Schwankungen [Dan07]. Temperaturunterschiede beeinflussen die Materialeigenschaften und können ganze Karosserien verziehen und insbesondere Einfluss auf integrierte Sensoren und deren relative Anordnung nehmen. Ständige Vibrationen im laufenden Betrieb, Kollisionen oder Erschütterungen, wie sie z. B. beim Be- und Entladen der Plattformen oder unebenem Gelände entstehen können, beanspruchen das mechanische Grundgerüst u. U. massiv und können zu Verschleißerscheinungen des Materials führen.

Mehrkamerasyteme zur dreidimensionalen Szenenauswertung auf mobilen Plattformen im Langzeiteinsatz benötigen eine dauerhaft korrekt parametrierte Kalibrierung. D. h. um keine Leistungseinbußen bei längeren Einsatzzeiten hinnehmen zu müssen, muss sichergestellt sein, dass sich die Kalibrierung einer Änderung der Systemanordnung anpasst. Erfolgt dies nicht, so kann u. U. bereits eine leichte Abweichung in der Kameraanordnung (z. B. eine leichte Drehung oder ein Versatz der Anordnung) zu einer völligen Unbrauchbarkeit der Auswerteergebnisse führen. Dies kann teure und aufwendige Werkstattbesuche für Rekalibrierungen nach sich ziehen.

Aus industrieller Sicht, bei der Produktion von kamerabestückten Fahrzeugen, stellt die Kalibrierung einen nicht zu vernachlässigenden Kostenpunkt dar. Beispielsweise ist die Kalibrierung von Kameras in kommerziellen Straßenfahrzeugen auf einem Prüfstand bei jedem Fahrzeug mit Zeit, ggf. Personal und damit mit Kosten verbunden. Bei Systemen mit derart hohen Anforderungen ist daher eine kontinuierliche Rekalibrierung zur Laufzeit des Systems sinnvoll und wünschenswert, d. h. eine Kalibrierung der Kameras im laufenden Fahrbetrieb, ohne separate Kalibriervorrichtungen.

Die vorliegende Arbeit beschreibt und untersucht ein Verfahren zur kontinuierlichen, extrinsischen Kalibrierung verteilter Kameras mit nichtüberlappenden Sichten in Fahrzeugen. Der Ansatz nutzt hierbei die Mobilität der Plattform aus, so dass eine Unabhängigkeit von Kalibermustern erreicht wird. Die Kalibrierung erfolgt rein bildbasiert. Die Hinzunahme zusätzlicher Sensoren wird zwar an geeigneter Stelle diskutiert, doch

ist es als Ziel dieser Arbeit zu verstehen, die Möglichkeiten der reinen Videoauswertung zu erforschen.

Die im folgenden Kontext diskutierten Kameraanordnungen werden als **MMD**-Anordnungen bezeichnet (s. Abb. 1.3). MMD steht hierbei für **Multiokular**, **Mobil** und **Disjunkt**. Solche Kameraanordnungen zeichnen sich zunächst dadurch aus, dass mehrere Kameras auf einer mobilen Plattform, z. B. einem PKW, einem Landroboter, oder aber auch Luftfahrzeugen wie UAVs<sup>2</sup>, fest justiert sind. D. h. die gesamte Kameraanordnung ist zwar mobil, die Kameras selbst jedoch sind zueinander fest justiert. Die Sichten der Kameras sind hierbei disjunkt, d. h. nichtüberlappend<sup>3</sup>. Für die Praxis bedeutet dies, dass sich keine Punktkorrespondenzen zwischen den Kamerabildern bestimmen lassen.

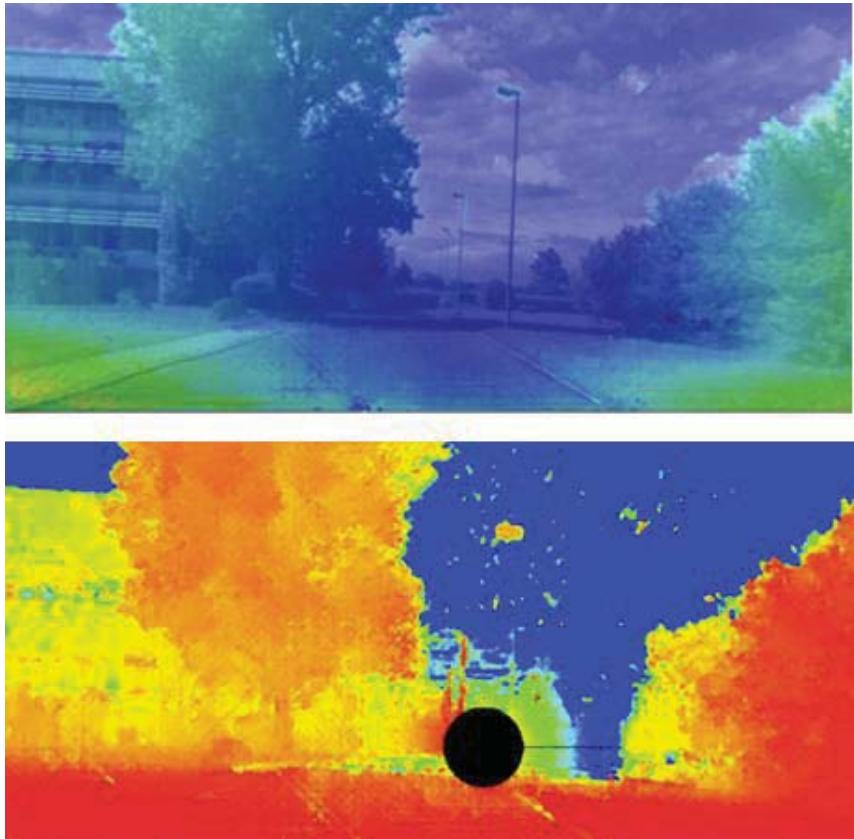
Wenn gleich in der vorliegenden Arbeit ein allgemeines Modell zur Kalibrierung von disjunkten Kameraanordnungen auf beliebigen mobilen Plattformen vorgestellt wird, so liegt der Fokus im Folgenden auf dem Einsatz in Straßenfahrzeugen. Die Fortbewegung auf einer Bodenebene führt zu zusätzlichen Restriktionen, die einer gesonderten Betrachtung bedürfen. Der Umstand der angenommenen Disjunktheit der Sichtfelder stellt die größte Schwierigkeit und Herausforderung der kontinuierlichen Kalibrierung von MMD-Systemen dar. Dieser Umstand fällt umso mehr ins Gewicht, wenn die Anzahl der Kameras *a priori* variabel gestaltet wird (wie dies in der vorliegenden Arbeit der Fall ist). Der Aufwand zur Kalibrierung von  $\mathfrak{N}$  Kameras soll sich also nicht nur linear bezüglich  $\mathfrak{N}$  verhalten, sondern annähernd konstant sein. Im Zuge der sich rasant entwickelnden Hardwaretechnologien und parallelen Rechnerarchitekturen soll dadurch die Möglichkeit zur Parallelisierung der Rechenprozesse gewährleistet bleiben. Obwohl die Auswertung von Mehrkamerasyystemen einen deutlichen Mehraufwand mit sich bringt, so bietet sich doch aufgrund der Mehrinformation verteilter Kameras auch

---

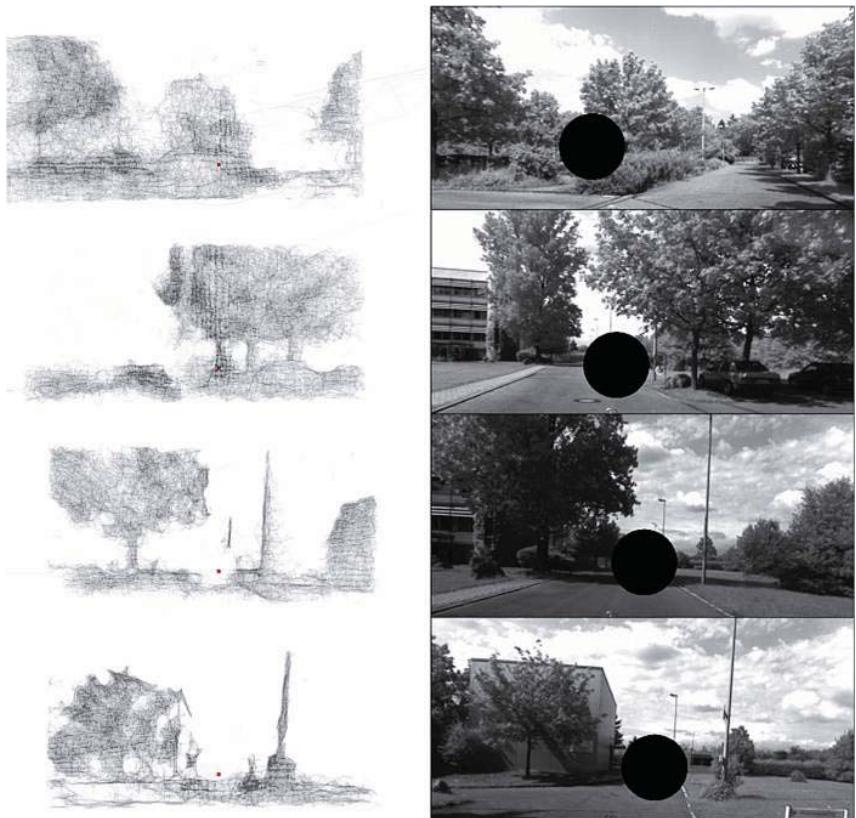
2 engl. *Unmanned Aerial Vehicles*

3 Die Disjunktheit der Sichtfelder bezieht sich hierbei nicht auf die Sichtpyramiden der Kameras, sondern auf die in den Kamerabildern abgebildeten Szenen. D. h. für ein disjunktes Kamerapaar gibt es keine Szenenpunkte, die in beiden Kameras gleichzeitig sichtbar sind.

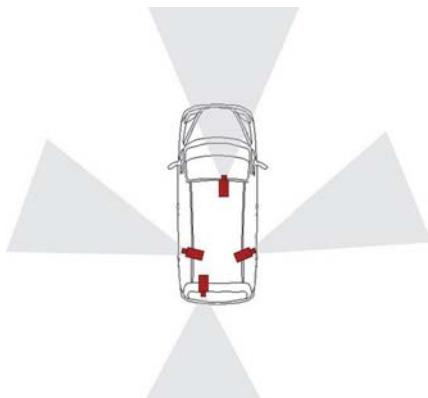
die Chance einer gesteigerten Informationsnutzung. Somit stehen letztlich neben der Kontinuität der Parameterschätzung, der Problematik der disjunkten Kameraansichten und der Skalierbarkeit des Kalibrierverfahrens hinsichtlich der Anzahl der Kameras, die optimierte Informationsnutzung in MMD-Systemen im Fokus der hier vorliegenden Arbeit.



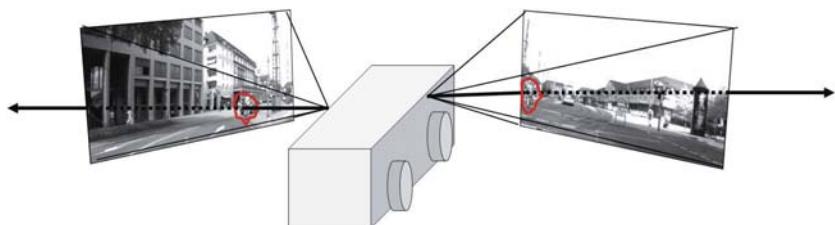
**Abbildung 1.1:** Monokulare 3D-Rekonstruktion. Oben: Dichte Disparitätskarte der Ver-schiebung der Bildstrukturen. Unten: Tiefenkarke (rot=nah, blau=fern). Der schwarze Kreis indiziert die Bewegungsrichtung.



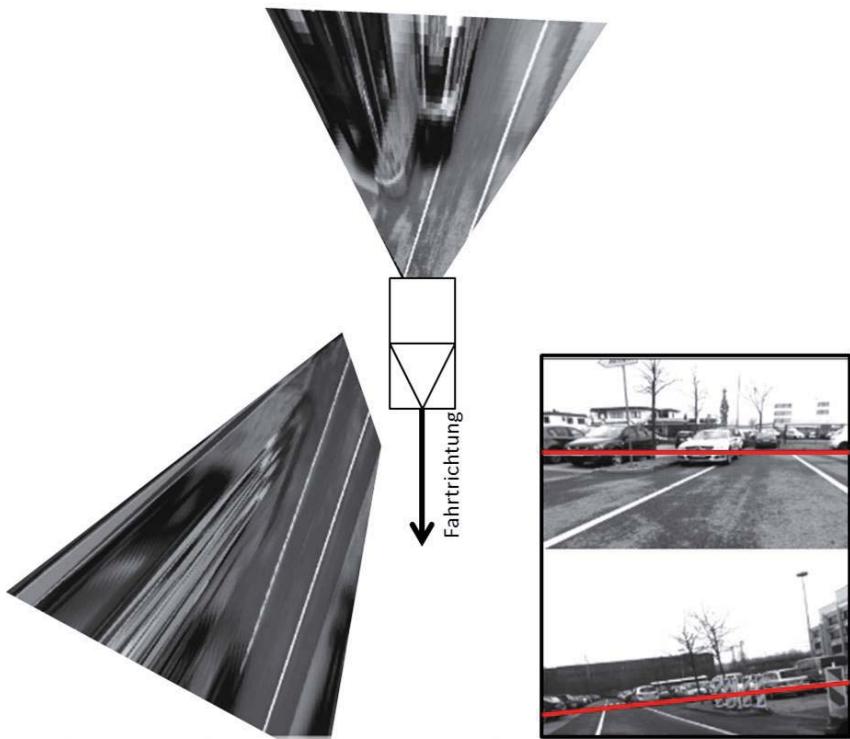
**Abbildung 1.2:** Monokulare 3D-Modellierung. Links: Polygonisiertes 3D-Modell aus einer monokular generierten, verrauschten 3D-Punktwolke wie aus Abb. 1.1. Rechts: Ausschnitte aus der originalen Bildsequenz. Der schwarze Kreis indiziert die Bewegungsrichtung.



**Abbildung 1.3:** Beispiel einer mobilen Plattform mit mehreren nichtüberlappenden Kameras.



**Abbildung 1.4:** Objektdetection in einem Multikamerasystem. Ist die Kalibrierung, d. h. die relative Lage der Kameras zueinander bekannt, so kann der Sichtstrahl bestimmt werden, entlang dessen sich das Objekt in der dreidimensionalen Szene befinden muss.



**Abbildung 1.5:** Synthetische Vogelperspektive mit zwei Kameras in einem Fahrzeug. Die Kamerabilder werden mit Hilfe der Kalibrierinformation auf eine Bodenebene projiziert. Rechts unten: Die Originalbilder mit eingezeichnetem Horizont.

## 1.2 Stand der Forschung

Für die Kalibrierung und Selbstkalibrierung von Mehrkamerasystemen existieren verschiedenste Lösungsansätze für die unterschiedlichsten Kamerakonstellationen. Dang et al. [Dan09], [Dan04] beispielsweise stellte eine kontinuierliche Selbstkalibrierung von Stereokameras vor. Basierend auf korrespondierenden Merkmalen wurden verschiedene Bedingungsgleichungen in ein iteratives, erweitertes Kalmanfilter eingebettet und die Parameter fortlaufend aktualisiert. Dieser Ansatz ließe sich prinzipiell auch auf Mehrkamerasysteme mit überlappenden Sichten beliebiger Mächtigkeit übertragen. Die Bedeutung dieser Arbeit liegt insbesondere in der schritthaltenden Schätzung der Kameraparameter, wie es auch in dieser Arbeit für die extrinsischen Kameraparameter erfolgen soll.

Die Kalibrierung von Kameras mit nichtüberlappenden Sichten gestaltet sich ungleich schwieriger, da sich keine unmittelbaren<sup>4</sup> Korrespondenzen zwischen den Kamerabildern ermitteln lassen. Für die Kalibrierung von nichtüberlappenden, statischen Kameranetzwerken existieren Lösungen [Kum08], [Rah04], die sich für eine kontinuierliche Selbstkalibrierung allerdings nicht umsetzen lassen.

Die Recherche fokussiert sich daher gezielt auf Anwendungen auf mobilen Plattformen. Lebraly [Leb11] und Davison [Car11] nutzen die Mobilität der Plattform, um korrespondierende Merkmale zu detektieren. Innerhalb eines geschlossenen Raums fährt die Plattform einen vorgegebenen Parcours ab, so dass alle Kameras zeitversetzt dieselbe Szene sehen. Mit Hilfe eines Bündelausgleichs oder eines SLAM<sup>5</sup>-Verfahrens

---

<sup>4</sup> Die Unimittelbarkeit bezieht sich hierbei auf die Gleichzeitigkeit der Korrespondenzfindung. Temporär versetzt lassen sich prinzipiell durchaus dieselben Szenenobjekte in den verschiedenen Kamerasichten wiederfinden.

<sup>5</sup> *Simultaneos Localization and Mapping*. In der Literatur werden im Kontext der bewegungsbasierten 3D-Rekonstruktion die Bezeichnungen SLAM und SFM des öfteren synonym verwendet.

erfolgt die simultane Berechnung der Punktkoordinaten sowie der Kameraanordnung unter Minimierung des Projektionsfehlers.

In einem ausgedehnteren räumlichen Diskurs versuchten Lamprecht et al. [Lam07] Verkehrsschilder als Kalibrierobjekte zu verwenden. Dadurch ist das Verfahren nicht mehr an bestimmte räumliche Umgebungen gebunden. Jedoch wird die Leistungsfähigkeit dieses Ansatzes durch die unterschiedlichen perspektivischen Erscheinungsformen der Objekte und der damit verbundenen sinkenden Zuordnungsgenauigkeit beeinträchtigt.

In [Pag10b] wird ein musterbasiertes Verfahren vorgeschlagen, bei dem mittels mehrerer, gezielt platzierte Kalibermuster Korrespondenzen detektiert werden (s. Abb. 1.6). Durch die Kenntnis der metrischen Maße der Muster ist es möglich, die Transformation zwischen Kamera- und Musterkoordinatensystem zu bestimmen. Darauf aufbauend können die Positionen und Orientierungen der Muster sowie der Kameras berechnet werden. Ein solcher Ansatz eignet sich jedoch lediglich als Initialkalibrierung und erfordert wegen des Einsatzes der Muster einen hohen Vorbereitungsaufwand.

Ein anderer Ansatz stammt aus dem Fachbereich der Robotik. Die sog. *Hand-Augen-Kalibrierung* versucht die Euklidische Transformation zwischen einer Kamera (Auge) und dem Endglied eines Roboterarms (Hand) zu ermitteln [Shi89], [Tsa89], [Par94], [Str06]. Kapitel 4.1 befasst sich eingehender mit diesem Thema. Dieses Problem wurde in den vergangenen Jahrzehnten ausgiebig untersucht. In den klassischen Anwendungsfällen wurde mittels eines Kalbiermusters die Position der Kamera geschätzt, die Lage des Roboterarms ist i. A. aus der Steuerung bekannt. Anhand verschiedener Positionen der Kamera und des Roboterarms lässt sich auf die relative Anordnung und Orientierung rückschließen. Auf diese Weise lassen sich Kameras auch ohne korrespondierende Merkmale extrinsisch kalibrieren. Die rotatorische Orientierung des Kamerakoordinatensystems relativ zum Roboterarm lässt sich prinzipiell getrennt von der translatorischen Anordnung berechnen. In [Hor95] und [Dan98] wurde eine simultane Schätzung von sowohl den Rotations- als auch den Translationsparametern vorgestellt und untersucht.



**Abbildung 1.6:** Anordnung von Kalibermustern mit unterschiedlichen aber bekannten Eigenschaften zur Kalibrierung einer MMD-Zweikameraanordnung.

Auch im Automotive-Bereich findet die Hand-Augen-Kalibrierung mit leicht modifizierter Aufgabenstellung Anwendung. Hierbei handelt es sich um mehrere Kameras, deren relative, extrinsische Lage zueinander bestimmt werden soll. Ruland et al. [Rul11a], [Rul11b] minimieren den Projektionsfehler von extrahierten Markmalen in Abhängigkeit von sowohl den extrinsischen Parametern, als auch der Fahrzeugbewegung. Die Schätzung bezieht sich hierbei allerdings primär auf die rotatorischen Parameter. In [Rul10] wird für langsame Einparkszenarien auch die Position zweier Kameras zueinander geschätzt.

Die reine Orientierungsschätzung von Kameras stellt eine Unterdisziplin dar. Hierfür werden u. a. gezielt Bodenmerkmale, Fahrbahnmarkierungen oder Bewegungseigenschaften herangezogen [Mik10], [Cou02], [Sim03].

Die Transformation zwischen Kamerapositionen kann nicht nur durch bekannte Kalibriermuster erreicht werden. Gerade im Bereich mobiler Plattformen ist die Bedeutung der musterlosen Positionsschätzung der Kameras groß, um Aufwand zu vermeiden. Die Schätzung der Bewegung der einzelnen Kameras kann mittels sog. *Structure from Motion*-Verfahren (SFM) erfolgen. Auf diese Weise gelöste Hand-Augen-Kalibrierungen werden in der Literatur oft auch als erweiterte Hand-Augen-Kalibrierung bezeichnet [And01], [Sch05c]. Esquivel et al. [Esq07] kalibrierten eine nichtüberlappende Zweikameraanordnung durch händisches Schwenken der Kameras. Allerdings wurde die Besonderheit rein planarer Bewegungen, wie sie bei Straßenfahrzeugen vorwiegend auftreten, noch nicht berücksichtigt.

Die bildgestützte Bewegungsschätzung, auch oft als visuelle Odometrie bezeichnet, stellt im Bereich der Bildfolgengenauswertung eine eigene Disziplin dar. SFM-Verfahren können gemeinhin nur bei auf einen skalaren Faktor hin bestimmt werden, d. h. der Absolutbetrag der Geschwindigkeit lässt sich ohne weitere sensorische Information nicht aus den reinen Bilddaten extrahieren [Har06]. Zur monokularen Bewegungsschätzung existieren bereits vielfältige Ansätze, z. B. [Nis04], [Dav07], [Pag09], [SÖ5], [Cli08]. Sie soll daher auch nicht im Zentrum der hier vorliegenden Arbeit stehen.

Seo et al. [Seo09] schlagen eine optimale Lösung der Hand-Augen-Kalibrierung zur Schätzung der Orientierung vor, basierend auf dem sog. *Second-Order Cone Programming*-Optimierungsverfahren [Har09]. Heller et al. [Hel11] greifen dieses Prinzip auf, um sowohl die Orientierung als auch die Translation zweier Kameras an einem Roboterarm zueinander zu berechnen. Hierbei berücksichtigen sie auch die Besonderheit von SFM-Verfahren, dass die translatorische Bewegung nämlich nur bis einen skalaren Faktor werden kann.

Auf dem Stand der aktuellen Forschung aufbauend behandelt die vorliegende Arbeit insbesondere folgende Aspekte:

**Kontinuierliche Schätzung extrinsischer Parameter von Mehrkamerasystemen**  
Gerade bei längeren Fahrten mobiler Kameraplatzformen ist es wünschenswert, die

Kameraparameter kontinuierlich mitzuführen und zu aktualisieren. Änderungen in der extrinsischen Kameraanordnung, wie sie durch äußere Einflüsse bewirkt werden können, sollen dadurch kompensiert und aufwändige Rekalibrierungen ersetzt werden.

**Erweiterte Hand-Augen-Kalibrierung bei planaren Bewegungen** Straßenfahrzeuge vollführen i. A. planare Bewegungen, d. h. sie fahren innerhalb einer Ebene. Eine derartige Einschränkung der Bewegungsart begrenzt die Schätzbarkeit aller extrinsischen Freiheitsgrade und bedarf daher einer gesonderten Betrachtung.

**Nutzung des erhöhten Informationspotentials von Mehrkamerasystemen** Das Gros der in der Fachliteratur behandelten Kalibrierverfahren behandelt Systeme mit zwei bis maximal drei nichtüberlappenden Kameras. Hierbei liefert ein 3-Kamerasystem mehr Information als ein 2-Kamerasystem. Die gezielte Ausnutzung dieser Mehrinformation sollte idealerweise im Systementwurf berücksichtigt werden.

**Integration des Problems skaliertener Bewegungen bei der erweiterten Hand-Augen-Kalibrierung** Es ist bekannt, dass die rein bildbasierte Schätzung der absoluten Geschwindigkeit einer einzelnen Kamera nicht möglich ist. Selbst die relative Geschwindigkeit lässt sich über längere Zeiten hinweg nicht immer robust schätzen. Die Beträge der zeitlichen, translatorischen Verschiebung jeder Kamera, insbesondere die relativen Beträge zwischen den Kameras, sind jedoch essentiell für die erweiterte Hand-Augen-Kalibrierung. Die Handhabung dieser Translationsbeträge wird in der Literatur bislang kaum berücksichtigt.

**Skalierbarkeit und Parallelisierbarkeit bei Mehrkamerasystemen** Je mehr Kameras in das System integriert sind, desto mehr steigt die Rechenlast. Von daher ist ein parallelisierbarer Systementwurf wünschenswert, der eine Teilung der Rechenlast auf einzelne Rechenmodule zulässt. In der vorliegenden Arbeit wird eine Kamera als eine eigenständige Recheneinheit betrachtet.

## 1.3 Die Kalibrierung als »Henne-Ei-Problem«

### 1.3.1 Prämissen

Inhalt dieser Arbeit ist die Schätzung der extrinsischen Parameter, d. h. der Euklidischen Transformationen zwischen den einzelnen Kameras einer Mehrkameraanordnung. Hierbei gelten die folgenden sechs Prämissen:

- Die Kameras seien relativ zueinander fix angeordnet.
- Die intrinsischen Parameter der Kameras (Brennweite, Bildhauptpunkt, Linsenverzerrung) seien bekannt.
- Die Kameras befinden sich auf einer beweglichen Plattform.
- Die Sichtfelder der Kameras seien nichtüberlappend.
- Die Anzahl der Kameras  $N > 1$  sei nach oben beliebig erweiterbar<sup>6</sup>.
- Die Euklidischen Transformationsparameter zur Beschreibung des Systemzustands seien mit einem normalverteilten, additiven Rauschen behaftet.

Basierend auf diesen sechs Annahmen wird in dieser Arbeit ein allgemeines Kalibriermodell zur Bestimmung der extrinsischen Sensoranordnung vorgestellt.

Bezüglich der Art der Plattform lassen sich drei weitere Prämissen anführen:

- Die Plattform bewege sich planar auf einer zweidimensionalen Ebene.
- Alle Kameras sehen die gemeinsame Bodenebene.

---

<sup>6</sup> Ein realistische und sinnvolle Anordnung nichtüberlappender Kameras besteht aus etwa zwei bis sechs Kameras oder Kameraparen – im Fall von Stereokameras. Mehr Kameras zielen tendenziell auf eine überlappende Anordnung ab, für die anderweitige Kalibrierverfahren in der Literatur existieren. Mit Weitwinkelobjektiven kann z. B. bereits mit vier Kameras eine in den Randbereichen überlappende Anordnung (Rundumsicht) erreicht werden.

- Die Bewegung der mobilen Plattform gehorche einer nichtholonomen Bedingung, die rotatorische Bewegungen nur in Abhängigkeit einer translatorischen Bewegung zulässt (siehe Abschnitt 2.2).

Die drei erweiterten Prämissen sind insbesondere für Straßenfahrzeuge relevant. Sie erfordern auch ein erweitertes Kalibriermodell, da das allgemeine Modell durch die Einschränkung auf planare Bewegungen nicht mehr alle zu bestimmenden Freiheitsgrade abdeckt.

### 1.3.2 Lösungsansatz und Beitrag der Arbeit

Ohne überlappende Sichtfelder ist keine direkte Korrespondenzbestimmung zwischen den einzelnen Bildern möglich<sup>7</sup>. Stattdessen macht man sich die Mobilität der Plattform zunutze. Aufbauend auf Ansätzen aus der Robotik, konkret der Hand-Augen-Kalibrierung (siehe Abschnitt 4.1), lassen sich aus der Bewegung der einzelnen Kameras Rückschlüsse auf deren Anordnung und Orientierung ziehen. Somit ist eine robuste Bewegungsschätzung Grundlage des hier vorgestellten Kalibrierverfahrens. Eine intrinsische Kalibrierung der Kameras lässt sich mit musterbasierten Standardverfahren (z. B. [Zha00]) durchführen oder auch als Selbstkalibrierung realisieren (z. B. [Fau92]). Die Schätzung der intrinsischen Parameter ist nicht Gegenstand dieser Arbeit.

Im Folgenden wird der wesentliche Kern und Beitrag dieser Arbeit zusammengefasst.

**Bewegungsbasierte Kalibrierung** Der hier vorgestellte Lösungsansatz baut auf dem Prinzip der sog. *Hand-Augen-Kalibrierung* auf. Aus den Bewegungen der einzelnen

---

<sup>7</sup> Zeitliche Integration der Punkte kann u. U. dazu eingesetzt werden, zeitübergreifend Bildpunktcorrespondenzen oder 3D-Punkte zwischen einzelnen Sichten zu referenzieren, siehe z. B. [Dav07]. Im regulären Fahrbetrieb eines Straßenfahrzeugs jedoch, führen perspektivische Verzerrungen, unterschiedliche Erscheinungen von Objekten in den einzelnen Sichten, eine begrenzte, quantisierte Tiefenauflösung und Mehrdeutigkeiten von Korrespondenzzuordnungen zu massiven praktischen Schwierigkeiten. D. h. es ist zu erwarten, dass zeitlich versetzt ermittelte Korrespondenzen zwischen den Sichten mit starken Unsicherheiten behaftet sind. Eine solche Vorgehensweise ist zudem sehr speicherintensiv, da möglichst viele Punkte und Merkmale zeitlich mitgeführt werden müssen, um das Wiedererkennen korrespondierender Objekte überhaupt gewährleisten zu können. Außerdem wäre eine Skalierung und Parallelisierung hinsichtlich des Rechenaufwandes nur bedingt möglich, da die Fusion üblicherweise auf Sensordatenebene stattfindet (ein Objekt, das in einer Kamera gesehen wurde, kann prinzipiell wieder in allen andern Kameras auftauchen).

Kameras lassen sich Bedingungen formulieren, mittels derer sich – unter bestimmten Umständen – auf die extrinsische Anordnung der Kameras rückschließen lässt. Die aus der Literatur bekannten erzielten Ergebnisse rechtfertigten die Übertragung dieser Methodik auf den Kontext der Kalibrierung von Kameras in Fahrzeugen. Die Lösung des Problems unterteilt sich demnach in die Schätzung der Kamerabewegungen anhand extrahierter und verfolgter Bildmerkmale, und – darauf aufbauend – die Bestimmung der extrinsischen Kalibrierparameter. Die Bewegungsschätzung erfolgt mittels Bündelausgleich (kurz: BA, [Tri00]), d. h. der simultanen Rekonstruktion der Szene sowie der Euklidischen Bewegungsparameter der Kamera.

**Erweitertes Modell für planare Bewegungen** Der Fokus in der vorliegenden Arbeit liegt auf dem Einsatz der kontinuierlichen Kalibrierung in Straßenfahrzeugen. Straßenfahrzeuge zeichnen sich u. a. dadurch aus, dass sie – bedingt durch die Bodenebene, auf der sie sich bewegen – eine weitestgehend planare Bewegung ausführen. Wie sich später noch zeigen wird, hängt der generelle Erfolg des Einsatzes der Hand-Augen-Kalibrierung jedoch davon ab, dass sich die Kameraplatform nicht nur um eine Rotationsachse bewegt, sondern um mehrere, um insbesondere den Höhenunterschied zwischen den Kameras bestimmen zu können. Planare Bewegungen besitzen aber nur eine Rotationsachse. Aus diesem Grund wird in der vorliegenden Arbeit zusätzlich zum allgemeinen Kalibrierprinzip ein erweitertes Modell vorgestellt und untersucht. Dieses Modell berücksichtigt explizit die Grundebene, auf der sich die Plattform bewegt, um dadurch möglichst viele Freiheitsgrade schätzen zu können.

**Lösungsansatz bei unbekannten Geschwindigkeitsbeträgen: Der Bewegungsausgleich** Im Zuge der Implementierung und Auswertung des Verfahrens hat sich gezeigt, dass die Bewegungsschätzung der einzelnen Kameras zwar in der Bestimmung der rotatorischen Bewegung sowie der Bewegungsrichtung brauchbare Ergebnisse lieferten. Der Geschwindigkeitsbetrag der einzelnen Kameras lieferte in Experimenten mit Realdaten vor allem bei längeren Beobachtungszeiträumen starke Abweichungen von den Sollwerten. Gerade dieser Geschwindigkeitsbetrag ist aber von großer Bedeutung,

um gute Kalibrierergebnisse zu erzielen.

Um dennoch mit den geschätzten Bewegungsdaten eine Kalibrierung durchführen zu können, wurde ein Verfahren zur simultanen Schätzung der relativen Beträge der zeitlichen, translatorischen Verschiebung der Kameras sowie deren extrinsischen Parameter entwickelt und eingesetzt. Die Parameterbestimmung lässt sich mit nichtlinearen Optimierungsmethoden lösen. Durch eine geschickte Implementierung, angelehnt an den sog. dünnen Bündelausgleich, lässt sich das Problem zudem recheneffizient lösen. Das Verfahren wird daher auch *dünner Bewegungsausgleich* oder *SMA* genannt (engl.: *Sparse Motion Adjustment*).

Darüber hinaus steht das hier vorgestellte Lösungskonzept im Fokus dreier Gesichtspunkte:

- **Fusionierbarkeit** - Wie nutzt man die Mehrinformation bedingt durch den Einsatz mehrerer Sensoren geschickt aus, um das Gesamtergebnis zu verbessern?
- **Skalierbarkeit** - Wie hält man den Rechenaufwand in Abhängigkeit zur Anzahl der Kameras in einem vernünftigen Rahmen? Die Parallelisierbarkeit des Problems soll gewährleistet bleiben. Ebenfalls soll die Datenrate zwischen den einzelnen Kameramodulen gering gehalten werden.
- **Kontinuität** - Die Schätzung der Parameter soll fortwährend erfolgen.

**Kontinuierliche Parameterschätzung** Die Kontinuität wird dadurch erreicht, dass die Zustände auf Basis der gemessenen Sensordaten zeitabhängig geschätzt werden. Von Bedeutung ist insbesondere die Schätzung der Unsicherheiten der Transformationsparameter. Durch die fortlaufende Schätzung ist das System in der Lage, sich selbst zu rekalibrieren. Die nichtlineare Filterung erfolgt durch ein iteratives erweitertes Kalmanfilter (IEKF).

**Skalierbarkeit und modulares Kalibriermodell** Eine ganzheitliche Schätzung über alle Parameter – also Bewegungs- und Kalibrierparameter – bringt den Nachteil mit

sich, dass der Zustandsvektor linear mit der Anzahl der Kameras wächst. Für ein Sensorsystem mit  $N$  Kameras müssen  $N$  Bewegungstransformationen und  $N$  extrinsische Transformationen geschätzt werden<sup>8</sup>. Die gesamtheitliche Verwendung aller Sensordaten in einem Schätzschritt führt zwangsläufig zu einem erhöhten Rechenaufwand. Aus diesem Grund werden die Bewegungs- und die Kalibrierparameter nacheinander und zunächst unabhängig von jedem Einzelmodul separat (d. h. lokal) berechnet. Jedes Kameramodul  $M$  agiert als eigenständige Rechen- und Auswerteeinheit [Pag10a]. Deshalb wird in der vorliegenden Arbeit eine Kamera häufig auch schlicht als *Modul* bezeichnet.

**Globales Fusionsmodell zur ganzheitlichen Informationsnutzung** Die Informationsfusion setzt – wie die Skalierbarkeit – auf einen modularen Systementwurf. In dem hier untersuchten Multisensorsystem häuft sich eine Redundanz an Informationen an, die sinnvollerweise effektiv genutzt werden soll (man vergleiche auch Abb. 1.7). Diese Redundanz röhrt daher, dass jedes der  $N$  Kameramodule seinen eigenen Systemzustand schätzt. Durch die Fusion redundanter Schätzungen können einzelne Fehlschätzungen abgefangen und die Unsicherheit des Gesamtsystems verringert werden. Im vorliegenden Fall lassen sich sowohl die Bewegungsschätzungen der Module, die die Grundlage für die eigentliche Kalibrierung darstellen und daher von besonderer Wichtigkeit sind, als auch die gesuchten extrinsischen Parameter fusionieren. Die Fusion überführt die *lokalen* Zustände der Einzelmodule in *globale* Zustandsparameter. Im *globalen Zustand* eines Moduls sind die Ergebnisse aller lokalen Einzelschätzungen berücksichtigt. Im Gegensatz hierzu ist ein Zustand *lokal* bezüglich eines Moduls, wenn zur Schätzung des Zustands nur die lokalen Sensordaten des entsprechenden Moduls verwendet und ausgewertet wurden, unabhängig von der Information der anderen Module.

Für die Zusammenführung der lokalen Schätzungen kann man sich des Konzepts der

---

<sup>8</sup> Die  $N$  extrinsischen Transformationen ergeben sich z. B. aus einer ringförmigen oder sternförmigen Topologie. Eine vollständige, redundante extrinsische Vernetzung zwischen allen Kameras bestünde aus  $\frac{N(N-1)}{2}$  Transformationen.

Gauß'schen Fehlerpropagierung (GFP) und Fusion (GF) bedienen. Das Konzept des Zusammenführens räumlicher Transformationen und deren Unsicherheit wurde bereits von Smith und Cheeseman [Smi86] vorgestellt und kann für das vorliegende Problem adaptiert werden. Eine wesentliche Bedeutung spielen hierbei die Unsicherheiten der Transformationszustände. Propagierung bedeutet, dass man unter Kenntnis der Kalibrierparameter und -unsicherheiten die Zustände der anderen Module schätzt. So ist ein Modul entsprechend der propagierten Unsicherheiten der Zustände in der Lage, mehr seiner eigenen oder eher der Schätzung der anderen Module zu vertrauen.

Weiteres Augenmerk fällt hinsichtlich der Fusion, Modularität und Skalabilität auf die Kommunikation zwischen den Modulen. Um die Teilergebnisse der Sensoren zum Zwecke der Fusion zusammenführen zu können, muss daher eine Kommunikation zwischen den Sensoren bestehen. In der Praxis ist es aufgrund begrenzter Rechenressourcen und -kapazitäten oftmals von Bedeutung, den Kommunikationsaufwand bzw. die Kommunikationsbandbreite gering zu halten. Eine Fusion auf Datenebene – in diesem Fall Bilder, Merkmals- oder Verschiebungsvektoren – würde sowohl den Rechenaufwand als auch die Bandbreite unnötig stark beanspruchen. Dies motiviert einen Fusionsentwurf auf Zustandsebene.

**Zusammenfassung** Bei den zu schätzenden Zuständen handelt es sich in dieser Arbeit um Euklidische Transformationsparameter, bestehend aus drei Rotations- und drei Translationsparametern. Die gemessenen Daten, die verwendet werden, um die Zustände zu schätzen, sind entweder Bildmerkmale, Szenenpunkte oder andere Transformationsparameter. Diese Messungen werden von jedem Modul lokal, also separat erzeugt, so dass jedes Modul seine eigenen Messdaten besitzt. Mit Hilfe dieser Messdaten ist jedes Modul in der Lage, die gesuchten Zustände zu schätzen. Es handelt sich dabei um eine lokale Schätzung. D. h. von ein und demselben Zustand gibt es nun  $N$  unterschiedliche, voneinander unabhängige Schätzungen, nämlich die der  $N$  Module. Führt man diese Information, also die  $N$  separaten Zustandsschätzungen, in geeigneter Weise zusammen, so erhält man eine globale Zustandsschätzung – global deshalb, da die lokalen Informationen aller Module in der neuen Schätzung zusammenfließen.

Der Algorithmus zur Kalibrierung eines  $N$ -Kamerasytems arbeitet wie folgt (s. Abb. 1.7):

Zunächst schätzt jedes Kameramodul anhand der (eigenen) Sensordaten die Eigenbewegung mittels Bündelausgleich. Anhand der Kalibrierparameter und der lokal geschätzten Bewegungsparameter schätzt jedes Modul die Bewegungszustände der anderen Module mittels Zustands- und Fehlerfortpflanzung<sup>9</sup>. Zu diesem Zeitpunkt hat jedes Modul eine lokale Schätzung der Bewegung jedes anderen Moduls. D. h. es existieren für die Bewegung einer Kamera  $N$  lokale Schätzungen. Diese Parameter und ihre Unsicherheiten werden zwischen den einzelnen Modulen ausgetauscht und fusioniert, so dass nun für jedes der  $N$  Module eine robustere, globale Schätzung vorliegt. Anschließend werden die globalen Bewegungszustände von jedem Modul benutzt, um die extrinsischen Transformationsparameter zu den  $N - 1$  übrigen Kameras zu schätzen. Dies geschieht lokal basierend auf dem Prinzip der Hand-Augen-Kalibrierung. Der Bewegungsausgleich schätzt hierbei explizit die relativen Beträge der translatorischen Bewegung der einzelnen Module mit. Die extrinsischen Parameter werden mit Hilfe eines erweiterten Kalmanfilters fortlaufend prädiziert und aktualisiert. Sind die Transformationen zu den  $N - 1$  übrigen Kameras bekannt, so lässt sich mittels Propagierung ein vollständiger Transformationsgraph berechnen, so dass jedem Modul eine vollständige Schätzung aller extrinsischen Parameter zwischen allen Modulen vorliegt. Durch Austauschen der lokalen Kalibrierparameter und anschließender Fusion erhält man die globale Schätzung der extrinsischen Parameter, also die Anordnung der Kameras unter Berücksichtigung aller lokalen Schätzergebnisse und damit der Information aller Kamerasensoren.

Das Vorgehen gemäß

1. Lokale Schätzung
2. Propagierung

---

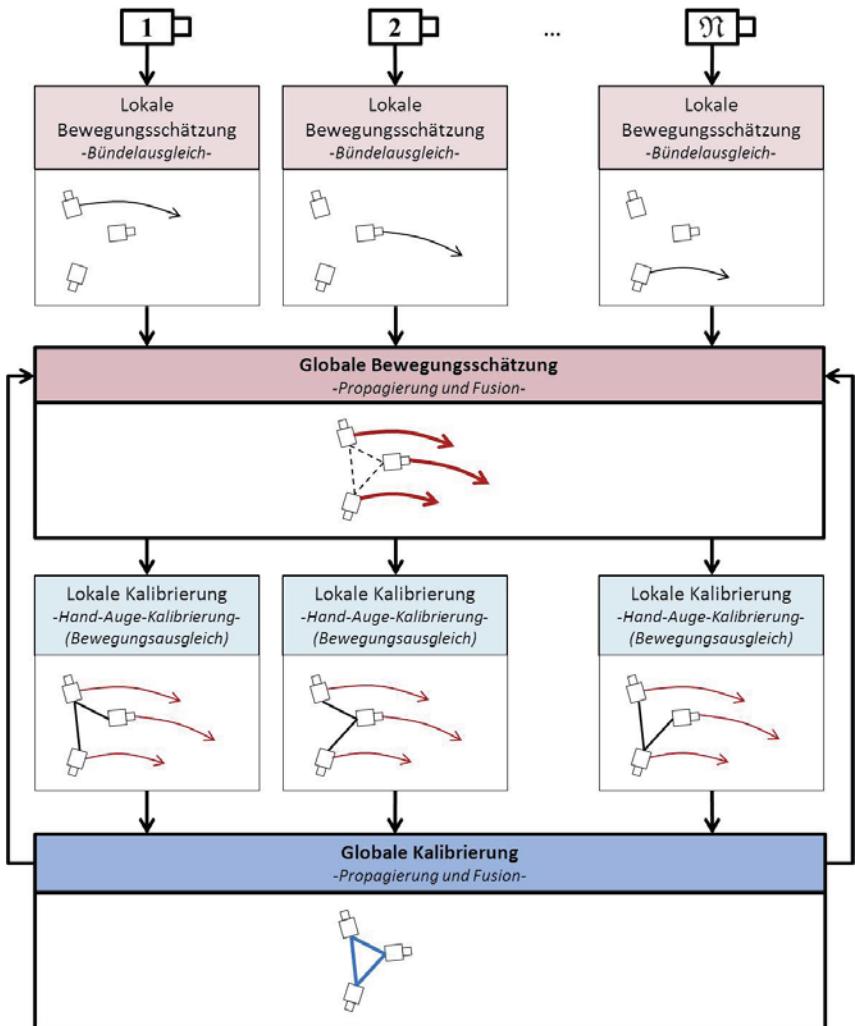
<sup>9</sup> Sind die Kalibrierparameter initial unbekannt, wird auch die Unsicherheit entsprechend hoch gewählt, so dass die propagierten Zustände bei der Fusion nur gering ins Gewicht fallen.

### 3. Fusion

kann sowohl zur Schätzung der Bewegungs- als auch der Kalibrierparameter angewendet werden. Daher wird ein Schätzzyklus unterteilt in einen Bewegungsschätzungs- und einen Kalibrierparameterschätzzyklus (s. Abb. 1.7).

Die hier beschriebene Fusion ist zweigeteilt in eine Fusion der Bewegungszustände und in eine Fusion der Kalibrierzustände. Kennt man die globale Kalibrierung des Systems, lassen sich die Modulbewegungen fusionieren. Sind die globalen Bewegungen aller Module bekannt, so kann man die Kalibrierparameter lokal schätzen und global fusionieren. Aus dieser globalen Sichtweise heraus gestaltet sich das hier vorgestellte Fusionskonzept wie ein »Henne-Ei-Problem«. Ist die Kalibrierung unbekannt, lässt sich keine Bewegungsfusion durchführen. Umgekehrt kann man ohne Kenntnis der Kamerabewegungen keine Kalibrierparameter schätzen oder fusionieren.

An dieser Stelle wird die Bedeutung der Unsicherheiten der Systemparameter ersichtlich. Die Modellierung der Unsicherheiten in Form von Kovarianzmatrizen umgeht dieses Problem. Durch die Verwendung von Kovarianzmatrizen lässt sich selbst eine völlige Unkenntnis mathematisch modellieren und in das Fusionsschema integrieren.



**Abbildung 1.7:** Schematische Darstellung des globalen, bewegungsbasierten Lösungsansatzes zur Kalibrierung von  $\mathfrak{N}$  Kameramodulen. Von oben nach unten: Bildeinzug und Merkmalsextraktion; Lokale Bewegungsschätzung mittels Bündelausgleich; Propagierung mittels der globalen Kalibrierparameter und anschließende Fusion; Lokale extrinsische Kalibrierung je Kameramodul mittels Hand-Augen-Kalibrierung, basierend auf den globalen Bewegungsschätzungen; Propagierung und Fusion der lokalen Schätzungen.

## 1.4 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich wie folgt:

Zunächst erfolgt eine Vorstellung der der Kalibrierung zugrundeliegenden algorithmischen Grundlagen (Kapitel 2). Der Bewegungsschätzung monokularer Kameras (Kapitel 3) sowie der Hand-Auge-Kalibrierung (Kapitel 4) wird wegen ihrer großen Bedeutung für das vorgestellte Kalibrierverfahren jeweils ein eigenes Kapitel gewidmet. Aufbauend auf den Grundlagen der Hand-Auge-Kalibrierung wird in Kapitel 4 das allgemeine Kalibriermodell für allgemeine Bewegungen eingeführt.

Im folgenden Kapiteln wird das erweiterte Kalibriermodell für planare Bewegungen vorgestellt (Kapitel 5). Da der Anwendungsfokus auf Multikamerasystemen in Straßenfahrzeugen liegt, beziehen sich die weiteren Ausführungen auf das erweiterte Modell. Der Bewegungsausgleich (Kapitel 6) sowie die kontinuierliche Zustandsfilterung (Kapitel 7) werden in den folgenden beiden Kapiteln behandelt. Die Vorstellung des ganzheitlichen Fusionsansatzes erfolgt in Kapitel 8.

Kapitel 9 präsentiert die Experimente und deren Auswertung. Eine abschließende Zusammenfassung und ein Ausblick auf mögliche Weiterentwicklungen runden die Arbeit ab (Kapitel 10).



# 2

---

## Allgemeine Grundlagen

---

### 2.1 Das Kameramodell

Für die in dieser Arbeit berücksichtigten Kameras wird die Abbildungsvorschrift des Lochkameramodells verwendet (s. Abb. 2.1). Das bedeutet, dass eine Kamera  $\mathcal{M}$  einen Szenenpunkt  $\mathbf{X} \in \mathbb{R}^3$  in Abhängigkeit der intrinsischen und der extrinsischen Parameter  $\mathbf{K}$  bzw.  $\mathbf{R}, \mathbf{t}$  auf den Bildpunkt  $\mathbf{x} \in \mathbb{R}^2$  abbildet gemäß

$$\tilde{\mathbf{x}} = \frac{1}{\lambda} \Pi \tilde{\mathbf{X}} \quad (2.1)$$

mit den homogenen Punktkoordinaten

$$\begin{aligned} \tilde{\mathbf{x}} &= [x, y, 1]^T \\ \tilde{\mathbf{X}} &= [X, Y, Z, 1]^T \end{aligned} \quad (2.2)$$

und der Projektionsmatrix

$$\begin{aligned}\Pi_{3 \times 4} &= \mathbf{K}[\mathbf{R}|\mathbf{t}] \\ &= \begin{bmatrix} \Pi_1^T \\ \Pi_2^T \\ \Pi_3^T \end{bmatrix},\end{aligned}\tag{2.3}$$

wobei  $\Pi_n^T$  die  $n$ -te Zeile von  $\Pi$  ist. Für den Skalierungsfaktor  $\lambda$  gilt  $\lambda = \Pi_3^T \tilde{\mathbf{X}}$ . Der Bildpunkt  $\mathbf{x}$  ist der Schnittpunkt der Bildebene  $\mathcal{I}$  mit der Sichtgeraden zwischen dem Szenenpunkt  $\mathbf{X}$  und dem Kameraursprung bzw. Projektionszentrum  $\mathbf{o}$ . Die Matrix  $\Pi \in \mathbb{R}^{3 \times 4}$  wird fortan als Kameramatrix bezeichnet, wobei  $\mathbf{R}$  eine  $3 \times 3$ -Rotationsmatrix ist, mit  $\mathbf{R}^T = \mathbf{R}^{-1}$  und  $\det(\mathbf{R}) = 1$ . Die Rotationsmatrix folgt hierbei der Konvention

$$\begin{aligned}\mathbf{R}(r_x, r_y, r_z) &= \mathbf{R}(0, 0, r_z) \mathbf{R}(0, r_y, 0) \mathbf{R}(r_x, 0, 0) \\ &= \begin{bmatrix} \cos(r_z) & -\sin(r_z) & 0 \\ \sin(r_z) & \cos(r_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} \cos(r_y) & 0 & \sin(r_y) \\ 0 & 1 & 0 \\ -\sin(r_y) & 0 & \cos(r_y) \end{bmatrix} \cdot \\ &\quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(r_x) & -\sin(r_x) \\ 0 & \sin(r_x) & \cos(r_x) \end{bmatrix}.\end{aligned}\tag{2.4}$$

Alternativ werden die Eulerwinkel auch in kompakterer Weise im Parametervektor  $\mathbf{r} = [r_x, r_y, r_z]^T$  dargestellt. Weitere Darstellungsarten von Rotationen finden sich in Anhang A.1.

Die intrinsischen Parameter sind durch die Kameramatrix

$$\mathbf{K} = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2.5)$$

gegeben, bestehend aus der Brennweite  $f$  und dem Durchstoßpunkt  $\mathbf{c} = [c_x, c_y]$ <sup>10</sup>.  $\mathbf{K}$  wird auch als intrinsische Matrix bezeichnet.

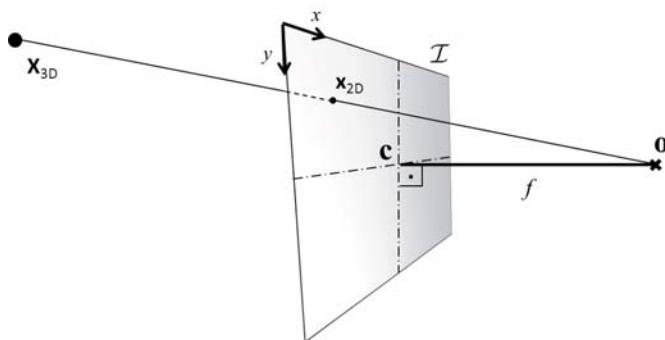


Abbildung 2.1: Lochkameramodell mit den intrinsischen Kameraparametern.

Die Projektion eines Szenenpunktes  $\mathbf{X}$  auf den Bildpunkt  $\mathbf{x}$  aus Gl. 2.1 wird in Abhängigkeit von den intrinsischen und extrinsischen Parameter durch die Abbildung  $\pi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  beschrieben, mit

$$\mathbf{x} = \pi(\mathbf{X}; \Pi) \text{ oder } \mathbf{x} = \pi(\mathbf{X}; \mathbf{K}, \mathbf{R}, \mathbf{t}). \quad (2.6)$$

Allgemeine Transformationen zwischen Koordinatensystemen sind durch Euklidische Transformationen modelliert. Diese setzen sich zusammen aus einer Rotationsmatrix  $\mathbf{R}$

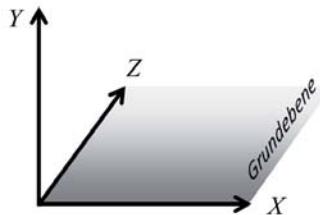
---

<sup>10</sup> Die Verzerrungseffekte des Objektivs [Bro86] werden in der vorliegenden Arbeit allerdings vernachlässigt. Diese Annahme ist durchaus gerechtfertigt, da sich mit Standardkalibrierverfahren die Brennweite und Linsenverzerrungsparameter bereits im Vorfeld bestimmen lassen, z. B. [Zha00]. Die akquirierten Kamerabilder lassen sich somit bereits in einem Vorverarbeitungsschritt entzerrn.

und einem Translationsvektor  $\mathbf{t}$ . D. h. auch die gesuchte extrinsische Kameraanordnung lässt sich vollständig mit Euklidischen Transformationen beschreiben. Eine extrinsische, homogene Euklidische Transformation, die das Koordinatensystem einer Kamera in das einer anderen Kamera überführt, hat die Form

$$\mathbf{C} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}. \quad (2.7)$$

Alle Transformationen erfolgen innerhalb eines linkshändigen Koordinatensystems (s. Abb. 2.2).



**Abbildung 2.2:** Koordinatensystem mit Grundebene.

Transformationen in eine zweidimensionale Ebene, wie sie an späterer Stelle relevant sind, werden mit  $\mathbf{G}$  bezeichnet:

$$\mathbf{G} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}. \quad (2.8)$$

Für spätere Rechenoperationen und Schätzungen ist es wichtig, die Transformationsmatrix in Form eines Zustandsvektors kompakter darzustellen. Der Zustandsvektor einer Transformationsmatrix lässt sich somit darstellen als

$$\mathbf{x}_C = \zeta_C(C) = [\mathbf{t}^T, \mathbf{r}^T]^T, \quad (2.9)$$

wobei die Funktion  $\zeta_C : \mathbb{R}^{4 \times 4} \rightarrow \mathbb{R}^6$  die homogene Transformationsmatrix  $C$  auf den

Zustandsvektor  $\mathbf{x}_G$  abbildet. Da für die Transformation eines Kamerakoordinatensystems in eine Ebene nur drei Freiheitsgrade notwendig sind, nämlich zwei rotatorische um die  $X$ - und  $Z$ -Achse sowie ein translatorischer entlang der  $Y$ -Achse, verhält sich die Abbildung der homogenen Matrix auf den Zustandsvektor entsprechend gemäß

$$\mathbf{x}_G = [h_y, \alpha, \gamma]^T . \quad (2.10)$$

Extrinsische Transformationen innerhalb einer 2D-Ebene, genauer der  $XZ$ -Ebene des Weltkoordinatensystems, seien des weiteren mit  $\underline{C}$  und dem Parametervektor  $\mathbf{x}_{\underline{C}} = [t_x, t_z, \beta]^T$  bezeichnet.

