

Kalibrierung von Multi-Kamera-Systemen

**Kombinierte Schätzung von intrinsischem Abbildungsverhalten
der einzelnen Kameras und deren relativer Lage zueinander
ohne Erfordernis sich überlappender Sichtbereiche**

Zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

der Fakultät für Maschinenbau
Karlsruher Institut für Technologie (KIT)

genehmigte

Dissertation

von

DIPL.-ING. TOBIAS STRAUSS

Tag der mündlichen Prüfung: 30. Januar 2015
Hauptreferent: Prof. Dr.-Ing. C. Stiller
Korreferent: Prof. Dr.-Ing. S. Hinz

Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit am Institut für Mess- und Regelungstechnik des Karlsruher Instituts für Technologie (KIT). Ich danke Herrn Prof. Dr.-Ing. C. Stiller und der Hans L. Merkle-Stiftung für die vielfältige Unterstützung und die Freiheiten im Hinblick auf die Ausrichtung meiner Forschung, ohne die die vorliegende Arbeit so nicht entstanden wäre.

Des Weiteren danke ich Herrn Prof. Dr.-Ing. S. Hinz für sein Interesse an dieser Arbeit und die Übernahme des Koreferats.

Besonderer Dank gilt an dieser Stelle auch meinem Mentor Dr.-Ing. S. Simon der Robert Bosch GmbH, der mich immer wieder in meinem Tun bestätigt und durch das in mich gesetzte Vertrauen gestärkt hat. Bei der Umsetzung des in dieser Arbeit vorgestellten Kalibrierverfahrens in C++ hat mein Kollege Johannes Beck für mich sehr wertvolle Unterstützung geleistet. Mein Kollege Julius Ziegler hat bei der Entwicklung der in Kapitel 5 vorgestellten globalen Assoziation entscheidend mitgewirkt. Ihnen und meinen weiteren Kollegen danke ich für die zahlreichen Diskussionen und Anregungen, die mit zum Gelingen dieser Arbeit beigetragen haben.

Bedanken möchte ich mich auch beim Sekretariat des Instituts für die Unterstützung bei Verwaltungsaufgaben, bei den Werkstätten des Instituts für die Umsetzung verschiedenster Kalibrierbauten sowie bei Herrn Paal für die Gewährleistung eines reibungslosen IT-Betriebs.

Abschließend möchte ich mich noch bei meiner Mutter für die Arbeit des Korrekturlesens im Hinblick auf sprachliche Korrektheit bedanken.

Karlsruhe, im September 2014

Tobias Strauß

Kurzfassung

In dieser Arbeit wird ein Verfahren zur Kalibrierung von Multi-Kamera-Systemen vorgestellt. Die Kalibrierung umfasst hierbei die Beschreibung des geometrischen Abbildungsverhaltens der einzelnen Kameras sowie deren relativer Lage zueinander. Als Messung wird die Beobachtung eines zeitlich bewegten Kalibrierobjekts verwendet, welches aus einer starren Anordnung mehrerer Schachbrettmuster besteht. Die Kamera-Kalibrierung wird als vollständiger Bündelausgleich formuliert, sodass sowohl die zeitliche Konsistenz der für die Kalibrierung verwendeten Struktur als auch die des Multi-Kamera-Systems berücksichtigt werden kann.

Eine Besonderheit des Kalibrierverfahrens ist die voll-automatische Extraktion der für die Kalibrierung erforderlichen Beobachtungen von Schachbrettmustern und deren globale Assoziation anhand eines die Schachbrettmuster umgebenden Binärcodes. Hierbei werden auch nur teilweise sichtbare Schachbrettmuster an den Bildrändern erkannt. Dies ist wichtig, um eine gute Abdeckung des gesamten Kamerabildbereichs mit Beobachtungen gewährleisten zu können.

Eine weitere wichtige Neuerung stellt die Implementierung eines universellen Kameramodells dar, welches neben unterschiedlichen Abbildungsfunktionen eine blickwinkelabhängige Verschiebung des Projektionszentrums berücksichtigt. Damit können neben konventionellen Objektiven auch bislang nur schwer zu kalibrierende Weitwinkel-Objektive zuverlässig beschrieben werden.

Dank der globalen Assoziation der Kalibrierbeobachtungen und der Verwendung einer Kalibriersequenz mit zeitlicher Relativbewegung zwischen Kamera-System und Kalibrierobjekt, ermöglicht das vorgestellte Verfahren auch die Kalibrierung von Multi-Kamera-Systemen mit sich nicht überlappenden Sichtbereichen.

Für den Bündelausgleich werden verschiedene Kostenmaße vorgestellt, mit denen Abweichungen zwischen Modellprädiktion und Beobachtung beschrieben werden können. Die Kostenmaße unterscheiden sich insbesondere hinsichtlich Berechnungs-Effizienz und Konvergenzverhalten. Insgesamt entspricht die Formulierung des Bündelausgleichs einem nichtlinearen, dünnbesetzten Least-Squares-Problem, für dessen effiziente Lösung die C++ Bibliothek CERES verwendet wird.

Des Weiteren wird ein Maß für die Beobachtbarkeit des Optimums als Beurteilung der Robustheit von Kalibrierungen vorgeschlagen. Ergebnisse einzelner Kalibrierungen und Anwendungen runden die Arbeit ab.

Schlagnworte: Kalibrierung – Kameramodell – Nicht überlappende Sichtbereiche

Abstract

This thesis presents a method for calibrating multi camera setups. The calibration comprises the geometric imaging characteristic of the individual cameras and their relative position to each other. The temporal movement of a calibration target consisting of a fixed arrangement of multiple checkerboard patterns is captured and used as measurement. The process of camera calibration is formulated as a full bundle adjustment problem. This allows for the consideration of temporal consistency regarding the calibration target and the multi camera rig.

A unique feature of the calibration framework is its full automated extraction of checkerboard patterns which are going to be globally associated using a binary code surrounding the checkerboard patterns. The extraction process works for partially visible checkerboard patterns at the image boundaries as well. This is important to achieve a good measurement coverage of the full imaging area.

A further improvement is the implementation of a universal camera model. Besides different projection functions this camera model can describe non-single viewpoint cameras as well. This allows for the calibration of conventional lenses as well as wide-angle lenses that are very difficult to calibrate up to now.

Due to the global association of the calibration markers and the use of a calibration sequence with temporal relative movement between camera system and calibration target, the presented calibration approach even allows calibrating multi camera systems with non-overlapping fields of view.

Several cost functions which describe the deviation between measurement predictions and observations are presented to be used in the bundle adjustment. The various cost functions differ mostly in calculation efficiency and convergence properties. In total, the formulation of the bundle adjustment problem corresponds to a sparse nonlinear least squares problem which is solved efficiently using the open source C++ library CERES.

Furthermore, a measure of the observability of the optimum is proposed as a benchmark for robustness of the calibration. Exemplary calibration results and applications complete the thesis.

Keywords: Calibration – Camera Model – Non-overlapping Fields of View

Inhaltsverzeichnis

Symbolverzeichnis	IX
1 Einleitung	1
1.1 Gegenstand der Arbeit	1
1.2 Aufbau der Arbeit	3
2 Stand der Technik	5
3 Kameramodell	9
3.1 Festlegung des Kamera-Koordinatensystems	9
3.2 Abbildungsfunktion	10
3.3 Verzeichnungsmodell	11
3.4 Verschiebung des Projektionszentrums	14
3.5 Universelles Kameramodell	15
4 Kamera-Kalibrierung als Bündelausgleich	17
4.1 Allgemeine Formulierung	18
4.2 Modellierung der Kalibrierstruktur	19
4.3 Umrechnung ins Kamera-Koordinatensystem	20
4.4 Kostenfunktion	21
4.5 Nicht überlappende Sichtbereiche	25
4.6 Implementierung	25
5 Extraktion global assoziierter Schachbrettmuster	29
5.1 Detektion von Schachbrettecken	30
5.2 Subpixel-Verfeinerung der Eckenposition	31
5.3 Schachbrettecken-Deskriptor	32

5.4	Vernetzung der Schachbrettecken	33
5.5	Globale Assoziation mittels Binärcode	35
6	Ablauf der Kalibrierung	39
6.1	Initialisierung der Schachbrettposen	41
6.2	Kamera-Kalibrierung ohne globale Assoziation	44
6.3	Globale Assoziation der Schachbrettmuster	44
6.4	Reduktion der Posen-Beschreibung	45
6.5	Multi-Kamera-Kalibrierung mit global assoziierten Beobachtungen	49
7	Beurteilung der Kalibrierung	51
7.1	Interpretation des Rückprojektionsfehlers	52
7.2	Beobachtbarkeit der Abbildung	53
7.3	Beurteilung von Kalibrierkonfigurationen	58
8	Anwendung der Kalibrierung	59
8.1	Vereinfachende Zentralprojektion	60
8.2	Bildentzerrung	61
8.3	Erzeugung virtueller Vogelperspektiven	64
8.4	Rektifizierung nach dem Lochkameramodell	65
8.5	Allgemeine Rektifizierung	66
9	Zusammenfassung und Ausblick	71
	Literaturverzeichnis	73

Symbolverzeichnis

Notationsvereinbarungen

Skalare	nicht fett, kursiv: a, b, c, \dots
Vektoren	fett, nicht kursiv, klein: $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$
Matrizen	fett, nicht kursiv, groß: $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$
Funktionen	kalligraphisch, groß: $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$

Symbole

$\mathbf{x}^T, \mathbf{A}^T$	Transposition eines Vektors bzw. einer Matrix
\mathbf{A}^{-1}	Inverse einer Matrix
$\mathbf{a} \cdot \mathbf{b}$	Skalarprodukt der Vektoren \mathbf{a} und \mathbf{b}
$\mathbf{a} \times \mathbf{b}$	Kreuzprodukt der Vektoren \mathbf{a} und \mathbf{b}
\sim	Proportionalität
$\stackrel{!}{=}$	geforderte Gleichheit
$:=$	Definition
$\sum a_i$	Summe über die Elemente a_i
$ \cdot $	Betrag einer Zahl, Mächtigkeit einer Menge
$\ \cdot\ $	euklidische Norm
$\arg \cdot$	Argument einer Funktion
$[a, b]$	Intervall von a bis b
\in	Element eine Menge bzw. eines Intervalls
\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{R}	Menge der reellen Zahlen
ϵ	kleine Konstante zur Verbesserung der numerischen Stabilität
\mathbf{I}	Einheitsmatrix
$\mathbf{0}$	Nullvektor

Bezeichner

Kameramodell

u, v	Spalten- bzw. Zeilenkoordinaten im Kamerabild
u_0, v_0	Spalten- bzw. Zeilenkoordinate des Kamerabild-Hauptpunkts
u', v'	Bildkoordinaten im Hauptpunkt-Koordinatensystem (verzeichnet)
u_n, v_n	normierte Bildkoordinaten im Hauptpunkt-KS (verzeichnungsfrei)
$\Delta u, \Delta v$	Verschiebung zwischen idealen und verzeichneten Koordinaten
f	Brennweite / Kammerkonstante der Kamera
r	euklidischer Abstand eines Bildpunktes zum Kamerabild-Hauptpunkt
r_n	brennweiten-normierter Abstand zum Kamerabild-Hauptpunkt
θ	Zenitwinkel (Winkel zwischen Sichtstrahl und optischer Achse)
φ	Azimutwinkel
$\mathcal{A}(\theta)$	Abbildungsfunktion ($\theta \rightarrow r_n$)
\mathbf{b}	Bildpunkt im Kamerabild
\mathbf{x}^K	Weltpunkt im Kamera-Koordinatensystem
\mathbf{p}_I	Vektor der intrinsischen Kameraparameter
$\mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$	Projektionsfunktion (Weltpunkt \rightarrow Bildpunkt)
$\mathcal{P}^{-1}(\mathbf{b}, \mathbf{p}_I)$	Inverse Projektionsfunktion (Bildpunkt \rightarrow Sichtstrahl)
\mathbf{z}	Aufpunkt des Sichtstrahls
\mathbf{r}	Richtungsvektor des Sichtstrahls

Bündelausgleich

\mathcal{D}_i	Datenterm (Abweichung zwischen Modellprädiktion und Messung)
\mathcal{R}_i	Regularisierungsterm für die geschätzten Parameter
$\mathbf{R}_S, \mathbf{t}_S$	Schachbrettmuster-Pose im Struktur-Koordinatensystem
$\mathbf{R}_R, \mathbf{t}_R$	Pose des Kamerasystems (Rig) im Struktur-KS
$\mathbf{R}_E, \mathbf{t}_E$	Kamerapose im Koordinatensystem des Kameraverbunds
$\mathbf{R}_K, \mathbf{t}_K$	Kamerapose im Struktur-KS
$\mathbf{R}_B, \mathbf{t}_B$	Schachbrettmuster-Pose im Kamera-KS
\mathbf{x}^B	Weltpunkt im Schachbrettmuster-KS
\mathbf{x}^S	Weltpunkt im Struktur-KS
\mathbf{x}^K	Weltpunkt im Kamera-KS
\mathbf{e}	Kosten-/Residuenvektor

Schachbrettmuster-Extraktion und globale Assoziation

W	Fenster für die Merkmalsberechnung
s	relative Position eines Bildpunkts innerhalb des Merkmalsfensters
\mathcal{I}	Grauwertbild der Kamera
$\nabla\mathcal{I}$	vektorwertiges Gradientenbild der Kamera
\mathbf{g}_p	Grauwert-Gradient
$\bar{\mathcal{I}}$	mittlerer Grauwert im Merkmalsfenster
M_O	Merkmalsbild der radialen Orientierung der Grauwert-Gradienten
M_S	Merkmalsbild für die Ausprägung der Punktsymmetrie
$\tilde{\mathbf{R}}_B, \tilde{\mathbf{t}}_B, \tilde{\mathbf{x}}^B$	Größen im lokalen Schachbrett-KS
$\mathbf{R}_B, \mathbf{t}_B, \mathbf{x}^B$	Größen im global assoziierten Schachbrett-KS
$\mathbf{R}_G, \mathbf{t}_G$	Umrechnung zw. lokalem und global assoziiertem Schachbrett-KS

Beobachtbarkeit der Abbildung

\mathbf{p}_S	Vektor der sekundären Parameter (Kalibrierstruktur, Bewegung etc.)
$\mathcal{T}(\mathbf{q})$	Operator zur Koordinatensystem-Transformation
\mathcal{K}	Kostenfunktion für die Differenz zweier Abbildungsmodelle
\mathcal{C}	Anstieg des Rückprojektionsfehler nach Kompensation durch \mathbf{p}_S
\mathbf{J}	Jacobi-Matrix
\mathbf{C}	Matrix zur Beschreibung des Anstieg des Rückprojektionsfehler
\mathbf{K}	Matrix zur Beschreibung der Änderung des Abbildungsmodells
B	Maß für die Beobachtbarkeit des Kameramodells

Allgemeine Rektifizierung

γ	Winkel der Epipolarebene zur z -Achse
β	Winkel auf der Epipolarebene
ρ	euklidischer Abstand eines beobachteten Punktes
\mathcal{G}, \mathcal{H}	beliebige Funktion zur Beschreibung der Zielabbildung

1 Einleitung

Das Auge als optisches System ist unser wohl wichtigstes Sinnesorgan zur Erfassung der Umwelt. Die besondere Bedeutung der visuellen Wahrnehmung zeigt sich auch an Redewendungen wie *Etwas wie seinen Augapfel hüten* oder *Ein Bild sagt mehr als tausend Worte*.

Dank optischer Systeme ist es uns seit langem möglich, persönliche Erinnerungen oder geschichtliche Ereignisse in Form von Bildern und Filmen festzuhalten. Die enormen Fortschritte im Bereich der digitalen Bilderfassung und -verarbeitung in den letzten Jahrzehnten haben darüber hinaus dazu geführt, dass optische Systeme auch aus vielen industriellen Anwendungen kaum mehr wegzudenken sind - beispielsweise der Überwachung von Fertigungsprozessen, der Qualitätskontrolle oder im Bereich von Fahrerassistenzsystemen. Insbesondere für messtechnische Anwendungen ist die genaue Kenntnis des Abbildungsverhaltens von entscheidender Bedeutung, denn nur so können aus Beobachtungen im Kamerabild sinnvoll Rückschlüsse auf Eigenschaften in der realen Welt gezogen werden.

1.1 Gegenstand der Arbeit

Die vorliegende Arbeit befasst sich mit der Bestimmung des Abbildungsverhaltens von Multi-Kamera-Systemen, wobei der Begriff Multi-Kamera-System einen konstruktiven Verbund aus mehreren Kameras bezeichnet. Das Abbildungsverhalten der einzelnen Kamera wird primär durch das jeweils eingesetzte Objektiv und dessen Positionierung relativ zum Bildsensor der Kamera bestimmt.

Das Abbildungsverhalten kann unterteilt werden in das geometrische und das photometrische Abbildungsverhalten. Das geometrische Abbildungsverhalten beschreibt, wo Objekte der Szene im Kamerabild wiederzufinden sind, während das photometrische Abbildungsverhalten den Zusammenhang zwischen den Helligkeiten im Bild und dem von einer Oberfläche abgestrahlten Lichtstrom wiedergibt. Diese Arbeit widmet sich nur dem geometrischen Aspekt.

Bei der Abbildung mittels eines Objektivs mit endlicher Blende, im Gegensatz zu einer idealen Lochkamera, wird jedem Punkt im Kamerabild ein Bündel von Sichtstrahlen zugeordnet. Die geometrische Verteilung der Sichtstrahlen innerhalb eines solchen Strahlenbündels (konvergent, parallel, divergent) ist eine allgemeine

Beschreibung der Fokussierung. In dieser Arbeit wird vereinfachend das jeweils einem Bildpunkt zugeordnete Strahlenbündel durch einen einzelnen Sichtstrahl angenähert, der Aspekt der Fokussierung bleibt somit unberücksichtigt.

Ziel dieser Arbeit ist es, anhand von Beobachtungen der Welt in Form von Kamerabildern, eine möglichst genaue Beschreibung des geometrischen Abbildungsverhaltens zu gewinnen. Dieser Vorgang wird als Kamera-Kalibrierung bezeichnet. Die Herausforderung hierbei besteht darin, dass die Lage der Objekte relativ zur Kamera im Allgemeinen nicht bekannt ist. Daher muss im Zuge der Kalibrierung gleichzeitig sowohl das Abbildungsverhalten der einzelnen Kameras als auch die Geometrie der beobachteten Szene geschätzt werden.

Das in dieser Arbeit vorgestellte Verfahren nutzt für die Kalibrierung die Beobachtung eines relativ zum Kamerasystem bewegten Kalibrierobjekts. Das Kalibrierobjekt besteht hierbei aus der dreidimensionalen Anordnung mehrerer codierter Schachbrettmuster. Die zum Zwecke der Kalibrierung weit verbreiteten Schachbrettmuster zeichnen sich insbesondere dadurch aus, dass die Position der darin enthaltenen Schachbrettecken im Kamerabild sehr genau bestimmt werden kann. Die Verwendung codierter Muster ermöglicht zusätzlich eine sehr einfache Assoziation der beobachteten Muster sowohl zwischen verschiedenen Kameras als auch im zeitlichen Verlauf der Bildsequenz.

Unter der Annahme eines starren Kalibrierobjekts kann durch die Vielzahl unterschiedlicher Beobachtungsposen die Geometrie des Kalibrierobjekts sehr genau geschätzt werden. Dadurch sinken die Anforderungen an die Herstellungsgenauigkeit, sodass die normalerweise vorausgesetzte Planarität der Schachbrettmuster nur noch eine untergeordnete Rolle spielt.

Eine wichtige Grundlage des vorgestellten Verfahrens bildet ein neu entwickelter Ansatz zur effizienten Detektion und Assoziation von Schachbrettmustern, auch wenn diese nur teilweise sichtbar sind. Durch die Auswertung eines Binärcodes, welcher jedes Schachbrettmuster umgibt, können die Beobachtungen global assoziiert und Fehldetektionen erkannt werden. Für die Kalibrierung wird ein modulares Abbildungsmodell vorgestellt, welches auch für Weitwinkel-Objektive ohne eindeutiges Projektionszentrum geeignet ist. Durch die Kombination aus zeitlicher Bewegung und globaler Assoziation können auch Multi-Kamera-Systeme kalibriert werden, bei denen sich die Sichtbereiche der einzelnen Kameras nicht überlappen.

Als Anwendung der Kalibrierung wird die Umrechnung der Bilddaten eines Multi-Kamera-Systems in eine gemeinsame Vogelperspektive gezeigt. Für den Fall der Kalibrierung von Stereo-Kamerasystemen wird ein allgemeines Konzept der Bild-Rektifizierung vorgestellt. Dieses ermöglicht für Konfigurationen mit Weitwinkel-Objektiven eine bessere Wiedergabe des erfassten Bildbereichs sowie eine Verringerung der durch die Rektifizierung eingeführten Verzerrungen.

1.2 Aufbau der Arbeit

Kapitel 2 gibt zunächst einen Überblick über den Stand der Technik hinsichtlich bereits verfügbarer Kalibrieransätze. Hierbei werden insbesondere auch die Aspekte des verwendeten Kalibriermodells, der Markerextraktion sowie des Abbildungsmodells betrachtet.

In Kapitel 3 wird dann ein modular aufgebautes Kameramodell eingeführt, welches sich am Objektiv-Design orientiert. Da für den Objektiv-Entwurf meist Rotationssymmetrie bezüglich einer optischen Achse angenommen wird, kommt der Modellierung des rotationssymmetrischen Falls besondere Bedeutung zu. Ein wesentlicher Aspekt ist hierbei die Berücksichtigung der blickwinkelabhängigen Verschiebung des Projektionszentrums, die insbesondere bei Weitwinkel-Objektiven relevant ist. Fertigungsbedingte Montageungenauigkeiten werden in Form eines Verzeichnungsmodells abgebildet.

In Kapitel 4 wird der der Kalibrierung zugrunde liegende Bündelausgleich formuliert. Das Kalibrierobjekt wird hierbei als dreidimensionale Anordnung von näherungsweise planaren Schachbrettmustern modelliert, wobei Abweichungen von der Planarität in Form einer zusätzlichen 3D-Verschiebung der einzelnen Schachbrettecken berücksichtigt werden können. Eine wesentliche Rolle im Hinblick auf Konvergenz und benötigte Rechenzeit zur Lösung des Bündelausgleichs spielt die verwendete Kostenfunktion. Der häufig als Kostenfunktion verwendete Rückprojektionsfehler ist zwar aus messtechnischer Sicht optimal, jedoch je nach Kameramodell nicht geschlossen berechenbar. Ein weiteres Problem ist der auf den Sichtbereich der Kamera beschränkte Gültigkeitsbereich der Projektionsvorschrift. Als Abhilfe werden verschiedene Kostenfunktionen vorgestellt, für die das Kameramodell nur an der Stelle der beobachteten Markerposition ausgewertet werden muss. Am Ende des Kapitels wird auf einige Implementierungsaspekte eingegangen, die für die effiziente Lösung des Bündelausgleichs wichtig sind. Hierbei wird insbesondere berücksichtigt, dass der Bündelausgleich als ein dünn-besetztes nichtlineares quadratisches Problem formuliert ist.

In Kapitel 5 wird im Detail beschrieben, wie die für den Bündelausgleich erforderlichen Beobachtungen global assoziierter Schachbrettecken gewonnen werden. Der vorgestellte Ansatz gliedert sich in mehrere Schritte. Zunächst werden potentielle Schachbrettecken extrahiert und anhand eines Deskriptors zu Schachbrettmustern vernetzt. Anschließend werden die Beobachtungen durch Auswertung eines die einzelnen Schachbrettmuster umgebenden Binärcodes verifiziert und global assoziiert. Eine Besonderheit bei der Schachbrettmuster-Extraktion ist, dass auch nur teilweise sichtbare Schachbrettmuster am Bildrand erkannt werden. Dadurch wird die Abdeckung des Bildbereichs mit Messungen deutlich verbessert.

Nachdem in den Kapiteln 3, 4 und 5 die grundlegenden Modelle für den Bündelausgleich und das Vorgehen zur Markerextraktion erläutert wurden, widmet sich Kapitel 6 dem für die praktische Anwendung sehr wichtigen Aspekt der Initialisierung und Implementierung des nichtlinearen Optimierungsproblems. Hierbei wird der vollständige Bündelausgleich in mehreren Teilschritten berechnet. Zunächst wird ein vereinfachter Bündelausgleich für jede Kamera gerechnet, welcher keine globale Assoziation erfordert. Das Ergebnis dieser Einzel-Kamera-Kalibrierungen ist die Grundlage, um den für die globale Assoziation verwendeten Binärcode auslesen zu können, da hierzu vom Schachbrett-Koordinatensystem ins Bild-Koordinatensystem umgerechnet werden muss. Anhand der globalen Assoziation kann die überbestimmte Menge von kamera-relativen Schachbrett-Posen in einem einheitlichen Koordinatensystem zusammengefasst werden. Hieraus werden Initialisierungen für die Kalibrierstruktur, die Posen des Kamerasystems sowie für die extrinsischen Kameraposen gewonnen.

Die Bewertung von Kalibrierergebnissen im Hinblick auf deren Genauigkeit erfolgt bislang meist direkt anhand des Rückprojektionsfehlers. In Kapitel 7 wird jedoch gezeigt, dass der Rückprojektionsfehler vielmehr ein Maß für die Genauigkeit des Ecken-Detektors ist als für die Genauigkeit der Kalibrierung selbst. Für die Bewertung der Kalibrierung wird ein Maß für die Beobachtbarkeit des Kameramodells eingeführt. Dieses bringt zum Ausdruck, in welchem Maße Änderungen am Kameramodell in Form eines Anstiegs des Rückprojektionsfehlers sichtbar werden. Hierbei wird insbesondere berücksichtigt, dass Änderungen des Projektionsmodells oftmals durch Anpassung der Beobachtungsposen ausgeglichen werden können.

In Kapitel 8 werden abschließend einige Anwendungen der Kalibrierung betrachtet. Da viele Anwendungen in ihrer bisherigen Form ein eindeutiges Projektionszentrum je Kamera erfordern, wird gezeigt, wie ein allgemeines Kameramodell auf den Spezialfall einer Zentralprojektion vereinfacht werden kann. Ausgehend vom Spezialfall der Zentralprojektion wird beschrieben, wie Kamerabilder auf ein anderes Abbildungsmodell umgerechnet werden können. Anwendung findet eine solche Umrechnung auf ein anderes Abbildungsmodell beispielsweise bei der Rektifizierung zur Vereinfachung der Bestimmung von Bildpunkt-Korrespondenzen für die Stereo-Tiefenrekonstruktion. Ausgehend von der klassischen Rektifizierung nach dem Lochkameramodell wird eine allgemeine Formulierung der Rektifizierung eingeführt, welche insbesondere beim Einsatz von Weitwinkel-Objektiven gewisse Vorteile bietet. Ein weiteres Anwendungsbeispiel ist die Berechnung einer virtuellen Vogelperspektive anhand der einzelnen Kamerabilder eines Multi-Kamera-Aufbaus.

2 Stand der Technik

Um ein optisches System, bestehend aus Kamera und Objektiv, für messtechnische Aufgaben verwenden zu können, muss zunächst dessen Abbildungsverhalten ermittelt werden. Dieser Vorgang wird als Kalibrierung bezeichnet. Das Ergebnis der Kalibrierung ist im Allgemeinen ein parametrisches Modell, welches beschreibt, wie Punkte aus der dreidimensionalen Welt im Kamerabild abgebildet werden. Da das Abbildungsverhalten durch Fertigungs- und Montageungenauigkeiten, aber auch durch die Einstellungen von Fokussierung und Blende beeinflusst wird, muss jedes System individuell kalibriert werden. Bei Multi-Kamera-Systemen umfasst die Kalibrierung neben der Bestimmung des Abbildungsverhaltens der einzelnen Kameras (intrinsische Kalibrierung) auch die Bestimmung der relativen Anordnung der Kameras zueinander (extrinsische Kalibrierung).

In der Literatur gibt es eine Vielzahl unterschiedlicher Ansätze zur Kalibrierung, die sich anhand der verwendeten Beobachtungen in Gruppen unterteilen lassen. Der vordergründig einfachste, aber in der Realisierung aufwändigste Ansatz besteht in der Verwendung genau bekannter 2D- oder 3D-Kalibrierobjekte, deren Position im Raum exakt vermessen und daher als bekannt angenommen wird [Tsa87, Bar08]. Der in der Praxis wohl am häufigsten verwendete Ansatz beruht auf der Verwendung planarer Kalibriermuster bekannter Größe, deren Position im Raum nicht als bekannt vorausgesetzt, sondern im Zuge der Kalibrierung mitgeschätzt wird [Stu99, Zha00, Ern99]. Die letzte wichtige Gruppe von Kalibrierverfahren beruht auf der Verwendung beliebiger Punkt-Korrespondenzen aus einer statischen Szene. Da hierbei kein spezielles Kalibriermuster erforderlich ist, wird diese Form der Kamera-Kalibrierung auch als Selbst-Kalibrierung bezeichnet [Dan09, Pol99a, Fau92]. Darüber hinaus lassen sich Kameras auch anhand kollinearere Punkte im Raum [Zha04b] oder anhand der Fluchtpunkte [Cap90] kalibrieren. Ein guter Überblick über die verschiedenen Verfahren ist in [Zha04a] zu finden.

