
Genauere und robuste Online-Kalibrierung für
Stereo-Kamera-Systeme mit fester Basisbreite

Diplom-Inf. Anton Feldmann

Dissertation

12. April 2016

Dissertation zur Erlangung des akademischen Grades *Doktors der Ingenieurwissenschaften*
(Dr.-Ing.) an der technischen Fakultät der Universität Bielefeld.

Anton Feldmann
email: anton.feldmann@gmail.com

Universität Bielefeld
Technische Fakultät
Arbeitsgruppe für angewandte Informatik
Universitätsstr. 25
33615 Bielefeld

Danksagung

Ich danke Herrn Professor Dr. Rainer Ott für sein Engagement, mich zum Schreiben zu bewegen und seine wertvollen Anmerkungen zu meinen Arbeiten.

Meinem Doktorvater Herrn Professor Dr. Franz Kummert möchte ich für seine Anmerkungen danken, die mich stets zu weiteren Ideen und zum weiteren Schreiben motiviert haben.

Ich danke Herrn Dr. Lars Krüger, der mich zum Schreiben motiviert und mich in meinen Ideen unterstützt hat. Herr Dr. Krüger hatte immer einen kritischen Blick auf die angewandten Methoden und hat mich sehr mit seinen Fähigkeiten im Bereich der 3D-Modellierung zusammen mit dem von ihm betreuten Studenten Herrn Christian Werner unterstützt.

Herr Dr. Frank Lindner hat mich ermutigt, alle meine Konzepte und Ideen kritisch zu beleuchten und diese aus einem anderen Blickwinkel zu betrachten.

Einen besonderen Dank möchte ich den von mir betreuten Studenten aussprechen. Andrea-Stephania Herta, Peter Kabiri, Pengfei Wang und Di Chen haben mich sehr in den Evaluierungen und in der Implementierung unterstützt. Timo Bailer hat sich sehr in die Idee der „Optimalen Positionierung des Kalibrationstargets für die Initialkalibrierung“ eingebracht. Matthias Rapp spreche ich besonderen Dank aus, da er unter meiner Betreuung eine neue Art der Online-Kalibrierung von Stereo-Kameras untersucht hat und dies in einer Masterarbeit niederschrieb.

Dank spreche ich auch allen meinen Kollegen bei der Daimler AG, der Universität Bielefeld und der Continental AG aus.

Abschließend möchte ich meiner Familie danken. Meinem Vater Anton, meiner Mutter Theresa, meinem Bruder Thomas und meiner Schwester Regina Theresia danke ich sehr für ihre Zuneigung und ihre Unterstützung.

Inhaltsverzeichnis

1. Einleitung	2
1.1. Systemaufbau für eine Online-Kalibrierung für Stereo-Kamera-Systeme . . .	10
1.2. Die Bedeutung der Online-Kalibrierung für Stereo-Kamera-Systeme . . .	17
2. Stand der Technik	20
2.1. Korrespondenzbestimmung	20
2.2. Ausreißerdetektion	22
2.3. Methoden zur Online-Kalibrierung	27
2.4. Triangulierung	32
3. Initialkalibrierung	33
4. Genaue und robuste Online-Kalibrierung	38
4.1. Online-Kalibrierung auf Basis von Entfernungsdaten	40
4.2. Online-Kalibrierung mit Bezug auf die Fluchtpunkte in beiden Bildern . .	43
4.3. Analyse von Einflüssen	48
5. Experimentelle Untersuchungen	59
5.1. Einfluss der Komponenten einer Korrespondenz auf die Kalibrierung . . .	62
5.2. Verteilung der Korrespondenzen	66
5.3. Korrespondenzauswahl mit Binning entspricht der Gleich-Verteilung . . .	71
5.4. Korrespondenzauswahl mit der <i>inneren Genauigkeit</i>	77
5.5. Rauscheinfluss auf die Online-Kalibrierung	82
5.6. Die Auswertung der Online-Kalibrierungsmethoden	85
6. Zusammenfassung und Ausblick	96
A. Anhang	99
A.1. Die erforderliche Rechenleistung auf dem DSP zur Online-Kalibrierung .	99
A.2. Herleitung des Prädiktionsintervalls	99
A.3. Kalibrierung des Velodyne-Laserscanners und des Stereo-Kamera-Systems	100
A.4. Differentialgeometrischer Ansatz für die Scheibenverzeichnung	102
Literaturverzeichnis	109

Zusammenfassung

In modernen Fahrzeugen werden verschiedene Fahrerassistenzsysteme genutzt, die den Fahrer während der Fahrt unterstützen. Durch den Einsatz von Fahrerassistenzsystemen wird die Fahrt sicherer und die Anzahl der Unfälle deutlich reduziert.

Um die Umgebung zu erfassen und Verkehrsteilnehmer, Hindernisse oder die Straßenoberfläche zu erkennen, werden verschiedene Sensoren genutzt. Die Kamera ist einer der Sensoren, die in diesen Fahrzeugen eingesetzt werden. In den Bilddaten können Objekte detektiert und klassifiziert werden.

Durch die großen Fortschritte in der Mikroelektronik und den Einsatz kleiner, leistungsfähiger Rechner ist der Einsatz eines Stereo-Kamera-Systems für neue Fahrerassistenzsysteme möglich. Aus den Sensordaten der Stereo-Kamera kann eine dreidimensionale Rekonstruktion der Fahrzeugumgebung erstellt werden. Durch die Umgebungsrekonstruktion kann die Position und die Orientierung von Objekten relativ zum Fahrzeug bestimmt werden. Mit diesen zusätzlichen Informationen wird die Bestimmung des Typs eines Objekts in der Szene verbessert oder gar erst ermöglicht.

Mit einer Stereo-Kamera kann eine 3D-Rekonstruktion der Fahrzeugumgebung errechnet werden. Um diese 3D-Rekonstruktion zu erstellen, wird eine Triangulierung genutzt. Selbst kleine Abweichungen in den Parametern der Stereo-Kamera führen zu merkbaren Abweichungen von der tatsächlichen Position relativ zu den Weltkoordinaten in dieser 3D-Rekonstruktion. Diese Abweichungen führen dann zu Fehlern in der Ausführung der eingesetzten Fahrerassistenzsysteme. Um diese Abweichungen minimal zu halten, wird das Stereo-Kamera-System initial kalibriert. Mit einer Initialkalibrierung eines Stereo-Kamera-Systems werden die intrinsischen Parameter (Brennweite, Bildhauptpunkt, Seitenverhältnis) und die extrinsischen Parameter (relative Positionierung und Orientierung der Kameras zueinander) bestimmt. Die Kameraparameter der Stereo-Kamera müssen während der gesamten Fahrzeit überwacht und angepasst werden, da sich diese Parameter des Kamerasystems durch die thermischen und mechanischen Einflüsse während der Fahrt verändern. Um eine kontinuierliche Überprüfung der Kameraparameter zu sichern, wird eine Online-Kalibrierung genutzt. Die Herausforderung der Online-Kalibrierung eines Stereo-Kamera-Systems in einem Fahrzeug ist die kontinuierliche Überprüfung der Stereo-Kamera-Parameter, sodass die 3D-Rekonstruktion der Fahrzeugumgebung genau und fehlerfrei berechnet werden kann. Wissenschaftliche Untersuchungen an realen

Stereo-Kamera-Systemen haben gezeigt, dass die relative Orientierung beider Kameras in einem Stereo-Kamera-System die größte Änderung durch die thermischen und mechanischen Einflüsse erfährt. Weitere Untersuchungen haben gezeigt, dass die Veränderungen der intrinsischen Parameter sehr klein sind und sich durch kleine Änderungen der Orientierungsparameter (ϕ_1, ϕ_2, ϕ_3) approximieren lassen. Dabei ist

$$\phi_j = \phi_{j_i} - \phi_{j_o} \text{ wobei } j = 1, 2, 3 \quad (1)$$

mit i dem Wert für den Winkel j , der durch die Initialkalibrierung bestimmt wurde, der Referenzwert und o dem entsprechenden Winkel aus der Online-Kalibrierung. Dabei ist ϕ_1 die Änderung des Nick-, ϕ_2 die Änderung des Schiel- und ϕ_3 die Änderung des Rollwinkels. In der Literatur ist keine Online-Kalibrierung für Stereo-Kameras bekannt, die mit den gegebenen Anforderungen die Änderungen der Kameraparameter berücksichtigt und eine hochgenaue 3D-Rekonstruktion über die Lebensdauer des Fahrzeug sicherstellt.

Im Rahmen dieser Arbeit werden zwei Methoden zur Online-Kalibrierung vorgestellt, die unter den gegebenen Bedingungen eines Fahrerassistenzsystems rentabel realisiert werden können. Wissenschaftliche Untersuchungen im Rahmen dieser Arbeit haben gezeigt, dass der Schielwinkel bei kleinen Abweichungen ϕ_2 durch mechanische oder thermische Einflüsse am Stereo-Kamera-System die größte Abweichung in der 3D-Rekonstruktion verursacht. Beide Methoden zur Online-Kalibrierung nutzen die Odometriedaten des Fahrzeugs. Diese Daten sind zum einen der Fahrzeug-Gierwinkel, der einen Einfluss auf genutzte Methoden zur Korrespondenzbestimmung hat, und zum anderen die Fahrzeug-Geschwindigkeit. Die Sensordaten aus dem Steuergerät („Elektronisches Stabilitätsprogramm“, ESP) stellen diese Odometriedaten zur Verfügung. Die Daten aus dem ESP-Steuergerät haben ein Rauschen, das bei Berechnung der Online-Kalibrierung berücksichtigt werden muss.