## 2.2 Das Bewegungsmodell und die nichtholome Bewegungsbedingung

Das Bewegungsmodell beschreibt zeitliche Transformationen eines Kamerakoordinatensystems zwischen zwei Zeitpunkten  $k, k'$  mit  $|k' - k| > 0$ . Wenngleich es sich bei diesen Transformationen – wie bei den extrinsischen – um Euklidische Transformationen handelt, so ist es für den weiteren Verlauf und eine bessere Übersicht lohnenswert, diese bereits in der Notation zu unterscheiden.

Die Bewegung einer Kamera wird analog zu Gl. 2.7 durch die homogene Matrix

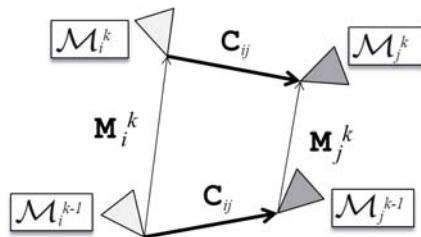
$$\mathbf{M} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{v} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (2.11)$$

beschrieben, wobei  $\mathbf{\Omega} \in \mathbb{R}^{3 \times 3}$  den rotatorischen und  $\mathbf{v} = [v_x, v_y, v_z]^T$  den translatorischen Anteil der Euklidischen Bewegung bezeichnet. Der Betrag der translatorischen Bewegung  $v$  entspricht der Euklidischen Norm von  $\mathbf{v}$ . Die entsprechende Geschwindigkeit  $\frac{v}{\Delta t}$  ergibt sich aus dem Betrag der translatorischen Bewegung relativ zur vergangenen Zeit  $\Delta t$  zwischen den Zeitpunkten  $k$  und  $k'$ . Um Verwirrungen vorzubeugen sei in der vorliegenden Arbeit  $\Delta t := 1$  definiert, so dass der Absolutbetrag der Geschwindigkeit

keit einer Kamera dem Betrag  $v$  des Translationsvektors der Bewegungstransformation entspricht:

$$v = ||\mathbf{v}|| = \frac{v}{\Delta t}. \quad (2.12)$$

Der Zusammenhang zwischen den räumlichen und zeitlichen Transformationen eines Mehrkamerasystems ist in Abb. 2.3 illustriert.



**Abbildung 2.3:** Grundlegende räumlich-zeitliche Anordnung für zwei Kameras  $\mathcal{M}_i$  und  $\mathcal{M}_j$  zu den Zeitpunkten  $k$  und  $k + 1$ .

Analog zu den Gl. 2.9 und 2.10 ergibt sich der Zustandsvektor aus der Transformationsmatrix gemäß

$$\mathbf{x}_{\mathbb{M}} = \zeta_{\mathbb{M}}(\mathbb{M}) = [\mathbf{v}, \boldsymbol{\omega}]^T \quad (2.13)$$

mit  $\boldsymbol{\omega} = [\omega_x, \omega_y, \omega_z]^T$ . Bewegungstransformationen innerhalb der  $XZ$ -Ebene seien mit  $\underline{\mathbb{M}}$  bezeichnet. Der zugehörige Parametervektor hat die Gestalt  $\underline{\mathbf{x}}_{\underline{\mathbb{M}}} = [\underline{v}_x, \underline{v}_z, \psi]^T$ .

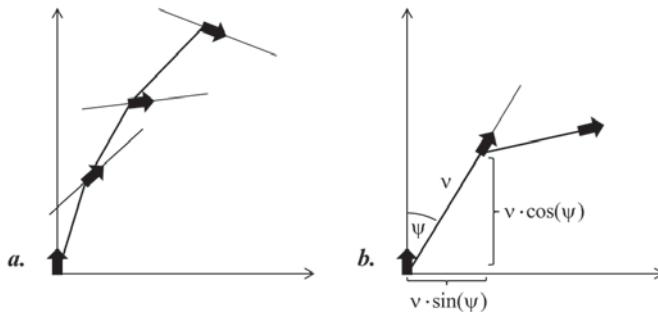
Über eine Vorderachse gelenkte mobile Plattform gehorcht i. A. der einer nichtholonomen Bewegungsbedingung der Art, dass eine rotatorische Bewegung der Plattform nur dann möglich ist, wenn sich die Plattform auch translatorisch bewegt [Zom87], [Gol78]. Diese Bedingung besagt, dass die Geschwindigkeit senkrecht zur Fahrtrichtung direkt vom Betrag der Geschwindigkeit in Fahrtrichtung abhängt. Eine Nichtbeachtung dieser Bedingung ließe also zu, dass sich Fahrzeuge im Stillstand um ihre vertikale Achse drehen. Dies ist zwar bei Landrobotern durchaus möglich, stellt bei kommerziellen

Straßenfahrzeugen aber i. d. R. eine unsinnige Annahme dar<sup>11</sup>.

Die Berücksichtigung dieser Eigenschaft ist für die spätere Implementierung des Lösungsansatzes erforderlich. Die Bewegung von Plattformen, die der nichtholonen Bewegungsbedingung gehorchen, lässt sich z. B. mit Hilfe des so genannten Einspurmodells modellieren [Leu00], [Haa98]. Für die Bewegung eines Fahrzeugs in der  $XZ$ -Grundebene mit der Winkelgeschwindigkeit oder Gierrate  $\psi$  bedeutet dies, dass sich die translatorische Verschiebung in Richtung der Fahrzeulgängsachse gemäß

$$\mathbf{v} = v \cdot \begin{bmatrix} \sin(\psi) \\ 0 \\ \cos(\psi) \end{bmatrix} \quad (2.14)$$

verhält, wobei  $v$  den Betrag des translatorischen Verschiebungsvektors zwischen zwei Zeitpunkten, bezeichnet (s. Abb. 2.4).



**Abbildung 2.4:** Beispiele zweier Bewegungen eines Fahrzeugs (als Pfeil dargestellt). Links: Bewegung eines Fahrzeugs unter Nichteinhaltung der nichtholonen Bewegungsbedingung. Das Fahrzeug (über-)dreht sich um die eigene Achse. Rechts: Bewegung unter Beachtung der nichtholonen Bewegungsbedingung.  $v$  bezeichnet den Betrag der translatorischen Verschiebung.

<sup>11</sup> Rutschen und Gleiten eines Fahrzeugs sind Spezialfälle, die in der vorliegenden Arbeit nicht adressiert werden.

## 2.3 Nichtlineare Levenberg-Marquardt-Optimierung

Der Levenberg-Marquardt-Algorithmus (LM-Algorithmus) [Lev44], [Mar63] ist ein iteratives Verfahren zum Finden lokaler Minima in nichtlinearen Funktionen mit Hilfe der Methode der kleinsten Fehlerquadrate. Es handelt sich hierbei um eine Variante des Gauss-Newton-Verfahrens und ist Grundlage vieler in der vorliegenden Arbeit durchgeführten Schätzprozesse. Im Folgenden wird das LM-Verfahren und seine Funktionsweise kurz beschrieben. Eine ausführlichere Beschreibung findet sich in [Pre07].

Sei  $\mathbf{h}(\mathbf{x}, \mathbf{z})$  eine nichtlineare Residuumsfunktion mit Parametervektor  $\mathbf{x} \in \mathbb{R}^N$  und Messung  $\mathbf{z} \in \mathbb{R}^M$ , die es in Abhängigkeit von  $\mathbf{x}$  zu minimieren gilt. Bei gegebenem Startwert  $\mathbf{x}_0$  ist der Vektor  $\boldsymbol{\delta}$  gesucht, so dass gilt

$$\mathbf{h}(\mathbf{x}, \mathbf{z}) = \mathbf{h}(\mathbf{x}_0 + \boldsymbol{\delta}, \mathbf{z}) \stackrel{!}{=} \mathbf{0}. \quad (2.15)$$

Mit Hilfe einer Taylor-Entwicklung bis zum ersten Glied erhält man

$$\mathbf{h}(\mathbf{x}_0 + \boldsymbol{\delta}, \mathbf{z}) \approx \underbrace{\mathbf{h}(\mathbf{x}_0, \mathbf{z})}_{=\boldsymbol{\epsilon}} - \mathbf{J} \underbrace{[\mathbf{x}_0 - \mathbf{x}]}_{=\boldsymbol{\delta}} \quad (2.16)$$

mit dem Residuum  $\boldsymbol{\epsilon}$  und  $\mathbf{J} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_0, \mathbf{z}}$ . Durch Minimieren des Funktional

$$\|\mathbf{h}(\mathbf{x}_0 + \boldsymbol{\delta}, \mathbf{z})\|^2 = [\boldsymbol{\epsilon} - \mathbf{J}\boldsymbol{\delta}]^2 \stackrel{!}{\rightarrow} \min \quad (2.17)$$

und Ableiten nach  $\boldsymbol{\delta}$  folgt

$$2\mathbf{J}^T [\boldsymbol{\epsilon} - \mathbf{J}\boldsymbol{\delta}] \stackrel{!}{=} 0, \quad (2.18)$$

woraus sich die sog. *Normalengleichung*

$$\mathbf{J}^T \mathbf{J} \boldsymbol{\delta} = \mathbf{J}^T \boldsymbol{\epsilon} \quad (2.19)$$

ergibt.

Die Bezeichnung Normalengleichung röhrt daher, dass Gl. 2.19 das Minimum erreicht, wenn der Vektor  $\mathbf{J}\boldsymbol{\delta} - \boldsymbol{\epsilon}$  orthogonal zu dem durch die Spaltenvektoren von  $\mathbf{J}$  aufgespannten Vektorraum steht, so dass  $\mathbf{J}^T[\mathbf{J}\boldsymbol{\delta} - \boldsymbol{\epsilon}] = \mathbf{0}$  gilt. Das LM-Verfahren erweitert Gl. 2.19 um einen Summanden zur *erweiterten Normalengleichung*

$$[\mathbf{J}^T\mathbf{J} + \rho \cdot \mathbf{I}]\boldsymbol{\delta} = \mathbf{J}^T\boldsymbol{\epsilon}, \rho > 0. \quad (2.20)$$

Falls der aktualisierte Parametervektor  $\mathbf{x}_0 + \boldsymbol{\delta}$  eine Verbesserung bringt bzw. der quadratische Fehler  $\boldsymbol{\epsilon}^T\boldsymbol{\epsilon}$  kleiner wird, so wird der neue Parametervektor beibehalten,  $\rho$  verkleinert und die Iteration fortgesetzt. Ansonsten wird  $\rho$  vergrößert. Gl. 2.20 wird solange wiederholt gelöst, bis ein Wert für  $\boldsymbol{\delta}$  gefunden wurde, der das Residuum verkleinert.

Für große Werte von  $\rho$  verhält sich das LM-Verfahren wie ein Gradientenabstieg in Richtung  $\mathbf{J}^T\boldsymbol{\epsilon}$  mit einem kleinen Betrag von  $\boldsymbol{\delta}$ . Umgekehrt verhält sich der Algorithmus für kleine  $\rho$  wie ein Gauss-Newton-Verfahren und konvergiert in diesem Fall schneller zum Minimum.

Die LM-Iterationen können abgebrochen werden, falls die maximale Anzahl an Iterationen erreicht wurde oder der Betrag von  $\boldsymbol{\delta}$  unter einen Schwellwert fällt.

## 2.4 Das iterative erweiterte Kalmanfilter

Dieser Abschnitt beschreibt das iterierte erweiterte Kalmanfilter (IEKF), wie es auch in [Sim06] beschrieben wird. Das Kalmanfilter von seiner klassischen linearen Form [Kal60] bis hin zu seinen nichtlinearen Adaptionen [Jul97], [Sor85] ist eines der am häufigsten eingesetzten Werkzeuge zur Zustandsfilterung. Das erweiterte Kalmanfilter ist ein nichtlinearer Schätzer zur temporären Filterung eines Systemzustandes mittels rauschbehafteter Messungen. Die Unsicherheit sowohl der Messungen als auch des Zustands werden als normalverteilt angenommen, man spricht in diesem Fall von

System- und Messrauschen.

Die Kalmangleichungen leiten sich aus dem Bayesschen Theorem

$$p(\mathbf{x}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{z})} \quad (2.21)$$

ab, mit dem Ziel, die Wahrscheinlichkeit  $p(\mathbf{x}|\mathbf{z})$ , dass sich das System bei der Beobachtung  $\mathbf{z}$  im Zustand  $\mathbf{x}$  befindet, zu maximieren.

Das Kalmanfilter führt im Wesentlichen zwei Schritte aus: Eine Prädiktion des Zustands für den Folgezeitpunkt anhand eines Systemmodells und einen Aktualisierungsschritt auf der Grundlage von rauschbehafteten Messungen. In der vorliegenden Arbeit liegt die große Bedeutung des Kalmanfilters insbesondere in seiner Eigenschaft begründet, zu jedem Zustand eine Unsicherheit in Form einer Kovarianzmatrix zu schätzen. Nicht-lineare Weiterentwicklungen des klassischen linearen Kalmanfilters sind das erweiterte Kalmanfilter (EKF) und das iterative erweiterte Kalmanfilter (IEKF). Das EKF resultiert aus dem linearen Kalmanfilter durch eine Linearisierung der Bedingungsfunktion mittels einer Taylorentwicklung bis zum ersten Glied. Durch diesen Linearisierungsschritt und das Verwerfen der Restglieder entsteht ein Fehler, dem man durch wiederholte Linearisierungen in den Iterationsschritten beizukommen versucht. Das IEKF, wie es in der vorliegenden Arbeit zur Schätzung der Kovarianzmatrizen und Zustandsfilterung eingesetzt wird, wird im Folgenden erläutert. Für eine detaillierte Herleitung des erweiterten Kalmanfilters sei an dieser Stelle an [Sim06] verwiesen.

Das Zustands- und das Messmodell zum Zeitpunkt  $k$  sei gegeben durch

$$\hat{\mathbf{x}}_k = \mathbf{f}(\hat{\mathbf{x}}_{k-1}, \mathbf{w}_{k-1}) \quad \text{mit} \quad \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{w}_k}), \quad (2.22)$$

$$\hat{\mathbf{z}}_k = \mathbf{z}_k + \mathbf{e}_k \quad \text{mit} \quad \mathbf{e}_k \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{e}_k}). \quad (2.23)$$

Weiter sei die nichtlineare Bedingungsgleichung  $\mathbf{h}$

$$\mathbf{h}(\mathbf{x}, \mathbf{z}) = \mathbf{0} \quad (2.24)$$

gegeben.

Des weiteren gelte

$$\mathbb{E}(\mathbf{w}_m \mathbf{w}_n^T) = \mathbb{E}(\mathbf{e}_m \mathbf{e}_n^T) = \mathbf{0}, \quad \forall m, n \in \mathbb{N}, m \neq n \quad (2.25)$$

$$\mathbb{E}([\mathbf{x} - \hat{\mathbf{x}}][\mathbf{x} - \hat{\mathbf{x}}]^T) = \mathbf{P}, \quad (2.26)$$

wobei  $\mathbf{x}$  den wahren Zustand,  $\hat{\mathbf{x}}$  dessen Schätzung und  $\mathbf{P}$  die entsprechende Kovarianzmatrix bezeichne. Nach der Initialisierung von Zustand und Kovarianzmatrix

$$\hat{\mathbf{x}}_0^- = \mathbb{E}(\hat{\mathbf{x}}_0) \quad (2.27)$$

$$\mathbf{P}_0 = \mathbb{E}([\mathbf{x}_0 - \hat{\mathbf{x}}_0][\mathbf{x}_0 - \hat{\mathbf{x}}_0]^T) \quad (2.28)$$

erfolgt der Prädiktionsschritt für Zustand und Kovarianzmatrix mittels der nichtlinearen Funktion  $\mathbf{f}$  des Systemmodells gemäß

$$\hat{\mathbf{x}}_k^- = \mathbf{f}(\hat{\mathbf{x}}_{k-1}^+, \mathbf{w}_{k-1}) \quad (2.29)$$

$$\mathbf{P}_k^- = \mathbf{F}_{\mathbf{x}} \mathbf{P}_{k-1}^+ \mathbf{F}_{\mathbf{x}}^T + \Sigma_{\mathbf{w}_k} \quad (2.30)$$

mit

$$\mathbf{F}_{\mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}_{k-1}^+, \mathbf{w}_{k-1}}. \quad (2.31)$$

Die Iteration des Innovationsschritts wird nach der Initialisierung durch

$$\hat{\mathbf{x}}_{k,0} := \hat{\mathbf{x}}_k^- \quad (2.32)$$

$$\hat{\mathbf{z}}_{k,0} := \hat{\mathbf{z}}_k \quad (2.33)$$

wie folgt ausgeführt:

Für  $l = 0, \dots, L-1$ :

$$\hat{\mathbf{x}}_{k,l+1} = \hat{\mathbf{x}}_k^- - \mathbf{G}_{k,l} \mathbf{r}_{k,l} \quad (2.34)$$

$$\hat{\mathbf{z}}_{k,l+1} = \hat{\mathbf{z}}_k - \Sigma_{\mathbf{e}_k} \mathbf{H}_{\mathbf{z},l}^T [\mathbf{H}_{\mathbf{z},l} \Sigma_{\mathbf{e}_k} \mathbf{H}_{\mathbf{z},l}^T]^{-1} \mathbf{r}_{k,l} \quad (2.35)$$

mit

$$\mathbf{H}_{\mathbf{x},l} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} \Big|_{\hat{\mathbf{x}}_{k,l}, \hat{\mathbf{z}}_{k,l}} \quad (2.36)$$

$$\mathbf{H}_{\mathbf{z},l} = \frac{\partial \mathbf{h}}{\partial \mathbf{z}} \Big|_{\hat{\mathbf{x}}_{k,l}, \hat{\mathbf{z}}_{k,l}} \quad (2.37)$$

$$\mathbf{G}_{k,l} = \mathbf{P}_k^- \mathbf{H}_{\mathbf{x},l}^T [\mathbf{H}_{\mathbf{x},l} \mathbf{P}_k^- \mathbf{H}_{\mathbf{x},l}^T + \mathbf{H}_{\mathbf{z},l} \Sigma_{\mathbf{e}_k} \mathbf{H}_{\mathbf{z},l}^T]^{-1} \quad (2.38)$$

$$\mathbf{r}_{k,l} = \mathbf{h}(\hat{\mathbf{x}}_{k,l}, \hat{\mathbf{z}}_{k,l}) + \mathbf{H}_{\mathbf{x},l} \cdot [\hat{\mathbf{x}}_k^- - \hat{\mathbf{x}}_{k,l}] + \mathbf{H}_{\mathbf{z},l} \cdot [\hat{\mathbf{z}}_k - \hat{\mathbf{z}}_{k,l}] \quad (2.39)$$

Nach  $L$  Iterationen ergibt sich der aktualisierte Zustand und die Kovarianzmatrix gemäß

$$\hat{\mathbf{x}}_k^+ := \hat{\mathbf{x}}_{k,L} \quad (2.40)$$

$$\mathbf{P}_k^+ := [\mathbf{I} - \mathbf{G}_{k,L-1} \mathbf{H}_{\mathbf{x},L-1}] \mathbf{P}_k^- \quad (2.41)$$

$$= [\mathbf{I} - \mathbf{G}_{k,L-1} \mathbf{H}_{\mathbf{x},L-1}] \mathbf{P}_k^- [\mathbf{I} - \mathbf{G}_{k,L-1} \mathbf{H}_{\mathbf{x},L-1}]^T + \mathbf{G}_{k,L-1} \Sigma_{\mathbf{e}_k} \mathbf{G}_{k,L-1}^T. \quad (2.42)$$

Die sog. *Joseph Form* in Gl. 2.42 ist numerisch stabiler und garantiert symmetrische Kovarianzmatrizen, weshalb sie in der Praxis trotz höheren Rechenaufwandes Gl. 2.41 häufig vorgezogen wird.

## 2.5 Zustands- und Fehlerpropagierung

Dieser Abschnitt beschreibt das Prinzip und die Herleitung der Gauß'schen Fehlerpropagierung (GFP) [Hoe80].

Sei eine lineare Funktion  $\Phi : \mathbb{R}^N \rightarrow \mathbb{R}^M$  gegeben mit

$$\Phi(\mathbf{x}) = \mathbf{a} + \mathbf{A}\mathbf{x} = \mathbf{z}. \quad (2.43)$$

Hierbei wird angenommen, dass sowohl die Schätzung  $\hat{\mathbf{x}}$  des Zustands  $\mathbf{x}$  als auch die daraus resultierenden Funktionswerte  $\hat{\mathbf{z}}$  einer Normalverteilung folgen, also  $\hat{\mathbf{x}} \sim \mathcal{N}(\mathbf{x}, \Sigma_{\mathbf{x}})$  und  $\hat{\mathbf{z}} \sim \mathcal{N}(\mathbf{z}, \Sigma_{\mathbf{z}})$ , wobei die wahren Werte  $\mathbf{x}$  und  $\mathbf{z}$  die Erwartungswerte der Verteilungen sind. Daraus folgt für die Beobachtung  $\hat{\mathbf{z}}$  und deren Erwartungswert

$$\hat{\mathbf{z}} = \mathbf{a} + \mathbf{A}\hat{\mathbf{x}} \quad (2.44)$$

$$\mathbf{z} = \hat{\mathbf{z}} + \boldsymbol{\delta}_{\mathbf{z}} = \mathbf{a} + \mathbf{A}\mathbf{x} = \mathbf{a} + \mathbf{A}[\hat{\mathbf{x}} + \boldsymbol{\delta}_{\mathbf{x}}]. \quad (2.45)$$

Subtrahieren von  $\hat{\mathbf{z}}$  führt zu

$$\boldsymbol{\delta}_{\mathbf{z}} = \mathbf{A}\boldsymbol{\delta}_{\mathbf{x}}. \quad (2.46)$$

Da die resultierenden Fehler ebenfalls normalverteilt sind mit  $\boldsymbol{\delta}_{\mathbf{x}} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{x}})$  und  $\boldsymbol{\delta}_{\mathbf{z}} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{z}})$ , gilt ferner

$$\Sigma_{\mathbf{z}} = \mathbb{E}(\boldsymbol{\delta}_{\mathbf{z}}\boldsymbol{\delta}_{\mathbf{z}}^T) \stackrel{2.46}{=} \mathbb{E}(\mathbf{A}\boldsymbol{\delta}_{\mathbf{x}}\boldsymbol{\delta}_{\mathbf{x}}^T\mathbf{A}^T) = \mathbf{A}\mathbb{E}(\boldsymbol{\delta}_{\mathbf{x}}\boldsymbol{\delta}_{\mathbf{x}}^T)\mathbf{A}^T = \mathbf{A}\Sigma_{\mathbf{x}}\mathbf{A}^T. \quad (2.47)$$

Darauf aufbauend folgt die Herleitung für den nichtlinearen Fall. Sei nun eine nichtlineare Funktion  $\Phi': \mathbb{R}^{N \times M} \rightarrow \mathbb{R}^O$  gegeben mit

$$\Phi'(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M) = \mathbf{z} \quad (2.48)$$

und statistisch unabhängigen Zufallsvariablen  $\mathbf{x}_1, \dots, \mathbf{x}_M$ . Dann kann  $\Phi'$  durch eine Taylor-Reihe approximiert werden gemäß

$$\begin{aligned}
 \mathbf{z} &= \hat{\mathbf{z}} + \boldsymbol{\delta}_{\mathbf{z}} \\
 &= \Phi'(\hat{\mathbf{x}}_1 + \boldsymbol{\delta}_{\mathbf{x}_1}, \dots, \hat{\mathbf{x}}_M + \boldsymbol{\delta}_{\mathbf{x}_M}) \\
 &= \Phi'(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M) \\
 &\quad + \frac{\partial \Phi'}{\partial \mathbf{x}_1}(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M) \cdot \boldsymbol{\delta}_{\mathbf{x}_1} \\
 &\quad + \dots \\
 &\quad + \frac{\partial \Phi'}{\partial \mathbf{x}_M}(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M) \cdot \boldsymbol{\delta}_{\mathbf{x}_M} \\
 &\quad + G \\
 &= \Phi'(\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_M) + \mathbf{J}_1 \boldsymbol{\delta}_{\mathbf{x}_1} + \dots + \mathbf{J}_M \boldsymbol{\delta}_{\mathbf{x}_M} + G
 \end{aligned} \tag{2.49}$$

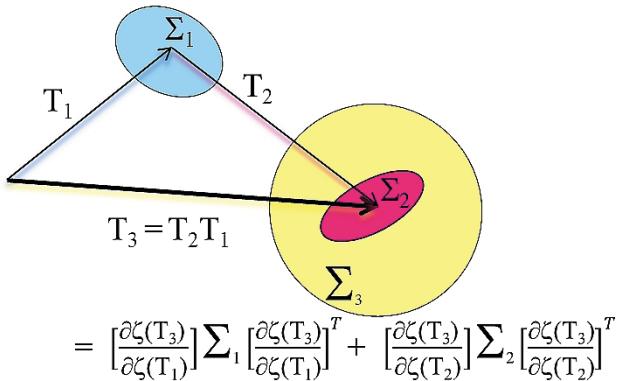
wobei  $\mathbf{J}_{\{1, \dots, M\}}$  die Jacobi-Matrizen und  $G$  das Restglied der Terme höherer Ordnung bezeichnen. Unter der Annahme, dass die Fehler  $\boldsymbol{\delta}_{\mathbf{x}_{\{1, \dots, M\}}}$  sehr klein sind, lässt sich die Taylor-Entwicklung nach dem ersten Glied abbrechen. Subtraktion von  $\hat{\mathbf{z}}$  führt zu

$$\boldsymbol{\delta}_{\mathbf{z}} \approx \sum_{i=1}^M \mathbf{J}_i \boldsymbol{\delta}_{\mathbf{x}_i}. \tag{2.50}$$

Analog zu (2.47) erhält man somit die Kovarianzmatrix von  $\mathbf{z}$  gemäß

$$\Sigma_{\mathbf{z}} = \sum_{i=1}^M \mathbf{J}_i \Sigma_{\mathbf{x}_i} \mathbf{J}_i^T. \tag{2.51}$$

In der vorliegenden Arbeit werden ausschließlich Euklidische Transformationen propagiert. Abb. 2.5(a) veranschaulicht die propagierte Unsicherheit bei der Verkettung von normalverteilten Transformationen.



**Abbildung 2.5:** Schema einer propagierten Transformation und deren Unsicherheit.  $T_3$  bzw.  $\Sigma_3$  lassen sich mit  $T_1, T_2$  und  $\Sigma_1, \Sigma_2$  bestimmen.

## 2.6 Gauß'sche Fusion

Die Fusion normalverteilter Zustandsvariablen kann auf vielfältige Weise erreicht werden. Schlosser [Sch05a] und Chong [Cho79] beispielsweise verknüpften mehrere Kalmanfilter hierarchisch. D. h. die Ergebnisse der Teilfilter wurden in einem separaten Filter zusammengeführt und gefiltert. Smith und Cheeseman [Smi86] stellten ein Schema zur paarweisen Fusion zweier normalverteilten Zustandsvariablen vor, das sich an den Kalmangleichungen aus Abschnitt 2.4 orientiert. Darin wird im Aktualisierungsschritt die Messung und die Zustandsprädiktion sowie deren Unsicherheiten zu einem neuen Zustand und einer neuen Kovarianzmatrix fusioniert. Hierbei wird die Unsicherheit im Laufe des Filterprozesses i. A. kleiner. Durch den Bezug zum Kalmanfilter kann die Gauß'sche Fusion auch als Bayessche Fusion bezeichnet werden.

Gegeben seien zwei stochastisch unabhängige, normalverteilte Zufallsvariablen  $\mathbf{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1)$  und  $\mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \Sigma_2)$ . Die Kovarianzmatrix  $\Sigma_3$  und der Mittelwert  $\boldsymbol{\mu}_3$  für die

Fusion der Zufallsvariablen  $\mathbf{x}_1 \cup \mathbf{x}_2$  ergibt sich gemäß [Smi86] zu

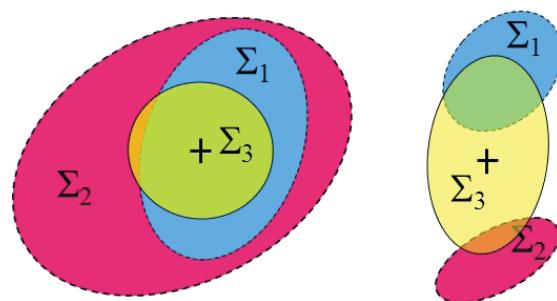
$$\Sigma_3 = \Sigma_1 - \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} \cdot \Sigma_1 \quad (2.52)$$

bzw.

$$\boldsymbol{\mu}_3 = \boldsymbol{\mu}_1 + \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} [\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1]. \quad (2.53)$$

Die Gln. 2.52 und 2.53 lassen sich aus den Kalmangleichungen (siehe Abschnitt 2.4) herleiten unter der Annahme, dass die Vektoren  $\mathbf{x}, \mathbf{z}$  Punkte in selben Parameterraum beschreiben, so dass folglich  $\dim(\mathbf{x}) = \dim(\mathbf{z})$ . Sei die Anzahl der Iterationen  $L = 1$  und die Bedingungsgleichung  $\mathbf{h}$  gegeben durch die Differenz der Messung und des Zustands  $\mathbf{h}(\mathbf{x}, \mathbf{z}) = (\mathbf{x} - \mathbf{z})$ . Das Kalmanfilter vereint in diesem Fall die Messung  $\mathbf{z}$  und den Zustandsvektor  $\mathbf{x}$  in Abhängigkeit von deren Unsicherheit in der Erwartung, dass die Werte beider Vektoren gleich sein sollen. Die Jacobi-Matrizen aus den Gln. 2.36 und 2.37 haben folglich die Form  $\mathbf{H}_{\mathbf{x}} = \mathbf{I}$  bzw.  $\mathbf{H}_{\mathbf{z}} = -\mathbf{I}$ . Somit ergeben sich die Gln. 2.52 und 2.53 aus den Gln. 2.41 bzw. 2.34.

Die Fusion einer endlichen Anzahl an Zufallsvariablen und den daraus resultierenden Erwartungswerten und Kovarianzmatrizen lässt sich durch schlichte Hintereinanderausführung der paarweisen Fusionsgleichungen erreichen. Die scheinbare Asymmetrie der beiden Gln. 2.53 und 2.52 hat keinerlei Einfluss auf das Fusionsergebnis und lässt sich leicht beweisen (siehe Anhang A.2). D. h. die Fusion der Normalverteilungen ist kommutativ.



**Abbildung 2.6:** Zwei schematische Beispiele der Fusion zweier normalverteilter Zufallsvariablen mit Kovarianzmatrizen  $\Sigma_1, \Sigma_2$  und der fusionierten Kovarianzmatrix  $\Sigma_3$ .



# 3

---

## Visuelle Odometrie

---

Unter visueller Odometrie versteht man gemeinhin die bildgestützte Bestimmung der Bewegung bzw. der Position und Orientierung eines Fahrzeugs. Im vorliegenden Fall wird die Bewegung mehrerer Kameras auf einer mobilen Plattform berechnet. Hierfür müssen in einem ersten Schritt signifikante Merkmalspunkte aus den Bildsequenzen detektiert und über mehrere Zeitschritte hinweg verfolgt werden. Mit Hilfe solcher Punktcorrespondenzen lässt sich auf die Bewegung der Kamera einerseits, sowie die Szenenstruktur andererseits rückschließen. Hierbei eignen sich nicht alle Szenenstrukturen oder Bewegungen gleichermaßen für die Berechnung der Kamerabewegung. Genaue Analysen finden sich u. a. in den Arbeiten [Wen87], [Stu97], [Kah99], [Pol00], [Har07]. Da die Bewegungsparameter der Kameras in den folgenden Kapiteln die Grundlage für die Schätzung der extrinsischen Parameter und somit der Selbstkalibrierung darstellen, kommt deren Schätzung besondere Bedeutung zu. Das folgende Kapitel beschreibt das hier eingesetzte Verfahren zur monokularen Bewegungsschätzung.

### 3.1 Extraktion und Verfolgung von Bildpunkten

Grundlage der bildbasierten Bewegungsschätzung der Kameramodule ist die Detektion signifikanter Bildpunkte und deren zeitliche Verfolgung. Hierbei steht die zeitliche

Verfolgung von Punkten innerhalb einer Kamerasequenz im Vordergrund. Die Wiedererkennung von Merkmalen in Bildern verschiedener Kameras gestaltet sich insbesondere bei MMD-Anordnungen schwierig, je nach Anordnung und Kontext sogar unmöglich. Perspektivische Verzerrungen und dreidimensionale Strukturverhältnisse müssen berücksichtigt werden, zudem soll die Konfiguration der Kamerakonstellation so allgemein wie möglich gehalten werden<sup>12</sup>.

Prinzipiell wird bei der Messung der Bildmerkmale in zwei Schritten verfahren. Zunächst werden signifikante Merkmale detektiert. Unterschiedliche Ansätze zur Merkmalsextraktion und -verfolgung und deren Evaluierung finden sich z. B. bei Schmid et al. [Sch05b], [Sch00]. Der nächste Schritt ist die Schätzung des zeitlichen Verschiebungsfeldes der Bildstrukturen (i. A. auch *Optischer Fluss* [Jae97] genannt). Hier kann man zwischen sog. dichten und dünnen Schätzverfahren unterscheiden [Bak07], [Bar94]. D. h. Verfahren, bei denen mittels Regularisierungs- und Glättungstechniken Korrespondenzen für alle Pixel gesucht werden, oder nur regionenbasiert für einzelne, signifikante Merkmale. Für die Bewegungsschätzung bzw. die Kalibrierung wird lediglich eine limitierte Anzahl an Flussvektoren benötigt, weshalb hier ein regionenbasiertes Verfahren sowohl zur Detektion und Verfolgung von Eckpunkten, als auch für deren Validierung eingesetzt wird [Har88], [Shi94], [Tom91], [Luc81].

Zur Verfolgung von signifikanten Bildpunkten unterstellen Shi und Tomasi [Shi94], dass sich die Regionen um zwei korrespondierende Punkte in zwei Bildern  $\mathcal{I}_1$  und  $\mathcal{I}_2$  durch eine affine 2D-Transformation  $[\mathbf{D}_a \ \mathbf{d}]_{2 \times 3}$  in Deckung bringen lassen

$$\mathcal{I}_1(\mathbf{x}) = \mathcal{I}_2(\mathbf{D}_a\mathbf{x} + \mathbf{d}). \quad (3.1)$$

[Shi94] formulieren daher das Tracking eines Punkts als die Schätzung der sechs Pa-

---

<sup>12</sup> So kann beispielsweise bei einem Kraftfahrzeug mit zwei seitlich entgegengesetzt ausgerichteten Kameras während einer gewöhnlichen Fahrt (beispielsweise im Stadtverkehr mit Häuserschluchten) i. A. nicht garantiert werden, dass die eine Kamera jemals dieselbe Szene beobachten wird, die die andere Kamera aktuell im Blickfeld hat.

rameter der Matrix  $\mathbf{D}_a$  und des Verschiebungsvektors  $\mathbf{d}$ . Hierbei spielt die Größe des Fensters  $\mathcal{W}$  um die Bildpunkte eine Rolle. Auf der einen Seite ist ein großes Fenster wünschenswert, um die affinen Parameter besser schätzen zu können. Andererseits sind kleinere Fenstergrößen für das Verfolgen eines Punktes besser geeignet, da sie robuster gegenüber Diskontinuitäten in der Tiefenstruktur der Szene sind. Daher wird für das Verfolgen der Punkte ein rein translatorisches Bewegungsmodell angenommen mit  $\mathbf{D}_a = \mathbf{I}$ . Die vollständige Schätzung der affinen Parameter wird lediglich zur Validierung der Korrespondenzen verwendet.

Die Schätzung der Parameter erfolgt über die Minimierung des Funktionalen

$$\varepsilon = \sum_{\mathbf{x} \in \mathcal{W}} c(\mathbf{x}) [\mathcal{I}_1(\mathbf{x}) - \mathcal{I}_2(\mathbf{D}_a \mathbf{x} + \mathbf{d})]^2 \quad (3.2)$$

wobei  $c$  eine Gewichtung des Fensters darstellt. Der vorgestellte Ansatz zur Lösung des daraus resultierenden Gleichungssystems mittels eines iterativen kleinsten-Quadrat-Verfahrens fordert zu dessen Lösbarkeit implizit eine gute Konditionierung des lokalen Strukturtensors

$$\mathcal{T}(\mathbf{x}) = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix}, \quad (3.3)$$

wobei  $g_x$  und  $g_y$  die lokalen Gradienten um den Aufpunkt  $\mathbf{x}$  bezeichnen. Das bedeutet, dass  $\mathcal{T}$  zwei große Eigenwerte besitzen muss, was einem ausgeprägten lokalen Gradienten um  $\mathbf{x}$  in zwei Richtungen entspricht. Diese Forderung dient dazu, das sog. Blendenproblem zu umgehen [Jae97]. Eine Genauigkeitsabschätzung der gefundenen Korrespondenzen kann durch eine gradientenbasierte Kovarianzmatrixschätzung auf den beiden Grauwertregionen erfolgen [Sti01].

Im Zuge der Kalibrierung dienen die Punktkorrespondenzen als Grundlage für die Bewegungsschätzung, der Schätzung von 3D-Szenenpunkten sowie zur Berechnung des Projektionsfehlers.

### 3.1.1 Ausreißereliminierung

Für die Bewegungsschätzung ist es essentiell, potentielle Ausreißer aus der Menge der Punktkorrespondenzen zu detektieren und zu entfernen. Ausreißer können falsche Korrespondenzzuordnungen aufgrund von Rauschen, Spiegelungen, periodischen Strukturmustern o. ä. sein, aber auch bewegte Objekte, die die Schätzung der Eigenbewegung beeinflussen. Dang [Dan07] schlägt zur Ausreißerdetektion einen robusten Hypothesentest vor, der in ein Kalmanfilter eingebettet ist. Im vorliegenden Fall geschieht die Ausreißerbestimmung mit Hilfe der Epipolargeometrie (siehe Anhang A.3). Die Bestimmung der Epipolargeometrie bzw. die Schätzung der Fundamentalmatrix  $\mathbf{F}$  für zwei Bilder kann direkt aus den Punktkorrespondenzen heraus erfolgen [Har97]. Zur robusten Schätzung mittels eines RANSAC-Verfahrens [Fis81] werden wiederholt zufällig 8 Punkte-Tupel aus der Korrespondenzmenge gewählt, um aus diesen die  $\Gamma$ -F-Matrizen der  $\Gamma + 1$  aufeinander folgenden Bilder zu schätzen und die zugehörigen *Inlier* zu zählen<sup>13</sup>. Die Menge der *Inlier*, also die Menge der zu einer F-Matrix konformen Korrespondenzpaare, ergibt sich aus der Prüfung des quadrierten Abstands

$$\frac{\|\mathbf{l}_1^T \tilde{\mathbf{x}}_1\|^2}{(l_1^1)^2 + (l_1^2)^2} \quad \text{und} \quad \frac{\|\mathbf{l}_2^T \tilde{\mathbf{x}}_2\|^2}{(l_2^1)^2 + (l_2^2)^2} \quad (3.4)$$

des Punktpaares  $\{\mathbf{x}_1, \mathbf{x}_2\}$  zu den entsprechenden Epipolgeraden

$$\begin{aligned} \mathbf{l}_1 &= \mathbf{F}\tilde{\mathbf{x}}_1 &= [l_1^1, l_2^1, l_3^1]^T \quad \text{bzw.} \\ \mathbf{l}_2 &= \mathbf{F}^T \tilde{\mathbf{x}}_2 &= [l_1^2, l_2^2, l_3^2]^T. \end{aligned} \quad (3.5)$$

---

<sup>13</sup> Im vorliegenden Fall erfolgt die Schätzung von  $\mathbf{F}_1$  nur für Korrespondenzen  $\{\mathbf{x}_{p0}, \mathbf{x}_{p1}\}_{p=1, \dots, \mathfrak{P}}$  des aktuellen Bildpaares  $t = 0$ . Für die restlichen Bildpaare sind die Bewegungstransformationen der letzten  $\Gamma - 1$  Zeitpunkte aus dem Bündelausgleich bekannt. Somit lassen sich die restlichen  $\Gamma - 1$  Fundamentalmatrizen  $\mathbf{F}_{r=2, \dots, \Gamma}$ , für die Korrespondenzpaare  $\{\mathbf{x}_{p1}, \mathbf{x}_{p2}\}, \dots, \{\mathbf{x}_{p\Gamma-1}, \mathbf{x}_{p\Gamma}\}_{p=1, \dots, \mathfrak{P}}$  parametrisch bestimmen (siehe Anhang A.3).

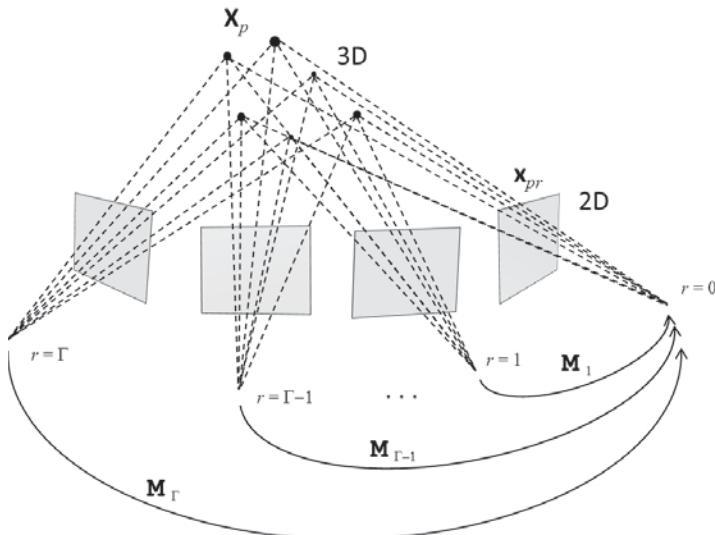


Abbildung 3.1: Extrahierte Verschiebungsvektoren: Ausreißer sind rot eingefärbt.

## 3.2 Der Bündelausgleich

Der Bündelausgleich (BA) ist ein Verfahren zur simultanen Schätzung der Kameraparameter und der Szenenstruktur. Diese Problematik wird i. A. auch als *Structure from Motion* (SfM) bezeichnet. Wenngleich in der Literatur etwaige Verfahren und Untersuchungen zur Bestimmung der extrinsischen *und* intrinsischen Parameter existieren [LH81], [LH87], [May92], [Fau87], [Aza95], so werden im Folgenden lediglich die Transformationsparameter der Kameras berücksichtigt.

Für den hier vorliegenden Zweck der Kamerakalibrierung wird der Bündelausgleich zur Bewegungsschätzung einer einzelnen Kamera benutzt. D. h. es werden *nicht* die extrinsischen Parameter zwischen  $\mathfrak{N}$  Kameras berechnet (was nur bei mehreren Kameras mit gemeinsamen Sichtbereichen möglich ist, wodurch eine direkte Berechnung der Kalibrierung möglich wäre), sondern die Bewegungstransformationen *eines* Kame-



**Abbildung 3.2:** Schematische Darstellung des Bündelausgleichs. Basierend auf korrespondierenden Bildpunkten  $\mathbf{x}_{pr}, p = 1, \dots, \mathfrak{P}, r = 0, \dots, \Gamma$ , kann sowohl auf die Euklidischen Bewegungsparameter  $\mathbf{M}_1, \dots, \Gamma$  als auch auf die Szenenstruktur  $\mathbf{X}_p$  rückgeschlossen werden.

ramoduls in  $\Gamma + 1$  aufeinanderfolgenden Messungen. Die Parameterschätzung mittels Bündelausgleich eignet sich deshalb besonders gut, weil die Bewegungsschätzung über einem beliebig großen Zeitfenster erfolgen kann. Diese Eigenschaft wird an späterer Stelle, beim Bewegungsausgleich (Kapitel 6), noch von Bedeutung sein.

Bei der Bewegungsschätzung mittels Bündelausgleich wird versucht, aus  $\Gamma + 1$  Kamerapositionen und  $\mathfrak{P}$  betrachteten Szenenpunkten  $\mathbf{X}_p$  sowohl die  $\Gamma$  Bewegungstransformationen als auch die Koordinaten der Szenenpunkte zu bestimmen. Grundlage hierfür sind Punktkorrespondenzen  $\{\mathbf{x}_{p0}, \dots, \mathbf{x}_{p\Gamma}\}$  und die Projektionsvorschrift

$$\tilde{\mathbf{x}}_{pr} = \mathbf{K}[\boldsymbol{\Omega}_r \mathbf{v}_r] \tilde{\mathbf{X}}_p. \quad (3.6)$$

Die intrinsische Matrix  $\mathbf{K}$  sei hierbei bekannt und konstant. Je Kameraposition ist also ein Vektor  $[\mathbf{v}_r^T, \boldsymbol{\omega}_r^T]^T, r = 0, \dots, \Gamma$ , mit drei translatorischen und drei rotatorischen Freiheitsgraden gesucht. Die Bewegungstransformationen seien o. B. d. A. relativ zur ersten Kameraposition  $r = 0$  orientiert (s. Abb. 3.2), so dass die erste Transformation mit  $[0, 0, 0, 0, 0, 0]^T$  definiert und fixiert werden kann.

Der Bündelausgleich bestimmt die Parameter durch Minimierung des quadratischen Projektionsfehlers

$$\sum_{p=1}^{\mathfrak{P}} \sum_{r=0}^{\Gamma} \|\pi(\mathbf{X}_p; \mathbf{K}, \mathbf{v}_r, \boldsymbol{\omega}_r) - \mathbf{x}_{pr}\|^2 \xrightarrow{!} \min_{\mathbf{v}_r, \boldsymbol{\omega}_r, \mathbf{X}_p} \quad (3.7)$$

Die Stärke des Bündelausgleichs liegt insbesondere darin begründet, dass mit dem Projektionsfehler ein physikalisch relevantes Residuum minimiert wird. Um den Parameterraum weiter zu reduzieren, wird je Szenenpunkt lediglich ein Skalar geschätzt, nämlich die Tiefenkoordinate  $\lambda$ . Ein Szenenpunkt  $\mathbf{X}$  lässt sich unter Kenntnis von  $\mathbf{K}$  sowie dessen projizierten Bildpunkts  $\mathbf{x}$  in Abhängigkeit der extrinsischen Parameter  $\boldsymbol{\Omega}, \mathbf{v}$  in Anlehnung an Gl. 2.1 beschreiben gemäß

$$\mathbf{X} = \lambda \cdot \boldsymbol{\Omega}^T \mathbf{K}^{-1} \tilde{\mathbf{x}} - \boldsymbol{\Omega}^T \mathbf{v}. \quad (3.8)$$

Der Parametervektor  $\mathbf{x} \in \mathbb{R}^{3+6 \cdot \Gamma}$  hat demnach die Form

$$\mathbf{x} = [\mathbf{v}_1^T, \boldsymbol{\omega}_1^T, \dots, \mathbf{v}_\Gamma^T, \boldsymbol{\omega}_\Gamma^T, \lambda_1, \dots, \lambda_{\mathfrak{P}}]^T. \quad (3.9)$$

Der Messvektor  $\mathbf{z} \in \mathbb{R}^{2 \cdot \mathfrak{P} \cdot \Gamma}$  beinhaltet die Punktkorrespondenzen

$$\mathbf{z} = [\mathbf{x}_{11}^T, \dots, \mathbf{x}_{1\Gamma}^T, \dots, \mathbf{x}_{\mathfrak{P}1}^T, \dots, \mathbf{x}_{\mathfrak{P}\Gamma}^T]^T. \quad (3.10)$$

Des weiteren sei eine Funktion  $\theta$  gegeben, die aus den Parameterwerten  $\mathbf{x}$  die Messung prädiziert

$$\hat{\mathbf{z}} = \theta(\mathbf{x}). \quad (3.11)$$

Die prädizierten Messpunkte ergeben sich aus der Projektionsvorschrift

$$\hat{\mathbf{x}}_{pr} = \pi(\mathbf{X}_p; \mathbf{K}, \mathbf{v}_r, \boldsymbol{\omega}_r). \quad (3.12)$$

Daraus lässt sich die Bedingungsgleichung

$$\begin{aligned} \mathbf{h}_{\mathbb{M}}(\mathbf{x}, \mathbf{z}) &= \mathbf{z} - \theta(\mathbf{x}) \\ &=: [\varepsilon_{11}, \dots, \varepsilon_{1\Gamma}, \dots, \varepsilon_{\mathfrak{P}1}, \dots, \varepsilon_{\mathfrak{P}\Gamma}]^T = \boldsymbol{\varepsilon} \end{aligned} \quad (3.13)$$

herleiten. Ferner sei  $\mathbf{J}$  die Jacobimatrix  $\frac{\partial \mathbf{h}_{\mathbb{M}}}{\partial \mathbf{x}}$ . Somit kann  $\mathbf{x}$  bei gegebener Startlösung  $\mathbf{x}_0$  mit einer LM-Optimierung (Abschnitt 2.3) bestimmt werden. Der Aktualisierungsvektor  $\boldsymbol{\delta}$  mit  $\mathbf{x}_{l+1} = \mathbf{x}_l - \boldsymbol{\delta}$  ergibt sich durch iteratives Lösen der erweiterten Normalengleichung 2.20.

Lourakis et al. [Lou09] stellten eine effiziente Berechnung des Bündelausgleichs vor, den sog. *dünnen Bündelausgleich* (SBA)<sup>14</sup>. Der Bündelausgleich und insbesondere die Implementierung des dünnen Bündelausgleichs sind im Zuge der erweiterten Hand-Augen-Kalibrierung für den an späterer Stelle vorgestellten Bewegungsausgleich von wesentlicher Bedeutung.

---

<sup>14</sup> engl. Sparse Bundle Adjustment

# 4

---

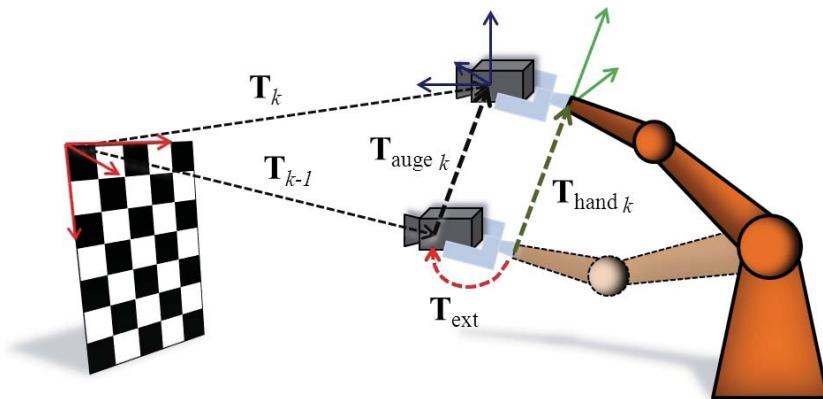
## Das allgemeine Kalibriermodell

---

### 4.1 Die Hand-Augen-Kalibrierung

Das Prinzip der Hand-Augen-Kalibrierung hat seinen Ursprung in der Robotik. Man stelle sich einen Roboterarm vor, an dessen Ende eine Kamera starr befestigt ist. Durch die Robotersteuerung lässt sich die Position des sog. Tool Center Points (TCP) eines Roboterarms recht genau bestimmen. Diese stimmt aber i. d. R. nicht mit dem Kamerazentrum überein. Für eine präzise Ansteuerung der Kameraposition mit dem Roboterarm ist es daher notwendig, die extrinsischen Transformationsparameter zwischen dem TCP (»Hand«) und der Kamera (»Auge«) zu kennen.

Die klassische Vorgehensweise zur Berechnung der Transformationsparameter, also zur extrinsischen Kalibrierung der Kamera hinsichtlich des Roboters, besteht darin, den Arm mit der Kamera vor einem bekannten Muster derart zu bewegen, dass sich die Lage und Orientierung des Kamerazentrums zum Koordinatensystem des Roboterarms ableiten lassen. Hierbei seien sowohl die Bewegungstransformationen  $T_{\text{hand}_k}$  des Roboters zum Zeitpunkt  $k$  aus der Steuerung als auch die Lage des Musters relativ zur



**Abbildung 4.1:** Schematische Darstellung der Hand-Augen-Kalibrierung aus der Robotik. Anhand eines wohlbekannten Musters kann die Position der Kameras und damit deren relative Verschiebungstransformation  $T_{auge}$  bestimmt werden. Die Position des Roboterarms lässt sich meist präzise ansteuern, so dass  $T_{hand}$  daher ebenfalls bekannt ist. Gesucht ist die Transformation zwischen der Kamera und dem Roboterarm  $T_{ext}$ .

Kamera zu jedem Aufnahmezeitpunkt bekannt<sup>15</sup>. Aus den bekannten Transformationen  $T_k$  und  $T_{k-1}$  zwischen den Koordinatensystemen der Kamera und des Musters zu zwei Zeitpunkten  $k$  und  $k - 1$ , lässt sich die Transformation  $T_{auge_k}$  zwischen den Kameras an  $k$  und  $k - 1$  direkt berechnen. Die gesuchte Transformation  $T_{ext}$  zwischen dem Kamera- und dem Roboterkoordinatensystem wird durch die Gleichung

$$T_{ext} \cdot T_{auge_k} = T_{hand_k} \cdot T_{ext} \quad (4.1)$$

ausgedrückt. Gl. 4.1 besagt prinzipiell, dass sich die extrinsischen Parameter  $T_{ext}$  mit Hilfe der Kamerabewegung  $T_{auge_k}$  und der Bewegung des Roboterarms  $T_{hand_k}$  bestimmen lassen.

Wollen wir nun anstatt der Transformation zwischen einer Hand und einem Auge die

---

<sup>15</sup> Bei bekannten Ausmaßen des Musters lässt sich die relative Lage der Kamera bestimmen, siehe Anhang A.7.

Transformation zwischen zwei Kameramodulen  $\mathcal{M}_i$  und  $\mathcal{M}_j$  anhand deren Bewegung bestimmen, lässt sich Gl. 4.1 allgemein umformulieren zu

$$\mathbf{M}_i = \mathbf{C}_{ij}^{-1} \mathbf{M}_j \mathbf{C}_{ij} \quad (4.2)$$

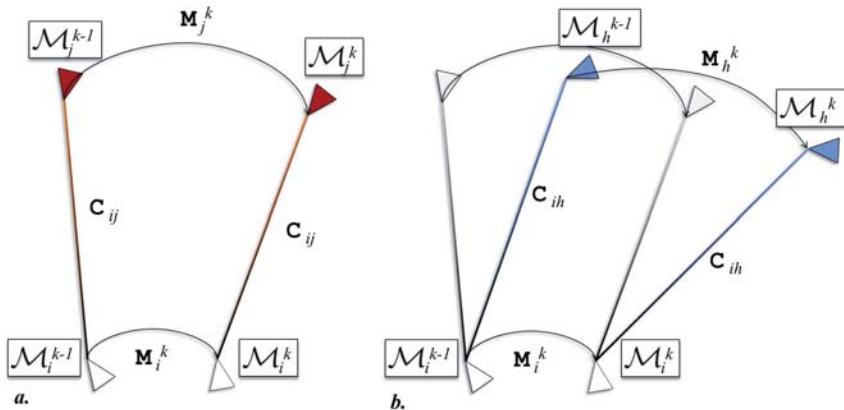
oder detaillierter

$$\underbrace{\begin{bmatrix} \boldsymbol{\Omega}_i & \mathbf{v}_i \\ \mathbf{0}^T & 1 \end{bmatrix}}_{\mathbf{M}_i} = \underbrace{\begin{bmatrix} \mathbf{R}_{ij}^T & -\mathbf{R}_{ij}^T \mathbf{t}_{ij} \\ \mathbf{0}^T & 1 \end{bmatrix}}_{\mathbf{C}_{ij}^{-1}} \underbrace{\begin{bmatrix} \boldsymbol{\Omega}_j & \mathbf{v}_j \\ \mathbf{0}^T & 1 \end{bmatrix}}_{\mathbf{M}_j} \underbrace{\begin{bmatrix} \mathbf{R}_{ij} & \mathbf{t}_{ij} \\ \mathbf{0}^T & 1 \end{bmatrix}}_{\mathbf{C}_{ij}} \quad (4.3)$$

und durch Ausmultiplizieren ergibt sich aus der resultierenden Translationskomponente die Bedingungsgleichung

$$[\mathbf{I} - \boldsymbol{\Omega}_i] \mathbf{t}_{ij} + \mathbf{R}_{ij} \mathbf{v}_j - \mathbf{v}_i = \mathbf{0}, \quad (4.4)$$

wobei  $\mathbf{R}_{ij}$  und  $\mathbf{t}_{ij}$  den gesuchten Euklidischen Transformationsparametern zwischen den zwei Kameramodulen entsprechen. Gl. 4.4 ist die Basisgleichung für die im Folgenden vorgestellte bewegungsbasierte, extrinsische Kalibrierung multiokularer Kameraanordnungen. Abb. 4.2 veranschaulicht die Basiskonstellation für  $\mathfrak{N} \geq 2$  Kameras.