Die Selbst-Kalibrierung zeichnet sich insbesondere durch ihre Einfachheit in der Anwendung aus, führt jedoch im Allgemeinen zu weniger genauen Kalibrierungen, da aufgrund der freien 3D-Punkte sehr viele Freiheitsgrade geschätzt werden müssen [Zha04a]. Die Selbst-Kalibrierung bietet jedoch insbesondere bei einer zeitlichen Veränderung des Kamerasystems beispielsweise durch Schwingungen oder Wärmeausdehnung den Vorteil, dass im Betrieb laufend nachkalibriert werden kann. Hierbei ist insbesondere für Multi-Kamera-Systeme der Fall interessant, nur einen Teil der Parameter wie beispielsweise manche der extrinsischen Para-

meter online nachzukalibrieren. Die Kalibrierung mit speziellen Markern bietet im Gegensatz dazu den Vorteil, dass die Lage der Marker im Bild meist genauer bestimmt werden kann. Darüber hinaus kann durch die Verwendung bekannter Kalibrieremuster Vorwissen über die Skale eingebracht werden – z.B. anhand der Größe von Schachbrettkacheln. Ohne Einbringen von Vorwissen über die Skale ist keine vollständige Kalibrierung möglich, da eine gleichzeitige Skalierung von beobachteter Szene und Kameraaufbau keine Änderung der Projektion bewirkt.

Im Bereich der Kalibrierung von Stereo-Kamerasystemen hat sich gezeigt, dass es vorteilhaft ist, intrinsische und extrinsische Kalibrierung gemeinsam durchzuführen [Gei12]. Dennoch werden im Bereich von Multi-Kamera-Systemen mit sich nicht überlappenden Sichtbereichen der einzelnen Kameras häufig intrinsische und extrinsische Kalibrierung getrennt durchgeführt [Esq07, Kum08, Cli08, Car11].

Neben den für die Kalibrierung verwendeten Beobachtungen unterscheiden sich die Kalibrierverfahren auch im jeweils für die intrinsische Kalibrierung verwendeten Projektionsmodell. Für perspektivische Kameras wird in der Regel das Lochkammermodell verwendet. Bei den zunehmend verbreiteten Fischaugen-Objektiven und katadioptrischen Objektiven wird der Zusammenhang zwischen Blickwinkel und Bildradius meist auf eine andere Art beschrieben, beispielsweise als Projektion auf eine Kugel [Mei07], direkt als funktionaler Zusammenhang [Sca06] oder modellfrei [Mir13]. Andere schlagen vor, sich an bekannten Projektionsfunktionen zu orientieren [Fle95, Xio97, Bak02]. Gemeinsam ist all diesen Abbildungsmodellen, dass sie von einem optischen System mit eindeutigem Projektionszentrum ausgehen, für das es ausreichend ist, jedem Pixel im Kamerabild einen Raumwinkel zuzuweisen. Bislang wenig verbreitet sind Projektionsmodelle, die auch eine blickwinkelabhängige Verschiebung der Eintrittspupille berücksichtigen [Gen06].

Eng mit dem Projektionsmodell verknüpft ist das Verzeichnungsmodell. Das Verzeichnungsmodell beschreibt die Abweichung von einem idealen Projektionsmodell, wobei meist das Lochkammermodell als Referenz verwendet wird. Sehr weit verbreitet ist die Aufteilung in radiale und tangential Verzeichnung, wobei die beiden Anteile durch einen polynomialen Ansatz beschrieben werden [Wen92, Zha00]. Insbesondere die radiale Verzeichnung [Tsa87] spielt eine wichtige Rolle, da es im Objektiventwurf im Allgemeinen nicht möglich ist, ein gewünschtes Projektionsmodell exakt zu realisieren und mithilfe der radialen Verzeichnung diese Abweichung weitgehend beschrieben werden kann. Im Gegensatz dazu liegt die Ursache für tangential Verzeichnung in Fertigungs- und Montageungenauigkeiten. In manchen Ansätzen werden Projektionsmodell und (radiales) Verzeichnungsmodell zusammengefasst [Sca06]. Für das Verzeichnungsmodell wird oftmals versucht, die Basisfunktionen physikalisch zu motivieren [Wen92]. Alternative Ansätze konzentrieren sich auf die numerischen Eigenschaften der Basisfunktionen [Abr97].

Für Stereo-Kamerasysteme schließt sich an die Kalibrierung üblicherweise der Schritt der Rektifizierung an. Hierbei werden die Bilder der beiden Kameras so verzerrt, dass der entfernungsabhängige Parallaxe-Effekt zwischen linkem und rechtem Kamerabild nur zu einer Verschiebung entlang der Bildzeile führt [Fus00]. Dies ist sehr vorteilhaft für die schnelle Suche von Bildpunkt-Korrespondenzen. Meist wird für die Rektifizierung eine ideale Anordnung zweier virtueller Lochkameras gewählt. Andere Ansätze wählen das Projektionsmodell der virtuellen Kameras so, dass die durch die Rektifizierung eingeführten Verzerrungen möglichst gering ausfallen [Pol99b, Glu01]. Ebenso gibt es für Fischaugen-Stereo-Anordnungen [Abr05] oder trinokulare Anordnungen [Hei06] spezielle Rektifizierungsmodelle. Grundlegende Voraussetzung für die Bildrektifizierung ist ein eindeutiges Projektionszentrum des Ausgangsmodells, da nur so unabhängig von der Tiefenstruktur der Szene in ein anderes Projektionsmodell umgerechnet werden kann. Ist diese Voraussetzung verletzt, so kann das Modell zwar vorab durch ein Kameramodell mit eindeutigem Projektionszentrum approximiert werden [Sch14b], allerdings geht dieser Schritt mit einem gewissen Informationsverlust einher.

3 Kameramodell

Die grundlegende Aufgabe eines Kameramodells ist die Beschreibung einer projektiven Abbildung, d.h. die Beschreibung des Zusammenhangs zwischen Punkten im dreidimensionalen Raum und deren Projektion auf den Bildsensor der Kamera. Das Kameramodell umfasst also die Kombination von Kamera und Objektiv. Da dem mit Abstand größten Teil der kommerziell verfügbaren Objektive ein rotationssymmetrischer Entwurf zugrunde liegt, beschränken sich die Betrachtungen in dieser Arbeit, abgesehen von kleinen Abweichungen, auf rotationssymmetrische Modelle.

Das im Folgenden vorgestellte Kameramodell ist modular aufgebaut. Es besteht aus einer Kombination von Abbildungsfunktion, Verzeichnungsmodell und einem Modell für die blickwinkelabhängige Verschiebung des Projektionszentrums.

3.1 Festlegung des Kamera-Koordinatensystems

Für die Beschreibung einer projektiven Abbildung ist es erforderlich, ein Kamera-Koordinatensystem festzulegen. Normalerweise wird dieses so gewählt, dass der Ursprung des Koordinatensystems mit dem Projektionszentrum der Abbildung zusammenfällt. Die Drehlage des Koordinatensystems wird so definiert, dass der Bildhauptpunkt (u_0, v_0) die z -Achse abbildet und die Bildachsen u' und v' im Hauptpunkt die Orientierung der x - und y -Achse festlegen (siehe Abbildung 3.1).

Für die weiteren Betrachtungen wird im Bildhauptpunkt ein neues Koordinatensystem eingeführt und der Bildradius $r := \sqrt{u'^2 + v'^2}$ definiert. Im Falle einer rotationssymmetrischen Abbildung stellt der Bildhauptpunkt das Rotationszentrum dar. Ist der Bildsensor bezüglich der optischen Achse des Kameraobjektivs zentriert, so liegt der Bildhauptpunkt in der Mitte des Kamerabildes.

Bei Verletzung der Rotationssymmetrie aufgrund von Fertigungs- oder Montageungenauigkeiten ist die Lage des Bildhauptpunkts nicht mehr klar bestimmt. Ebenso zeichnet sich für eine Abbildung ohne eindeutiges Projektionszentrum kein Punkt für die Festlegung des Koordinatensystemursprungs besonders aus. Es ist allerdings auch nicht erforderlich, das Koordinatensystem vorab explizit festzulegen, da die Bestimmung der Lage des Bildhauptpunkts Bestandteil der Kalibrierung ist. Hieraus ergibt sich automatisch dasjenige Koordinatensystem, in dem die Abbildung am besten mit den freien Parametern des Kameramodells beschrieben werden kann.

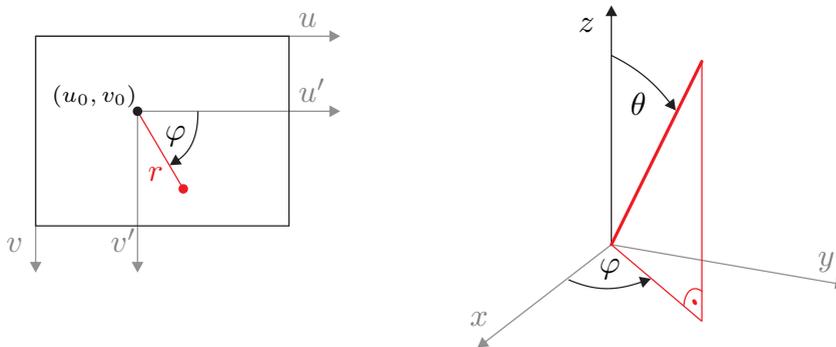


Abbildung 3.1: Festlegung des Kamera-Koordinatensystems. Die Abbildung verdeutlicht den Zusammenhang zwischen Kamerabild (links) und Kamera-Koordinatensystem (rechts). Die Orientierung des Kamera-Koordinatensystems ist so festgelegt, dass der Bildhauptpunkt (u_0, v_0) die z -Achse abbildet. Der rote Punkt im linken Bild korrespondiert mit dem rechts gezeigten Sichtstrahl.

Der Einfluss des verwendeten Koordinatensystems auf die parametrische Darstellung der Abbildung spielt insbesondere beim Vergleich verschiedener Kameramodelle oder Parametrisierungen eine Rolle (Abschnitt 7.2).

3.2 Abbildungsfunktion

Für rotationssymmetrische Optiken, wie sie in überwiegender Anzahl in der Praxis eingesetzt werden, existiert ein direkter Zusammenhang zwischen Bildradius r , Brennweite f und Einfallswinkel θ zwischen Sichtstrahl und optischer Achse. Dieser Zusammenhang wird mit dem normierten Bildradius $r_n = \frac{r}{f}$ über die Abbildungsfunktion $r_n = \mathcal{A}(\theta)$ beschrieben. Einige Abbildungsfunktionen zeichnen sich durch besondere Eigenschaften aus:

Verzeichnungsfrei:	$\mathcal{A}(\theta) = \tan \theta$
Winkeltreu:	$\mathcal{A}(\theta) = 2 \tan \frac{\theta}{2}$
Äquidistant:	$\mathcal{A}(\theta) = \theta$
Flächentreu:	$\mathcal{A}(\theta) = 2 \sin \frac{\theta}{2}$

Bei der verzeichnungsfreien Abbildung werden Geraden im Raum als Geraden abgebildet. Bei der winkeltreuen Abbildung werden die Gitternetzlinien einer Kugelfläche so abgebildet, dass die Schnittwinkel erhalten bleiben. Die Besonderheit der äquidistanten Abbildung ist der lineare Zusammenhang zwischen Bildradius und Blickwinkel. Die flächentreue Abbildung zeichnet sich dadurch aus, dass der

Flächeninhalt proportional zum abgebildeten Raumwinkel ist. Die Wirkung der unterschiedlichen Abbildungsfunktionen ist in Abbildung 3.2 veranschaulicht.

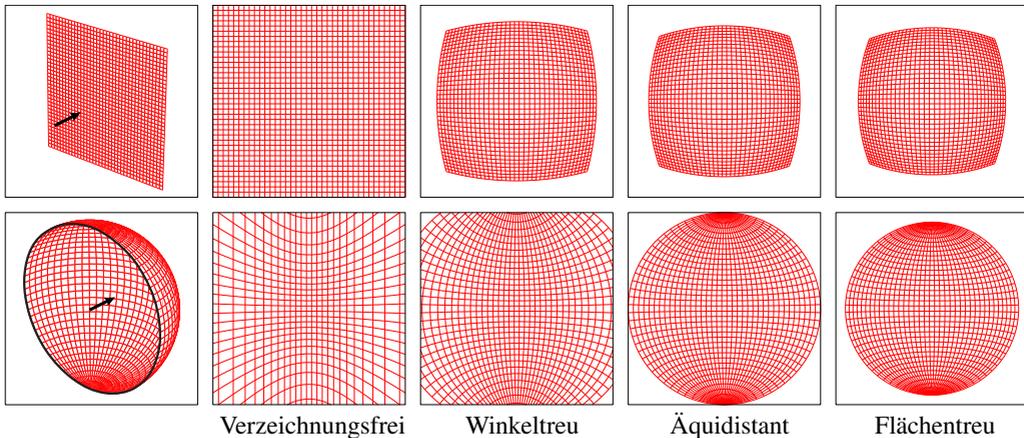


Abbildung 3.2: Veranschaulichung unterschiedlicher Abbildungsfunktionen. Auswirkungen der Abbildungsfunktion beim Blick auf eine Ebene (oben) sowie in eine Halbkugel (unten). Links ist die jeweilige Anordnung im dreidimensionalen Raum zu sehen, wobei der schwarze Pfeil die Blickrichtung der Kamera verdeutlicht. Rechts daneben ist jeweils für verschiedene Abbildungsfunktionen gezeigt, wie die Gitternetzlinien der Ebene bzw. Halbkugel abgebildet werden.

Im Objektiv-Entwurf wird üblicherweise versucht, je nach geplantem Einsatzzweck eine dieser Abbildungsfunktionen zu realisieren. Beim Entwurf müssen jedoch auch noch weitere Abbildungseigenschaften berücksichtigt werden, wie beispielsweise die Fokussierung oder eine geringe chromatische Aberration, auf die in dieser Arbeit nicht weiter eingegangen wird. Um die teilweise widersprüchlichen Entwurfsziele in Einklang bringen zu können, muss in der Regel eine Abweichung von der gewünschten Abbildungsfunktion in Kauf genommen werden. Diese Abweichungen können als radiale Verzeichnung beschrieben werden (Abschnitt 3.3).

3.3 Verzeichnungsmodell

Üblicherweise wird unter dem Begriff Verzeichnung die Abweichung einer Abbildung von der projektiven (verzeichnungsfreien) Abbildung einer Lochkamera verstanden. Diese Abweichung lässt sich als Verschiebungsfeld beschreiben, welches meist mithilfe von Polynomen dargestellt wird. Oftmals wird zwischen radialer und tangentialer Verzeichnung unterschieden, wobei die radiale Verzeichnung den rotationssymmetrischen Anteil des Verschiebungsfelds repräsentiert und die tangentiale Verzeichnung den nicht-rotationssymmetrischen Anteil.

Während die tangentielle Verzeichnung durch Fertigungs- und Montageungenauigkeiten bedingt ist, ist die radiale Verzeichnung auf eine von der projektiven Abbildung abweichende Abbildungsfunktion zurückzuführen. Insbesondere bei Weitwinkel-Objektiven führt daher die Beschreibung der Abweichung bezüglich des Lochkammermodells oftmals zu einer großen radialen Verzeichnung. Um diese dennoch hinreichend genau beschreiben zu können, wäre eine sehr hohe Ordnung der Polynome erforderlich, was jedoch die Extrapolationseigenschaften des Modells in der Regel negativ beeinflusst. Daher wird in dieser Arbeit die Verzeichnung etwas allgemeiner interpretiert und als Verschiebungsfeld bezüglich einer beliebig wählbaren Abbildungsfunktion definiert. Idealerweise wird als Referenz die Abbildungsfunktion aus dem Objektiv-Entwurf verwendet. Diese ist jedoch meist nur dem Hersteller des Objektivs bekannt. Da davon ausgegangen werden kann, dass beim Objektiv-Design versucht wurde, eine der Abbildungsfunktionen aus Abschnitt 3.2 zu realisieren, stellt deren Verwendung einen sinnvollen Ausweg dar.

Zur Beschreibung der Verzeichnung existieren unterschiedliche Ansätze, die zum Teil physikalisch und zum Teil mathematisch motiviert sind. Das in dieser Arbeit vorgestellte Verfahren ermöglicht prinzipiell die Verwendung einer beliebigen Modellierung der Verzeichnung. Im Folgenden wird zur Beschreibung der Verzeichnung der physikalisch motivierte Ansatz von Weng gewählt, welcher neben der radialen Verzeichnung auch Effekte durch Dezentrierung, Verkippung sowie Fertigungsungenauigkeiten der einzelnen Linsen modelliert [Wen92]. Alternativ sind beispielsweise auch ein direkter Monomansatz oder die Verwendung von Tschebyschow-Polynomen [Abr97] denkbar. Hierbei wird die Ursache der einzelnen Anteile außer Acht gelassen und der Fokus auf eine mathematisch günstige Modellierung gelegt. Da die tatsächliche Verzeichnung auch vom konkreten Anwendungsfall abhängig ist - beispielsweise beim Blick der Kamera durch eine gewölbte Windschutzscheibe - kann die Frage, welche Modellierung am besten geeignet ist, nicht pauschal beantwortet werden.

Bei der Wahl der Parametrisierung der Verzeichnung ist darauf zu achten, dass keine Mehrdeutigkeiten mit anderen Modellparametern wie Brennweite oder Bildhauptpunkt vorhanden sind, da dies für die spätere Optimierung nachteilig wäre. Zur Verbesserung der numerischen Genauigkeit ist es darüber hinaus sinnvoll, das Verzeichnungsmodell mit normierten Bildkoordinaten zu formulieren. Die Normierung kann hierbei zum Beispiel über die Bildgröße oder die Brennweite erfolgen.

Mit $r^2 = u'^2 + v'^2$ gilt nach dem Ansatz von Weng für die Verschiebungen eines Bildpunktes (u', v') bezüglich des Referenzmodells:

$$\Delta u = u'(k_1 r^2 + k_2 r^4) + q_1 r^2 + u'(p_1 u' + p_2 v') \quad (3.1)$$

$$\Delta v = v'(k_1 r^2 + k_2 r^4) + q_2 r^2 + v'(p_1 u' + p_2 v') \quad (3.2)$$

Die einzelnen Anteile des Verzeichnungsmodells sind in Abbildung 3.3 dargestellt.

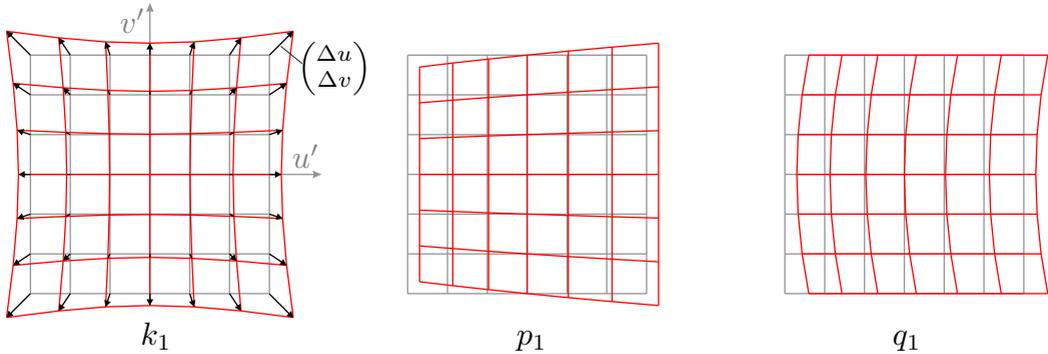


Abbildung 3.3: Basisfunktionen zur Beschreibung der Verzeichnung. Die Wirkung der einzelnen Parameter des Verzeichnungsmodells nach Gln. (3.1) und (3.2) lässt sich durch deren verzerrende Wirkung auf ein regelmäßiges Gitter veranschaulichen. Mit k_1 wird der quadratische Anteil der radialen Verzeichnung beschrieben, p_1 und q_1 beschreiben nicht-rotationssymmetrische Anteile. Die schwarzen Pfeile verdeutlichen die Verschiebungsvektoren zwischen dem unverzerrten Ausgangsgitter (grau) und dem verzerrten Zielgitter (rot). Je nach Festlegung wird das Ausgangsgitter entweder in realen oder idealen Bildkoordinaten definiert.

In der Literatur existieren unterschiedliche Festlegungen bezüglich der Richtung des Verschiebungsfelds. Meist wird dieses in Abhängigkeit von den idealen (verzeichnungsfreien) Koordinaten definiert – dies hat den Vorteil, dass unter Voraussetzung eines eindeutigen Projektionszentrums die Projektion eines Weltpunktes geschlossen berechnet werden kann. In anderen Quellen, wie auch in dieser Arbeit, wird das Verschiebungsfeld ausgehend von den realen Koordinaten beschrieben. Somit kann der zu einem Bildpunkt gehörende Sichtstrahl geschlossen berechnet werden. Motiviert ist diese Festlegung dadurch, dass für Abbildungsmodelle ohne eindeutiges Projektionszentrum die Projektion ohnehin nicht mehr geschlossen berechnet werden kann. Die Festlegung des Verschiebungsfelds in Abhängigkeit von den idealen Koordinaten würde in diesem Fall dazu führen, dass sowohl die Projektion von Weltpunkten als auch die zu einem Bildpunkt gehörenden Sichtstrahlen nur noch implizit definiert wären. Die beiden Varianten zur Beschreibung des Verschiebungsfelds sind bezüglich der Genauigkeit als gleichwertig zu erachten, wenn davon ausgegangen wird, dass die Verschiebungen klein sind. In diesem Fall unterscheiden sich die beiden Verschiebungsfelder primär im Vorzeichen der Verschiebungsrichtung. Geringfügige betragsmäßige Unterschiede der Verschiebungsfelder werden nahezu vollständig durch die Parameter-Optimierung kompensiert [Wen92].

3.4 Verschiebung des Projektionszentrums

In den allermeisten Anwendungen ist es wünschenswert, dass die Abbildung ein eindeutiges Projektionszentrum besitzt. Aus denselben Gründen wie bei der Umsetzung einer vorgegebenen Abbildungsfunktion ist es jedoch nicht ohne weiteres möglich, ein Objektiv mit eindeutigem Projektionszentrum zu konstruieren. Insbesondere bei weitwinkligen Fischaugen-Objektiven unterliegt das Projektionszentrum daher oftmals einer beachtlichen Verschiebung in Abhängigkeit vom Einfallswinkel in der Größenordnung von etwa einem Viertel des Objektiv-Durchmessers (siehe Abbildung 3.4).

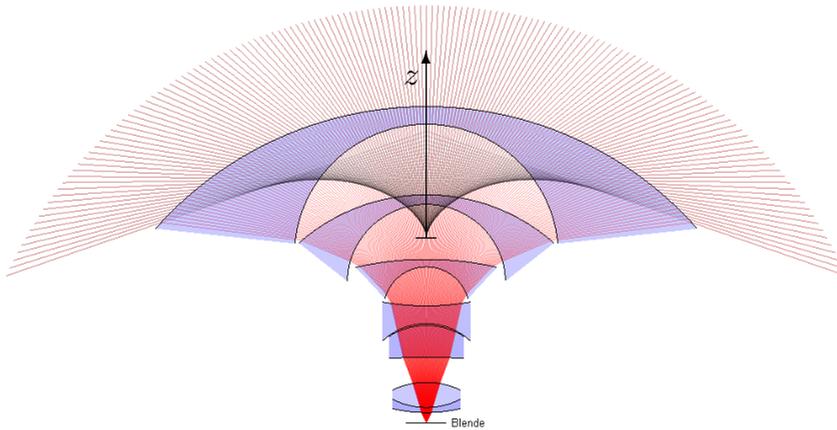


Abbildung 3.4: Verschiebung des Projektionszentrums in der Simulation. Die Strahlengangsimulation eines Fischaugen-Objektivs [Shi73] zeigt deutlich, wie sich das Projektionszentrum in Abhängigkeit vom Einfallswinkel nach oben verschiebt. Die einfallenden Sichtstrahlen (rot) sind zur Verdeutlichung bis zum Schnitt mit der optischen Achse nach innen verlängert (schwarz). Die einzelnen Linsenelemente des Objektivs sind blau eingefärbt. Die Simulation zeigt den Strahlengang bis zur Blende des Objektivs (unten im Bild). Der sensorseitige Teil des Objektivs ist für die hier betrachtete Verschiebung des Projektionszentrums irrelevant.

Zur Modellierung der Verschiebung des Projektionszentrums in Abhängigkeit vom Einfallswinkel wird in dieser Arbeit der Ansatz nach [Gen06] verwendet:

$$z(\theta) = \left(\frac{\theta}{\sin \theta} - 1 \right) \sum_{i=0}^N s_i \theta^{2i} \quad (3.3)$$

Der Ausdruck $\left(\frac{\theta}{\sin \theta} - 1 \right)$ beschreibt nach [Gen06] bereits für viele Objektive

relativ gut den qualitativen Verlauf der Verschiebung des Projektionszentrums, sodass zur Anpassung an ein bestimmtes Objektiv zusätzlich nur noch wenige Monom-Parameter s_i erforderlich sind. Werden alle Parameter s_i zu Null gewählt, so ergibt sich der Spezialfall einer Abbildung mit eindeutigem Projektionszentrum.

Mit der Modellierung der Verschiebung als Funktion $z(\theta)$ kann nur der rotationssymmetrische Fall exakt beschrieben werden, bei dem alle Sichtstrahlen die optische Achse schneiden. Für eine allgemeine Beschreibung kann als Aufpunkt eines Sichtstrahls der Punkt auf der Sichtgeraden mit dem kürzesten Abstand $|d|$ zur z -Achse gewählt werden. Mit den beiden Funktionen $d(\theta, \varphi)$ und $z(\theta, \varphi)$ gilt damit für den Aufpunkt:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -d(\theta, \varphi) \sin \varphi \\ d(\theta, \varphi) \cos \varphi \\ z(\theta, \varphi) \end{pmatrix} \quad (3.4)$$

Diese Beschreibung ist zwar allgemeingültig, jedoch sind sehr viele zusätzliche Parameter zur Beschreibung der Funktionen $d(\theta, \varphi)$ und $z(\theta, \varphi)$ erforderlich, um auch den nicht rotationssymmetrischen Fall abzubilden. Da in der Regel der Anteil der entwurfsbedingten Verschiebung des Projektionszentrums entlang der optischen Achse gegenüber zusätzlichen Verschiebungen bedingt durch Fertigungs- und Montageungenauigkeiten dominiert, wird in dieser Arbeit die vereinfachende rotationssymmetrische Modellierung der Verschiebung gemäß Gl. (3.3) verwendet.

3.5 Universelles Kameramodell

Nachdem in den vorhergehenden Abschnitten die einzelnen Bestandteile der Modellierung einer näherungsweise rotationssymmetrischen Projektion beschrieben wurden, werden diese nun zusammengesetzt. Das im Folgenden vorgestellte universelle Kameramodell ist explizit unterteilt in Abbildungsfunktion, Verzeichnungsmodell und ein Modell für die Verschiebung des Projektionszentrums. Das Kameramodell weist jedem Bildpunkt (u, v) im Kamerabild einen Sichtstrahl im Kamera-Koordinatensystem zu. Durch den modularen Aufbau kann die Beschreibung der Kameraabbildung künftig sehr einfach angepasst oder erweitert werden.

Für die Berechnung der Sichtstrahlen werden zunächst die Bildkoordinaten bezüglich des Bildhauptpunktes (u_0, v_0) zentriert und mit der Brennweite f normiert:

$$\begin{pmatrix} u' \\ v' \end{pmatrix} = \frac{1}{f} \begin{pmatrix} u - u_0 \\ v - v_0 \end{pmatrix} \quad (3.5)$$

Die Normierung mit der Brennweite dient zur Verbesserung der numerischen Eigenschaften des Verzeichnungsmodells. Ohne Normierung würden bei der Auswertung

der Polynome des Verzeichnungsmodells Zahlenwerte sehr unterschiedlicher Größenordnungen auftreten.

Über das Verzeichnungsmodell nach Gl. (3.1) und Gl. (3.2) wird die Verschiebung zwischen normierten, verzeichneten Koordinaten (u', v') und normierten, idealen Koordinaten (u_n, v_n) beschrieben. Für die idealen Bildkoordinaten gilt damit:

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} u' + \Delta u(u', v') \\ v' + \Delta v(u', v') \end{pmatrix} \quad (3.6)$$

Mit dem normierten, idealen Bildradius $r_n = \sqrt{u_n^2 + v_n^2}$ können durch Invertieren der Abbildungsfunktion $r_n = \mathcal{A}(\theta)$ die zugehörigen Blickwinkel ermittelt werden:

$$\theta(u, v) = \mathcal{A}^{-1} \left(\sqrt{u_n^2 + v_n^2} \right) \quad (3.7)$$

$$\varphi(u, v) = \arctan \left(\frac{v_n}{u_n} \right) \quad (3.8)$$

Das Modell für die Verschiebung des Projektionszentrums $z(\theta)$ nach Gl. (3.3) beschreibt abschließend die Position des Sichtstrahl-Aufpunktes, womit für den Sichtstrahl im Kamera-Koordinatensystem gilt:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ z(\theta) \end{pmatrix} + \lambda \begin{pmatrix} \cos \theta \cos \varphi \\ \cos \theta \sin \varphi \\ \sin \theta \end{pmatrix} \quad (3.9)$$

Mit den im Parametervektor \mathbf{p}_I zusammengefassten intrinsischen Parametern wird über das explizit definierte inverse Projektionsmodell

$$\mathcal{P}^{-1}(\mathbf{b}, \mathbf{p}_I) := \{\mathbf{z} + \lambda \mathbf{r}, \lambda \in \mathbb{R}^+\} \quad \text{mit} \quad \|\mathbf{r}\| = 1 \quad (3.10)$$

somit einem gegebenen Bildpunkt $\mathbf{b} = (u, v)^\top$ ein Sichtstrahl im Kamera-Koordinatensystem zugeordnet. Alle Punkte \mathbf{x}^K auf diesem Sichtstrahl werden gemäß dem Projektionsmodell \mathcal{P} auf den Bildpunkt $\mathbf{b} = \mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$ abgebildet. Der Bildpunkt \mathbf{b} kann jedoch im allgemeinen Fall nicht geschlossen ermittelt werden, da das Projektionsmodell $\mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$ nur implizit über das explizit formulierte inverse Projektionsmodell $\mathcal{P}^{-1}(\mathbf{b}, \mathbf{p}_I)$ definiert ist.