Die erste Methode zur Online-Kalibrierung nutzt die Information über die Gierrate und über die Geschwindigkeit des Fahrzeugs bei der Bildaufnahme. Wie schon erwähnt, kann aus dem ESP-Steuergerät aus der Information über die Geschwindigkeit die Strecke zwischen zwei Bildaufnahmen bestimmt werden. So kann in der 3D-Rekonstruktion die Distanz zu den vorausfahrenden Fahrzeugen und Objekten im sichtbaren Bereich des Stereo-Kamera-Systems auf Basis der zeitlich aufeinanderfolgenden Rekonstruktionen berechnet werden. Stimmt die Distanzangabe aus den Fahrzeugdaten mit der Distanz aus den 3D-Rekonstruktionen überein, ist das Kamerasystem kalibriert, ansonsten muss das Kamerasystem neu kalibriert werden.

Die zweite Methode hingegen nutzt die Fahrzeuginformation, um zu bestimmen, wann das Fahrzeug geradeaus fährt. Bei einer Fahrt geradeaus schneiden sich alle Vektoren aus der optischen Flussberechnung in einem Fluchtpunkt. Untersuchungen haben gezeigt, dass Fahrzeuge häufig geradeaus fahren und so ist das zeitliche Intervall zwischen

Kalibrierungen klein. Wird der Fluchtpunkt in beiden Bildern bestimmt, so müssen die Koordinaten dieser Fluchtpunkte übereinstimmen. Diese Übereinstimmung folgt aus der Projektion der Korrespondenzen vom linken ins rechte Bild. Stimmen diese Koordinaten nicht überein, dann gibt die Differenz der Koordinaten in der horizontalen Bildausrichtung die Abweichung im Schielwinkel an. Die Fluchtpunkte in den beiden Bildern werden in der zweiten Methode mit einem Histogramm bestimmt. Dieses Histogramm bildet die Schnittpunkte der Flussvektoren über die horizontale Koordinate ab.

Die erste Methode wurde auf der Basis von fünf Stunden Videomaterial getestet. Das entspricht rund 300 000 Bildpaaren, die für die Analyse genutzt wurden. Da das Stereo-Kamera-System, wie es ins Fahrzeug beispielhaft integriert wurde, für diese Arbeit sehr spät zur Verfügung stand, wurden rund 250 000 Bildpaare mit einem Versuchsaufbau aufgezeichnet. Abschließend wurde eine Untersuchung mit einem Nachtsicht-Stereo-Kamera-System durchgeführt, bei dem die hier vorgestellten Methoden zur Online-Kalibrierung robust eingesetzt werden konnten. Bei allen Stereo-Kamera-Systemen und allen entwickelten Methoden im Test wurde ein Fehler von weniger als $0.2 [Pix]$ Abweichung zwischen der Soll- und Ist-Position des Pixels auf dem Imager nach der Kalibrierung gemessen. Dabei wurde eine Standardabweichung von $0.1 [Pix]$ zur Ground Truth gemessen. Dies ist eine Verbesserung gegenüber allen Methoden, die in der Literatur angegeben sind und von uns getestet wurden. Dies entspricht einer Verbesserung von mehreren Metern in der 3D-Rekonstruktion in der Arbeitsdistanz jedes der drei Kamera-Systeme.

Die Methoden, die im Rahmen dieser Arbeit zu Vergleichszwecken nachimplementiert wurden, sind in [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] beschrieben. Diese nachimplementierten Methoden wurden auf den gleichen Sequenzen getestet und haben eine Abweichung zur Ground Truth von mehr als $0.8 [Pix]$. Die Methoden zur Kalibrierung des Stereo-Kamera-Systems, die im Rahmen dieser Arbeit entwickelt wurden, haben eine Abweichung von $0.19 [Pix]$ für die erste Methode und $0.12 [Pix]$ für die Methode 2.

1. Einleitung

Die sehr großen Fortschritte der Mikroelektronik, der Computertechnologie und der Sensortechnologie sowie der damit verbundenen preiswerten Bereitstellung sehr hoher Rechenleistung und Speicherkapazität auf sehr kleinem Raum ermöglicht es erst, intelligente Systeme in Fahrzeuge zu integrieren. Durch die Verfügbarkeit von Videokameras sehr kleinen Ausmaßes und höchster Bildqualität sowie großer Rechenleistung ist der Einsatz eines Stereo-Kamera-Systems mit der identischen Rechenleistung wie im Versuchsaufbau möglich. Ein allgemeiner Versuchsaufbau, mit dem die Methoden für neue Assistenzsysteme entwickelt werden, ist in Abb. 1.1 dargestellt. Die Algorithmen werden auf einem Versuchsrechner entwickelt, um in kurzer Zeit die Methoden und die eingesetzten Algorithmen zu untersuchen. Die Entwicklung findet immer mit Blick auf zukünftige Entwicklungen in der Mikroelektronik und Sensortechnologie statt.

Die Stereo-Kamera, die nach dem Entwicklungsprozess der Assistenzsysteme entstanden ist, verarbeitet die identischen Daten wie das Versuchssystem. Durch die Fortschritte in der Mikroelektronik kann die Geometrie des Entwicklungssystems aus Abb. 1.1 mit einem Volumen von ca. $0.5 [m^3]$ bei einer Breite von ca. $1.60 [m]$ auf ein Volumen von ca. $0.000648 [m^3]$ bei einer Breite von ca. $0.27 [m]$, wie in Abb. 1.2, verringert werden. Durch die eingesetzte Hardware in dem Stereo-Kamera-System ist die Leistung in der Verarbeitung der eingesetzten Methoden des Versuchsaufbaus mit der Leistung des entwickelten Stereo-Kamera-Systems vergleichbar.

Basis eines jeden Assistenzsystems sind Sensoren, die das Fahrzeugumfeld erfassen und eine Vielzahl von Daten liefern. Diese Daten müssen ausgewertet, verarbeitet und interpretiert werden. Das Ziel dieser Schritte ist das Erfassen der aktuellen Umgebungssituation des Fahrzeugs mit den relevanten Objekten wie vorausfahrende Fahrzeuge, Fußgänger oder Radfahrer, um nur einige zu nennen. Durch die korrekte Interpretation der Fahrzeugumgebung mit den relevanten Objekten kann ein Assistenzsystem optimal eingesetzt werden.

Die eingesetzten Sensoren, die das Fahrzeugumfeld überwachen, werden in aktive und passive Sensoren eingeteilt. Zu den aktiven Sensoren gehören die Ultraschall- und die Radar-Sensoren. Ein passiver Sensor ist z.B. die Kamera. Das Radar, die Ultraschallsensoren und die Kameras decken verschiedene geometrische Bereiche um das Fahrzeug ab. Diese überwachten Bereiche bei einem beispielhaften Sensoraufbau eines Fahrzeugs sind in Abb. 1.3 dargestellt.

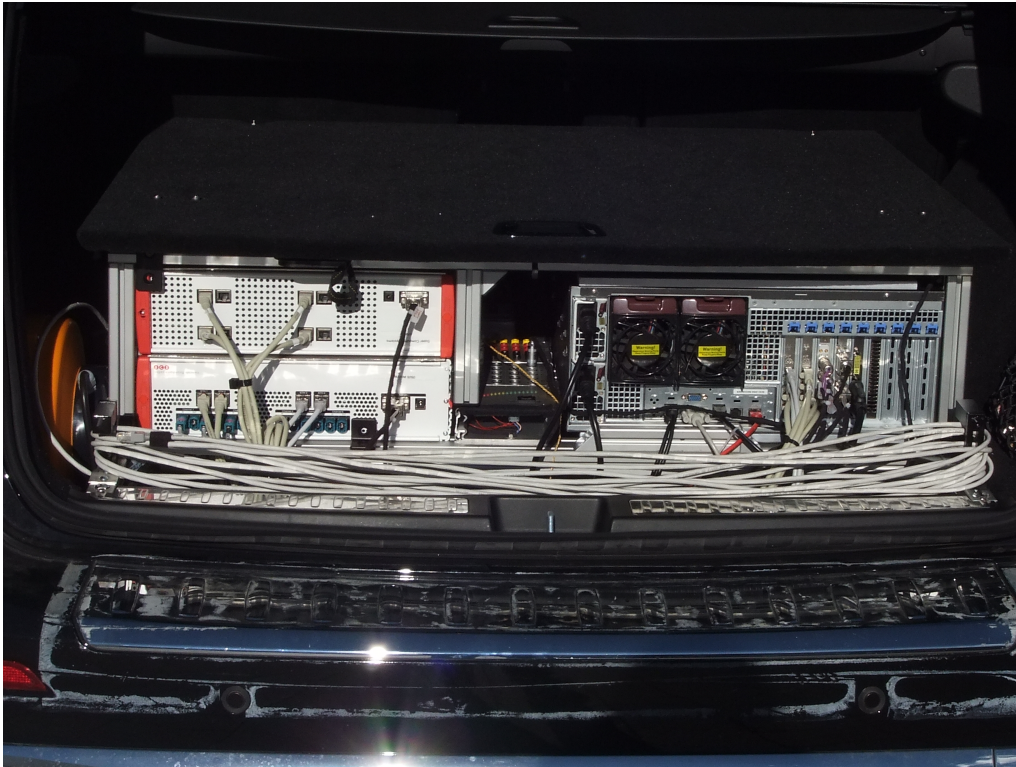


Abbildung 1.1.: Der Versuchsaufbau für die Entwicklung der Assistenzsysteme, die ins Fahrzeug integriert werden, hat ein Volumen von ca. $0.5 [m^3]$ mit einer Breite von ca. $1.60 [m]$. Damit die Echtzeit der Verarbeitung aller Methoden garantiert ist, werden spezielle Hardwarekomponenten eingesetzt. Diese eingesetzten Hardwarekomponenten unterstützen die Entwicklung, indem auf den Komponenten häufig genutzte Algorithmen integriert sind. Die auf der Hardware realisierten Methoden sind bezüglich Speicher und Geschwindigkeitsbedarf sowie auf Kamerageometrie optimiert.