**Abbildung 4.2:** Grundlegendes Kalibriermodell für (a.) zwei und (b.) drei Kameramodule  $\mathcal{M}_i$  und  $\mathcal{M}_j$  bzw.  $\mathcal{M}_h$ .

## 4.2 Generelle Einschränkungen bewegungsbasierter Kalibrierungen

Die Hand-Augen-Kalibrierung unterliegt gewissen Einschränkungen, insbesondere hinsichtlich der Art der Bewegungen, die zur Lösung des Problems verwendet werden dürfen. Tsai und Lenz [Tsa89] oder auch Horaud und Dornaika [Hor95] haben hierzu bereits analytische Untersuchungen durchgeführt. Für die vorliegende Aufgabenstellung ergeben sich folgende relevante Lemmata:

**Lemma 1** *Zur Bestimmung von  $\mathbf{t}_{ij}$  anhand Gl. 4.4 ist eine rotatorische Bewegung erforderlich.*

**BEWEIS** Bei rein translatorischer Bewegung ist die Rotationsmatrix der Bewegung gleich der Einheitsmatrix, also  $\boldsymbol{\Omega}_i = \mathbf{I}$ . Durch einsetzen in Gl. 4.4 erhält man

$$\mathbf{R}_{ij} \mathbf{v}_j = \mathbf{v}_i \quad (4.5)$$

Der Translationsvektor  $\mathbf{t}_{ij}$  ist aus Gl. 4.5 verschwunden. D.h. die Basisgleichung ist für alle Werte von  $\mathbf{t}$  erfüllt, und  $\mathbf{t}_{ij}$  ist somit unbestimmt. ■

**Lemma 2** *Zur Bestimmung von  $\mathbf{t}_{ij}$  anhand Gl. 4.4 sind mindestens zwei Bewegungspaare der Module  $\mathcal{M}_i, \mathcal{M}_j$  erforderlich.*

**BEWEIS** Gl. 4.4 formuliert ein lineares Gleichungssystem zur Lösung von  $\mathbf{t}_{ij}$ .  $[\mathbf{I} - \boldsymbol{\Omega}]$  ist hierbei die Koeffizientenmatrix. Weil  $\det(\mathbf{I} - \boldsymbol{\Omega}) = 0$  und somit  $[\mathbf{I} - \boldsymbol{\Omega}]$  singulär ist, sind mindestens zwei Bewegungspaare notwendig. ■

**Lemma 3** *Die Rotationsachsen der Bewegungsrotationen  $\boldsymbol{\Omega}_{i_t}, t = 1, \dots, \mathfrak{H}$ , zur Bestimmung von  $\mathbf{t}$  anhand Gl. 4.4 müssen unterschiedliche Richtungen aufweisen.*

**BEWEIS** Sei die Rotation  $\mathbf{R}_{ij}$  bekannt. Dann muss zur Bestimmung der Translation  $\mathbf{t}_{ij}$  und wegen Lemma 2 aus Gl. 4.4 ein (überbestimmtes) Gleichungssystem mit mehreren Bewegungspaaren für Module  $\mathcal{M}_i, \mathcal{M}_j$  gebildet werden:

$$\begin{bmatrix} \mathbf{I} - \boldsymbol{\Omega}_{i_1} \\ \vdots \\ \mathbf{I} - \boldsymbol{\Omega}_{i_{\mathfrak{H}}} \end{bmatrix} \mathbf{t}_{ij} = \underbrace{\begin{bmatrix} \mathbf{v}_{i_1} - \mathbf{R}_{ij} \mathbf{v}_{j_1} \\ \vdots \\ \mathbf{v}_{i_{\mathfrak{H}}} - \mathbf{R}_{ij} \mathbf{v}_{j_{\mathfrak{H}}} \end{bmatrix}}_{=: \mathbf{b}}. \quad (4.6)$$

Hierbei sind  $\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_{\mathfrak{H}}}$  und  $\boldsymbol{\Omega}_{i_1}, \dots, \boldsymbol{\Omega}_{i_{\mathfrak{H}}}$  die translatorischen bzw. rotatorischen Transformationskomponenten von Modul  $i$  innerhalb einer Bewegungshistorie der Größe  $\mathfrak{H}$ . Sei  $\boldsymbol{\alpha}_R$  die Rotationsachse zu  $\boldsymbol{\Omega}_{i_1}$ , also der Eigenvektor zum Eigenwert 1 von  $\boldsymbol{\Omega}_{i_1}$  mit  $\boldsymbol{\Omega}_{i_1} \boldsymbol{\alpha}_R = \boldsymbol{\Omega}_{i_1}^T \boldsymbol{\alpha}_R = \boldsymbol{\alpha}_R$  und  $\boldsymbol{\alpha}_R^T \boldsymbol{\Omega}_{i_1} = \boldsymbol{\alpha}_R^T$ . Durch linksseitiges Multiplizieren von Gl. 4.6 mit der Matrix

$$\mathbf{A} = \begin{bmatrix} \boldsymbol{\alpha}_R^T & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \boldsymbol{\alpha}_R^T \end{bmatrix}_{\mathfrak{H} \times 3\mathfrak{H}} \quad (4.7)$$

erhält man

$$\left[ \begin{array}{c} \underbrace{\boldsymbol{\alpha}_R^T - \underbrace{\boldsymbol{\alpha}_R^T \boldsymbol{\Omega}_{i_1}}_{= \boldsymbol{\alpha}_R^T}}_{=: \mathbf{B}} \\ = \mathbf{0} \\ \vdots \\ \underbrace{\boldsymbol{\alpha}_R^T - \underbrace{\boldsymbol{\alpha}_R^T \boldsymbol{\Omega}_{i_{\mathfrak{H}}}}_{=: \mathbf{B}}} \end{array} \right] \mathbf{t}_{ij} = \mathbf{Ab}. \quad (4.8)$$

Aus Gl. 4.8 wird ersichtlich, dass die Rotationsachsen der  $\boldsymbol{\Omega}_{i_t}$  unterschiedliche Richtungen aufweisen müssen, da  $\mathbf{B}$  ansonsten keinen vollen Spaltenrang mehr hätte und das LGS somit nicht mehr eindeutig lösbar, d. h.  $\mathbf{t}_{ij}$  nicht bestimmbar wäre. Falls alle Rotationen  $\boldsymbol{\Omega}_{i_t}$  dieselbe Rotationsachse  $\boldsymbol{\alpha}_R$  besitzen, gilt  $\mathbf{B} = \mathbf{0}$ . ■

Lemma 3 ist insbesondere bei der Kalibrierung von Plattformen, die sich ausschließlich planar, d. h. auf einer Grundebene bewegen, von wesentlicher Bedeutung.

**Lemma 4** *Die Bewegungsrotationen  $\boldsymbol{\Omega}_{i_t}, t = 1, \dots, \mathfrak{H}$  aus Gl. 4.6 dürfen nicht identisch sein, da  $\mathbf{t}_{ij}$  ansonsten unbestimmt ist.*

**BEWEIS** Der Beweis für Lemma 4 ist ein Spezialfall von Lemma 3. Konstante rotatorische Bewegungen, wie beispielsweise Kreisfahrten, bewirken, dass in Gl. 4.8 der Fall  $\mathbf{B} = \mathbf{0}$  eintritt. ■

Zusammenfassend lässt sich sagen, dass zur Bestimmung aller extrinsischen Freiheitsgrade mit dem allgemeinen Kalibriermodell

- eine Bewegungshistorie der Länge  $\mathfrak{H} > 2$  verwendet werden muss,
- die Bewegung rotatorisch sein muss,
- die Bewegungen nicht komplanar sein dürfen und
- die rotatorischen Anteile der Bewegungen innerhalb einer Historie nicht gleich sein dürfen.

In Kapitel 5 wird ein erweitertes Kalibriermodell vorgestellt, das dem besonderen Fall der planaren Bewegungen Rechnung trägt.

### 4.3 Zur Verwendung skalierter Translationsbeträge der Kamerabewegung

Der Betrag der translatorischen Bewegung eines Moduls ist eine essentiell wichtige Größe zur Bestimmung des translatorischen Abstands zwischen den Kameras. Ein einfaches Beispiel veranschaulicht diesen Sachverhalt. Man stelle sich ein Fahrzeug mit je einer Kamera an seiner linken und rechten Außenseite vor. Fährt dieses Fahrzeug auf einer ebenen Straße um eine Rechtskurve, so ist die Wegstrecke, die die linke Kamera zurücklegt, größer als die der kurveninneren rechten. An diesem Beispiel erkennt man nun auch, warum der Einsatz von handelsüblicher, odometrischer Fahrzeugsensorik zur Geschwindigkeitsmessung hier nur geringfügig weiterhilft. Entscheidend für die Kalibrierung ist der Translationsbetrag der Bewegung jedes einzelnen Moduls. Vielmehr müsste jedes Kameramodul um einen eigenen Geschwindigkeitsmesser (z. B. Intertial-sensorik) ergänzt werden, so dass die Geschwindigkeit jedes Kameramoduls, und nicht nur der Plattform, direkt gemessen werden kann<sup>16</sup> [Bae05].

In praktischen Versuchen im Vorfeld der vorliegenden Arbeit hat sich gezeigt, dass sich selbst bei einer integrierten Prädiktion des Geschwindigkeitsbetrags im Bündelausgleich der Betrag der translatorischen Geschwindigkeit nicht stabil über längere Zeiträume hinweg schätzen lässt [Pag12]. Hierbei wurde der Geschwindigkeitsbetrag des vorangegangenen Zeitpunkts fixiert, so dass die Bewegungsschätzung, insbesondere die Geschwindigkeitsschätzung, relativ zum Vorgängerzeitpunkt erfolgte. Abb. 4.3 zeigt einen Bildausschnitt und eine Skizze der Messfahrt.

Abb. 4.4 zeigt die Translationsparameter der Bewegungsschätzung zweier Kameras

---

<sup>16</sup> Eine zentrale Intertialsensorik zur Bestimmung der Rotation und Translation des Fahrzeugs wäre in diesem Falle nicht ausreichend. Die gesuchten Kalibrierparameter, also die relative Lage der Kameras zum Bezugspunkt des Fahrzeugs, lassen sich erst berechnen, wenn auch die Bewegungen der Kameras – relativ zum Fahrzeug – bekannt sind.

in einem Fahrzeug. Auf den ersten Blick ergeben die Translationsparameter keinen Sinn, da die Schätzungen der beiden Module augenscheinlich in keiner Weise zusammenhängen. Tatsächlich ist dies jedoch nicht der Fall. Lediglich die kontinuierliche Schätzung des Translationsbetrags erweist sich aufgrund von Fehlerfortpflanzungen als nutzlos. Wie Abb. 4.5 zeigt, liefert die auf 1 skalierte Bewegungsrichtung

$$\mathbf{v}^s = \frac{\mathbf{v}}{||\mathbf{v}||} \quad (4.9)$$

plausible Werte. Man erkennt deutlich die Richtungsänderungen bei der U-förmigen Testfahrt des Fahrzeugs (s. Abb. 4.3).

Die rotatorischen Bewegungsparameter sind von dieser Problematik nicht betroffen. Abb. 4.6 zeigt, dass die Schätzungen beider Kameras zueinander konsistent sind. Auch hier sind die vom Fahrzeug durchgeführten zwei leichten Rechtskurven klar erkennbar (positive Ausschläge im  $r_y$ -Parameter).

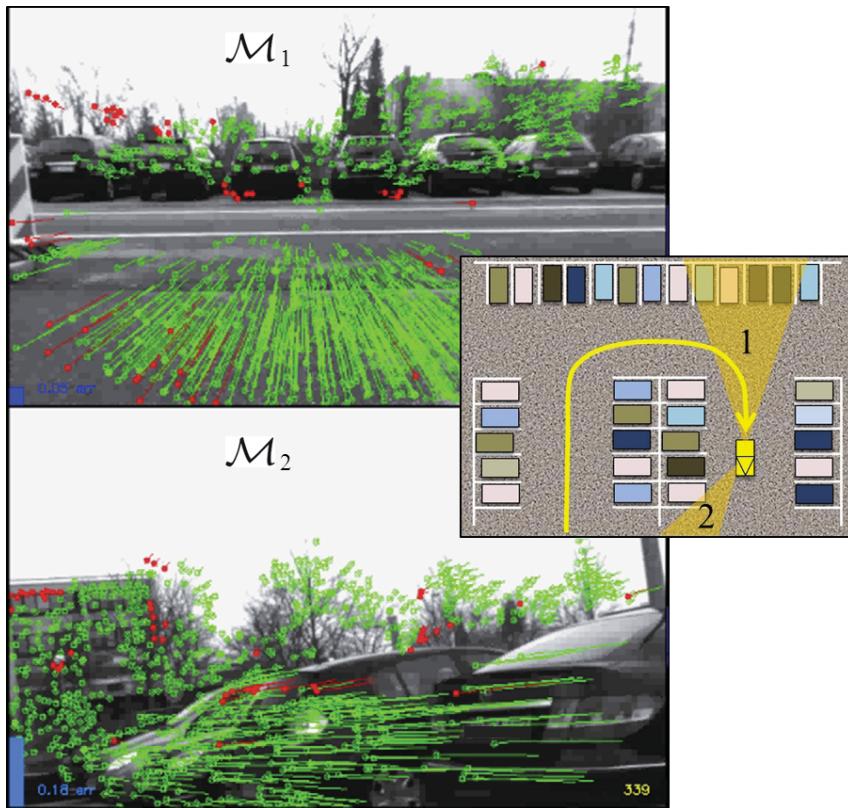
Gängige Geschwindigkeitssensorik ist in der vorliegenden Arbeit allenfalls von Nutzen, wenn alle Module dieselbe Bewegungsgeschwindigkeit haben (insbesondere bei reinen Geradeausfahrten ohne rotatorischen Bewegungsanteil)<sup>17</sup> Die relativen, unterschiedlichen Beträge der Bewegungstranslationen der einzelnen Kameras können damit selbst bei rotatorischen Bewegungen – bei Unkenntnis der Kameraanordnung – nicht erfasst werden. D. h. die Kenntnis der Plattformbewegung alleine ist noch nicht ausreichend, um eine bewegungsbasierte Kalibrierung durchzuführen.

In dieser Arbeit wird das Skalierungsproblem durch den Bewegungsausgleich gelöst (Kapitel 6). Der Bewegungsausgleich schätzt sowohl die extrinsischen Parameter als auch die Skalierungsparameter der Modulgeschwindigkeiten simultan. Hierbei gehen

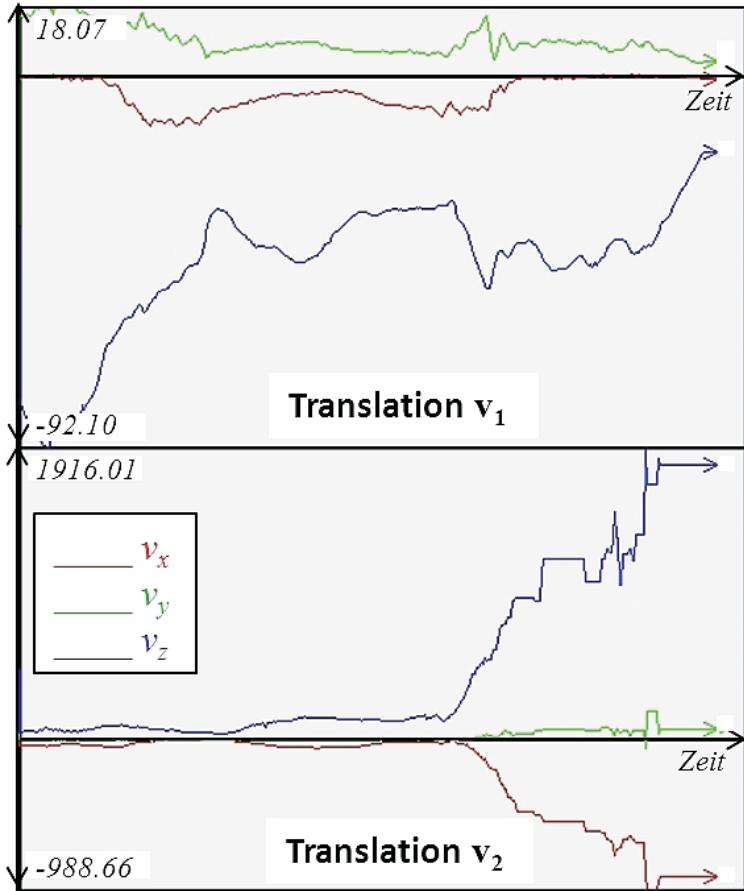
---

<sup>17</sup> Odometrische Geschwindigkeitsmessungen können auch verwendet werden, um das Abdriften der geschätzten Geschwindigkeiten wie in Abb. 4.4 einzudämmen. Eine solche Integration externer Sensorik zur Stabilisierung der bildgestützten Berechnungen wurde in der vorliegenden Arbeit nicht realisiert. Prinzipiell lassen sich aber externe Stellgrößen recht einfach in nichtlineare Schätzer wie das Kalmanfilter integrieren [Dan07].

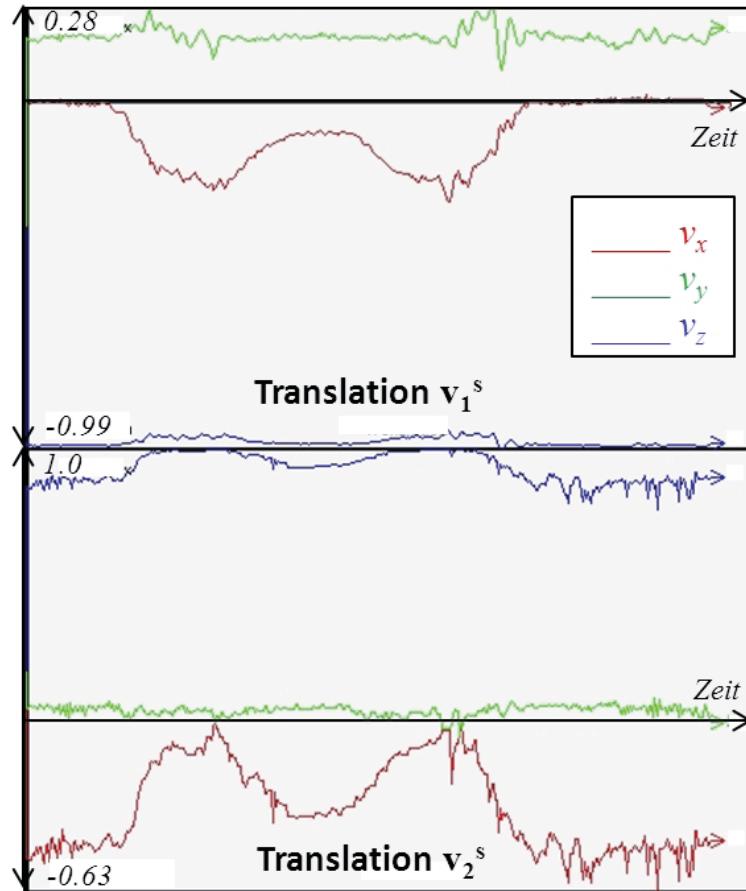
die Translationsvektoren der Bewegungstransformationen in skalierte Form in die Berechnungen ein.



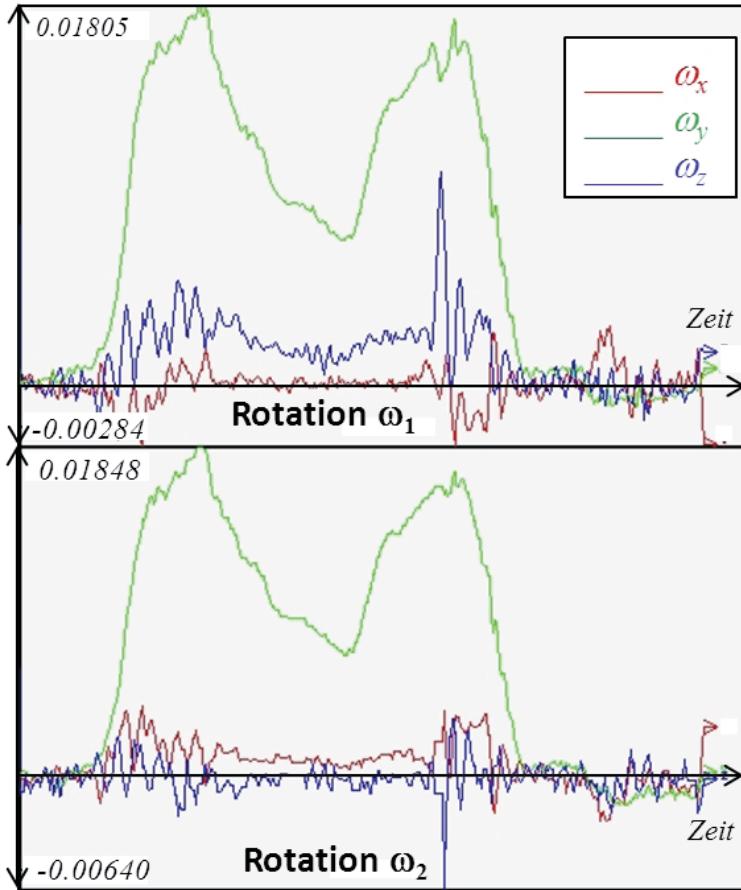
**Abbildung 4.3:** Kamerabilder mit Flussvektoren einer Zweikameraanordnung in einem Fahrzeug. Mitte: Schematische Darstellung der zurückgelegten Strecke auf dem Testgelände und die ungefähre Anordnung der beiden Kameras mit deren Sichtfeldern. Rote Flussvektoren markieren Ausreißer (siehe Abschnitt 3.1.1).



**Abbildung 4.4:** Translationsparameter der Bewegungsschätzung der zwei Kameramodule aus Abb. 4.3. Die Translationswerte liefern keine plausiblen Werte und driften auseinander.



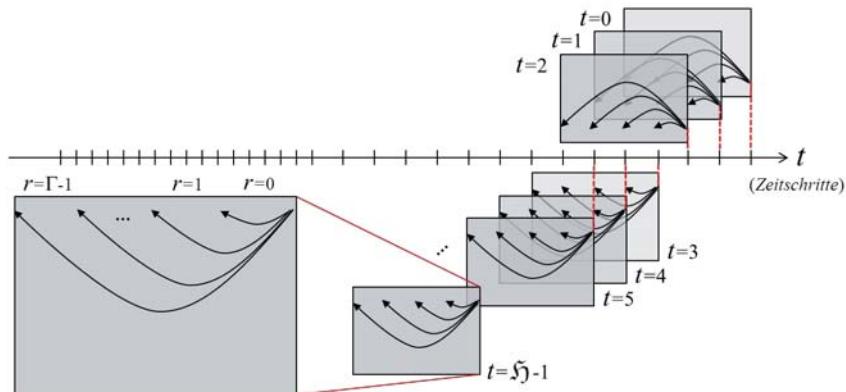
**Abbildung 4.5:** Normierte Translationsparameter der Bewegungsschätzung der zwei Kameramodule aus Abb. 4.3. Ähnlich wie in Abb. 4.6 lässt sich auch hier die Richtungsänderung der Kameras in den Kurven erkennen.



**Abbildung 4.6:** Rotationsparameter der Bewegungsschätzung der zwei Kameramodule aus Abb. 4.3. Man erkennt deutlich die positiven Ausschläge von  $\omega_y$  an den beiden gefahrenen Rechtskurven.

## 4.4 Notwendigkeit und Einführung einer Historie

Für die folgenden Betrachtungen und Berechnungen ist die Einführung einer Historie der Kamerabewegungen vonnöten. Hierbei sind zwei Arten von Historien zu unterscheiden: Die *Bewegungshistorie* berücksichtigt  $\mathfrak{H}$  zurückliegende Zeitschritte. Jede Messeinheit der Bewegungshistorie besteht aus einer Art Unterhistorie, einem sog. *Bewegungsblock*. Der Bewegungsblock ist zusammenhängend und enthält explizit die letzten  $\Gamma$  Bewegungstransformationen, ausgehend vom Zeitpunkt  $t$  der Bewegungshistorie. Eine Bewegungshistorie besteht somit aus der Menge der Bewegungstransformationen  $\{\mathbf{M}_{tr}\}_{t=0, \dots, \mathfrak{H}-1, r=0, \dots, \Gamma-1}$ . Die Bewegungsblöcke sind Resultat der Berechnung des Bündelausgleichs (siehe Abschnitt 3.2). Abb. 4.7 veranschaulicht eine aus Bewegungsblöcken bestehende Bewegungshistorie.



**Abbildung 4.7:** Historie  $t = 0, \dots, \mathfrak{H} - 1$  und Bewegungsblock  $r = 0, \dots, \Gamma - 1$ . Zu jedem zurückliegenden Zeitpunkt  $t$  wurden Blöcke der letzten  $\Gamma$  Bilder für die Bewegungsschätzung verwendet. Für die Berechnung der Kalibrierparameter verwendet man eine Historie von  $\mathfrak{H}$  Bewegungsblöcken.

Die Notwendigkeit einer Bewegungshistorie wurde bereits in Lemma 2 dargelegt. Die Notwendigkeit der Bewegungsblöcke hingegen ergibt sich erst an späterer Stelle bei der Einführung des Bewegungsausgleichs für skalierte Bewegungen, sei aber an dieser

Stelle schon vorweg genommen.

## **Teil II**

# **Bewegungsbasierte, kontinuierliche extrinsische Kalibrierung**



# 5

---

## **Das erweiterte Kalibriermodell für Straßenfahrzeuge**

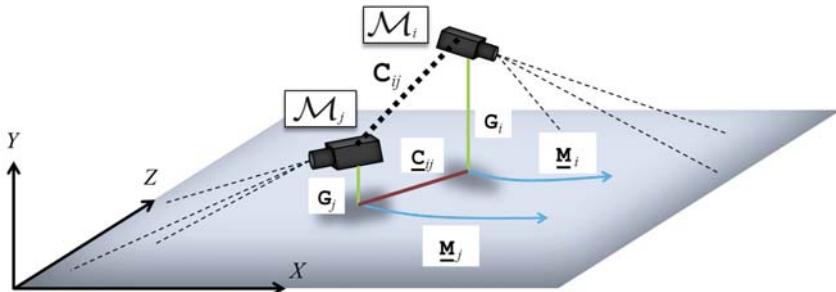
---

Das allgemeine Kalibriermodell beruht im Kern auf den Überlegungen und Prinzipien der Hand-Auge-Kalibrierung. In Abschnitt 4.2 wurden hierzu die wichtigsten Eigenschaften vorgestellt. Da in der vorliegenden Arbeit die Zielplattform allerdings kein Roboterarm, sondern ein Straßenfahrzeug ist, ergeben sich andere Rahmenbedingungen. Straßenfahrzeuge bewegen sich wesentlich in einer Bodenebene, d. h. planar. Im vorliegenden Fall entspricht dies einer Bewegung in der  $XZ$ -Ebene des Weltkoordinatensystems (s. Abb. 5.1). Dies impliziert, dass sich das Fahrzeug, und damit die Kameras, um eine Rotationsachse senkrecht zur Bodenebene drehen und translatorisch in der  $X - Z$ -Ebene bewegen. Da die rotatorische Bewegung grundlegend für die Bestimmung der translatorischen Kalibrierparameter ist (siehe Abschnitt 4.2), reicht die Fahrzeugbewegung lediglich aus, um die relativen Kamerapositionen und -orientierungen innerhalb der  $XZ$ -Ebene zu bestimmen. Der Höhenunterschied bleibt damit allerdings unbestimmt [Pag11].

Um auch diesen Freiheitsgrad abzufangen, zielt der in den folgenden Abschnitten vorgestellte Lösungsansatz darauf ab, die extrinsischen Parameter nicht direkt im 3D-

Raum zu schätzen, sondern einen Umweg über die Bodenebene zu nehmen. Diese Bodenebene ist für alle Kameramodule gleich.

In einem ersten Schritt berechnet daher jedes Modul die Transformation in die Bodenebene. Anschließend werden die Kameras und deren Bewegungstransformationen in diese Bodenebene überführt, um danach die Lage der Kameras in der  $XZ$ -Ebene berechnen zu können. In einem letzten Schritt werden die Teiltransformationen wieder zusammengeführt, so dass man im Idealfall die extrinsischen Parameter im 3D-Raum mit allen sechs Freiheitsgraden erhält (s. Abb. 5.1).



**Abbildung 5.1:** Erweitertes Kalibriermodell durch Einbinden der Bodenebene in die Beschreibung der Kameraanordnung.

Zusammengefasst stellt sich das Vorgehen im erweiterten Modell wie folgt dar:

1. Bestimmung der Transformationen in die Bodenebene  $G_i, G_j$  für alle Module  $\mathcal{M}_i, \mathcal{M}_j$
2. Transformation der Bewegungstransformationen  $M_i, M_j$  in die Bodenebene zu  $\underline{M}_i, \underline{M}_j$
3. Schätzung der relativen Anordnung  $C_{ij}$  innerhalb der Bodenebene anhand der Transformationen  $\underline{M}_i, \underline{M}_j$
4. Berechnung der extrinsischen Parameter  $C_{ij}$  durch Zusammenführung der Transformationen  $G_i, C_{ij}$  und  $G_j$

In den folgenden Abschnitten dieses Kapitels wird insbesondere der erste Punkt behandelt. Ausführungen zu Punkt 2 findet man in Anhang A.5, Punkt 3 wird ein gesondertes Kapitel gewidmet 6. Die Zusammenführung der Transformationsparameter aus Punkt 4 ist Inhalt von Kapitel 7.

## 5.1 Erweiterte Annahmen

Das erweiterte Kalibriermodell ergibt sich aus dem allgemeinen Modell durch die Hinzunahme weiterer Annahmen:

**Planare Bewegungen** Im folgenden wird davon ausgegangen, dass sich der mobile Sensorträger auf einer Ebene bewegt. Es wird unterstellt, dass alle Module dieselbe Ebene sehen. Darüber hinaus sei diese Ebene für alle Kameramodule – aus verschiedenen Perspektiven und zumindest zeitweise – sichtbar<sup>18</sup>.

Wenngleich auch Straßenfahrzeuge Unebenheiten wie Brücken oder Schwankungen durch Straßenschäden begegnen müssen, so ist dies doch im Allgemeinen eine sinnvolle Annahme<sup>19</sup>.

**Abhängigkeit der rotatorischen von der translatorischen Bewegung** Es wird unterstellt, dass die mobile Plattform ein zweiachsiges, vorderachsengelenktes Fahrzeug sei. Derartige Fahrzeuge folgen i. A. einer nichtholonomen Bewegungsbedingung, sodass rotatorische Bewegungen nur bei gleichzeitiger translatorischer Bewegung zugelässig sind (siehe Abschnitt 2.2)<sup>20</sup>.

Die Planaritätsannahme wird zusammen mit der Annahme der nichtholonomen Be-

---

<sup>18</sup> In der Literatur ist diese Annahme häufig unter dem Begriff der sog. *flat world assumption* zu finden.

<sup>19</sup> Die Detektion von Verletzungen der Planaritätsannahme wird in der vorliegenden Arbeit nicht explizit behandelt. Dennoch lassen die meisten Schätzverfahren über ein verbleibendes Residuum Aussagen über die Güte der Schätzung und damit der Erfüllung der Annahmen zu. Eine explizite Detektion und Behandlung derartiger Ausnahmefälle ist das Ziel weiterführender Arbeiten.

<sup>20</sup> Auch hier sind Ausnahmefälle wie Rutschen auf glatter Fahrbahn nicht berücksichtigt.

wegungen dafür verwendet, die Position eines zentralen Fahrzeugbezugspunkts in Abhängigkeit der rotatorischen Bewegung und der Geschwindigkeit zu bestimmen.

## 5.2 Fallspezifische Auswertung der Bewegung

Prinzipiell ist es möglich, anhand von Gleichung 4.4 die extrinsischen Parameter aus den Bewegungsstrajektorien der Kameras zu rekonstruieren [Esq07], [Tsa89]. Da aber insbesondere beim erweiterten Kalibriermodell davon ausgegangen werden muss, dass sich die Bewegungsparameter weder allzu schnell noch allzu stark ändern (wie dies beispielsweise mit einem Roboterarm möglich wäre), bietet es sich an, die Besonderheiten bestimmter Bewegungsarten auszunutzen.

Im folgenden werden zwei Bewegungstypen unterschieden und separat ausgewertet:

- **Rein translatorische Bewegung mit  $\Omega = \mathbf{I}$ .**

Bei nichtholonomen Bewegungen, wie sie beim erweiterten Kalibriermodell unterstellt sind, entspricht dies einer Geradeausfahrt der mobilen Plattform. Dieser Bewegungstyp wird u. a. dazu verwendet, die relative Orientierung der Kameras zueinander zu bestimmen (siehe Abschnitt 5.2.1). Für die Schätzung des Abstands der Kameras von der Bodenebene ist es gar eine Notwendigkeit, auf rein translatorische Bewegungen zurückzugreifen (siehe Abschnitt 5.3.4). Darüber hinaus wird eine Geradeausfahrt zur Bewegungsinitialisierung verwendet (siehe Anhang A.4).

- **Allgemeine (rotatorische) Bewegung mit  $\Omega \neq \mathbf{I}$ .**

Anhand rotatorischer Bewegungen kann die Rotationsachse bestimmt werden, die ihrerseits für die Schätzung der Bodenebene wertvolle Informationen bereit stellt (siehe Abschnitt 5.3.1 und 5.3.3). Dem nichtholonomen Bewegungsmodell zufolge entsprechen rotatorische Bewegungen einer Kurvenfahrt. Diese sind Grundlage für die bewegungsbasierte Kalibrierung und den Bewegungsausgleich in Abschnitt 6.

Eine explizite Auflistung der Bewegungsphasen und die damit verbundenen Schätzungen findet sich in Tabelle 8.1.

### 5.2.1 Schätzung der Kameraorientierungen bei nichtrotatorischer Bewegung

Ausgehend von der Basisgleichung 4.2 für zwei Module  $\mathcal{M}_i, \mathcal{M}_j$ , erhält man bei rein translatorischer Bewegung mit  $\boldsymbol{\Omega} = \mathbf{I}$  die Gleichung

$$\mathbf{R}_{ij}\mathbf{v}_i - \mathbf{v}_j = \mathbf{0}. \quad (5.1)$$

Aus dieser Konstellation heraus lässt sich die Orientierung, d. h. die Rotation  $\mathbf{R}_{ij}$ , zwischen den beiden Modulen leicht bestimmen. Ersetzt man die Rotationsmatrix  $\mathbf{R}_{ij}$  durch die entsprechenden Quaternionen, so lässt sich Gl. 5.1 umformulieren zu

$$\mathbf{q} \star \mathbf{v}_i^q \star \bar{\mathbf{q}} - \mathbf{v}_j^q = \mathbf{q} \star \mathbf{v}_i^q - \mathbf{v}_j^q \star \mathbf{q} = \mathbf{0} \quad (5.2)$$

wobei  $\bar{\mathbf{q}}$  die zu  $\mathbf{q}$  konjugierte Quaternion ist, und  $\mathbf{v}^q = [0, v_x, v_y, v_z]$  die dem Vektor  $\mathbf{v}$  assoziierte Quaternion mit einem nullgesetzten Realteil. Der besseren Übersichtlichkeit wegen wurde bei den Quaternionen  $\mathbf{q}_{ij}$  auf die Indexierung verzichtet.

Der  $\star$ -Operator multipliziert zwei Quaternionen (siehe Anhang A.1). Die Quaternionenmultiplikationen lassen sich durch Matrixprodukte ersetzen, und Gl. 5.2 lässt sich darstellen als

$$\mathbf{Q}_{\mathbf{v}_i}^\zeta \mathbf{q} - \mathbf{Q}_{\mathbf{v}_j}^\zeta \mathbf{q} = \underbrace{\left[ \mathbf{Q}_{\mathbf{v}_i}^\zeta - \mathbf{Q}_{\mathbf{v}_j}^\zeta \right]}_{=: \mathbf{A}} \mathbf{q} \quad (5.3)$$

mit

$$\mathbf{Q}_{\mathbf{v}_i}^\zeta = \begin{bmatrix} 0 & -v_{x_i} & -v_{y_i} & -v_{z_i} \\ v_{x_i} & 0 & v_{z_i} & -v_{y_i} \\ v_{y_i} & -v_{z_i} & 0 & v_{x_i} \\ v_{z_i} & v_{y_i} & -v_{x_i} & 0 \end{bmatrix} \quad (5.4)$$

und

$$\mathbf{Q}_{\mathbf{v}_j}^> = \begin{bmatrix} 0 & -v_{x_j} & -v_{y_j} & -v_{z_j} \\ v_{x_j} & 0 & -v_{z_j} & v_{y_j} \\ v_{y_j} & v_{z_j} & 0 & -v_{x_j} \\ v_{z_j} & -v_{y_j} & v_{x_j} & 0 \end{bmatrix} \quad (5.5)$$

Will man mehrere Bewegungen einer Historie der Länge  $\mathfrak{H}$  berücksichtigen, so dass

$$\mathbf{A}_t = [\mathbf{Q}_{\mathbf{v}_i^t}^< - \mathbf{Q}_{\mathbf{v}_j^t}^>], \quad t = 1, \dots, \mathfrak{H}, \quad (5.6)$$

dann lassen sich die  $\mathbf{A}_t$  stapeln zu

$$\mathbf{A} = [ \dots, \mathbf{A}_t^T, \dots ]^T. \quad (5.7)$$

Die gesuchte Orientierung zwischen den Modulen erhält man durch Lösen des homogenen Gleichungssystems

$$[\mathbf{A}^T \mathbf{A}] \mathbf{q} = \mathbf{0} \quad (5.8)$$

mittels Eigenwertzerlegung. Die gesuchte Orientierung  $\mathbf{q}$  ist somit der Eigenvektor zum kleinsten Eigenwert von  $\mathbf{A}^T \mathbf{A}$  [Esq07].

### 5.3 Einbeziehung der Bodenebene

Die Schätzung der Bodenebene in Form der Transformation und Ausrichtung des Kamerasystemkoordinatensystems in die Bodenebene erfolgt, wie die Berechnung der Orientierung der Kamera, bei einer rein translatorischen Bewegung ( $\boldsymbol{\Omega} = \mathbf{I}$ ). Hierfür werden zunächst 3D-Punkte selektiert, die in der Bodenebene liegen. Anschließend werden mittels nicht-linearer Optimierung die Transformationsparameter ermittelt, die die Koordinaten der Bodenpunkte in das zweidimensionale Koordinatensystem der Ebene überführen.

Kennt man bereits die Rotationsachse eines Kameramoduls, so ist durch die Orthogona-

lität der Achse zum Boden auch die Orientierung der Bodenebene bekannt und es bleibt lediglich der Abstand zwischen Kamera und Ebene zu bestimmen. Die Rotationsachse ist sehr hilfreich, wenn es darum geht, Bodenpunkte zu selektieren. Die Achse kann mittels rotatorischer Bewegungen, d. h. Kurvenfahrten berechnet werden und erfolgt günstigerweise vor der Schätzung der Bodenebene.

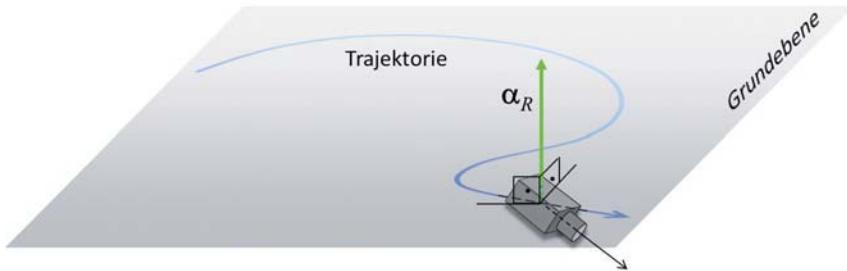
Im Folgenden werden die Teilberechnungen, die zur Bestimmung der Bodenebene herangezogen werden, erläutert.

### 5.3.1 Schätzung der Rotationsachse

Bewegt sich eine mobile Plattform kurvenförmig in einer Ebene, so rotiert die Plattform um genau eine Achse senkrecht zur Ebene. Folglich rotieren auch die Kameras auf der Plattform um diese Achse, abhängig von der Orientierung der jeweiligen Kamera (s. Abb. 5.2). Eine Rotationsachse  $\alpha_R$  zur Rotationsmatrix  $\Omega$  ist rotationsinvariant bezüglich  $\Omega$ . Somit ist  $\alpha_R$  der Eigenwert zu  $\Omega$  mit  $\Omega\alpha_R = \alpha_R$ . Mit Hilfe der Darstellung einer Rotationsmatrix durch ein Quaternion  $\mathbf{q} = [q_w, q_x, q_y, q_z]$ , lässt sich die Rotationsachse  $\alpha_R = [a_x, a_y, a_z]^T$  anhand des rotatorischen Bewegungsanteils einfach bestimmen:

$$a_* = \frac{q_*}{\|\mathbf{q}\| \sin(\frac{\phi}{2})} \quad \text{mit } * \in \{x, y, z\} \text{ und } \phi = 2 \arccos\left(\frac{q_w}{\|\mathbf{q}\|}\right) \quad (5.9)$$

Unter der Annahme einer rein planaren Bewegung der Plattform mit fest montierten Kameras, rotiert jede Kamera  $\mathcal{M}_i$  prinzipiell zu jedem Zeitpunkt um dieselbe, jeweilige Rotationsachse  $\alpha_{R_i}$  (s. Abb. 5.2). Bei bekannter Bewegungshistorie der Länge  $\mathfrak{H}$  kann diese zur Schätzung von  $\alpha_{R_i}$  genutzt werden, um alle  $\mathfrak{H}$  Teilachsen von  $\mathcal{M}_i$  zu mitteln. Der Vorteil einer Mittelung liegt darin, dass sich über der Historie eine Varianz berechnen lässt, die Aufschluss über die Güte der Schätzung der Achse gibt. So kann es beispielsweise bei starker Bremsung oder durch Fahrbahnunebenheiten dazu kommen, dass die Annahme der planaren Bewegung verletzt wird. Eine solche Verletzung drückt sich dann durch eine große Varianz der Rotationsachse aus.



**Abbildung 5.2:** Bestimmung der Rotationsachse aus einer planaren, rotatorischen Bewegung.

### 5.3.2 Selektion von Punkten der Bodenebene

Bevor die Bodenebene mittels nichtlinearer Optimierung berechnet werden kann, ist es wichtig, die Menge der 3D-Punkte zu bestimmen, die für die Optimierung verwendet werden sollen. Im Folgenden sind dies 3D-Punkte, die in der zu schätzenden Bodenebene liegen. Zur Bestimmung dieser Punktmenge, der *Inlier*, bietet sich das RANSAC Verfahren an [Fis81]. Die gesamte Punktmenge  $\mathcal{S}$  aller 3D-Punkte ist das Ergebnis des Bündelausgleichs. Ausreißer sind demzufolge alle Punkte, die nicht in der Bodenebene liegen.

Gesucht ist also die Menge der Bodenpunkte  $\underline{\mathcal{S}} \subseteq \mathcal{S}$ . Um die Menge vorab einzuschränken, werden nur Punkte für das RANSAC Verfahren berücksichtigt, die in einer Ebene orthogonal zur Rotationsachse  $\alpha_R$  der Kamera liegen (siehe Abschnitt 5.3.1). Hierfür werden zufällig drei Punkte  $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$  aus  $\mathcal{S}$  gewählt und die Normale  $\mathbf{n}$  berechnet

$$\mathbf{n} = [\mathbf{X}_1 - \mathbf{X}_2] \times [\mathbf{X}_1 - \mathbf{X}_3]. \quad (5.10)$$

Ist der Winkel zwischen der Normalen und der Rotationsachse klein genug, d. h. ist die Bedingung

$$\arccos\left(\frac{\alpha_R^T \mathbf{n}}{\|\mathbf{n}\| \cdot \|\alpha_R\|}\right) < \Lambda_\varphi \quad (5.11)$$

für einen Schwellwert  $\Lambda_\varphi > 0$  erfüllt, so wird geprüft, wie viele Punkte aus  $\mathcal{S}$  in der durch das Tripel repräsentierten Ebene liegen. Hierfür wird gemäß der Hesse-Normalform der Abstand des Punktes zur Ebene bestimmt. Ein Punkt  $\mathbf{X}$  gilt als *Inlier*, sofern der Betrag des Skalarprodukts  $d = |\mathbf{X}^T \mathbf{n}|$  kleiner als ein Schwellwert  $\Lambda_d > 0$  ist. Die Menge der *Inlier* der  $\xi$ -ten Iteration sei mit  $\mathcal{S}_\xi$  bezeichnet. Dieses Vorgehen wird mit  $\mathfrak{P}_{\mathcal{S}}$  zufällig gewählten Tripeln wiederholt. Die gesuchte Menge an Bodenpunkten ist die Konsensusmenge  $\mathcal{S}_\xi$  mit der größten Kardinalität

$$\underline{\mathcal{S}} = \operatorname{argmax}_{\mathcal{S}_{\xi=1,\dots,\mathfrak{P}_{\mathcal{S}}}} (|\mathcal{S}_\xi|). \quad (5.12)$$

und  $\mathfrak{P}_{\mathcal{G}} = |\underline{\mathcal{S}}|$ .

### 5.3.3 Parameterbestimmung mittels nichtlinearer Optimierung

Für ein Modul  $\mathcal{M}_i$  sei eine Menge  $\underline{\mathcal{S}}$  an komplanaren 3D-Punkten  $\mathbf{X}_{p_i} = [X_{p_i}, Y_{p_i}, Z_{p_i}]^T \in \underline{\mathcal{S}}$  gegeben mit  $p_i = 1, \dots, \mathfrak{P}_{\mathcal{G}} = |\underline{\mathcal{S}}|$ . Diese Punktmenge ist das Ergebnis der robusten Merkmalsauswahl von Abschnitt 5.3.2. Dann ist die Euklidische Transformation  $\mathbf{G}_i \equiv \mathbf{G}_i$  mit Rotation  $\mathbf{R}_{\mathbf{G}_i}(\alpha_i, 0, \gamma_i)$  und Translation  $\mathbf{t}_{\mathbf{G}} = [0, h_{y_i}, 0]^T$  gesucht, mit der alle Punkte der Bodenebene vom Kamera- in das Bodenkoordinatensystem überführt werden können. D. h. die resultierenden Höhen- (d. h.  $y$ -Koordinaten) der transformierten Punkte sind gleich Null. Die Punktmenge lässt sich somit in einem reduzierten, zweidimensionalen Koordinatensystem beschreiben. Um dies zu erreichen sind lediglich drei Parameter notwendig, nämlich zwei Eulerwinkel  $\alpha_i, \gamma_i$  um die  $X$ - bzw.  $Z$ -Achse, und einen translatorischen Parameter  $h_{y_i}$ , der die Höhe des Kamerakoordinatensystems von  $\mathcal{M}_i$  beschreibt. Bei Unterdrückung des Modulindexes  $i$  soll also gelten

$$\mathbf{R}_{\mathbf{G}} \mathbf{X}_p + \mathbf{t}_{\mathbf{G}} = \begin{bmatrix} \underline{X}_p \\ 0 \\ \underline{Z}_p \end{bmatrix}, \underline{X}_p, \underline{Z}_p \in \mathbb{R}, p = 1, \dots, \mathfrak{P}_{\mathcal{G}}. \quad (5.13)$$

Aus der zweiten Zeile des Gleichungssystems 5.13 lässt sich eine Bedigungsgleichung für eine LM-Optimierung (Anhang 2.3) formulieren

$$\mathbf{h}_G(\mathbf{G}, \mathbf{X}_p) = \left[ \underbrace{\sin(\gamma)}_{g_1} X_p + \underbrace{\cos(\alpha)\cos(\gamma)}_{g_2} Y_p - \underbrace{\sin(\alpha)\cos(\gamma)}_{g_3} Z_p + \underbrace{h_y}_{g_4} \right]^2 \stackrel{!}{=} 0. \quad (5.14)$$

Durch Minimieren von Gl. 5.14 über alle  $\mathfrak{P}_G$  Bodenpunkte  $\mathbf{X}_{p=1,\dots,\mathfrak{P}_G}$ , erhält man letztlich die gesuchten drei Parameter. Aus Gl. 5.14 lässt sich gleichsam die Ebene in parametrisierter, homogener Form als Vektor  $\mathbf{g} = (g_1, g_2, g_3, g_4)^T$  darstellen mit der zugehörigen Ebenennormalengleichung

$$\mathbf{g}^T \tilde{\mathbf{X}}_p = 0. \quad (5.15)$$

### 5.3.4 Anmerkungen zur Abhängigkeit des Höhenparameters von der Geschwindigkeit

Der absolute Geschwindigkeitsbetrag spielt eine wichtige Rolle bei der Schätzung der Bodenebene. Man stelle sich ein Fahrzeug vor mit je einer Kamera an seiner linken und rechten Seite, senkrecht zur Fahrtrichtung blickend. Des weiteren fahre das Fahrzeug zum Zeitpunkt  $k_1$  konstant mit der translatorischen Geschwindigkeit  $\|\mathbf{v}_1\| = v_1$  geradeaus ( $\boldsymbol{\Omega}_1 = \mathbf{I}$ )<sup>21</sup>.

Sei weiter angenommen, dass dasselbe Fahrzeug zu einem anderen Zeitpunkt  $k_2$  dieselbe Strecke in identischer Art und Weise fahre, jedoch nun mit einer konstanten Geschwindigkeit  $\|\mathbf{v}_2\| = v_2 < v_1$ . Wären die Geschwindigkeitsbeträge  $v_1$  und  $v_2$  in ihrer absoluten Größe bekannt, z. B. aus Odometriesensoren, so ließe sich mit Hilfe des Bündelausgleichs die Szene (d. h. einzelne Szenenpunkte) zu beiden Zeitpunkten absolut, metrisch rekonstruieren, d. h. zu beiden Zeitpunkten wäre die rekonstruierte Szene identisch.

Sind die Geschwindigkeitsbeträge aufgrund der Beobachtungen in Abschnitt 4.3 nur

---

<sup>21</sup> An dieser Stelle sei darauf hingewiesen, dass nach der Definition in Gl. 2.12 die Geschwindigkeit einer Kamera deren Verschiebung  $\|\mathbf{v}\| = v$  entspricht.

skaliert bekannt, so dass  $v_1 = v_2 = \|\mathbf{v}_1^s\| = \|\mathbf{v}_2^s\| = \left\| \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \right\| = \left\| \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} \right\| = 1$ , so kann auch die Szene zu beiden Zeitpunkten nur noch bis auf einen skalaren Faktor metrisch rekonstruiert werden. Im Beispiel würde dies bedeuten, dass zum Zeitpunkt  $k_2$  dieselbe Szene entsprechend größer erschiene, als an  $k_1$ .

Dieser Sachverhalt lässt sich an dem vereinfachten Beispiel auch mathematisch zeigen. In Anlehnung an das Lochkameramodell aus Abschnitt 2.1 ergibt sich für senkrecht zur Bewegungsrichtung stehende Kameras die Tiefe  $\lambda$  eines Szenenpunktes aus dem Strahlensatz in Abhängigkeit der Brennweite  $f$  und dem Geschwindigkeitsbetrag gemäß

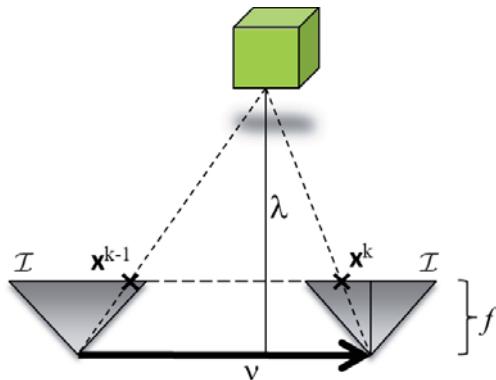
$$\lambda_1 = \frac{v_1 \cdot f}{D_1} \text{ und } \lambda_2 = \frac{v_2 \cdot f}{D_2} \quad (5.16)$$

mit der Disparität der Bildpunktcorrespondenzen  $D_1 = |\mathbf{x}_1^k - \mathbf{x}_1^{k-1}|$  und  $D_2 = |\mathbf{x}_2^k - \mathbf{x}_2^{k-1}|$  (s. Abb. 5.3). Bei konstanter Brennweite  $f$  und skalierter Geschwindigkeit  $v_1 = v_2 = 1$  hängt der Abstand eines Szenenpunktes von der Kamera antiproportional von der Disparität ab. Bleibt der Abstand der Kameras zu der Szene gleich, so ergibt sich bei geringerer Geschwindigkeit eine kleinere Disparität für dieselben Szenenpunkte. Also gilt  $D_2 < D_1$  und damit  $\lambda_2 > \lambda_1$ .

Dies hat erheblich Einfluss auf die Schätzung der Bodenebene. Denn aus dem obigen Beispiel wird ersichtlich, dass auch der Abstand der Kameras zum Boden, und damit der Höhenparameter  $h_{y_i}$  je Modul  $\mathcal{M}_i$ , wesentlich von dem Geschwindigkeitsbetrag abhängt. Ist die Plattformgeschwindigkeit nur skaliert und nicht absolut bekannt, so ergeben sich aus der Bodenschätzung in Abschnitt 5.3.3 bei unterschiedlichen Fahrgeschwindigkeiten unterschiedliche Punktkoordinaten und damit auch unterschiedliche Höhenparameter. Da aber der Abstand der Kameras zur Bodenebene in Wahrheit (mehr oder minder) konstant bleibt, führt die Schätzung der Transformationsparameter bei unterschiedlichen Geschwindigkeiten zu Inkonsistenzen und damit zu Folgefehlern bei der späteren Zusammenführung der Teiltransformationen in Kapitel 7.

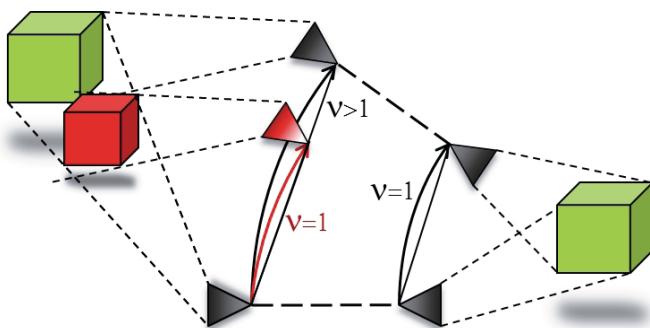
Ausgehend von Abschnitt 5.3 ist die Bodenschätzung nur bei rein translatorischen

Bewegungen sinnvoll, da sich die einzelnen Kameramodule bei Kurvenfahrten i. A. mit unterschiedlichen Geschwindigkeiten bewegen. Aber auch bei Kurvenfahrten würden, bedingt durch die Skalierung der translatorischen Geschwindigkeiten, gleich große Objekte in verschiedenen Kameras unterschiedlich groß rekonstruiert. Abb. 5.4 veranschaulicht den Einfluss skalierter Geschwindigkeiten bei Kurvenfahrten auf die Rekonstruktion der Szene.