4 Kamera-Kalibrierung als Bündelausgleich

Das Ergebnis der Kamera-Kalibrierung umfasst die intrinsischen Parameter zur Beschreibung des Abbildungsverhaltens einer einzelnen Kamera sowie die extrinsischen Parameter zur Beschreibung der relativen Lage mehrerer Kameras zueinander. Dadurch wird jedem Bildpunkt ein Sichtstrahl im dreidimensionalen Raum zugeordnet. Da diese Zuordnung nicht direkt messbar ist, werden üblicherweise projektive Abbildungen in Form von Kamerabildern als Beobachtungen herangezogen. Meist ist hierbei jedoch die beobachtete Struktur oder zumindest deren Lage zur Kamera nicht exakt bekannt und muss im Zuge der Optimierung ebenfalls geschätzt werden. Die gemeinsame Optimierung von beobachteter Struktur sowie extrinsischen und intrinsischen Kamera-Parametern wird als Bündelausgleich bezeichnet.

Für die Beschreibung des Abbildungsverhaltens einer einzelnen Kamera wird in dieser Arbeit das in Abschnitt 3.5 beschriebene universelle Kameramodell verwendet. Bei der Multi-Kamera-Kalibrierung kann das Kameramodell für jede Kamera individuell angepasst werden – beispielsweise durch Wahl der jeweils verwendeten Abbildungsfunktion (Abschnitt 3.2) sowie darüber, ob eine Verschiebung des Projektionszentrums modelliert wird (Abschnitt 3.4).

Die beobachtete Struktur wird als Anordnung näherungsweise planarer Schachbrettmuster modelliert, deren Kachelgröße als bekannt vorausgesetzt wird (Abschnitt 4.2). Zur Verbesserung der Beobachtbarkeit wird für den Bündelausgleich eine Bildsequenz mit Relativbewegung zwischen Kamerasystem und Kalibrierstruktur verwendet. Um die Kalibrierstruktur im jeweiligen Kamera-Koordinatensystem beschreiben zu können, wird eine zeitlich veränderliche Beobachtungspose für das Kamerasystem eingeführt (Abschnitt 4.3).

Eine entscheidende Rolle für das Konvergenzverhalten des Optimierungsproblems spielt die Kostenfunktion, die bewertet, wie gut mit einem gegebenen Parametersatz die Beobachtungen erklärt werden können. In Abschnitt 4.4 werden hierfür verschiedene Möglichkeiten vorgestellt.

Bedingt durch die verwendete Modellierung zur Beschreibung der Struktur im jeweiligen Kamera-Koordinatensystem, können mit geeigneten Beobachtungsposen ohne weitere Anpassungen auch Kamerasysteme mit nicht überlappenden Sichtbereichen kalibriert werden (Abschnitt 4.5).

4.1 Allgemeine Formulierung

Im Allgemeinen kann das für die Lösung des Bündelausgleichs erforderliche Minimierungsproblem als eine Summe quadratischer Terme formuliert werden. Diese können zusätzlich unterteilt werden in Daten- und Regularisierungsterme:

$$\mathbf{p} = \arg \min_{\mathbf{p}'} \left(\sum_i \|\mathcal{D}_i(\mathbf{b}_i, \mathbf{p}')\|^2 + \sum_r \|\mathcal{R}_r(\mathbf{p}')\|^2 \right) \quad (4.1)$$

Der Parametervektor $\mathbf{p} = (\mathbf{p}_I, \mathbf{p}_S)$ ist zusammengesetzt aus dem Parametervektor \mathbf{p}_I zur Beschreibung des Abbildungsmodells und dem Parametervektor \mathbf{p}_S für die sekundären Parameter (Beschreibung der Kalibrierstruktur, Bewegung des Kameraaufbaus etc.). Über die Datenterme \mathcal{D}_i werden die Abweichungen zwischen Modellprädiktion und zugehöriger Beobachtung $\mathbf{b}_i = (u_i, v_i)^\top$ bewertet, während die Regularisierungsterme \mathcal{R}_r nur von den zu schätzenden Parametern abhängen und dazu dienen, die Konsistenz bzw. Glattheit der Parametrisierung zu verbessern. Im Rahmen dieser Arbeit wird der Regularisierungsterm nur verwendet, wenn die Modellierung nichtplanarer Schachbrettmuster verwendet wird. In diesem Fall werden die entsprechenden Verschiebungsparameter $\mathbf{p}_{\eta(r)}$ mit einer Gewichtung γ gegen Null reguliert, wobei mit $\eta(r)$ der Index der einzelnen Verschiebungsparameter beschrieben ist:

$$\mathcal{R}_r(\mathbf{p}) = \gamma \mathbf{p}_{\eta(r)} \quad (4.2)$$

In Abschnitt 4.4 werden für die Berechnung der Datenterme unterschiedliche Kostenfunktionen vorgestellt. Die Auswertung dieser Kostenfunktionen basiert auf der Darstellung der beobachteten Marker der Kalibrierstruktur im Kamera-Koordinatensystem $\mathbf{x}_i^K(\mathbf{p}_S)$. Wie diese berechnet werden, wird in den Abschnitten 4.2 und 4.3 gezeigt. Mit $\mathcal{P}(\mathbf{x}_i^K(\mathbf{p}_S), \mathbf{p}_I)$ wird die Projektion eines im Kamera-Koordinatensystem gegebenen Punktes beschrieben. Damit ergeben sich bei Verwendung des Rückprojektionsfehlers als Kostenfunktion die 2D-Residuenvektoren:

$$\mathcal{D}_i(\mathbf{b}_i, \mathbf{p}) := \mathbf{b}_i - \mathcal{P}(\mathbf{x}_i^K(\mathbf{p}_S), \mathbf{p}_I) \quad (4.3)$$

Die restlichen in Abschnitt 4.4 eingeführten Kostenfunktionen können zusammenfassend als vektorwertige Funktion \mathcal{F} in Abhängigkeit des zu einem Bildpunkt gehörenden Sichtstrahls $\mathcal{P}^{-1}(\mathbf{b}_i, \mathbf{p}_I)$ verallgemeinert werden:

$$\mathcal{D}_i(\mathbf{b}_i, \mathbf{p}) := \mathcal{F}[\mathcal{P}^{-1}(\mathbf{b}_i, \mathbf{p}_I), \mathbf{x}_i^K(\mathbf{p}_S)] \quad (4.4)$$

Die Verwendung der Formulierung nach Gl. (4.4) hat den Vorteil, dass das Projektionsmodell nur im Bildbereich der Kamera ausgewertet wird, während die Berechnung der Projektion nach Gl. (4.3) für Weltpunkte außerhalb des Kamera-Sichtbereichs zu undefinierten Ergebnissen führt.

4.2 Modellierung der Kalibrierstruktur

Für die Kalibrierung werden Beobachtungen planarer Schachbrettmuster verwendet. Schachbrettmuster zeichnen sich dadurch aus, dass die Lage der Kreuzungspunkte zwischen benachbarten Kacheln (im Folgenden als Schachbrettecken bezeichnet) im Kamerabild sehr genau ermittelt werden kann. Darüber hinaus erlaubt die Verwendung von punktförmigen Merkmalen, im Gegensatz zu Kurven oder Flächen, eine einfache Beschreibung der Korrespondenzen zwischen Bild und Struktur.

Oftmals wird ein einzelnes Schachbrettmuster relativ zur Kamera bewegt, um so die Anzahl der Beobachtungen zu erhöhen und um über das ganze Bild verteilt Beobachtungen zu erhalten. In [Gei12] wird als alternativer Ansatz eine Einzelbild-Kalibrierung vorgestellt, bei der mehrere Schachbrettmuster gleichzeitig beobachtet werden. Wird nicht zwischen zeitlicher Bewegung und räumlicher Anordnung unterschieden, so sind beide Ansätze prinzipiell identisch. Werden jedoch, wie für diese Arbeit, Abweichungen vom idealen Muster modelliert, so muss zwischen der räumlichen Struktur an sich und unterschiedlichen Beobachtungen dieser Struktur aufgrund von Relativbewegung unterschieden werden.

Um im Multi-Kamera-Fall die Lage der Kameras zueinander vollständig bestimmen zu können, ist zusätzlich Vorwissen über die Größe der Struktur erforderlich. Andernfalls sind die Verschiebungen zwischen den einzelnen Kameras nur bis auf einen gemeinsamen Skalierungsfaktor bestimmt. Für Kameramodelle ohne eindeutiges Projektionszentrum ist bereits die Einzel-Kamera-Kalibrierung ohne zusätzliches Vorwissen über die Strukturgröße unterbestimmt. Meist, wie auch in dieser Arbeit, wird die Größe der Schachbrettkacheln vorgegeben.

Für die folgenden Untersuchungen wird als Kalibrierstruktur eine statische, dreidimensionale Anordnung mehrerer nahezu planarer Schachbrettmuster verwendet. Abweichungen von der Planarität der Einzelmuster werden als zusätzliche Verschiebung $\Delta \mathbf{x}^B = (\Delta x^B, \Delta y^B, \Delta z^B)^T$ der einzelnen Schachbrettecken modelliert:

$$\mathbf{x}^B = \begin{pmatrix} x^B \\ y^B \\ 0 \end{pmatrix} + \begin{pmatrix} \Delta x^B \\ \Delta y^B \\ \Delta z^B \end{pmatrix} \quad (4.5)$$

Beschreiben \mathbf{R}_S und \mathbf{t}_S die dreidimensionale Lage des jeweiligen Schachbrettmusters im Koordinatensystem der Kalibrierstruktur, gilt für die Ecke eines Schachbrettmusters im globalen Struktur-Koordinatensystem:

$$\mathbf{x}^S = \mathbf{R}_S \mathbf{x}^B + \mathbf{t}_S \quad (4.6)$$

Die Verschiebungen $\Delta \mathbf{x}^B$ werden gemäß Gl. (4.2) gegen Null regularisiert, um die

Mehrdeutigkeit der Beschreibung aufzulösen, welche sich ansonsten in Kombinationen mit den Schachbrettposen ergeben würde. Durch die Regularisierung wird aus der Menge ansonsten äquivalenter Beschreibungen die Parametrisierung mit den geringsten Verschiebungen ausgewählt, sodass die über die Kachelgröße des planaren Grundmodells eingebrachte Skaleninformation erhalten bleibt.

An dieser Stelle sei erwähnt, dass die Regularisierung auch zu einem gewissen Bias in Richtung des planaren Grundmodells führt. Da die Regularisierung jedoch lediglich dazu dient, Mehrdeutigkeiten aufzulösen, kann durch Wahl eines betragsmäßig kleinen Gewichtungsfaktors γ in Gl. (4.2) der Bias vernachlässigbar klein gehalten werden.

4.3 Umrechnung ins Kamera-Koordinatensystem

Ein wesentlicher Bestandteil des Kalibrieransatzes ist die Verwendung einer Kalibriersequenz, bei der die Kalibrierstruktur relativ zum Kamerasystem bewegt wird. Um einen im Struktur-Koordinatensystem gegebenen Punkt \mathbf{x}^S mit der Projektion ins Kamerabild in Beziehung setzen zu können, muss dieser ins jeweilige Kamera-Koordinatensystem umgerechnet werden. Hierbei muss sowohl die Relativbewegung des Kamerasystems $(\mathbf{R}_R, \mathbf{t}_R)$ berücksichtigt werden als auch die relative Lage der jeweiligen Kamera im Koordinatensystem des Kameraverbunds $(\mathbf{R}_E, \mathbf{t}_E)$.

Unter Verwendung von homogenen Koordinaten gilt für einen Punkt im Kamera-Koordinatensystem \mathbf{x}^K :

$$\begin{pmatrix} \mathbf{x}^K \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_E & \mathbf{t}_E \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R}_R & \mathbf{t}_R \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R}_S & \mathbf{t}_S \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}^B \\ 1 \end{pmatrix} \quad (4.7)$$

Die Zerlegung der Koordinatensystem-Transformationen in ihre einzelnen Bestandteile ist elementar, um die zeitliche Konsistenz in Form einer Starrkörper-Annahme berücksichtigen zu können (siehe Abbildung 4.1).

Die Verwendung vieler unterschiedlicher Beobachtungsposen für die Kalibrierung ermöglicht zum einen eine robustere Kalibrierung (sofern die Starrkörper-Annahme zutreffend ist) und zum anderen eine angemessene Beobachtbarkeit nicht-planarer Schachbrettmuster. Darüber hinaus ist es unter Umständen erst durch Hinzunahme von Bewegung möglich, Kamerasysteme mit nicht überlappenden Sichtbereichen zu kalibrieren (Abschnitt 4.5).

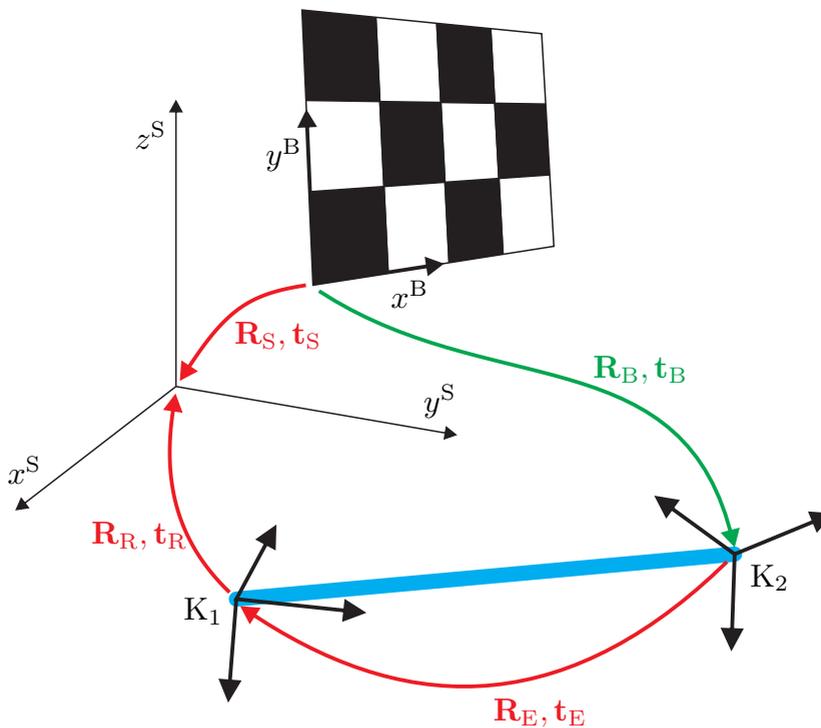


Abbildung 4.1: Umrechnung ins Kamera-Koordinatensystem. Die Abbildung veranschaulicht die verwendeten Koordinatensystem-Transformationen, um eine im Schachbrett-Koordinatensystem gegebene Schachbrettecke (x^B, y^B) im jeweilige Kamera-Koordinatensystem zu beschreiben. Die dicke blaue Linie zwischen den beiden Kamera-Koordinatensystemen stellt das Kamerasystem dar. Durch die Zerlegung der beobachteten Pose des Schachbrettmusters (grün) in mehrere Transformationsschritte (rot) kann berücksichtigt werden, dass sich Schachbrett- und Kameraposen zeitlich nicht verändern. Die Bewegung des Kamerasystems als Ganzes wird hierbei über die Kamera-System-Pose $(\mathbf{R}_R, \mathbf{t}_R)$ beschrieben. Im Beispiel dient das Kamera-Koordinatensystem K_1 als Referenz für die Pose des Kamerasystems.

4.4 Kostenfunktion

Für die Beurteilung einer Parametrisierung im Verlauf der Optimierung ist eine Kostenfunktion erforderlich. Diese Kostenfunktion sollte verschiedene Eigenschaften erfüllen. Elementar ist die Forderung, dass die Kosten bei korrekter Beschreibung der Abbildung zu Null verschwinden, also für den Fall idealer Beobachtungen. Im Hinblick auf das Konvergenzverhalten sollte die Funktion zusätzlich in einem möglichst großen Bereich um das Minimum herum konvex sein. Außerdem ist

es darüber hinaus wünschenswert, dass die Funktion schnell ausgewertet werden kann. Im Hinblick auf die Laufzeit der Optimierung entscheidend ist hierbei die Kombination aus Rechenzeit für eine einzelne Auswertung der Kostenfunktion und der erforderlichen Anzahl an Funktionsauswertungen. Letztere wird insbesondere von der Glattheit der Kostenfunktion beeinflusst. Das Ergebnis der Kostenfunktion ist der Kostenvektor

$$\mathbf{e}_i := \mathcal{D}_i(\mathbf{b}_i, \mathbf{p}) \quad (4.8)$$

Für das Konvergenzverhalten ist es vorteilhaft, bei der Optimierung direkt die einzelnen Komponenten des Kostenvektors zu verwenden und nicht nur dessen Norm. Dies lässt sich damit begründen, dass es einfacher ist, die Nullstellen der Vektorkomponenten zu finden als die Nullstelle der Betragsfunktion.

Im Folgenden werden nun verschiedene Kostenfunktionen vorgestellt. Diese verwenden zur Berechnung entweder die Projektion $\mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$ oder den zu einer Beobachtung gehörenden Sichtstrahl:

$$\mathcal{P}^{-1}(\mathbf{b}, \mathbf{p}_I) := \{\mathbf{z} + \lambda \mathbf{r}, \lambda \in \mathbb{R}^+\} \quad \text{mit} \quad \|\mathbf{r}\| = 1 \quad (4.9)$$

Die Verwendung des Sichtstrahls hat den Vorteil, dass die Projektion nur für tatsächliche Beobachtungen und damit Punkte innerhalb des Kamerabildes ausgewertet wird. Im Gegensatz dazu müssen bei der Berechnung der Projektion $\mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$ unter Umständen während der Optimierung Punkte außerhalb des Sichtbereichs der Kamera projiziert werden, für die die Projektion jedoch undefiniert ist.

Rückprojektionsfehler

Wird davon ausgegangen, dass der am Ende der Optimierung verbleibende Fehler ausschließlich auf einen normalverteilten und mittelwertfreien Fehler des Schachbrettecken-Detektors zurückzuführen ist, so ist im messtechnischen Sinne die Verwendung des Rückprojektionsfehlers als Kostenfunktion optimal. Hierzu wird der einer Beobachtung zugeordnete Punkt \mathbf{x}^K mithilfe des entsprechenden Kameramodells ins Kamerabild projiziert und dann der Fehler im Bildraum ermittelt:

$$\mathbf{e} = \begin{pmatrix} u \\ v \end{pmatrix} - \mathcal{P}(\mathbf{x}^K, \mathbf{p}_I) \quad (4.10)$$

Insbesondere für Kameras ohne eindeutiges Projektionszentrum kann die Projektion $\mathcal{P}(\mathbf{x}^K, \mathbf{p}_I)$ jedoch nicht geschlossen berechnet werden. Ein weiterer Nachteil der Verwendung des Rückprojektionsfehlers liegt darin, dass die Projektionsfunktion nur für Punkte innerhalb des Sichtbereichs der Kamera sinnvoll definiert ist. Dies kann insbesondere bei einer schlechten Initialisierung der Parameter oder bei Ausreißern ein Problem darstellen.

Abstand zwischen Weltpunkt und Sichtgerade

Eine Möglichkeit, um sicherzustellen, dass das Kameramodell nur innerhalb des gültigen Bereichs ausgewertet wird, ist die Verwendung des euklidischen Abstands des zu projizierenden Punkts \mathbf{x}^K von der zur Beobachtung gehörenden Sichtgerade. Als Kostenfunktion kann in diesem Fall der Lotvektor verwendet werden:

$$\mathbf{e} = (\mathbf{x}^K - \mathbf{z}) - (\mathbf{r}^\top (\mathbf{x}^K - \mathbf{z})) \mathbf{r} \quad (4.11)$$

Die Verwendung des Kreuzprodukts ermöglicht eine noch etwas kompaktere Darstellung des Kostenvektors, die sich zwar in der Richtung des Vektors unterscheidet, aber dieselbe Norm besitzt:

$$\tilde{\mathbf{e}} = \mathbf{r} \times (\mathbf{x}^K - \mathbf{z}) \quad (4.12)$$

Nachteil dieser beiden Kostenfunktionen ist, dass auch ein auf der falschen Seite der Kamera liegender Punkt zu verschwindenden Kosten führt. Dies ist auf die vereinfachende Verwendung der Sichtgerade anstelle des Sichtstrahls zurückzuführen. Darüber hinaus bewirkt die Verwendung des euklidischen Abstands eine unerwünschte Gewichtung des Fehlers mit dem Abstand des zu projizierenden Punktes \mathbf{x}^K vom Projektionszentrum. Durch die Verwendung des Blickwinkelfehlers können diese Probleme umgangen werden.

Blickwinkelfehler

Mit dem Richtungsvektor \mathbf{r} des Sichtstrahls und dem Richtungsvektor vom Projektionszentrum \mathbf{z} zum zu projizierenden Punkt \mathbf{x}^K wird ein weiterer Fehlervektor definiert:

$$\mathbf{e} = \frac{(\mathbf{x}^K - \mathbf{z})}{\|\mathbf{x}^K - \mathbf{z}\|} - \mathbf{r} \quad (4.13)$$

Der Betrag des Fehlervektors \mathbf{e} ist im gesamten Intervall $\alpha \in [0, \pi]$ eine monoton steigende Funktion in Abhängigkeit vom Blickwinkelfehler α :

$$\|\mathbf{e}\| = \sqrt{(1 - \cos \alpha)^2 + (\sin \alpha)^2} = \sqrt{2(1 - \cos \alpha)} \in [0, 2] \quad (4.14)$$

Die erste Ableitung $\frac{d}{d\alpha} \|\mathbf{e}\|$ ist eine monoton fallende Funktion im Intervall $\alpha \in [0, \pi]$ mit Steigung eins für $\alpha = 0$. Somit ist das Fehlermaß nach Gl. (4.13) eine geeignete Näherung für den Blickwinkelfehler α . Falls die abnehmende Gewichtung für größere Blickwinkelfehler nicht erwünscht ist, kann diese entsprechend

kompensiert werden:

$$\mathbf{e}' = \frac{\mathbf{e}}{\|\mathbf{e}\|} \arccos \left(1 - \frac{\|\mathbf{e}\|^2}{2} \right) \quad (4.15)$$

Es hat sich gezeigt, dass die direkte Verwendung des Blickwinkelfehlers zur Optimierung nicht geeignet ist, da hierbei die Brennweite gegen Unendlich konvergiert. Dies liegt daran, dass für $f \rightarrow \infty$ alle Richtungsvektoren \mathbf{r}_i gegen den Richtungsvektor der optischen Achse und somit für Schachbrettmuster, die in Richtung der optischen Achse unendlich weit entfernt sind die Blickwinkelfehler gegen Null konvergieren. Diese Konvergenz-Problematik ergibt sich daraus, dass auch die Lage der Schachbrettmuster während der Kalibrierung optimiert werden.

Linearisierter Rückprojektionsfehler

Eine Möglichkeit, dem Konvergenzproblem des Blickwinkelfehler-Kostenmaßes zu begegnen, ist die Berücksichtigung der lokalen Linearisierung des Blickwinkelfehlers im Punkt der Beobachtung (u, v) :

$$\mathbf{e} = \mathbf{e}_0 + \mathbf{e}_u \Delta u + \mathbf{e}_v \Delta v \quad (4.16)$$

Als Approximation für den Rückprojektionsfehler folgt damit:

$$\begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} = \arg \min_{\Delta u, \Delta v} \{ \mathbf{e}_0 + \mathbf{e}_u \Delta u + \mathbf{e}_v \Delta v \} \quad (4.17)$$

Durch die Dimensionsreduktion bei der Umrechnung auf den tangentialen Unterraum führt jedoch auch für dieses Kostenmaß ein auf der falschen Seite der Kamera liegender Punkt zu verschwindenden Kosten.

Gewichteter Blickwinkelfehler

Wird der Blickwinkelfehler nach Gl. (4.13) mit der Brennweite gewichtet, so kann das Konvergenzproblem mit $f \rightarrow \infty$ beseitigt werden, welches bei direkter Verwendung des Blickwinkelfehlers auftritt. Das Fehlermaß

$$\mathbf{e} = \left(\frac{(\mathbf{x}^K - \mathbf{z})}{\|\mathbf{x}^K - \mathbf{z}\|} - \mathbf{r} \right) f \quad (4.18)$$

entspricht einer groben Approximation des Rückprojektionsfehlers, die jedoch lokale Änderungen des Abbildungsmaßstabs nicht berücksichtigt. Im Gegensatz zum

linearisierten Rückprojektionsfehler ist für die Umrechnung ins Pixelmaß allerdings keine Dimensionsreduktion erforderlich, sodass der volle Konvergenzbereich erhalten bleibt. Punkte auf der falschen Seite des Sichtstrahls stellen hierbei Maxima dar. Aufgrund des optimalen Konvergenzverhaltens wird für den Bündelausgleich in dieser Arbeit der gewichtete Blickwinkelfehler nach Gl. (4.18) verwendet.

Um die messtechnische Optimalität sicherzustellen, kann im Anschluss an eine Optimierung mit dem gewichteten Blickwinkelfehler noch ein Optimierungsdurchlauf unter Verwendung des Rückprojektionsfehlers erfolgen (nach Eliminierung von Ausreißern). Auch ein kombiniertes Gütemaß aus linearisiertem Rückprojektionsfehler für $\alpha \rightarrow 0$ und gewichtetem Blickwinkelfehler für $\alpha \rightarrow \pi$ ist denkbar.

4.5 Nicht überlappende Sichtbereiche

Ist die Kalibrierstruktur bekannt, so ist es ausreichend, mit zwei Kameras zeitgleich jeweils einen Teil der Kalibrierstruktur zu beobachten, um die relative Lage der Kameras zueinander bestimmen zu können. Für die Bestimmung der Kalibrierstruktur, bestehend aus mehreren Schachbrettmustern, kann ausgenutzt werden, dass aus den Beobachtungen mehrerer Schachbrettmuster mit einer Kamera (oder mit Kameras, deren Anordnung zueinander bereits bekannt ist) die relative Lage dieser Schachbrettmuster zueinander bestimmt werden kann, sodass in der Regel, über die Kalibriersequenz hinweg, die gesamte Kalibrierstruktur bestimmt werden kann (siehe Abbildung 4.2).

Bei gegebener Kalibrierstruktur ist es unerheblich, ob sich die jeweils sichtbaren Bereiche von zwei Kameras überlappen, sodass ohne weiteres auch Kameras mit nicht überlappenden Sichtbereichen extrinsisch zueinander kalibriert werden können. Anschaulich sind die Sichtbereiche über die bekannte (aber nicht überlappend beobachtete) Kalibrierstruktur miteinander verbunden. Im Gegensatz zur Herstellung eines sehr großen verbindenden Musters, ist es im Allgemeinen einfacher, eine Anordnung aus mehreren kleinen Schachbrettmustern zu erzeugen und diese dann im Verlauf der Kalibrierung zu einem Verbund zusammenzufassen.