Die Daten der Sensoren sind in einem Verbund zu einer Fusion zusammengeschlossen. Eine Fusion beschreibt das Verarbeiten von unterschiedlichen Sensordaten, die zu unterschiedlichen Zeitpunkten zur Verfügung stehen und zu einem gemeinsamen Ergebnis zusammengefasst werden. So können die Nachteile der einzelnen Sensoren in einer Fusion durch andere Sensoren ausgeglichen werden. Im Falle der Assistenzsysteme wird durch die Fusion die Erkennungsleistung der einzelnen Sensoren in einem Sensorverbund erhöht, um auf stationäre und bewegte Objekte in der Fahrzeugumgebung, die zu einer kritischen oder gefährlichen Situation führen können, schnell zu reagieren.

In Tab. 1.1 sind einige Beispiele moderner Fahrerassistenzsysteme aufgelistet.

Mit der Erweiterung des Mono-Kamera-Systems zu einem Stereo-Kamera-System eröffnen sich die für Assistenzsysteme so fundamental wichtigen Auswertungen und Analysen des dreidimensionalen Fahrzeugumgebungsszenarios. Diese Erweiterung bezieht sich auf das Berechnen von metrischen Größen aller Objekte, die in beiden Kamerabildern des Stereo-Kamera-Sensors sichtbar sind. Die Berechnung der metrischen Distanzen basiert auf der Brennweite, der Basisbreite und der Disparität. Das Produkt von Brennweite und

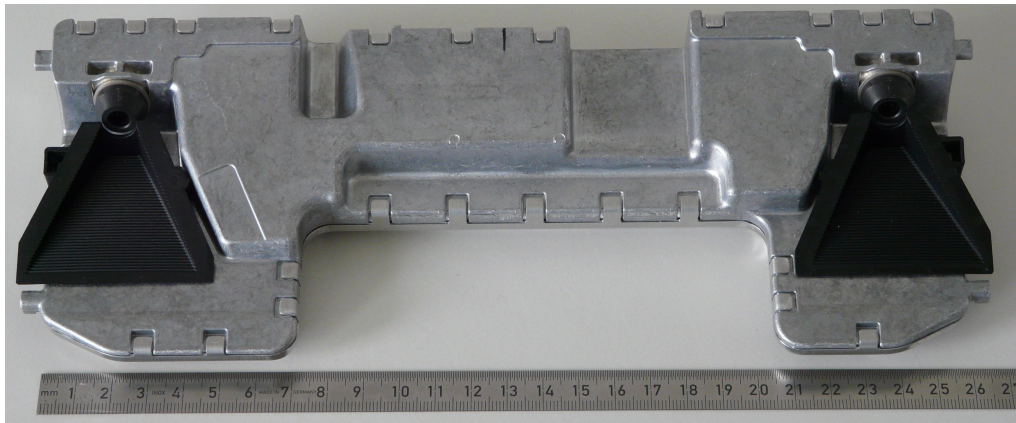


Abbildung 1.2.: Durch den Einsatz spezieller Mikrocontroller im Stereo-Kamera-System ist die Leistung in der Ausführung der Methoden in beiden Systemen vergleichbar. Alle Methoden, die auf dem Versuchssystem entwickelt wurden, werden auf den Schaltkreisen und Mikrocontrollern des Stereo-Kamera-Systems realisiert. So konnte das Volumen auf ca. $0.000648 [m^3]$ bei einer Breite des Stereo-Kamera-Systemaufbaus von ca. $0.27 [m]$.

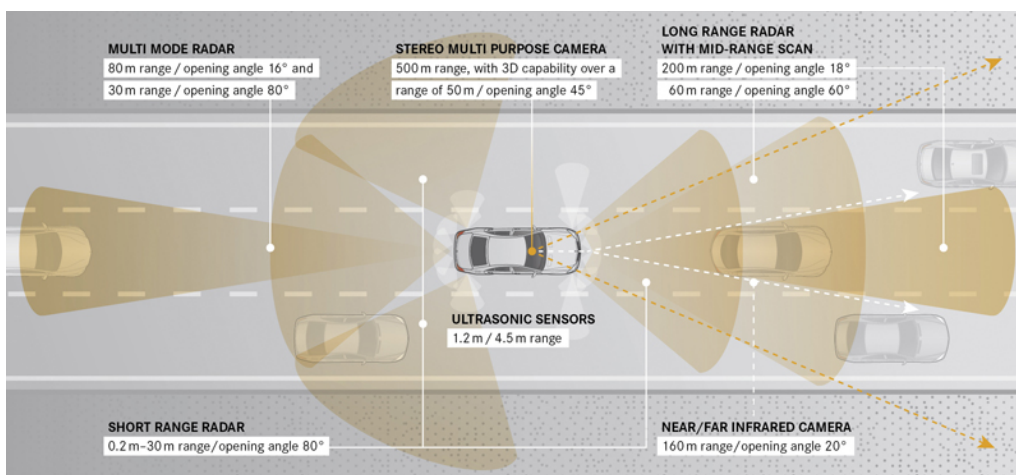


Abbildung 1.3.: Moderne Fahrzeuge nutzen eine Vielzahl verschiedener Sensoren, die zur Unterstützung des Fahrers eingesetzt werden. Mit der Fusion der Sensordaten kann die Umgebung des Fahrzeugs analysiert werden, sodass die relevanten Objekte für den Fahrer in der Umgebung bestimmt werden können. Bildquelle: [Daimler, 2013]

Basisbreite wird Kamerakonstante der Stereo-Kamera genannt. Durch die Bestimmung von Distanzen aus den Bilddaten lassen sich die Objekte in der Fahrzeugumgebung in Breite, Höhe und Entfernung bestimmen und eine robustere Klassifikation der Objekte realisieren. Auf der Basis der Auswertung der Daten der Stereo-Kamera können weitere Anwendungsszenarien analysiert werden, die mit Mono-Kameras nicht oder nur eingeschränkt möglich sind.

Verkehrszeichenerkennungssystem: Ein Verkehrszeichenerkennungssystem, wie es in modernen Fahrzeugen vorhanden ist, wird in [Lindner, 2012] beschrieben. Der Fahrer

Tabelle 1.1.: In der letzten Dekade sind viele verschiedene Assistenzsysteme auch mit neu entwickelten Sensoren in moderne Fahrzeuge integriert worden. In dieser Tabelle sind einige Assistenzsysteme angegeben, die in modernen Fahrzeugen integriert sind. Diese Aufzählung teilt die Assistenzsysteme nach den genutzten Sensoren und den Zielen des Systems ein. Der Fokus der angegebenen Systeme in dieser Tabelle liegt auf Assistenzsystemen mit optischen Sensoren. Diese Systeme werden im Folgenden näher beschrieben.

Assistenzsystem	Eingesetzte Sensorik	Ziel des Assistenzsystem
Verkehrszeichen-erkennung	Mono-Kamera	Anzeigen der vorgeschriebenen Geschwindigkeitsbegrenzungen
Spurhalteassistent	Mono-Kamera	Führen des Fahrzeugs in der Spur
Parkassistent	Mono-Kamera und Ultraschallsensor	Erleichtern des Findens von Parklücken und des Rangierens
Abstandsregeltempomat	Stereo-Kamera und Radarsensor	Einhalten des richtigen Abstandes zum vorausfahrenden Fahrzeug
Bremsassistent	Stereo-Kamera und Radarsensor	Vermeiden von Unfällen durch Querverkehr
Fußgängerwarnassistent	Stereo-Kamera	Vermeiden von Unfällen mit Fußgängern
Fahrbahnoberflächenassistent	Stereo-Kamera	Stabilisieren des Fahrzeugs trotz Fahrbahnunebenheiten
6D-Vision	Stereo-Kamera	Reagieren auf künftige Verkehrssituationen

erhält durch das System die Information über die zulässige Höchstgeschwindigkeit. Durch eine Triangulierung wird dabei die Position des Schildes relativ zum Fahrzeug bestimmt. Aber auch die Größe und die Entfernung des Schildes werden [Lindner, 2010] berechnet, um zu bestimmen, ob das Verkehrszeichen für den Fahrer relevant ist und deshalb angezeigt werden muss [Lindner, 2012]. Abb. 1.4 zeigt, dass durch die Triangulierung die Information zur Verfügung steht, die es ermöglicht zu unterscheiden, ob es sich um ein reales Zeichen handelt, das relevant ist für den Fahrer, oder um ein auf z.B. einen LKW aufgeklebtes Zeichen, das für den Fahrer nicht von Bedeutung ist.

Die Eingangsgrößen dieser Methode werden während der Fahrt bestimmt und sind durch Nicken und Wanken des Fahrzeugs fehlerbehaftet. Dadurch kann die Berechnung der Eingangsgrößen für das System ungenau werden. Durch das Nutzen der Daten der Stereo-Kamera kann die Entfernung und Größe schnell und robust aus jedem Bildpaar bestimmt werden. Somit kann durch den Einsatz eines Stereo-Kamera-Systems die Leistung des Verkehrszeichenerkennungssystems verbessert werden.