**Abbildung 5.3:** Triangulationsschema für rektifizierte Kameras.

Zusammenfassend gilt, dass sich das Fahrzeug für eine sinnvolle Schätzung der Bodenebene zum einen in einer Geradeausfahrt befinden muss, zum anderen muss der absolute Geschwindigkeitsbetrag bekannt sein (ggf. aus anderen Sensorquellen).



**Abbildung 5.4:** Veranschaulichung des Unterschieds zwischen skaliert (rot) und unskaliert (schwarz) translatorischer Bewegung in Kurven an einem Zwei-Kamera-Beispiel. Obwohl die Objekte, die beide Kameras sehen, identisch sind, erscheint das Objekt der äuferen Kamera bedingt durch die Skalierung kleiner (roter Würfel).



# 6

---

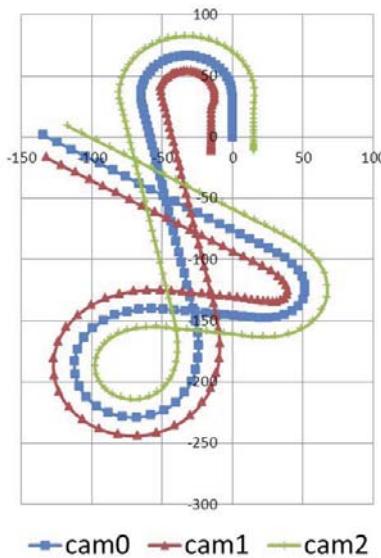
## Der Bewegungsausgleich

---

Wie in Abschnitt 4.3 bereits erwähnt, hat sich in der Praxis die rein bildgestützte Bewegungsschätzung über längere Zeiträume hinweg hinsichtlich des Geschwindigkeitsbeitrags als instabil erwiesen. D. h. wenngleich die rotatorische Bewegung jedes Kameramoduls hinreichend genau geschätzt werden konnte, so kann es doch vorkommen, dass die Geschwindigkeitsbeträge der einzelnen Module derart stark divergieren, dass eine Kalibrierung gemäß der Basisgleichung 4.4 nicht mehr möglich ist. Die resultierende Untauglichkeit der Bewegungsschätzungen liegt in der Aussage der Basisgleichung begründet. Die Bedingung, die implizit gefordert wird, besagt u. a. , dass der Abstand zwischen je zwei Kameramodulen während der gesamten Fahrt konstant bleibt. D. h. an  $\tilde{\mathfrak{H}}$  gemessenen Zeitpunkten soll die Lage der Kameras zueinander unverändert sein. Abb. 6.1 zeigt beispielhaft die Trajektorien dreier Module zur Verdeutlichung der unterschiedlich starken Translationsbeträge in Kurven.

Aus den in Abschnitt 4.3 aufgeführten Gründen wird Gl. 4.4 um einen Skalierungsfaktor  $\kappa$  erweitert und die Geschwindigkeitsvektoren auf Länge 1 normiert.

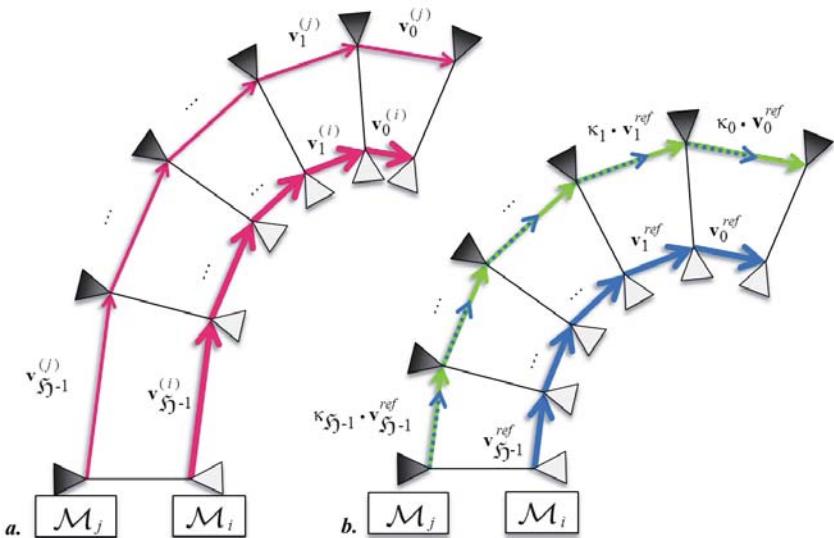
$$[\mathbf{I} - \boldsymbol{\Omega}_{i_t}] \mathbf{t}_{ij} + \kappa_{ij_t} \mathbf{R}_{ij} \mathbf{v}_{j_t}^s - \mathbf{v}_{i_t}^s = \mathbf{0} \quad (6.1)$$



**Abbildung 6.1:** Trajektorien einer simulierten Drei-Kamera-Anordnung zur Veranschaulichung der unterschiedlich großen Translationsbeträge in den Kurven.

für zwei Module  $\mathcal{M}_i, \mathcal{M}_j$  mit  $t = 1, \dots, \mathfrak{H}$ .  $\mathbf{v}^s$  bezeichnet hierbei den normierten Vektor von  $\mathbf{v}$  mit  $\mathbf{v}^s = \frac{\mathbf{v}}{\|\mathbf{v}\|}$  und  $\|\mathbf{v}^s\| = 1$ . Dem Problem der nur skaliert bekannten Geschwindigkeiten wird derart beigekommen, dass der Geschwindigkeitsbetrag zusammen mit den extrinsischen Parametern geschätzt wird. Der Parameterraum vergrößert sich demnach abhängig von der Größe  $\mathfrak{H}$  der betrachteten Bewegungshistorie. Dies bedeutet, dass der relative Geschwindigkeitsunterschied zwischen den beiden Modulen mitgeschätzt wird. Es wird also die Frage beantwortet, wie viel schneller oder langsamer sich die äußere relativ zur inneren Kamera bzw. die innere sich relativ zur äußeren Kamera zu jedem Zeitpunkt  $t$  innerhalb der betrachteten Historie bewegt.

Wie beim allgemeinen, so ist auch beim erweiterten Kalibriermodell eine rotatorische Bewegung  $\boldsymbol{\Omega}_i \neq \mathbf{I}$  aller Module  $\mathcal{M}_i$  Voraussetzung für die Bestimbarkeit von  $\mathbf{t}_{ij}$ . Man



**Abbildung 6.2:** Schematische Darstellung des Bewegungsausgleichs für skalierte Bewegungen. *a.* Tatsächlich zurückgelegte Trajektorie zweier Kameras  $\mathcal{M}_i, \mathcal{M}_j$  mit unterschiedlich großen Translationsvektoren der Bewegungstransformationen. *b.* Skalierte Translationsvektoren vor (blau) und nach (grün) dem Bewegungsausgleich mit den ermittelten Skalierungsparametern  $\kappa_t$ ,  $t = 0, \dots, \mathfrak{H} - 1$ . Modul  $\mathcal{M}_i$  ist hierbei das Referenzmodul, d. h. die neuen Translationsvektoren für  $\mathcal{M}_j$  ergeben sich zu  $\kappa_t \cdot v_t^{ref}$ , wobei  $v_t^{ref} = \frac{v_t^{(i)}}{\|v_t^{(i)}\|}$ .

muss sich an dieser Stelle bewusst machen, dass durch die Skalierung der Geschwindigkeiten auch der absolute Betrag von  $\mathbf{t}_{ij}$  verloren geht. D. h. die relative Anordnung der Kameras, d. h. aller  $\mathbf{t}_{ij}$ , kann lediglich bis auf einen skalaren Faktor rekonstruiert werden. Allerdings könnte hier die Nutzung eines Odometriesensors helfen, den Absolutbetrag zumindest anzunähern. Darauf wird an späterer Stelle näher eingegangen (Abschnitt 6.2).

Im erweiterten Kalibriermodell wird eine Reduktion von der dreidimensionalen Szene in die zweidimensionale Grundebene vorgenommen, so dass sich hier für den

Bewegungsausgleich die Anzahl der zu schätzenden extrinsischen Parameter auf drei verringert. Dies ist mit Hilfe der Rotationsachse auch ohne explizite Kenntnis der translatorischen Höhe der Kameras relativ zur Bodenebene möglich (siehe Anhang A.5).

Für die weiteren Ausführungen in diesem Abschnitt wird auf die Modulindizes » $i j$ « bei den extrinsischen Parametern  $\underline{C}$  mit  $\underline{\mathbf{t}}_{xij} = [\underline{x}_{ij}, \underline{z}_{ij}]^T$  und  $\underline{\mathbf{R}}_{ij}$  mit  $\beta_{ij}$ , sowie den Skalierungsparametern  $\kappa_{ij_t}$  zugunsten einer besseren Lesbarkeit verzichtet.

Die extrinsische Transformation  $\underline{C}$  zwischen zwei Modulen  $\mathcal{M}_i, \mathcal{M}_j$  innerhalb der Grundebene ist parametriert durch zwei translatorische Parameter  $\underline{x}, \underline{z}$  sowie einen Eulerwinkel  $\beta$ . Zur Berechnung werden die Bewegungstransformationen  $\underline{\mathbf{M}}_*$  herangezogen mit den Parametern  $\underline{v}_{xi}, \underline{v}_{zi}, \psi_i$  und  $\underline{v}_{xj}, \underline{v}_{zj}, \psi_j$  (siehe Anhang A.5). Prinzipiell lässt sich der Bewegungsausgleich aber auch zur Kalibrierung mittels 3D-Bewegungen einsetzen.

Um das Problem noch weiter zu vereinfachen, kann die berechnete Orientierung zwischen den Modulen aus Abschnitt 5.2.1 verwendet werden, um die Rotation  $\mathbf{R}$  während des Schätzprozesses zu fixieren. Die Anzahl der Freiheitsgrade reduziert sich dadurch zusätzlich.

Im Folgenden wird dem erweiterten Kalibriermodell Rechnung getragen und Gl. 6.1 auf den zweidimensionalen Fall übertragen, d. h. die Bestimmung der relativen Lage der Kameras erfolgt nun in der Grundebene. Mittels Gl. 6.1 lässt sich ein unterbestimmtes lineares Gleichungssystem formulieren

$$\underline{\mathbf{y}}_{it}^s = \begin{bmatrix} [\mathbf{I} - \underline{\boldsymbol{\Omega}}_{it}] & \underline{\mathbf{R}}\underline{\mathbf{v}}_{jt}^s \end{bmatrix} \begin{bmatrix} \underline{\mathbf{t}} \\ \kappa_t \end{bmatrix} \quad (6.2)$$

mit

$$\underline{\mathbf{t}} = [\underline{t}_x, \underline{t}_z]^T \quad (6.3)$$

$$\underline{\mathbf{R}} = \begin{bmatrix} \cos(\beta) & \sin(\beta) \\ -\sin(\beta) & \cos(\beta) \end{bmatrix} \quad (6.4)$$

$$\underline{\mathbf{v}}_{i_t} = [\underline{v}_{x_{i_t}}, \underline{v}_{z_{i_t}}]^T \quad (6.5)$$

$$\underline{\Omega}_{i_t} = \begin{bmatrix} \cos(\psi_{i_t}) & \sin(\psi_{i_t}) \\ -\sin(\psi_{i_t}) & \cos(\psi_{i_t}) \end{bmatrix}. \quad (6.6)$$

Aus Gl. 6.2 lassen sich zwei Lemmata ableiten:

**Lemma 5** *Zur eindeutigen Bestimmung sowohl der extrinsischen Translation  $\underline{\mathbf{t}}$  als auch des Faktors  $\kappa_t$  aus Gl. 6.2 zwischen zwei Modulen  $\mathcal{M}_i, \mathcal{M}_j$  sind mindestens zwei Bewegungspaare je Module notwendig. D. h. ein Bewegungsblock muss mindestens  $\Gamma > 1$  Bewegungstransformationen umfassen.*

**Beweis** Sei  $\underline{\mathbf{R}}(\beta)$  bekannt und  $\underline{\Omega}_i := [\mathbf{w}_1 \ \mathbf{w}_2] \in \mathbb{R}^{2 \times 2}$ . Die Koeffizientenmatrix

$$\mathbf{A} = [\mathbf{e}_1 - \mathbf{w}_1 \ \mathbf{e}_2 - \mathbf{w}_2 \ \underline{\mathbf{R}} \mathbf{v}_{j_1}^s]_{2 \times 3} \quad (6.7)$$

zur Bestimmung des Lösungsvektors  $\mathbf{x} = [\underline{t}_x, \underline{t}_z, \kappa_t]^T$  mit Hilfe des LGSs  $\mathbf{Ax} = \underline{\mathbf{v}}_{i_t}^s$  besitzt höchstens Rang 2 und benötigt daher mindestens zwei Bewegungspaare, um ein überbestimmtes LGS mit einem Rang>2 zur Berechnung der Lösung für  $\mathbf{x}$  formulieren zu können. ■

Lemma 5 zeigt die Notwendigkeit von zeitlich zusammenhängenden Bewegungsblöcken der Größe  $\Gamma > 1$ , innerhalb derer der Skalierungsfaktor  $\kappa_t = \kappa_{t0} = \dots = \kappa_{t\Gamma-1}$  zwischen den Modulen als hinreichend konstant angenommen werden kann. D. h. die relative Geschwindigkeit innerhalb eines Blocks muss bekannt oder ihre Varianz hinreichend klein sein.

Hinsichtlich der Bestimmung der Translation und des Skalierungsfaktors ergibt sich in Anlehnung an Lemma 4 eine weitere Restriktion:

**Lemma 6** Zur eindeutigen Bestimmung sowohl der extrinsischen Translation  $\underline{\mathbf{t}}$  als auch des Faktors  $\kappa_t$  aus Gl. 6.2 zwischen zwei Modulen  $\mathcal{M}_i, \mathcal{M}_j$  dürfen die Bewegungen  $\underline{\mathcal{M}}_{i,tr}, \underline{\mathcal{M}}_{j,tr}, r = 0, \dots, \Gamma - 1$ , der Module  $\mathcal{M}_i, \mathcal{M}_j$  innerhalb eines Bewegungsblocks keiner Kreisfahrt folgen.

**BEWEIS** Eine reine Kreisfahrt zeichnet sich durch einen zeitlich konstanten Rotationswinkel  $\psi_{i,tr} = \psi_{j,tr}$  aus. Dann gilt aufgrund der Normierung  $\underline{\mathbf{Rv}}_j^s = \underline{\mathbf{v}}_i^s$  und Gl. 6.2 reformuliert sich gemäß

$$\begin{bmatrix} [\mathbf{I} - \underline{\boldsymbol{\Omega}}_{i,t}] & \underline{\mathbf{v}}_{i,t}^s \end{bmatrix} \begin{bmatrix} \underline{\mathbf{t}} \\ \kappa_t \end{bmatrix} = \underline{\mathbf{v}}_{i,t}^s. \quad (6.8)$$

In diesem Fall ergibt sich ohne zusätzliche Bedingungen stets die triviale Lösung  $[\mathbf{0}^T, 1]^T$ . ■

Die Erfüllung von Lemma 6 lässt sich durch eine einfache Überprüfung der Winkelgeschwindigkeiten erreichen. Es wird also gefordert, dass innerhalb eines Bewegungsblocks die Winkelgeschwindigkeiten eine Mindestvarianz aufweisen müssen.

Aus Lemma 5 ergibt sich die Notwendigkeit von Bewegungsblöcken, innerhalb derer die relativen Geschwindigkeiten bekannt sind. Dem kommt die Bewegungsschätzung mit Hilfe des Bündelausgleichs zu Gute. Der Bündelausgleich verwendet  $\Gamma + 1$  aufeinander folgende Bilder, um mit den darin detektierten Punktkorrespondenzen  $\Gamma$  Euklidische Bewegungstransformationen zu berechnen. D. h. die relativen Geschwindigkeiten  $\frac{\underline{\mathbf{v}}_{i,tr}}{\|\underline{\mathbf{v}}_{i,t,0}\|}, r = 0, \dots, \Gamma - 1$  für ein Modul  $\mathcal{M}_i$  bezüglich des aktuellen Geschwindigkeitsbetrags  $\|\underline{\mathbf{v}}_{i,t,0}\|$  sind also bekannt<sup>22</sup>. Somit ist die Grundlage gegeben, um einen Bewegungsausgleich, also die simultane Berechnung der extrinsischen Kameraparameter und deren Bewegungsskalierung, durchzuführen.

---

<sup>22</sup> Die Notwendigkeit der Skalierung der Geschwindigkeitsbeträge wie in Abschnitt 4.3 ergab sich aus der Fehlerfortpflanzung bei der Schätzung des Geschwindigkeitsbetrags relativ zur Schätzung des vorangegangenen Zeitpunkts. Die Schätzung der relativen Teilbewegungen innerhalb des Bündelausgleichs sind hiervon nicht betroffen.

## 6.1 Effiziente Berechnung des dünnen Bewegungsausgleichs

Der Bewegungsausgleich lässt sich mit Hilfe einer LM-Optimierung berechnen. Die besondere Struktur des Problems lässt eine effizientere Berechnung zu, ähnlich zu der Implementierung des dünnen Bündelausgleichs von Lourakis et al. [Lou09]. Darum wird hier im Folgenden in Anlehnung an den dünnen Bündelausgleich in [Lou09] vom *dünnen Bewegungsausgleich* gesprochen (SMA)<sup>23</sup>.

Das Problem des Bewegungsausgleichs zwischen zwei Modulen  $\mathcal{M}_i, \mathcal{M}_j$  ist parametrisiert durch den gesuchten Zustand  $\mathbf{x}$  und den Messvektor  $\mathbf{z}$

$$\mathbf{x} = [\underline{\mathbf{t}}^T, \beta, \boldsymbol{\kappa}^T]^T \quad (6.9)$$

$$\mathbf{z} = [..., \underline{\mathbf{v}}_{i_{tr}}^s, \psi_{i_{tr}}, \underline{\mathbf{v}}_{j_{tr}}^s, \psi_{j_{tr}}, ...]^T \quad (6.10)$$

mit  $\underline{\mathbf{t}}_C = [\underline{t}_x, \underline{t}_z]^T$  und  $\boldsymbol{\kappa} = [..., \kappa_t, ...]^T$ ,  $t = 0, \dots, \mathfrak{H}-1$ ,  $r = 0, \dots, \Gamma-1$ .  $\underline{\mathbf{v}}_{i_{tr}}^s, \psi_{i_{tr}}$  und  $\underline{\mathbf{v}}_{j_{tr}}^s, \psi_{j_{tr}}$  sind die Parameter der Bewegungs transformationen  $\underline{\mathbf{M}}_{i_{tr}}^s$  bzw.  $\underline{\mathbf{M}}_{j_{tr}}^s$ . Die Transformationen beziehen sich hier im Kontext des erweiterten Kalibriermodells auf den planaren Fall, können aber prinzipiell auch für den 3D-Fall übertragen werden.  $\underline{\mathbf{M}}_{i_{tr}}^s$  ist hierbei die Bewegung über Grund von  $\mathcal{M}_i$  mit normiertem Geschwindigkeitsvektor  $\underline{\mathbf{v}}_{i_{tr}}^s$ , und zwar die  $r$ -te Transformation im  $t$ -ten Historienblock.

Aus Gl. 6.1 ergibt sich die Funktion

$$\underline{\mathbf{h}}_C(\mathbf{x}, \mathbf{z}) = \begin{bmatrix} \vdots \\ h_{C_{tr}}(\mathbf{x}, \mathbf{z}) \\ \vdots \end{bmatrix} =: \boldsymbol{\epsilon} \quad (6.11)$$

---

<sup>23</sup> engl. *Sparse Motion Adjustment*

mit

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \vdots \\ \varepsilon_{tr} \\ \vdots \end{bmatrix}_{(\mathfrak{H} \cdot \Gamma \times 1)} = \begin{bmatrix} \vdots \\ \boldsymbol{\varepsilon}_t \\ \vdots \end{bmatrix}, \quad t = 0, \dots, \mathfrak{H}-1, \quad r = 0, \dots, \Gamma-1, \quad (6.12)$$

und

$$h_{\underline{C}_{tr}}(\mathbf{x}, \mathbf{z}) = \left\| [\mathbf{I} - \underline{\Omega}_{it_r}] \underline{\mathbf{t}} + \kappa_t \underline{\mathbf{R}} \underline{\mathbf{v}}_{j_{tr}}^s - \underline{\mathbf{v}}_{i_{tr}}^s \right\|^2 \stackrel{!}{=} 0. \quad (6.13)$$

An dieser Stelle greife man auf die LM-Normalengleichung 2.20 zurück

$$[\mathbf{J}^T \mathbf{J} - \rho \mathbf{I}] \begin{bmatrix} \boldsymbol{\delta}_{\underline{C}} \\ \boldsymbol{\delta}_{\kappa} \end{bmatrix} = \mathbf{J}^T \boldsymbol{\varepsilon}. \quad (6.14)$$

Die Jacobi-Matrix  $\frac{\partial h_{\underline{C}}}{\partial \underline{x}}$  hat dann die Form

$$\mathbf{J} = \begin{bmatrix} \mathbf{B}_0 & \mathbf{s}_1 & \mathbf{0} & \cdots & \cdots & \cdots & \cdots & \cdots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \ddots & & & & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \mathbf{B}_t & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{s}_t & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & & & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{B}_{\mathfrak{H}-1} & \mathbf{0} & \mathbf{0} & \cdots & \cdots & \cdots & \cdots & \cdots & \mathbf{s}_{\mathfrak{H}} \end{bmatrix} \quad (6.15)$$

mit den Matrizen der partiellen Ableitungen

$$\mathbf{B}_t = \begin{bmatrix} \frac{\partial h_{\underline{C}_t 0}}{\partial \underline{t}} & \frac{\partial h_{\underline{C}_t 0}}{\partial \beta} \\ \vdots & \vdots \\ \frac{\partial h_{\underline{C}_t \Gamma-1}}{\partial \underline{t}} & \frac{\partial h_{\underline{C}_t \Gamma-1}}{\partial \beta} \end{bmatrix} \quad \text{und} \quad \mathbf{s}_t = \begin{bmatrix} \frac{\partial h_{\underline{C}_t 0}}{\partial \kappa_t} \\ \vdots \\ \frac{\partial h_{\underline{C}_t \Gamma-1}}{\partial \kappa_t} \end{bmatrix}. \quad (6.16)$$

Der linksseitige Term der Normalengleichung hat dann die Gestalt

$$\begin{aligned} [\mathbf{J}^T \mathbf{J} - \rho \mathbf{I}] &= \left[ \begin{array}{c|ccc} \left[ \sum_{t=0}^{j-1} \mathbf{B}_t^T \mathbf{B}_t \right] - \rho \mathbf{I} & \mathbf{B}_0^T \mathbf{s}_0 & \cdots & \mathbf{B}_{j-1}^T \mathbf{s}_{j-1} \\ \mathbf{s}_0^T \mathbf{B}_0 & \mathbf{s}_0^T \mathbf{s}_0 - \rho & & \\ \vdots & & \ddots & 0 \\ \mathbf{s}_{j-1}^T \mathbf{B}_{j-1} & 0 & & \mathbf{s}_{j-1}^T \mathbf{s}_{j-1} - \rho \end{array} \right] \quad (6.17) \\ &=: \begin{bmatrix} \mathbf{U} & \mathbf{W} \\ \mathbf{W}^T & \mathbf{V} \end{bmatrix}. \end{aligned}$$

Damit lässt sich die Normalengleichung darstellen als

$$\begin{bmatrix} \mathbf{U} - \mathbf{WV}^{-1}\mathbf{W}^T & \mathbf{0} \\ \mathbf{W}^T & \mathbf{V} \end{bmatrix} \begin{bmatrix} \boldsymbol{\delta}_C \\ \boldsymbol{\delta}_K \end{bmatrix} = \begin{bmatrix} \boldsymbol{\varepsilon}_C - \mathbf{WV}^{-1}\boldsymbol{\varepsilon}_K \\ \boldsymbol{\varepsilon}_K \end{bmatrix} \quad (6.18)$$

mit

$$\boldsymbol{\varepsilon}_C = \sum_{t=0}^{j-1} \mathbf{B}_t^T \boldsymbol{\varepsilon}_t \quad \text{und} \quad \boldsymbol{\varepsilon}_K = [\dots, \mathbf{s}_t^T \boldsymbol{\varepsilon}_t, \dots]^T. \quad (6.19)$$

Aufgrund der blockdiagonalen Struktur von  $\mathbf{V}$  lässt sich auch die entsprechende Inverse schnell und einfach berechnen. Man sieht in Gl. 6.18, dass sich  $\boldsymbol{\delta}_C$  unabhängig von  $\boldsymbol{\delta}_K$  lösen lässt:

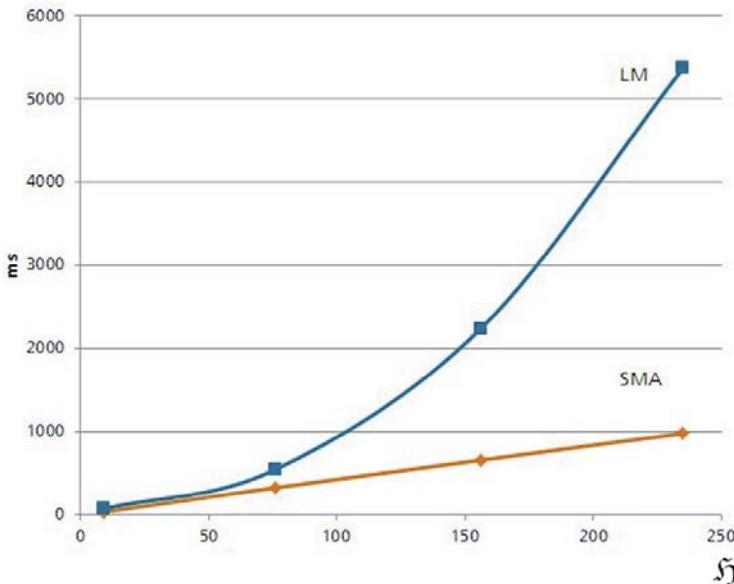
$$\mathbf{S}\boldsymbol{\delta}_C = \boldsymbol{\varepsilon}_C - \mathbf{WV}^{-1}\boldsymbol{\varepsilon}_K \quad (6.20)$$

mit  $\mathbf{S} = \mathbf{U} - \mathbf{WV}^{-1}\mathbf{W}^T$ , wobei  $\mathbf{S}$  symmetrisch und positiv definit ist, so dass Gl. 6.20 mit einer Cholesky-Zerlegung von  $\mathbf{S}$  effizient lösbar ist. Ist  $\boldsymbol{\delta}_C$  erst einmal bekannt, so lässt sich auch  $\boldsymbol{\delta}_K$  bestimmen gemäß

$$\boldsymbol{\delta}_K = \mathbf{V}^{-1}[\boldsymbol{\varepsilon}_K - \mathbf{W}^T \boldsymbol{\delta}_C]. \quad (6.21)$$

Durch diese Implementierung des SMA können signifikante Laufzeitverbesserungen gegenüber einer nicht-modifizierten Implementierung einer Standard-LM-Optimierung

erreicht werden (siehe Abb. 6.3).



**Abbildung 6.3:** Vergleich der Laufzeiten einer regulären Implementierung einer LM-Optimierung mit der effizienten Implementierung des SMA.

## 6.2 SMA relativ zu einem zentralen Bezugspunkt

Unterstellt man der mobilen Plattform die Einhaltung der nichtholonen Bewegungsbedingung aus Abschnitt 2.2, wie sie insbesondere bei zweiachsigen Fahrzeugen mit einer lenkbaren Vorderachse erwartet werden kann [Leu00], [Haa98], so lässt sich aus dem allen Kameramodulen gemeinsamen, rotatorischen Bewegungsanteil  $\psi$  und dem (skalierten) Geschwindigkeitsbetrag  $v$  die translatorische Bewegung der Plattform herleiten. Die daraus resultierende Bewegung bezieht sich auf einen Punkt  $C$ , der nicht mit einem der Kamerakoordinatensysteme zusammenfallen muss.  $C$  bezeichnet daher einen zentralen Bezugspunkts, auf den sich die relativen Bewegungen der Kameramodu-

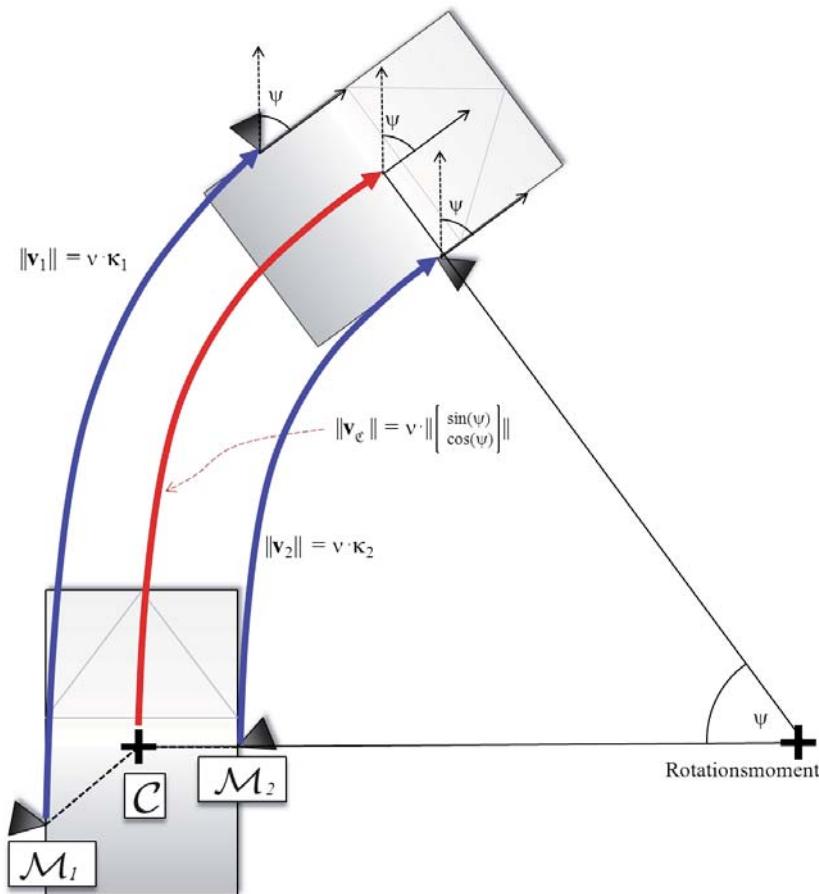
le beziehen können. Der Bewegungsausgleich erfolgt folglich nicht paarweise zwischen den Kameramodulen und relativ zu einem Referenzmodul, sondern zentrumsbezogen, d. h. relativ zu  $\mathcal{C}$ .

Als Resultat eines zentrumsbezogenen Bewegungsausgleichs erhält man eine sternförmige Transformationsanordnung mit einem zentralen Knotenpunkt, der nicht mit Koordinatenursprung eines Kameramodul zusammenfallen muss. Durch einfache Transformationsverkettungen lassen sich die zentrumsbezogenen, extrinsischen Transformationen aber auf die jeweiligen Kameramodule referenzieren, so dass das in Kapitel 5 vorgestellte Kalibriermodell o. B. d. A. beibehalten werden kann.

Bei einer Bewegung in einer 2D-Ebene ist der Winkel des rotatorischen Anteils der Bewegungstransformation bei allen Kameras gleich und ist daher – aus den kamerabasierten Bewegungsschätzungen – auch für den zentralen Bezugspunkt bekannt<sup>24</sup>. D. h. unter Kenntnis des Rotationswinkels und der Geschwindigkeit des Fahrzeugs, wie sie z. B. aus den Odometriedaten bekannt sein könnte, lässt sich sowohl die Translation des Koordinatenursprungs entsprechend Gl. 2.14 als auch die Euklidische Bewegung aller Kameramodule i. Abh. der extrinsischen Parameter zum Bezugspunkt  $\mathcal{C}$  bestimmen. In der vorliegenden Arbeit wurde der Bewegungsausgleich, insbesondere bei den experimentellen Versuchen, stets bezüglich des Bezugspunktes  $\mathcal{C}$  referenziert, da diese Implementierung gerade bei der Propagierung und Fusion mehrerer Parameterschätzungen Vorteile bietet. Dieser Sachverhalt wird im folgenden Kapitel 8 nochmals aufgegriffen.

---

<sup>24</sup> Genauer gesagt ist der Rotationswinkel für jeden beliebigen Punkt innerhalb der Ebene gleich. Die konkrete Position des zentralen Bezugspunkts ist als solche – zunächst – unbekannt.



**Abbildung 6.4:** Zentrumsbasierter Bewegungsausgleich beim erweiterten Kalibriermodell. Alle Kameramodule referenzieren sich relativ zum zentralen Bezugspunkt  $C$ . Dieser ergibt sich gemäß den nichtholonomen Bewegungsbedingungen aus den rotatorischen Bewegungen  $\psi$  der Kameras und dem (ggf. skalierten) Translationsbetrag der Bewegung.

### 6.3 Anmerkungen zur Schätzbarkeit der Kalibrierparameter

Im Fokus dieses Abschnitts steht die Frage, inwiefern die mittels SMA berechneten extrinsischen Translationsparameter, d. h. der Positionen der Kameras von den Bewegungen derselben abhängen. Bei Versuchen mit simulierten Bewegungsdaten zur Untersuchung des Zusammenhangs zwischen den Kamerapositionen und der Bewegungstrajektorie der Plattform, konnte ein Einfluss der Plattformbewegung und der Kameraposition auf das Schätzergebnis beobachtet werden. Bei gleicher Bewegung wurden verschiedene Kamerakonstellationen unterschiedlich gut rekonstruiert.

Für die Simulation wurde der Orientierungswinkel  $\beta$  zwischen den beiden Kameras als bekannt und fixiert angenommen, so dass der Fokus ausschließlich auf der Schätzung der zweidimensionalen, translatorischen Anordnung lag. Als Bewegungsart wurden hierbei verschieden starke Kurvenfahrten gewählt. Zum einen eine Bewegung mit linear steigenden Rotationswinkeln mit  $\psi(t) = t \cdot \frac{\psi_{max}}{5}$ , zum anderen eine Bewegung mit sinusförmiger Rotationskurve  $\psi(t) = \psi_{max} \cdot \sin(t \cdot \frac{\pi}{5})$ , mit  $t = 1, \dots, 5$  und  $\psi_{max} \in \{0,01; 0,02; 0,03; 0,04; 0,05\}$ .

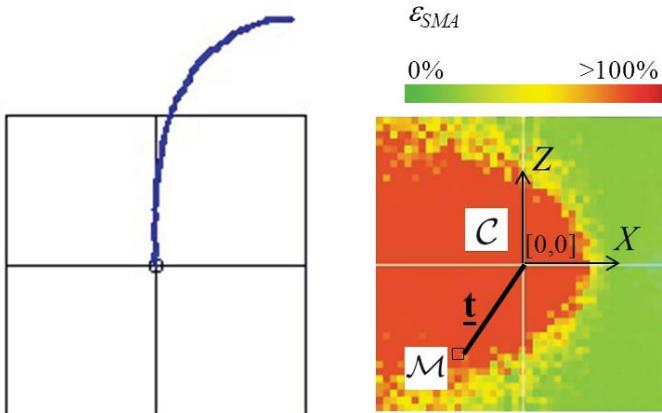
Die translatorischen Bewegungen des Bezugspunkts und der Kamera berechnen sich gemäß Gl. 2.14. Bei bekannter Grundwahrheit  $\underline{\mathbf{t}}^*$  erhält man den Restfehler als prozentuale Abweichung des Schätzwerts  $\hat{\underline{\mathbf{t}}}$  von der Grundwahrheit

$$\varepsilon_{SMA} = \frac{\|\underline{\mathbf{t}}^* - \hat{\underline{\mathbf{t}}}\|}{\|\underline{\mathbf{t}}^*\|}. \quad (6.22)$$

In den Versuchen wurde die translatorische Verschiebung  $\underline{\mathbf{t}}$  zwischen einem Kamерamodul  $\mathcal{M}$  und dem zentralen Bezugspunkt  $\mathcal{C}$  mit Hilfe des zweidimensionalen Bewegungsausgleichs geschätzt. Als Startwert für die translatorischen Parameter wurde der Nullvektor gewählt, alle Skalierungsfaktoren wurden auf 1 gesetzt, sowie  $\mathfrak{H} = 100$  und  $\Gamma = 6$ . Die simulierten zweidimensionalen Bewegungsparameter wurden mit einem additiven, normalverteilten Rauschen versehen mit  $\sigma_y = 0,05$  (entspricht 5% des Translationsbetrags) und  $\sigma_\psi = 0,0087 \text{ rad} = 0,5^\circ$ .

Abb. 6.5 zeigt solch eine Fehlerkarte für eine Bewegung mit linear steigender Rotation.

Jeder Punkt der Karte entspricht der Position der Kamera, während sich der Bezugspunkt in der Mitte der Karte befindet. In den Abb. 6.7 und 6.9 ist auffällig, dass die Karten systematische Fehlschätzungen aufzeigen. D. h. die translatorischen extrinsischen Parameter sind i. Abh. der Bewegung nicht für alle Kamerapositionen schätzbar.



**Abbildung 6.5:** Fehlerkarte aus dem Bewegungsausgleich mit simulierten Daten zwischen einem Kameramodul  $\mathcal{M}$  und dem zentralen Bezugspunkt  $\mathcal{C}$  zur Aufdeckung des Zusammenhangs zwischen Bewegung und Position der Kamera. Links: Trajektorie des Bezugspunkts. Rechts: Visualisierung des Fehlers  $\epsilon_{SMA}$  in Abhängigkeit der Translation  $\underline{t}$  nach Berechnung des Bewegungsausgleichs für jede Position des Moduls innerhalb der Karte. Rot indiziert schlecht schätzbare Bereiche; in den grünen Bereichen ließen sich die Translationsparameter gut schätzen.

Im Zuge dieser Beobachtung stellt sich die Frage, inwieweit man solche Fehlschätzungen schon vorab detektieren kann. Durch eine Analyse der Basisgleichung ließ sich ein solcher Zusammenhang bislang nicht nachweisen. Um jedoch zu einer Aussage über die Schätzbarkeit eines Parametervektors i. Abh. einer gegebenen Messung zu gelangen, ist das Konzept der Beobachtbarkeit relevant [Foe94], [Ham83]. Ein zeitdiskretes System gilt als vollständig beobachtbar, wenn sich der Anfangszustand eines Systems anhand bekannter Stellgrößen und Ausgangsmessungen eindeutig rekonstruieren lässt. Um gerade bei kontinuierlichen Schätzproblemen, wie der hier vorliegenden Kalibrierung,

statt einer rein binären Aussage ein quantitatives Gütemaß über die Schätzbarkeit eines Zustands zu erhalten, wird in [Dan07] eine Hauptkomponentenanalyse vorgeschlagen. Sei  $\underline{\mathbf{P}}_C$  die Kovarianzmatrix<sup>25</sup> zum normalverteilten Zustandsparametervektor

$$\underline{\mathbf{x}}_C = [\underline{t}_x, \underline{t}_z, \kappa_1, \dots, \kappa_{\mathfrak{H}}]^T =: [x_1, \dots, x_{\mathfrak{H}+2}]^T. \quad (6.23)$$

Dann sei  $\mathbf{D} = \text{Diag}(\eta_1, \dots, \eta_{\mathfrak{H}+2})$  die Diagonalmatrix mit den sortierten Eigenwerten  $\eta_1 \geq \eta_2 \geq \dots \geq \eta_{\mathfrak{H}+2} \geq 0$  von  $\underline{\mathbf{P}}_C$  und  $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_{\mathfrak{H}+2}]$  die Matrix mit den entsprechenden Eigenvektoren  $\mathbf{v}_{m=1, \dots, \mathfrak{H}+2} = [v_{m1}, \dots, v_{mn}, \dots, v_{m, \mathfrak{H}+2}]^T$ .  $\mathbf{V}$  ist orthogonal und wird im vorliegenden Kontext häufig auch als *Ladungsmatrix* bezeichnet [Krz00]. Es gilt also

$$\mathbf{D} = \mathbf{V}^T \underline{\mathbf{P}}_C \mathbf{V} \quad (6.24)$$

D. h. mit Hilfe der Matrix  $\mathbf{V}$  lässt sich  $\underline{\mathbf{x}}_C$  in einen Vektor  $\mathbf{y} = \mathbf{V}^T \underline{\mathbf{x}}_C$  überführen, dessen Kovarianzmatrix  $\Sigma_y$  Diagonalgestalt hat und deren Varianzen entlang der Diagonalen den Eigenwerten entsprechen, also  $\Sigma_y = \mathbf{D}$ . Der Vektor  $\mathbf{y}$  entspricht den Hauptkomponenten. Somit ist  $\eta_m$  interpretierbar als die verbleibende Unsicherheit des Zustands  $\underline{\mathbf{x}}_C$  in Richtung des Richtungsvektors  $\mathbf{v}_m$ . Die Komponenten  $v_{mn}$  der Eigenvektoren beschreiben den Beitrag des Parameters  $x_n$  entlang der Hauptachse  $\mathbf{v}_m$  mit der Unsicherheit  $\eta_m$ . Durch die Eigenwertzerlegung der Kovarianzmatrix erhält man also die Hauptachse des Zustandsraums, und damit die Dimension, in der die höchste Unsicherheit vorliegt.

Bei einem gut schätzbaren Problem sollten die Eigenwerte idealerweise gleichverteilt sein. Allerdings ergab sich bei Versuchen ein sehr dominanter größter Eigenwert. In Abb. 6.6 sind die ersten beiden Komponenten der ersten beiden Eigenvektoren jeweils mit den zwei größten Eigenwerten gewichtet. Das Ergebnis zeigt eine große Restunsicherheit insbesondere für den Parameter  $\underline{t}_x$ . Auffallend ist auch die systematische

---

25 Die Kovarianzmatrix  $\underline{\mathbf{P}}_C$  erhält man aus Gl. 2.41 oder 2.42, wobei hier die prädizierte Kovarianzmatrix (da nur eine Schätzung an einem einzelnen Zeitpunkt betrachtet wird) durch eine initiale Kovarianzmatrix  $\underline{\mathbf{P}}_C^0$  ersetzt wird.

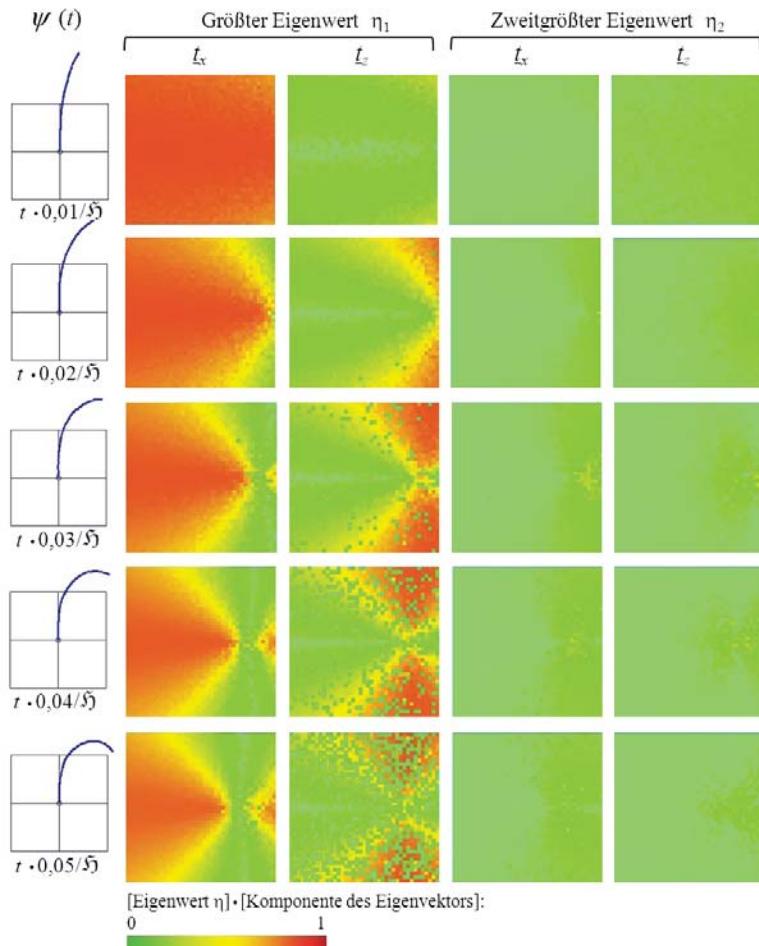
Verteilung der Unsicherheit.

Bei näherer Betrachtung einer Fehlerkarte wie in Abb. 6.8 oder 6.10 stellt man fest, dass in den Fällen mit großem Restfehler insbesondere die Schätzung der Skalare  $\kappa$  fehlerhaft ist. Die Vermutung liegt daher nahe, dass in solchen Fällen die extrinsischen Parameter  $\underline{t}_x, \underline{t}_z$  und die Skalare  $\kappa$  stark korreliert sind. Diese Vermutung wird durch eine Betrachtung der Einträge der Kovarianzmatrizen bestätigt. Somit lässt sich schon direkt aus der Kovarianzmatrix eine Größe extrahieren, die auf ein möglicherweise schlechtes Schätzergebnis hinweist. Die gemittelte Summe der Kovarianzmatrixeinträge, die die Korrelation der Translation und der Skalierungsfaktoren modellieren,

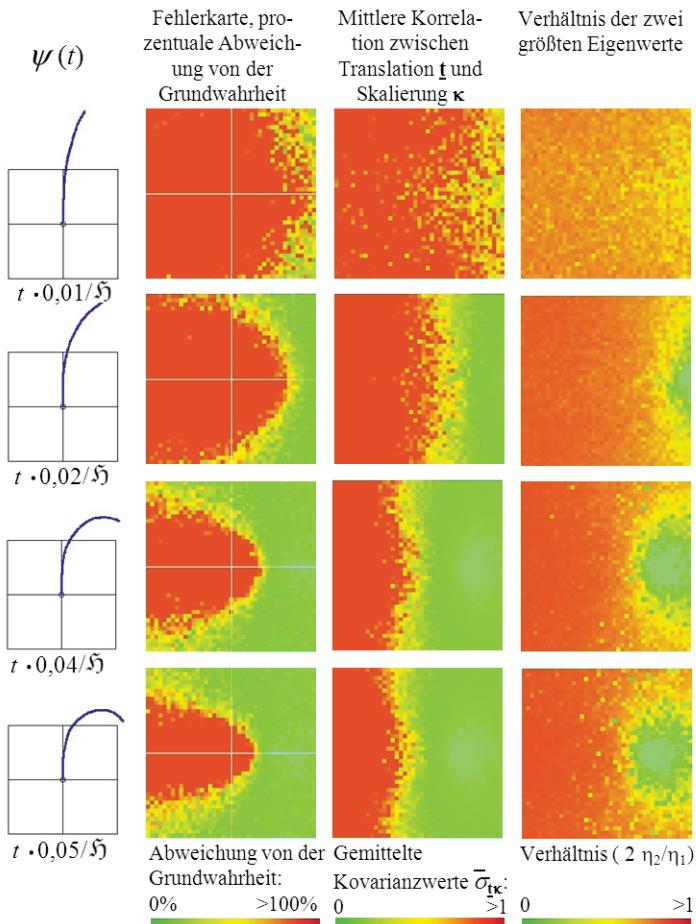
$$\bar{\sigma}_{\underline{\mathbf{t}}\kappa} = \frac{1}{\mathfrak{H}} \sum_{m=3}^{\mathfrak{H}+2} \left[ |(\underline{\mathbf{P}}_{\mathcal{C}})_{1m}| + |(\underline{\mathbf{P}}_{\mathcal{C}})_{2m}| \right] \quad (6.25)$$

formuliert solch eine Größe, wobei in Anlehnung an Gl. 6.23  $(\underline{\mathbf{P}}_{\mathcal{C}})_{nm}$  das Element von  $\underline{\mathbf{P}}_{\mathcal{C}}$  in der  $n$ -ten Zeile und der  $m$ -ten Spalte ist.

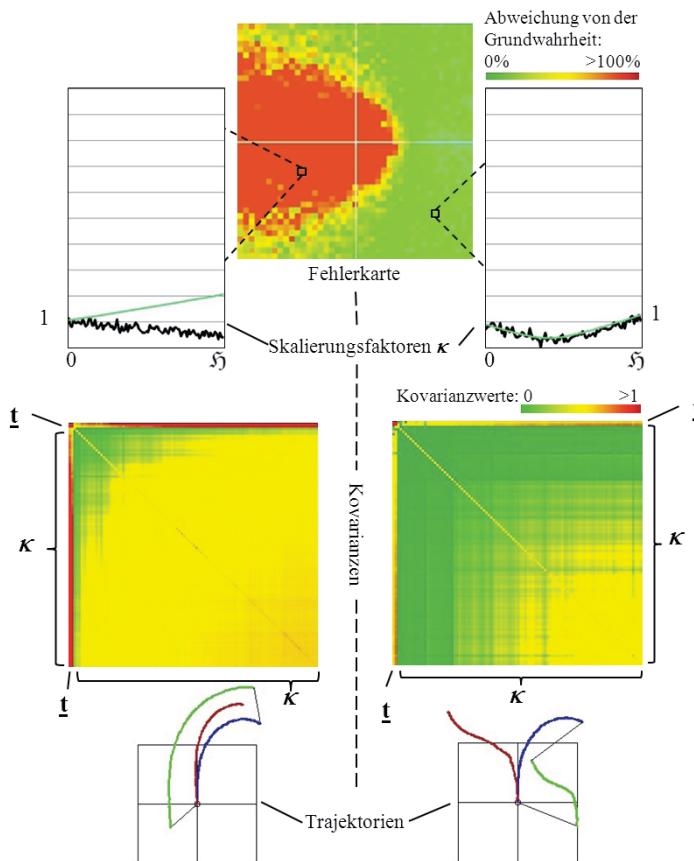
Abb. 6.7 und 6.9 zeigen die Fehlerkarten  $\varepsilon_{SMA}$  bezüglich der Grundwahrheit, die Karten der Korrelationswerte  $\bar{\sigma}_{\underline{\mathbf{t}}\kappa}$  sowie die Karten der Verhältnisse des zweitgrößten zum größten Eigenwert. Es fällt auf, dass bei Kurvenfahrten der Außenbereich der Kurve nur schlecht geschätzt werden kann. Dies tritt bei den hier durchgeführten Versuchen vor allem dann auf, wenn die Eingangstrajektorien der Kameras sehr ähnlich sind (wie z. B. in Abb. 6.8, links). Man erkennt ferner, dass der Wert von  $\bar{\sigma}_{\underline{\mathbf{t}}\kappa}$  bereits gute Hinweise auf die zu erwartende Güte des Schätzergebnisses zurück liefert. D. h. es ist möglich, unmittelbar aus der Kovarianzmatrix Rückschlüsse auf das Schätzverhalten ziehen, ohne eine aufwändige Eigenwertzerlegung durchführen zu müssen.



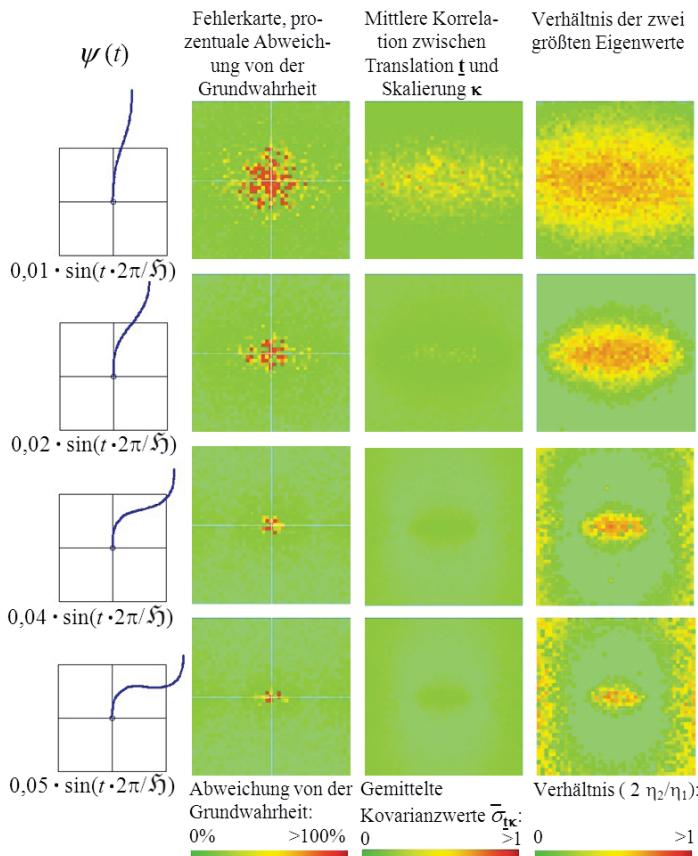
**Abbildung 6.6:** Ergebnis der Hauptkomponentenanalyse für planare Bewegungen mit linear steigender Rotation  $\psi$ . Die Karten zeigen die ersten beiden Komponenten der Eigenvektoren (in Richtung  $t_x$  und  $t_z$ ) gewichtet mit den jeweils ersten beiden Eigenwerten. In den meisten Fällen ist der zweite Eigenwert verschwindend gering.



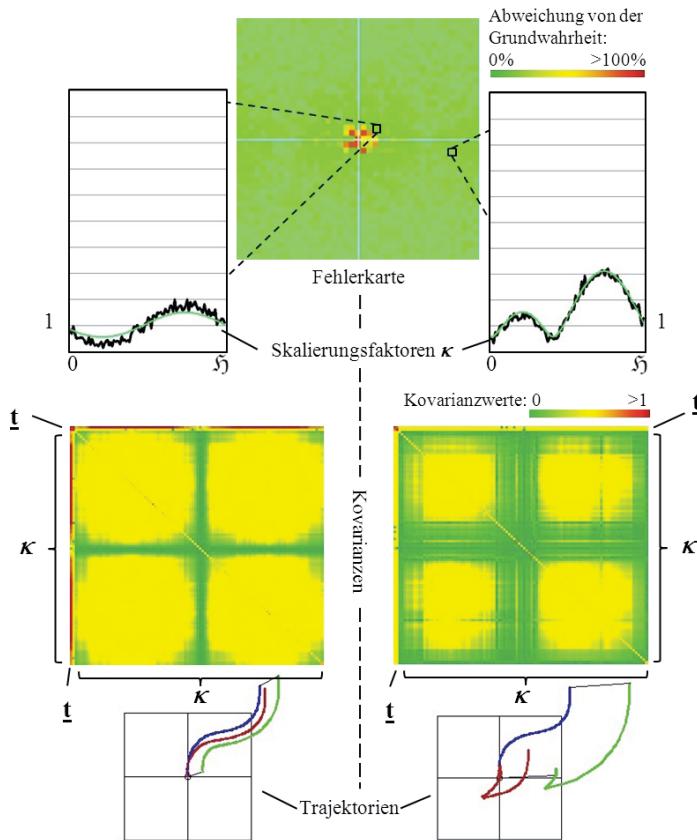
**Abbildung 6.7:** Auswertung verschiedener linearer Parameterkurven  $\psi(t)$  (1. Spalte) und Vergleich der kovarianz- und der eigenwertbasierten Kenngrößen (die beiden rechten Spalten) mit der Fehlerkarte der Grundwahrheit (2.Spalte). Rote Bereiche markieren schlecht schätzbare Kamerapositionen relativ zum Bezugspunkt  $C$ . Sowohl die 3. als auch die 4. Spalte spiegeln den Sachverhalt aus Spalte 2 wider, dass Kamerapositionen links vom Bezugspunkt  $C$  bei der durchgeführten Rechtskurve schlechter geschätzt werden können.  $\bar{\sigma}_{tx}$  weist hierbei stärkere Kanten zwischen den roten und grünen Bereichen auf, als das Eigenwertverhältnis.