4.6 Implementierung

In der Regel werden die einzelnen Kostenterme in Gl. (4.1) jeweils nur von wenigen Komponenten des Parametervektors \mathbf{p} beeinflusst. Dies hat die wichtige Eigenschaft zur Folge, dass die während der Optimierung häufig zu berechnende Jacobi-Matrix

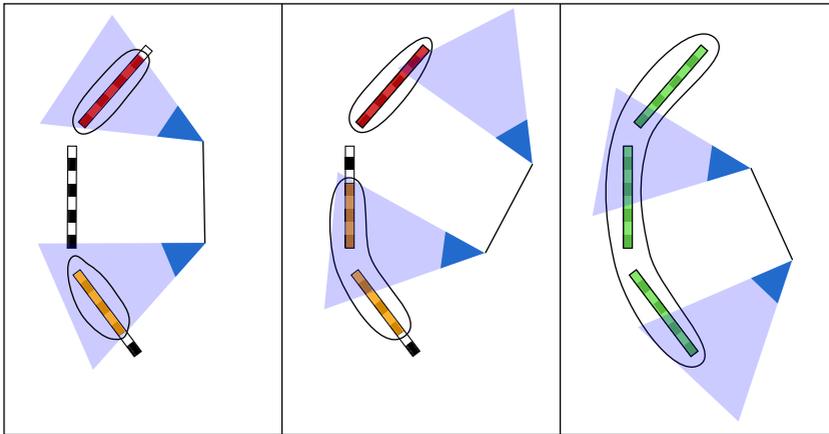


Abbildung 4.2: Kalibrierung mit nicht überlappenden Sichtbereichen. Schematische Darstellung einer Bildsequenz zur Kalibrierung zweier Kameras (blau) mit nicht überlappenden Sichtbereichen. Die Kalibrierstruktur besteht aus drei Schachbrettmustern, deren relative Lage zueinander vorab nicht bekannt ist. Ausgehend von der linken Aufnahmeposition, wird die Kalibrierstruktur in Form von zwei unabhängigen Teilstrukturen geschätzt (jeweils dargestellt in einer Farbe). Durch Hinzunahme der mittleren Aufnahmeposition können bereits zwei Schachbrettmuster zueinander registriert werden. Durch die rechte Aufnahmeposition können dann die beiden bis dahin unabhängigen Teile zusammengefügt werden. Durch Kombination der drei Aufnahmepositionen ist die Kalibrierstruktur zusammenhängend und eindeutig bestimmt. Die Darstellung als sequentieller Ablauf dient lediglich zur Verdeutlichung – für die Kalibrierung werden stets alle Aufnahmepositionen gemeinsam verarbeitet.

lediglich dünn besetzt ist. Auf diese Art von Ausgleichsproblemen ist die von Google entwickelte und frei verfügbare C++ Bibliothek CERES spezialisiert, auf die in dieser Arbeit zurückgegriffen wird. CERES stellt verschiedene Verfahren sowohl für die nichtlineare Optimierung als auch die unterlagerte Lösung der linearen quadratischen Probleme zur Auswahl. Der bekannteste Vertreter für die nichtlineare Optimierung ist der Levenberg-Marquardt Algorithmus [Lev44, Mar63]. Auf eine detaillierte Betrachtung der Verfahren zur Lösung großer nichtlinearer Optimierungsprobleme wird an dieser Stelle verzichtet. Ein Überblick über einige der von CERES verwendeten Techniken ist in [Aga10] zu finden.

Ein großer Vorteil von CERES ist die sehr komfortable Schnittstelle für die Problemformulierung, bei der lediglich die einzelnen Kostenvektoren in Abhängigkeit von den relevanten Parametern zur Verfügung gestellt werden müssen sowie eine Initialisierung dieser Parameter. Hierbei ist es möglich, für die einzelnen Kosten-

vektoren zusätzlich individuell robuste Gewichtsfunktionen zu wählen, die auf die Vektornorm angewendet werden. Auch Unterraum-Parametrisierungen oder das Festhalten einzelner Parameter während der Optimierung sind sehr einfach möglich. CERES kümmert sich intern um den Aufbau der zugehörigen dünn-besetzten Problemstruktur. Für die Berechnung der Jacobi-Matrizen werden verschiedene Techniken bereitgestellt: Die partiellen Ableitungen können analytisch vorgegeben, durch numerische Differentiation angenähert oder mit Hilfe von automatischem Differenzieren [Gri00] berechnet werden. Im Hinblick auf Wartbarkeit des Quellcodes, Konvergenzverhalten und Rechenzeit ist insbesondere das Verfahren der automatischen Differentiation hervorzuheben.

Nachdem in diesem Kapitel der für die Kalibrierung eingesetzte Bündelausgleich grundlegend beschrieben wurde, befassen sich die folgenden zwei Kapitel mit den praktischen Aspekten auf dem Weg zur Berechnung des Bündelausgleichs. Hierzu wird in Kapitel 5 zunächst beschrieben, wie die erforderlichen Beobachtungen von global assoziierten Schachbrettecken automatisiert gewonnen werden können. Neben der Bereitstellung der Beobachtungen spielt auch die Initialisierung der zu schätzenden Parameter eine wichtige Rolle, um in das globale Minimum zu konvergieren. Kapitel 6 beschreibt eine Möglichkeit, wie anhand von vereinfachten Formulierungen des Bündelausgleichs, welche ohne global assoziierte Beobachtungen auskommen und noch keine zeitliche Konsistenz berücksichtigen, Initialisierungen für die Parameter des vollen Bündelausgleichs gewonnen werden können.

5 Extraktion global assoziierter Schachbrettmuster

Für die Kamera-Kalibrierung werden in dieser Arbeit Beobachtungen einer Anordnung von näherungsweise planaren Schachbrettmustern in Form von Kamerabildern verwendet. Um im Verlauf der Optimierung nicht mehr auf die Kamerabilder zugreifen zu müssen, werden in einem Vorverarbeitungsschritt zunächst die Kalibriermuster aus den Bilddaten extrahiert. Da die Muster durch die Projektion ins Kamerabild einer unbekanntem Verzerrung unterliegen, ist es nicht ohne weiteres möglich, direkt die genaue Position der flächigen Strukturen im Bild zu beschreiben. Daher werden für die Kalibrierung anstelle der flächigen Schachbrettmuster nur die Kreuzungspunkte zwischen jeweils vier benachbarten Kacheln verwendet, welche auch als Schachbrettecken bezeichnet werden. Die Ecken zeichnen sich dadurch aus, dass sie im Bild vergleichsweise einfach und robust detektiert und genau lokalisiert werden können. Die Beobachtungen der Schachbrettecken können durch einfache Punkt-zu-Punkt-Beziehungen dargestellt werden:

$$(u, v) \leftrightarrow (x^B, y^B) \tag{5.1}$$

In diesem Kapitel wird der Vorgang beschrieben, wie aus Kamerabildern die darin zu sehenden Schachbrettmuster extrahiert werden können. Die Extraktion von Schachbrettmustern ist in mehrere Teilschritte unterteilt, die sequentiell ausgeführt werden. Zunächst wird, losgelöst von der Topologie eines Schachbretts, im ganzen Kamerabild nach Schachbrettecken gesucht (Abschnitt 5.1). Für diese Detektionen wird anschließend ein Deskriptor berechnet, der die jeweilige Ecke beschreibt (Abschnitt 5.3). Die Deskriptoren dienen dazu, die Anordnung der Ecken zu Schachbrettmustern wiederzufinden. Hierzu werden die Ecken anhand eines Vernetzungsalgorithmus zu Kacheln und schließlich zu Schachbrettmustern zusammengefasst (Abschnitt 5.4). Für die Vernetzung ist es hierbei unerheblich, ob ein Schachbrettmuster vollständig oder nur teilweise sichtbar ist.

Für den Bündelausgleich muss den einzelnen Beobachtungen im Kamerabild neben den Schachbrettkoordinaten (x^B, y^B) auch zugeordnet werden, welches der Schachbrettmuster beobachtet wurde. Außerdem ist es wichtig, eine einheitliche Indizierung der Schachbrettecken zu verwenden. Um dies gewährleisten zu können, werden die Beobachtungen anhand eines die einzelnen Schachbrettmuster umgebenden Binärcodes global assoziiert (Abschnitt 5.5).

5.1 Detektion von Schachbrettecken

Die Detektion von Schachbrettecken erfolgt anhand von zwei Merkmalsbildern M_O und M_S . Diese liefern ein Maß für die radiale Orientierung der Intensitätsgradienten bzw. die Symmetrie in einer lokalen Umgebung. Die Merkmalsbilder werden durch Bildfilterung berechnet, wobei für die lokale Umgebung jeweils die Pixel aus einem Fenster

$$W = \left\{ (du, dv)^T \in \mathbb{N}^2 \mid \sqrt{du^2 + dv^2} < R \right\} \quad (5.2)$$

verwendet werden, welches um das Zielpixel $\mathbf{b} = (u, v)^T$ herum zentriert wird.

Im Folgenden bezeichnet $\nabla \mathcal{I}$ den Gradienten des Intensitätsbilds \mathcal{I} und ϵ eine kleine Konstante, die zur Verbesserung der numerischen Stabilität eingeführt wird.

Zur Bewertung der Gradientenorientierungen wird berechnet, wie gut die Grauwertgradienten der lokalen Umgebung radial ausgerichtet sind.

$$M_O(\mathbf{b}) := \frac{\epsilon + \sum_{\mathbf{s} \in W} |\nabla \mathcal{I}(\mathbf{b} + \mathbf{s}) \cdot \mathbf{s}|}{\epsilon + \sum_{\mathbf{s} \in W} \|\nabla \mathcal{I}(\mathbf{b} + \mathbf{s})\| \|\mathbf{s}\|} \in (0,1] \quad (5.3)$$

Mit dem mittleren Grauwert des Fensters

$$\bar{\mathcal{I}}(\mathbf{b}) := \frac{1}{|W|} \sum_{\mathbf{s} \in W} \mathcal{I}(\mathbf{b} + \mathbf{s}) \quad (5.4)$$

gilt für das Symmetriemaß:

$$M_S(\mathbf{b}) := \frac{\epsilon + \frac{1}{2} \sum_{\mathbf{s} \in W} |\mathcal{I}(\mathbf{b} + \mathbf{s}) - \mathcal{I}(\mathbf{b} - \mathbf{s})|}{\epsilon + \sum_{\mathbf{s} \in W} |\mathcal{I}(\mathbf{b} + \mathbf{s}) - \bar{\mathcal{I}}(\mathbf{b})|} \in (0,1] \quad (5.5)$$

Die Menge der Schachbrettecken ergibt sich nun aus denjenigen Positionen, an denen das Merkmalsbild M_O ein lokales Minimum der 8er-Nachbarschaft besitzt und gleichzeitig vorgegebene Schwellwerte für beide Merkmalswerte unterschritten werden (siehe Abbildung 5.1). Durch Verwendung einer Bildpyramide, welche das Bild in verschiedenen Skalierungen enthält, kann zusätzlich die Erkennungsrate für unscharfe oder stark verrauschte Bilder verbessert werden.

Die sehr einfache und effiziente Variante zur Bestimmung von Schachbrettecken mittels zweier Merkmalsbilder hat keinen Anspruch auf Vollständigkeit, besitzt aber eine ausreichend gute Detektionsrate für die nachfolgende Vernetzung. Auch Fehldetektionen beeinflussen die Vernetzung im Allgemeinen nicht.

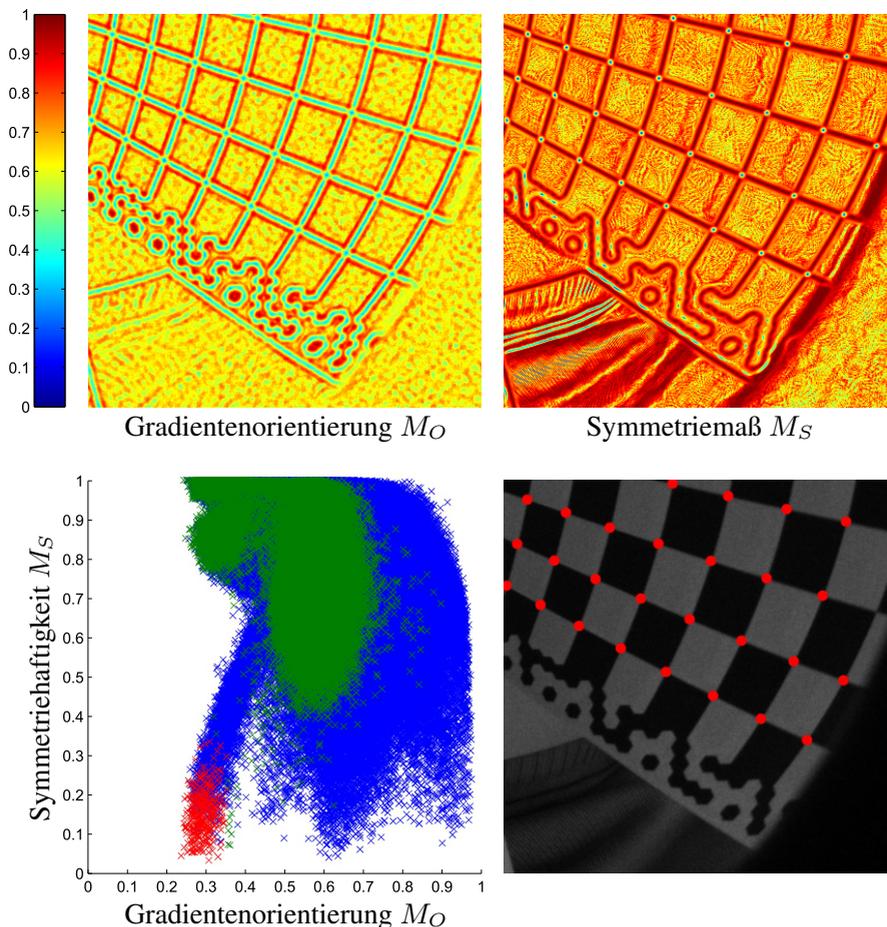


Abbildung 5.1: Merkmalsbasierte Eckendetektion. Die Detektion von Schachbrettecken erfolgt anhand der Merkmalsbilder für Gradientenorientierung (oben links) und Symmetriemaß (oben rechts). Für die Berechnung der Merkmalsbilder wurde ein Fenster mit einem Radius von $R = 5$ Px verwendet. Unten links sind die beiden verwendeten Merkmale gegeneinander aufgetragen. Lokale Minima der Gradientenorientierung sind rot bzw. grün eingefärbt, wobei sich die Farbe anhand der Klassifikation mittels Schwellwert-Filterung ergibt. Die restlichen Merkmalspunkte (keine lokalen Minima der Gradientenorientierung) sind blau eingefärbt. Das Bild unten rechts zeigt das zugehörige Kamerabild mit den mittels Schwellwert-Filterung detektierten Schachbrettecken (rot).

5.2 Subpixel-Verfeinerung der Eckenposition

Mittels einer Optimierung nach der Methode der kleinsten Fehlerquadrate wird die Lage der Detektionen auf Subpixel-Genauigkeit verfeinert. Hierzu wird für die

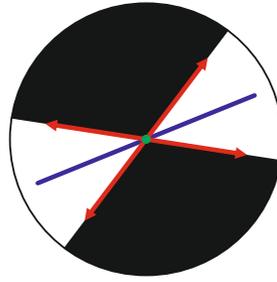


Abbildung 5.2: Schematische Darstellung des Schachbrettecken-Deskriptors. Der Schachbrettecken-Deskriptor beinhaltet die Schwarz-Weiß-Orientierung in Form der geschätzten Winkelhalbierenden der weißen Kacheln (blau) sowie die Richtungen der vier auslaufenden Kanten (rot).

optimal bestimmte Lage einer Schachbrettecke angenommen, dass die Gradienten tangential bezüglich dieses Punktes orientiert sind [Gei12].

Mit $\mathbf{g}_p := \nabla \mathcal{I}(\mathbf{b} + \mathbf{s})$ ergibt sich daraus das Optimierungsproblem

$$\Delta \mathbf{b} = \arg \min_{\Delta \mathbf{b}'} \sum_{\mathbf{s} \in W} (\mathbf{g}_p^\top (\mathbf{s} - \Delta \mathbf{b}'))^2 \quad (5.6)$$

Für die Subpixel-Verschiebung relativ zur Detektion folgt damit:

$$\Delta \mathbf{b} = \left(\sum_{\mathbf{s} \in W} \mathbf{g}_p \mathbf{g}_p^\top \right)^{-1} \sum_{\mathbf{s} \in W} \mathbf{g}_p \mathbf{g}_p^\top \mathbf{s} \quad (5.7)$$

5.3 Schachbrettecken-Deskriptor

Um die Schachbrettecken später richtig vernetzen zu können, werden diese über einen Deskriptor genauer beschrieben. Dieser beinhaltet die Schwarz-Weiß-Orientierung zur Beschreibung der Drehlage sowie die Richtungen der vier auslaufenden Kanten (siehe Abbildung 5.2).

Zur Beschreibung der Schwarz-Weiß-Orientierung wird die Orientierung der Hauptachse des Intensitätsbildes verwendet. Mit $\mathcal{I}_s := \mathcal{I}(\mathbf{b} + \mathbf{s})$ und $\mathbf{s} = (du, dv)^\top$ gilt hierfür:

$$o = \arctan \left(\frac{2 \sum_{\mathbf{s} \in W} \mathcal{I}_s du dv}{\sum_{\mathbf{s} \in W} \mathcal{I}_s dv^2 - \sum_{\mathbf{s} \in W} \mathcal{I}_s du^2} \right) \quad (5.8)$$

Anhand der Schwarz-Weiß-Orientierung kann die lokale Nachbarschaft in vier Quadranten W_i unterteilt werden, die jeweils eine der auslaufenden Kanten beinhalten. Getrennt für jeden Quadranten werden dann die Gradienten gemittelt, um

eine Schätzung für die Richtung der jeweils beinhalteten Kante zu erhalten. Hierbei wird eine zusätzliche Gewichtung der Gradienten mit dem Abstand zum Zentrum (Position der Ecke) durchgeführt:

$$\bar{\mathbf{g}}_i = \sum_{s \in W_i} \|s\| \mathbf{g}_p \quad (5.9)$$

Die Richtung der gesuchten Kante ist dann derjenige Richtungsvektor im untersuchten Quadranten, welcher senkrecht auf dem mittleren Gradientenvektor $\bar{\mathbf{g}}_i$ steht. Zur Verbesserung der Robustheit ist es zweckmäßig, gegenüberliegende Quadranten für die Schätzung der Kantenorientierung zusammenzufassen.

5.4 Vernetzung der Schachbrettecken

Um Fehldetektionen zu erkennen und um die Nachbarschaftsbeziehungen zwischen den Schachbrettecken abzubilden, werden die Schachbrettecken anhand der Ecken-Deskriptoren vernetzt. Diese Vernetzung ist in mehrere Teilschritte untergliedert. Zunächst werden für jede Ecke die vier Nachbarecken gesucht und zwischen jeweils paarweise benachbarten Ecken eine Kante eingeführt. Anschließend werden anhand der Kanten Schachbrettkacheln gebildet. Diese werden dann im letzten Schritt zu Schachbrettmustern zusammengefasst und indiziert. Im Folgenden werden nun die einzelnen Schritte ausführlicher beschrieben.

Suche benachbarter Schachbrettecken

Der erste Schritt der Vernetzung besteht darin, anhand der extrahierten Schachbrettecken Paare von Schachbrettecken zu bestimmen, die entlang der Kanten der Schachbrettmuster benachbart sind. Hierzu wird zunächst, basierend auf den Schachbrettecken-Deskriptoren (Abschnitt 5.3), für jede Ecke in einem schmalen Winkelbereich in der Umgebung der vier angrenzenden Kanten nach der jeweils nächstgelegenen Ecke gesucht. Für die Ermittlung der nächstgelegenen Ecke werden nur diejenigen Ecken berücksichtigt, deren Schwarz-Weiß-Orientierung zur Ausgangs-Ecke um etwa 90 Grad gedreht ist. Somit wird jeder Ecke für jede der vier Vernetzungsrichtungen, sofern vorhanden, eine Ecke zugewiesen. Anschließend wird überprüft, welche Ecken paarweise miteinander verbunden sind – Verbindungen in nur eine Richtung werden entfernt.

Extraktion von Schachbrettkacheln

Im zweiten Schritt der Vernetzung wird in der Menge der paarweise benachbarten Ecken nach 4er-Tupeln von Ecken gesucht, welche ringförmig benachbart sind. Bei der Suche kann ausgenutzt werden, dass die vier Nachbarn jeder Ecke in einem festen Umlaufsinn gespeichert sind. Dadurch kann, ausgehend von einer Ecke, direkt die nächste Ecke gemäß dem Umlaufsinn bestimmt werden. Kommt man nach vier Schritten wieder bei der Startecke an, so wurde eine Kachel gefunden. Unabhängig davon, ob nach vier Schritten eine Kachel gefunden wurde oder nicht, können die untersuchten (gerichteten) Kanten als abgearbeitet markiert werden, da sie nicht Teil einer (weiteren) Kachel sein können. Die Suche ist beendet, wenn alle Kanten markiert sind.

Im Zuge der Suche benachbarter Schachbrettecken werden teilweise auch Schachbrettecken mit Fehldetektionen außerhalb des Musters verbunden. Durch die Vernetzung zu Kacheln werden diese falschen Verbindungen in der Regel eliminiert.

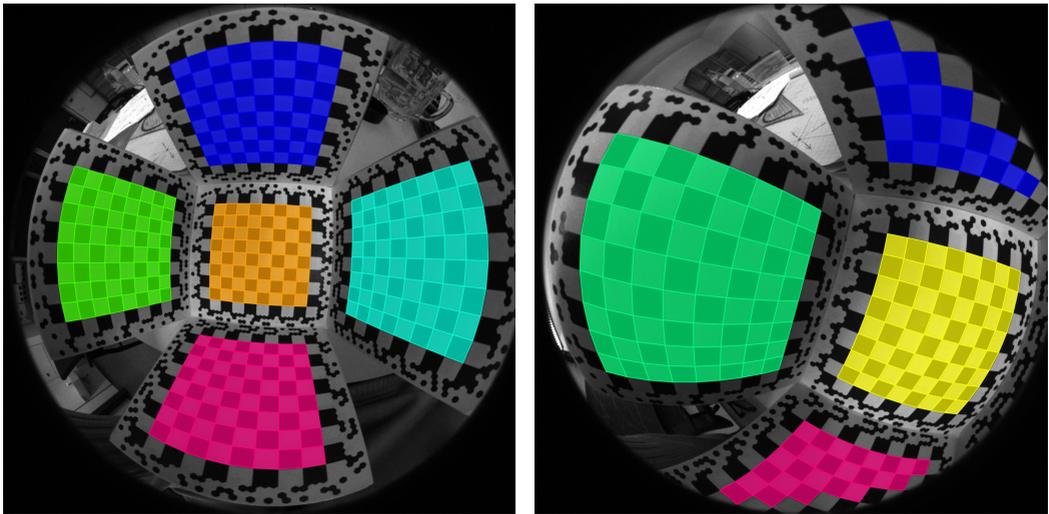


Abbildung 5.3: Extraktion von Schachbrettmustern. Anhand von Deskriptoren der einzelnen Schachbrettecken (Abschnitt 5.3) werden zunächst benachbarte Ecken durch Kanten verbunden. Anschließend wird in der Menge der Kanten nach ringförmig verbundenen Kanten gesucht, um diese zu Kacheln zusammenzufassen. Im letzten Schritt werden dann benachbarte Kacheln zu Schachbrettmustern vereint. Die Abbildung zeigt das Ergebnis der Schachbrettmuster-Bildung. Alle zu einem Schachbrettmuster zusammengefassten Kacheln sind in derselben Farbe dargestellt. Ohne globale Assoziation (Abschnitt 5.5) ist noch keine Zuordnung der Schachbrettmuster zwischen unterschiedlichen Bildern möglich.

Aufbau und Indizierung der Schachbrettmuster

Im letzten Schritt werden die gefundenen Kacheln zu Schachbrettmustern zusammengesetzt. Hierzu wird eine beliebige Kachel als Startkachel ausgewählt. Diese bildet den Ursprung für das Koordinatensystem des zugehörigen Schachbrettmusters. Anschließend wird über die Kanten der Kachel hinweg nach einer angrenzenden Kachel gesucht. Wird eine solche gefunden, so können anhand der gemeinsamen Kante direkt die Koordinaten aller Ecken der angrenzenden Kacheln ermittelt werden. Zum Schachbrettmuster hinzugefügte Kacheln werden entsprechend als abgearbeitet markiert. Die Prozedur wird entsprechend für die hinzugefügten Kacheln wiederholt, bis in der Menge der unmarkierten Kacheln keine angrenzende Kachel mehr gefunden wird. Solange die Menge der unmarkierten Kacheln nicht leer ist, wird daraus wieder eine beliebige Kachel als neue Startkachel ausgewählt, um das nächste Schachbrettmuster aufzubauen. Das Ergebnis der Schachbrettmuster-Bildung ist in Abbildung 5.3 zu sehen.

5.5 Globale Assoziation mittels Binärkode

Mit zu Schachbrettmustern vernetzten Ecken ist es bereits möglich, einzelne Kameras unabhängig voneinander intrinsisch zu kalibrieren und die Lage der Schachbrettmuster relativ zur jeweiligen Kamera zu bestimmen. Um jedoch die Anordnung mehrerer Kameras in einem Multi-Kamera-System bestimmen oder die zeitliche Konsistenz der Kalibrierstruktur berücksichtigen zu können, ist es erforderlich, Beobachtungen derselben Ecke einander zuzuordnen zu können. Prinzipiell ist ein Zusammenfassen von Beobachtungen derselben Ecke durch Anwenden des Iterative Closest Point (ICP) Algorithmus' [Bes92] möglich, indem die aus Einzelbildern geschätzten Teile der Kalibrierstruktur zueinander registriert werden. Insbesondere die Mehrdeutigkeiten bei periodischen Strukturen stellen für den ICP-Algorithmus allerdings ein Problem dar.

In dieser Arbeit wird daher ein alternativer Ansatz gewählt. Dieser basiert auf der Verwendung eines zusätzlichen Binärkodes, welcher eine globale Assoziation der Schachbrettecken ermöglicht [Str14]. Der Binärkode ist um die Schachbrettmuster herum verteilt und weist einen hohen Grad an Redundanz auf. Dadurch ist es möglich, auch teil-verdeckte Schachbrettmuster noch korrekt zu assoziieren. Der für die globale Assoziation verwendete Binärkode ist, wie auch das Schachbrettmuster selbst, in Kacheln unterteilt. Diese haben dieselben Abmessungen wie die Schachbrettkacheln und beinhalten jeweils 9 Codebits (siehe Abbildung 5.4). Die Form der Codebits wurde hexagonal gewählt, um Punkte zu vermeiden, welche die Symmetrieeigenschaften des Schachbrettecken-Detektors erfüllen.

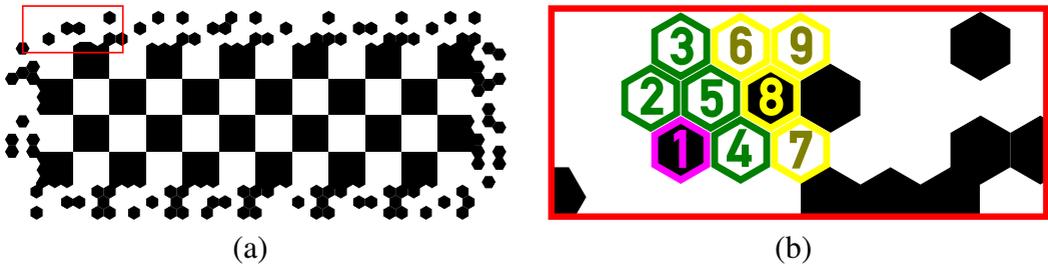


Abbildung 5.4: Binärcode für die globale Assoziation der Schachbrettecken. Für die globale Assoziation der Schachbrettecken ist jedes Schachbrettmuster mit einem zusätzlichen Binärcode umgeben. (a) Codiertes Schachbrettmuster. (b) Detailansicht von (a) mit dem Binärlayout einer einzelnen Code-Kachel. Die Code-Kacheln codieren abwechselnd Zeilen- und Spalteninformation (magentafarbenes Spalten-Bit). Die vier grünen Index-Bits repräsentieren den binär codierten Zeilen- bzw. Spaltenindex, welcher die relative Lage auf dem Schachbrettmuster wiedergibt. Die Kennzahl des Schachbrettmusters ist in zwei 4-Bit-Gruppen (gelb) zerlegt und auf Zeilen- und Spaltenkachel verteilt abgelegt.

Die Code-Kacheln codieren abwechselnd ihre Zeilen- bzw. Spaltenposition auf dem Schachbrettmuster sowie eine eindeutige Kennzahl für das Schachbrettmuster. Durch den Aufbau des Binärcodes genügen prinzipiell zwei korrekt gelesene, benachbarte Code-Kacheln, um alle Ecken dieses Schachbrettmusters global assoziieren zu können. Über die Lage der Code-Kacheln relativ zum untersuchten Schachbrettmuster kann die Positionsinformation jeder Code-Kachel in eine Schachbretthypothese umgerechnet werden. Die Schachbretthypothese umfasst hierbei die Kennzahl des Schachbrettmusters sowie die Transformation, wie das initiale Schachbrett-Koordinatensystem der Vernetzung in das globale Koordinatensystem des Schachbrettmusters umgerechnet werden kann (Abschnitt 6.3). Alle konsistent gelesenen Code-Kacheln führen hierbei zu derselben Schachbretthypothese, während falsch gelesene Code-Kacheln zu davon abweichenden Schachbretthypothesen führen. Als Ergebnis der Code-Auswertung wird diejenige Hypothese gewählt, für welche die meisten Code-Kacheln konsistent gelesen wurden. Um ein gewisses Maß an Robustheit zu gewährleisten, wird eine Mindestanzahl korrekt gelesener Code-Kacheln gefordert. Das Ergebnis der Assoziation zeigt Abbildung 5.5.