Spurhalteassistent: Das System, das das Fahrzeug in der Spur hält, kann den Fahrer



Abbildung 1.4.: Das System zur Verkehrszeichenerkennung nach [Lindner, 2012] berechnet zusätzlich zur Bedeutung die Position, die Größe und die Entfernung der Verkehrszeichen zur Kamera anhand einer Triangulierung. Bildquelle: [Lindner, 2012, S. 122]

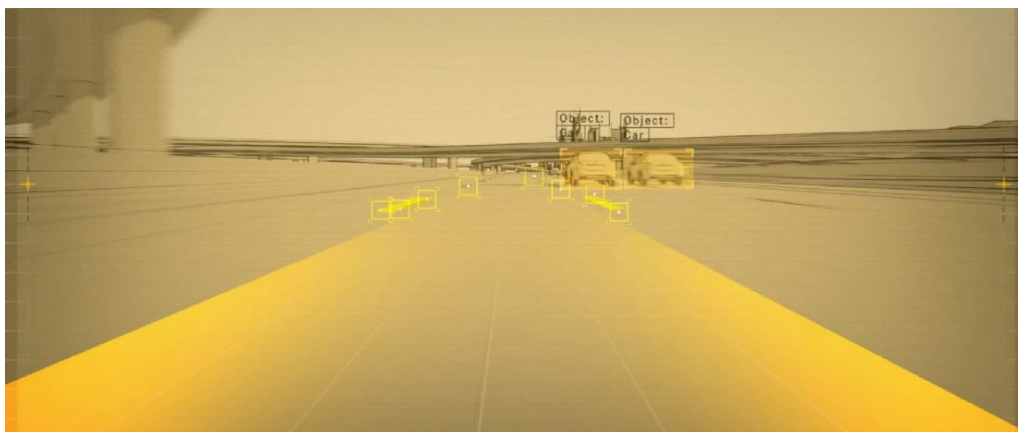


Abbildung 1.5.: Mit einer Mono-Kamera kann die Spur im nahen Umfeld des Fahrzeugs berechnet werden. So kann das Fahrzeug auf der Spur geführt werden und beim unbeabsichtigten Spurwechsel eine Warnung geben. Bildquelle: [Jordan, 2012]

vor unabsichtlichem Berühren der Fahrbahnmarkierung warnen oder dieses gegebenenfalls verhindern. Der Spurhalteassistent errechnet den Mittelpunkt der Spur mittels einer Mono-Kamera, sodass das Fahrzeug entlang dieser Spur geführt werden kann. Dies ist in Abb. 1.5 dargestellt.

Dieses Assistenzsystem kann durch den Einsatz einer Stereo-Kamera verbessert werden. Dazu wird das bestehende Klothoidenmodell auf ein dreidimensionales Modell erweitert [Loose, 2013]. Dabei wird die Spur in der Ebene zu einem Geländemodell erweitert und so die Spur dreidimensional interpretiert.

Parkassistent: Dem Fahrer wird durch den Parkassistenten das Finden und Manövrieren des Fahrzeugs in einen Parkbereich abgenommen. Dazu wird ein Freiraum detektiert, der den Ausmaßen des Fahrzeugs entspricht. Nachdem das System einen solchen Freiraum festgestellt hat, kann der Fahrer das Einparken starten. Das Finden des freien Platzes und das Einparken des Fahrzeugs wird mit Mono-Kameras und Ultraschallsensoren des Fahrzeugs durchgeführt. Wegen der geringen Auflösung der Ultraschallsensoren werden Objektdetektion und -klassifikation bei dem System mit der Mono-Kamera durchgeführt. Wird statt der Mono-Kamera eine Stereo-Kamera eingesetzt, können Objekte im Sichtbereich der Stereo-Kamera schneller und robuster erkannt werden, da dreidimensionale Ausmaße der Objekte berechnet werden können. Zudem ist es mit der Stereo-Kamera möglich, Passanten zu detektieren. Bei Gefahren für die Fahrzeuginsassen und Passanten kann der Vorgang dann abgebrochen werden.

Abstandsregeltempomat: Das System, das den Abstand zum vorausfahrenden Objekt regelt, wird Abstandsregeltempomat genannt und wird zur Sicherheit und zum Komfort des Fahrers in modernen Fahrzeugen eingesetzt. Dieses System nutzt die Daten des Radarsensors, um im weiten Umfeld des Fahrzeugs dem vorausfahrenden Fahrzeug zu folgen. Durch den Einsatz einer Stereo-Kamera können die Ergebnisse des Spurhalteassistenten mit in die Auswertung einbezogen werden, sodass selbst bei hohen Geschwindigkeiten das Fahrzeug sicher in der Spur geführt werden kann, wie in Abb. 1.6 dargestellt.

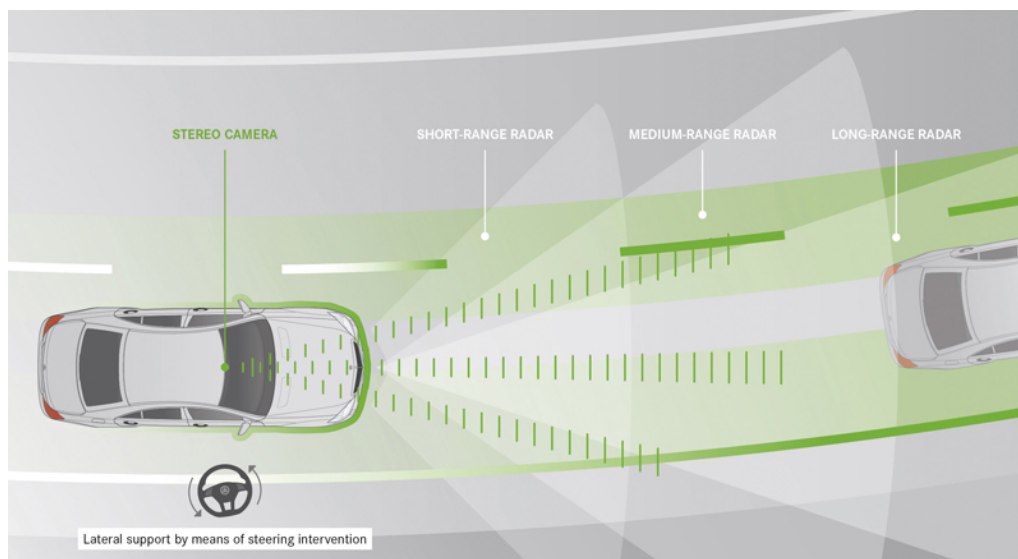


Abbildung 1.6.: Der Abstandsregeltempomat realisiert die Funktion ausschließlich auf der Radarinformation. Durch den Einsatz einer Stereo-Kamera kann der Spurhalteassistent integriert werden. Dies sorgt für eine sichere Führung innerhalb der Fahrspur. Bildquelle: [Daimler, 2013]

Insbesondere kann die Spur als 3D-Spur interpretiert werden und so die Spurführung verbessert werden.

Bremsassistent: Das Regelsystem, das das Bremsen unterstützt, reagiert auf querenden Verkehr. Um dieses System zu realisieren, werden die Sensordaten der Stereo-Kamera und des Radars zusammen ausgewertet. Mit den Radardaten ist es möglich, schon sehr früh Objekte, die in der Kamera noch nicht zu erkennen sind, zu detektieren. Mit der Mono-Kamera können die Fahrzeuge dann detektiert und klassifiziert werden. Durch den Einsatz eines Stereo-Kamera-Systems kann die Distanz zu Objekten in naher Umgebung aus dem System errechnet werden und bei einer Übereinstimmung mit den Daten aus dem Radarsystem kann der Bremsvorgang sicher eingeleitet werden.

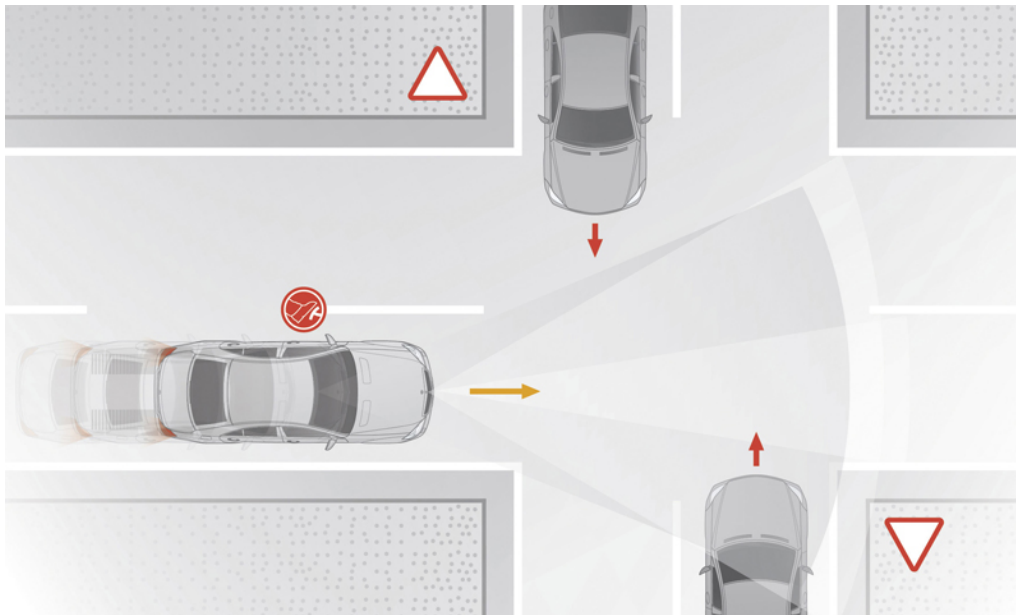


Abbildung 1.7.: Das Bremsassistentensystem unterstützt den Fahrer in nahezu unausweichlichen Gefahrensituationen, die ohne diesen Assistenten unweigerlich zum Unfall durch den querenden Verkehr führen würden. Dazu werden die Daten der Stereo-Kamera und des Radarsensors genutzt, sodass der Bremsprozess rechtzeitig und mit der richtigen Kraft eingeleitet wird. Quelle: [Daimler, 2013]

Durch die Auswertung der Daten der Stereo-Kamera kann der Assistent nicht nur Fahrzeuge, sondern auch Fußgänger erkennen und darauf reagieren.