**Abbildung 6.8:** Ausgewählte Testergebnisse einer Fehlerkarte bei einer einfachen Kurvenfahrt mit linearer Parameterkurve  $\psi(t)$  und einer schlechten Schätzung (links) und einer guten Schätzung (rechts). Oben: Fehlerkarte mit den geschätzten Skalaren  $\kappa$  zweier exemplarischer Auswertungen (die Sollwerte sind grün eingefärbt). Mitte: Die Kovarianzmatrizen der geschätzten Parameter  $[t, \kappa]$ . Die Kovarianzen von  $t$  und  $\kappa$  sind links sehr hoch (roter Rand) und deuten auf eine starke Korrelation der Skalierungsfaktoren und der Translationsparameter hin. Unten: Die Trajektorien des Bezugspunkts  $C$  (blau), die initiale, skalierte Trajektorie des Kameramoduls  $M$ , wie sie eingangs in den Schätzprozess eingeht (rot), die wahre, unskalierte Solltrajektorie des Kameramoduls  $M$  (grün).



**Abbildung 6.9:** Auswertung verschiedener sinusförmiger Parameterkurven  $\psi(t)$  (1. Spalte) und Vergleich der kovarianz- und der eigenwertbasierten Kenngrößen (die beiden rechten Spalten) mit der Fehlerkarte der Grundwahrheit (2. Spalte). Rote Bereiche markieren schlecht schätzbare Kamerapositionen relativ zum Bezugspunkt  $\mathcal{C}$ . Sowohl die 3. als auch die 4. Spalte spiegeln den Sachverhalt aus Spalte 2 wider, dass Kamerapositionen um den Bezugspunkt  $\mathcal{C}$  herum bei der durchgeführten Rechtskurve schlechter geschätzt werden können. Die Korrelationswerte sind schwächer ausgeprägt als die Eigenwertverhältnisse. Letztere jedoch weisen zunehmend hohe Werte im äußeren Bereich auf, welche sich nicht mit der Grundwahrheit in Spalte 2 in Deckung bringen lässt.



**Abbildung 6.10:** Ausgewählte Testergebnisse einer Fehlerkarte bei einer doppelten Kurvenfahrt mit sinusförmiger Parameterkurve  $\psi(t)$  und einer schlechten Schätzung (links) und einer guten Schätzung (rechts). Oben: Fehlerkarte mit den geschätzten Skalaren  $\kappa$  zweier exemplarischer Auswertungen (die Sollwerte sind grün eingefärbt). Mitte: Die Kovarianzmatrizen der geschätzten Parameter  $[t, \kappa]$ . Die Kovarianzen von  $t$  und  $\kappa$  sind links sehr hoch (roter Rand) und deuten auf eine starke Korrelation der Skalierungsfaktoren und der Translationsparameter hin. Unten: Die Trajektorien des Bezugspunkts  $C$  (blau), die initiale, skalierte Trajektorie des Kameramoduls  $M$ , wie sie eingangs in den Schätzprozess eingeht (rot), die wahre, unskalierte Solltrajektorie des Kameramoduls  $M$  (grün).



## Kontinuierliche Zustandsfilterung

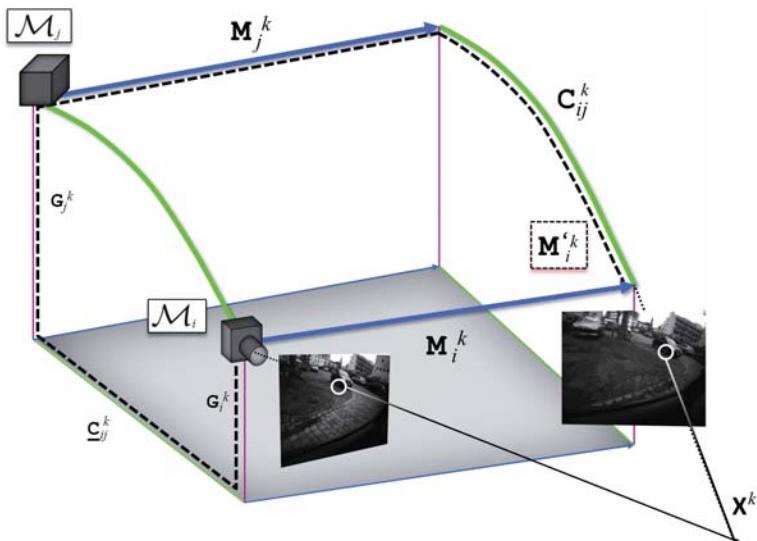


Abbildung 7.1: Projektion eines Szenenpunktes

In diesem Abschnitt wird erläutert, wie die extrinsischen Parameter zwischen zwei

Modulen in Abhängigkeit der lokalen Messungen, d. h. der gemessenen Punktkorrespondenzen eines Moduls, und der Bewegung der entsprechenden Module kontinuierlich geschätzt werden. Hierbei wird vorausgesetzt, dass die Bewegungstransformationen aller Module aus vorangegangenen Berechnungen bekannt seien. Ziel ist es, unter Berücksichtigung von Gl. 4.2, die Kalibrierparameter und deren Unsicherheit zu bestimmen, sodass der Projektionsfehler minimal wird (s. Abb. 7.1). Das bedeutet konkret: Die aus dem Bündelausgleich bekannten 3D-Punkte im Koordinatensystem von  $\mathcal{M}_i$  zum Zeitpunkt  $k - 1$  werden in das Koordinatensystem von  $\mathcal{M}_i$  zum Zeitpunkt  $k$  projiziert. Die Verschiebung des Koordinatensystems resultiert aus der Kamerabewegung  $M_i$  von  $\mathcal{M}_i$  zwischen den beiden Zeitpunkten.  $M_i$  lässt sich mit Hilfe der Bewegung  $M_j$  von  $\mathcal{M}_j$  in Abhängigkeit der gesuchten extrinsischen Parameter  $C_{ij}$  gemäß Gl. 4.2 modellieren gemäß

$$M_i = C_{ij}^{-1} M_j C_{ij}. \quad (7.1)$$

Für eine kontinuierliche Schätzung sollen sinnvollerweise sowohl die Zwischenergebnisse aus dem Bewegungsausgleich (siehe Kapitel 6) und ggf. der Bodenschätzung (siehe Abschnitt 5.3.3) als auch der prädizierte Zustand eingehen. Aus diesem Grund wird Gl. 7.1 dahingehend modifiziert, sodass eine extrinsische Transformation die Messung  $C_{z_{ij}}$ , und die andere Transformationsmatrix den zu schätzenden Zustand  $C_{x_{ij}}$  modelliert. Dadurch formuliert sich Gl. 7.1 um zu

$$M'_i := C_{z_{ij}}^{-1} M_{z_j} C_{x_{ij}}. \quad (7.2)$$

wobei der Index  $x$  signalisiert, dass die entsprechenden Transformationsparameter Teil des zu schätzenden Zustandsvektors sind, wohingegen der Index  $z$  die Transformation als Eingangsmessung deklariert. Der Projektionsfehler, über den es zu minimieren gilt, ergibt sich dann gemäß

$$\sum_{p=1}^{\mathfrak{P}} \left\| \pi(\mathbf{x}_{i_p}^k; \mathbf{K}_i, M'_i) - \mathbf{x}_{i_p}^k \right\|^2 \xrightarrow{!} \min_{C_{x_{ij}}} \quad (7.3)$$

Das erweiterte Kalibriermodell vollzieht eine Aufteilung der Transformationen  $C_{ij}$  zwischen den Kameramodulen  $\mathcal{M}_i, \mathcal{M}_j$  in Teiltransformationen, wie sie in Kapitel 6 und den Abschnitten 5.3.3 und 5.3.1 berechnet werden können. Demnach existieren für jedes Kameramodul  $\mathcal{M}_i$  eine Transformation  $G_i$ , die das Koordinatensystem der Kamera in die Bodenebene transformiert, sowie Transformationen  $\underline{C}_{ij}$  aus dem Bewegungsausgleich, die die Lage und Orientierung zweier Module innerhalb der Bodenebene beschreiben.  $C_{ij}$  setzt sich in diesem Fall gemäß

$$C_{ij} = G_i^{-1} \underline{C}_{ij} G_j, \quad (7.4)$$

zusammen. Darüber hinaus existieren zu jeder dieser Transformationen Kovarianzmatrizen, zur Modellierung der Unsicherheiten der geschätzten Transformationen. D.h. die Zustände der Teiltransformationen sowie der Bewegungstransformation gehen als Messung in die Filterung des Zustands der extrinsischen Parameter ein. Um eine einheitliche Beschreibung der extrinsischen Kameraparameter zu erhalten, müssen diese Teiltransformationen nun in Abhängigkeit ihrer Unsicherheiten zusammengeführt werden.

## 7.1 Bestimmung der Kovarianzmatrizen der Teiltransformationen

Für die Zustände der Teiltransformationen  $\underline{C}$  und  $G$  und der Bewegungstransformationen  $M$  wird je eine  $3 \times 3$ -Kovarianzmatrix mitgeführt. Das bedeutet, dass auch für diese Transformationen, wenngleich sie nicht explizit gefiltert werden, eine Kovarianzmatrix zu Beginn des Algorithmus initialisiert und zu jedem Zeitschritt aktualisiert wird. Grundlage hierfür sind die Kalmangleichungen 2.41 oder 2.42.  
Die prädizierte Kovarianzmatrix  $P_{*k}^-$  ergibt sich durch Addition einer konstanten Rauschma-

trix  $\mathbf{W} = \text{Diag}\sigma_{*_1}^2, \dots, \sigma_{*\dim(x_*)}^2$  gemäß

$$\mathbf{P}_{*k}^- := \mathbf{P}_{*k-1}^+ + \mathbf{W}, \quad * \in \{\underline{C}, G, M\} \quad (7.5)$$

Die Aktualisierung einer Kovarianzmatrix erfolgt gemäß

$$\mathbf{P}_{*k}^+ := \mathbf{P}_{*k}^- - \mathbf{P}_{*k}^- \mathbf{J}_*^T \left[ \mathbf{J}_* \mathbf{P}_{*k}^- \mathbf{J}_* + \sigma_*^2 \cdot \mathbf{I} \right]^{-1} \mathbf{J}_* \mathbf{P}_{*k}^-, \quad * \in \{\underline{C}, G, M\} \quad (7.6)$$

mit  $\mathbf{J}_* = \frac{\partial \mathbf{h}_*}{\partial \mathbf{x}_*}$ . Das Messrauschen  $\sigma_*^2$  ergibt sich je nach Transformationstyp aus der jeweiligen Schätzung und der entsprechenden Bedingungsfunktion:

- $\mathbf{h}_{\underline{C}}$  (6.11, 6.13) aus dem Bewegungsausgleich für  $\underline{C}$
- $\mathbf{h}_G$  (5.14) aus der Bodenschätzung für  $G$
- $\mathbf{h}_M$  (3.13) aus dem Bündelausgleich für  $M$

Diese drei Bedingungsgleichungen sind zugleich Residuumsfunktionen, in dem Sinne, dass sie den Fehler in Abhängigkeit der jeweiligen Transformation beschreiben. Unter der Annahme, dass der Fehler jedes Messpunktes normal um Null verteilt ist, lassen sich anhand der geschätzten Transformationsparameter und der verbleibenden Residuen die Varianzen  $\sigma_{\underline{C}}^2$ ,  $\sigma_G^2$  und  $\sigma_M^2$  berechnen.

## 7.2 Zusammenführen der Teilergebnisse und Filterung der extrinsischen Parameter

Sowohl die Minimierung des Projektionsfehlers als auch die Zusammenführung der Teiltransformationen, lässt sich geeigneterweise mit einem nichtlinearen Filter, wie dem iterativen erweiterten Kalmanfilter realisieren (siehe Abschnitt 2.4).

Um die folgenden Darstellungen übersichtlich zu halten, wird im vorliegenden Abschnitt auf die Indexierung der Zustände, Messungen sowie Kovarianzmatrizen mit dem Zeitindex  $k$  verzichtet.

Sei der zu schätzende Zustand und dessen Kovarianzmatrix gegeben durch

$$\underline{\mathbf{x}}_{\mathcal{C}ij} = [t_{x_{ij}}, t_{y_{ij}}, t_{z_{ij}}, r_{x_{ij}}, r_{y_{ij}}, r_{z_{ij}}]^T \quad (7.7)$$

und  $\mathbf{P}_{\mathcal{C}ij}$ . Des weiteren seien die Messungen gegeben durch die berechneten Teiltransformationen und der Bewegung von  $\mathcal{M}_j$ ,

$$\underline{\mathbf{x}}_{\mathcal{C}ij} = [\underline{t}_{x_{ij}}, \underline{t}_{y_{ij}}, \beta_{x_{ij}}]^T, \quad (7.8)$$

$$\mathbf{x}_{\mathcal{G}i} = [h_{y_i}, \alpha_i, \gamma_i]^T, \quad (7.9)$$

$$\mathbf{x}_{\mathcal{G}j} = [h_{y_j}, \alpha_j, \gamma_j]^T, \quad (7.10)$$

$$\mathbf{x}_{\mathcal{M}j} = [v_{x_j}, v_{y_j}, v_{z_j}, \omega_{x_j}, \omega_{y_j}, \omega_{z_j}]^T \quad (7.11)$$

sowie deren Kovarianzmatrizen  $\mathbf{P}_{\mathcal{C}ij}$ ,  $\mathbf{P}_{\mathcal{G}i}$ ,  $\mathbf{P}_{\mathcal{G}j}$  bzw.  $\mathbf{P}_{\mathcal{M}j}$ . Als Prädiktionsfunktion  $f$  wird die Identitätsfunktion verwendet. Zusammen mit den von  $\mathcal{M}_i$  im SBA errechneten Szenenpunkten  $\mathbf{X}_{i_p}$  und deren gemessenen Projektionen  $\mathbf{x}_{i_p}$  ergibt sich der Messvektor  $\mathbf{z}$  zum Zeitpunkt  $k$  gemäß

$$\mathbf{z} = \begin{bmatrix} \underline{\mathbf{x}}_{\mathcal{C}ij} \\ \mathbf{x}_{\mathcal{G}i} \\ \mathbf{x}_{\mathcal{G}j} \\ \mathbf{x}_{\mathcal{M}j} \\ \mathbf{x}_{i_1} \\ \mathbf{X}_{i_1} \\ \vdots \\ \mathbf{x}_{i_{\mathfrak{P}}} \\ \mathbf{X}_{i_{\mathfrak{P}}} \end{bmatrix} . \quad (7.12)$$

Die Kovarianzmatrix des Messvektors hat die blockdiagonale Gestalt

$$\Sigma_z = \left[ \begin{array}{cccc|cccc} P_{C_{ij}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & P_{G_i} & \mathbf{0} & \mathbf{0} & \vdots & \ddots & & \vdots \\ \mathbf{0} & \mathbf{0} & P_{G_j} & \mathbf{0} & \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & P_{M_j} & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \hline \mathbf{0} & \cdots & \cdots & \mathbf{0} & D & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & & \vdots & \mathbf{0} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots & \vdots & & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \cdots & D \end{array} \right] \quad (7.13)$$

mit  $D = \text{Diag}(\sigma_x^2, \sigma_x^2, \sigma_X^2, \sigma_X^2, \sigma_X^2, \sigma_X^2)$ , wobei  $D$  die unkorrelierte, diagonale Kovarianzmatrix mit dem Messrauschen der Bild- und Szenenpunkte bezeichne. Die Bedingungsgleichung lässt sich in Anlehnung an Gln. 7.3 bis 7.13 formulieren gemäß

$$\mathbf{h}_C(x_{C_{ij}}, z, K) = \left[ \begin{array}{c} \vdots \\ h_{C_p}(x_{C_{ij}}, z, K) \\ \vdots \end{array} \right] \stackrel{!}{=} \mathbf{0} \quad (7.14)$$

mit

$$h_{C_p}(x_{C_{ij}}, z, K) = \left\| \pi\left(X_{i_p}; K_i, \{x_{C_{ij}}, x_{C_{ij}}, x_{G_i}, x_{G_j}, x_{M_j}\}\right) - x_{i_p} \right\|^2. \quad (7.15)$$

Die Jacobi-Matrizen mit den partiellen Ableitungen zu Gl. 7.14 und 7.15 nach  $x_{C_{ij}}$  und  $z_p$  haben die Gestalt

$$\mathbf{H}_x = \frac{\partial \mathbf{h}_C}{\partial x_{C_{ij}}} \Big|_{\hat{x}_{C_{ij}}, z} = \left[ \begin{array}{c} \vdots \\ \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial x_{C_{ij}}} \right]^T \\ \vdots \end{array} \right] \quad (7.16)$$

bzw.

$$\mathbf{H}_z = \frac{\partial \mathbf{h}_C}{\partial z} \Big|_{\hat{x}_{C_{ij}}, z} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix}, \text{ mit} \quad (7.17)$$

$$\mathbf{A} = \left[ \begin{array}{cccc} \vdots & \vdots & \vdots & \vdots \\ \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{x}_{C_{ij}}} \right]_{1 \times 3}^T & \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{x}_{G_i}} \right]_{1 \times 3}^T & \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{x}_{G_j}} \right]_{1 \times 3}^T & \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{x}_{M_j}} \right]_{1 \times 6}^T \\ \vdots & \vdots & \vdots & \vdots \end{array} \right]_{\mathfrak{P} \times 15}, \quad (7.18)$$

$$\mathbf{B} = \left[ \begin{array}{cccc} \mathbf{B}_1 & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \mathbf{0}^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0}^T \\ \mathbf{0}^T & \cdots & \mathbf{0}^T & \mathbf{B}_{\mathfrak{P}} \end{array} \right]_{\mathfrak{P} \times \mathfrak{P} \cdot 5}, \quad (7.19)$$

$$\mathbf{B}_p = \left[ \begin{array}{cc} \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{x}_{i_p}} \right]_{1 \times 2}^T & \left[ \frac{\partial \mathbf{h}_{C_p}}{\partial \underline{\mathbf{X}}_{i_p}} \right]_{1 \times 3}^T \end{array} \right]_{1 \times 5}. \quad (7.20)$$

Zu Beginn des Schätzprozesses ist die Kalibrierung  $\hat{x}_{C_{ij}}$  noch unbekannt, sodass die Zustandskovarianzmatrix  $\mathbf{P}_{C_{ij}}$  mit entsprechend hohen Werten initialisiert werden muss. Durch Einsetzen der Gln. 7.12 bis 7.17 in die Kalmangleichungen aus Abschnitt 2.4 lassen sich die extrinsischen Parameter je Kamerapaar mit Hilfe der Teiltransformationen und in Abhängigkeit der gemessenen Bildmerkmale über die Zeit hinweg filtern. D. h. durch Minimierung des Projektionsfehlers sind die Bildpunktcorrespondenzen der einzelnen Module, und damit die fundamentale, der Kalibrierung zugrunde liegende Messung, explizit in den Optimierungsprozess integriert.

Bei der Zusammensetzung der Teiltransformationen beeinflusst die Höhenkomponente  $h_{y_i}, h_{y_j}$  der Bodentransformationen der jeweiligen Module den resultierenden Translationsvektor  $\mathbf{t}$ . Ausmultiplizieren von Gl. 7.4 ergibt die Linearkombination

$$\mathbf{t} = \underline{t}_{x_{ij}} \begin{bmatrix} \cos(\gamma_i) \\ -\cos(\alpha_i)\sin(\gamma_i) \\ \sin(\alpha_i)\sin(\gamma_i) \end{bmatrix} + \underline{t}_{z_{ij}} \begin{bmatrix} 0 \\ \sin(\alpha_i) \\ \cos(\alpha_i) \end{bmatrix} + (h_{y_j} - h_{y_i}) \begin{bmatrix} \sin(\gamma_i) \\ \cos(\alpha_i)\cos(\gamma_i) \\ -\sin(\alpha_i)\cos(\gamma_i) \end{bmatrix}$$

(7.21)

Aus Gl. 7.21 wird ersichtlich, dass die translatorische Komponente der extrinsischen Parameter ohne Kenntnis der Höhenkomponenten nur bis auf ein lineares Restglied bestimbar sind, dessen Einfluss umso stärker ist, je größer der vertikale Versatz der Kameras auf der Plattform ist.

# 8

---

## Multimodulare Fusion

---

Der Einsatz von mehreren, verteilten Kameras führt zwangsläufig zu einem gesteigerten Informationsgehalt gegenüber Systemen mit weniger Kameras. Aus dieser einfachen Überlegung heraus wird nun versucht, die Information aller Module in einer Art und Weise ganzheitlich zu verarbeiten, sodass sowohl Teilergebnisse der einzelnen Module als auch das fusionierte Gesamtergebnis davon profitieren.

Ein wichtiges Paradigma, das in diesem Kontext beachtet werden soll, ist die Skalierbarkeit des Systems. Vor allem bei Mehrkamerasytstemern unbestimmter Anzahl ist es von Bedeutung, dass der Rechenaufwand nicht proportional mit der Anzahl der Kameras zunimmt. Insbesondere die sich stetig weiterentwickelnde Technologie paralleler Rechnerplattformen macht eine Lösung attraktiv, die eine parallele Verarbeitung der Algorithmen unterstützt. Dies schließt eine Informationsfusion auf unterster Datenebene, d. h. eine gemeinsame Verarbeitung aller Kamerabilder oder Bildmerkmale, weitestgehend aus. Bedenkt man beispielsweise eine gemeinsame Verarbeitung der in den  $N$  Modulen jeweils berechneten  $P$  Bildpunktcorrespondenzen, so hätte der Auswertealgorithmus bei  $N$  Kameras nunmehr  $N \cdot P$  Eingangsdaten je Zeitschritt zu prozessieren statt – wie bei einer parallelen Verarbeitung – nur  $P$  Merkmale. Bei einer simultanen Schätzung aller Transformationsparameter des Gesamtsystems würde die

Dimension des Zustandsraumes auf  $\mathfrak{N} \cdot 6$  (Bewegungsparameter) +  $[\mathfrak{N} - 1] \cdot 6$  (extrinsische Parameter) =  $12\mathfrak{N} - 6$  ansteigen. Auch die Datentransferrate ist eine nicht zu vernachlässigende Größe. Der Austausch von Bilddaten oder -merkmalen kann sich bei stark belasteten Bussystemen leicht zu einem Flaschenhals entwickeln.

**Propagierung und Fusion** Der hier vorgeschlagene Lösungsansatz modelliert die Fusion auf Zustandsebene. D. h. statt einer ganzheitlichen Auswertung der Bilddaten bzw. -merkmale werden einzelne Zustände und deren Unsicherheiten zwischen den einzelnen Modulen ausgetauscht. Dieser Ansatz ist auch hinsichtlich der Datenrate günstig, da lediglich Zustandsvektoren und Kovarianzmatrizen sowie einzelne Skalare transferiert werden müssen.

Im vorliegenden Fall handelt es sich bei den Zuständen um Euklidische Transformationen. Bei bekannten extrinsischen Parametern wäre die Bewegungsschätzung aller  $\mathfrak{N}$  Kameramodule redundant, da man von der Bewegung einer Kamera auf die Bewegung aller anderen Kameras rückschließen könnte. In der Praxis ist eine bildgestützte Bewegungsschätzung fehlerbehaftet. Die  $\mathfrak{N}$  redundanten Schätzungen  $M_i^{(j)}$  der Module  $\mathcal{M}_{j=1,\dots,\mathfrak{N}}$ , die die Bewegung  $M_i$  von Modul  $\mathcal{M}_i$  beschreiben, sollten daher die Unsicherheit von  $M_i$  verringern und das Schätzergebnis verbessern. Hierbei wird unterstellt, dass die Unsicherheiten der Transformationen normalverteilt und daher mittels Kovarianzmatrizen modellierbar sind. Somit lassen sich durch Propagierung der Bewegungszustände und deren Unsicherheiten sowie einer anschließenden Gauß'schen Fusion die je Modul berechneten Kamerabewegungen zusammenführen.

Dies gilt auch umgekehrt für die Kalibrierparameter. Bei bekannter Bewegung jedes einzelnen Moduls lassen sich gemäß den Ausführungen in Kapitel 4 je Modul die Transformationen der relativen Lage und Orientierung zwischen allen Modulen berechnen. Als Ergebnis liegt folglich jedem Modul eine »eigene« Schätzung des vollständigen Transformationsgraphen der extrinsischen Anordnung vor. Auch hier tritt wieder eine Redundanz auf, die fusioniert und dadurch konstruktiv genutzt werden kann.

Das Propagierungs- bzw. Fusionskonzept der Bewegungs- und extrinsischen Parameterschätzung wird in den folgenden Abschnitten eingehend erläutert.

**Lokale und globale Zustände** Im folgenden werden Zustände bzw. Euklidische Transformationen und Kovarianzmatrizen, deren Berechnungen auf den Sensordaten eines einzelnen Moduls beruhen, als *lokal* bezeichnet. Werden die redundanten Zustände aller Module mittels Propagierung und Fusion zusammengeführt, so spricht man im vorliegenden Kontext von *globalen* Zuständen. In den globalen Zuständen sind die Informationen aller Module auf Zustands-, d. h. abstrakter Ebene entsprechend ihrer Unsicherheit miteinander vereint. Lokale Bewegungszustände eines Moduls sind unabhängig von den Sensordaten der anderen Module. Lokale Kalibrierzustände hängen von den zuvor berechneten globalen Bewegungszuständen ab, sowie den lokalen Bildpunktcorrespondenzen zur Berechnung des Projektionsfehlers (siehe Kapitel 7). Globale Transformationen werden mit einem hochgestellten  $\mathfrak{c}$  gekennzeichnet:  $T^G$ . Lokale Transformationen haben den Index des Moduls, auf dessen Schätzung die Transformation beruht, in Klammern hochgestellt.  $T_i^{(j)}$  ist demnach die lokale Schätzung bzw. Propagierung von Modul  $\mathcal{M}_j$ , die die Transformation von Modul  $\mathcal{M}_i$  beschreibt. Die gleichen Notationen gelten auch für Zustände und Kovarianzmatrizen.

Die kommenden beiden Abschnitte beschreiben das Konzept und die Realisierung der Fusion lokaler Bewegungs- und Kalibriertransformationen in Abhängigkeit der Zustandsunsicherheiten. Sowohl bei der Bewegungs- als auch der extrinsischen Fusion wird nach folgendem Grundschema verfahren:

1. Lokale Transformationsschätzung
2. Propagierung der Transformationen mittels Gauß'scher Fehlerfortpflanzung
3. Gauß'sche Fusion der redundanten Transformationsparameter

## 8.1 Globales Schema zur Bewegungsschätzung

In diesem Abschnitt wird die Fusion der in den einzelnen Modulen geschätzten Bewegungen beschrieben. Hierfür werden die extrinsischen Parameter und vor allem deren

Unsicherheit benötigt. Hierbei gestaltet sich die Vorgehensweise für die Bewegungsschätzung gemäß dem Grundschema wie folgt:

0. System- bzw. Bewegungsinitialisierung
1. Lokale Bewegungsschätzung mittels Bündelausgleich
2. Propagierung der Bewegungen
3. Prüfen der Bewegungskonsistenz
4. Gauß'sche Fusion der Bewegungsparameter

### 8.1.1 Systeminitialisierung durch initiale Bewegungsschätzung

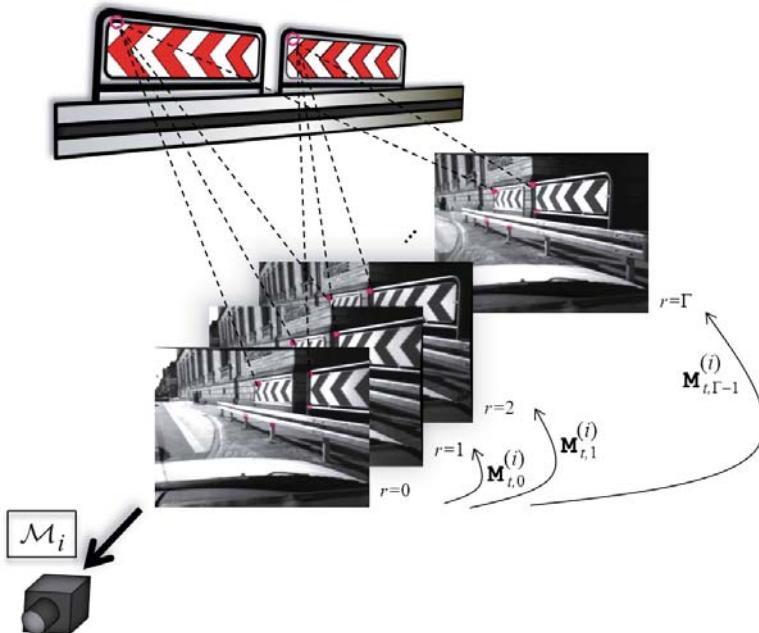
Der Schätzprozess wird mittels einer Geradeausfahrt initialisiert (siehe Anhang A.4). D.h. es werden drei translatorische Bewegungsparameter je Kameramodul geschätzt. Anhand der geschätzten Bewegungsparameter kann die 3D-Struktur der verwendeten Bildpunktcorrespondenzen rekonstruiert werden [Har95]. Sowohl die geschätzte translatorische Bewegung als auch die damit berechnete Punktwolke dient der Initialisierung des Bündelausgleichs im Folgezeitschritt<sup>26</sup>.

### 8.1.2 Lokale Bewegungsschätzung

Die lokale Bewegungsschätzung erfolgt durch die Berechnung des Bündelausgleichs auf einem zeitlich zusammenhängenden Satz von  $\Gamma + 1$  Bildern (3.2). Die eigentliche lokale Messung, d.h. die direkte Auswertung der Bildsensorinformation, erfolgt zu Beginn eines jeden Zeitschritts  $k$ . Durch die in Abschnitt 3.1 beschriebene regionenbasierte Korrespondenzanalyse werden in den  $\Gamma + 1$  letzten Bildern der laufenden Sequenz in

---

<sup>26</sup> Ein erneutes Aufsetzen des Algorithmus kann auch während der Fahrt anhand einer translatorischen Bewegung erfolgen. Dies kann u.U. dann erforderlich sein, wenn in den Modulen zu wenige Bildpunktcorrespondenzen gemessen werden (beispielsweise wegen homogener Bildregionen, Gegenlicht oder Bewegungsunschärfe). Aber auch, wenn zu viele Modulbewegungen aufgrund der Kovarianzmatrix als unsicher oder inkonsistent erachtet werden (siehe Abschnitt 8.1.5).



**Abbildung 8.1:** Lokale Schätzung eines Bewegungsblocks mittels Bündelausgleich.

je zwei aufeinanderfolgenden Aufnahmen paarweise korrespondierende Merkmale extrahiert. In einer Datenvorauswahl werden Punktetupel, die der Epipolarbedingung nicht entsprechen, verworfen (3.1.1 und A.3). Für die Bewegungsschätzung mittels Bündelausgleich werden  $\mathfrak{P}$  Bildpunktetupel  $\{\mathbf{x}_{0,p}, \dots, \mathbf{x}_{\Gamma-1,p}\}_{p=1, \dots, \mathfrak{P}}$  verwendet. Eine Initialisierung der 3D-Struktur erfolgt mittels Triangulation anhand der Bewegungsparameter und Korrespondenzen aus dem vorangegangenen Zeitschritt [Har95]. Nach der Berechnung werden die Bewegungstranslationsvektoren relativ zum Translationsbetrag von  $M_{t,0}$  skaliert gemäß

$$\mathbf{v}'_{t,r} = \frac{\mathbf{v}_{t,r}}{\|\mathbf{v}_{t,0}\|} \cdot v, \quad (8.1)$$

wobei  $v$  der aktuellen Fahrzeuggeschwindigkeit (bekannt aus z. B. Odometriesensoren) oder einer Konstanten entspricht<sup>27</sup>. Die Berechnung der entsprechenden Kovarianzmatrizen erfolgt wie in Abschnitt 7.1 beschrieben.

### 8.1.3 Propagierung und Fusion der Bewegungsparameter

Für die nun folgende Fusion werden die extrinsischen Parameter benötigt. Diese sind zu Beginn des Schätzprozesses im schlechtesten Fall völlig unbekannt. Um dennoch mit Hilfe der extrinsischen Parameter die Transformationen propagieren und somit fusionieren zu können, muss die Kovarianzmatrix in diesem Fall mit entsprechend hohen Werten initialisiert sein.

Seien nach dem Bündelausgleich die lokalen Bewegungstransformationen und die extrinsischen Parameter sowie deren Unsicherheiten bekannt. D. h. jedes Modul kennt die eigene Bewegung und die extrinsische Anordnung. Dann lässt sich gemäß Gl. 4.2 aus der lokalen Betrachtung von Modul  $\mathcal{M}_i$  auch auf die Bewegung aller anderen Module  $\mathcal{M}_{j \neq i}$  rückschließen

$$\mathbf{M}_{jr}^{(i)} = \mathbf{C}_{ij}^G \mathbf{M}_{ir}^{(i)} \mathbf{C}_{ij}^G {}^{-1}. \quad (8.2)$$

Die entsprechende Unsicherheit  $\mathbf{P}_{\mathbf{M}_{jr}}^{(i)}$  lässt sich entsprechend Gl. 2.51 propagieren

$$\mathbf{P}_{\mathbf{M}_{jr}}^{(i)} = \mathbf{J}_{\mathbf{M}_{ir}} \mathbf{P}_{\mathbf{M}_{ir}} \mathbf{J}_{\mathbf{M}_i}^T + \mathbf{J}_{\mathbf{C}_{ij}} \mathbf{P}_{\mathbf{C}_{ij}} \mathbf{J}_{\mathbf{C}_{ij}}^T \quad (8.3)$$

mit

$$\mathbf{J}_{\mathbf{M}_{ir}} = \frac{\partial \zeta_{\mathbf{M}}(\mathbf{M}_{jr}^{(i)})}{\partial \mathbf{x}_{\mathbf{M}_{ir}}} \quad , \quad \mathbf{J}_{\mathbf{C}_{ij}} = \frac{\partial \zeta_{\mathbf{M}}(\mathbf{M}_{jr}^{(i)})}{\partial \mathbf{x}_{\mathbf{C}_{ij}}} \quad (8.4)$$

---

<sup>27</sup> Ist die Fahrzeuggeschwindigkeit unbekannt so gilt i. d. R.  $v = 1$ .

und  $\zeta_M(M) = x_M$ ,  $\zeta_C(C) = x_C$ . Diese Propagierung ist für jede der Transformationen  $M_{j_r=0,...,\Gamma-1}^{(i)}$  aus dem Historienblock durchführbar. Man kann auch sagen, jedes Modul hat zu diesem Zeitpunkt eine »eigene Vorstellung« von der Welt bzw. von der Anordnung aller Kameras und deren Bewegung. Um die lokalen Zustände fusionieren zu können, müssen diese zunächst zwischen den einzelnen Modulen ausgetauscht werden<sup>28</sup>.

Nach der Propagierung und dem Datenaustausch liegt jedem Modul dieselbe (redundante) Information vor. Da nun jedes der  $\mathfrak{N}$  Module eine eigene Schätzung der  $\mathfrak{N}$  Bewegungen hat, existieren folglich für jede Bewegung  $\mathfrak{N}$  normalverteilte Schätzungen. Um diese  $\mathfrak{N}$  mal  $\mathfrak{N}$  lokalen Bewegungen zu fusionieren<sup>29</sup>, bedienen wir uns der Gauß'schen Fusion aus Abschnitt 2.6

$$M_{i_r}^G = \bigcup_{i \neq j}^{\mathfrak{N}} M_{i_r}^{(j)} \quad (8.5)$$

$$\Sigma_{M_{i_r}}^G = \bigcup_{i \neq j}^{\mathfrak{N}} \Sigma_{M_{i_r}}^{(j)}, \quad r = 0, \dots, \Gamma - 1, \quad (8.6)$$

wobei  $\bigcup$  den Fusionsoperator symbolisiert.

Der hochgestellte Bezeichner  $G$  identifiziert die Transformation bzw. den entsprechenden Zustand und seine Kovarianzmatrix als *global*.

Als Ergebnis liegen nun  $\mathfrak{N}$  globale Bewegungs transformationen vor, die, unter Berücksichtigung der Kovarianzmatrizen, die Informationen aller Module berücksichtigen. Für den weiteren Kalibrierprozess ist die Art der Bewegung relevant. Insbesondere, ob eine rotatorische Bewegung vorliegt oder nicht. Durch eine einfache Evaluierung der

---

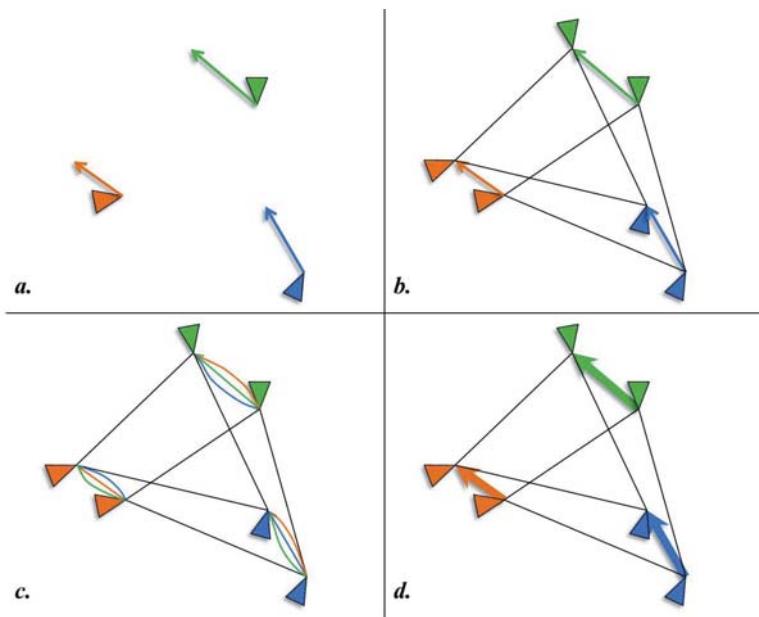
<sup>28</sup> Die Daten, die jedes Modul auf den Datenbus legen würde, hätte folglich die Größe  $108 \cdot \Gamma$  Bytes, 6 32-bit-Fließkommazahlen für den Bewegungszustand, und aufgrund der Symmetrie der Kovarianzmatrix, 21 Fließkommazahlen zur Beschreibung der Unsicherheit. Hier liegt der Grund für den Ansatz einer Fusion auf Zustandsebene. Allein der Austausch von 150 Bildpunkten brächte schon ein Datenaufkommen von 1200 Bytes.

<sup>29</sup> Bei unbekannten Translationsbeträgen werden lediglich die  $\Gamma$  Translationsrichtungen fusioniert.

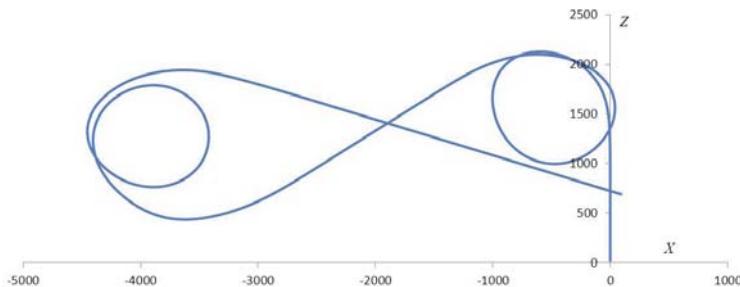
Eulerwinkel mit Hilfe eines Schwellwerts  $\Lambda_M$  kommt man zu folgender Klassifizierung:

$$\text{Bewegungstyp}(M_i^G) = \begin{cases} \text{rotatorisch,} & \sqrt{\omega_{x_i}^2 + \omega_{y_i}^2 + \omega_{z_i}^2} > \Lambda_M \\ \text{rein translatorisch,} & \text{sonst} \end{cases}. \quad (8.7)$$

Abb. 8.2 veranschaulicht den Prozess der Bewegungsfusion von der lokalen Schätzung bis zur Fusion der propagierten Bewegungstransformationen.



**Abbildung 8.2:** Propagierungs- und Fusionsschema der Bewegungsschätzung. *a.* Lokale Schätzung. *b.* Berücksichtigung der Kalibrierinformation. *c.* Propagierung der lokalen Bewegungsschätzungen basierend auf der Kalibrierung. *d.* Fusion der redundanten, propagierten Bewegungsschätzungen zu globalen Bewegungsschätzungen.



**Abbildung 8.3:** Simulierte Trajektorie der Plattform zur Untersuchung des Einflusses der Fusion bei der Bewegungsschätzung.

#### 8.1.4 Ergebnisse zur fusionierten Bewegungsschätzung

Zur Demonstration des Nutzens der Bewegungsfusion bei bekannter Kalibrierung wurde anhand einer Simulation eine Messfahrt mit einem Zweikamerasystem rekonstruiert (s. Abb. 8.3). Hierbei wurde die erste Kamera nach vorne, die zweite zur Seite schauend modelliert. Eingangsdaten waren simulierte Flussvektoren, die mit einem additiven, normalverteilten Rauschen  $\sigma = 1.0$  behaftet wurden. Die der Simulation zu Grunde liegende Bewegung kann zu Vergleichszwecken als Grundwahrheit herangezogen werden. Abb. 8.4 zeigt die sechs Euklidischen Bewegungsparameter jedes Kameramoduls ohne durchgeführte Fusion. Man sieht, dass die zweite Kamera größere Schwierigkeiten hat, die korrekte Rotation zu schätzen, als dies bei der ersten Kamera der Fall ist. Das Fusionsergebnis in Abb. 8.5 zeigt insbesondere bei den rotatorischen Parametern des zweiten Moduls  $\omega_2^G$  eine deutliche Verbesserung gegenüber der unfusionierten Schätzung  $\omega_2^{(2)}$ . Auch das Rauschen der Translation  $v_2^G$ , insbesondere des Parameters  $v_z$ , konnte reduziert werden.

#### 8.1.5 Prüfen der Bewegungskonsistenz

Sind Fahrzeugdaten wie Lenkwinkel und Geschwindigkeit des Fahrzeugs bekannt, so sind zumindest grobe Bewegungfehlschätzungen leicht zu detektieren. Stehen diese

Informationen nicht zur Verfügung, so lässt sich aber dennoch mit Hilfe der Kalibrierinformation eine Aussage darüber treffen, ob sich die Bewegung eines Kameramoduls konsistent zu den Bewegungen der anderen Kameras verhält.

Um zu entscheiden, ob eine Bewegung für eine weitere Verwendung im Kalibrierprozess geeignet ist, kann jede Bewegung relativ zur gesamten Anordnung unter einem globalen Gesichtspunkt evaluiert werden. Es sei angenommen, dass jedes Modul Kenntnis von den Bewegungsparametern der anderen Module hat. Dann lässt sich mit Hilfe einer Transformationspropagierung der Bewegungstransformation  $\mathbf{M}_j$  ( $= \mathbf{M}_{j_{r=0}}$ ) von Modul  $\mathcal{M}_j$  und der globalen Kalibriertransformation  $C_{ij}^G$  zwischen den Modulen  $\mathcal{M}_i$  und  $\mathcal{M}_j$  die Bewegung von Modul  $\mathcal{M}_i$  umformulieren zu

$$\mathbf{M}_i^{(j)} = C_{ij}^G \mathbf{M}_j^{(j)} C_{ij}^{G^{-1}}. \quad (8.8)$$

Die Unsicherheit berechnet sich gemäß Abschnitt 2.5

$$\mathbf{P}_{\mathbf{M}_i}^{(j)} = \mathbf{J}_M \mathbf{P}_{\mathbf{M}_j}^{(j)} \mathbf{J}_M^T + \mathbf{J}_C \mathbf{P}_{C_{ij}}^G \mathbf{J}_C^T \quad \text{mit} \quad (8.9)$$

$$\mathbf{J}_M = \frac{\partial \zeta_M(\mathbf{M}_i^{(j)})}{\partial \mathbf{x}_{\mathbf{M}_j}} \quad (8.10)$$

$$\mathbf{J}_C = \frac{\partial \zeta_M(\mathbf{M}_i^{(j)})}{\partial \mathbf{x}_{C_{ij}}}. \quad (8.11)$$

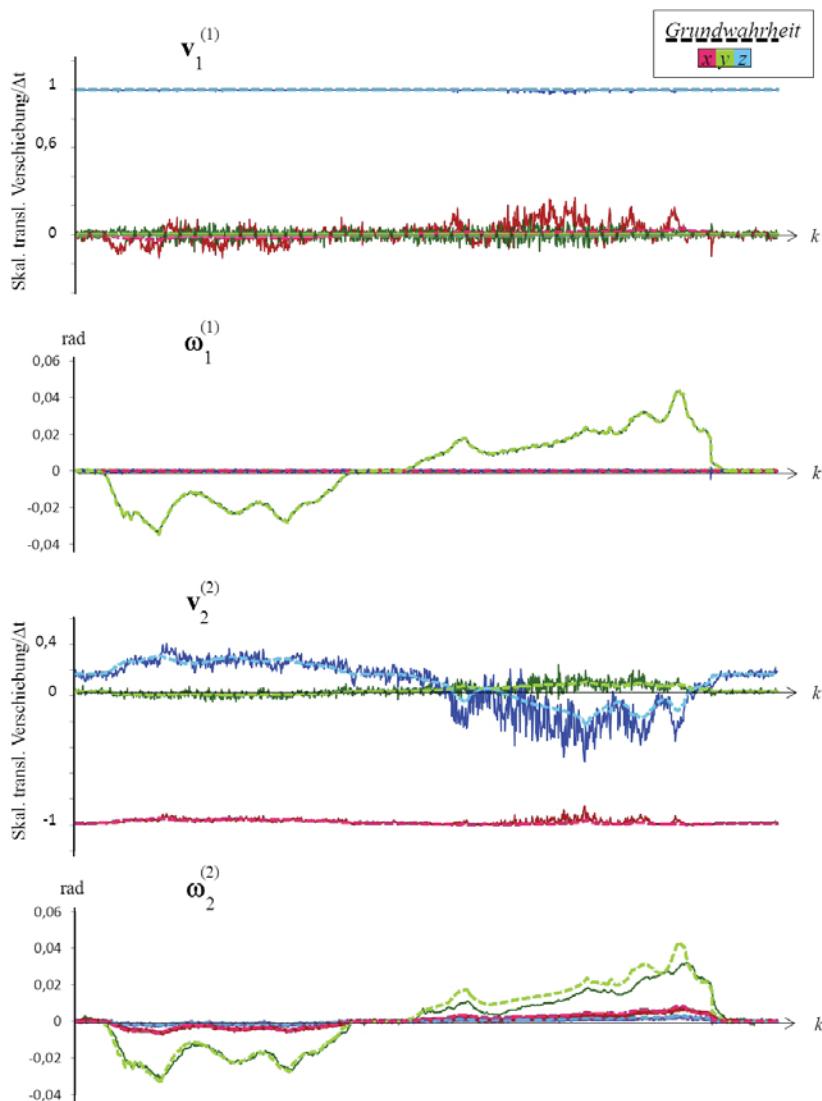
$\zeta_M(\mathbf{M}) = \mathbf{x}_M \in \mathbb{R}^6$  bzw.  $\zeta_C(\mathbf{C}) = \mathbf{x}_C \in \mathbb{R}^6$  seien hierbei die Parametervektoren mit den je drei translatorischen und drei rotatorischen Euklidischen Parametern.  $\mathbf{M}_i^{(j)}$  ( $= \mathbf{M}_{i_{r=0}}^{(j)}$ ) beschreibt die Bewegung von Modul  $\mathcal{M}_i$  in Abhängigkeit der Bewegung von Modul  $\mathcal{M}_j$ , der extrinsischen Parameter  $C_{ij}^G$  sowie der propagierten Unsicherheit  $\mathbf{P}_{\mathbf{M}_i}^{(j)}$ . Mit Hilfe von  $\mathbf{P}_{\mathbf{M}_i}^{(j)}$  lässt sich die Mahalanobisdistanz berechnen

$$d_{ij} = \sqrt{\left[ \mathbf{x}_{\mathbf{M}_i}^{(i)} - \mathbf{x}_{\mathbf{M}_i}^{(j)} \right]^T \mathbf{P}_{\mathbf{M}_i}^{(j)-1} \left[ \mathbf{x}_{\mathbf{M}_i}^{(i)} - \mathbf{x}_{\mathbf{M}_i}^{(j)} \right]} \quad (8.12)$$

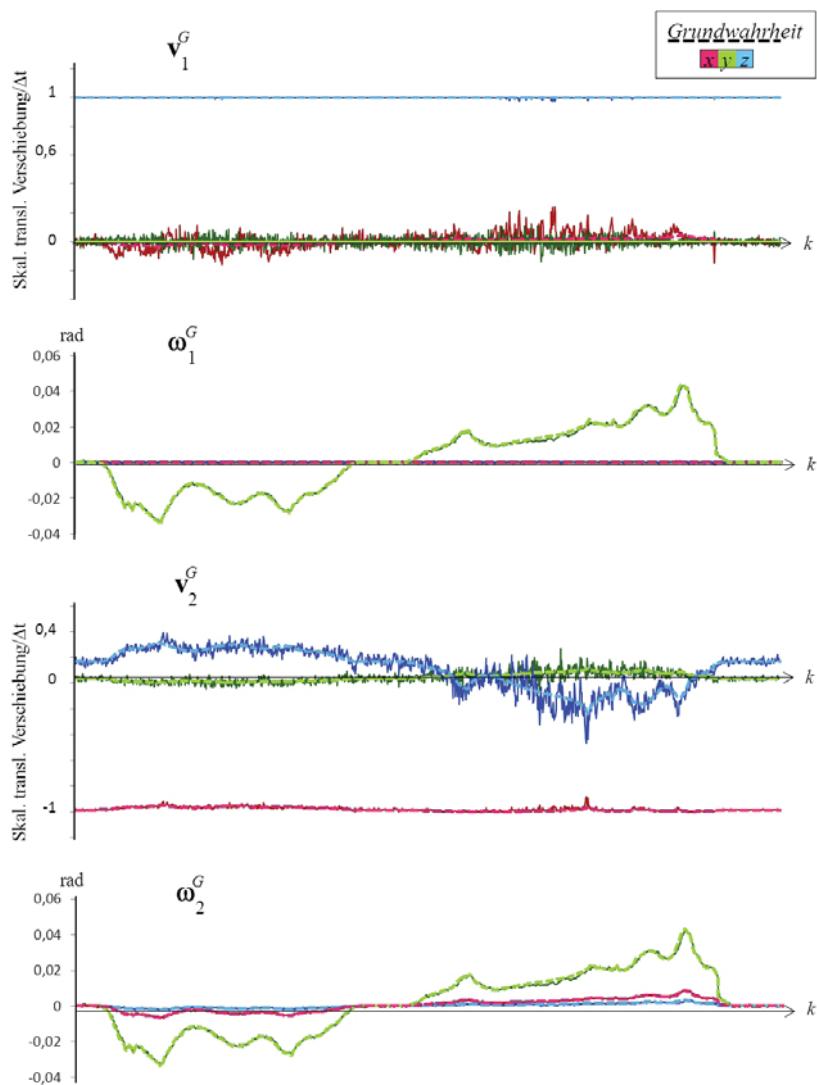
mit den Zustandsvektoren der propagierten Bewegungstransformationen  $\mathbf{x}_{\mathcal{M}_i}^{(i)} = \zeta_{\mathcal{M}}(\mathcal{M}_i^{(i)})$  und dem lokal berechneten Parametervektor  $\mathbf{x}_{\mathcal{M}_i}^{(j)} = \zeta_{\mathcal{M}}(\mathcal{M}_i^{(j)})$ . Diese Metrik berücksichtigt explizit die Unsicherheit der extrinsischen Parameter. Falls z. B. die Schätzung der extrinsischen Parameter stark von den wahren Werten abweicht, die Unsicherheit aber sehr groß ist (wie es in der Initialisierungsphase der Fall ist), dann bleibt  $d_{ij}$  klein, auch wenn der Euklidische Fehler  $\sqrt{[\mathbf{x}_{\mathcal{M}_i}^{(i)} - \mathbf{x}_{\mathcal{M}_i}^{(j)}]^T [\mathbf{x}_{\mathcal{M}_i}^{(i)} - \mathbf{x}_{\mathcal{M}_i}^{(j)}]}$  sehr groß wäre.  $d_{ij}$  bleibt auch dann klein, wenn die lokale Bewegung  $\mathbf{x}_{\mathcal{M}_j}^{(j)}$  sehr unsicher ist. Diese Eigenschaft ist wichtig, da sich bei großen Unsicherheiten der restlichen Modulbewegungen oder der extrinsischen Parameter keine Rückschlüsse über die Genauigkeit der Bewegung von  $\mathcal{M}_i$  ziehen lassen. In diesem Fall wäre  $\mathcal{M}_i$  aufgrund der vorherrschenden Unsicherheit konsistent zu den anderen Modulen. Das Konsistenzmaß für  $\mathcal{M}_i$  bezüglich aller anderen Module lässt sich dann formulieren gemäß

$$d_{M_i} = \frac{1}{\mathfrak{N}-1} \sum_{i \neq j}^{\mathfrak{N}} d_{ij}. \quad (8.13)$$

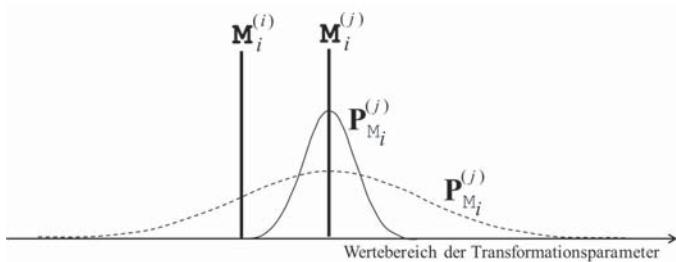
Abb. illustriert beispielhaft den Abstand zweier Zustände in Abhängigkeit der Unsicherheiten.



**Abbildung 8.4:** Geschätzte Bewegungsparameter ohne Fusion. Oben: Translation; Unten: Rotation. Die aus der Simulation bekannte Grundwahrheit ist gestrichelt eingezeichnet.



**Abbildung 8.5:** Geschätzte Bewegungsparameter mit Fusion. Oben: Translation; Unten: Rotation. Die aus der Simulation bekannte Grundwahrheit ist gestrichelt eingezeichnet.



**Abbildung 8.6:** Normalverteilte Zustände zweier lokaler, propagierter Bewegungstransformationen mit zwei beispielhaften Unsicherheiten  $P_{\mathbf{M}_i}^{(j)-1}$  für  $x_{\mathbf{M}_i}^{(j)}$ ; beispielhafte, eindimensionale Illustration.