Um den Binärcode auszulesen, müssen die Positionen der einzelnen Codebits im Kamerabild ermittelt werden. Hierzu wird eine Einzel-Kamera-Kalibrierung der jeweiligen Kamera verwendet, welche ohne Berücksichtigung der globalen Assoziation auskommt. Diese liefert in der Regel eine für diesen Zweck hinreichend gute Schätzung der einzelnen Schachbrettposen sowie des Abbildungsmodells. Durch Extrapolation im Schachbrett-Koordinatensystem können dann die 3D-Positionen

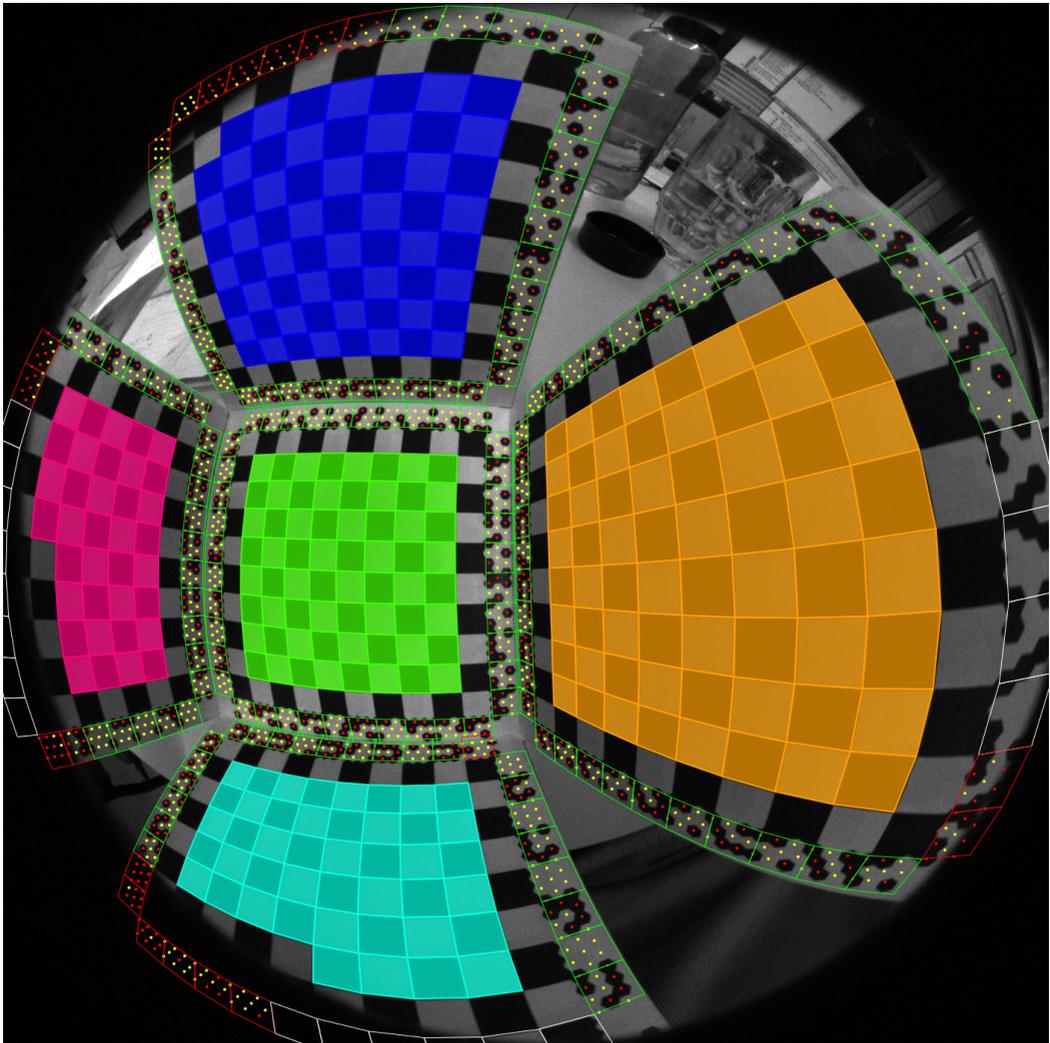


Abbildung 5.5: Globale Assoziation der Schachbrettmuster. Unter Verwendung einer vereinfachten Einzel-Kamera-Kalibrierung, ohne Berücksichtigung der globalen Assoziation, ist es möglich, beliebige Schachbrett-Koordinaten ins Bild zu projizieren. Dies wird genutzt, um die Positionen der Codebits im Kamerabild zu ermitteln. Falls ein Teil der projizierten Code-Kachel nicht mehr im Kamerabild zu liegen kommt, wird diese ignoriert (weiße Kachel). Ansonsten werden die Bits einer Code-Kachel anhand der Intensitätswerte der Bits dieser Kachel als schwarz oder weiß klassifiziert (rote bzw. gelbe Punkte). Liefert die Auswertung des Codes eine zur gewählten Schachbrettmuster-Hypothese konsistente Information, wird diese als gültig (grüne Kachel) oder ungültig (rote Kachel) markiert.

der potentiellen Codebits ermittelt und anhand des Abbildungsmodells ins Bild projiziert werden. Anhand der Intensitätswerte des Kamerabilds an den Positionen der Codebits werden für jede Code-Kachel unabhängig die Bits als schwarz oder weiß klassifiziert. Durch den Entwurf des Binärcodes kann hierbei angenommen werden, dass jede Kachel sowohl schwarze als auch weiße Bits enthält.

Die Lage der auszulesenden Code-Kacheln ist zunächst unbekannt. Es wird jedoch angenommen, dass zumindest ein Teil der Schachbrett-Randkacheln detektiert wurden. Die Position der Code-Kacheln ergibt sich damit aus dem nach außen erweiterten umschließenden Rechteck der detektierten Schachbrett-Kacheln. Fallen Rand-Kacheln der Vernetzung mit Rand-Kacheln eines Musters zusammen, so werden durch die Extrapolation tatsächliche Codebit-Positionen ermittelt. Andernfalls liefert die Extrapolation Positionen auf Schachbrett-Kacheln. Aufgrund der Redundanz des Codes stellt es jedoch kein Problem dar, wenn nur ein Teil der Code-Kacheln sichtbar ist und korrekt gelesen werden kann.

Als Ergebnis der globalen Assoziation sind alle Beobachtungen eindeutig Schachbrettecken der Kalibrierstruktur zugeordnet. Beobachtungen von Schachbrettmustern, die trotz der hohen Redundanz des Binärcodes nicht assoziiert werden können, werden für die Kalibrierung verworfen. Damit führt die Auswertung des Binärcodes neben der globalen Assoziation auch zur Erhöhung der Robustheit im Falle von falsch extrahierten Schachbrettmustern.

6 Ablauf der Kalibrierung

Das der Kalibrierung zugrundeliegende Optimierungsproblem wurde bereits in Kapitel 4 eingeführt. Die hierfür erforderlichen Messungen stellen global assoziierte Beobachtungen von Schachbrettecken dar (Kapitel 5). Das in Abschnitt 5.5 beschriebene Verfahren zur globalen Assoziation setzt bereits eine Kalibrierung der intrinsischen Kameraparameter sowie Kenntnis der einzelnen Schachbrettposen im Kamera-Koordinatensystem voraus. Aus diesem Grund ist der Ablauf der Kamera-Kalibrierung in mehrere Schritte unterteilt. Abbildung 6.1 gibt einen Überblick über die implementierte Ablaufstruktur. Zunächst werden vernetzte Schachbrettecken aus den Kamerabildern extrahiert, wobei für jedes Schachbrettmuster Ursprung und Orientierung des jeweiligen Koordinatensystems willkürlich gewählt werden. Unter Verwendung dieser nicht global assoziierten Beobachtungen werden zunächst alle Kameras einzeln kalibriert (Abschnitt 6.2), um damit die Voraussetzungen für die globale Assoziation zu schaffen.

Für die der globalen Assoziation vorausgehende Einzel-Kamera-Kalibrierung (Abschnitt 6.2) müssen neben den intrinsischen Kameraparametern auch die Schachbrettposen initialisiert werden. Für die intrinsischen Kameraparameter werden hierzu manuell Startwerte vorgegeben, wobei eine ungefähre Angabe der Brennweite und Annahme des Hauptpunkts in der Bildmitte im Allgemeinen bereits ausreichend ist. Die Initialisierung der Schachbrettposen erfolgt automatisch anhand der Beobachtungen (Abschnitt 6.1). Generell hat die Initialisierung der zu optimierenden Parameter einen nicht unerheblichen Einfluss auf die Laufzeit des Verfahrens und dessen Konvergenzverhalten. Eine triviale Null-Initialisierung der Parameter ist in der Regel nicht zielführend.

Um die Parameter des in Kapitel 4 beschriebenen vollständigen Optimierungsproblems zu initialisieren, können die Ergebnisse der Einzel-Kamera-Kalibrierung herangezogen werden. Hierbei werden die Posen-Parameter anhand der Information aus der globalen Assoziation in ein einheitliches Bezugssystem transformiert (Abschnitt 6.3) und anschließend fusioniert (Abschnitt 6.4), während die Kamera-Parameter direkt übernommen werden.

Für die Lösung der einzelnen Optimierungsprobleme wird die von Google entwickelte Software-Bibliothek CERES verwendet (Abschnitt 4.6). Diese stellt eine effiziente Implementierung für die Lösung nichtlinearer dünnbesetzter Least-Squares-Probleme zur Verfügung.

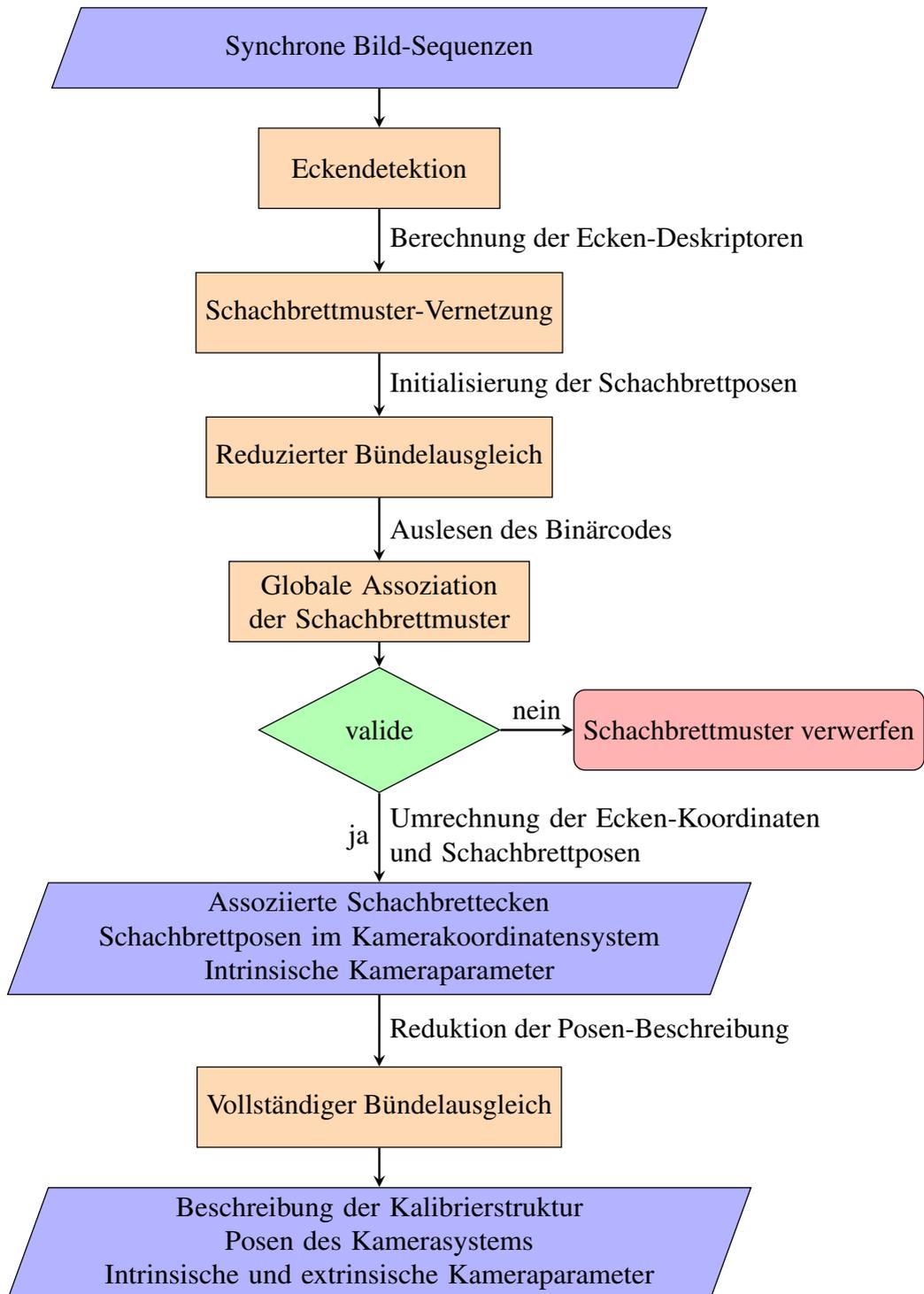


Abbildung 6.1: Ablaufdiagramm der Kalibrierung.

6.1 Initialisierung der Schachbrettposen

Im ersten Schritt wird ein reduzierter Bündelausgleich gerechnet, der die Beobachtungen der einzelnen Schachbrettmuster als voneinander unabhängig annimmt. Hierfür ist keine globale Assoziation der Schachbrettmuster erforderlich und der Bündelausgleich kann für jede Kamera unabhängig berechnet werden. Für die nichtlineare Optimierung ist es hierbei sinnvoll, die Schachbrettposen relativ zur jeweiligen Kamera zu initialisieren. Dadurch wird die Gefahr verringert, in ein falsches lokales Minimum zu konvergieren. Entscheidend bei der Initialisierung ist nicht die absolute Genauigkeit der Pose, sondern dass die ermittelte Modellbeschreibung in Kombination mit dem Kameramodell eine möglichst konsistente Beschreibung ist und somit einen geringen Startfehler aufweist.

Mit der Annahme eines planaren Musters lässt sich der Punkt mit den Schachbrettkoordinaten x_i und y_i im Kamera-Koordinatensystem ausdrücken als

$$\mathbf{x}^K = \mathbf{R}_B \begin{pmatrix} x^B \\ y^B \\ 0 \end{pmatrix} + \mathbf{t}_B = \underbrace{[\mathbf{r}_{B1} \ \mathbf{r}_{B2} \ \mathbf{t}_B]}_{\mathbf{H}} \begin{pmatrix} x^B \\ y^B \\ 1 \end{pmatrix} \quad (6.1)$$

wobei \mathbf{R}_B die Rotationsmatrix und \mathbf{t}_B der Translationsvektor der gesuchten Beobachtungspose des Schachbrettmusters sind. Zur Vereinfachung der Schreibweise werden die ersten beiden Spalten von \mathbf{R}_B sowie \mathbf{t}_B zu einer neuen Matrix \mathbf{H} zusammengefasst, welche homogene Schachbrettkoordinaten auf einen Punkt im dreidimensionalen Raum abbildet.

Werden die intrinsischen Kameraparameter \mathbf{p}_I als bekannt vorausgesetzt, so kann anhand der Beobachtungen im Kamerabild $\mathbf{b}_i = (u_i, v_i)$ und des Kameramodells der zum Punkt i gehörende Sichtstrahl

$$\mathcal{P}^{-1}(\mathbf{b}_i, \mathbf{p}_I) := \{\mathbf{z}_i + \lambda \mathbf{r}_i, \lambda \in \mathbb{R}^+\} \quad \text{mit} \quad \|\mathbf{r}_i\| = 1 \quad (6.2)$$

bestimmt werden. Der Abstand zwischen Schachbrettzentrum \mathbf{x}^K und Sichtstrahl kann über die Norm von \mathbf{e}_i berechnet werden - siehe auch Gl. (4.12):

$$\mathbf{e}_i = \mathbf{r}_i \times (\mathbf{x}_i^K - \mathbf{z}_i) = \mathbf{r}_i \times \left[\mathbf{H} \begin{pmatrix} x_i^B \\ y_i^B \\ 1 \end{pmatrix} - \mathbf{z}_i \right] \quad (6.3)$$

Für den Fall $\mathbf{e}_i = \mathbf{0}$ stimmt die Projektion des Schachbrettzentrums mit der Beobachtung im Kamerabild überein.

Werden die Einträge der Matrix \mathbf{H} in Form eines Spaltenvektors umsortiert zu

$$\mathbf{h} = (h_{11} \ h_{12} \ \cdots \ h_{33})^T \quad (6.4)$$

so führt Gl. (6.3) auf das unterbestimmte Gleichungssystem $\mathbf{A}_i \mathbf{h} = \mathbf{y}_i$ mit 3 Gleichungen und 9 Unbekannten. Mit $\mathbf{r}_i = (r_{i1} \ r_{i2} \ r_{i3})^\top$ und $\mathbf{w}_i = (x_i^B \ y_i^B \ 1)$ ergeben sich

$$\mathbf{A}_i = \begin{pmatrix} \mathbf{0} & -r_{i3} \mathbf{w}_i & r_{i2} \mathbf{w}_i \\ r_{i3} \mathbf{w}_i & \mathbf{0} & -r_{i1} \mathbf{w}_i \\ -r_{i2} \mathbf{w}_i & r_{i1} \mathbf{w}_i & \mathbf{0} \end{pmatrix} \quad \text{und} \quad \mathbf{y}_i = \mathbf{r}_i \times \mathbf{z}_i \quad (6.5)$$

Durch Zusammenfassen der Beobachtungen $i = 1 \dots n$ eines Schachbretts mit

$$\mathbf{A} = (\mathbf{A}_1^\top \ \mathbf{A}_2^\top \ \dots \ \mathbf{A}_n^\top)^\top \quad \text{und} \quad \mathbf{y} = (\mathbf{y}_1^\top \ \mathbf{y}_2^\top \ \dots \ \mathbf{y}_n^\top)^\top \quad (6.6)$$

ergibt sich das überbestimmte Gleichungssystem

$$\mathbf{A} \mathbf{h} = \mathbf{y} \quad (6.7)$$

Für den Fall einer Abbildung mit eindeutigem Projektionszentrum gilt $\mathbf{z}_i = \mathbf{0}$ und somit $\mathbf{y} = \mathbf{0}$. Damit degeneriert Gl. (6.7) zu dem homogenen Gleichungssystem $\mathbf{A} \mathbf{h} = \mathbf{0}$. Die nicht-triviale Lösung dieses Gleichungssystems kann mit Hilfe der Singulärwertzerlegung bestimmt werden [Har04]. Für den Fall $\mathbf{y} \neq \mathbf{0}$ kann die Lösung des Gleichungssystems direkt mittels des Least-Squares-Schätzers bestimmt werden. Als Anmerkung sei erwähnt, dass durch Verschieben des Koordinatensystemursprungs auch der Fall einer Abbildung mit eindeutigem Projektionszentrum in den Fall $\mathbf{y} \neq \mathbf{0}$ überführt werden kann.

Die Vorgehensweise zur Ermittlung der Matrix \mathbf{H} ist angelehnt an die Schätzung der Homographie-Beziehung zwischen zwei projektiven Abbildungen [Har04], mit dem Unterschied, dass anstelle von homogenen Bildkoordinaten direkt die zugehörigen Richtungsvektoren \mathbf{r}_i verwendet werden. Dies umgeht die Normierung auf $r_{i3} = 1$, welche für Blickwinkel nahe 180° ($r_{i3} \approx 0$) numerisch ungünstig ist. Für die Schachbrettkoordinaten x^B und y^B kann weiterhin, wie in [Har04] empfohlen, zur Verbesserung der numerischen Eigenschaften eine mittelwertbereinigende und varianznormalisierende Transformation eingeführt werden.

Die Matrix \mathbf{H} ist mit 9 Freiheitsgraden eine überbestimmte Darstellung der Pose, da diese nur 6 Freiheitsgrade besitzt. Nach der Definition in Gl. (6.1) sind die ersten beiden Spalten von $\mathbf{H} = (\mathbf{a} \ \mathbf{b} \ \mathbf{c})$ Teil einer Rotationsmatrix. Hieraus ergeben sich die drei Nebenbedingungen

$$\|\mathbf{a}\| = 1 \quad \|\mathbf{b}\| = 1 \quad \mathbf{a} \cdot \mathbf{b} = 0 \quad (6.8)$$

Sowohl bei der Singulärwertzerlegung als auch bei der Least-Squares-Schätzung können diese Nebenbedingungen nicht direkt berücksichtigt werden. Daher ist im

Allgemeinen keine exakte Zerlegung von \mathbf{H} in \mathbf{R} und \mathbf{t} möglich. Es kann also auch nicht davon ausgegangen werden, dass eine berechnete Zerlegung die Lösung des Optimierungsproblems unter den Nebenbedingungen darstellt – dennoch liefert sie meist eine brauchbare Näherungslösung, die als Initialisierung ausreichend ist.

Bei Verwendung der Singulärwertzerlegung wird \mathbf{H} mit nur 8 Freiheitsgraden geschätzt und nur bis auf einen Skalierungsfaktor bestimmt. Gegebenenfalls muss daher das Vorzeichen von \mathbf{H} umgedreht werden. Dies ist der Fall, wenn mittels \mathbf{H} abgebildete Punkte bezüglich \mathbf{x}_i^K auf der falschen Seite liegen, wenn also gilt:

$$\mathbf{x}_i^K \cdot \mathbf{H}\mathbf{w}_i < 0 \quad (6.9)$$

Als zusätzlicher Skalierungsfaktor kann beispielsweise der Kehrwert von $\|\mathbf{a}\|$ verwendet werden, um somit die erste Nebenbedingung zu erfüllen. Für den Translationsanteil von \mathbf{H} folgt damit:

$$\mathbf{t}_B = \frac{\mathbf{c}}{\|\mathbf{a}\|} \quad (6.10)$$

Eine Schätzung für die Rotationsmatrix kann durch Orthonormalisierung der Matrix $\mathbf{R}' = (\mathbf{a} \ \mathbf{b} \ \mathbf{a} \times \mathbf{b})$ gewonnen werden. Die Orthonormalisierung ergibt sich anhand der Singulärwertzerlegung $\mathbf{R}' = \mathbf{U}\mathbf{S}\mathbf{V}^T$ zu [Zha00]:

$$\mathbf{R} = \mathbf{U}\mathbf{V}^T \quad (6.11)$$

Die Rotationsmatrix ist eine weit verbreitete Darstellungsform für Rotationen. Im Rahmen der Optimierungsrechnung hat diese direkte Darstellungsform jedoch den Nachteil, dass sie mit ihren neun Matrixkomponenten mehr Freiheitsgrade besitzt als die Rotation selbst mit nur drei Freiheitsgraden. Eine Möglichkeit, diesem Problem zu begegnen, ist die Einführung von sechs zusätzlichen Zwangsbedingungen an die Komponenten der Rotationsmatrix. In dieser Arbeit wird jedoch ein alternativer Ansatz basierend auf der Rodrigues-Darstellung von Rotationen verwendet. Hierbei wird eine Rotation durch einen Vektor der Länge drei beschrieben, wobei die Richtung des Vektors die Rotationsachse beschreibt und die Länge des Vektors den Rotationswinkel. Mit genau drei Freiheitsgraden ist die Rodrigues-Darstellung somit eine minimalparametrische Repräsentation einer Rotation.

An dieser Stelle wird auf eine nichtlineare Optimierung der Schachbrettpose verzichtet, da die Pose lediglich als Initialisierung verwendet wird. Während der Einzel-Kamera-Kalibrierung (Abschnitt 6.2) werden dann alle Schachbrettposen gemeinsam mit den intrinsischen Parametern der Kamera optimiert.

6.2 Kamera-Kalibrierung ohne globale Assoziation

Die Kamera-Kalibrierung ohne globale Assoziation kann für jede Kamera getrennt durchgeführt werden und ist vergleichbar mit bislang üblicherweise verwendeten Kalibrierverfahren. Als Eingangsdaten dienen die zu Schachbrettmustern vernetzten, jedoch noch nicht global assoziierten Ecken (Abschnitt 5.4). Aufgrund der noch nicht global assoziierten Beobachtungen kann die zeitliche Konsistenz der Kalibrierstruktur noch nicht berücksichtigt werden.

Die Kalibrierung der einzelnen Kameras liefert neben einer groben Schätzung der intrinsischen Kameraparameter die Posen der beobachteten Schachbrettmuster im jeweiligen Kamera-Koordinatensystem. Hierbei beschreibt eine Schachbrettpose die relative Lage des während der Vernetzung willkürlich gewählten Schachbrett-Koordinatensystems im Kamera-Koordinatensystem.

6.3 Globale Assoziation der Schachbrettmuster

Unter Verwendung des Kameraabbildungsmodells und der Schachbrettposen aus den einzelnen Kamera-Kalibrierungen (Abschnitt 6.2) ist es möglich, die Projektionen beliebiger Punkte aus der Ebene eines Schachbrettmusters ins Kamerabild zu berechnen. Dies ist die Voraussetzung, um den zur globalen Assoziation der Schachbrettbeobachtungen verwendeten Binärcode lesen zu können, welcher die Schachbrettmuster umgibt (Abschnitt 5.5).

Das Auslesen der Binärcodes liefert eine Kennzahl, die die Beobachtungen global eindeutig den Schachbrettmustern der Kalibrierstruktur zuordnet. Darüber hinaus können anhand der Informationen des Binärcodes alle Schachbrettecken eindeutig adressiert werden. Hieraus ergibt sich eine Berechnungsvorschrift zur Transformation des während der Vernetzung willkürlich gewählten Schachbrett-Koordinatensystems in ein zeitlich konsistent gewähltes und für alle Kameras einheitliches Schachbrett-Koordinatensystem. Die Umrechnung eines Punktes $\tilde{\mathbf{x}}^B$ vom ursprünglichen Schachbrett-Koordinatensystem in das globale Schachbrett-Koordinatensystem lässt sich in Matrix-Vektor-Schreibweise darstellen als:

$$\mathbf{x}^B = \underbrace{\begin{pmatrix} R_{11} & R_{12} & 0 \\ R_{21} & R_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\mathbf{R}_G} \tilde{\mathbf{x}}^B + \underbrace{\begin{pmatrix} t_1 \\ t_2 \\ 0 \end{pmatrix}}_{\mathbf{t}_G} \quad (6.12)$$

Hierbei beschreibt \mathbf{R}_G die Rotation zwischen den beiden Koordinatensystemen und \mathbf{t}_G deren Verschiebung zueinander.

Im Zuge einer Änderung des Schachbrett-Koordinatensystems müssen auch die während der initialen Kamera-Kalibrierung ermittelten Posen entsprechend transformiert werden, da diese jeweils die Beziehung zweier Koordinatensysteme zueinander beschreiben. Für die Pose in Verbindung mit den global assoziierten Schachbrettkoordinaten muss gelten:

$$\begin{pmatrix} \tilde{\mathbf{R}}_B & \tilde{\mathbf{t}}_B \\ \mathbf{0} & 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_G & \mathbf{t}_G \\ \mathbf{0} & 1 \end{pmatrix} \quad (6.13)$$

Damit gilt für die Umrechnung der Schachbrettpose in das neue Koordinatensystem:

$$\begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{R}}_B & \tilde{\mathbf{t}}_B \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_G & \mathbf{t}_G \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \quad (6.14)$$

Die gemeinsame Transformation von Schachbrett-Koordinaten und Schachbrettpose hat keinen Einfluss auf das Fehlermaß der Kamera-Kalibrierung. Lediglich die Schachbrettecken sind anschließend für alle Beobachtungen einheitlich indiziert.

Durch die global assoziierten Beobachtungen ist es möglich, zusätzlich auch Abweichungen von der Annahme eines idealen planaren Schachbrettmusters zu modellieren, da nun jede Beobachtung einer Schachbrettecke eindeutig einer Schachbrettecke der statischen Kalibrierstruktur zugeordnet werden kann.

Schlägt die globale Assoziation für eine Schachbrett-Beobachtung fehl, so handelt es sich in der Regel um eine fehlerhafte Beobachtung, die entfernt werden kann. Die Durchführung der globalen Assoziation trägt somit gleichzeitig auch zu einer Erhöhung der Robustheit der Kalibrierung bei.

6.4 Reduktion der Posen-Beschreibung

Bei der Verwendung der Kalibrierung ohne globale Assoziation wird für jede Beobachtung eines Schachbrettmusters dessen Pose im Kamera-Koordinatensystem geschätzt. Im Allgemeinen ergibt sich hieraus eine überbestimmte Beschreibung der tatsächlichen Posenzusammenhänge. Um die Robustheit und Güte der Schätzungen zu erhöhen, sollte die Anzahl der zu schätzenden Posen auf den Fall der eindeutigen Beschreibung reduziert werden. Ziel der Posen-Reduktion ist es, gegebene Schachbrettposen im Kamera-Koordinatensystem $(\mathbf{R}_B, \mathbf{t}_B)$ darzustellen als Kette nacheinander angewandter Transformationen (siehe auch Abbildung 4.1).

Die Posen der einzelnen Schachbrettmuster der Kalibrierstruktur werden hierzu in einem globalen Referenzsystem beschrieben $(\mathbf{R}_S, \mathbf{t}_S)$. Über die Systempose

$(\mathbf{R}_R, \mathbf{t}_R)$ kann diese dann in ein mit dem Kameraaufbau mitbewegtes Koordinatensystem transformiert werden. Die extrinsische Kamerapose $(\mathbf{R}_E, \mathbf{t}_E)$ beschreibt die Transformation vom bewegten Koordinatensystem des Kamerasystems in das Koordinatensystem der jeweiligen Kamera. Insgesamt soll gelten:

$$\begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{R}_E & \mathbf{t}_E \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R}_R & \mathbf{t}_R \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R}_S & \mathbf{t}_S \\ \mathbf{0} & 1 \end{pmatrix} \quad (6.15)$$

Zu beachten ist an dieser Stelle, dass das Schachbrett-Koordinatensystem entgegengesetzt zur Transformationsrichtung der anderen Posen festgelegt wurde.