Fußgängerwarnassistent: Mit den Daten der Stereo-Kamera zusammen mit den Daten des Radarsensors wird das Fußgängerwarnassistentensystem realisiert. Dieses System macht auf Passanten in der Fahrspur des Fahrzeugs aufmerksam und kann auf die Passanten reagieren. Dazu werden die Daten des Radarsystems und des Stereo-Kamera-Systems genutzt, wie in Abb. 1.8 dargestellt. Durch die Neuinterpretation der Spur als 3D-Spur kann der Passant besser im Bild detektiert werden, da in einem Bild der Mono-Kamera nicht ermittelt werden kann, wo der Passant steht. So ist die Detektion von Passanten in einem Mono-Kamera-Bild fehleranfälliger als die Detektion in den Daten einer Stereo-

Kamera, da der Passant in der dreidimensionalen Punktwolke aus den Daten der Stereo-Kamera gut zu erkennen ist.



Abbildung 1.8.: Der Fußgängerwarnassistent meldet, wenn ein Fußgänger durch das Fahrzeug gefährdet ist. Bildquelle: [Daimler, 2013]

Fahrbahnoberflächenassistent: Dieses Beispiel zeigt, dass die Assistenzsysteme, die die Daten der Mono-Kamera nutzen, entscheidend durch den Einsatz einer Stereo-Kamera verbessert werden können. Weiter sind durch den Einsatz von Stereo-Kamera-Systemen Assistenzsysteme mit einer Vielzahl von neuen Anwendungen realisierbar. Mit den Daten der Stereo-Kamera kann die Fahrbahnoberfläche vermessen werden, um Unebenheiten zu detektieren. Wenn Unebenheiten auf der Fahrbahnoberfläche erkannt wurden, wird durch das System das Fahrwerk aktiv in der Weise gesteuert, dass die Fahrzeugkarosse vollkommen ruhig gehalten wird und der Fahrer keine Unebenheiten bemerkt, wie in Abb. 1.9 dargestellt.

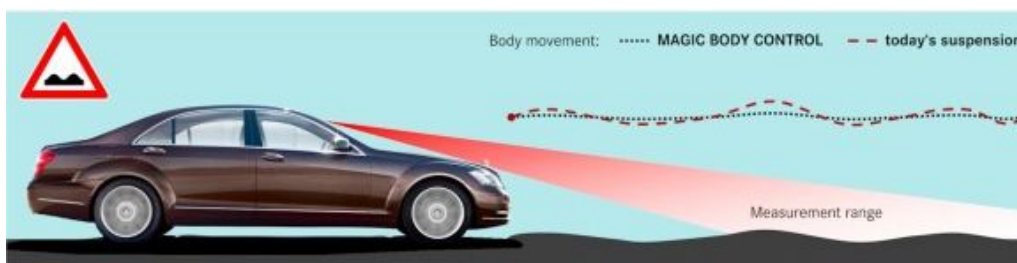


Abbildung 1.9.: Der Fahrbahnoberflächenassistent kaschiert Fahrbahnunebenheiten vor dem Fahrer, sodass sich der Fahrer wesentlich bequemer fühlt und dadurch mehr auf die Fahrt konzentrieren kann. Dabei wird das Fahrzeug auf einer Ebene gehalten. Bildquelle: [Meiners, 2010]

6D-Vision: Das 6D-Vision-System nutzt die Daten aus der Stereo-Kamera zur Berechnung der 3D-Koordinaten $(x, y, z)_{\mathcal{W}}$ im Weltkoordinatensystem \mathcal{W} und der drei Geschwindigkeitsparameter $(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}, \frac{\partial z}{\partial t})_{\mathcal{W}}$ des Punktes $(x, y, z)_{\mathcal{W}}$. Dazu wird aus den Sensordaten eine 3D-Rekonstruktion generiert. In der 3D-Rekonstruktion können sich bewegende Objekte durch den optischen Fluss als Cluster von Punkten zusammengefasst werden. Durch die Information der eigenen Odometrie und das Tracken dieser bewegten Objekte in der 3D-Rekonstruktion können die Geschwindigkeiten der Objekte aus der 3D-Rekonstruktion bestimmt werden. Da die Position der Objekte und die Geschwindigkeit der Ob-

jekte bekannt sind, kann die zukünftige Position geschätzt werden. Diese Methode wird in [Müller et al., 2011; Franke et al., 2007] vorgestellt.

1.1. Systemaufbau für eine Online-Kalibrierung für Stereo-Kamera-Systeme

Die in Tab. 1.1 beschriebenen Assistenzsysteme nutzen Sensordaten aus optischen Sensoren wie der Mono-Kamera oder der Stereo-Kamera. Alle Sensorsysteme in einem Fahrzeug müssen kalibriert werden, da sonst alle Assistenzsysteme fehlerhafte Daten erzeugen und so die gewünschten Funktionen nicht realisiert werden können.

Die Initialkalibrierung bleibt während der Fahrt wegen äußerer Einflüsse, wie z.B. Wärme oder Stöße durch das Fahrwerk, nicht erhalten. Deshalb ist eine kontinuierliche Überwachung der Kameraparameter und eine ständige Nachkalibrierung wichtig. Ein solches ständiges Nachkalibrieren wird Online-Kalibrierung genannt. Damit die Kameras während der ganzen Fahrt kalibriert sind, ist eine Online-Kalibrierung (in der Literatur auch Selbstkalibrierung genannt) essenziell.

Bei der Online-Kalibrierung werden die Parameter der Fundamentalmatrix \mathbb{F} (siehe [Faugeras, 1993] und [Wöhler, 2009, S.32,S.39])

$$\hat{x}_l^T \mathbb{F} \hat{x}_r = 0 \quad (1.1)$$

oder die Essentialmatrix \mathbb{E} aus der Essential-Gleichung

$$x_l^T \mathbb{E} x_r = 0 \quad (1.2)$$

geschätzt. Dies ist der Kern von $x_l^T \mathbb{E} x_r$. Wir schreiben dann $\ker(x_l^T \mathbb{E} x_r)$ für den Kern der Abbildung, die durch die Essential-Gleichung beschrieben ist. Dabei gilt nach [Luong, 1992]

$$\mathcal{K}^T \mathbb{E} \mathcal{K} = \mathbb{F}, \quad (1.3)$$

mit \mathcal{K} die Kameramatrix für Kameras mit rechteckigen Pixel

$$\mathcal{K} = \begin{pmatrix} f \cdot \alpha & 0 & u_o \\ 0 & f \cdot \beta & v_o \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.4)$$

mit f der Kamerakonstante $\frac{\alpha}{\beta}$ das Aspektverhältnis und $(u_o, v_o)^T$ der Bild- oder Kamerahauptpunkt. Diese Parameter werden als intrinsische Kameraparameter oder intrinsische Parameter bezeichnet. Sind die Pixel quadratisch, dann folgt für die Kamerakonstanten

$f := f_u = f_v$. Die Essentialmatrix ist $\mathbb{E} = \mathbf{T}_\times \mathbf{R}$, mit \mathbf{T}_\times die schief-symmetrische Matrix der Translation und \mathbf{R} die Rotationsmatrix. Dabei sind die Komponenten der Essentialmatrix die extrinsischen Kameraparameter oder extrinsischen Parameter, die die relative Position und die relative Orientierung der Kameras zueinander beschreiben.

Die Kalibrierung eines Kamerasystems kann in zwei Teile unterteilt werden. Zum einem müssen die intrinsischen Parameter und zum anderen müssen die extrinsischen Parameter bestimmt werden. Die intrinsischen Parameter werden im Rahmen dieser Arbeit in einer Initialkalibrierung (s. Kap. 3) bestimmt. Die Parameter, die in der extrinsischen Kalibrierung bestimmt werden, werden in dieser Arbeit in zwei Gruppen unterteilt. Die drei Parameter, die die relative Position beider Kameras in einem Stereo-Kamera-System zueinander beschreiben, werden als fest angenommen und die drei Winkeländerungen (ϕ_1, ϕ_2, ϕ_3) der Orientierung werden in einer Online-Kalibrierung ständig geschätzt.

Da alle Assistenzfunktionen auf einer gemeinsamen Hardware realisiert werden, ist die verfügbare Rechenleistung aufgeteilt, sodass jeder auf der Hardware realisierten Methode ein Teil der Gesamtrechenleistung zusteht. Da für die Online-Kalibrierung nur ein Teil der Rechenleistung zur Verfügung steht, ist eine speicher- und rechenzeiteffiziente Online-Kalibrierung eine Herausforderung im Rahmen dieser Arbeit. Genauere Angaben über die Rechenleistung für die Online-Kalibrierung sind in App. A.1 dokumentiert. In wissenschaftlichen Untersuchungen wurde gezeigt, dass sich in einem Fahrzeug die Orientierung am stärksten unter den auf das Fahrzeug einwirkenden Kräften ändert, wohingegen sich die intrinsischen Parameter nur geringfügig ändern.

Durch die Anordnung der Kameras in einer achsparallelen Position, d.h. so, dass die Z-Achse entlang der Hauptachse ausgehend vom Kamerahauptpunkt zeigt und die y-Achse zum Boden zeigt, haben kleine Abweichungen im Schielwinkel einen großen Einfluss auf die Entfernungsberechnung, da wie in [Hartley und Zisserman, 2004] beschrieben, die Epipole im Unendlichen liegen. In Kap. 4.3.1 wird dies näher untersucht. Durch diese Anordnung ist eine 3D-Rekonstruktion mit nicht-linearen Methoden zu berechnen. In Abb. 1.11 ist dargestellt, dass selbst bei kleinen Abweichungen des Schielwinkels von weniger als einem Pixel sich eine große Abweichung in der 3D-Rekonstruktion ergibt.