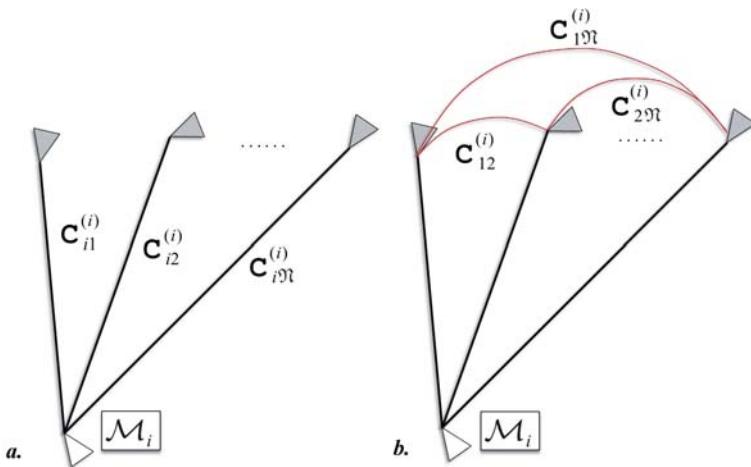
## 8.2 Globales Kalibrierschema

Der Fokus dieses Abschnitts liegt in der Systembeschreibung eines skalierbaren Fusionsansatzes zur kontinuierlichen, extrinsischen Kalibrierung eines Mehrkamerasytems. Das globale Kalibrierschema folgt demselben Prinzip wie das globale Bewegungsschema. Im ersten Schritt werden die Messungen der einzelnen Module zunächst lokal ausgewertet. Als Ergebnis dieses Schritts kennt ein Modul  $\mathcal{M}_i$  die lokalen Transformationen zu allen anderen Modulen, relativ zu  $\mathcal{M}_i$ , sowie deren Kovarianzmatrizen. Anschließend kann jedes Modul mittels Propagierung aus den modulbezogenen Transformationen einen vollständigen Transformationsgraphen berechnen. D. h. für jedes Modul existiert eine entsprechende Ausprägung der gesamten Kamerakonstellation bzw. den extrinsischen Transformationen zwischen allen Modulen. Zuletzt werden diese lokalen Schätzungen anhand ihrer Unsicherheit fusioniert. Zusammenfassend zeichnet sich für das globale Kalibrierschema folgendes Vorgehensmodell ab:

1. Lokale Schätzung der Anordnung der Module relativ zum Referenzmodul
2. Propagierung der Transformationen, zur Erstellung eines vollständigen Transformationsgraphens je Modul
3. Gauß'sche Fusion der extrinsischen Transformationsparameter

### 8.2.1 Lokale Kalibrierung

Bei der *lokalen* Kalibrierung eines Moduls  $\mathcal{M}_i$ , d. h. der Bestimmung der extrinsischen Transformationsparameter  $C_{i,1}, \dots, C_{i,N}$  basierend auf den lokalen Messungen von  $\mathcal{M}_i$  (sowie der bereits allen Modulen gleichermaßen bekannten globalen Zustandsparameter aus der Bewegungsschätzung), wird zur Bestimmung der notwendigen (Teil-)Transformationen bzw. Parameter eine Unterscheidung zwischen zwei grundsätzlichen Bewegungsarten gemacht: Einer rein translatorischen Bewegung, was bei Straßenfahrzeugen, die der Nicht-Holonomie-Bedingung gehorchen, einer Geradeausfahrt entspricht, und einer allgemeinen bzw. rotatorischen Bewegung. Die lokale



**Abbildung 8.7:** Schematische Darstellung der lokalen Kalibrierung von Modul  $\mathcal{M}_i$ . a. Lokale Schätzungen der extrinsischen Transformationen  $\mathbf{C}_{ij}^{(i)}, i \neq j$ . b. Erweiterung von Abb. a. durch die propagierten Transformationen (rot).

Kalibrierung beruht auf den in Kapiteln 5 bis 7 erläuterten Verfahren.

Bei der lokalen Berechnung der extrinsischen Parameter kann je nach Szenario auf das allgemeine oder erweiterte Kalibriermodell zurückgegriffen werden. Der Fokus der folgenden Ausführungen liegt auf dem erweiterten Kalibriermodell, da dieses auch für die spätere Evaluierung im Straßenfahrzeug relevant ist.

Abb. 8.7.a illustriert schematisch die lokalen Kalbriertransformationen, die von einem Modul  $\mathcal{M}_i$  berechnet werden.

Die einzelnen Teilprozesse werden im Folgenden in Anlehnung an Abschnitt 5 rekapituliert.

**Schätzung der Orientierung der Kameras bei Geradeausfahrten** Im Zuge der Kalibrierung wird eine Geradeausfahrt u. a. dazu genutzt, die relative Orientierung der Kameras zu schätzen (siehe Abschnitt 5.2.1). Prinzipiell lassen sich somit bereits alle Rotationsparameter  $\mathbf{r}_{ij}, i \neq j$ , zwischen allen Modulen bestimmen.

Im erweiterten Modell werden die Bewegungstransformationen  $M_i$  und  $M_j$  in die Bodenebene transformiert (siehe Anhang A.5). Eine relative Orientierungsschätzung auf den Bewegungen  $\underline{M}_i$  und  $\underline{M}_j$  ergibt den Eulerwinkel  $\beta_{ij}$  um die Hoch- bzw. Y-Achse des Weltkoordinatensystems. Die vollständige Bestimmung aller drei Eulerwinkel erfolgt anschließend durch die Zusammenführung der Teiltransformationen im IEKF (Kapitel 7).

**Schätzung der Bodenebene bei Geradeausfahrten** Bei rein translatorischer Bewegung kann man zur Schätzung der Bodenebene von dem Umstand, dass alle Module denselben Geschwindigkeitsbetrag haben, effektiv Gebrauch machen. Tatsächlich wird nicht die Bodenebene selbst, sondern die Transformation eines Moduls in die allen Modulen gemeinsame Bodenebene geschätzt. Die Transformationsschätzung wird nur dann durchgeführt, wenn die Rotationsachse des Moduls bereits bekannt ist (Abschnitt 5.3.1). Diese wird in dem RANSAC-Verfahren zur Selektion von Bodenpunkten dafür verwendet, Punktmengen für die Schätzung auszuschließen, welche in einer Ebene liegen, die nicht senkrecht zur Rotationsachse steht.

Die Schätzung der Bodenebene bei Geradeausfahrten anhand von 3D-Punkten innerhalb der Bodenebene erfolgt nur dann, wenn die Geschwindigkeit des Fahrzeugs bekannt ist (z. B. aus Odometriesensoren, siehe Abschnitt 5.3.4).

**Schätzung der Rotationsachse bei Kurvenfahrten** Mit Hilfe einer Historie an rotatorischen Bewegungen lässt sich die Rotationsachse der mobilen Plattform schätzen (Abschnitt 5.3.1). In der Rotationsachse stecken bereits Informationen über die Lage der Bodenebene, die bei der Schätzung der drei Freiheitsgrade der Transformation  $G$  in die Bodenebene eingesetzt wird (siehe Abschnitt 5.3.3). Die Rotationsachse kann auch verwendet werden, um Bewegungstransformationen in die Grundebene zu transformieren (siehe Abschnitt A.5).

**Bewegungsausgleich bei Kurvenfahrten** Der dünne Bewegungsausgleich wird – wie die Bestimmung der Rotationsachse – ebenfalls anhand einer rotatorischen Be-

wegungshistorie berechnet. Im erweiterten Modell wird an dieser Stelle ein Zwischen-schritt durchgeführt. Noch vor Berechnung des SMA berechnet man die skalierte Bewe-gung des zentralen Bewegungsbezugspunkts  $\mathcal{C}$  des Fahrzeugs. Entscheidend ist hierbei die Erfüllung der Nicht-Holonomie-Bedingung bezüglich der Plattformbewegung. Nach der Berechnung der Bewegungstransformationen eines zentralen Bezugspunkts  $\mathcal{C}$  be-rechnet nun jedes der  $N$ Module  $\mathcal{M}_i$   $N$  Bewegungsausgleiche zur Bestimmung von  $\underline{\mathcal{C}}_i^{(i)} \mathcal{C}$ . Das Ergebnis ist die relative Anordnung der Module innerhalb der Bodenebene relativ zum Bezugspunkt  $\mathcal{C}$ . Die extrinsischen, planaren Transformationen zwischen den Modulen lassen sich einfach aus den bezüglich  $\mathcal{C}$  referenzierten Transformationen überführen.

**Zusammenführen der Teilinformationen** Die extrinsischen Parameter im 3D-Raum, d. h. mit sechs Freiheitsgraden je Transformation, werden temporär durch ein iteratives erweitertes Kalmanfilter geschätzt. Die Schätzung erfolgt durch Minimierung des Projektionsfehlers (siehe Kapitel 7), welcher in Abhängigkeit zu den globalen Bewe-gungszuständen der Module sowie den zuvor mittels SMA ermittelten extrinsischen Parametern steht.

Das erweiterte Kalibriermodell erfordert als Folge der Berechnung des SMA in der Bodenebene die Zusammenführung der Transformationverkettungen<sup>30</sup>. Dies verlangt zusätzlichen Kommunikationsaufwand, da jedes Modul  $\mathcal{M}_i$  Kenntnis von den Boden-transformationen  $\mathbf{G}_j$  aller anderen Module haben muss, um die Zusammenführung via Kalmanfilterung durchzuführen. Für das Kalmanfilter stellt die Integration des er-weiterten Modells keine größere Schwierigkeit dar, da die einzelnen Transformationen schlicht als zusätzliche Messung in das Kalmanmodell mit aufgenommen werden. Zu-sätzlich muss die Bedingungsgleichung des Kalmanfilters  $\mathbf{h}_C$  entsprechend Abschnitt 7

---

<sup>30</sup> Im allgemeinen Kalibriermodell ist aufgrund der allgemeinen Bewegung, mit rotatorischen Freiheitsgraden um alle drei räumlichen Achsen, das Ergebnis der Bewegungsschätzung als Eingangsmessung für den Bewegungsausgleich ausreichend.

erweitert werden.

Aus den lokalen Berechnungen geht hervor, dass jedes Modul  $\mathcal{M}_i$  lediglich die extrinsischen Transformationen zu allen anderen Modulen berechnet. Ergebnis der lokalen Berechnungen sind die extrinsischen Transformationen eines Moduls  $\mathcal{M}_i$  zu allen anderen Modulen  $\mathcal{M}_{j \neq i}$ . Somit muss jedes Modul  $\mathfrak{N} - 1$  Kalman-Aktualisierungen berechnen. Das Resultat sind die lokalen extrinsischen Transformationsparameter  $C_{ij}^{(i)}, i \neq j$ . Am Ende der lokalen Kalibrierung liegt also für jedes Modul eine auf sich referenzierte Transformationstopologie vor (s. Abb. 8.7.a).

### 8.2.2 Propagierung und Fusion der Kalibrierparameter

Am Ende der lokalen Kalibrierung kennt jedes Modul  $\mathcal{M}_i$  die Transformationen zwischen sich und den anderen Modulen (Abb. 8.7.b), also  $C_{ij}^{(i)}, j = 1, \dots, \mathfrak{N}$ . Hat ein Modul lokale Kenntnis von den Transformationen  $C_{i1}^{(i)}, \dots, C_{i\mathfrak{N}}^{(i)}$ , so kennt es implizit auch die Transformation  $C_{hj \neq i}^{(i)}$  durch Propagierung der bekannten Transformationen (Abb. 8.8.a)

$$C_{hj}^{(i)} = C_{ih}^{(i)-1} \cdot C_{ij}^{(i)}. \quad (8.14)$$

Mit Hilfe von Gl. 8.14 lassen sich mittels Fehlerpropagierung nicht nur die fehlenden Transformationen ergänzen, sondern auch deren Kovarianzmatrizen

$$\mathbf{P}_{C_{hj}}^{(i)} = \mathbf{J}_{C_{ih}} \mathbf{P}_{C_{ih}}^{(i)} \mathbf{J}_{C_{ih}}^T + \mathbf{J}_{C_{ij}} \mathbf{P}_{C_{ij}}^{(i)} \mathbf{J}_{C_{ij}}^T \quad (8.15)$$

mit

$$\mathbf{J}_{C_{ih}} = \frac{\partial \zeta_C(C_{hj}^{(i)})}{\partial \mathbf{x}_{C_{ih}}} \quad , \quad \mathbf{J}_{C_{ij}} = \frac{\partial \zeta_M(C_{hj}^{(i)})}{\partial \mathbf{x}_{C_{ij}}} \quad (8.16)$$

und  $\mathbf{x}_{C_{ih}} = \zeta_C(C_{ih})$ ,  $\mathbf{x}_{C_{ij}} = \zeta_C(C_{ij})$ . Jedes Modul hat nun also nach der Propagierung Kenntnis von dem vollständigen Transformationsnetz, also insgesamt  $\frac{(\mathfrak{N}-1)(\mathfrak{N}-2)}{2}$  extrinsischen Transformationen. Das bedeutet wiederum, da jedes Modul seine eigene

Schätzung unabhängig von den lokalen Schätzungen der andern Module hat, dass für jede extrinsische Transformation  $\mathfrak{N}$  redundante Schätzungen vorliegen.

Damit die redundanten Schätzungen zu einer einheitlichen Schätzung fusioniert werden können, werden die Schätzergebnisse an alle Module versendet, so dass jedes Modul Kenntnis von den lokalen, propagierten Zuständen aller anderen Module hat. Die  $\mathfrak{N}$  redundanten Schätzungen von jeder der  $\frac{(\mathfrak{N}-1)(\mathfrak{N}-2)}{2}$  extrinsischen Transformationen werden wie bei der Bewegungsfusion zusammengeführt (siehe Abb. 8.8). Auch hier bedienen wir uns der Gauß'schen Fusion aus Abschnitt 2.6 und vereinen die Transformationen gemäß

$$\mathbf{C}_{hj}^G = \bigcup_{i=1}^{\mathfrak{N}} \mathbf{C}_{hj}^{(i)} \quad (8.17)$$

$$\mathbf{P}_{\mathbf{C}_{hj}}^G = \bigcup_{i=1}^{\mathfrak{N}} \mathbf{P}_{\mathbf{C}_{hj}}^{(i)} \quad (8.18)$$

Bei der Fusion der extrinsischen Transformationen zeigt sich der Vorteil der Berücksichtigung der nichtholomen Bewegungsbedingung (siehe Abschnitt 2.2 und 6.2). Für mobile Plattformen lässt sich ein zentraler Bezugspunkt  $\mathcal{C}$  der Plattform bestimmen. D. h. beim Bewegungsausgleich berechnen alle Module die extrinsischen Parameter relativ zu diesem Bezugspunkt, anstatt die Eigenbewegung als Referenz heranzuziehen. Der Vorteil dieser Vorgehensweise ist folgender:

Beim Bewegungsausgleich werden die Geschwindigkeiten aller Module normiert bzw. auf einen Odometriewert  $v$  gesetzt. Berechnen die Module die extrinsischen Parameter nun *nicht* relativ zu einem allen gemeinsamen Bezugspunkt, so sind die relativen Geschwindigkeiten der Module stets auf das jeweilige Modul referenziert (s. Abb. 8.9). Dies bedeutet, dass der Translationsvektor der resultierenden lokalen extrinsischen Parameter ebenfalls auf den (normierten) Geschwindigkeitsbetrag des entsprechenden Moduls referenziert ist. Somit haben alle lokal propagierten Transformationen zwischen je zwei Modulen unterschiedliche Translationsbeträge. Es lassen sich folglich nur noch die Translationsrichtungen und die Orientierungen fusionieren, nicht aber die

Abstände zwischen den Modulen. D. h. selbst bei bekannter Odometrie, also der Fahrzeuggeschwindigkeit als externe, zusätzliche Messung, ließe sich rein theoretisch der absolute Betrag der Translationsvektoren der extrinsischen Kalibriertransformationen nicht mehr rekonstruieren. Ein allen Modulen gemeinsamer Bewegungsbezugspunkt (siehe Abschnitt 6.2) bei einer dem Einspurmodell folgenden Bewegungsannahme (man vergleiche die nichtholonom Bewegungsbedingung in Abschnitt 2.2) löst diese Mehrdeutigkeit auf. Dieser Sachverhalt ist in Abb. 8.9 veranschaulicht.

Ist die absolute Geschwindigkeit der Plattform nicht bekannt, so lässt sich die extrinsische Anordnung ohnehin nur bis auf einen skalaren Faktor bestimmen. In diesem Fall kann z. B. der skalare Freiheitsgrad dadurch fixiert werden, dass ein Kamerapaar, beispielsweise die Module  $\mathcal{M}_1$  und  $\mathcal{M}_2$ , als Referenzpaar definiert wird, mit  $\|\mathbf{t}_{12}\| = 1$ .

### 8.2.3 Zusammenfassung: Bewegungsphasen und Systementwurf

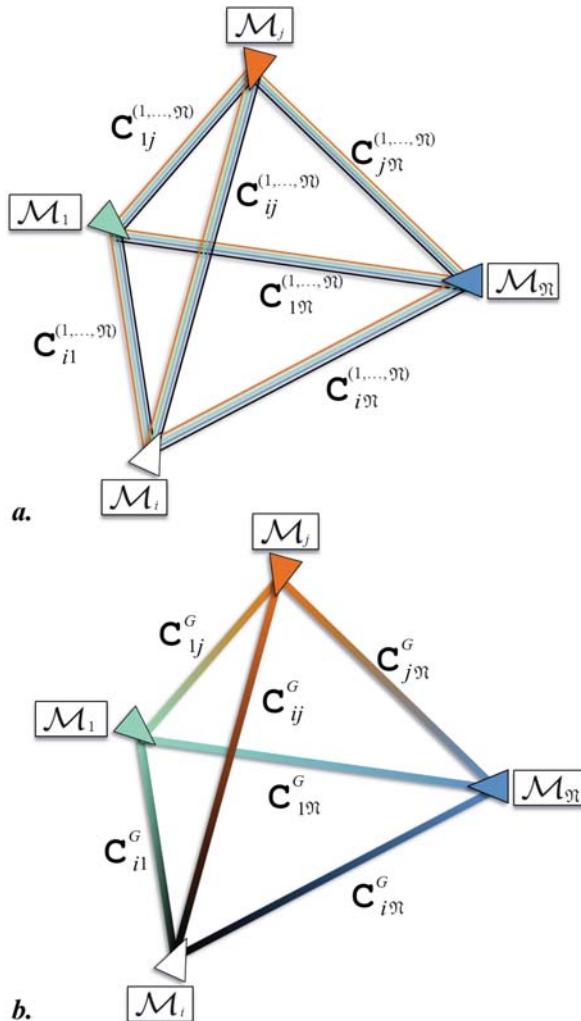
Aus den vorangegangenen Abschnitten ergibt sich zusammenfassend, dass das hier beschriebene Kalibrierverfahren in vier Bewegungsphasen unterteilbar ist, wie in Tabelle 8.1 dargestellt.

Ab Phase 4 ist das System in der Lage, die Kalibrierparameter kontinuierlich zu filtern. Dies bedeutet (insbesondere für die Experimente in Kapitel 9), dass mit dem hier vorliegenden Systemmodell jede Kalibrierfahrt aus einer Abfolge von einer Gerdaeausfahrt, einer Kurvenfahrt, einer Geradeausfahrt und wieder gefolgt von einer Kurvenfahrt bestehen muss. Da dies allerdings so elementare Bewegungen und sehr allgemein formulierten Forderungen sind, stellt dies keine wesentliche Einschränkung dar.

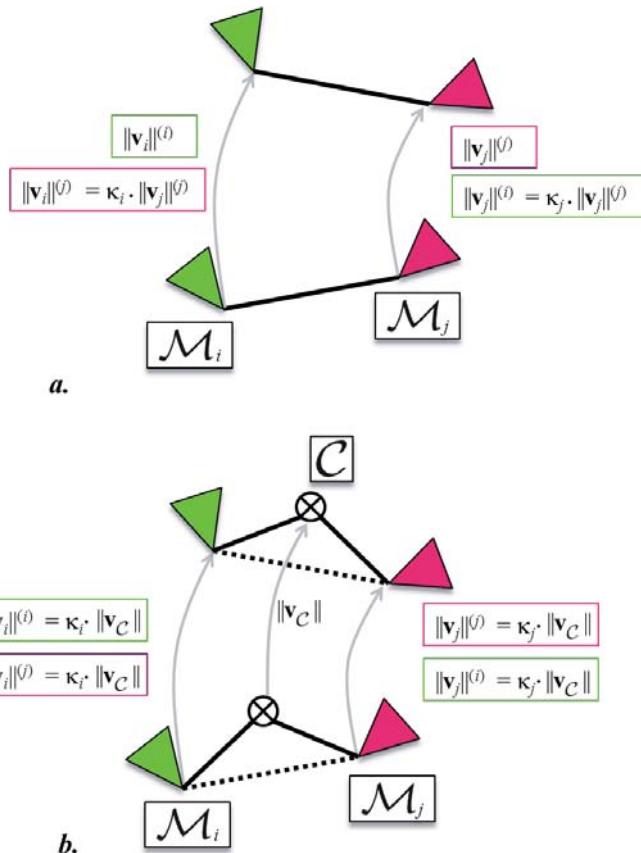
Die Flussdiagramme in Abb. 8.10 und 8.11 veranschaulichen den implementierten Algorithmus zur bewegungsbasierten Kalibrierung und Fusion der redundanten Informationen. Die einzelnen Rechenschritte des Algorithmus sind in Kästen, deren Ausgabe kursiv in Parallelogrammen dargestellt. Entscheidungsknoten sind grau und Kommunikationsschnittstellen zum Austausch lokaler Informationen sind gelb hinterlegt.

Bewegungsphase	Durchgeführte Berechnung	Verwendung für...
1. Gerade	Initialisierung der lokalen Bewegungsschätzungen zu Beginn des Verfahrens (Anhang A.4)	die Initialschätzung des Bündelausgleichs (Abschnitt 8.1.2)
2. Kurve	Berechnung der Rotationsachsen der Kameramodule (Abschnitt 5.3.1)	die Transformation der Bewegungstransformationen in die Grundebene (Anhang A.5); die Selektion von Bodenpunkten (Abschnitt 5.3.2)
3. Gerade	Berechnung der Transformationen in die Bodenebene (Abschnitt 5.3.3); Berechnung der relativen Kameraorientierungen (Abschnitt 5.2.1)	den Bewegungsausgleich (Kapitel 6); das Zusammenführen der Teiltransformationen und Zustandsfilterung (Kapitel 7)
4. Kurve	Bewegungsausgleich und Zustandsfilterung (Kapitel 6 und 7)	die Zustandsschätzung der Kalibrierparameter; die Fusion der extrinsischen und Bewegungstransformationen (Abschnitt 8.2 und 8.1.3)

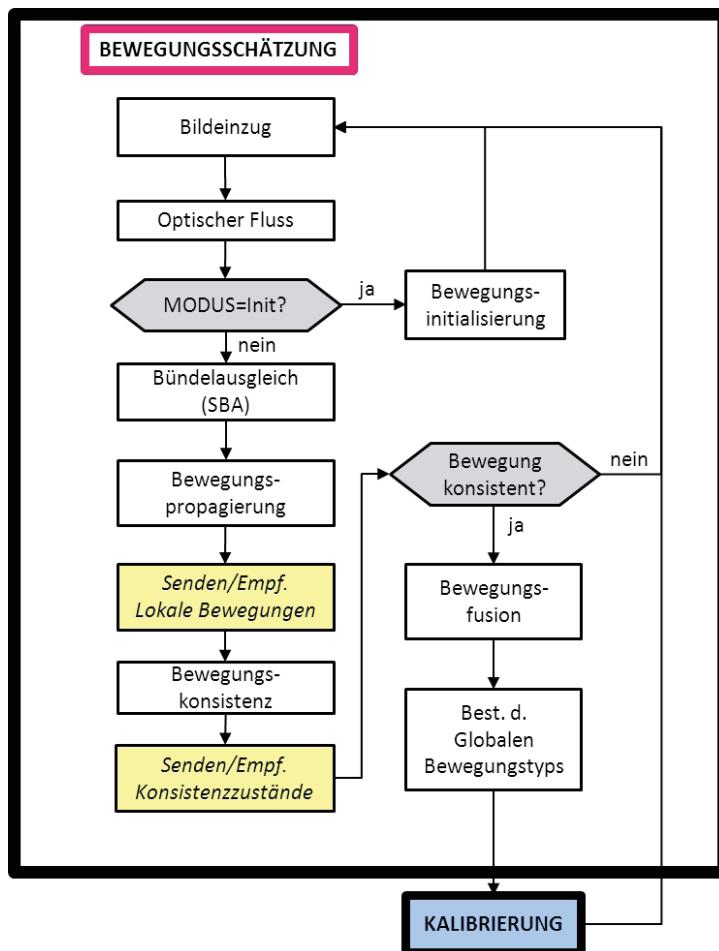
**Tabelle 8.1:** Bewegungsphasen der Kalibrierung. Die ersten drei Phasen dienen der Initialisierung des Systems. Ab Phase 4 erfolgt die reguläre kontinuierliche Schätzung der Zustandsparameter im Kalmanfilter.



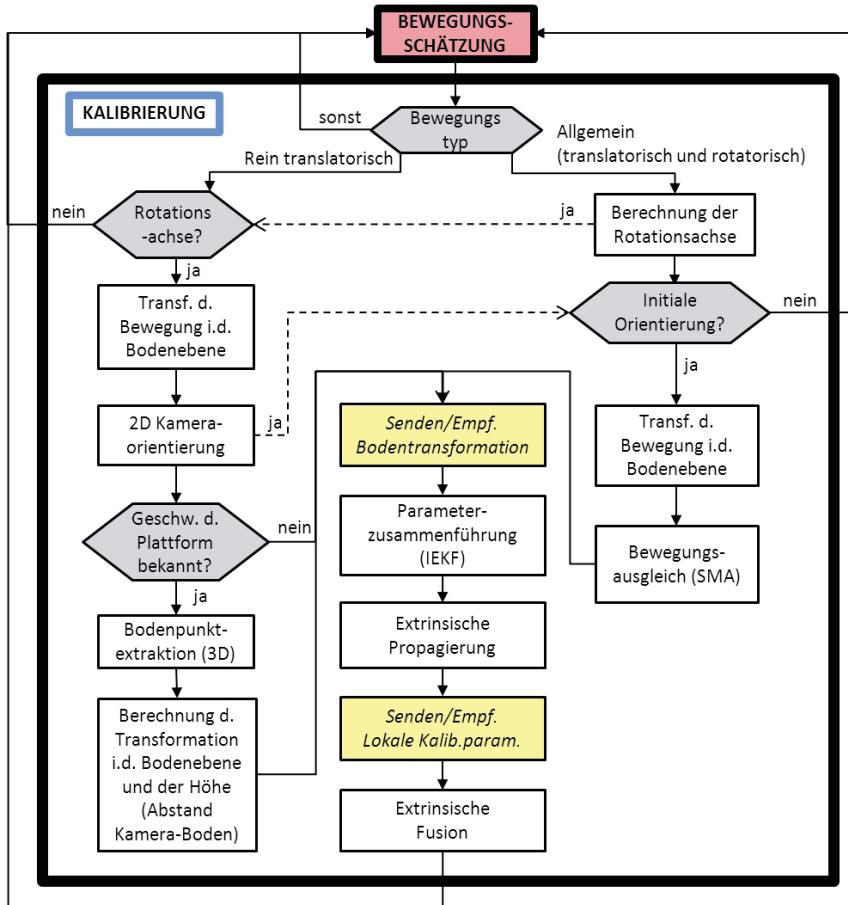
**Abbildung 8.8:** a. Redundante, lokale Kalibriertransformationen nach der Propagierung. Bei  $\mathfrak{N}$  Modulen existieren je Modulpaar  $\mathfrak{N}$  redundante Transformationen. b. Fusionierte, globale Transformationskonstellation für  $\mathfrak{N}$  Module.



**Abbildung 8.9:** a. Ergebnis des Bewegungsausgleichs ohne Bezug zum zentralen Bezugs-  
punkt  $C$ . Die beim Bewegungsausgleich geschätzten Skalare  $\kappa_i, \kappa_j$  sind jeweils relativ  
zu den einzelnen Modulen (der Skalar jedes Moduls  $\mathcal{M}_i$  entspricht stets der Referenz-  
geschwindigkeit für die Berechnung der lokalen Schätzungen von  $\mathcal{M}_i$ , sodass  $\kappa_i^{(i)} = 1$ ).  
b. Ergebnis des Bewegungsausgleichs relativ zu  $C$ . Alle Module haben dieselbe relative  
Skalierung (also  $\kappa_i^{(i)} = \kappa_i^{(j)}$  und  $\kappa_j^{(i)} = \kappa_j^{(j)}$ ).



**Abbildung 8.10:** Flussdiagramm der Implementierung des globalen Bewegungsschätzungsschemas.



**Abbildung 8.11:** Flussdiagramm der Implementierung des globalen Kalibrierschemas. Die gestrichelten Linien bewirken eine Aktivierung der grauhinterlegten Entscheidungskästen.

## **Teil III**

# **Experimentelle Ergebnisse und Bewertung**



# 9

---

## Experimentelle Ergebnisse

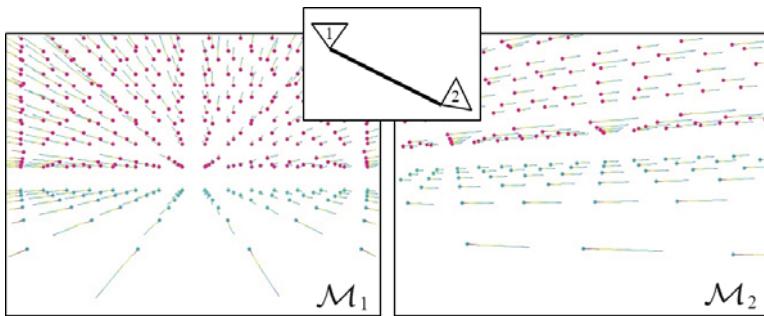
---

### 9.1 Ergebnisse mit simulierten Daten

Zur Demonstration der allgemeinen Funktionsfähigkeit des erweiterten Kalibriermodells und des ganzheitlichen Fusionsansatzes, wurde die Fahrt einer Zwei-Kamera-Plattform simuliert. Die Szene wurde hierbei als gleichverteilte 3D-Punktwolke im Raum sowie in der Bodenebene modelliert, sodass spezifische Szenenkonstellationen nicht berücksichtigt werden mussten. Die Projektionen der Szenenpunkte, aus denen der Optische Fluss berechnet wird (s. Abb. 9.1), und die somit Grundlage für die Berechnung des SBA sowie der Filterung der Kalibrierparameter sind, wurden mit einem additiven, normalverteilten Rauschen mit der Varianz  $\sigma^2 = 0,5$  Pixel behaftet. Darüber hinaus wurden 10% Ausreißer hinzugefügt. Für die allgemeinen Systemparameter gelten die Werte  $\mathfrak{H} = 100$ ,  $\Gamma = 5$ ,  $\mathfrak{P} = 300$ .

Simuliert wurden zwei Bewegungsarten. Bei der ersten bewegt sich die Plattform spiralförmig in der Ebene (Abb. 9.2), d. h. mit linear zunehmender rotatorischer Bewegung. Die zweite Simulationsfahrt beschreibt eine Rechts- und eine Linkskurve. Abb. 9.5 zeigt hierzu die geschätzten Bewegungsparameter.

Abb. 9.3 bis 9.7 dokumentieren die Auswerteergebnisse. In beiden Szenarien konnten



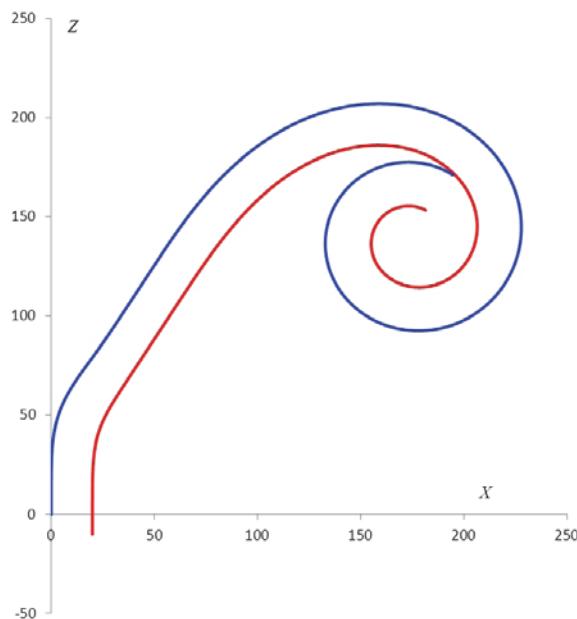
**Abbildung 9.1:** Simulierte, unverrauschte Flussvektoren zweier Module. Mitte: Schematische Draufsicht auf die Kameraanordnung.

die extrinsischen Rotationsparameter gut bestimmt werden. Während bei der Spiralfahrt die Kalibrierparameter noch zu den wahren Werten konvergieren (Abb. 9.3), ist dies bei der zweiten Simulation nicht der Fall (Abb. 9.6). Vielmehr entfernen sich insbesondere die translatorischen Parameter mit zunehmender Dauer der Schätzung wieder von der Grundwahrheit.

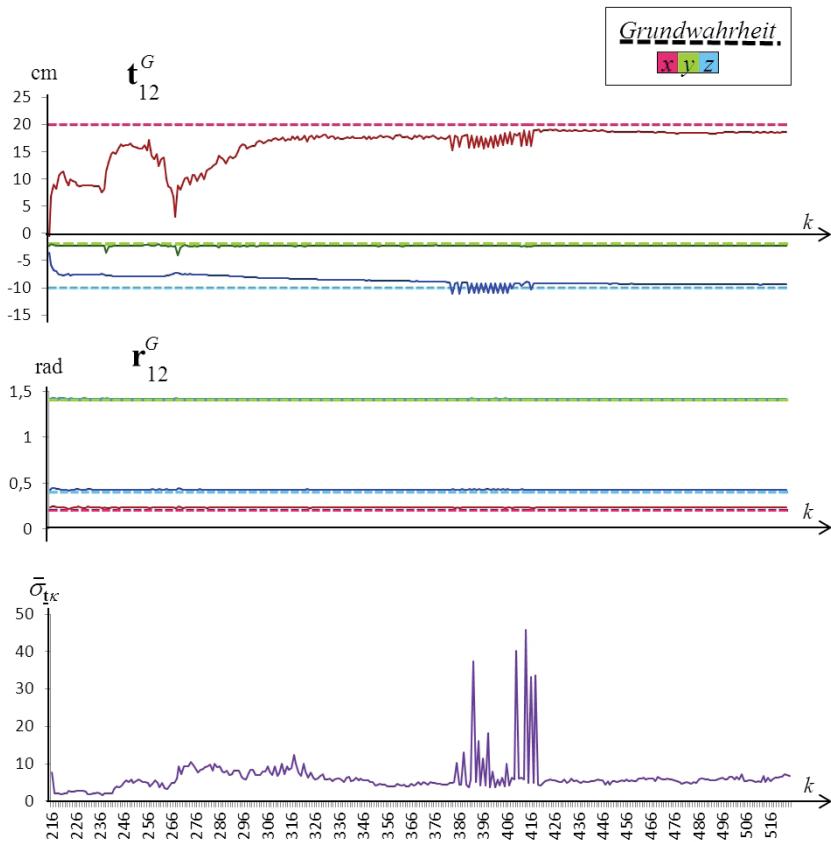
Man sieht aus Abb. 9.8, dass die in Abb. 9.6 dargestellten translatorischen Parameter wesentlich von den Ergebnissen des Bewegungsausgleichs abhängen. Während der Berechnung des Bewegungsausgleichs wird die Bodenebene nicht geschätzt (dies erfolgt in der vorausgegangenen rotationsfreien Bewegungsabschnitt, vgl. Abschnitt 8.2.3). Die Bodenebene selbst konnte zuverlässig geschätzt werden (Abb. 9.9).

Weiter erkennt man auch, dass die Korrelationswerte  $\tilde{\sigma}_{tk}$  aus Gl. 6.25 mit zunehmender Abweichung der Translationsparameter von der Grundwahrheit ansteigen. Eine ähnliche Beobachtung lässt sich auch in Abb. 9.3 machen. Ein zwischenzeitliches Flattern der translatorischen Parameterschätzwerte spiegelt sich in der  $\tilde{\sigma}_{tk}$ -Kurve. Dies wiederum lässt sich auf die Erkenntnisse aus Abschnitt 6.3 zurückführen.

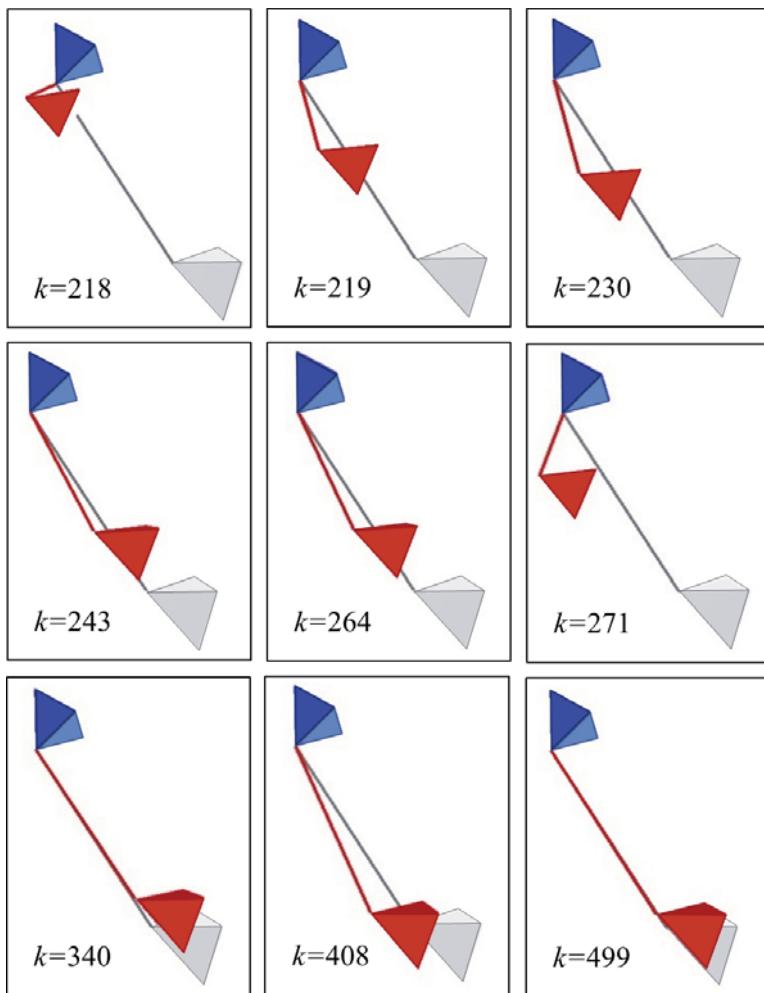
Abb. 9.4 und 9.7 zeigen eine perspektivische Visualisierung der geschätzten extrinsischen Parameter aus Abb. 9.3 bzw. 9.6 der beiden Kameramodule und deren Grundwahrheit.



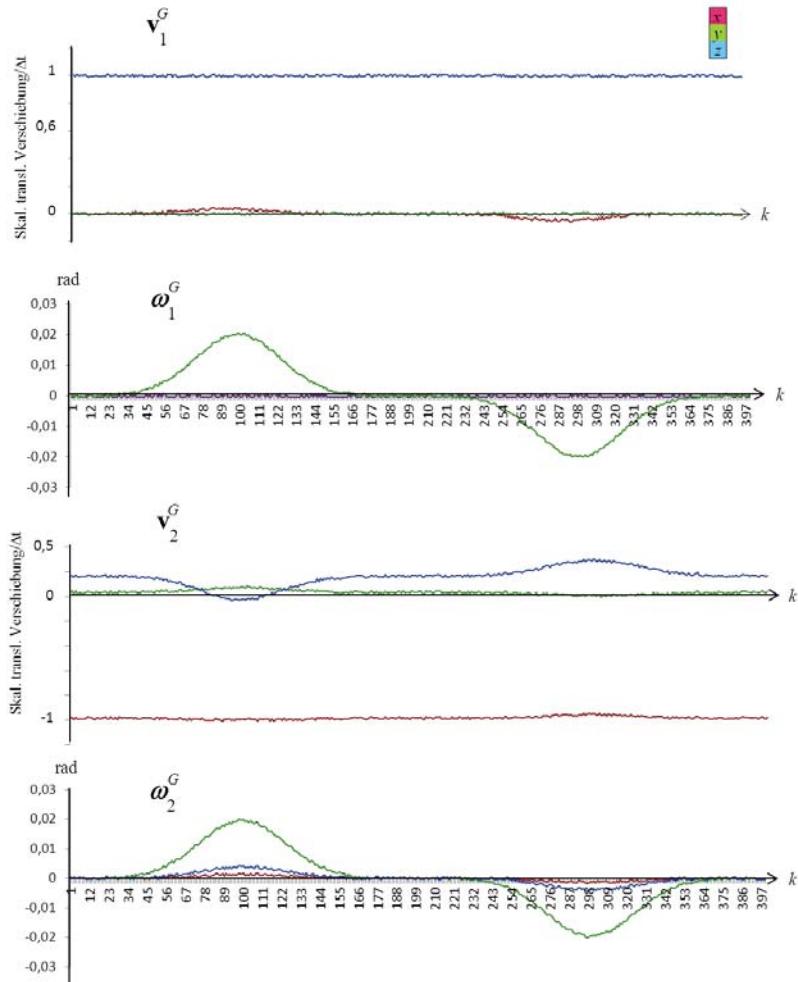
**Abbildung 9.2:** Trajektorie einer simulierten Spiralfahrt (linear ansteigende rotatorische Bewegung in der Grundbene) einer Zwei-Kamera-Anordnung. Die erste Kurve dient der Initialisierung der Rotationsachsen, das anschließende Geradenstück der Initialisierung der Orientierung der beiden Kameras.



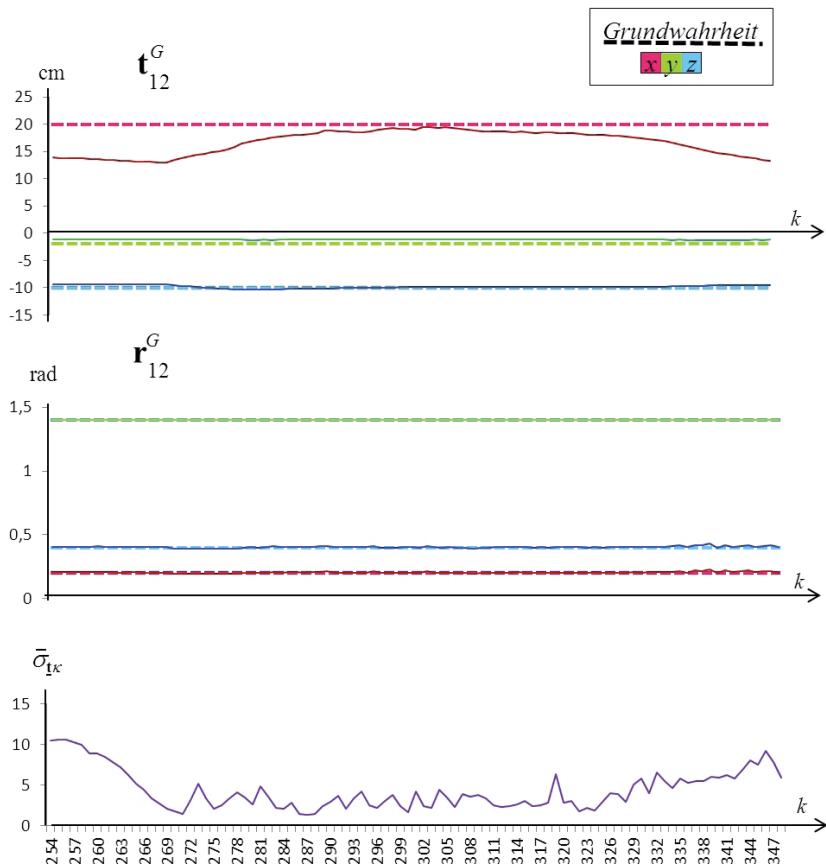
**Abbildung 9.3:** Ergebnis der der Spiral-Simulation: Geschätzte extrinsische Parameter (Translation  $\mathbf{t}$  und die Eulerwinkel  $\mathbf{r}$ ) und Korrelationskenngröße  $\bar{\sigma}_{\mathbf{t}\mathbf{x}}$  zwischen den Modulen  $\mathcal{M}_1$  und  $\mathcal{M}_2$ .



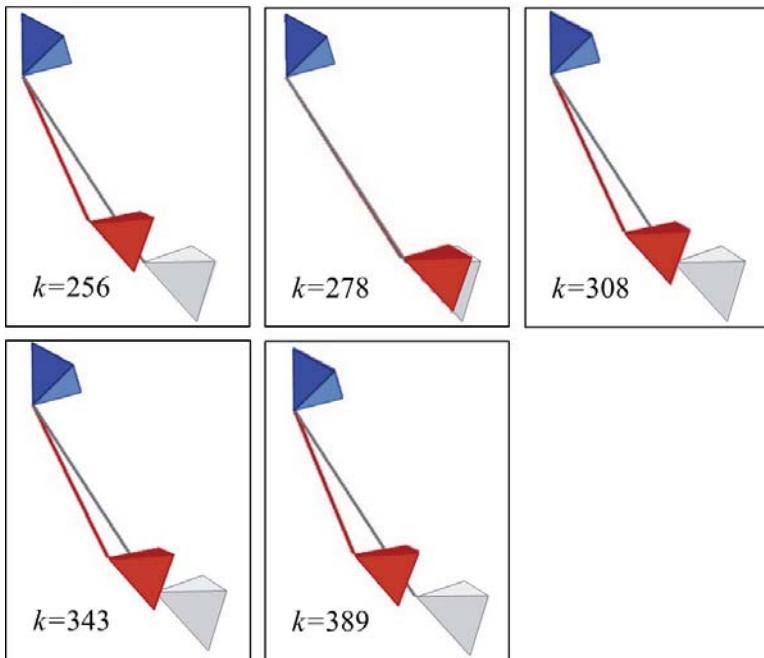
**Abbildung 9.4:** Perspektivische Darstellung der Ergebnisse aus Abb. 9.3 (Spirale-Simulation) mit Modul  $\mathcal{M}_1$  (=Referenzmodul, blau), der geschätzten Lage und Orientierung von  $\mathcal{M}_2$  (rot), sowie der Grundwahrheit von  $\mathcal{M}_2$  (grau).



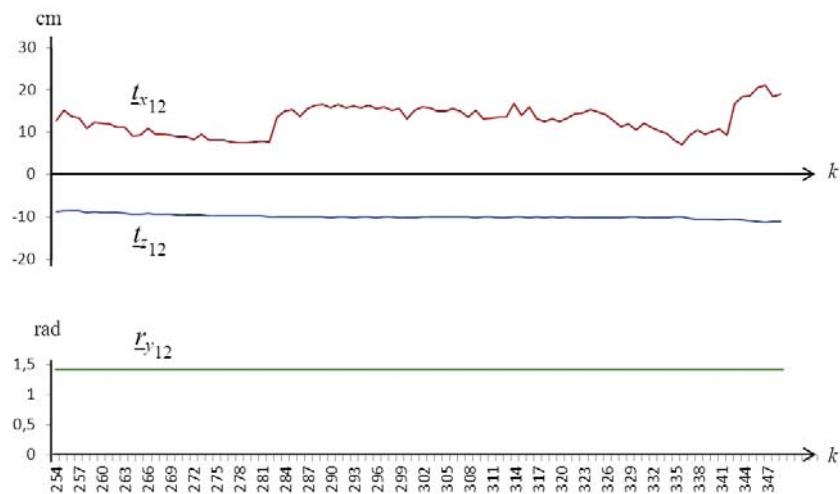
**Abbildung 9.5:** Geschätzte Bewegungsparameter der Zwei-Kurven-Simulation für die Kameramodule  $\mathcal{M}_1$  (links) und  $\mathcal{M}_2$  (rechts) mit skalierter Translation (oben) und Eulerwinkeln (unten).



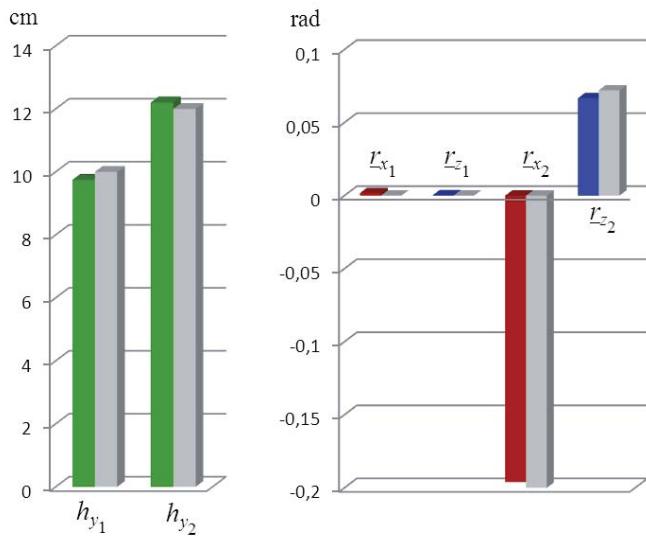
**Abbildung 9.6:** Ergebnis der der Zwei-Kurven-Simulation: Geschätzte extrinsische Parameter (Translation  $\mathbf{t}$  und die Eulerwinkel  $\mathbf{r}$ ) und Korrelationskenngröße  $\bar{\sigma}_{\mathbf{t}k}$  zwischen den Modulen  $\mathcal{M}_1$  und  $\mathcal{M}_2$ .



**Abbildung 9.7:** Perspektivische Darstellung der Ergebnisse aus Abb. 9.6 (Zwei-Kurven-Simulation) mit Modul  $\mathcal{M}_1$  (=Referenzmodul, blau), der geschätzten Lage und Orientierung von  $\mathcal{M}_2$  (rot), sowie der Grundwahrheit von  $\mathcal{M}_2$  (grau).



**Abbildung 9.8:** Geschätzte Parameter des Bewegungsausgleichs ( $\underline{C}$ , zwei translatorische un ein rotatorischer Parameter) der der Zwei-Kurven-Simulation zwischen den Modulen  $\mathcal{M}_1$  und  $\mathcal{M}_2$ .



**Abbildung 9.9:** Geschätzte Parameter des Bodenebene ( $G$ ) der der Zwei-Kurven-Simulation der Module  $\mathcal{M}_1$  und  $\mathcal{M}_2$ .

## 9.2 Versuche mit realen Messdaten

### 9.2.1 Referenzkalibrierung



**Abbildung 9.10:** Testfahrzeug mit Kameraaufbau.

Zum Testen der Kalibrierung mit Realdaten wurden auf dem Betriebsgelände des IOSB mit einer auf einem PKW montierten 3-Kamera-Anordnung Sequenzen aufgenommen und ausgewertet (s. Abb. 9.10). Die drei Kameras haben im folgenden die Bezeichnung  $\mathcal{M}_1$ ,  $\mathcal{M}_2$  und  $\mathcal{M}_3$ . Um die Ergebnisse der kontinuierlichen Kalibrierung vergleichbar zu machen, wurde die Kameraanordnung zunächst statisch kalibriert. Die intrinsische Kalibrierung zur Bestimmung der Kameramatrizen  $\mathbf{K}_i$  sowie die extrinsische Kalibrierung, d. h. die Bestimmung der relativen Lage und Orientierung der Kameras, erfolgen beide musterbasiert [Zha00], [Pag10b] (vergleiche Anhang A.7). Die Verwendung eines Kalibriermusters mit bekannten Ausmaßen ermöglicht die Bestim-

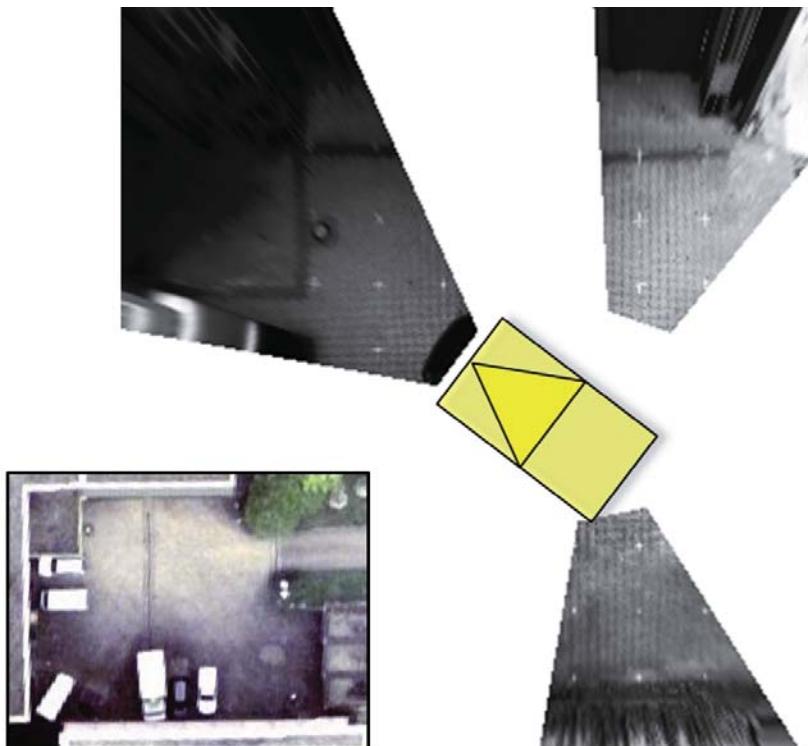
mung der absoluten Abstände zwischen Kameras. Da vom Testfahrzeug keine odometrischen Daten verfügbar waren, können die translatorischen Beträge zwischen den Kameras lediglich relativ zueinander ermittelt werden. In der folgenden Auswertung geschieht dies relativ zum Kamerapaar  $\{\mathcal{M}_1, \mathcal{M}_2\}$ . Darüber hinaus bleibt der Parameter  $t_y$ , der longitudinale Abstand zwischen den Kameras, unbestimmt (man vergleiche Abschnitt 5.3.4).

Der verbleibende Restfehler der Referenzkalibrierung ergibt sich aus dem Mittel der Differenz der für die Kalibrierung verwendeten Bildpunkte (bzw. der Abbildung des Musters in der Bildebene) mit der Projektion der Szenenpunkte des Musters in die Bildebene der jeweiligen Kamera in Abhängigkeit der extrinsischen Parameter relativ zum Koordinatensystems des Musters. Der Projektionsfehler beträgt 1,68 Pixel. Abb. 9.11 zeigt den Innenhof auf dem Gelände des Fraunhofer IOSB Karlsruhe, der zur Kalibrierung verwendet wurde. Die Referenzparameter sind in Tabelle 9.1 aufgeführt.

Modulpaar	$t_x$	$t_y$	$t_z$	$[r_x] = \text{rad}$	$[r_y] = \text{rad}$	$[r_z] = \text{rad}$
$1 \rightarrow 2$	-0,39	0,16	-0,91	-0,339	-2,612	0,074
$1 \rightarrow 3$	0,71	0,22	-1,19	-0,448	2,441	0,178

**Tabelle 9.1:** Referenzwerte aus der musterbasierten Initialkalibrierung. Die Translationsvektoren sind relativ zum (Referenz-)Translationsvektor  $\mathbf{t}_{12}$  mit  $\|\mathbf{t}_{12}\| = 1$  skaliert und daher ohne Einheitsangabe.

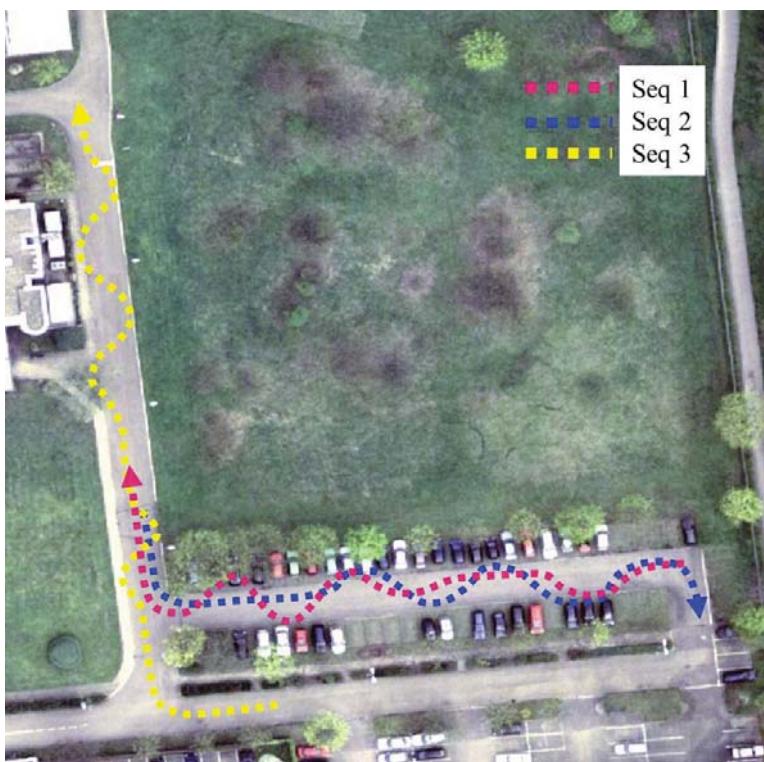
Da die Kameras in der vorliegenden Messkampagne komplanar angeordnet waren, wurde der Höhenparameter  $t_y$  fixiert und die Höhenparameter  $h_{y_i}$  der jeweiligen Module der Referenzkalibrierung entnommen. Somit fokussiert sich die Auswertung auf die Bestimmung der rotatorischen Parameter  $r_x, r_y, r_z$  und den (relativen) extrinsischen Parametern  $t_x, t_z$ . Die Systemparameter wurden hierfür auf die Werte  $\mathfrak{H} = 100$ ,  $\Gamma = 5$ ,  $\mathfrak{P} = 300$  gesetzt mit den initialen Zustandsparametern  $\mathbf{t}_{ij} = \mathbf{r}_{ij} = \mathbf{0}$ . Abb. 9.10 zeigt den Versuchsträger und die drei Sensoren.