Sowohl das globale Referenzsystem als auch das Koordinatensystem für den Kameraaufbau sind nicht eindeutig festgelegt. Das globale Referenzsystem kann beispielsweise so gewählt werden, dass eine der Schachbrettposen der Nullpose ($\mathbf{R}_S = \mathbf{I}$ und $\mathbf{t}_S = \mathbf{0}$) entspricht. Analog kann das Koordinatensystem für den Kameraaufbau so gewählt werden, dass eine der extrinsischen Kameraposen der Nullpose entspricht.

Die Anzahl der globalen Schachbrettposen und der extrinsischen Kameraposen sind für einen gegebenen Kalibrieraufbau konstant, während die Anzahl der Systemposen an die Länge der Kalibriersequenz gekoppelt ist. In der Regel kann zu einem Zeitpunkt der Kalibriersequenz für alle Kameras gemeinsam mehr als nur ein einzelnes Schachbrettmuster beobachtet werden, während für die Beschreibung nur eine einzige Systempose erforderlich ist, welche die zeitliche Bewegung beschreibt. Vorausgesetzt, dass zwei Posen fest gewählt wurden, führt dieser Zusammenhang dazu, dass die Gesamtheit der Transformationsketten in der Regel mit weniger Posen beschrieben wird als Beobachtungsposen vorhanden sind.

Im Folgenden wird nun ein Verfahren beschrieben, wie aus den Beobachtungsposen die Posen der Transformationsketten ermittelt werden können. Dieser Vorgang ist mit einer Informationsreduktion verbunden. Ziel des Algorithmus ist hierbei nicht die Bestimmung der optimalen Posenkette, sondern lediglich die robuste Ermittlung einer geeigneten Initialisierung für die anschließende Multi-Kamera-Kalibrierung (Abschnitt 6.4).

Das Verfahren zur Posenreduktion gliedert sich in zwei Schritte. Im ersten Schritt wird eine Beschreibung der Kalibrierstruktur in einem globalen Referenzsystem aufgebaut und der zeitliche Verlauf der Bewegung der einzelnen Kameras ermittelt. Im zweiten Schritt werden dann aus dem zeitlichen Verlauf der Kameraposen die extrinsischen Kameraposen ermittelt, welche die relative Lage der einzelnen Kameras zueinander beschreiben. Prinzipiell kann die Posenreduktion auch direkt als Posengraph-Optimierungsproblem formuliert werden.

Aufbau der globalen Kalibrierstruktur

Sind die Schachbrettposen der Kalibrierstruktur in einem globalen Referenzsystem $(\mathbf{R}_S, \mathbf{t}_S)$ gegeben und der zeitliche Verlauf der Bewegung der einzelnen Kameras in diesem Referenzsystem $(\mathbf{R}_K, \mathbf{t}_K)$ bekannt, so können die im jeweiligen Kamera-Koordinatensystem beobachteten Schachbrettposen $(\mathbf{R}_B, \mathbf{t}_B)$ wie folgt beschrieben werden:

$$\begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{R}_K & \mathbf{t}_K \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{R}_S & \mathbf{t}_S \\ \mathbf{0} & 1 \end{pmatrix} \quad (6.16)$$

Dies entspricht der Beschreibung eines zeitlich flexiblen Kameraaufbaus, da im Gegensatz zu Gl. (6.15) die einzelnen Kameraposen als voneinander unabhängig beschrieben sind. Eine Möglichkeit, eine solche Zerlegung zu finden, besteht darin, die globale Kalibrierstruktur iterativ aufzubauen. Hierzu wird zunächst dasjenige Schachbrettmuster gesucht, welches am häufigsten beobachtet wurde und dieses willkürlich als Ursprung des globalen Koordinatensystems gewählt. Ist eine Schachbrettpose im globalen Koordinatensystem bekannt, so können anhand dieser die Kameraposen ermittelt werden:

$$\begin{pmatrix} \mathbf{R}_K & \mathbf{t}_K \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_S & \mathbf{t}_S \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \quad (6.17)$$

Für das als Referenz gewählte Schachbrettmuster gelten:

$$\mathbf{R}_S = \mathbf{I} \quad \mathbf{t}_S = \mathbf{0} \quad (6.18)$$

Anschließend können alle Beobachtungen von Schachbrettmustern, die von diesen Kameraposen aus gemacht wurden, in das globale Koordinatensystem transformiert werden:

$$\begin{pmatrix} \mathbf{R}_S & \mathbf{t}_S \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_K & \mathbf{t}_K \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_B & \mathbf{t}_B \\ \mathbf{0} & 1 \end{pmatrix} \quad (6.19)$$

Im Allgemeinen existieren nun aufgrund von Beobachtungs- bzw. Schätzunsicherheiten verschiedene Posen für die einzelnen Schachbrettmuster. Um diese robust zu fusionieren, wird eine komponentenweise Medianfilterung für alle zusammenhängenden Posen durchgeführt. Für die komponentenweise Filterung ist es sinnvoll, eine nicht-periodische Darstellung der Rotation zu verwenden. Da die minimal-parametrische Darstellung in Form von Rodrigues-Vektoren aus diesem Grund ungeeignet ist, werden für die Mittelung direkt die Komponenten der Rotationsmatrizen verwendet. Die aus der Mittelung resultierende Matrix \mathbf{R}'_S erfüllt in der Regel nicht mehr die Orthonormalitätseigenschaft einer Rotationsmatrix. Eine

gültige Rotationsmatrix kann aber anhand der Singulärwertzerlegung $\mathbf{R}'_S = \mathbf{USV}^\top$ erzeugt werden:

$$\mathbf{R}_S := \mathbf{UV}^\top \quad (6.20)$$

Im nächsten Schritt muss ein neues Schachbrettmuster als Referenz gewählt werden. Hierzu wird dasjenige Schachbrettmuster gesucht, für welches bereits die meisten Beobachtungen in das globale Koordinatensystem transformiert wurden und welches bislang noch nicht als Referenz verwendet wurde. Anschließend wird die Iterationsschleife erneut durchlaufen, wobei jedoch nur noch diejenigen Beobachtungen transformiert werden, welche nicht bereits im globalen Koordinatensystem dargestellt sind. Dieser Vorgang wird solange wiederholt, bis alle Kameraposen ermittelt sind.

In seltenen Fällen kann es vorkommen, dass auf die beschriebene Weise die Kalibrierstruktur nicht vollständig ermittelt werden kann. Dies ist dann der Fall, wenn die Kalibrierstruktur ohne Berücksichtigung der Annahme eines festen Kameraaufbaus nicht zusammenhängend beobachtbar ist. Abhilfe schafft hier die Berücksichtigung der extrinsischen Kameraposen aus dem nächsten Schritt, mit Hilfe derer dann die Beobachtungen aller Kameras zu einem Zeitpunkt in das globale Koordinatensystem transformiert werden können, wenn mit mindestens einer Kamera ein Schachbrettmuster beobachtet wird, dessen globale Position bereits bekannt ist.

Bestimmung der extrinsischen Kameraposen

Nachdem im letzten Abschnitt gezeigt wurde, wie die zeitliche Konsistenz der Kalibrierstruktur berücksichtigt werden kann, soll nun auch für den zeitlichen Verlauf der Kameraposen die Eigenschaft eines festen Kameraaufbaus berücksichtigt werden. Hierzu wird versucht, den im ersten Schritt ermittelten zeitlichen Verlauf der Kameraposen ($\mathbf{R}_K, \mathbf{t}_K$) in Form von zeitabhängig bewegten Systemposen und zeitlich unveränderlichen extrinsischen Kameraposen darzustellen:

$$\begin{pmatrix} \mathbf{R}_K & \mathbf{R}_K \\ \mathbf{0} & 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} \mathbf{R}_R & \mathbf{t}_R \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_E & \mathbf{t}_E \\ \mathbf{0} & 1 \end{pmatrix} \quad (6.21)$$

Prinzipiell ist zur Bestimmung der extrinsischen Kameraposen ein Vorgehen analog zum Aufbau der globalen Kalibrierstruktur möglich. Dennoch soll hier noch ein alternativer Ansatz erläutert werden. Hierzu werden zunächst die Posentransformationen zwischen allen Paaren von Kameras ermittelt und diese dann, wie im letzten Abschnitt beschrieben, über Medianfilterung zusammengefasst. Hieraus

ergibt sich zunächst ein überbestimmtes Netz von Posenbeziehungen. Die Kameras können hierbei als Knoten in einem Graph aufgefasst werden, während die Posentransformationen zwischen jeweils zwei Kameras Kanten darstellen. Um die Posenbeziehungen auf den eindeutigen Fall zu reduzieren, werden anschließend die Kanten des minimalen Spannbaums ausgewählt. Hierbei wird als Gewicht einer Kante das Inverse der Anzahl der Beobachtungen dieser Kante verwendet. Anschließend wird ein beliebiger Knoten als Referenz für den Kameraaufbau gewählt und der Graph durchlaufen, wobei über die Posentransformationen zwischen jeweils zwei Kameras schrittweise die Posen aller Kameras ermittelt werden.

Sind alle extrinsischen Posen bekannt, kann aus den Kameraposen im globalen Koordinatensystem der Verlauf der Posen des Kameraaufbaus ermittelt werden:

$$\begin{pmatrix} \mathbf{R}_R & \mathbf{t}_R \\ \mathbf{0} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R}_K & \mathbf{R}_K \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_E & \mathbf{t}_E \\ \mathbf{0} & 1 \end{pmatrix}^{-1} \quad (6.22)$$

Meist wird eine zeitliche Pose des Kameraaufbaus durch mehrere Kameraposen beobachtet. In diesem Fall werden die einzelnen Beobachtungen wieder mittels Medianfilterung zusammengefasst und anschließend orthonormalisiert.

6.5 Multi-Kamera-Kalibrierung mit global assoziierten Beobachtungen

Unter Verwendung der in Abschnitt 6.1 beschriebenen Posen-Initialisierung können die einzelnen Kameras eines Multi-Kamera-Aufbaus bereits grob kalibriert werden (Abschnitt 6.2). Mithilfe dieser Kalibrierung kann die in Abschnitt 5.5 beschriebene globale Assoziation durchgeführt werden.

Die einzelnen Kamera-Kalibrierungen liefern im Allgemeinen eine überbestimmte Beschreibung der Posenzusammenhänge. Aufgrund von Beobachtungs- oder Modellfehlern hat dies eine inkonsistente Schätzung zur Folge, sodass für die verschiedenen Kameras unterschiedliche Kalibrierstrukturen rekonstruiert werden. Darüber hinaus werden auch die Annahmen einer statischen Kalibrierstruktur sowie eines statischen Kameraaufbaus bei den einzelnen Kamera-Kalibrierungen nicht berücksichtigt.

Durch die Posen-Fusion (Abschnitt 6.4) wird die überbestimmte Beschreibung der Posenzusammenhänge auf den Fall einer eindeutigen Beschreibung reduziert. Diese reduzierte Beschreibung wird als Initialisierung der verschiedenen Posen des vollständigen Bündelausgleichs (Kapitel 4) verwendet. Da die Reduktion im Allgemeinen mit Informationsverlust verbunden ist, ist eine erneute Optimierung

erforderlich. Wie in Abschnitt 6.4 beschrieben, können für die Festlegung der Transformationskette zwei Posen beliebig gewählt werden. Im Rahmen der Optimierung ist es sinnvoll, diese Posen konstant zu halten, um numerische Probleme zu vermeiden.

7 Beurteilung der Kalibrierung

In der Messtechnik gilt der Grundsatz, dass ein Messwert erst in Kombination mit einer Angabe über dessen Unsicherheit brauchbare Information liefert [iso95]. Die Formulierung der Kamera–Kalibrierung als Bündelausgleich (Kapitel 4) umfasst zwar weitgehend anschauliche Kostenfunktionen, die im Rahmen der Optimierung minimiert werden, liefert aber bislang keine Aussage darüber, wie verlässlich die damit erzielten Ergebnisse sind und wie das Residuum zu interpretieren ist.

Zur Bewertung der Genauigkeit einer Kalibrierung wird oftmals der verbleibende mittlere Rückprojektionsfehler herangezogen. Da jedoch die betrachtete Szene im Rahmen der Optimierung mitgeschätzt wird, ist ein kleiner Rückprojektionsfehler unter Umständen auch auf Überanpassung zurückzuführen. Wird beispielsweise nur ein einzelnes Kamerabild und ein nicht-planares Schachbrettmodell verwendet, so kann das Schachbrettmodell stets so angepasst werden, dass der Rückprojektionsfehler verschwindet. Somit kann ein geringer Rückprojektionsfehler allenfalls als Indiz für eine genaue Kalibrierung gewertet werden. In Abschnitt 7.1 wird gezeigt, dass der Rückprojektionsfehler maßgeblich von der Unsicherheit des Schachbrettecken-Detektors beeinflusst wird. Im Umkehrschluss deutet ein großer Rückprojektionsfehler auf eine möglicherweise unzureichende Modellierung oder fehlerhaft detektierte Schachbrettecken hin. Auch starkes Bildrauschen kann Ursache für einen großen Rückprojektionsfehler sein.

Die Frage nach der Genauigkeit einer Kalibrierung lässt sich ohne Kenntnis der wahren Abbildung nicht ohne weiteres beantworten. Teilweise wird versucht, anhand von zusätzlichen Anwendungsexperimenten auf die Genauigkeit einer Kalibrierung zu schließen. Dies ist jedoch meist mit hohem Aufwand verbunden und birgt die Gefahr in sich, dass das zur Bewertung herangezogene Anwendungsexperiment möglicherweise invariant gegenüber gewissen Fehlern des Abbildungsmodells ist und diese somit unentdeckt bleiben.

In Abschnitt 7.2 wird die Beobachtbarkeit der Abbildung eingeführt und zur Beurteilung von Kalibrierergebnissen vorgeschlagen. Die Beobachtbarkeit setzt eine Änderung der Kameraabbildung mit der jeweiligen Änderung des Rückprojektionsfehlers in Beziehung. Bei einer schlechten Beobachtbarkeit, wie im erwähnten Beispiel mit Überanpassung, kann das Abbildungsmodell in einem großen Bereich variiert werden, ohne dass dies zu einem nennenswerten Anstieg des Rückprojektionsfehlers führt. Dies ist ein Anzeichen für eine wenig verlässliche Kalibrierung.

7.1 Interpretation des Rückprojektionsfehlers

In Abschnitt 4.4 wurde bereits erwähnt, dass die Minimierung des mittleren quadratischen Rückprojektionsfehlers im messtechnischen Sinne optimal ist, falls der Messfehler ausschließlich auf einen normalverteilten und mittelwertfreien Fehler des Schachbrettecken-Detektors zurückzuführen ist. Im Sinne dieser Aussage liefert der Rückprojektionsfehler also keine unmittelbare Information über die Güte eines Kalibrierergebnisses, sondern nur über die des Schachbrettecken-Detektors. Dennoch kann der Rückprojektionsfehler unter Umständen auch darüber hinaus wichtige Informationen liefern.

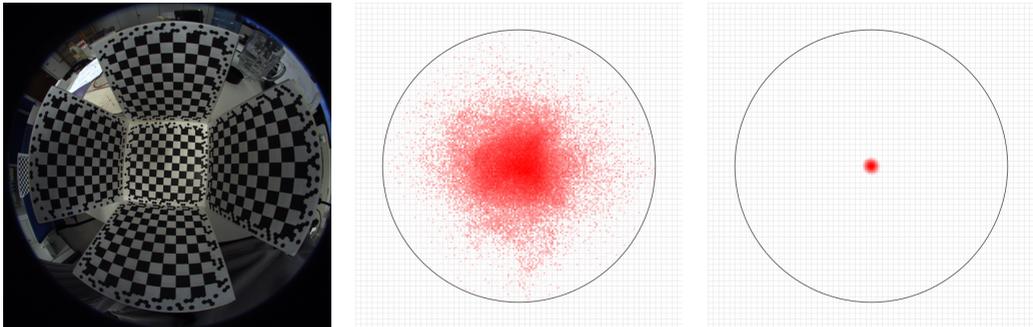


Abbildung 7.1: Modellierung der Verschiebung des Projektionszentrums. Links: Einzelbild einer Kalibriersequenz für ein Fischaugen-Objektiv mit einem aufgebogenen Kalibrierwürfel mit etwa 12cm Seitenlänge. Mitte und rechts: Histogramm des Rückprojektionsfehlers nach der Kalibrierung, jeweils normiert anhand des Histogramm-Maximums ohne Verschiebung des Projektionszentrums (Mitte) bzw. mit Modellierung einer Verschiebung (rechts). Der schwarze Kreis entspricht einem Rückprojektionsfehler von fünf Pixeln.

Vorausgesetzt, die Normalverteilungshypothese für den Fehler des Schachbrettecken-Detektors ist zutreffend, so deutet eine Abweichung der Fehlerverteilung des Rückprojektionsfehlers von der Normalverteilungshypothese auf einen systematischen Fehler hin. Durch die Validierung der Beobachtungen mittels der globalen Assoziation (Abschnitt 5.5) können Ausreißer als Ursache weitgehend ausgeschlossen werden. Meist ist daher ein systematischer Fehler auf eine unzureichende Modellierung zurückzuführen. Dies kann beispielsweise eine nicht modellierte Verschiebung des Projektionszentrums sein (siehe Abbildung 7.1). Auch die Nichtberücksichtigung von Verletzungen der Planaritätsannahme für die Kalibrierstruktur (siehe Abbildung 7.2) führt zu einem höheren Rückprojektionsfehler, ist allerdings in der Abbildung nicht direkt als Abweichung von der Normalverteilung zu erkennen. Sehr schön zu sehen ist die Fehlerverteilung des Schachbrettecken-Detektors,

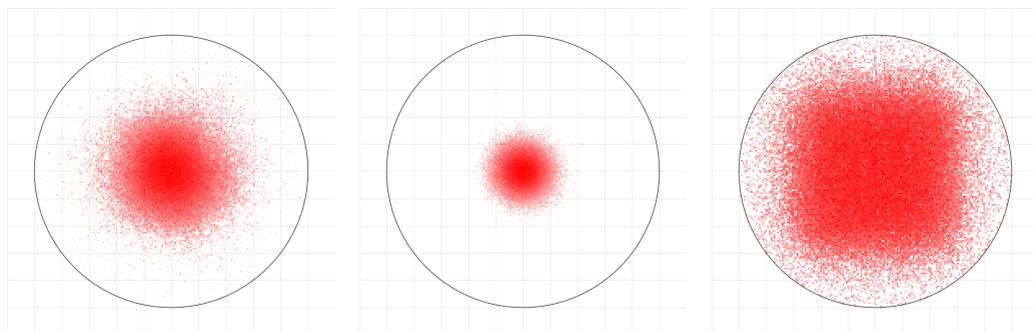


Abbildung 7.2: Modellierung von nicht-planaren Schachbrettmustern und Einfluss von Subpixel-Verfeinerung. Histogramm des Rückprojektionsfehlers nach der Kalibrierung, jeweils normiert anhand des Histogramm-Maximums. Der schwarze Kreis hat, im Unterschied zu Abbildung 7.1, einen Durchmesser von nur einem Pixel. Links: Modellierung mit planaren Schachbrettmustern, Mitte: Modellierung mit nicht-planaren Schachbrettmustern, rechts: Modellierung mit nicht-planaren Schachbrettmustern ohne Subpixel-Verfeinerung bei der Extraktion der Schachbrettecken.

wenn für die Extraktion der Schachbrettecken die Subpixel-Verfeinerung deaktiviert wird. In diesem Fall verbleibt, wie zu erwarten, als Rückprojektionsfehler nach der Kalibrierung näherungsweise eine Gleichverteilung über einem Quadrat der Kantenlänge von 1Px .

7.2 Beobachtbarkeit der Abbildung

Die Beobachtbarkeit wird im Folgenden als Bewertungsmaß verstanden, welches ausdrückt, wie robust das während der Kalibrierung ermittelte Abbildungsmodell bestimmt werden kann. Hierzu wird untersucht, inwieweit eine Änderung des Abbildungsmodells aus dem Optimum heraus auch mit einem Anstieg des Rückprojektionsfehlers verbunden ist. Bei schlechter Beobachtbarkeit führen auch größere Änderungen des Abbildungsmodells zu einem nur geringen Anstieg des Rückprojektionsfehlers.

Um auch Kalibrierungen mit unterschiedlichen Kameramodellen miteinander vergleichen zu können, wird die Beobachtbarkeit des Abbildungsmodells als Ganzes untersucht und nicht die Beobachtbarkeit einzelner Parameter. Die Parameter-Beobachtbarkeit ist zwar numerisch ebenfalls wichtig, ihr liegt jedoch zusätzlich die an dieser Stelle weniger relevante Fragestellung zugrunde, ob das Abbildungsmodell durch einen Satz von Parametern eindeutig beschrieben wird.

Im Folgenden wird die Beobachtbarkeit für den Fall der Kalibrierung einer einzelnen Kamera untersucht. Für die Betrachtungen wird der im Zuge der Kalibrierung ermittelte Parametervektor \mathbf{p} aufgespalten in den intrinsischen Parametervektor \mathbf{p}_I zur Beschreibung des Abbildungsmodells und den Parametervektor \mathbf{p}_S für die sekundären Parameter (Beschreibung der Kalibrierstruktur, Bewegung des Kameraaufbaus etc.).

Mit $\mathcal{P}(\mathbf{p}_I)$ wird das Abbildungsverhalten der Kamera beschrieben. Als Maß für die Beobachtbarkeit wird nun der Kostenanstieg des Kalibrierungskostenmaßes in Abhängigkeit von einer Änderung der intrinsischen Parameter $\Delta\mathbf{p}_I$ in Beziehung gesetzt zur entsprechenden Änderung des Abbildungsmodells.

Für die Beschreibung des Unterschieds zweier Abbildungsmodelle wird untersucht, wie sich die Projektion gegebener Sichtstrahlen unterscheidet. Hierzu wird eine skalare Kostenfunktion $\mathcal{K}'(\mathcal{P}_1(\mathbf{p}_{I1}), \mathcal{P}_2(\mathbf{p}_{I2}))$ eingeführt, mit der der Unterschied zwischen zwei Kameramodellen bewertet werden kann. Als Kostenfunktion kann beispielsweise die durch gleichmäßige Abtastung des Sichtbereichs der Kamera bestimmte mittlere quadratische Differenz der Rückprojektionen verwendet werden. Da während der Kalibrierung ein beliebiges Koordinatensystem gewählt wird, um das Abbildungsmodell zu beschreiben, ist es wichtig, zusätzlich ein geändertes Koordinatensystem berücksichtigen zu können. Beispielsweise kann eine geringfügige Verschiebung des Bildhauptpunktes weitgehend durch eine Verdrehung des Koordinatensystems kompensiert werden. Zur Berücksichtigung dieser Kompensationsmöglichkeit wird eine zusätzliche Pose \mathbf{q} zur Beschreibung einer Koordinatensystem-Transformation $\mathcal{T}(\mathbf{q})$ eingeführt. Das transformierte Abbildungsmodell wird dann durch $\mathcal{T}(\mathbf{q}) \circ \mathcal{P}(\mathbf{p}_I)$ beschrieben. Unter Berücksichtigung eines möglicherweise veränderten Koordinatensystems gilt für die effektive Änderung des Abbildungsmodells $K(\Delta\mathbf{p}_I)$:

$$\mathcal{K}(\Delta\mathbf{p}_I) = \min_{\mathbf{q}} \{\mathcal{K}'(\mathcal{P}(\mathbf{p}_I), \mathcal{T}(\mathbf{q}) \circ \mathcal{P}(\mathbf{p}_I + \Delta\mathbf{p}_I))\} \quad (7.1)$$

Analog zur Beschreibung des Unterschieds zweier Abbildungsmodelle wird für das Kalibrierungskostenmaß eine skalarwertige Kostenfunktion $\mathcal{C}'(\mathbf{p}_I, \mathbf{p}_S)$ eingeführt. Als Kostenfunktion kann beispielsweise der mittlere quadratische Rückprojektionsfehler verwendet werden. Da während der Kalibrierung der Parametervektor \mathbf{p}_S beliebig gewählt werden kann, ist es wichtig, diese Kompensationsmöglichkeit ebenfalls zu berücksichtigen. Für die effektive Änderung des Kostenmaßes gilt:

$$\mathcal{C}(\Delta\mathbf{p}_I) = \min_{\Delta\mathbf{p}_S} \{\mathcal{C}'(\mathbf{p}_I + \Delta\mathbf{p}_I, \mathbf{p}_S + \Delta\mathbf{p}_S)\} - \mathcal{C}'(\mathbf{p}_I, \mathbf{p}_S) \quad (7.2)$$

Die Beobachtbarkeit \mathcal{B}' einer Parameteränderung $\Delta\mathbf{p}_I$ wird damit definiert als:

$$\mathcal{B}'(\Delta\mathbf{p}_I) = \frac{\mathcal{C}(\Delta\mathbf{p}_I)}{\mathcal{K}(\Delta\mathbf{p}_I)} \quad (7.3)$$

Entscheidend für die Beobachtbarkeit des Abbildungsmodells B ist diejenige Parameteränderung $\Delta \mathbf{p}_I$, die am schlechtesten beobachtbar ist:

$$B = \min_{\Delta \mathbf{p}_I} \left\{ \lim_{k \rightarrow 0} \frac{\mathcal{C}(k\Delta \mathbf{p}_I)}{\mathcal{K}(k\Delta \mathbf{p}_I)} \right\} \quad (7.4)$$

Wichtig für die Vergleichbarkeit der Beobachtbarkeit ist die Verwendung normierter Kostenmaße. Das Kostenmaß \mathcal{C} für die Kalibrierungskosten sollte daher anhand der Anzahl der Beobachtungen, das Kostenmaß \mathcal{K} anhand der Anzahl der zur Auswertung gewählten Abtastpunkte normiert werden.

Im Folgenden soll nun die Bestimmung des Beobachtbarkeitsmaßes nach Gl. (7.4) im Optimum des Bündelausgleichs gemäß der Formulierung aus Gl. (4.1) näher untersucht werden. Unabhängig von der verwendeten Kostenfunktion gelte für den Residuenvektor die Linearisierung:

$$\mathcal{E}(\mathbf{p}_I + \Delta \mathbf{p}_I, \mathbf{p}_S + \Delta \mathbf{p}_S) = \mathcal{E}(\mathbf{p}_I, \mathbf{p}_S) + (\mathbf{J}_i \ \mathbf{J}_s) \begin{pmatrix} \Delta \mathbf{p}_I \\ \Delta \mathbf{p}_S \end{pmatrix} \quad (7.5)$$

und damit für die quadratische Näherung des Anstiegs der Kostenfunktion:

$$\mathcal{C}'(\mathbf{p}_I + \Delta \mathbf{p}_I, \mathbf{p}_S + \Delta \mathbf{p}_S) = (\Delta \mathbf{p}_I^\top \ \Delta \mathbf{p}_S^\top) \begin{pmatrix} \mathbf{J}_i^\top \mathbf{J}_i & \mathbf{J}_i^\top \mathbf{J}_s \\ \mathbf{J}_s^\top \mathbf{J}_i & \mathbf{J}_s^\top \mathbf{J}_s \end{pmatrix} \begin{pmatrix} \Delta \mathbf{p}_I \\ \Delta \mathbf{p}_S \end{pmatrix} \quad (7.6)$$

Da im Optimum der Gradient dem Null-Vektor gleicht, tauchen die linearen Terme bezüglich $\Delta \mathbf{p}_I$ und $\Delta \mathbf{p}_S$ in Gl. (7.6) nicht auf.

Für die effektive Änderung des Kostenmaßes $\mathcal{C}(\Delta \mathbf{p}_I)$ gilt nun:

$$\mathcal{C}(\Delta \mathbf{p}_I) = \min_{\Delta \mathbf{p}_S} \left\{ \Delta \mathbf{p}_I^\top \mathbf{J}_i^\top \mathbf{J}_i \Delta \mathbf{p}_I + 2\Delta \mathbf{p}_S^\top \mathbf{J}_s^\top \mathbf{J}_i \Delta \mathbf{p}_I + \Delta \mathbf{p}_S^\top \mathbf{J}_s^\top \mathbf{J}_s \Delta \mathbf{p}_S \right\} \quad (7.7)$$

Durch Ableiten nach $\Delta \mathbf{p}_S$ folgt damit für das Minimum:

$$\Delta \mathbf{p}_S = -(\mathbf{J}_s^\top \mathbf{J}_s)^{-1} \mathbf{J}_s^\top \mathbf{J}_i \Delta \mathbf{p}_I \quad (7.8)$$

An dieser Stelle wird vorausgesetzt, dass die Jacobi-Matrix \mathbf{J}_s keine Nullspalten aufweist, da ansonsten die Matrix $\mathbf{J}_s^\top \mathbf{J}_s$ nicht regulär und somit deren Inverse nicht bestimmbar ist. Falls \mathbf{J}_s Nullspalten enthält (z.B. durch Festhalten von Parametern in der Optimierung), so müssen diese vorab eliminiert werden, indem die entsprechenden Komponenten des Parametervektors entfernt werden.