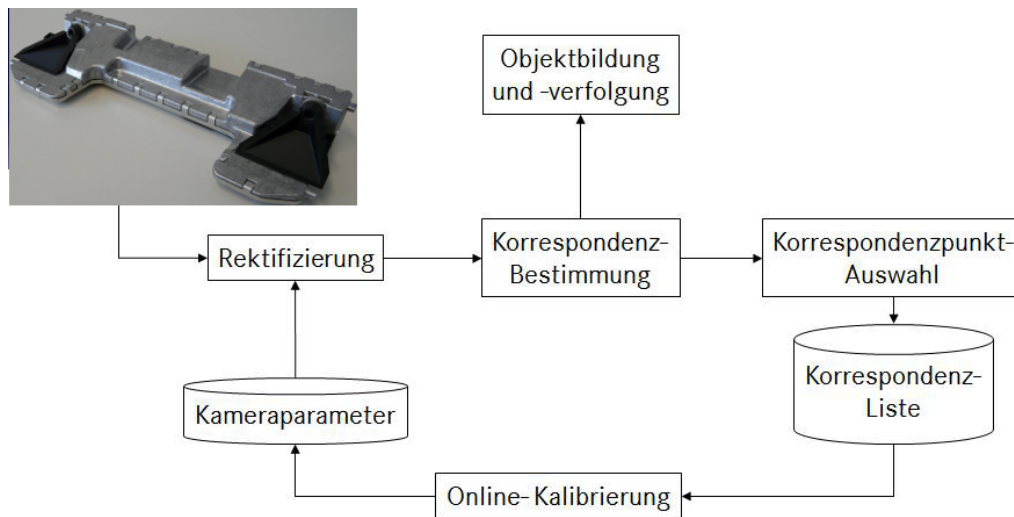


Abbildung 1.10.: Die Kameramodellparameter des Stereo-Kamera-Systems sind durch eine Initialkalibrierung ermittelt worden und liefern für die Online-Kalibrierung die initialen Werte. Mit den eingestellten Kameraparametern können mittels verschiedener Methoden ausgewählte Korrespondenzen bestimmt werden. Diese Korrespondenzen werden dann genutzt, um die aktualisierten Kameramodellparameter zu schätzen.

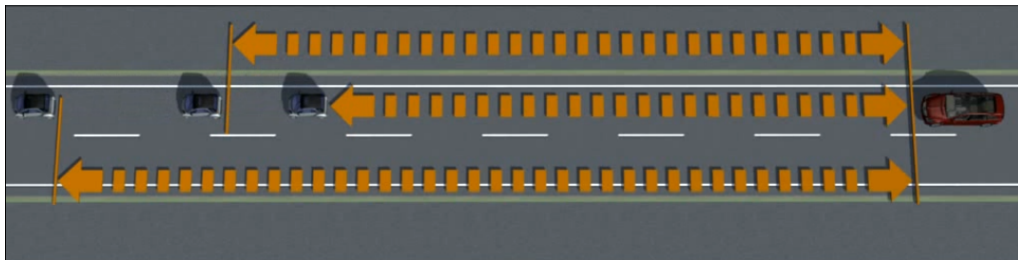


Abbildung 1.11.: Abweichungen von wenigen Pixeln der Disparität bringen eine große Abweichung in der Rückprojektion und damit in der Entfernungsschätzung zum vorausfahrenden Fahrzeug. Oben ist der korrekte Abstand. Unten und in der Mitte die Entfernungsmessung mit Abweichungen von wenigen Pixeln.

Durch Interaktion des Fahrzeugs mit der Fahrbahn und thermische Einflüssen verändern sich die Orientierung beider Kameras zueinander. Dadurch stimmen die Kameraparameter nicht mehr genau, was zu Abweichungen in der 3D-Rekonstruktion der Szene, wie in Abb. 1.11 dargestellt, führt.

Durch die Anbringung des Gehäuses der Stereo-Kamera an der Scheibe werden die mechanischen und termischen Einflüsse unmittelbar auf das Gehäuse der Stereo-Kamera übertragen. Um diesen Einflüssen entgegenzuwirken, muss das Intervall zwischen zwei Kalibrierungen der Online-Kalibrierung möglichst kurz sein.

Eine Stereo-Kamera kann auf der Basis von Sensordaten weiterer Sensoren kalibriert werden. Im Fahrzeug eignen sich hierfür die Sensordaten des Radars. Bei einer Kalibrierung

mit den Sensordaten des Radars werden die Daten des Radarsensors mit den Objekten aus der 3D-Rekonstruktion der Stereo-Kamera abgeglichen. Um eine Stereo-Kamera zu kalibrieren, wird in dieser Arbeit auf das Nutzen weiterer Sensoren zur Entfernungsmessung verzichtet. Erstens, da bei weiteren Sensoren darauf geachtet werden muss, dass die Systeme aufeinander und initial kalibriert sein müssen. Zweitens, weil bei einem Wechsel des Versuchsfahrzeugs darauf geachtet werden muss, dass die zusätzlichen Komponenten vorhanden sind. Die Kalibrierung des Stereo-Kamera-Systems wird aus den Informationen der Bilder und aus der Fahrzeug-Odometrie bestimmt, die in jedem Fahrzeug vorhanden ist.

Um eine Stereo-Kamera auf der Basis von Bilddaten allein zu kalibrieren, wird im Rahmen dieser Arbeit das Lochkameramodell mit Linsenverzeichnung genutzt. Ein solches Kameramodell wird im Rahmen dieser Arbeit Linsenkameramodell genannt. Dabei ist der Öffnungswinkel beider Kamerasysteme 56° . Mit der in [Brown, 1966] angegebenen Methode wird die Linsenverzeichnung aus dem Bild herausgerechnet. Bilder ohne eine solche Verzeichnung werden verzeichnungsfreie Bilder genannt. Verzeichnungsfreie Bilder sind essenziell für die in dieser Arbeit entwickelten Methoden zur Online-Kalibrierung, da diese Bilder zur Rekonstruktion genutzt werden.

Das in dieser Arbeit genutzte Stereo-Kamera-System ist achsenparallel ausgerichtet, damit die 3D-Rekonstruktion einfach bestimmt werden kann. Bei einer solchen Anordnung der Kameras ist das Bestimmen der Korrespondenzen erleichtert, da sich diese auf einer Pixelzeile befinden, wenn das Stereo-Kamera-System kalibriert ist. Das Ausrichten der Bilder aus einem Stereo-Kamera-System, sodass die Korrespondenzen beider Bilder auf einer Pixelzeile zu finden sind, nennen wir Rektifizierung. Um die Bilder rektifizieren zu können, werden die intrinsischen und extrinsischen Parameter genutzt [Fusiello et al., 2000].

Der Ausgangspunkt einer jeden Online-Kalibrierung ist eine durchgeführte Initialkalibrierung. Jede Online-Kalibrierung hat einen Schritt zur Optimierung, der mit Startwerten beginnt. Diese Startwerte werden in der Initialkalibrierung bestimmt. Bilder, die mit den Kameraparametern ausgerichtet wurden, nennen wir rektifizierte Bilder. Da der Einfluss von Verzeichnungen auf die Bildern minimiert wurde, sind Messungen sehr genau.

Um ein Stereo-Kamera-System online zu kalibrieren, werden Korrespondenzen genutzt. Korrespondenzen ist die Bezeichnung von korrespondierenden Punkten in verschiedenen Bildern. Es gibt unterschiedliche Methoden, um Korrespondenzen auf rektifizierten Bildern zu bestimmen (siehe dazu Abschn. 2.1). Um eine Kalibrierung zu verbessern oder zu unterstützen, können Odometrie-Daten zur Berechnung neuer Parameter herangezogen werden. Die in der Online-Kalibrierung neu errechneten Parameter werden dann genutzt,

um das nächste Bildpaar zu rektifizieren, um dann auf diesem Bildpaar neue Korrespondenzen zu bestimmen. Dieses ist in Abb. 1.10 dargestellt.