**Abbildung 9.11:** Links; Luftbild vom Gelände des Fraunhofer IOSB in Karlsruhe. Rechts: Nach der Kalibrierung generierte Vogelperspektive. Man erkennt die Markierungen, die für die Kalibrierung eingesetzt wurden.

### 9.2.2 Ergebnisse

Abb. 9.12 gibt einen groben Überblick über die Szene und die gefahrenen Teststrecken. Im folgenden werden drei Sequenzen (Seq1, Seq2, Seq3) und deren Auswertung exemplarisch dargestellt. Die extrinsischen Kalibrierparameter wurden mit Null und deren Kovarianzmatrizen mit sehr großen Werten initialisiert. Die Berechnung des SMA



**Abbildung 9.12:** Übersicht über die ausgewerteten Testfahrten auf dem Gelände des Fraunhofer IOSB in Karlsruhe. Alle Fahrten erfolgten nach dem Bewegungsschema 8.1.

erfolgte bezüglich eines allen Modulen gemeinsamen Bezugspunkts (man vergleiche hierzu Abschnitt 6.2).

In Abb. 9.13 und 9.14 sind die geschätzten und fusionierten Bewegungsparameter der drei Kameras sowie die geschätzten extrinsischen Kalibrierparameter der ersten Sequenz auf einer gemeinsamen Zeitachse dargestellt. Man erkennt deutlich, dass der Kalibrierprozess nach der zweiten Geradenstrecke startet. Im ersten Geradenabschnitt wird die Bewegungsschätzung initialisiert, im ersten rotatorischen Abschnitt wird die

Rotationsachse bestimmt, und im zweiten Geradenabschnitt werden die Bewegungs- transformationen mit Hilfe der Rotationsachse in eine Grundebene überführt und die Orientierung der Kameras initialisiert. Abb. 9.15 vergleicht die Kalibrierergebnisse mit und ohne Fusion<sup>31</sup>.

Die Ergebnisse der drei Sequenzen sind in den Abb. 9.14 bis 9.17 zu sehen<sup>32</sup>.

In allen Fällen konnten die Orientierungsparameter der Kameras zueinander gut geschätzt werden. Die translatorischen Kalibrierparameter konvergieren nicht – wie man es vielleicht erwartet hätte – gegen die Grundwahrheit. Vielmehr bewegen sich die Parameter zwar im Bereich der Grundwahrheit, stabilisieren sich dort aber nicht. Der Grund hierfür liegt maßgeblich in den schwankenden Ergebnissen des SMA, die als wesentliche Messung in die Zustandsfilterung einfließen (siehe Abschnitt 6.3). Eine Verringerung des Systemrauschens würde die Schwankungen der Parameterkurven auf Kosten der Anpassungsfähigkeit reduzieren.

Abb. 9.18 bis 9.20 zeigt die in Abschnitt 6.3 vorgestellten Korrelationsparameter der Kovarianzmatrizen der SMA-Parameter. Diese dienen als Indikator für die Korrelation der extrinsischen Translatorischen Parameter  $\underline{t}_x$ ,  $\underline{t}_z$  mit den Skalaren der Bewegungs- translationen  $\kappa$  und sind ein Indiz für die Qualität der SMA-Schätzung<sup>33</sup>. Vergleicht man die Abb. 9.18 bis 9.20 mit den Abb. 9.14, 9.16 bzw. 9.17, so erkennt man in allen Sequenzen, in denen sich die translatorischen Parameter von der eingezeichneten Grundwahrheit entfernen, auch einen leichten Anstieg der entsprechenden Korrelationskurve. Zusätzlich sind die mittleren Projektionsfehler in aus Abschnitt 7.2 bzw.

---

31 Da ohne Gauß'sche Fusion jedes Kameramodul eigene, unabhängige Schätzwerte berechnet, wurden für die Erstellung des Diagramms am Ende des Kalibrierschritts die Mittelwerte berechnet. Eine Bewegungsfusion fand nicht statt.

32 Bei den Diagrammen 9.15, 9.18 bis 9.20 sowie 9.14 bis 9.17 sind nur die Zeitpunkte dargestellt, in denen auch tatsächlich eine Schätzung des Zustands der Kalibrierparameter durchgeführt wurde. Man beachte daher, dass die Zeitachse der Diagramme nicht mehr stetig fortlaufend ist.

33 Da die Berechnung des Bewegungsausgleichs im vorliegenden Fall zentrumsbasiert erfolgte, existieren für jedes Modul drei SMA-Schätzungen, nämlich die Transformationen zwischen Bezugspunkt und den jeweiligen drei Kameramodulen, sowie deren jeweiligen Kovarianzmatrizen. Somit lassen sich auch drei Korrelationsparameter  $\hat{\sigma}_{tx}$  gemäß Abschnitt 6.3 bestimmen. Die Parameterkurve in den Abb. 9.18 bis 9.20 zeigt das Maximum aus den drei Teilkurven. Dies ist u. a. damit zu begründen, dass bereits eine schlechte Teilschätzung großen Einfluss auf die Richtigkeit der zusammengeführten Parameter haben kann.

Gl. 7.15 und Gl. 7.16 auf einer Zeitachse aufgeführt. Der mittlere Projektionsfehler  $\varepsilon_C$  in Abhängigkeit der extrinsischen Parameter bzw. Transformation  $C_{ij}$  berechnet sich hierbei gemäß

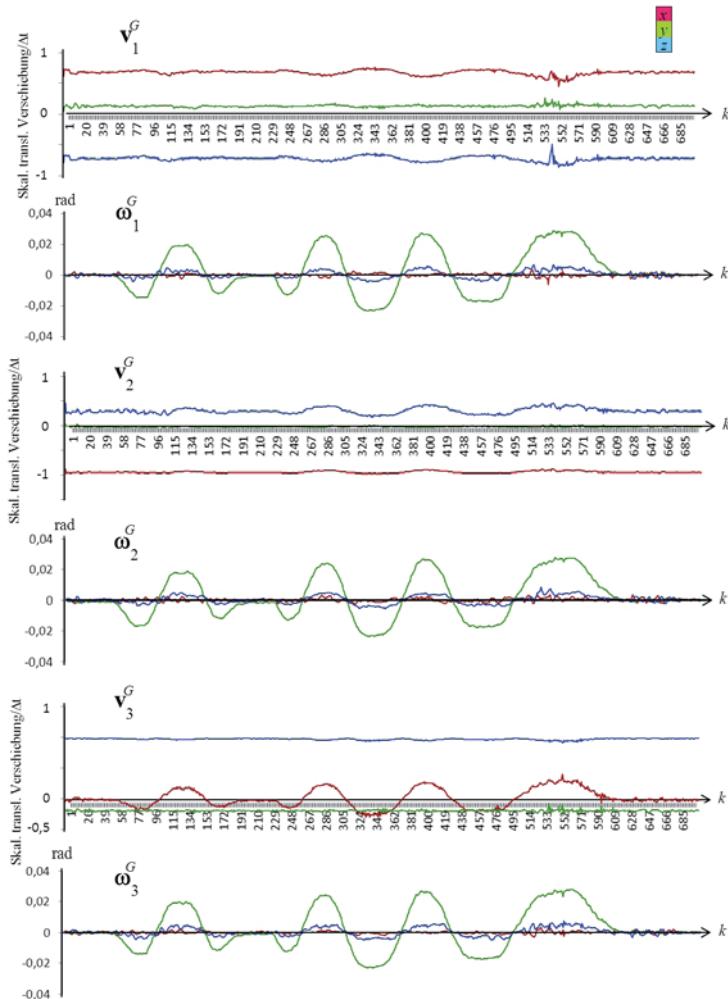
$$\varepsilon_C = \sum_{i \neq j}^{\mathfrak{N}} \frac{1}{\mathfrak{P}_i} \sum_{p=1}^{\mathfrak{P}_i} \left\| \pi(\mathbf{X}_{ip}; \mathbf{K}_i, C_{ij}) - \mathbf{x}_{ip} \right\|^2. \quad (9.1)$$

### Anwendungsbeispiel: Vogelperspektive

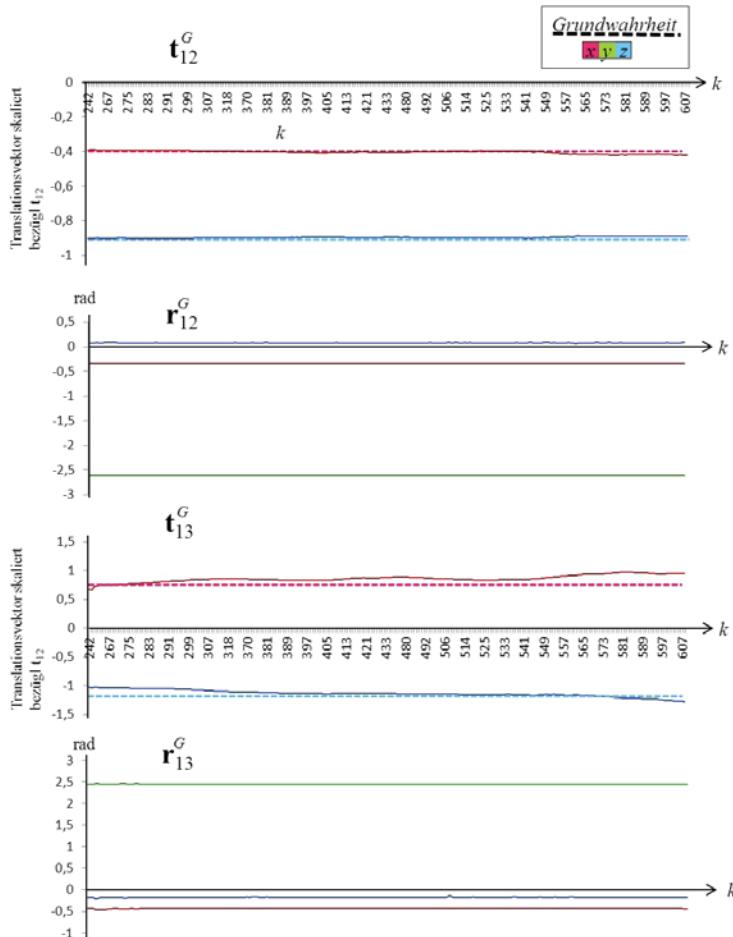
Eine Anwendung im Zusammenhang mit der extrinsischen Kalibrierung ist die räumliche Darstellung der Kamerabilder aus der Vogelperspektive. Bei dieser Anwendung werden die Kamerabilder unter der Annahme einer flachen Welt mit Hilfe einer Homographie, in Abhängigkeit der geometrischen Anordnung der Kameras, auf die Bodenebene projiziert.

Aus der Referenzkalibrierung sind die Transformationen in die Bodenebene bereits bekannt. In der folgenden Implementierung wurde lediglich die Transformation in die Bodenebene des Referenzmoduls fixiert (in den Abb. 9.21 bis 9.23 ist dies die Kamera links unten). Die Homographien der anderen Module wurden ausschließlich auf Basis der geschätzten extrinsischen Parameter (und als bekannt angenommenen intrinsischen Parametern) berechnet (siehe Anhang A.6).

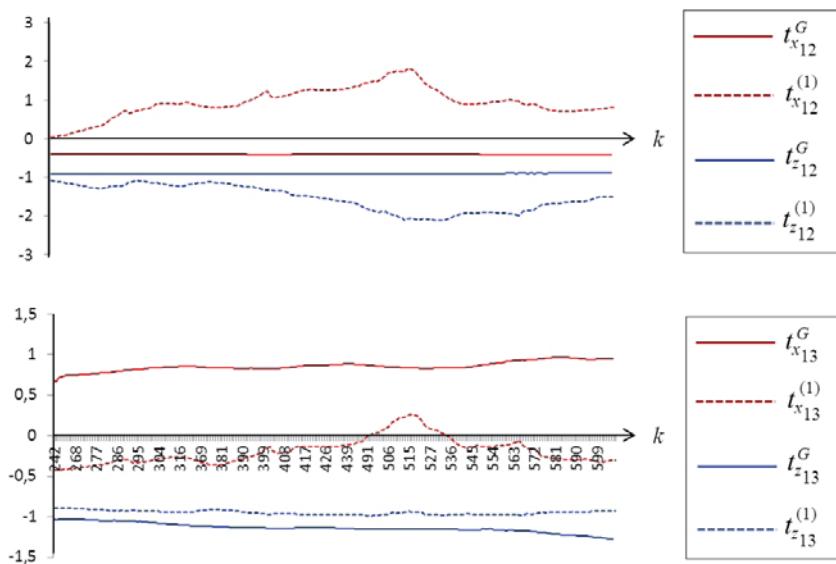
Abb. 9.21 bis 9.23 zeigen Ausschnitte der Sequenzen aus Abb. 9.12. Das erste Bild zeigt jeweils die Ausgangslage (da die extrinsischen Parameter jeweils mit Null initialisiert waren, unterscheiden sich die Transformationshomographien der drei Kameramodule jeweils lediglich durch die intrinsischen Parameter). Man erkennt, dass die prinzipielle Anordnung der Kameras durchaus getroffen wird. Bei näherer Betrachtung jedoch lassen sich Abweichungen erkennen, insbesondere bei bildübergreifenden Geradensegmenten.



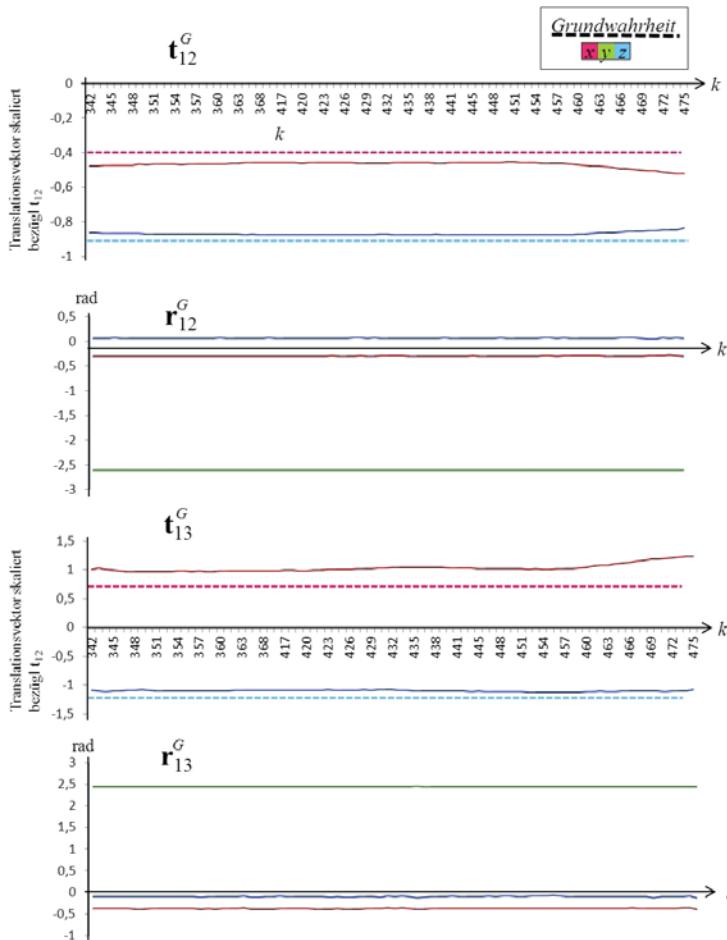
**Abbildung 9.13:** Sequenz 1: Ergebnis der Bewegungsschätzung der drei Module  $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3$ . Die Translationsvektoren sind skaliert, sodass  $\|\mathbf{v}^G\| = 1$ . Man erkennt die Ähnlichkeit der Rotationsparameter zwischen den Modulen, die sehr deutlich den Kurvenverlauf der Messfahrt widerspiegeln.



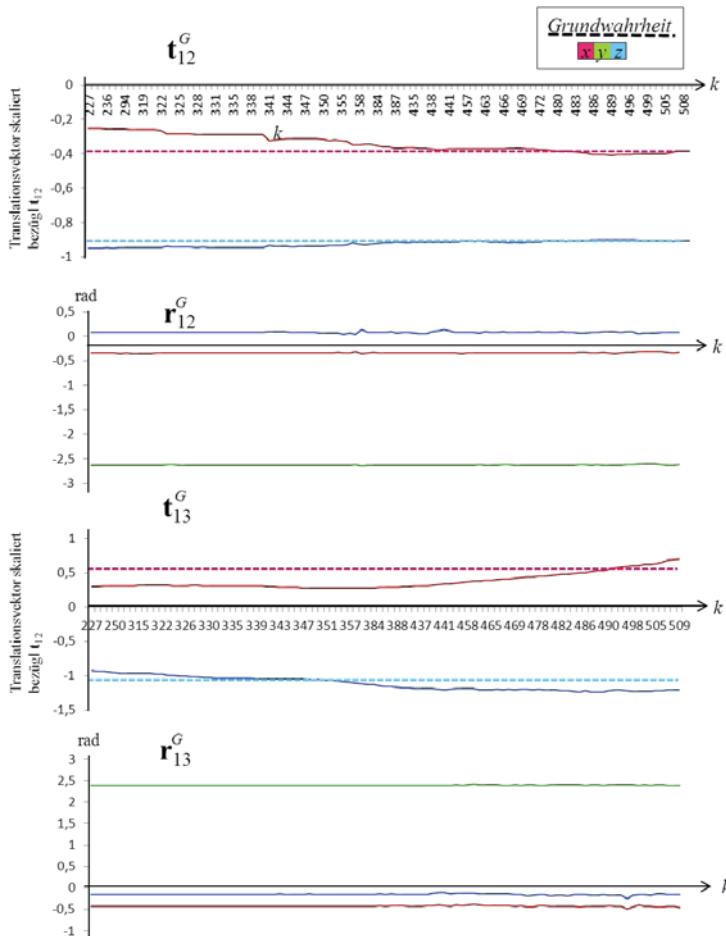
**Abbildung 9.14:** Sequenz 1: Ergebnis der Schätzung der extrinsischen Parameter zwischen  $\mathcal{M}_1, \mathcal{M}_2$  und  $\mathcal{M}_1, \mathcal{M}_3$ . Der Translationsvektor  $t_{13}^G$  wurde relativ zu  $t_{12}^G$  geschätzt mit  $\|t_{12}^G\| = 1$ . Der longitudinale Translationsparameter  $t_y$  ist nicht aufgeführt (siehe Abschnitt 5.3.4). Die Orientierungswinkel  $r_{12}^G$  und  $r_{13}^G$  konnten sehr gut geschätzt werden, ebenso  $t_{12}^G, t_{13}^G$  weist leichte Abweichungen von den Referenzparametern auf.



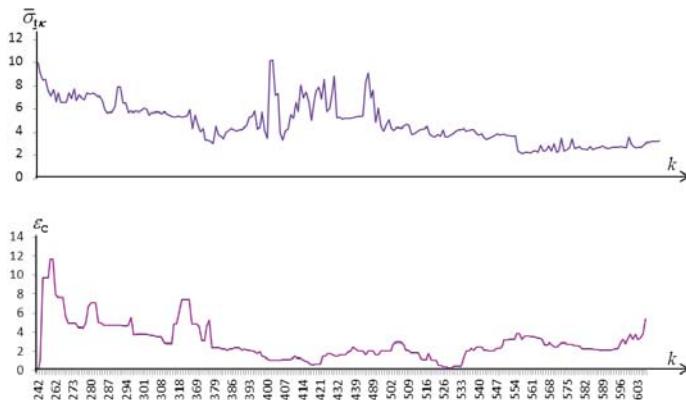
**Abbildung 9.15:** Vergleich der Schätzergebnisse der Kalibrierparameter  $t_{12}^G$  und  $t_{13}^G$  von Sequenz 1 mit (dünne Linie) und ohne Fusion (schraffierte Linie). Ein Vergleich mit Abb. 9.14 zeigt, dass das Fusionsergebnis wesentlich näher an der Grundwahrheit liegt, als das unfusionierte Ergebnis. Darüber hinaus weist das unfusionierte Ergebnis starke Schwankungen auf.



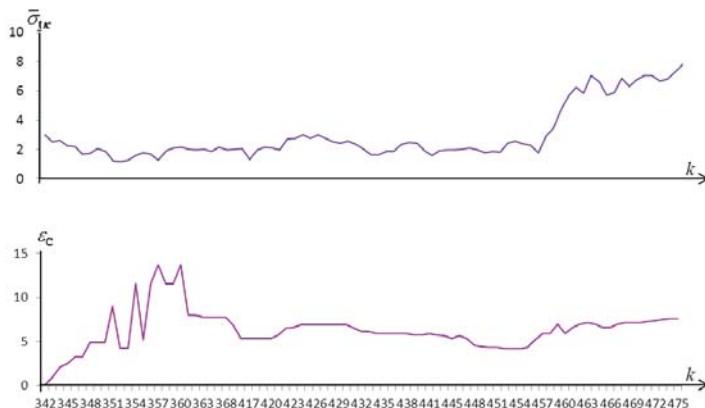
**Abbildung 9.16:** Sequenz 2: Geschätzte extrinsische Kalibrierparameter. Der longitudinalen Translationsparameter  $t_y$  ist nicht aufgeführt (siehe Abschnitt 5.3.4). Die Orientierungswinkel  $r_{G12}$  und  $r_{13}^G$  konnten sehr gut geschätzt werden. Sowohl die Schätzwerte für  $t_{12}^G$  als auch  $t_{13}^G$  entfernen sich zum Ende der Messsequenz von den Referenzwerten. Dieser Sachverhalt spiegelt sich in Abb. 9.19 wider.



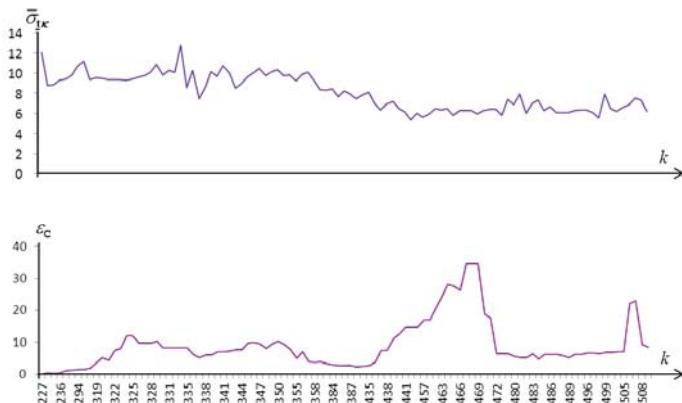
**Abbildung 9.17:** Sequenz 3: Geschätzte extrinsische Kalibrierparameter. Der longitudinalen Translationsparameter  $t_y$  ist nicht aufgeführt (siehe Abschnitt 5.3.4). Die Orientierungswinkel  $\mathbf{r}_{G12}$  und  $\mathbf{r}_{13}^G$  konnten sehr gut geschätzt werden. Sowohl  $\mathbf{t}_{12}^G$  als auch  $\mathbf{t}_{13}^G$  nähern sich ab etwa  $k = 350$  den Referenzwerten an. Abb. 9.20 (oben) gibt der hier vorliegenden Beobachtung Ausdruck: Ein Fallen der Kurve indiziert eine sinkende Unsicherheit.



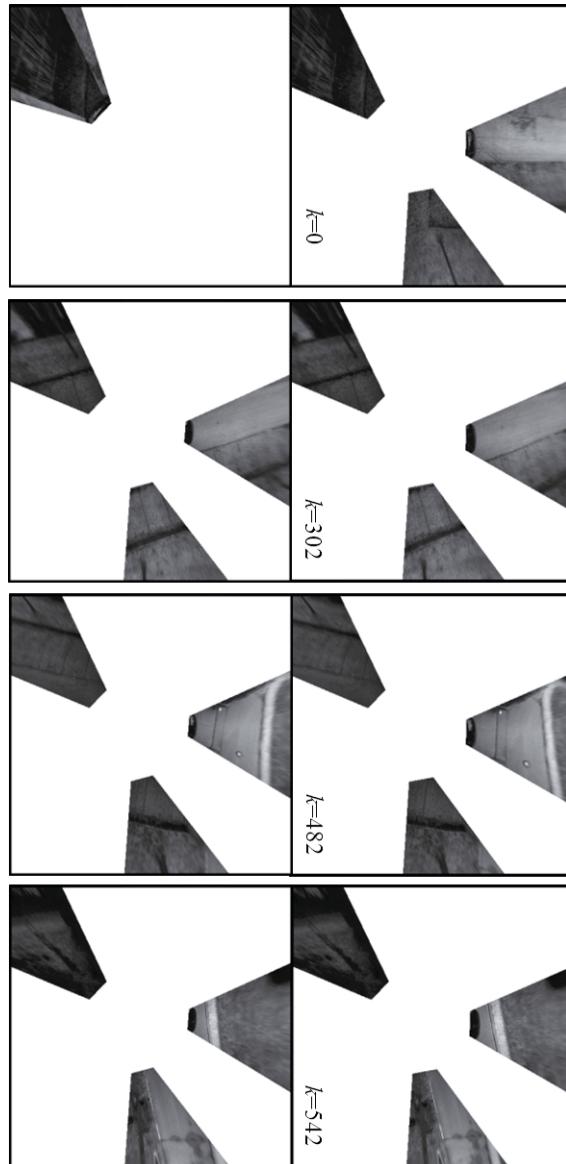
**Abbildung 9.18:** Diagramme der Korrelationskenngrößen  $\bar{\sigma}_{tk}$  sowie der Projektionsrestfehler  $\varepsilon_C$  für Sequenz 1. Die Kenngrößen geben u.a. Aufschluss über die momentane Güte der Messung (vergleiche mit Abb. 9.14). Beide Kurven deuten auf eine gute Schätzung hin, mit leichten Unsicherheiten in der Mitte der Sequenz.



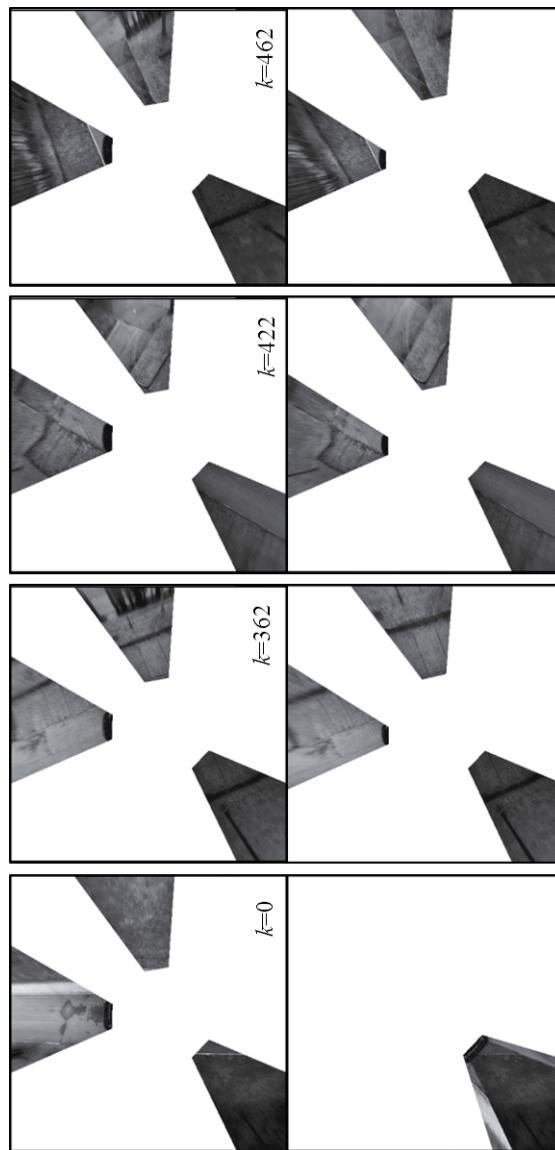
**Abbildung 9.19:** Diagramme der Korrelationskenngrößen  $\bar{\sigma}_{tk}$  sowie der Projektionsrestfehler  $\varepsilon_C$  für Sequenz 2. Die Kenngrößen geben u.a. Aufschluss über die momentane Güte der Messung (vergleiche mit Abb. 9.16). Die Güte des Schätzergebnisses verschlechtert sich am Ende der Sequenz ab etwa  $k = 455$ .



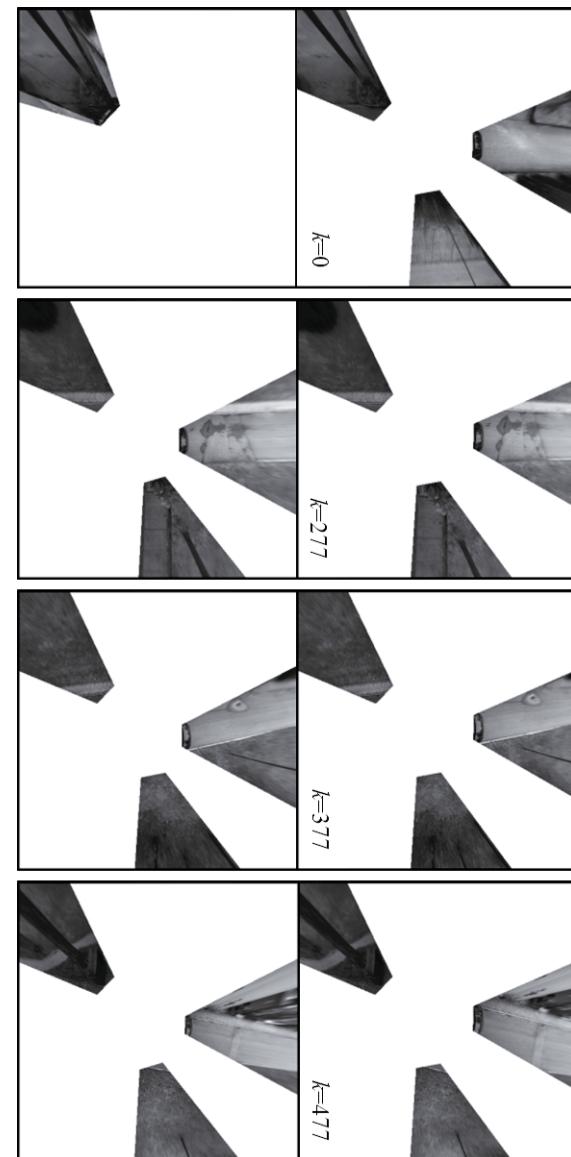
**Abbildung 9.20:** Diagramme der Korrelationskenngrößen  $\bar{\sigma}_{tk}$  sowie der Projektionsrestfehler  $\varepsilon_C$  für Sequenz 3. Die Kenngrößen geben u.a. Aufschluss über die momentane Güte der Messung (vergleiche mit Abb. 9.17). Die Unsicherheit des Schätzergebnisses nimmt ab etwa der Mitte der Messsequenz ab. Der Projektionsfehler weist temporäre Schwankungen auf, die im Laufe der Sequenz wieder abgefangen werden. Ein Vergleich mit Abb. 9.17 zeigt eine Verbesserung des Schätzergebnisses ab etwa  $k = 350$ .



**Abbildung 9.21:** Für Sequenz 1 generierte Vogelperspektive mit den Parametern der Referenzkalibrierung (oben) und den kontinuierlich, bewegungsbasiert geschätzten extrinsischen Parametern (unten). Im Idealfall soll die untere Draufsicht der oberen entsprechen. Der Vergleich der Positionen und Orientierungen Sichtkegel zwischen der Referenz- und der geschätzten Draufsicht erlaubt eine qualitative Bewertung der Kalibrierparameter.



**Abbildung 9.22:** Für Sequenz 2 generierte Vogelperspektive mit den Parametern der Referenzkalibrierung (oben) und den kontinuierlich, bewegungsbaasiert geschätzten extrinsischen Parametern (unten). Im Idealfall soll die untere Draufsicht der oberen entsprechen. Der Vergleich der Positionen und Orientierungen Sichtkegel zwischen der Referenz- und der geschätzten Draufsicht erlaubt eine qualitative Bewertung der Kalibrierparameter.



**Abbildung 9.23:** Für Sequenz 3 generierte Vogelperspektive mit den Parametern der Referenzkalibrierung (oben) und den kontinuierlich, bewegungsbasiert geschätzten extrinsischen Parametern (unten). Im Idealfall soll die untere Draufsicht der oberen entsprechen. Der Vergleich der Positionen und Orientierungen Sichtkegel zwischen der Referenz- und der geschätzten Draufsicht erlaubt eine qualitative Bewertung der Kalibrierparameter.

# 10

---

## Abschließende Bemerkungen

---

### 10.1 Zusammenfassung

Die vorliegende Arbeit präsentiert ein Verfahren zur bewegungsbasierten extrinsischen Selbstkalibrierung von Kameras mit nichtüberlappenden Sichten in Fahrzeugen. Hierfür wird insbesondere auf eine Methode aus der Robotik, die sog. Hand-Augen-Kalibrierung, zurückgegriffen.

Die Bewegungsschätzung findet im vorliegenden Fall kamerabasiert und monokular statt (Kapitel 3). Daher ist es wichtig, (selbst unter Einsatz von Odometriesensoren) die relativen Geschwindigkeitsbeträge zusätzlich zu den extrinsischen Parametern zu ermitteln. Hierfür wurde in Anlehnung an den wohlbekannten Bündelausgleich der Bewegungsausgleich vorgestellt.

Da für planare Bewegungen, wie sie vorwiegend von Straßenfahrzeugen durchgeführt werden, die rotatorische Bewegung i. A. nicht ausreicht, um alle Freiheitsgrade mit Hilfe der Hand-Augen-Kalibrierung bzw. des Bewegungsausgleichs zu bestimmen, wurde ein erweitertes Modell vorgestellt, das explizit die Existenz einer allen Kameras gemeinsamen Bodenebene explizit berücksichtigt. In diesem Zusammenhang wurde auch die Bedeutung zusätzlicher sensorischer Information zur Messung der Fahrzeuggeschwin-

digkeit herausgearbeitet, ohne die die longitudinalen Abstände zwischen den Kameras nicht schätzbar sind (Kapitel 5).

Der Bewegungsausgleich modelliert und schätzt die relativen Geschwindigkeiten der Kameramodule explizit. In Experimenten konnte dargelegt werden, dass die Abhängigkeit des Schätzergebnisses von der Lage der Kameras sowie deren Bewegung abhängt. Durch analytische Betrachtung der Kovarianzmatrizen der Zustandsparameter ließen sich Hinweise auf die Güte der Schätzung extrahieren (Kapitel 6).

Das Kalibriermodell wurde in einen ganzheitlichen, skalierbaren Ansatz eingebettet (Kapitel 8). Dieses erfüllt zwei eingangs gestellte Forderungen.

Zum einen ist die Kalibrierung skalierbar hinsichtlich der Anzahl der Kameras. D. h. jede Kamera kann als eigenständige Recheneinheit modelliert werden, was eine Parallelisierung der Rechenprozesse ermöglicht. Die Kommunikation zwischen den Kameramodulen und der Datenaustausch erfolgt auf Zustandsebene. D. h. im Gegensatz zu einer Fusion auf Merkmals- oder gar Rohdatenebene wird die Kommunikationsbandbreite dadurch gering gehalten.

Zum anderen wird die dem Schätzmodell innwohnende Redundanz durch ein Propagierungs- und Fusionsschema effektiv genutzt. Anhand von Simulationen konnte der Nutzen einer solchen Fusion gezeigt werden. Die Fusion beruht auf der Vereinigung normalverteilter Zustandsvariablen mit bekannter Kovarianzmatrix.

Das gesamte Kalibriersystem wurde mit simulierten und mit Realdaten evaluiert (Kapitel 9). Die Akquirierung der Realdaten erfolgte mit einem PKW, der mit drei Kameras ausgestattet war. Um die Auswertung der Realdaten vergleichbar zu machen, wurden die Kameras im Fahrzeug im Vorfeld mit Hilfe eines Musters kalibriert. Diese Kalibrierung konnte als Referenz zur Evaluierung der Ergebnisse herangezogen werden. In allen Versuchen wurden die rotatorischen Parameter, also die Orientierung der Kameras zueinander, zuverlässig geschätzt. Die Schätzung der relativen, translatorischen Lageparameter konnte nicht mit der gleichen Genauigkeit erreicht werden. Anstatt einer wünschenswerten Konvergenz der Zustandsparameter schwankten die Parameter

mehr oder minder stark um die wahren Werte. Dies lässt sich auf die Erkenntnisse aus der Analyse des Bewegungsausgleichs zurückführen. Eine Beispielapplikation, die Generierung einer Vogelperspektive, demonstriert die Auswirkung des in dieser Arbeit vorgestellten Kalibrierverfahrens im praktischen Einsatz.

## 10.2 Ausblick

Der Fokus der vorliegenden Arbeit liegt ausschließlich auf der Schätzung der extrinsischen Kameraparameter. Für praktische Anwendungen ist ebenfalls die Schätzung der intrinsischen Parameter, insbesondere der Brennweite, von Interesse. Daher ist es lohnenswert Überlegungen anzustellen, wie Ansätze zur Selbstkalibrierung in die kontinuierliche, extrinsische Kalibrierung integriert werden können.

Da die Lage der Kameras und die Bewegung der Plattform starke Auswirkung auf das translatorische Kalibrierergebnis haben (Kapitel 6), ist der Bewegungsausgleich bedingt für die Kalibrierung solcher Mehrkamerasysteme geeignet. Eine Weiterentwicklung muss daher darauf abzielen, kritische Schätzergebnisse bzw. Sensor-Bewegungs-Kombinationen unmittelbar zu erkennen und diese nicht in den weiteren Schätzprozess einfließen zu lassen. Die Schätzung der rotatorischen Kalibrierparameter ist davon unbeeinflusst. Diese lässt sich sehr wohl und zuverlässig mit den vorgestellten Methoden berechnen.

Es wäre daher denkbar, für eine bestimmte Kamerakonstellation *a priori* die optimale Bewegung zu definieren, die ein gutes Kalibrierergebnis garantiert. Die Plattform könnte dann ggf. sogar autonom den »Kalibrierparcours« abfahren. Eine solche Einschränkung ist jedoch nicht unbedingt wünschenswert. Vielmehr sollte eine solche Bewegung im laufenden Prozess erkannt werden, sobald sie durchgeführt wurde, um damit die Kameras zu kalibrieren.

Jüngere Optimierungsverfahren wie das sog. *Second-Order Cone Programming* [Har09] erscheinen vielversprechend und müssen hinsichtlich ihrer Leistungsstärke gegenüber den hier eingesetzten LM-Verfahren untersucht werden.

Alternative algorithmische Ansätze, wie SLAM bzw. ein multiokularer Bündelausgleich, gehen davon aus, dass dieselben Objekte von verschiedenen Kameras zu verschiedenen Zeitpunkten sichtbar sind. Derartige Lösungen existieren bisher lediglich in eingeschränkten Umgebung bzw. in geschlossenen Räumen mit vordefinierten Bewegungen der Plattform. Dadurch wird erreicht, dass jede Kamera annähernd dieselbe Szene wie die anderen Kameras vor Augen hatte. Die Schwierigkeit für den Einsatz in allgemeinen Verkehrsszenarien liegt in der unterschiedlichen perspektivischen Ausprägungen der Objekte im Bild. Hier müssen robuste und perspektivisch invariante Landmarken extrahiert und in den anderen Kamerabildern zeitversetzt zugeordnet werden.

Eine andere Möglichkeit bietet die direkte Fusionierung von 3D-Punktwolken, beispielsweise mit SLAM- oder *Iterative Closest Point*-Verfahren. Hierbei werden die von den einzelnen Sensoren generierten Punktwolken unter Minimierung eines Fehlerfunktionalen in ein gemeinsames Koordinatensystem eingefügt. Jedoch leiden monokulare 3D-Rekonstruktionen meist unter einer schlechten Tiefenauflösung, verglichen mit Stereo- oder Lidarsensoren.

Aufgrund der Tatsache, dass die meisten sensoriell ausgestatteten Fahrzeuge nicht nur mit Kameras, sondern vielmehr mit weiteren (heterogenen) Sensoren bestückt sind, wird die Fusion mit anderen sensorischen Informationsquellen ein wichtiger Schritt für derartige Multi-Sensor-Kalibrieraufgaben sein. Hierfür kommen sowohl einfache Intertialsensoren zur Beschleunigungsmessung in Frage als auch komplexe Sensoren zur direkten Erfassung von 3D-Informationen wie Stereokameras oder Lidarsensoren. Teilschätzungen, wie z. B. die Bewegungsschätzung, könnten dadurch robuster und genauer bestimmt werden. Allerdings kann der zusätzliche Einsatz von Sensorik zu einem zusätzlichen Kalibrieraufwand führen, falls die heterogenen, sensorischen Informationsquellen zueinander referenziert werden müssen. In einem solchen Fall gilt es den Aufwand gegen den Nutzen abzuwegen.

Wenngleich das hier vorgestellte Verfahren prinzipiell darauf ausgerichtet ist, in Echtzeit, d. h. mit der Bildeinzugsrate zu rechnen, so erreichte die im Zuge dieser Arbeit

angefertigte Implementierung diese Anforderungen nicht. Vielmehr handelt es sich um ein prototypisches Experimentalsystem. Bei den durchgeführten Versuchen auf einem 2,20 GHz Dual Core Prozessor lag die Performanz je Berechnugszyklus je Kameramodul bei 0,6-1 Hz (die Schwankungen sind bedingt durch die Anzahl der extrahierten und prozessierten Bildmerkmale). Die Flaschenhälse sind hierbei die Merkmalsextraktion, die Korrespondenzfindung sowie die Berechnung des Bündel- und Bewegungsausgleichs. Die Echtzeitfähigkeit eines schritthalrenden Kalibriersystems ist jedoch essentiell für den praktischen Nutzen. Hier gilt es in zukünftigen Arbeiten, die Kapazitäten moderner Rechnerhardware und softwaretechnischer Optimierung auszuloten, wie beispielsweise den Einsatz von Grafikkarten zur beschleunigten, parallelisierten Berechnung des Optischen Flusses oder des Bündelausgleichs.



# A

---

## Anhang

---

### A.1 Darstellung von Rotationen im dreidimensionalen Raum

**Eulerwinkel** Die drei Eulerwinkel  $r_x, r_y, r_z$  beschreiben die Drehung eines Koordinatensystems, wobei jeder der drei Winkel das Koordinatensystem jeweils um die  $X-, Y-$  bzw.  $Z$ -Achse dreht.

**Rotationsmatrix**  $\mathbf{R}$  bzw.  $\mathbf{R}(r_x, r_y, r_z)$  ist die  $3 \times 3$ -Rotationsmatrix zu den entsprechenden Eulerwinkeln, mit  $\mathbf{R}^T \mathbf{R} = \mathbf{R} \mathbf{R}^T = \mathbf{I}$  und  $\det(\mathbf{R}) = 1$ . Die Rotation eines Punktes innerhalb eines Koordinatensystems  $\mathbf{X} \in \mathbb{R}^3$  wird durch Multiplikation mit der Rotationsmatrix erreicht, also

$$\mathbf{X}' = \mathbf{R} \mathbf{X} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}. \quad (\text{A.1})$$

Die Rotationsachse ist der Eigenvektor von  $\mathbf{R}$  zum Eigenwert 1.

**Quaternionen** Eine Quaternion  $q_w + i_q \cdot q_x + j_q \cdot q_y + k_q \cdot q_z$  mit  $i_q^2 = j_q^2 = k_q^2 = i_q^2 j_q^2 k_q^2 = -1$  wird in der vorliegenden Arbeit anhand ihrer Koeffizienten  $\mathbf{q} = [q_w, q_x, q_y, q_z]$

dargestellt [Koe83]. Ähnlich zu den komplexen Zahlen bildet hierbei  $q_w$  den Realteil und  $[q_x, q_y, q_z]$  den Imaginärteil von  $\mathbf{q}$ .

Die Multiplikation zweier Quaternionen ergibt sich wie folgt:

$$\mathbf{q}_1 \star \mathbf{q}_2 = [q_{w3} \ q_{x3} \ q_{y3} \ q_{z3}]$$

mit

$$\begin{aligned} q_{w3} &= q_{w1}q_{w2} - q_{x1}q_{x2} - q_{y1}q_{y2} - q_{z1}q_{z2} \\ q_{x3} &= q_{w1}q_{x2} + q_{x1}q_{w2} + q_{y1}q_{z2} - q_{z1}q_{y2} \\ q_{y3} &= q_{w1}q_{y2} - q_{x1}q_{z2} + q_{y1}q_{w2} + q_{z1}q_{x2} \\ q_{z3} &= q_{w1}q_{z2} + q_{x1}q_{y2} - q_{y1}q_{x2} + q_{z1}q_{w2} \end{aligned} \quad (\text{A.2})$$

Gl. A.2 lässt sich auch als Matrixmultiplikation formulieren gemäß

$$\mathbf{q}_1 \star \mathbf{q}_2 = \mathbf{Q}_{\mathbf{q}_2}^{\prec} \cdot \mathbf{q}_1 = \mathbf{Q}_{\mathbf{q}_1}^{\succ} \cdot \mathbf{q}_2 \quad (\text{A.3})$$

mit

$$\mathbf{Q}_{\mathbf{q}_2}^{\prec} = \begin{bmatrix} q_{w2} & -q_{x2} & -q_{y2} & -q_{z2} \\ q_{x2} & q_{w2} & q_{z2} & -q_{y2} \\ q_{y2} & -q_{z2} & q_{w2} & q_{x2} \\ q_{z2} & q_{y2} & -q_{x2} & q_{w2} \end{bmatrix} \quad (\text{A.4})$$

und

$$\mathbf{Q}_{\mathbf{q}_1}^{\succ} = \begin{bmatrix} q_{w1} & -q_{x1} & -q_{y1} & -q_{z1} \\ q_{x1} & q_{w1} & -q_{z1} & q_{y1} \\ q_{y1} & q_{z1} & q_{w1} & -q_{x1} \\ q_{z1} & -q_{y1} & q_{x1} & q_{w1} \end{bmatrix} \quad (\text{A.5})$$

Einheitsquaternionen mit  $\|\mathbf{q}\| = \sqrt{q_w^2 + q_x^2 + q_y^2 + q_z^2} = 1$  können zur Darstellung von Rotationen verwendet werden. Ein Drehung äquivalent zu Gl. A.1 erhält man gemäß

$$\mathbf{q} \star \mathbf{X}^q \star \bar{\mathbf{q}} \quad (\text{A.6})$$

wobei  $\bar{\mathbf{q}}$  die zu  $\mathbf{q}$  konjugierte Quaternion  $[q_w, -q_x, -q_y, q_z]$  bezeichnet.  $\mathbf{X}^q$  kann hierbei als rein imaginäre Quaternion  $[0, X, Y, Z]$  betrachtet werden.

**Konvertierungen** Eine Rotationsmatrix  $\mathbf{R}$  erhält man aus drei Eulerwinkeln  $r_x, r_y, r_z$  gemäß Gl. 2.4 zu

$$\mathbf{R}(r_x, r_y, r_z) = \begin{bmatrix} \cos(r_y)\cos(r_z) & \sin(r_x)\sin(r_y)\cos(r_z) - \cos(r_x)\sin(r_z) & \sin(r_x)\sin(r_z) + \cos(r_x)\sin(r_y)\cos(r_z) \\ \cos(r_y)\sin(r_z) & \sin(r_x)\sin(r_y)\sin(r_z) + \cos(r_x)\cos(r_z) & \cos(r_x)\sin(r_y)\sin(r_z) - \sin(r_x)\cos(r_z) \\ -\sin(r_y) & \sin(r_x)\cos(r_y) & \cos(r_x)\cos(r_y) \end{bmatrix}. \quad (\text{A.7})$$

Der umgekehrte Weg gestaltet sich ungleich schwieriger, da die Abbildung einer Rotationsmatrix auf drei Eulerwinkel nicht eindeutig ist. Beispielsweise resultieren für den Winkel  $r_y$  aus Gl. A.7 auf dem Intervall  $[-\pi, \pi]$  zwei Lösungen, nämlich

$$\begin{aligned} r_{y1} &= \arcsin(-r_{33}) \quad \text{und} \\ r_{y2} &= \begin{cases} \pi - r_{y1} & , \quad \text{falls } r_{y1} \geq 0 \\ -\pi - r_{y1} & , \quad \text{sonst.} \end{cases} \end{aligned} \quad (\text{A.8})$$

Basierend auf Gl. A.8 lassen sich aus den Matrixelementen  $r_{21}$  und  $r_{32}$  aus Gl. A.1 Werte für  $r_x$  und  $r_z$  berechnen, sodass sich letztlich acht mögliche Kombinationen für die Eulerwinkel  $\mathbf{r} = [r_x, r_y, r_z]^T$  ergeben. Um zu gewährleisten, dass die Eulerwinkel wieder auf die ursprüngliche Rotationsmatrix abgebildet werden, wird für jede der acht Winkelkombinationen die Rotationsmatrix gemäß Gl. A.7 berechnet. Von den Kombinationen  $\mathbf{r}_{i=1\dots 8}$ , für deren Differenz gilt  $\|\text{vec}(\mathbf{R}) - \text{vec}(\mathbf{R}(r_{xi}, r_{yi}, r_{zi}))\| < \varepsilon$  mit  $\text{vec}(\mathbf{R}) = [r_{11}, r_{21}, \dots, r_{23}, r_{33}]_{9 \times 1}^T$ , wird diejenige mit dem kleinsten Betrag  $\min_i \|\mathbf{r}_i\|$  gewählt. Hierbei stapelt der vec-Operator die Spalten der Matrix übereinander ( $\text{vec} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \cdot n \times 1}$ ).

Die Konvertierung zwischen Eulerwinkeln und Quaternionen kann u. a. bei Horaud und Dornaika [Hor95] nachgelesen werden. Eine Quaternion lässt sich aus Eulerwinkeln

wie folgt beschreiben:

$$\begin{aligned} q_w &= \cos(r_x/2)\cos(r_y/2)\cos(r_z/2) + \sin(r_x/2)\sin(r_y/2)\sin(r_z/2) \\ q_x &= \sin(r_x/2)\cos(r_y/2)\cos(r_z/2) - \cos(r_x/2)\sin(r_y/2)\sin(r_z/2) \\ q_y &= \cos(r_x/2)\sin(r_y/2)\cos(r_z/2) + \sin(r_x/2)\cos(r_y/2)\sin(r_z/2) \\ q_z &= \cos(r_x/2)\cos(r_y/2)\sin(r_z/2) - \sin(r_x/2)\sin(r_y/2)\cos(r_z/2) \end{aligned} \quad (\text{A.9})$$

Aus den Gln. A.3 und A.6 lässt sich eine Rotationsmatrix ableiten. Sei

$$\begin{aligned} \mathbf{X}' &= \mathbf{q} \star \mathbf{X}^q \star \bar{\mathbf{q}} \\ &= [\mathbf{Q}_{\mathbf{q}}^{\succ} \mathbf{X}^q] \star \bar{\mathbf{q}} \\ &= \mathbf{Q}_{\bar{\mathbf{q}}}^{\prec} [\mathbf{Q}_{\mathbf{q}}^{\succ} \mathbf{X}^q] \\ &= [\mathbf{Q}_{\mathbf{q}}^{\prec T} \mathbf{Q}_{\mathbf{q}}^{\succ}] \mathbf{X}^q \end{aligned} \quad (\text{A.10})$$

dann ergibt sich die Rotationsmatrix  $\mathbf{R}$  aus dem Matrixprodukt

$$\begin{aligned} \mathbf{Q}_{\mathbf{q}}^{\prec T} \mathbf{Q}_{\mathbf{q}}^{\succ} &= \\ \left[ \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & q_w^2 + q_x^2 - q_y^2 - q_z^2 & 2[q_x q_y - q_w q_z] & 2[q_x q_z + q_w q_y] \\ 0 & 2[q_x q_y + q_w q_z] & q_w^2 - q_x^2 + q_y^2 - q_z^2 & 2[q_y q_z - q_w q_x] \\ 0 & 2[q_x q_z - q_w q_y] & 2[q_y q_z + q_w q_x] & q_w^2 - q_x^2 - q_y^2 + q_z^2 \\ 1 & \mathbf{0}^T & & \\ \mathbf{0} & \mathbf{R} & & \end{array} \right] &= \end{aligned} \quad (\text{A.11})$$

Quaternionen haben generelle Vorteile gegenüber dem Rechnen mit Eulerwinkeln. Bei der sog. kardianischen Blockade (engl. *Gimbal Lock*) kann durch eine ungünstige Reihenfolge der Drehungen um die drei Hauptachsen des Koordinatensystems mit den entsprechenden Eulerwinkeln eine Konstellation auftreten, in der bereits nach den ersten beiden Drehungen die Rotationsachse mit der dritten Hauptachse zusammenfallen [Gra98]. Dies hat zur Folge, dass die dritte Drehung wirkungslos bleibt und der Rotation insgesamt ein Freiheitsgrad fehlt. Im Gegenzug muss beim Rechnen mit

Einheitsquaternionen deren Normierung ggf. als Nebenbedingung mitmodellieren.

## A.2 Beweis der Kommutativität der paarweisen Gauß'schen Fusion

**Satz 1** Die Gauß'sche Fusion zweier normalverteilter Zufallsvariablen  $\mathbf{x}_1 \cup \mathbf{x}_2$  mit  $\mathbf{x}_1 \sim \mathcal{N}(\boldsymbol{\mu}_1, \Sigma_1)$  und  $\mathbf{x}_2 \sim \mathcal{N}(\boldsymbol{\mu}_2, \Sigma_2)$  gemäß Gln. 2.53 und 2.52 ist kommutativ.

BEWEIS Entsprechend der Gln. 2.53 und 2.52 muss für die Fusion der Mittelwerte gelten

$$\boldsymbol{\mu}_1 + \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} [\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1] \stackrel{!}{=} \boldsymbol{\mu}_2 + \Sigma_2 [\Sigma_2 + \Sigma_1]^{-1} [\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] \quad (\text{A.12})$$

woraus durch einfaches Umformen folgt

$$[\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] - \Sigma_2 [\Sigma_1 + \Sigma_2]^{-1} [\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] - \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} [\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] = \mathbf{0} \quad (\text{A.13})$$

$$[\mathbf{I} - \Sigma_2 [\Sigma_1 + \Sigma_2]^{-1} - \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1}] [\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] = \mathbf{0} \quad (\text{A.14})$$

$$\underbrace{[\mathbf{I} - \underbrace{[\Sigma_1 + \Sigma_2][\Sigma_1 + \Sigma_2]^{-1}}_{= \mathbf{I}}]}_{= \mathbf{0}} [\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2] = \mathbf{0}. \quad (\text{A.15})$$

In gleicher Weise muss für die Kovarianzmatrizen gelten

$$\Sigma_1 - \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} \Sigma_1 \stackrel{!}{=} \Sigma_2 + \Sigma_2 [\Sigma_2 + \Sigma_1]^{-1} \Sigma_2. \quad (\text{A.16})$$

Die Multiplikation zweier Kovarianzmatrizen ist aufgrund ihrer Symmetrie kommutativ. Somit lässt sich Gl. A.16 umformen zu

$$[\Sigma_1 - \Sigma_2] = \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} \Sigma_1 - \Sigma_2 [\Sigma_1 + \Sigma_2]^{-1} \Sigma_2 \quad (\text{A.17})$$

$$= \Sigma_1 \Sigma_1 [\Sigma_1 + \Sigma_2]^{-1} - \Sigma_2 \Sigma_2 [\Sigma_1 + \Sigma_2]^{-1} \quad (\text{A.18})$$

$$[\Sigma_1 - \Sigma_2] [\Sigma_1 + \Sigma_2] = \Sigma_1 \Sigma_1 - \Sigma_2 \Sigma_2, \quad (\text{A.19})$$

wobei Gl. A.19 das dritte Binomische Gesetz erfüllt. ■

### A.3 Die Epipolareometrie

Die Fundamentalmatrix (oder kurz: F-Matrix) ist von elementarer Bedeutung in der geometrischen Bild- und Bildfolgengenauisierung. Sie beschreibt implizit sowohl die Kamera- als auch die Szenenstruktur [Luo95]. Im Folgenden wird die mathematische Herleitung der F-Matrix gezeigt.