Einsetzen von Gl. (7.8) in Gl. (7.7) und Vereinfachen ergibt:

$$\mathcal{C}(\Delta \mathbf{p}_I) = \Delta \mathbf{p}_I^\top \left[\mathbf{J}_i^\top \mathbf{J}_i - \mathbf{J}_i^\top \mathbf{J}_s (\mathbf{J}_s^\top \mathbf{J}_s)^{-1} \mathbf{J}_s^\top \mathbf{J}_i \right] \Delta \mathbf{p}_I \quad (7.9)$$

Der Ausdruck $\mathbf{J}_s (\mathbf{J}_s^\top \mathbf{J}_s)^{-1} \mathbf{J}_s^\top \mathbf{J}_i$ entspricht anschaulich der Projektion von \mathbf{J}_i auf den von \mathbf{J}_s aufgespannten Unterraum. Die Berechnung der Inversen ist teilweise schlecht konditioniert, kann aber unter Verwendung der reduzierten QR-Zerlegung $\mathbf{J}_s = \mathbf{Q}_s \mathbf{R}_s$ mit $\mathbf{J}_s, \mathbf{Q}_s \in \mathbb{R}^{m \times n}$ umgangen werden. Unter der Voraussetzung, dass \mathbf{J}_s vollen Rang besitzt, gilt damit:

$$\mathbf{J}_s (\mathbf{J}_s^\top \mathbf{J}_s)^{-1} \mathbf{J}_s^\top \mathbf{J}_i = \mathbf{Q}_s \mathbf{Q}_s^\top \mathbf{J}_i \quad (7.10)$$

sodass Gl. (7.9) vereinfacht werden kann zu:

$$\mathcal{C}(\Delta \mathbf{p}_I) = \Delta \mathbf{p}_I^\top \underbrace{[\mathbf{J}_i^\top \mathbf{J}_i - (\mathbf{J}_i^\top \mathbf{Q}_s)(\mathbf{Q}_s^\top \mathbf{J}_i)]}_{\mathbf{C}} \Delta \mathbf{p}_I \quad (7.11)$$

Eine orthonormale Basis \mathbf{U}_s des Unterraums von \mathbf{J}_s kann auch über die Singulärwertzerlegung $\mathbf{J}_s = \mathbf{U}_s \mathbf{S}_s^\top \mathbf{V}_s^\top$ bestimmt werden. Alternativ kann der Teilausdruck $(\mathbf{J}_s^\top \mathbf{J}_s)^{-1} \mathbf{J}_s^\top \mathbf{J}_i$ als Lösung des Gleichungssystems $\mathbf{J}_s^\top \mathbf{J}_s \mathbf{x} = \mathbf{J}_s^\top \mathbf{J}_i$ ermittelt werden. Diese Vorgehensweise ist effizienter, aber numerisch etwas weniger stabil.

Für die effektive Änderung des Abbildungsmodells $\mathcal{K}(\Delta \mathbf{p}_I)$ kann analog eine quadratische Näherung der Form $\mathcal{K}(\Delta \mathbf{p}_I) = \Delta \mathbf{p}_I^\top \mathbf{K} \Delta \mathbf{p}_I$ bestimmt werden.

Damit kann nun das Maß für die Beobachtbarkeit B dargestellt werden als:

$$B = \min_{\Delta \mathbf{p}_I} \left\{ \frac{\Delta \mathbf{p}_I^\top \mathbf{C} \Delta \mathbf{p}_I}{\Delta \mathbf{p}_I^\top \mathbf{K} \Delta \mathbf{p}_I} \right\} \quad (7.12)$$

Minimierungsprobleme dieser Form treten zum Beispiel im Rahmen der Diskriminanzanalyse auf [Kie95]. Durch die Substitution $\mathbf{y} = \mathbf{K}^{\frac{1}{2}} \Delta \mathbf{p}_I$ lässt sich dieses Problem vereinfachen zu:

$$B = \min_{\mathbf{y}} \left\{ \frac{\mathbf{y}^\top \mathbf{K}^{-\frac{1}{2}} \mathbf{C} \mathbf{K}^{-\frac{1}{2}} \mathbf{y}}{\mathbf{y}^\top \mathbf{y}} \right\} \quad (7.13)$$

Der Wert für B entspricht dem kleinsten Eigenwert der Matrix $\mathbf{K}^{-\frac{1}{2}} \mathbf{C} \mathbf{K}^{-\frac{1}{2}}$. Der den Ausdruck minimierende Vektor \mathbf{y} ist der zugehörige Eigenvektor. Die am schlechtesten beobachtbare Änderung des Parametervektors ist $\Delta \mathbf{p}_I = \mathbf{K}^{-\frac{1}{2}} \mathbf{y}$.

Für den Fall, dass die Matrix \mathbf{K} nicht invertierbar ist, kann alternativ das entsprechende Maximierungsproblem gelöst werden:

$$B^{-1} = \max_{\Delta \mathbf{p}_I} \left\{ \frac{\Delta \mathbf{p}_I^\top \mathbf{K} \Delta \mathbf{p}_I}{\Delta \mathbf{p}_I^\top \mathbf{C} \Delta \mathbf{p}_I} \right\} \quad (7.14)$$

Die mit den Eigenvektoren aus Gl. (7.13) bzw. Gl. (7.14) korrespondierenden Änderungen des Parametervektors $\Delta \mathbf{p}_I$ bilden eine orthogonale Basis der Änderungsmöglichkeiten des Abbildungsmodells im Optimum. Abbildung 7.3 zeigt die am schlechtesten beobachtbaren Richtungen im Bildraum des Abbildungsmodells als betragsmäßige Änderung des Blickwinkels sowie als Verschiebungsfeld.

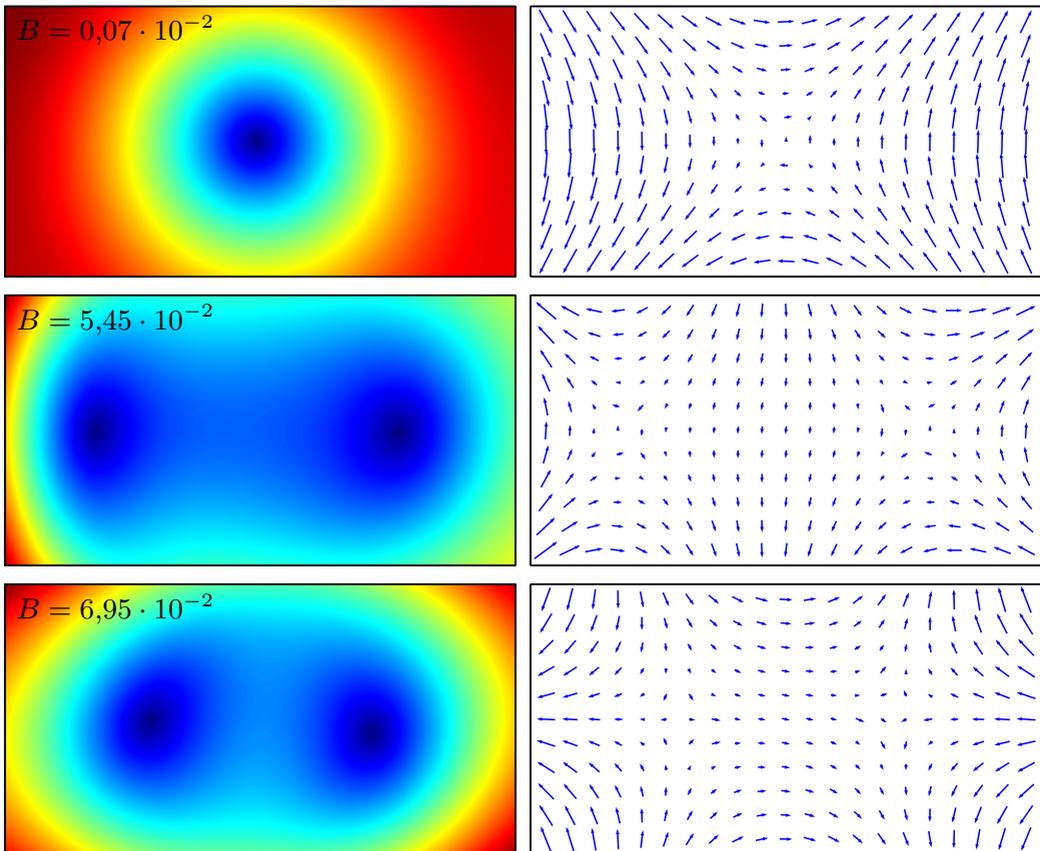


Abbildung 7.3: Basis des Tangentialraums des Abbildungsmodells. Die Eigenvektoren des für die Beobachtbarkeit formulierten Minimierungs-/Maximierungsproblems nach Gl. (7.13) bzw. Gl. (7.14) bilden eine orthonormale Basis für den tangentialen Unterraum des Abbildungsmodells im Optimum. Gezeigt sind die drei am schlechtesten beobachtbaren Richtungen im Bildraum der Kamera - von oben nach unten mit zunehmend besserer Beobachtbarkeit. Links ist die normierte betragsmäßige Änderung des Blickwinkels im Vergleich zum Blickwinkel im Optimum gezeigt sowie der zugehörige Wert der Beobachtbarkeit B , rechts die jeweiligen Blickwinkeländerungen als Verschiebungsfeld.

7.3 Beurteilung von Kalibrierkonfigurationen

In der praktischen Anwendung stellt sich die Frage, welche Modellierung bzw. Kalibrierkonfiguration zu den besten Ergebnissen führt. Diese Frage lässt sich allerdings nicht pauschal beantworten, da hierbei viele verschiedene Einflussfaktoren eine Rolle spielen – beispielsweise die Planarität der verwendeten Schachbrettmuster, deren Entfernung und Orientierung zur Kamera, die vorherrschenden Lichtverhältnisse während der Aufzeichnung der Kalibriersequenz oder das verwendete Objektiv.

Eine gute Kalibrierung zeichnet sich unter anderem durch einen geringen Rückprojektionsfehler aus. Dieser für sich genommen ist jedoch kein geeignetes Maß, da ein geringer Rückprojektionsfehler beispielsweise auch die Folge von Überanpassung sein kann. Da für Parametrisierungen mit zu vielen Freiheitsgraden in der Regel die Beobachtbarkeit nach Abschnitt 7.2 abnimmt, stellt die Beobachtbarkeit eine hilfreiche Ergänzung zur Beurteilung von Kalibrierkonfigurationen dar. Abbildung 7.4 zeigt exemplarisch den Zusammenhang von Beobachtbarkeit und mittlerem Rückprojektionsfehler für unterschiedliche Konfigurationen.

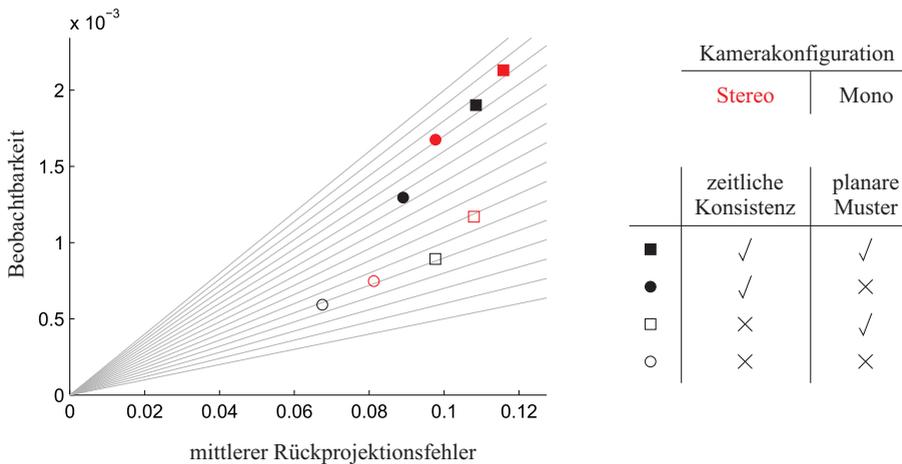


Abbildung 7.4: Einfluss der Kalibrierkonfiguration auf die Beobachtbarkeit.

Das Diagramm zeigt die Werte für Beobachtbarkeit und mittleren Rückprojektionsfehler, welche sich für unterschiedliche Kalibrierkonfigurationen ergeben. Für die Auswertung wurde die Sequenz eines Stereo-Aufbaus verwendet. Die Beobachtbarkeit bezieht sich in allen Fällen jedoch nur auf die linke Kamera. Generell ist zu beobachten, dass sowohl die Berücksichtigung der zeitlichen Konsistenz der Kalibrierstruktur als auch die Hinzunahme von Beobachtungen der zweiten Kamera die Beobachtbarkeit verbessern, aber zu einem etwas höheren Rückprojektionsfehler führen. Die grauen Linien entsprechen einem jeweils konstanten Verhältnis von Beobachtbarkeit und mittlerem Rückprojektionsfehler.

8 Anwendung der Kalibrierung

In der praktischen Anwendung von Bildverarbeitungsalgorithmen wird meist nicht direkt mit den erfassten Bilddaten gearbeitet. Stattdessen werden die Bilddaten vorab anhand der Kalibrierung entzerrt und auf ein vereinfachtes Kameramodell umgerechnet. Dies hat den Vorteil, dass im weiteren Verlauf der Bildverarbeitung keine Kenntnis über das Kalibriermodell und dessen Parametrisierung mehr erforderlich ist. Eine Umrechnung von Bilddaten zwischen verschiedenen Abbildungen mit jeweils eindeutigem Projektionszentrum ist, abgesehen von Interpolationsfehlern, verlustlos möglich, solange das Projektionszentrum beibehalten wird. Im Fall einer Abbildung ohne eindeutiges Projektionszentrum wäre für die exakte Umrechnung Kenntnis der Tiefeninformation der betrachteten Szene erforderlich, da eine Änderung des lokalen Projektionszentrums zu einer tiefenabhängigen Parallaxe führt. In Abschnitt 8.1 wird ein Verfahren vorgestellt, um für ein allgemeines Abbildungsmodell die optimale Lage des Projektionszentrums einer vereinfachenden Zentralprojektion zu bestimmen. Diese Reduktion auf eine Zentralprojektion ist zwar verlustbehaftet, jedoch erforderlich, um bereits vorhandene Bildverarbeitungsalgorithmen weiterhin anwenden zu können.

In Abschnitt 8.2 wird beschrieben, wie die Bilddaten einer allgemeinen Zentralprojektion so umgewandelt werden können, dass sie den Bilddaten eines einfach beschreibbaren Projektionsmodells genügen. In der Regel wird als Zielmodell das Modell einer Lochkamera verwendet. Prinzipiell kann aber auch jedes andere Modell einer Zentralprojektion verwendet werden. Die Verwendung eines vom Lochkameramodell abweichenden Zielmodells ist insbesondere für Weitwinkel-Objektive interessant, da für diese eine Umrechnung auf das Lochkameramodell zu starken Verzerrungen der ursprünglichen Bilddaten führt.

Die Bildentzerrung spielt auch als Vorverarbeitung für die Stereo-Tiefenschätzung eine wichtige Rolle. Zur Vereinfachung der Suche von Bildpunkt-Korrespondenzen werden hierbei die Bilder von linker und rechter Kamera so entzerrt, dass korrespondierende Punkte in beiden Kamerabildern in derselben Bildzeile zu finden sind. Dieser Schritt wird als Rektifizierung bezeichnet. Üblicherweise werden die Bilder auch hierfür so entzerrt, dass sie anschließend der Abbildung mit einer Lochkamera entsprechen (Abschnitt 8.4). In Abschnitt 8.5 wird die Rektifizierung allgemein formuliert und erläutert, welche Voraussetzungen das Zielmodell im Hinblick auf die Vereinfachung der Suche von Bildpunkt-Korrespondenzen erfüllen muss.

8.1 Vereinfachende Zentralprojektion

Die meisten der bislang verfügbaren Bildverarbeitungsalgorithmen gehen von einer Abbildung mit eindeutigem Projektionszentrum aus. Während diese Annahme für Standard-Objektive meist sehr gut erfüllt ist, weisen jedoch insbesondere Weitwinkel-Objektive in der Regel eine deutliche Verschiebung des Projektionszentrums in Abhängigkeit vom Blickwinkel auf. Diese Verschiebung liegt in etwa in der Größenordnung der Objektivabmessungen. Für Anwendungen, bei denen die betrachteten Objekte mehrere Meter weit entfernt sind, hat die Verschiebung des Projektionszentrums nur noch einen geringen Einfluss auf deren Abbildung im Kamerabild und kann oftmals vernachlässigt werden. Andererseits ist es aus Gründen der Handhabbarkeit meist günstiger, mit Kalibrieremustern im Nahfeld zu kalibrieren. Um hierbei gute Ergebnisse zu erzielen, ist es wichtig, die Verschiebung des Projektionszentrums für die Kalibrierung mit zu modellieren und erst im Anschluss das Modell auf den Fall der Zentralprojektion zu reduzieren.

Um ein allgemeines Abbildungsmodell auf den Fall einer Zentralprojektion zu vereinfachen, ist es im Grunde ausreichend, lediglich die Modellierung des Blickwinkels aus der Kalibrierung zu übernehmen. Für Multi-Kamera-Aufbauten müssen zusätzlich das Projektionszentrum der Zentralprojektion bestimmt und die extrinsischen Posen entsprechend angepasst werden. In [Sch14b] wird eine Möglichkeit beschrieben, wie die Lage des Projektionszentrums für die Abbildung mit einem katadioptrischen Objektiv bestimmt werden kann. Verallgemeinert ausgedrückt, wird hierbei das Projektionszentrums so bestimmt, dass es von den zum Kamerabild gehörenden Sichtstrahlen im quadratischen Mittel den geringsten Abstand aufweist. In Abschnitt 4.4 wurde bereits beschrieben, wie der Abstand zwischen einem Weltpunkt und der Sichtgerade ermittelt werden kann. Wird der Weltpunkt durch das gesuchte Projektionszentrum \mathbf{v} ersetzt, ergibt sich für die quadratische Norm des Lotvektors mit $\|\mathbf{r}_i\| = 1$:

$$\|\mathbf{e}_i\|^2 = \|\mathbf{v} - \mathbf{z}_i\|^2 - (\mathbf{r}_i^\top (\mathbf{v} - \mathbf{z}_i))^2 \quad (8.1)$$

Mit einer geeigneten Abtastung des Abbildungsmodells mit Sichtstrahlen $(\mathbf{z}_i + \lambda \mathbf{r}_i)$ gilt für die Wahl des Projektionszentrums:

$$\mathbf{v} = \arg \min_{\mathbf{v}'} \sum_i \left(\|\mathbf{v}' - \mathbf{z}_i\|^2 - (\mathbf{r}_i^\top (\mathbf{v}' - \mathbf{z}_i))^2 \right) \quad (8.2)$$

Hieraus folgt die Lösungsgleichung:

$$\sum_i (\mathbf{v} - \mathbf{z}_i - \mathbf{r}_i (\mathbf{v} - \mathbf{z}_i)^\top \mathbf{r}_i) = \mathbf{0} \quad (8.3)$$

Mit der Einheitsmatrix \mathbf{I} folgt durch Umformung von Gl. (8.3) das lineare Gleichungssystem:

$$\left[\sum_i (\mathbf{I} - \mathbf{r}_i \mathbf{r}_i^\top) \right] \mathbf{v} = \sum_i (\mathbf{I} - \mathbf{r}_i \mathbf{r}_i^\top) \mathbf{z}_i \quad (8.4)$$

Dessen Lösung ist die Lage des Projektionszentrums für das vereinfachte Modell.

Um für die vereinfachende Zentralprojektion das Projektionszentrum in den Ursprung zu verschieben, muss der Translationsanteil der extrinsischen Pose entsprechend angepasst werden:

$$\mathbf{t}'_E = \mathbf{t}_E - \mathbf{v} \quad (8.5)$$

Bei Annahme weit entfernter Objekte erfordert die Vereinfachung auf eine Zentralprojektion keine Umwandlung von Bilddaten (Abschnitt 8.2), da die Blickwinkel unverändert bleiben.

8.2 Bildentzerrung

Das Ziel der Bildentzerrung ist die Umrechnung der von einer Kamera erfassten Bilddaten auf ein anderes Abbildungsmodell. Sehr häufig wird als Zielmodell für die Umrechnung das Lochkamera-Modell verwendet, da dieses sehr einfach und mit wenigen Parametern beschrieben werden kann. Ohne Kenntnis über die Entfernung der erfassten Objekte ist eine exakte Umrechnung nur dann möglich, wenn das Ausgangs-Abbildungsmodell ein eindeutiges Projektionszentrum aufweist.

Die Bildentzerrung kann als eine zweidimensionale Interpolation aufgefasst werden, die jedem Punkt (u, v) im Zielbild denjenigen Punkt (u_i, v_i) des Ausgangsbildes zuweist, welcher einen gewünschten Blickwinkel $\theta(u, v)$, $\varphi(u, v)$ abbildet. Die Funktionen $\theta(u, v)$ und $\varphi(u, v)$ sind hierbei durch das für die Bildentzerrung vorgegebene Zielmodell definiert. Die Größen θ und φ zur Beschreibung des Blickwinkels sind in Abbildung 8.1 nochmals veranschaulicht. Mit Bildhauptpunkt (u_0, v_0) gilt für ein rotationssymmetrisches Zielmodell:

$$\varphi(u, v) = \arctan \left(\frac{v - v_0}{u - u_0} \right) \quad (8.6)$$

Mit dem Bildradius

$$r(u, v) = \sqrt{(u - u_0)^2 + (v - v_0)^2} \quad (8.7)$$

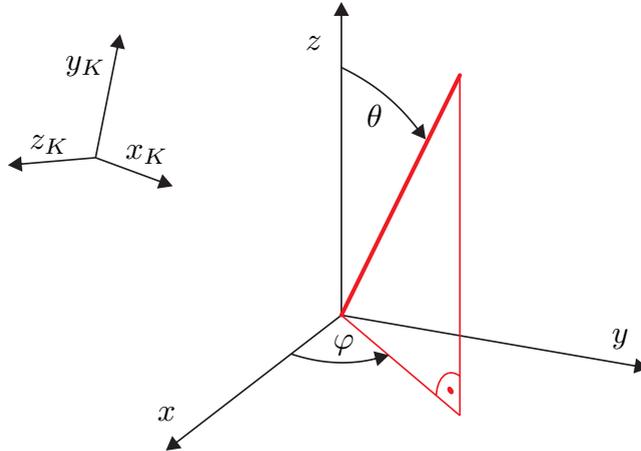


Abbildung 8.1: Bezugs-Koordinatensystem für die Bildentzerrung. Veranschaulichung der Größen θ und φ zur Beschreibung des Blickwinkels in Kugelkoordinaten. Die Orientierung der optischen Achse z im ursprünglichen Kamera-Koordinatensystem (x_K, y_K, z_K) kann beliebig gewählt werden. Die Verschiebung des Koordinatensystemursprungs in der Skizze dient nur der Verbesserung der Lesbarkeit. Tatsächlich kann der Koordinatensystemursprung und damit das Projektionszentrum der Abbildung nicht verändert werden, da ansonsten bereits Tiefeninformation über die abgebildete Szene erforderlich wäre, um die Bildentzerrung durchzuführen.

gilt beispielsweise für die Abbildung einer Lochkamera mit Brennweite f :

$$\theta(u, v) = \arctan\left(\frac{r(u, v)}{f}\right) \quad (8.8)$$

Sowohl die Orientierung des Bezugs-Koordinatensystems als auch die Funktionen $\theta(u, v)$ und $\varphi(u, v)$ können prinzipiell beliebig vorgegeben werden. Für die spätere Auswertung der entzerrten Bilddaten ist es vorteilhaft, wenn der Zusammenhang zwischen θ und r nach beiden Größen aufgelöst werden kann.

Die mit einem gegebenen Blickwinkel korrespondierenden Bildpunkt-Koordinaten im Quellbild können anhand des kalibrierten Kameramodells ermittelt werden:

$$\begin{pmatrix} u_i(u, v) \\ v_i(u, v) \end{pmatrix} = \mathcal{P}(\theta(u, v), \varphi(u, v)) \quad (8.9)$$

Die beiden Funktionen $u_i(u, v)$ und $v_i(u, v)$ entsprechen jeweils 2D-Lookup-Tabellen. In Abhängigkeit von dem für die Kalibrierung verwendeten Abbildungsmodell ist zur Berechnung der Lookup-Tabellen gegebenenfalls eine iterative Suche

erforderlich. Da diese Berechnung jedoch nur einmalig für eine gegebene Kalibrierung durchgeführt werden muss, ist dies jedoch keine nennenswerte Einschränkung.

Für die Intensitäten im entzerrten Zielbild \mathcal{U} gilt in Abhängigkeit von den Intensitäten des Quellbildes \mathcal{I} :

$$\mathcal{U}(u, v) = \mathcal{I}(u_i(u, v), v_i(u, v)) \quad (8.10)$$

Je nach verfügbarer Rechenleistung und erforderlicher Qualität der entzerrten Bilddaten können für das Auslesen des Quellbildes im Zuge der Bildentzerrung verschiedene Interpolationsverfahren verwendet werden. Normalerweise wird die bilineare oder bikubische Interpolation verwendet.

Eine beispielhafte Bildentzerrung für unterschiedliche Zielmodelle ist in Abbildung 8.2 zu sehen. Zuschneiden des Zielbildes anhand des einschließenden Rechtecks führt je nach gewähltem Zielmodell zu einer großen Einschränkung des Sichtbereichs. Falls für die weitere Verarbeitung nicht zwingend eine Abbildung nach dem Lochkameramodell benötigt wird, ist es daher sinnvoll, ein Zielmodell zu wählen, welches dem der tatsächlichen Abbildung möglichst ähnlich ist.



Abbildung 8.2: Bildentzerrung für unterschiedliche Zielmodelle. Entzerrung eines Kamerabilds unter Verwendung des Lochkamera-Modells (links) sowie eines Kameramodells mit winkeltreuer (Mitte) bzw. flächentreuer Abbildungsfunktion (rechts) als Vorgabe für das Zielmodell. Der Bereich mit gültigem Bildinhalt im Zielbild ist in der Regel nicht mehr rechteckig, sondern kissen- oder tonnenförmig. Oftmals wird das Zielbild daher anhand des eingeschlossenen Rechtecks beschnitten (rot umrandet). Dies führt allerdings zu einer teilweise nicht unerheblichen Einschränkung des Sichtbereichs, sodass es sinnvoll ist ein Zielmodell zu wählen, welches dem der tatsächlichen Abbildung bereits möglichst gut entspricht.

8.3 Erzeugung virtueller Vogelperspektiven

Eine Anwendung für Multi-Kamera-Systeme ist die Rundumsicht für Fahrerassistenzsysteme oder autonomes Fahren. Neben der Erfassung anderer Verkehrsteilnehmer ist für die Fahrzeugführung insbesondere die Straßengeometrie von Interesse. Diese lässt sich am besten anhand einer virtuellen Vogelperspektive untersuchen. Für deren Berechnung können unter Annahme einer Bodenebene die Ansichten der einzelnen Kameras entsprechend transformiert und zusammengesetzt werden. Für das Beispiel in Abbildung 8.3 wurden die Parameter der Bodenebene von Hand optimiert. Grundsätzlich kann die Ebene aber auch anhand der Stereo-Information im Überlappungsbereich der Kameras geschätzt werden.

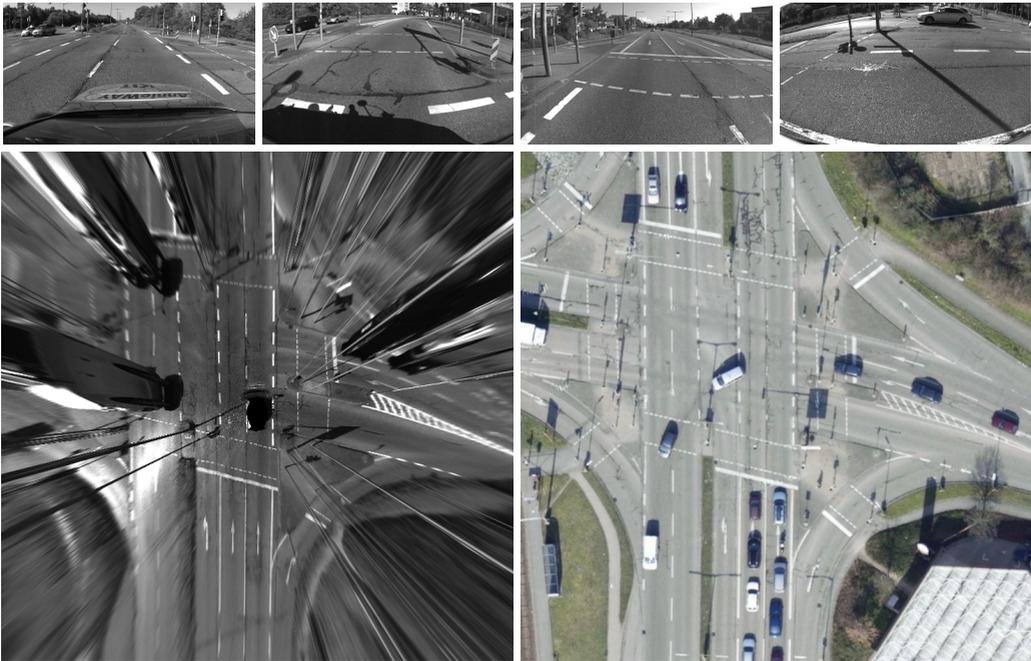


Abbildung 8.3: Berechnung einer virtuellen Vogelperspektive. Aus den Bildern von rund um ein Fahrzeug herum angeordneten Kameras (oben) kann anhand der Multi-Kamera-Kalibrierung eine virtuelle Vogelperspektive errechnet werden (unten links). Hierzu werden die Projektionen der Punkte aus einer vorgegebenen Bodenebene in allen Kameras berechnet, um zu ermitteln, wo und in welchen Kamerabildern der entsprechende Punkt sichtbar ist. Ist die Bodenebene im Kamerabild durch Objekte verdeckt, so ergeben sich in diesen Bereichen der Vogelperspektive näherungsweise radial strahlenförmige Störungen. Zur visuellen Bewertung der berechneten Vogelperspektive ist unten rechts der dem Kreuzungsbereich entsprechende Ausschnitt aus einer Luftbildaufnahme der Stadt Karlsruhe gezeigt.