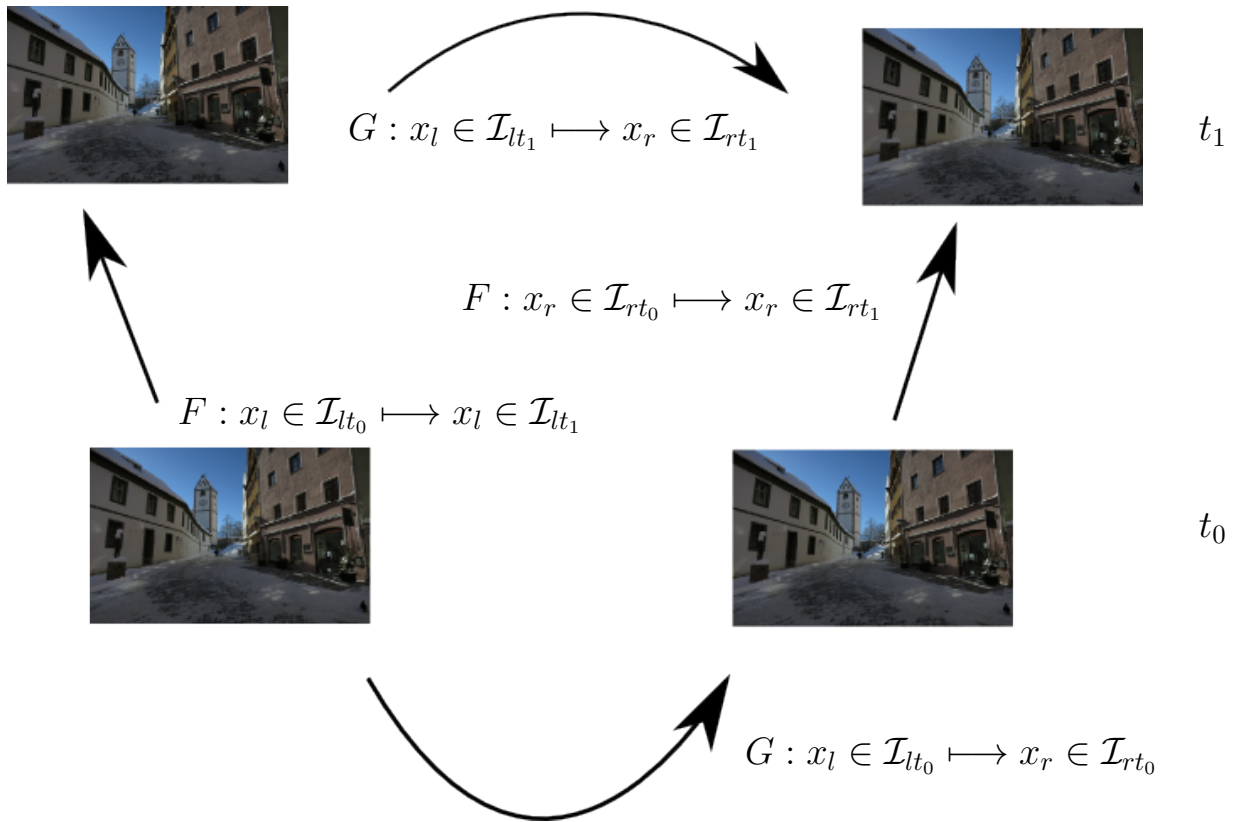


Abbildung 1.12.: Durch eine Funktion $F(x)$ werden Korrespondenzen in Bildern der gleichen Szene zu unterschiedlichen Zeitschritten gefunden. Die Funktion $G(x)$ findet Korrespondenzen in Bildern einer Szene, die den beiden Kameras eines Stereo-Kamera-Systems aufgenommen wurden.

In dieser Arbeit wird der optische Fluss – im folgenden Fluss genannt – genutzt, um Korrespondenzen zu bestimmen. Das Verfahren ist in einer Vielzahl von wissenschaftlichen Untersuchungen erarbeitet und optimiert worden [Stein, 2004]. Auf der Basis dieser Methode werden weitere Assistenzsysteme realisiert. Die Methode ist auf einer speziellen Hardware realisiert. Dabei ist diese Methode ein Bestandteil des genutzten Frameworks und kann in der Arbeit mit verwendet werden. Dies ist in Abb. 1.12 dargestellt und der Fluss ist eine Methode für die Funktion $F : x_{t_0} \mapsto x_{t_1}$. Dabei wird jeder Punkt $x \in \mathcal{I}_{t_0}$ zur Korrespondenz $x \in \mathcal{I}_{t_1}$ zugeordnet. Der Fluss wird beschrieben durch eine Menge von Verschiebungsvektoren zwischen zwei Bildern der gleichen Szene. Dabei werden die Objekte, die sich bewegen, bei stationären Kameras durch sogenannte Flussvektoren beschrieben. Wenn sich die Kamera bewegt, werden auch alle stationären Objekte abgebildet, da sich die Kameras bewegt haben. Ist in einem Bildpaar keine Veränderung zu finden, dann wird kein Flussvektor gefunden. Abb. 1.13 stellt dies dar. Dabei sind die Flussvektoren rot in Abb. 1.13 eingezeichnet.



Abbildung 1.13.: Der Fluss wird aus der Bewegung von zwei zeitlich aufeinanderfolgenden Bildern berechnet. Dabei stellt der Fluss die Bewegung von Objekten im Bild dar, wenn die Kamera sich nicht bewegt. Wenn sich die Kamera bewegt, dann werden alle Änderungen im Bild angegeben, so wie oben dargestellt. Zunächst werden miteinander korrespondierende Punkte ermittelt. Der Verschiebungsvektor, der einen Punkt in den anderen überführt, ist das Resultat. Dieser Verschiebungsvektor ist hier rot dargestellt. Die Auswertung einer Vielzahl von Punkten ergibt schließlich das Verschiebungsvektorfeld (vgl. [Paragios et al., 2005, S. 239]).

In dieser Arbeit werden räumliche Korrespondenzen zum gleichen Zeitpunkt mit der Semi-Global-Matching(SGM)-Methode, wie in [Hirschmüller, 2008] beschrieben, bestimmt. Die SGM-Methode schätzt für jeden Pixel eines Bildes die Disparität zum korrespondierenden Pixel im anderen Bild. In Abb. 1.12 kann die SGM-Methode die Funktion $G : x_{t_0} \mapsto x_{t_0}$ oder $G : x_{t_1} \mapsto x_{t_1}$ realisieren. Dabei werden die Korrespondenzen von einem Bild \mathcal{I}_l zum korrespondierenden Bild \mathcal{I}_r berechnet. Die Bilder müssen zum gleichen Zeitpunkt aufgenommen werden. Abb. 1.14 ist eine Darstellung der Tiefenkarte, die die Disparitätskarte des SGM repräsentiert.

Die SGM-Methode ist Bestandteil des verwendeten Frameworks und kann zur Kalibrierung genutzt werden. Die Bestimmung der Disparität mit der Kreuzkorrelation kann alternativ genutzt werden. Die Vor- und Nachteile der Methoden werden in Kap. 2 besprochen.



Abbildung 1.14.: Das Semi-Global-Matching(SGM) bestimmt für zwei gegebene Bilder einer Szene, die von unterschiedlichen Positionen zum gleichen Zeitpunkt aufgenommen wurden, die Korrespondenzen. Dieses Bild beschreibt einen Raum, der in [Hirschmüller, 2008] definiert und bestimmt wird. Die Größe der Disparität zweier miteinander korrespondierender Punkte ist farblich kodiert. Cyan sind kleine Disparitäten in einem Nahbereich im Sichtbereich und blau sind die Disparitäten in weiter Entfernung im Sichtbereich der Kamera.

Abb. 1.15 stellt die Beziehung der Methoden SGM und Fluss auf Basis von Abb. 1.12 zum Bestimmen der Korrespondenzen dar.

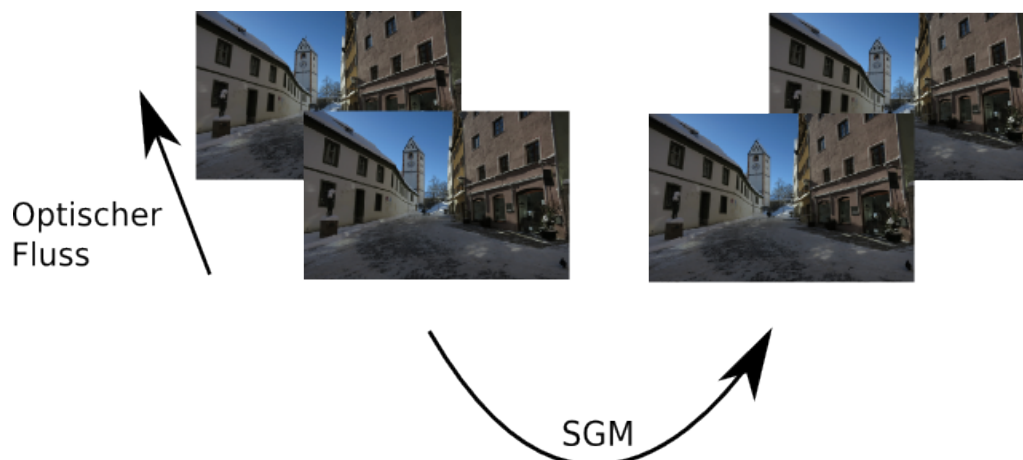


Abbildung 1.15.: In zwei zeitlich aufeinanderfolgenden Bildpaaren werden in dieser Arbeit die zwei Methoden SGM und Fluss genutzt, um Korrespondenzen in zeitlich aufeinanderfolgenden Bildpaaren zu bestimmen. Die Funktionen F und G in Abb. 1.12 werden im Rahmen dieser Arbeit ersetzt durch $F \equiv \text{Fluss}$ und $G \equiv \text{SGM}$, da diese Methoden wissenschaftlich untersucht und evaluiert wurden, auch für andere Verfahren eingesetzt werden und so die Resultate der Methoden der Online-Kalibrierung zur Verfügung stehen.

Auf der Basis der durch Fluss und SGM ermittelten Korrespondenzen kann ein Stereo-Kamera-System kalibriert werden. Ein Teil der Menge der extrinsischen Kameraparameter ist die relative Orientierung der beiden Kamerasysteme zueinander sowie deren

Änderungen nach der Initialkalibrierung, beschrieben durch die Parameter (ϕ_1, ϕ_2, ϕ_3) . In dieser Arbeit wird diese Orientierung bestimmt. Sie ist der Kern der in dieser Arbeit beschriebenen Verfahren. Es wird vorausgesetzt, dass die Kameras achsenparallel angeordnet sind, d.h. dass die z -Achsen beider Kameras parallel sind und jeweils vom Kamerahauptpunkt in die Rückprojektionsrichtung zeigen.

Die Methode zur Online-Kalibrierung soll mit den anderen Methoden auf einer gemeinsamen Hardware realisiert werden. Die Rechenleistung, um die Programme auszuführen, wird durch einen Digitalen Signalprozessor (DSP) bereitgestellt. Durch die Integration weiterer Programme auf den Hardwarekomponenten steht der Online-Kalibrierung nur ein Teil der Rechenleistung des DSPs zur Verfügung. Genauere Aussage über die verfügbare Rechenleistung wird in Kap. A.1 beschrieben.