Gegeben seien zwei Kameras und eine als starr angenommene Szene.  $\mathbf{X}_1$  sei ein Szenenpunkt im Koordinatensystem der ersten,  $\mathbf{X}_2$  im Koordinatensystem der zweiten Kamera. Die beiden Punkte lassen sich mittels der extrinsischen Parameter  $\mathbf{R}, \mathbf{t} = [t_x, t_y, t_z]^T$  zueinander in Relation setzen

$$\mathbf{X}_2 = \mathbf{R}\mathbf{X}_1 + \mathbf{t}. \quad (\text{A.20})$$

Bilden des Kreuzprodukts mit  $\mathbf{t}$  auf beiden Seiten ergibt

$$(\mathbf{X}_2 \times \mathbf{t}) = (\mathbf{R}\mathbf{X}_1 \times \mathbf{t}) + \underbrace{(\mathbf{t} \times \mathbf{t})}_{=0} \quad (\text{A.21})$$

und durch anschließendes Bilden des Skalarprodukts mit  $\mathbf{R}\mathbf{X}_1$  erhält man

$$(\mathbf{X}_2 \times \mathbf{t})^T \mathbf{R}\mathbf{X}_1 = (\mathbf{R}\mathbf{X}_1 \times \mathbf{t})^T \mathbf{R}\mathbf{X}_1 = \mathbf{t}^T \underbrace{(\mathbf{R}\mathbf{X}_1 \times \mathbf{R}\mathbf{X}_1)}_{=0} = 0. \quad (\text{A.22})$$

Der Vektor  $\mathbf{R}\mathbf{X}_1$  steht demnach senkrecht zu dem Normalenvektor der Ebene, die durch den Ursprung,  $\mathbf{X}_2$  und  $\mathbf{t}$  aufgespannt wird. Diese drei Vektoren sind koplanar, d. h. sie liegen einer gemeinsamen Ebene, in diesem Falle der sog. Epipolarebene. Gl. A.22 lässt

sich umformen zu

$$\mathbf{X}_2^T \begin{bmatrix} 0 & t_z & -t_y \\ -t_z & 0 & t_x \\ t_y & -t_x & 0 \end{bmatrix} \mathbf{R} \mathbf{X}_1 = \mathbf{X}_2^T [\mathbf{t}]_\times^T \mathbf{R} \mathbf{X}_1 = 0. \quad (\text{A.23})$$

Die Matrix

$$\mathbf{E} = [\mathbf{t}]_\times^T \mathbf{R} \quad (\text{A.24})$$

bezeichnet man als essentielle oder schlicht E-Matrix. Im unkalibrierten Fall erweitert sich Gl. A.23 unter Zuhilfenahme von Gl. 2.1 zu

$$[\lambda_2 \tilde{\mathbf{x}}_2^T \mathbf{K}_2^{-T}] [\mathbf{t}]_\times^T \mathbf{R} [\lambda_1 \mathbf{K}_1^{-1} \tilde{\mathbf{x}}_1] = 0 \quad (\text{A.25})$$

und man erhält die fundamentale oder F-Matrix in Abhängigkeit der intrinsischen Kameramatrizen  $\mathbf{K}_1, \mathbf{K}_2$  gemäß

$$\mathbf{F} = \mathbf{K}_2^{-T} [\mathbf{t}]_\times^T \mathbf{R} \mathbf{K}_1^{-1} \quad (\text{A.26})$$

mit der entsprechenden Epipolarbedingung

$$\tilde{\mathbf{x}}_2^T \mathbf{F} \tilde{\mathbf{x}}_1 = 0 \quad (\text{A.27})$$

für zwei korrespondierende Bildpunkte  $\mathbf{x}_1$  und  $\mathbf{x}_2$ .

Die Ebene, die durch die beiden Kamerazentren und den Szenenpunkt  $\mathbf{X}_1$  aufgespannt wird, nennt man Epipolarebene. Die Epipolargeraden

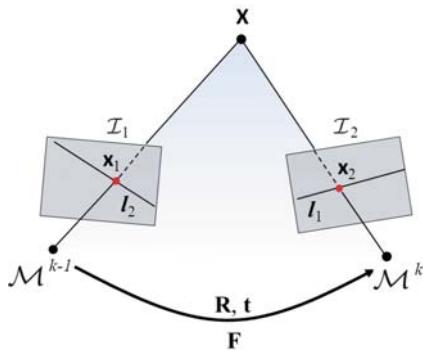
$$\mathbf{l}_1 = \mathbf{F} \tilde{\mathbf{x}}_1 \text{ und } \mathbf{l}_2 = \mathbf{F}^T \tilde{\mathbf{x}}_2 \quad (\text{A.28})$$

zu den beiden Punkten  $\mathbf{x}_1, \mathbf{x}_2$  bezüglich  $\mathbf{F}$  sind die Schnittgeraden der Epipolarebene mit den Bildebenen. Alle Epipolargeraden schneiden sich in einem Punkt, dem sog.

Eipol e mit

$$\mathbf{F}\mathbf{e} = \mathbf{0} \quad (\text{A.29})$$

Die Epipolarbedingung fordert letztlich, dass der Punkt  $\mathbf{x}_2$  auf der Epipolgeraden  $\mathcal{l}_1$  liegt bzw. der Punkt  $\mathbf{x}_1$  auf  $\mathcal{l}_2$ .



**Abbildung A.1:** Epipolarlinien und Eipolarebene zu zwei korrespondierenden Bildpunkten.

Die Epipolargeometrie ist in der vorliegenden Arbeit im Wesentlichen in zwei Bereichen von Bedeutung. Zum einen dient die Epipolarbedingung als Kriterium bei der Auswahl der Bildpunktcorrespondenzen und der Detektion von Ausreißern. Zum anderen wird das Schätzprinzip bei der Initialisierung der Bewegungsschätzung herangezogen.

#### A.4 Initialisierung bei rein translatorischer Bewegung

Da die kontinuierliche Bewegungsschätzung mittels SBA eine Parameterinitialisierung benötigt, wird hier der Einfachheit halber eine Initialisierung bei einer Geradeausfahrt vorgeschlagen. Wenngleich in der Literatur Verfahren und Methoden zur F-Matrix-Schätzung und Faktorisierung bei allgemeinen Bewegungen existieren (z. B. [Har97]),

so hat sich die rein translatorische Bewegungsschätzung als überaus robust und effektiv erwiesen. Als Eingangsdaten werden lediglich die Punktkorrespondenzen zweier aufeinander folgender Bilder benötigt.

Gesucht ist also der Translationsvektor  $\mathbf{v} = [v_x, v_y, v_z]^T$  bei bekannten intrinsischen Parametern  $\mathbf{K}$ . Da der Translationsbetrag initial unbekannt ist und sich auch ohne weitere Informations- bzw. Sensorquellen nicht näher bestimmen lässt, kann man  $\mathbf{v}$  skalieren, sodass der Translationsvektor o. B. d. A. die Form  $\mathbf{v} = [v_x, v_y, 1]^T$  annimmt, wodurch sich die Anzahl der Freiheitsgrade auf zwei verringert.

Ähnlich zur Ausreißereliminierung der Punktkorrespondenzen in Abschnitt 3.1.1 wird zunächst ein RANSAC-Verfahren [Fis81] durchgeführt, um die Punktpaare hinsichtlich der Einhaltung der Epipolarbedingung zu prüfen. Allerdings gestaltet sich hier die Berechnung der F-Matrix wesentlich simpler. Die Epipolarbedingung vereinfacht sich auf die Form

$$\tilde{\mathbf{x}}_2^T \mathbf{K}^{-T} \underbrace{\begin{bmatrix} 0 & -1 & v_y \\ 1 & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix}}_{[\mathbf{v}]_x^T} \mathbf{K}^T \tilde{\mathbf{x}}_1 = 0. \quad (\text{A.30})$$

Seien  $\tilde{\mathbf{x}}'_{1_o} = [x'_{1_o}, y'_{1_o}, 1]^T = \mathbf{K}^{-1} \tilde{\mathbf{x}}_{1_o}$  und  $\tilde{\mathbf{x}}'_{2_o} = [x'_{2_o}, y'_{2_o}, 1]^T = \mathbf{K}^{-1} \tilde{\mathbf{x}}_{2_o}$  die intrinsisch kalibrierten Projektionen der  $\mathfrak{P}_{\mathcal{S}}$  Punktpaare aus der RANSAC-Submenge  $\mathcal{S}_\xi$ , mit  $o = 1, \dots, \mathfrak{P}_{\mathcal{S}}$  und  $\xi = 1, \dots, \mathfrak{M}_{\mathcal{S}}$ . Dann muss für alle  $\mathfrak{P}_{\mathcal{S}}$  Punktkorrespondenzen je RANSAC-Submenge die Bedingung

$$\tilde{\mathbf{x}}'_{2_o}^T [\mathbf{v}]_x^T \tilde{\mathbf{x}}'_{1_o} = 0 \quad (\text{A.31})$$

in Abhängigkeit der Bewegungsrichtung  $\mathbf{v}_\xi$  gelten, was sich in das LGS

$$\underbrace{\begin{bmatrix} \vdots & \vdots \\ y'_{2_o} - y'_{1_o} & x'_{1_o} - x'_{2_o} \\ \vdots & \vdots \end{bmatrix}}_{=: \mathbf{A}} \underbrace{\begin{bmatrix} v_{x_\xi} \\ v_{y_\xi} \end{bmatrix}}_{=: \mathbf{x}} = \underbrace{\begin{bmatrix} \vdots \\ x'_{2_o} y'_{1_o} - x'_{1_o} y'_{2_o} \\ \vdots \end{bmatrix}}_{=: \mathbf{b}} \quad (\text{A.32})$$

umformulieren lässt, mit dem zu bestimmenden Parametervektor  $\mathbf{x}$ .

Die Anzahl der Punktkorrespondenzen  $\mathfrak{P}_{\mathcal{S}}$  je Submenge für das RANSAC-Verfahren zur Lösung von Gl. A.32 beträgt mindestens zwei. In Versuchen hat sich der Wert  $\mathfrak{P}_{\mathcal{S}} = 8$  bewährt.

Das i. d. R. überbestimmte Gleichungssystem A.32 lässt sich mit Hilfe der Pseudoinversen nach  $\mathbf{x}$  auflösen. Aus den so ermittelten Parametern  $v_{x_\xi}, v_{y_\xi}$  lässt sich wiederum eine F-Matrix

$$\mathbf{F} = \mathbf{K}^{-T} \begin{bmatrix} 0 & -1 & v_{y_\xi} \\ 1 & 0 & -v_{x_\xi} \\ -v_{y_\xi} & v_{x_\xi} & 0 \end{bmatrix} \mathbf{K}^{-1} \quad (\text{A.33})$$

berechnen. Um zu prüfen, welche Korrespondenzpaare aus der Menge aller  $\mathfrak{P}$  Punktepaare zur Konsensusmenge von  $v_{x_\xi}, v_{y_\xi}$  gehören, wird der quadrierte Euklidische Abstand zweier korrespondierender Punkte  $\{\mathbf{x}_{1_p}, \mathbf{x}_{2_p}\}$  zu den entsprechenden Epipolarlinien

$$\mathbf{l}_2^p = \mathbf{F} \tilde{\mathbf{x}}_1^p = [l_{1_p}^2, l_{2_p}^2, l_{3_p}^2]^T \quad \text{und} \quad (\text{A.34})$$

$$\mathbf{l}_1^p = \mathbf{F}^T \tilde{\mathbf{x}}_2^p = [l_{1_p}^1, l_{2_p}^1, l_{3_p}^1]^T, \quad p = 1, \dots, \mathfrak{P} \quad (\text{A.35})$$

als Gütekriterium herangezogen. Der Abstand zwischen einem Bildpunkt und der entsprechenden Epipolarlinie berechnet sich zu

$$d_1^p = \frac{[\mathbf{l}_1^p]^T \tilde{\mathbf{x}}_{1_p}]^2}{(l_{1_p}^1)^2 + (l_{2_p}^1)^2} \quad \text{bzw.} \quad d_2^p = \frac{[\mathbf{l}_2^p]^T \tilde{\mathbf{x}}_{2_p}]^2}{(l_{1_p}^2)^2 + (l_{2_p}^2)^2}. \quad (\text{A.36})$$

Der Wert für den Schwellwert  $\Lambda_F$  wird auf ca. 1 Pixel gesetzt. Ein Punktepaar  $\{\mathbf{x}_{1p}, \mathbf{x}_{2p}\}$  gehört genau dann zur Menge der *Inlier*, wenn die Bedingungen  $d_1^p < \Lambda_F$  und  $d_2^p < \Lambda_F$  erfüllt sind. Nach vollständigem Durchlaufen der RANSAC-Schleife sind letztlich die gesuchten Parameter  $v_x, v_y$  diejenigen  $v_{x_\zeta}, v_{y_\zeta}$ , die die größte Konsensusmenge erzielten. Die Parameter  $v_x$  und  $v_y$  werden anschließend mit einem Gauß-Newton-Verfahren derart optimiert, dass sich die Abstände aus Gl. A.36 minimieren [Pre07].

## A.5 Überführung der Bewegungstransformationen in die Grundebene

Für das erweiterte Kalibriermodell werden die Kamerabewegungen  $M_i$  und  $M_j$  der Module  $\mathcal{M}_i$  bzw.  $\mathcal{M}_j$  aus dem  $\mathbb{R}^3$  in eine gemeinsame Bodenebene transformiert, sodass sich das Problem der Bestimmung der extrinsischen Parameter auf den zweidimensionalen Raum reduziert. Hierfür ist die Kenntnis der Rotationsparameter notwendig, die das Koordinatensystem der Kamera derart drehen, sodass die  $X$ - und  $Z$ -Achsen des Kamerakoordinatensystems parallel zur Bodenebene stehen. Für diese Rotation reichen bereits zwei Eulerwinkel  $\alpha_i$  und  $\gamma_i$  zur Drehung um die  $X$ - bzw.  $Z$ -Achse. Die Winkel sind entweder aus den Transformationen  $G_i, G_j$  gemäß Abschnitt 5.3.3 bekannt oder aus der in Abschnitt 5.3.1 ermittelten Rotationsachse  $\alpha_{Ri}$ . Die resultierenden Bewegungstransformationen der Module sollen also nach der Transformation o. B. d. A. folgende Gestalt aufweisen

$$\underline{M}_i = \left[ \begin{array}{c|c} \begin{matrix} \frac{1}{\pm 0} & := \psi \\ \mathbf{R}(\underbrace{\omega_x}_{\mathbf{0}}, \underbrace{\omega_{x_i}}_{\mathbf{0}}, \underbrace{\omega_z}_{\mathbf{0}}) & \end{matrix} & \begin{matrix} v_x \\ v_y \\ v_z \end{matrix} \\ \hline \mathbf{0}^T & 1 \end{array} \right], \quad (\text{A.37})$$

d. h. die transformierten Rotationsachsen werden auf die kanonische Form  $\mathbf{e}_2 = [0, 1, 0]^T$  transformiert.

Sind  $\alpha_i$  und  $\gamma_i$  noch nicht explizit aus  $\mathbf{G}_i$  bekannt, so lassen sich die beiden Winkel mit Hilfe der Rotationsachse  $\boldsymbol{\alpha}_{Ri}$  bestimmen. Gesucht sind dann die Winkel, für die gilt

$$\boldsymbol{\alpha}_{Ri} = \mathbf{R}(\alpha_i, 0, \gamma_i) \mathbf{e}_2 = \begin{bmatrix} \sin(\gamma_i) \\ \cos(\alpha_i) \cos(\gamma_i) \\ -\sin(\alpha_i) \cos(\gamma_i) \end{bmatrix} \stackrel{5.15}{=} \begin{bmatrix} g_{1i} \\ g_{2i} \\ g_{3i} \end{bmatrix} \quad (\text{A.38})$$

woraus sich

$$\begin{aligned} \gamma_i &= \arcsin(a_{x_i}) \\ \alpha_i &= \arccos\left(\frac{a_{y_i}}{\cos(\gamma_i)}\right) \text{ bzw. } \arcsin\left(-\frac{a_{z_i}}{\cos(\gamma_i)}\right) \end{aligned} \quad (\text{A.39})$$

ergibt, mit zwei Lösungen für  $\alpha_i$ , wobei der gesuchte Winkel um die  $X$ -Achse dasjenige  $\alpha_i$  ist, für den der Euklidische Abstand zum Einheitsvektor  $\mathbf{e}_2$  minimal ist.

Die 2D-Bewegung  $\underline{\mathbf{M}}_i$  innerhalb der Grundebene erhält man aus der 3D-Bewegung  $\mathbf{M}_i$  durch einfache Rotation mit  $\mathbf{R}(\alpha_i, 0, \gamma_i)$ :

$$\underline{\mathbf{M}}_i = \begin{bmatrix} \mathbf{R}(\alpha_i, 0, \gamma_i) & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \cdot \mathbf{M}_i \quad (\text{A.40})$$

mit

$$\zeta_{\underline{\mathbf{M}}}(\underline{\mathbf{M}}_i) = (\underline{v}_{xi}, \underline{v}_{zi}, \psi_i)^T. \quad (\text{A.41})$$

Der translatorische Restfehler  $\varepsilon_{\underline{\mathbf{M}}}$  ergibt sich aus dem Quadrat der verbleibenden  $y$ -Koordinate des Translationsvektors in der Grundebene gemäß Gl. A.37 zu

$$\varepsilon_{\underline{\mathbf{M}}}(\alpha_i, \gamma_i) = \underline{v}_{yi}^2. \quad (\text{A.42})$$

An dieser Stelle sei erwähnt, dass die Transformation in die Grundebene unabhängig von der translatorischen Höhe  $h_{yi}$  ist, und nur von zwei Eulerwinkeln  $\alpha_i, \gamma_i$  abhängt.  $\alpha_i$  und  $\gamma_i$  lassen sich aus der Rotationsachse (Abschnitt 5.3.1) extrahieren, ohne explizit

auf 3D-Messdaten der Szene wie in Abschnitt 5.3.3 zurückgreifen zu müssen. D. h. auch ohne Kenntnis der Höhenkomponente (z. B. bei Unkenntnis der Absolutgeschwindigkeit der Plattform) ist eine Überführung der Bewegungstransformationen in eine Ebene parallel zur tatsächlichen Bodenebene möglich.

## A.6 Berechnung der Vogelperspektive

Im folgenden wird am Beispiel eines Zwei-Kamera-Systemaufbaus gezeigt, wie sich anhand der Kalibrierparameter eine Rektifizierung der Bodenebene erreichen lässt. Diese Rektifizierung wird im Kontext von Straßen häufig auch als Vogelperspektive (engl.: *Bird's Eye View*) bezeichnet.

Seien aus der Kalibrierung die extrinsischen Parameter  $\mathbf{R}_{12}$  sowie  $\mathbf{t}_{12}$  zwischen den Modulen  $\mathcal{M}_1$  und  $\mathcal{M}_2$  bekannt, sowie deren intrinsische Parameter  $\mathbf{K}_1$  und  $\mathbf{K}_2$  mit

$$\tilde{\mathbf{x}}_1 = \mathbf{K}_1 \mathbf{X}_1 \quad (\text{A.43})$$

$$\tilde{\mathbf{x}}_2 = \mathbf{K}_2 \mathbf{X}_2 \quad (\text{A.44})$$

$$\mathbf{X}_2 = \mathbf{R}_{12} \mathbf{X}_1 + \mathbf{t}_{12}. \quad (\text{A.45})$$

$\mathbf{X}_1$  und  $\mathbf{X}_2$  bezeichnen Szenenpunkte in den Koordinatensystemen von  $\mathcal{M}_1$  bzw.  $\mathcal{M}_2$ ,  $\mathbf{x}_1$  und  $\mathbf{x}_2$  die entsprechenden projizierten Bildpunkte. Durch Einfügen erhält man

$$\tilde{\mathbf{x}}_2 = \mathbf{K}_2 [\mathbf{R}_{12} \mathbf{X}_1 + \mathbf{t}_{12}]. \quad (\text{A.46})$$

$\mathcal{M}_1$  sei o. B. d. A. das Referenzmodul. Zusätzlich sei aus dem erweiterten Kalibriermodell die Bodenebene  $\mathbf{g} = [g_1, g_2, g_3, g_4]^T$  entsprechend den Gln. 5.15 und 5.14 in Abhängigkeit der Parameter  $\alpha, \gamma$  und  $h_y$  von  $\mathcal{M}_1$  bzw. die Euklidische Transformation  $[\mathbf{R}_G | \mathbf{t}_G]_{3 \times 4}$  bekannt, mit  $\mathbf{R}_G = \mathbf{R}(\alpha, 0, \gamma)$  und  $\mathbf{t}_G = [0, h_y, 0]^T$ . Die Transformation des Punktes  $\mathbf{X}_1$  in die Bodenebene erhält man aus Gl. 5.13 gemäß

$$\mathbf{X}_1 = \underbrace{\mathbf{R}_G^T}_{=: \mathbf{R}_G^{-1}} \underbrace{\mathbf{X}_1 - \mathbf{R}_G^T \mathbf{t}_G}_{=: \mathbf{t}_G^{-1}} \quad (\text{A.47})$$

wobei  $\mathbf{R}_G^{-1}$  durch seine Spaltenvektoren dargestellt werden kann

$$\mathbf{R}_G^{-1} = \begin{bmatrix} \mathbf{m}_{G_1} & \mathbf{m}_{G_2} & \mathbf{m}_{G_3} \end{bmatrix}. \quad (\text{A.48})$$

Da  $\underline{\mathbf{X}}_1$  in der  $X - Z$ -Ebene liegt und daher nur zwei Koordinaten besitzt, also  $\underline{\mathbf{X}}_1 = [X_1, 0, Z_1]^T$ , lässt sich Gl. A.47 umformen zu

$$\underline{\mathbf{X}}_1 = \underbrace{\begin{bmatrix} \mathbf{m}_{G_1} & \mathbf{m}_{G_3} & \mathbf{t}_G^{-1} \end{bmatrix}}_{=: \mathbf{H}_G} \underbrace{\begin{bmatrix} \underline{X}_1 \\ \underline{Z}_1 \\ 1 \end{bmatrix}}_{=: \tilde{\mathbf{x}}_G}. \quad (\text{A.49})$$

Aus Gl. 5.15 ergibt sich die Normalengleichung

$$\mathbf{n}^T \underline{\mathbf{X}}_1 = g_4 \iff \frac{1}{g_4} \mathbf{n}^T \underline{\mathbf{X}}_1 = 1 \quad (\text{A.50})$$

mit

$$\mathbf{n} = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} \stackrel{5.15}{=} \begin{bmatrix} \sin(\gamma) \\ \cos(\alpha)\cos(\gamma) \\ -\sin(\alpha)\cos(\gamma) \end{bmatrix}. \quad (\text{A.51})$$

$\mathbf{n}$  ist hierbei die mittlere Spalte von  $\mathbf{R}_G$ . Durch Einfügen der Gln. A.50 und A.49 in Gl. A.46 ergibt sich

$$\tilde{\mathbf{x}}_2 = \mathbf{K}_2 \left[ \mathbf{R}_{12} \underline{\mathbf{X}}_1 + \mathbf{t}_{12} \cdot \left[ \frac{\mathbf{n}^T \underline{\mathbf{X}}_1}{g_4} \right] \right] \quad (\text{A.52})$$

$$= \mathbf{K}_2 \left[ \mathbf{R}_{12} + \frac{\mathbf{t}_{12} \mathbf{n}^T}{g_4} \right] \underline{\mathbf{X}}_1 \quad (\text{A.53})$$

$$= \underbrace{\mathbf{K}_2 \left[ \mathbf{R}_{12} + \frac{\mathbf{t}_{12} \mathbf{n}^T}{g_4} \right]}_{=: \mathbf{H}_{BEV}} \mathbf{H}_G \tilde{\mathbf{x}}_G \quad (\text{A.54})$$

Entsprechend Gl. 5.15 hat  $g_4$  genau den Wert  $-h_y$ . Dies impliziert die durchaus sinnvolle Bedingung, dass  $h_y \neq 0$  gilt, sich die Position der Kamera also nicht innerhalb der Bodenebene befinden darf.

Mit Hilfe der Homographie  $\mathbf{H}_{BEV}$  lässt sich nun ein Bildpunkt von Modul  $\mathcal{M}_2$  in die Grundebene des Referenzmoduls  $\mathcal{M}_1$  abbilden [Har06].

## A.7 Zur musterbasierten extrinsischen Kalibrierung

Zhang [Zha00] beschreibt eine musterbasierte Methode zur intrinsischen Kalibrierung monokularer Kameras. Im Zuge dieser Kalibrierung werden sowohl die intrinsischen Parameter berechnet als auch die extrinsischen Transformationen zwischen der Kamera und dem Muster in den jeweiligen Ansichten. Im Folgenden soll kurz das Prinzip zur Bestimmung der extrinsischen Transformationen erläutert werden.

Bei gegebener Rotationsmatrix  $\mathbf{R} = [\mathbf{r}_1 \mathbf{r}_2 \mathbf{r}_3]$ , dem Translationsvektor  $\mathbf{t}$  und der intrinsischen Projektionsmatrix  $\mathbf{K}$  ergibt sich die Projektion des homogenen 3D Punkts  $\tilde{\mathbf{X}} = [X, Y, Z, 1]^T$  gemäß

$$\lambda \tilde{\mathbf{x}} = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}] \tilde{\mathbf{X}}. \quad (\text{A.55})$$

Der Punkt  $\mathbf{X}$  sei ein Punkt im Koordinatensystem des (wohlbekannten) Musters. Bei planaren Mustern liegen alle Objektpunkte innerhalb einer Ebene, sodass für alle Punkte des Musters o. B. d. A.  $Z = 0$  gelte. Damit lässt sich Gl.A.55 umformulieren zu

$$\lambda \tilde{\mathbf{x}} = \mathbf{K}[\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{t}] \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{K}[\mathbf{r}_1, \mathbf{r}_2, \mathbf{t}] \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \quad (\text{A.56})$$

Aus Gl. A.56 erhält man die Homographie

$$\mathbf{H} = \mathbf{K}[\mathbf{r}_1, \mathbf{r}_2, \mathbf{t}] = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]. \quad (\text{A.57})$$

Zur Bestimmung von  $\mathbf{H}$  sind mindestens vier Punktkorrespondenzen, d. h. Zuordnungen zwischen Bild- und Objektpunkt, notwendig (siehe z. B. [Har06]). Bei bekannten intrinsischen Parametern  $\mathbf{K}$  lassen sich die extrinsischen Parameter bestimmen gemäß

$$\begin{aligned}\mathbf{r}_1 &= c \cdot \mathbf{K}^{-1} \mathbf{h}_1 \\ \mathbf{r}_2 &= c \cdot \mathbf{K}^{-1} \mathbf{h}_2 \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 \\ \mathbf{t} &= c \cdot \mathbf{K}^{-1} \mathbf{h}_3\end{aligned} \quad (\text{A.58})$$

mit

$$c = \frac{1}{\mathbf{K}^{-1} \mathbf{h}_1} = \frac{1}{\mathbf{K}^{-1} \mathbf{h}_2}. \quad (\text{A.59})$$

Die daraus resultierenden Rotationsparameter werden als Startlösung  $\hat{\mathbf{R}}$  für eine weitere Optimierung verwendet, um sicherzugehen, dass die gesuchte Rotationsmatrix  $\mathbf{R}$  die Eigenschaft  $\mathbf{R}^T \mathbf{R} = \mathbf{I}$  erfüllt. Dies wird numerisch durch eine Minimierung der Frobeniusnorm

$$\min_{\mathbf{R}} \|\mathbf{R} - \hat{\mathbf{R}}\|_F^2 \quad (\text{A.60})$$

erreicht (weitere Details sind aus [Zha00] zu entnehmen).

---

## Literaturverzeichnis

---

- [And01] ANDREFF, N.; HORAUD, R. und ESPIAU, B.: Robot hand-eye calibration using structure-from-motion. *International Journal of Robotics Research* (2001), Bd. 20(3): S. 228–248 S. 16
- [Aza95] AZARBAYEJANI, A. und PENTLAND, A. P.: Recursive Estimation of Motion, Structure, and Focal Length. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1995), Bd. 17(6): S. 562–575 S. 51
- [Bac06] BACHMANN, A. und DANG, T.: Multiple Object Detection under the Constraint of Spatiotemporal Consistency. *Proceedings of the IEEE Intelligent Transportation Systems Conference* (2006): S. 295–300 S. 4
- [Bae05] BAEHRING, D.: Detection of close cut-in and overtaking vehicles for driver assistance based on planar parallax. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2005): S. 290–295 S. 3 und 61
- [Bak07] BAKER, S.; SCHARSTEIN, D.; LEWIS, J. P.; ROTH, S.; BLACK, M. J. und SZELISKI, R.: A Database and Evaluation Methodology for Optical Flow. *Proceedings of the IEEE International Conference on Computer Vision* (2007), Bd. 1–6: S. 588–595 S. 48

- [Bar94] BARRON, J. L. und BEAUCHEMIN, D. J. Fleet S.: Performance of Optical Flow Techniques. *International Journal of Computer Vision* (1994), Bd. 12: S. 43–77  
S. 48
- [Bro86] BROWN, D. C. und FRYER, J. G.: Lens Distortion for Close Range Photogrammetry. *Photogrammetric Engineering and Remote Sensing* (1986), Bd. 52(1): S. 51–58  
S. 31
- [Car11] CARRERA, G.; ANGELI, A. und DAVISON, A. J.: SLAM-based automatic extrinsic calibration of a multi-camera rig (2011): S. 2652–2659  
S. 13
- [Cho79] CHONG, C.-Y.: Hierarchical Estimation. *Proceedings of MIT/ONR Workshop on C3* (1979)  
S. 43
- [Cli08] CLIPP, B.; KIM, J. H.; FRAHM, J. M.; POLLEFEYS, M. und HARTLEY, R. I.: Robust 6DOF Motion Estimation for Non-Overlapping, Multi-Camera Systems. *Proceedings of the IEEE Workshop on Applications of Computer Vision* (2008): S. 1–8  
S. 16
- [Cou02] COULOMBEAU, P. und LAURGEAU, C.: Vehicle yaw, pitch, roll and 3d lane shape recovery by vision. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2002), Bd. 2: S. 619–625  
S. 15
- [Dan98] DANILIDIS, K.: Hand-eye calibration using dual quaternions. *International Journal of Robotics Research* (1998), Bd. 18: S. 286–298  
S. 14
- [Dan02] DANG, T.; HOFFMANN, C. und STILLER, C.: Fusing Optical Flow and StereoDisparity for Object Tracking. *Proceedings of the IEEE International Conference on Intelligent Transportation Systems* (2002): S. 112–117  
S. 4
- [Dan04] DANG, T. und HOFFMANN, C.: Stereo Calibration in Vehicles. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2004): S. 268–273  
S. 13

- [Dan07] DANG, T.: Kontinuierliche Selbstkalibrierung von Stereokameras. *Dissertation. Fakultät für Maschinenbau der Universität Karlsruhe (TH)* (2007) S. 6, 50, 62 und 99
- [Dan09] DANG, T.; HOFFMANN, C. und STILLER, C.: Continuous stereo self-calibration by camera parameter tracking. *IEEE Transactions on Image Processing* (2009), Bd. 18: S. 1536–1550 S. 13
- [DAR] DARPA: Grand Challenge, <http://archive.darpa.mil/grandchallenge/index.asp> S. 4
- [Dav07] DAVISON, A. J.; REID, I. D.; MOLTON, N. D. und STASSE, O.: MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2007), Bd. 29(6): S. 1052–1067 S. 16 und 19
- [Dic02] DICKMANNS, E. D.: The development of machine vision for road vehicles in the last decade. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2002), Bd. 1: S. 268–281 S. 3
- [Emt10] EMTER, T. und ULRICH, T.: Visuelle Information zur robusten Zuordnung von Landmarken für die Navigation mobiler Roboter. *VID Forum Bildverarbeitung* (2010): S. 95–106 S. 5
- [Esq07] ESQUIVEL, S.; WÖLK, F. und KOCH, R.: Calibration of a Multi-camera Rig from Non-overlapping Views. *Proceedings of the DAGM-Symposium* (2007) S. 16, 74 und 76
- [Fau87] FAUGERAS, O.; TOSCANI, G. und LUSTMANN, E.: Motion and Structure from Motion from Point and Line Matches. *Proceedings of the IEEE International Conference of Computer Vision* (1987): S. 25–34 S. 51
- [Fau92] FAUGERAS, O.; LUONG, Q.-T. und MAYBANK, S. J.: Camera Self-Calibration: Theory and Experiments. *Proceedings of the European Conference on Computer Vision* (1992): S. 321–334 S. 19

- [Fis81] FISCHLER, M. A. und BOLLES, R. C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* (1981), Bd. 24(6): S. 381–395  
S. 50, 78 und 183
- [Foe94] FOELLINGER, O.: Regelungstechnik. *Hüthig Verlag* (1994) S. 98
- [Gol78] GOLDSTEIN, H.: Klassische Mechanik. *Verlag Akademische Verlag-Gesellschaft, 5. Auflage* (1978) S. 34
- [Gra98] GRASSIA, S.: Practical Parameterization of Rotations Using the Exponential Map. *Journal of Graphic Tools* (1998), Bd. 3(3): S. 29–48 S. 178
- [Haa98] HAAG, M.: Bildfolgenauswertung zur Erkennung der Absichten von Straßenverkehrsteilnehmern. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (1998) S. 35 und 94
- [Ham83] HAM, F. M. und BROWN, R. G.: Observability, Eigenvalues, and Kalman Filtering. *IEEE Transactions on Aerospace and Electronic Systems* (1983), Bd. AES-19(2): S. 269–273 S. 98
- [Har88] HARRIS, C. und STEPHENS, M.: A Combined Corner and Edge Detector. *Proceedings of the 4th Alvey Vision Conference* (1988): S. 147–151 S. 48
- [Har95] HARTLEY, R. I. und STURM, P.: Triangulation. *Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns* (1995), Bd. 68(2): S. 146–157 S. 118 und 119
- [Har97] HARTLEY, R. I.: In Defense of the Eight-Point Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (1997), Bd. 19(6): S. 580–593  
S. 50 und 182
- [Har06] HARTLEY, R. und ZISSERMAN, A.: Multiple View Geometry in Computer Vision. *Cambridge University Press, 2. Auflage* (2006) S. 16, 189 und 190

- [Har07] HARTLEY, R.I. und KAHL, F: Critical Configurations for Projective Reconstruction from Multiple Views. *International Journal of Computer Vision* (2007), Bd. 71(1): S. 5–47 S. 47
- [Har09] HARTLEY, R. I. und KAHL, F: Global optimization through rotation space search. *International Journal of Computer Vision* (2009), Bd. 82(1): S. 64–79 S. 16 und 171
- [Hei00] HEIMES, F: Sichtsystemgestützte Fahrerassistenz im innerstädtischen Straßenverkehr. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (2000) S. 4
- [Hei02] HEIMES, F und NAGEL, H.-H.: Towards Active Machine-Vision-Based Driver Assistance for Urban Areas. *International Journal of Computer Vision* (2002), Bd. 50(1): S. 5–34 S. 4
- [Hel11] HELLER, J.; HAVLENA, M.; SUGIMOTO, A. und PAJDLA, T.: Structure from motion based hand-eye calibration using L1 minimization. *Proceedings of the IEEE International Conference of Computer Vision and Pattern Recognition* (2011): S. 3497–3503 S. 16
- [Hoe80] HOEPCKE, W.: Fehlerlehre und Ausgleichsrechnung. *Verlag Walter de Gruyter, Cambridge* (1980) S. 40
- [Hor95] HORAUD, R. und DORNAIKA, F: Hand-Eye Calibration. *International Journal of Robotic Research* (1995), Bd. 14(3): S. 195–210 S. 14, 58 und 177
- [Hum08] HUMMEL, B.; THIEMANN, W. und LULCHEVA, I.: Scene Understanding of Urban Road Intersections with Description Logic. *Logic and Probability for Scene Interpretation, Dagstuhl Seminar Proceedings* (2008) S. 4
- [Jae97] JAEHNE, B.: Digitale Bildverarbeitung. *Springer Verlag, 4.Auflage* (1997) S. 48 und 49

- [Jul97] JULIER, S. J. und UHLMANN, J. K.: A New Extension of the Kalman Filter to Nonlinear Systems (1997): S. 182–193 S. 37
- [Kah99] KAHL, F. und TRIGGS, B.: Critical Motions in Euclidean Structure from Motion. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (1999), Bd. 2: S. 366–372 S. 47
- [Kal60] KALMAN, R. E.: A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering* (1960), Bd. 82: S. 35–45 S. 37
- [Kan80] KANADE, T.: Region Segmentation: Signal vs. Semantics. *International Journal of Computer Graphics and Image Processing* (1980), Bd. 13: S. 279–297 S. 4
- [Kan05] KANG, J.; COHEN, I. und YUAN, C.: Detection and Tracking of Moving Objects from a Moving Platform in Presence of Strong Parallax. *Proceedings of the IEEE International Conference on Computer Vision* (2005), Bd. 1: S. 10–17 S. 4
- [Koe83] KOECHER, M. und REMMERT, R.: Hamiltonsche Quaternionen, in H.-D. Ebbinghaus et al.: Zahlen. *Springer-Verlag, Berlin* (1983) S. 176
- [Kol92] KOLLER, D.: Detektion, Verfolgung und Klassifikation bewegter Objekte in monokularen Bildfolgen am Beispiel von Straßenverkehrsszenen. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (1992) S. 4
- [Kru01] KRUEGER, W.: Robust and efficient map-to-image registration with line segments. *Machine Vision and Applications* (2001), Bd. 13(1): S. 38–50 S. 3
- [Krz00] KRZANOWSKI, W. J.: Principles of Multivariate Analysis. Rev. ed. *Oxford University Press, Oxford* (2000) S. 99
- [Kum08] KUMAR, R. K.; ILIE, A.; FRAHM, J. M. und POLLEFEYS, M.: Simple calibration of non-overlapping cameras with a mirror. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (2008): S. 1–7 S. 13

- [Lam07] LAMPRECHT, B.; RASS, S.; FUCHS, S. und KYAMAKYA, K.: Extrinsic Camera Calibration for an On-Board Two Camera System Without Overlapping Field of View. *Proceedings of the IEEE Intelligent Transportation Systems Conference* (2007): S. 265–270 S. 14
- [Leb11] LEBRALY, P.; ROYER, E.; AIT-AIDER, O.; DEYMIER, C. und DHOME, M.: Fast calibration of embedded non-overlapping cameras. *Proceedings of the IEEE International Conference on Robotics and Automation* (2011): S. 221–227 S. 13
- [Leu00] LEUCK, H.: Untersuchungen zu einer systematischen Leistungssteigerung in der modellbasierten Bildfolgenauswertung. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (2000) S. 35 und 94
- [Lev44] LEVENBERG, K.: A Method for the Solution of Certain Problems in Least Squares. *Quart. Appl. Math.* (1944), Bd. 2 S. 36
- [LH81] LONGUET-HIGGINS, H. C.: A computer algorithm for reconstructing a scene from two projections. *Nature* (1981), Bd. 293: S. 133–135 S. 51
- [LH87] LONGUET-HIGGINS, H. C.: A Computer Algorithm for Reconstructing a Scene from Two Projections. *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms* (1987): S. 61–62 S. 51
- [Liu08] LIU, Y. C.; LIN, K. Y. und CHEN, Y. S.: Bird's-Eye View Vision System for Vehicle Surrounding Monitoring. *International Workshop on Robot Vision* (2008): S. 207–218 S. 5
- [Lou09] LOURAKIS, M. I. A. und ARGYROS, A. A.: SBA: A software package for generic sparse bundle adjustment. *ACM Transactions on Mathematical Software* (2009), Bd. 36: S. 1–30 S. 54 und 91
- [Luc81] LUCAS, B. D. und KANADE, T.: An Iterative Image Registration Technique with an Application to Stereo Vision. *Proceedings of the International Joint Conference on Artificial Intelligence* (1981), Bd. 2: S. 674–679 S. 48

- [Luo95] LUONG, Q.-T. und FAUGERAS, O.: The Fundamental matrix: theory, algorithms, and stability analysis. *International Journal of Computer Vision* (1995), Bd. 17(1): S. 43–75 S. 180
- [Mar63] MARQUARDT, D.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *SIAM Journal on Applied Mathematics* (1963), Bd. 11(2): S. 431–441 S. 36
- [May92] MAYBANK, S. J. und FAUGERAS, O.: A theory of self-calibration of a moving camera. *International Journal of Computer Vision* (1992), Bd. 8(2): S. 123–151 S. 51
- [Mik10] MIKSCH, M.; YANG, B. und ZIMMERMANN, K.: Homography-based extrinsic self-calibration for cameras in automotive applications. *Workshop on Intelligent Transportation* (2010): S. 17–22 S. 15
- [Mor06] MORDOHAI, A. F.; AKBARZADEH, A.; FRAHM, J. M.; MORDOHAI, P.; ENGELS, C.; GALLUP, D.; MERRELL, P.; PHELPS, M.; SINHA, S.; TALTON, B.; WANG, L.; YANG, Q.; STEWENIUS, H.; YANG, R.; WELCH, G.; TOWLES, H.; NISTER, D. und POLLEFEYS, M.: Towards Urban 3D Reconstruction From Video. *3DPVT* (2006): S. 1–8 S. 4
- [MT11] M. THOM, L. Krüger, M. Ulken und RITTER, W.: Headlight Range Calibration During Driving Operation. *Proceedings of the International Symposium on Automotive Lighting* (2011) S. 3
- [Mue07] MUELLER, M.; KRÜGER, W. und SAUR, G.: Robust image registration for fusion. *Information Fusion* (2007), Bd. 8(4): S. 347–353 S. 3
- [Mue08] MUELLER, T. und MÜLLER, M.: CACAMO - Computer-Aided Camouflage Assessment of Moving Objects. *SPIE Defense and Security: Infrared Imaging Systems: Design, Analysis, Modeling, and Testing* (2008), Bd. 6941: S. 69410V–69410V–12 S. 4

- [New10] NEWCOMBE, R. A. und DAVISON, A. J.: Live dense reconstruction with a single moving camera. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (2010): S. 1498–1505 S. 4
- [Nis04] NISTER, D.; NARODITSKY, O. und BERGEN, J.: Visual Odometry. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (2004): S. 652–659 S. 16
- [Pag09] PAGEL, F.: Robust Monocular Egomotion Estimation Based on an IEKF. *Proceedings of the Canadian Conference on Computer and Robot Vision* (2009): S. 213–220 S. 16
- [Pag10a] PAGEL, F. und WILLERSINN, D.: Motion-based Online Calibration for Non-overlapping Camera Views. *Proceedings of the IEEE Intelligent Transportation Systems Conference* (2010): S. 843–848 S. 22
- [Pag10b] PAGEL, Frank: Calibration of Non-Overlapping Cameras in Vehicles. *IEEE Intelligent Vehicles Symposium* (2010): S. 1178–1183 S. 14 und 153
- [Pag11] PAGEL, F. und WILLERSINN, D.: Kontinuierliche Kalibrierung und Fusion mobiler Kamerasyteme mit nichtüberlappenden Sichten in Fahrzeugen. 8. *Workshop Fahrerassistenzsysteme* (2011): S. 95–104 S. 71
- [Pag12] PAGEL, F.: Motion Adjustment for Extrinsic Calibration of Cameras with Non-overlapping Views. *Proceedings of the Canadian Conference on Computer and Robot Vision* (2012): S. 94–100 S. 61
- [Par94] PARK, F. und MARTIN, B.: Robot sensor calibration: solving  $ax=xb$  on the euclidean group. *IEEE Transactions on Robotics and Automation* (1994), Bd. 10(5): S. 717–721 S. 14
- [Par10] PARK, S. Y.; SONG, Y. S.; PARK, J. und KIM, H. J.: Monocular-vision-based Parking Assistance for Constrained Environment. *Proceedings of the IEEE*

- International Conference on Intelligent Transportation Systems* (2010): S. 617–622 S. 3
- [Pol99a] POLLEFEYS, M.: Self-Calibration and Metric 3D Reconstruction from Uncalibrated Image Sequences. *Dissertation. K.U. Leuven, Belgien* (1999) S. 4 und 5
- [Pol99b] POLLEFEYS, M.; KOCH, R. und VAN GOOL, L.: A Simple and Efficient Rectification Method for General Motion. *Proceedings of the IEEE International Conference on Computer Vision* (1999): S. 496–501 S. 5
- [Pol00] POLLEFEYS, M. und VAN GOOL, L.: Some Issues on Self-Calibration and Critical Motion Sequences. *Proceedings of the Asian Conference on Computer Vision* (2000): S. 893–898 S. 47
- [Pre07] PRESS, W.H.; TEUKOLSKY, S.A.; VETTERLING, W.T. und FLANNERY, B.P.: Numerical Recipes. *Cambridge University Press, 3. Auflage* (2007) S. 36 und 185
- [Rah04] RAHIMI, A.; DUNAGAN, B. und DARRELL, T.: Simultaneous calibration and tracking with a network of non-overlapping sensor. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition* (2004), Bd. 1: S. I–187–I–194 S. 13
- [Rul10] RULAND, T.; LOOSE, H.; PAJDLA, T. und KRÜGER, L.: Hand-Eye Autocalibration of Camera Positions on Vehicles. *Proceedings of the IEEE Intelligent Transportation Systems Conference* (2010): S. 367–372 S. 15
- [Rul11a] RULAND, T.; PAJDLA, T. und KRÜGER, L.: Global Optimization of Extended Hand-Eye Calibration. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2011): S. 740–745 S. 15
- [Rul11b] RULAND, T.; PAJDLA, T. und KRÜGER, L.: Robust Hand-Eye Self-Calibration. *Proceedings of the IEEE Intelligent Transportation Systems Conference* (2011): S. 87–94 S. 15

- [SÖ5] SÜNDERHAUF, N.; KONOLIGE, K.; LACROIX, S. und PROTZEL, P.: Visual Odometry using Sparse Bundle Adjustment on an Autonomous Outdoor Vehicle. *Tagungsband Autonome Mobile Systeme* (2005): S. 157–163 S. 16
- [Sch00] SCHMID, C.; MOHR, R. und BAUCKHAGE, C.: Evaluation of Interest Point Detectors. *International Journal of Computer Vision* (2000), Bd. 37(2): S. 151–172 S. 48
- [Sch05a] SCHLOSSER, M.: Lineare Schätzer zur multisensoriellen Objektverfolgung in verteilten, nichtlinearen Systemen. *Dissertation. Fakultät für Elektrotechnik der Universität Karlsruhe (TH)* (2005) S. 43
- [Sch05b] SCHMID, C. und MIKOLAJCZYK, K.: A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence* (2005), Bd. 27(10): S. 1615–1630 S. 48
- [Sch05c] SCHMIDT, J.; VOGT, F. und NIEMANN, H.: Calibration-free hand-eye calibration: A structure-from-motion approach. *Pattern Recognition, ser. Lecture Notes in Computer Science* (2005), Bd. 3663: S. 67–74 S. 16
- [Sch06] SCHOEN, T.; EIDEHALL, A. und GUSTAFSSON, F.: Lane departure detection for improved road geometry estimation. *Proceedings of the IEEE Intelligent Vehicles Symposium* (2006): S. 546–551 S. 3
- [Seo09] SEO, Y.; CHOI, Y. J. und LEE, S. W.: Branch-and-bound algorithm for globally optimal calibration of a camera-and-rotation-sensor system. *Proceedings of the IEEE International Conference of Computer Vision* (2009): S. 1173–1178 S. 16
- [Shi89] SHIU, Y. und AHMAD, S.: Calibration of wrist-mounted robotic sensors by solving homogeneous transform equations of the form  $ax=xb$ . *IEEE Transactions on Robotics and Automation* (1989), Bd. 5(1): S. 16–29 S. 14

- [Shi94] SHI, J. und TOMASI, C.: Good Features to Track. *Proceedings of the IEEE International Conference on Compter Vision and Pattern Recognition* (1994): S. 593–600 S. 48
- [Sim03] SIMOND, N. und RIVES, P.: Homography from a vanishing point in urban scenes. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems* (2003), Bd. 1: S. 1005–1010 S. 15
- [Sim06] SIMON, D.: Optimal State Estimation: Kalman, H-infinity, and Non-linear Approaches. *John Wiley & Sons* (2006) S. 37 und 38
- [Smi86] SMITH, R. C. und CHEESEMAN, P.: On the Representation and Estimation of Spatial Uncertainty. *International Journal of Robotics Research* (1986), Bd. 5(4) S. 23, 43 und 44
- [Sor85] SORENSEN, H. W.: Kalman Filtering: Theory and Application. *IEEE Press, Cambridge* (1985) S. 37
- [Sti01] STILLER, C.; KAMMEL, S.; HORN, J. und DANG, T.: The Computation of Motion. *Digital Image Sequence Processing, Compression and Analysis, CRC Press* (2001) S. 49
- [Sti08] STILLER, C.; KAMMEL, S.; PITZER, B.; ZIEGLER, J.; WERLING, M.; GINDELE, T.; JAGSZENT, D.; SCHROEDER, J.; THUY, M.; GOEBL, M.; v. HUNDELSHAUSEN, F.; PINK, O. und FRESE, C.: Team AnnieWAY's Autonomous System. *Journal of Field Robotics - Special Issue on the 2007 DARPA Urban Challenge, Part II* (2008), Bd. 25: S. 615–639 S. 4
- [Str06] STROBL, K. H. und HIRZINGER, G.: Optimal Hand-Eye Calibration. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems* (2006): S. 4647–4653 S. 14

- [Stu97] STURM, P. F.: Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (1997): S. 1100–1105 S. 47
- [Tom91] TOMASI, C. und KANADE, T.: Detection and Tracking of Point Features. *Tech. Report CMU-CS-91-132, Carnegie Mellon University* (1991) S. 48
- [Tri00] TRIGGS, B.; MCLAUCHLAN, P. F.; HARTLEY, R. I. und FITZGIBBON, A. W.: Bundle Adjustment - A Modern Synthesis. *Proceedings of the IEEE International Workshop on Vision Algorithms* (2000): S. 298–372 S. 20
- [Tsa89] TSAI, R. Y. und LENZ, R. K.: A New Technique for Fully Autonomous and Efficient 3D Robotics Hand/Eye Calibration. *IEEE Transactions on Robotics and Automation* (1989), Bd. 5(3): S. 345–358 S. 14, 58 und 74
- [Vac09] VACEK, S.: Videogestützte Umfelderfassung zur Interpretation von Verkehrssituationen für kognitive Automobile. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (2009) S. 4
- [Wac97] WACHTER, S.: Verfolgung von Personen in monokularen Bildfolgen. *Dissertation. Fakultät für Informatik der Universität Karlsruhe (TH)* (1997) S. 4
- [Wen87] WENG, J.; HUANG, T. S. und AHUJA, N.: Error Analysis of Motion Parameters from Image Sequences. *Proceedings of the IEEE International Conference of Computer Vision* (1987): S. 703–707 S. 47
- [Zha00] ZHANG, Z.: A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2000), Bd. 22(11): S. 1330–1334 S. 19, 31, 153, 189 und 190
- [Zom87] ZOMOTOR, A.: Fahrwerktechnik: Fahrverhalten, Band 10. *Verlag Vogel* (1987) S. 34



---

## Bisherige Veröffentlichungen des Autors

---

- [Pag08] PAGEL, Frank; ELIAS, Björn; GIESLER, Björn und WILLERSINN, Dieter: Analysis of Binocular Stereo Systems. *Proceedings of the International Workshop on Intelligent Transportation, Hamburg* (2008): S. 1–5
- [Pag09a] PAGEL, Frank: Robust Monocular Egomotion Estimation Based on an IEKE. *Proceedings of the Canadian Conference on Computer and Robot Vision, Kelowna, B.C, Kanada* (2009): S. 213–220
- [Pag09b] PAGEL, Frank: A Segment and Fusion-Based Stereo Approach. *Proceedings of the Canadian Conference on Computer and Robot Vision, Kelowna, B.C, Kanada* (2009): S. 170–177
- [Pag09c] PAGEL, Frank und WILLERSINN, Dieter: Extrinsische Kalibrierung einer Multikameraanordnung mit nichtüberlappenden Sichtfeldern auf einer mobilen Plattform. *7. Workshop Fahrerassistenzsysteme, Löwenstein* (2009): S. 116–125
- [Pag10a] PAGEL, Frank: Calibration of Non-Overlapping Cameras in Vehicles. *Proceedings of the IEEE Intelligent Vehicles Symposium, San Diego, Kalifornien, USA* (2010): S. 1178–1183
- [Pag10b] PAGEL, Frank: Monocular Panoramic 3D Reconstruction Based on a Particle Filter. *SPIE Defence, Security and Sensing, Unmanned Systems Technology XI, Orlando, Florida, USA* (2010), Bd. 7692: S. 76921S–76921S–12

- [Pag10c] PAGEL, Frank und RING, Jochen: Online Visualization of Noisy 3D Point Clouds: From Monocular Image Sequences to Synthetic Views. *Proceedings of the IEEE Intelligent Vehicles Symposium, San Diego, Kalifornien, USA* (2010): S. 436–442
- [Pag10d] PAGEL, Frank und WILLERSINN, Dieter: Motion-Based Online Calibration for Non-Overlapping Camera Views. *Proceedings of the IEEE Intelligent Transportation Systems Conference, Madeira, Spanien* (2010): S. 843–848
- [Pag11a] PAGEL, Frank und WILLERSINN, Dieter: Extrinsic Camera Calibration in Vehicles with Explicit Ground Estimation. *Proceedings of the International Workshop on Intelligent Transportation, Hamburg* (2011)
- [Pag11b] PAGEL, Frank und WILLERSINN, Dieter: Kontinuierliche Kalibrierung und Fusion mobiler Kamerasytsteme mit nichtüberlappenden Sichten in Fahrzeugen. 8. *Workshop Fahrerassistenzsysteme, Walting* (2011): S. 95–104
- [Pag12] PAGEL, Frank: Motion Adjustment for Extrinsic Calibration of Cameras with Non-overlapping Views. *Proceedings of the Canadian Conference on Computer and Robot Vision, Toronto, Ontario, Kanada* (2012): S. 94–100
- [Rom09] ROMERA, Merce Mora und PAGEL, Frank: Monocular Motion Segmentation. *Proceedings of the International Workshop on Intelligent Transportation, Hamburg* (2009): S. 121–126
- [Woo07] WOOCK, Philipp; PAGEL, Frank; GRINBERG, Michael und WILLERSINN, Dieter: Odometry-Based Structure from Motion. *Proceedings of the IEEE Intelligent Vehicles Symposium, Istanbul, Türkei* (2007): S. 1112–1117
- [Woo09] WOOCK, Philipp und PAGEL, Frank: Analysis of a Kalman Filter-based Online Stereo Camera Calibration. *Proceedings of the International Workshop on Intelligent Transportation, Hamburg* (2009): S. 127–132