Für die Berechnung der Vogelperspektive wird zunächst ein regelmäßiges Gitter auf der geschätzten Bodenebene erzeugt. Anhand der Multi-Kamera-Kalibrierung werden dann die Punkte dieser Ebene in die Koordinatensysteme der einzelnen Kameras transformiert und durch Anwendung des jeweiligen Abbildungsmodells ermittelt, wo und in welchen Kamerabildern der entsprechende Punkt sichtbar ist. Da konkrete 3D-Punkte projiziert werden, ist es, im Gegensatz zur Bildentzerrung (Abschnitt 8.2), nicht erforderlich, dass die Abbildungsmodelle ein eindeutiges Projektionszentrum besitzen. Da ohne Stereo-Information nur schwer validiert werden kann, ob ein Bildpunkt im Kamerabild die Bodenebene oder ein Objekt zeigt, wird für alle Bildpunkte die Bodenebene angenommen. Dadurch entstehen in Bildbereichen, in denen die Annahme der Bodenebene nicht zutreffend ist, näherungsweise radial strahlenförmige Störungen im Bild der Vogelperspektive.

8.4 Rektifizierung nach dem Lochkameramodell

Bei einem Kameraaufbau mit sich überlappenden Sichtbereichen existieren Weltpunkte, deren Projektionen in mehreren Kamerabildern zu sehen sind. Die Bildpunkte desselben Weltpunkts in den verschiedenen Kamerabildern werden als korrespondierende Bildpunkte bezeichnet. Bei bekannter Kalibrierung kann aus einer gefundenen Bildpunkt-Korrespondenz die Entfernung des zugehörigen Weltpunkts rekonstruiert werden. In diesem Zusammenhang beschreibt die Rektifizierung eine spezielle Bildentzerrung, die das Ziel hat, das Auffinden von Bildpunkt-Korrespondenzen zu vereinfachen.

Allgemein ist der mit der Projektion eines Weltpunkts korrespondierende Punkt auf einer Kurve im Bild der anderen Kamera zu finden. Genügen die beiden Kameras dem Abbildungsmodell einer Lochkamera, so können die Korrespondenzen über die Epipolargeometrie [Har04] beschrieben werden. In diesem Fall existieren korrespondierende Epipolarlinien, die jeweils paarweise aufeinander abgebildet werden. Durch eine geeignete Ausrichtung der Lochkameras ist es darüber hinaus möglich, die Epipolarlinien entlang der Bildzeilen auszurichten. Eine solche Konfiguration ergibt sich, wenn die optischen Achsen beider Kameras parallel zueinander sind und senkrecht auf der Verbindungslinie der beiden Projektionszentren stehen.

In der Praxis ist es nicht möglich, direkt einen Kameraaufbau zu erreichen, der diesen Bedingungen genügt, da die Objektive in der Regel eine gewisse Verzeichnung aufweisen und eine exakte Ausrichtung der Kameras sehr aufwändig wäre. Für den Fall, dass die Kameras ein eindeutiges Projektionszentrum besitzen, ist es möglich, durch Anwendung einer bestimmten Bildentzerrung aus den Eingangsbildern die virtuellen Ansichten gemäß der idealen Kamerakonfiguration zu berechnen. Dieser

Schritt wird als Rektifizierung bezeichnet. Die Rektifizierung ist insbesondere für die dichte Stereo-Tiefenschätzung ein elementarer Vorverarbeitungsschritt, da die Ausrichtung der Epipolarlinien entlang der Bildzeilen die Korrespondenzen-Suche erheblich vereinfacht.

Das Lochkameramodell ist zur vereinfachenden Beschreibung von Kameraabbildungen sehr weit verbreitet. Seine Anwendung im Zusammenhang der Rektifizierung besticht insbesondere bei der Auswertung von Bildpunkt-Korrespondenzen durch ihre Einfachheit, da die Entfernung nur von der gemessenen Disparität abhängig ist. Die Disparität d bezeichnet hierbei den Pixel-Abstand korrespondierender Bildpunkte $d = u_l - u_r$ zwischen linkem und rechtem Kamerabild entlang der Suchzeile. Mit der Basisbreite B , welche den Abstand der Projektionszentren angibt, und der Brennweite f der rektifizierten Lochkamera gilt zur Berechnung der Entfernung eines Objekts der invers proportionale Zusammenhang

$$z = \frac{B \cdot f}{d} \quad (8.11)$$

wobei z die Entfernung in Richtung der optischen Achse angibt.

8.5 Allgemeine Rektifizierung

Für die Vereinfachung der Korrespondenzen-Suche entlang von Bildzeilen ist es nicht zwingend erforderlich, dass die rektifizierten Bilder dem Abbildungsmodell einer Lochkamera genügen. Entscheidend ist lediglich, dass alle einer Bildzeile zugeordneten Sichtstrahlen in einer gemeinsamen Epipolarebene liegen. Dies ist genau dann der Fall, wenn diese Sichtstrahlen eine Ebene aufspannen, welche auch die beiden Kamera-Projektionszentren beinhaltet. Die Auffächerung der einzelnen Sichtstrahlen innerhalb dieser Epipolarebene spielt für die Vereinfachung der Korrespondenzen-Suche keine Rolle. Für die Rektifizierung auf ein vom Lochkameramodell abweichendes Zielmodell wird im Folgenden der Begriff der allgemeinen Rektifizierung verwendet.

Für die Beschreibung der allgemeinen Rektifizierung wird zunächst ein Koordinatensystem eingeführt, auf dessen x -Achse die Projektionszentren der beiden Kameras liegen – die Rotation des Koordinatensystems bezüglich dieser Achse kann beliebig gewählt werden, wobei die z -Achse im Folgenden als Bezugsachse verwendet wird. In diesem Koordinatensystem können die Punkte einer Epipolarebene sehr einfach anhand des euklidischen Abstands ρ zum Projektionszentrum, des Winkels γ der Epipolarebene zur z -Achse und des Winkels β zur Projektion

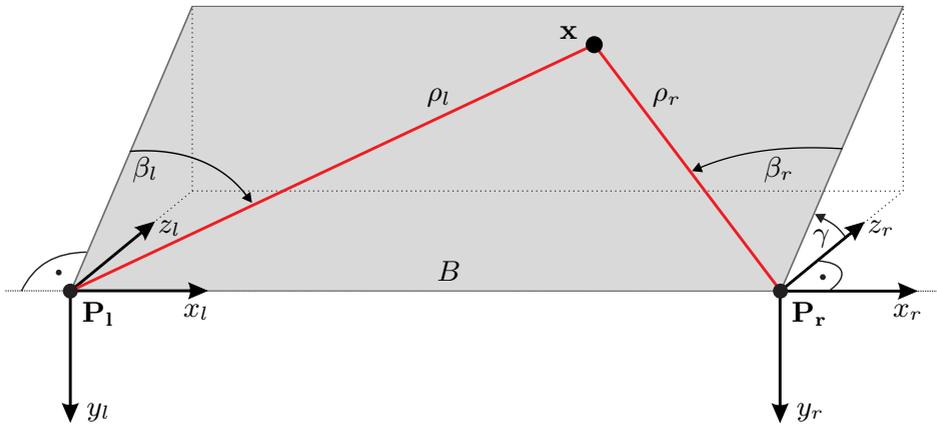


Abbildung 8.4: Allgemeine Rektifizierung. Grundidee für die Rektifizierung ist, dass alle Punkte einer Epipolarebene (grau) auf eine Bildzeile abgebildet werden. Alle Epipolarebenen beinhalten die x -Achse der rektifizierten Koordinatensysteme, welche durch die beiden Projektionszentren \mathbf{P}_l und \mathbf{P}_r definiert ist. Der Abstand der beiden Projektionszentren voneinander ist mit B bezeichnet. Ein Punkt \mathbf{x} kann über den Epipolarwinkel γ , den Winkel β sowie den euklidischen Abstand ρ zum Projektionszentrum beschrieben werden. Es gilt: $\beta_l < 0 < \beta_r$. Der Winkel zwischen Sichtstrahl und x -Achse beträgt $\beta + \frac{\pi}{2}$.

der z -Achse auf die Epipolarebene parametrisiert werden (siehe Abbildung 8.4):

$$\mathbf{x} = \rho \begin{pmatrix} \sin \beta \\ -\cos \beta \sin \gamma \\ \cos \beta \cos \gamma \end{pmatrix} \quad (8.12)$$

Damit alle Punkte der Epipolarebene in einer Bildzeile liegen, darf der Winkel γ nur von der Bildzeile v im rektifizierten Bild abhängen, während der Winkel β beliebig in Abhängigkeit von u und v gewählt werden kann:

$$\gamma := \mathcal{G}(v) \quad \beta := \mathcal{H}(u, v) \quad (8.13)$$

Für den Fall einer bijektiven Abbildung gilt ebenso:

$$v := \tilde{\mathcal{G}}(\gamma) \quad u := \tilde{\mathcal{H}}(\gamma, \beta) \quad (8.14)$$

Ein Spezialfall der allgemeinen Rektifizierung ist die sphärische Rektifizierung [Sch14a] mit $v \sim \gamma$ und $u \sim \beta$.

Es ist nicht zwingend erforderlich, für linke und rechte Kamera dieselbe Zuordnung $\beta := \mathcal{H}(u, v)$ zu wählen. Allerdings ist für unterschiedlich gewählte Zuordnungen bei der Beschreibung des gültigen Suchraums zu beachten, dass Punkte im Unendlichen nicht mehr die Disparität Null besitzen.

Für einen Punkt, der unter den Winkeln β_l und β_r beobachtet wird, kann der euklidische Abstand zum linken bzw. rechten Projektionszentrum nach dem Sinussatz berechnet werden:

$$\rho_l = \frac{B \cos \beta_r}{\sin(\beta_r - \beta_l)} \quad \rho_r = \frac{B \cos \beta_l}{\sin(\beta_r - \beta_l)} \quad (8.15)$$

Für die Entfernung in Richtung der Bezugsachse gilt dann:

$$z = \rho_l \cos \beta_l \cos \gamma = \rho_r \cos \beta_r \cos \gamma \quad (8.16)$$

In einem bereits rektifizierten Bild können die einzelnen Bildzeilen unabhängig voneinander nochmals nichtlinear transformiert werden, ohne die Vereinfachung einer zeilenweisen Suche zu beeinflussen. Eine Gruppe möglicher Transformationen sind Scherung und Streckung in horizontaler Richtung. Dies wird beispielsweise für Stereo-Blockmatching auf geneigten Ebenen genutzt [BR14] oder auch zur Vereinfachung der Korrespondenzen-Suche im trinokularen Fall [An04]. Neben der Transformation der einzelnen Bildzeilen kann auch die Abtastung in Spaltenrichtung (senkrecht zu den Epipolarlinien) beliebig angepasst werden.

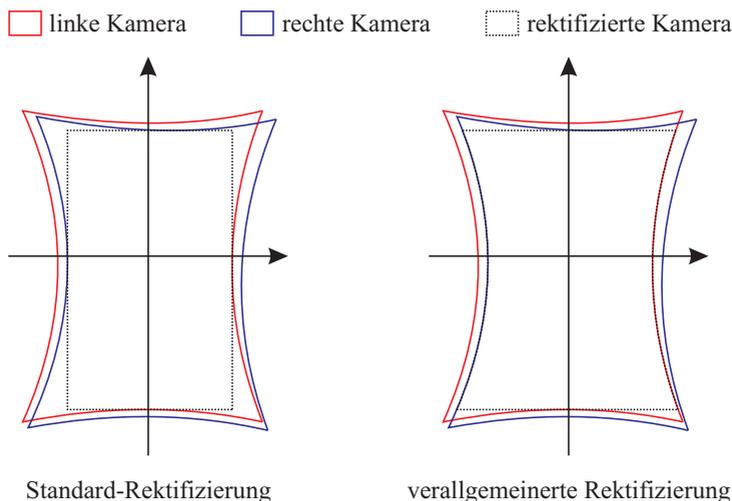


Abbildung 8.5: Bessere Ausnutzung des erfassten Sichtbereichs. Veranschaulichung der Sichtbereiche beider Kameras in einem gemeinsamen Lochkamera-Koordinatensystem. Durch Aufheben der Beschränkung auf das Lochkamera-Modell bei Verwendung der allgemeinen Rektifizierung kann insbesondere für Weitwinkel-Objektive der Sichtbereich teilweise deutlich erweitert werden. Dieser Effekt wird besonders sichtbar, wenn wie im Beispiel oben die Epipolarlinien entlang der kurzen Bildseite verlaufen.

Vorteile der allgemeinen Rektifizierung

Insbesondere bei Weitwinkel-Objektiven entsteht durch die Rektifizierung nach dem Lochkameramodell oftmals eine deutliche Einschränkung des Sichtbereichs, falls nur der rechteckige Bereich mit vollständig gültigen Bilddaten verwendet wird. Ähnlich wie bei der Bildentzerrung (Abschnitt 8.2) kann durch eine entsprechend angepasste Rektifizierung dieser Effekt verringert werden (siehe Abbildung 8.5). Neben einer besseren Erhaltung des Sichtbereichs kann durch die allgemeine Rektifizierung auch die Verzerrung der Eingangs-Bilddaten reduziert werden, indem ein Zielmodell gewählt wird, welches besser mit dem tatsächlichen Kameramodell übereinstimmt. So lassen sich beispielsweise durch eine Anpassung der sphärischen Rektifizierung insbesondere im Randbereich nahe den Epipolen deutliche Verbesserungen erzielen (siehe Abbildung 8.6). Die Zuordnungsvorschriften der beiden Ansichten genügen jeweils dem Modell nach Gl. (8.13). Die Zuordnungsvorschrift der optimierten allgemeinen Rektifizierung wurde manuell anhand des visuellen Eindrucks gewählt.



Abbildung 8.6: Einfluss der Rektifizierung auf die Bildverzerrungen. Rektifizierung von zwei nebeneinander angeordneten virtuellen Rundumsicht-Kameras mit Ausrichtung der Epipolarlinien in Zeilenrichtung. Gezeigt ist jeweils nur das Bild von einer der beiden Kameras – simulativ erzeugt anhand einer sphärischen Textur. Links: sphärische Rektifizierung, rechts: manuell anhand des visuellen Eindrucks optimierte allgemeine Rektifizierung. Durch Anpassung der Abtastung in Spaltenrichtung und nichtlineare Transformation der einzelnen Bildzeilen werden insbesondere die Randbereiche links und rechts weniger stark verzerrt.

9 Zusammenfassung und Ausblick

In der vorliegenden Arbeit wurde ein Verfahren zur Kalibrierung von Multi-Kamera-Systemen vorgestellt. Die für die Kalibrierung erforderlichen Beobachtungen von Schachbrettecken werden hierbei voll-automatisch extrahiert und global assoziiert. Durch die Verwendung von Kalibriersequenzen mit Relativbewegung zwischen Kamera-System und Kalibrierobjekt und global assoziierten Beobachtungen können auch Multi-Kamera-Systeme kalibriert werden, bei denen die Sichtbereiche der einzelnen Kameras nicht überlappend sind.

Für die Extraktion der Schachbrettecken werden basierend auf den Merkmalsbildern für Gradientenorientierung und Symmetrie zunächst potentielle Schachbrettecken-Pixel detektiert. Durch eine anschließende Vernetzung der Ecken zu Schachbrettkacheln können hierbei auftretende Fehldetektionen weitgehend eliminiert werden. Entscheidend für die Robustheit des Gesamtsystems ist jedoch primär die daran anschließende globale Assoziation der Beobachtungen anhand eines die Schachbrettmuster umgebenden Binärcodes, da sie eine sehr zuverlässige Validierung der Beobachtungen ermöglicht.

Zur Beschreibung des Abbildungsverhaltens der einzelnen Kameras wurde ein universelles Kameramodell entwickelt, welches neben verschiedenen Abbildungsfunktionen auch eine blickwinkelabhängige Verschiebung des Projektionszentrums berücksichtigt. Damit kann neben dem Abbildungsverhalten von konventionellen Objektiven auch das von Weitwinkel-Optiken und katadioptrischen Objektiven beschrieben werden.

Die verwendete Formulierung des vollständigen Bündelausgleichs mit global assoziierten Beobachtungen ermöglicht eine in sich konsistente Modellierung unter Verwendung einer Minimal-Parametrierung. Hierbei werden beispielsweise die Anordnung und Form der Schachbrettmuster, die Anordnung der Kameras zueinander sowie deren intrinsisches Abbildungsverhalten für alle Beobachtungen gemeinsam modelliert. Die Hinzunahme zusätzlicher Beobachtungen aus weiteren Zeitschritten der Bildsequenz führt dann lediglich zur einer zusätzlich zu schätzenden Pose für die Bewegung des Kamera-Systems. Dadurch können effektiv viele Beobachtungen gewonnen und eine gute Sensor-Abdeckung erzielt werden, ohne dass die Anzahl der zu schätzenden Parameter übermäßig ansteigt.

Im vorletzten Kapitel der Arbeit wurde ein Ansatz vorgestellt, um die numerischen Eigenschaften des Bündelausgleichs im Hinblick auf die Beobachtbarkeit quantita-

tiv zur erfassen und verschiedene Kalibrierkonfigurationen miteinander vergleichen zu können. Einige Anwendungen zeigen schließlich das Potential auf, welches sich für neuartige Kamerasysteme aus der Möglichkeit von deren Kalibrierung ergibt, wie beispielsweise eine optimierte Stereo-Rektifizierung und die Berechnung einer virtuellen Vogelperspektive.

Bislang werden für die Kalibrierung die Beobachtungen global assoziierter Schachbrettecken verwendet. Diese ermöglichen eine präzise Kalibrierung. In der Praxis ergibt sich jedoch häufig das Problem, dass der Kameraaufbau geringen zeitlichen Veränderungen durch Erschütterungen, Wärmeausdehnung etc. unterliegt. Daher ist es naheliegend, die Kalibrierverfahren künftig dahingehend zu ergänzen, dass ein Teil der Kalibrier-Parameter im Betrieb nachgeführt wird. Hierzu können prinzipiell beliebige Weltpunkte (oft auch als Landmarken bezeichnet) verwendet werden, solange deren Beobachtungen zeitlich bzw. zwischen den verschiedenen Kameras assoziiert werden können. Für die Modellierung nicht-planarer Schachbrettmuster wurden die einzelnen Schachbrettecken bereits als regularisierte Landmarken behandelt. Im Gegensatz zu den Schachbrettecken werden die freien Landmarken direkt, ohne Verwendung einer zusätzlichen Pose, durch deren Position in einem globalen Koordinatensystem beschrieben. Da ohne Regularisierung auch die Möglichkeit entfällt, Skalen-Information mit einzubeziehen, muss zumindest einer der Abstände zwischen den Kameras konstant gehalten oder anderweitig Skalen-Information zugeführt werden.

Neben einer Einbeziehung natürlicher Landmarken in den Kalibrierprozess stellt für die praktische Anwendung auch die kombinierte extrinsische Kalibrierung bezüglich anderer Sensorsysteme wie beispielsweise Inertialsensorik, Laserscanner etc. eine interessante Erweiterungsmöglichkeit dar.

Für den Einsatz von Farb-Kameras sollte das Kameramodell noch um einen farbkanal-abhängigen Anteil des Blickwinkels erweitert werden, um die insbesondere bei Weitwinkel-Objektiven häufig vorhandenen, durch chromatische Aberration bedingten, Farbsäume kompensieren zu können. Hierzu muss die Extraktion der Kalibriermarker für die einzelnen Farbkanäle getrennt voneinander durchgeführt werden. Durch die Kompensation der chromatischen Aberration kann unter Umständen mit Farb-Kameras sogar eine bessere Detailauflösung erreicht werden als bei Verwendung von Schwarz-Weiß-Kameras, bei denen sich die chromatische Aberration nur in Form von Unschärfe bemerkbar macht.

Literaturverzeichnis

- [Abr97] S. Abraham und T. Hau: *Towards autonomous high-precision calibration of digital cameras*. In: *Optical Science, Engineering and Instrumentation '97*, S. 82–93, International Society for Optics and Photonics, 1997.
- [Abr05] S. Abraham und W. Förstner: „Fish-eye-stereo calibration and epipolar rectification“. *ISPRS Journal of photogrammetry and remote sensing* **59** (5), S. 278–288, 2005.
- [Aga10] S. Agarwal, N. Snavely, S. M. Seitz und R. Szeliski: *Bundle adjustment in the large*. In: *Computer Vision–ECCV 2010*, S. 29–42. Springer, 2010.
- [An04] L. An, Y. Jia, J. Wang, X. Zhang und M. Li: *An efficient rectification method for trinocular stereovision*. In: *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, Bd. 4, S. 56–59, IEEE, 2004.
- [Bak02] H. Bakstein und T. Pajdla: *Panoramic mosaicing with a 180 field of view lens*. In: *Omnidirectional Vision, 2002. Proceedings. Third Workshop on*, S. 60–67, IEEE, 2002.
- [Bar08] J. Barreto, J. Santos, P. Menezes, F. Fonseca et al.: *Ray-based calibration of rigid medical endoscopes*. In: *The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras-OMNIVIS*, 2008.
- [Bes92] P. Besl und N. D. McKay: „A method for registration of 3-D shapes“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **14** (2), S. 239–256, Feb 1992.
- [BR14] T. S. Benjamin Ranft: *Hocheffizientes Stereosehen durch Modellierung geneigter Ebenen und Parallelverarbeitung*. In: *9. Workshop Fahrerassistenzsysteme (FAS 2014)*, S. 27–36, 2014.
- [Cap90] B. Caprile und V. Torre: „Using vanishing points for camera calibration“. *International journal of computer vision* **4** (2), S. 127–139, 1990.

- [Car11] G. Carrera, A. Angeli und A. J. Davison: *SLAM-based automatic extrinsic calibration of a multi-camera rig*. In: *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, S. 2652–2659, IEEE, 2011.
- [Cli08] B. Clipp, J.-H. Kim, J.-M. Frahm, M. Pollefeys und R. Hartley: *Robust 6dof motion estimation for non-overlapping, multi-camera systems*. In: *Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on*, S. 1–8, IEEE, 2008.
- [Dan09] T. Dang, C. Hoffmann und C. Stiller: „Continuous stereo self-calibration by camera parameter tracking“. *Image Processing, IEEE Transactions on* **18** (7), S. 1536–1550, 2009.
- [Ern99] S. Ernst, C. Stiller, J. Goldbeck und C. Roessig: *Camera calibration for lane and obstacle detection*. In: *Intelligent Transportation Systems, 1999. Proceedings. 1999 IEEE/IEEJ/JSAI International Conference on*, S. 356–361, IEEE, 1999.
- [Esq07] S. Esquivel, F. Woelk und R. Koch: *Calibration of a multi-camera rig from non-overlapping views*. In: *Pattern Recognition*, S. 82–91. Springer, 2007.
- [Fau92] O. D. Faugeras, Q.-T. Luong und S. J. Maybank: *Camera self-calibration: Theory and experiments*. In: *Computer Vision–ECCV 1992*, S. 321–334, Springer, 1992.
- [Fle95] M. M. Fleck: „Perspective projection: the wrong imaging model“. *Department of Computer Science, University of Iowa* S. 1–27, 1995.
- [Fus00] A. Fusiello, E. Trucco und A. Verri: „A compact algorithm for rectification of stereo pairs“. *Machine Vision and Applications* **12** (1), S. 16–22, 2000.
- [Gei12] A. Geiger, F. Moosmann, O. Car und B. Schuster: *Automatic Calibration of Range and Camera Sensors using a single Shot*. In: *International Conference on Robotics and Automation (ICRA)*, 2012.
- [Gen06] D. B. Gennery: „Generalized camera calibration including fish-eye lenses“. *International Journal of Computer Vision* **68** (3), S. 239–266, 2006.
- [Glu01] J. Gluckman und S. K. Nayar: *Rectifying transformations that minimize resampling effects*. In: *Computer Vision and Pattern Recognition*,

2001. *CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Bd. 1, S. I–111, IEEE, 2001.
- [Gri00] A. Griewank: *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*. Nr. 19 in *Frontiers in Appl. Math.*. SIAM, Philadelphia, PA, 2000.
- [Har04] R. I. Hartley und A. Zisserman: *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, Second. Aufl., 2004.
- [Hei06] M. Heinrichs und V. Rodehorst: *Trinocular rectification for various camera setups*. In: *Symp. of ISPRS Commission III-Photogrammetric Computer Vision PCV*, Bd. 6, S. 43–48, Citeseer, 2006.
- [iso95] „Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM)“. *ISO/IEC Guide 98-3:2008* 1995.
- [Kie95] H. Kiers: „Maximization of sums of quotients of quadratic forms and some generalizations“. *Psychometrika* **60**, S. 221–245, 1995.
- [Kum08] R. K. Kumar, A. Ilie, J.-M. Frahm und M. Pollefeys: *Simple calibration of non-overlapping cameras with a mirror*. In: *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, S. 1–7, IEEE, 2008.
- [Lev44] K. Levenberg: „A method for the solution of certain problems in least squares“. *Quarterly of applied mathematics* **2**, S. 164–168, 1944.
- [Mar63] D. W. Marquardt: „An algorithm for least-squares estimation of non-linear parameters“. *Journal of the Society for Industrial & Applied Mathematics* **11** (2), S. 431–441, 1963.
- [Mei07] C. Mei und P. Rives: *Single View Point Omnidirectional Camera Calibration from Planar Grids*. In: *ICRA*, 2007.
- [Mir13] P. Miraldo und H. Araujo: „Calibration of Smooth Camera Models“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **35** (9), S. 2091–2103, 2013.
- [Pol99a] M. Pollefeys, R. Koch und L. Van Gool: „Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters“. *International Journal of Computer Vision* **32** (1), S. 7–25, 1999.

- [Pol99b] M. Pollefeys, R. Koch und L. Van Gool: *A simple and efficient rectification method for general motion*. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Bd. 1, S. 496–501, IEEE, 1999.
- [Sca06] D. Scaramuzza und A. Martinelli: *A Toolbox for Easily Calibrating Omnidirectional Cameras*. In: *IROS*, 2006.
- [Sch14a] M. Schoenbein und A. Geiger: *Omnidirectional 3D Reconstruction in Augmented Manhattan Worlds*. In: *International Conference on Intelligent Robots and Systems (IROS)*, 2014.
- [Sch14b] M. Schoenbein, T. Strauss und A. Geiger: *Calibrating and Centering Quasi-Central Catadioptric Cameras*. In: *International Conference on Robotics and Automation (ICRA)*, 2014.
- [Shi73] Y. Shimizu: *Wide Angle Fisheye Lens*, Juni 5 1973. US Patent 3,737,214.
- [Str14] T. Strauss, J. Ziegler und J. Beck: *Calibrating Multiple Cameras with Non-overlapping Views Using Coded Checkerboard Targets*. In: *International Conference on Intelligent Transportation Systems (ITSC)*, 2014.
- [Stu99] P. F. Sturm und S. J. Maybank: *On plane-based camera calibration: A general algorithm, singularities, applications*. In: *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, Bd. 1, IEEE, 1999.
- [Tsa87] R. Y. Tsai: „A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses“. *Robotics and Automation, IEEE Journal of* **3** (4), S. 323–344, 1987.
- [Wen92] J. Weng, P. Cohen und M. Herniou: „Camera calibration with distortion models and accuracy evaluation“. *IEEE Transactions on pattern analysis and machine intelligence* **14** (10), S. 965–980, 1992.
- [Xio97] Y. Xiong und K. Turkowski: *Creating image-based VR using a self-calibrating fisheye lens*. In: *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, S. 237–243, IEEE, 1997.
- [Zha00] Z. Zhang: „A flexible new technique for camera calibration“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **22** (11), S. 1330–1334, Nov 2000.

-
- [Zha04a] Z. Zhang: *Camera Calibration*. In: *Emerging topics in computer vision*, G. Medioni und S. B. Kang (Hrsg.), S. 4–43. Prentice Hall PTR, 2004.
- [Zha04b] Z. Zhang: „Camera calibration with one-dimensional objects“. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **26** (7), S. 892–899, 2004.