1.2. Die Bedeutung der Online-Kalibrierung für Stereo-Kamera-Systeme

Eine kontinuierliche Überwachung und Anpassung der Parameter eines Stereo-Kamera-Systems über die Lebenszeit dieses Systems ist wichtig, da sonst die 3D-Rekonstruktion der Umgebung nicht genau genug gelingt und so die Fahrerassistenzsysteme zu ungenaue Daten erhalten. Sind die Daten aus dem Stereo-Kamera-System über die Lebensdauer des Stereo-Kamera-Systems nicht zuverlässig, dann werden alle Messungen, die in den Assistenzsystemen gemacht werden, fehlerhaft sein und die Systeme werden gar nicht oder ungenau reagieren. Um sich dieser, immer wieder neu auftretenden Herausforderung, zu stellen, ist eine Online-Kalibrierung des Stereo-Kamera-Systems essenziell.

Die Punkte in der 3D-Rekonstruktion der Fahrzeugumgebung haben die Repräsentation $X_{\mathcal{W}} = (x, y, z)_{\mathcal{W}}^T$ und können durch eine Triangulierung bei den gegebenen Bild-Koordinaten $(u, v)_{\mathcal{I}}$, bestimmt werden. Um einen Punkt $X_{\mathcal{W}}$ in einer 3D-Rekonstruktion aus den Daten der Stereo-Kamera zu bestimmen, wird Gl. (1.5) genutzt. Eine Auswertung der Funktion an einem Versuchsaufbau ist in Kap. 5.5 beschrieben.

$$\begin{pmatrix} u \\ v \end{pmatrix}_{\mathcal{I}} \rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{W}} = \begin{pmatrix} \frac{z}{f_u} \cdot (u - u_0) \\ \frac{z}{f_v} \cdot (v - v_0) \\ \frac{f \cdot b}{d} \end{pmatrix}, \quad (1.5)$$

wobei $(x, y, z)_{\mathcal{W}}^T$ der 3D-Punkt in der 3D-Rekonstruktion, $(u_0, v_0)^T$ der Bildhauptpunkt, d die Disparität und b die Basisbreite sind. Das Produkt $f \cdot b = \mathcal{F}$ wird Stereo-Kamera-Konstante oder Stereo-Konstante genannt. Im Rahmen dieser Arbeit wurden quadratische Pixel genutzt, so dass aus Gl. (1.5) die Gl. (1.6)

$$\begin{pmatrix} u \\ v \end{pmatrix}_{\mathcal{I}} \rightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\mathcal{W}} = \begin{pmatrix} \frac{b}{d} \cdot (u - u_0) \\ \frac{b}{d} \cdot (v - v_0) \\ \frac{f \cdot b}{d} \end{pmatrix}, \quad (1.6)$$

folgt. Auf der einen Seite ist zu sehen, dass das Rauschen in der Variable \mathcal{F} oder der Basisbreite sowie gaußsches Rauschen im Bildhauptpunkt nur linear in die Berechnung des 3D-Punktes eingehen, aber auf der anderen Seite ein Rauschen oder eine Abweichung der Disparität mit dem Faktor $\frac{1}{d}$ in die Berechnung des 3-D Punktes eingehen. Dies zeigt, dass selbst bei kleinen Änderungen in der Disparität, die in Pixel gemessen wird, ein großer Fehler in der 3D-Rekonstruktion entstehen kann.

In Abb. 1.16 ist dargestellt, welche Entfernung (auf der Ordinate in Metern) sich zwischen dem Stereo-Kamera-System und dem beobachteten Objekt bei einer Disparität (Abszisse in Pixel) und einer gegebenen Stereo-Kamera-Konstante $\mathcal{F} = 204.8 [Pix]$ ergibt. Diese Stereo-Kamera-Konstante wurde für den Versuchsaufbau des Stereo-Kamera-Systems genutzt. Aus Abb. 1.16 kann herausgelesen werden, wie sich eine Abweichung dann darstellt.

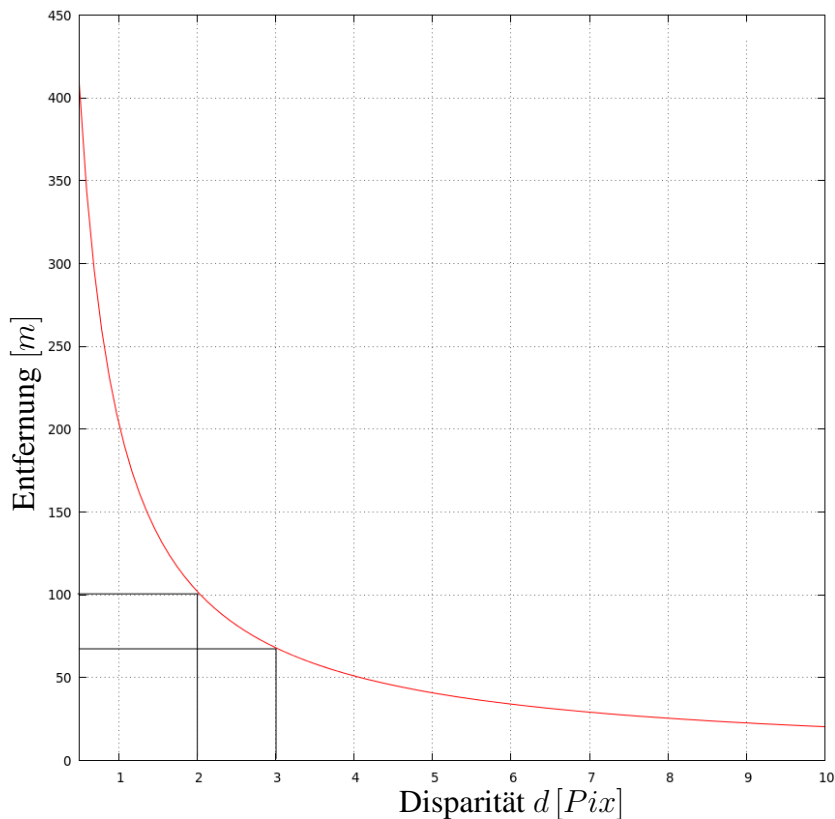


Abbildung 1.16.: Sei die Stereo-Kamera-Konstante \mathcal{F} gegeben, dann kann die Entfernung zu einem Objekt mit der Disparität d errechnet werden. Weicht dieser Wert in einem geringen Umfang ab, wird eine große Abweichung in der 3D-Rekonstruktion bestimmt.

Wird die Disparität d fälschlicherweise mit $d = 3 [Pix]$ statt mit 2 Pixeln berechnet, so wird die Entfernung mit $68 [m]$ statt mit $102 [m]$ berechnet.

Gliederung der Arbeit

Die Arbeit gliedert sich in sechs Kapitel. Im ersten Kapitel, wird eine Motivation zum Thema gegeben. Es wird grob angerissen, welches die erwarteten physikalischen Einflüsse sind und wie sich durch diese Einflüsse das Stereo-Kamera-System die Umgebungsrekonstruktionen ändern.

Anschließend wird in Kap. 2 der Stand der Technik aufgezeigt. Der Stand der Technik stellt eine Vielzahl von Methoden bereit. Alle Methoden haben unterschiedliche Vorteile und auch Nachteile, die diese für eine Anwendung in dieser Arbeit charakterisieren oder ausschließen.

Im darauf folgenden Kap. 3 wird dann erklärt, wie im Rahmen dieser Arbeit eine Initialkalibrierung vorgenommen wird. Die Werte aus dieser Initialkalibrierung stellen die Parameter für die Rektifizierung bereit.

Dem folgend wird im nächsten Kap. 4 die Methoden zur Schätzung des Schielwinkels beschrieben und anschließend werden die Methoden, zur Korrespondenzfilterung und Korrespondenzselektion erklärt. Die erste Methode, die in Abschn. 4.1 beschrieben ist, nutzt die Informationen über den zurückgelegten Weg des Fahrzeugs und prüft, ob die stationären Objekte der Scene in zwei aufeinanderfolgenden Bildern sich auf gleicherweise geändert haben. Dabei wird immer vorausgesetzt, dass das Fahrzeug sich in einer geradeausfahrt befindet. Die zweite Methode wird in Abschn. 4.2 beschrieben. Auch bei dieser Methode wird vorausgesetzt, dass sich das Fahrzeug in einer geradeausfahrt befindet. In dieser Methode wird der Fluchtpunkt geschätzt, sodass die Differenz der u-Koordinate des Fluchtpunktes zwischen beiden Bildern die Abweichung im Schielwinkel angibt.

Das Kapitel Kap. 5 gibt dann die Ergebnisse aller wissenschaftlichen Untersuchungen.

Diese Arbeit wird abgeschlossen mit Kap. 6. In dem Kapitel werden die Kapitel zusammengefasst und es werden Anregungen zu weiteren Entwicklungen gegeben.

2. Stand der Technik

Zur Online-Kalibrierung eines Stereo-Kamera-Systems werden Korrespondenzen genutzt. Um Korrespondenzen zu bestimmen und ihren Einfluss auf die Kalibrierung zu untersuchen, gibt es verschiedene Methoden. Dies wird in Abschn. 2.1 beschrieben.

Um ein Stereo-Kamera-System zu kalibrieren, können verschiedene Methoden genutzt werden. Eine Übersicht gibt Abschn. 2.3. Alle in Abschn. 2.3 beschriebenen Methoden nutzen eine Optimierung, die dann die Komponenten der Fundamentalmatrix oder die Komponenten der Essentialmatrix schätzt. Dazu wird ein Startwert genutzt, der aus der Initialkalibrierung genommen wird.

2.1. Korrespondenzbestimmung

Anders als bei einer Initialkalibrierung kann bei einer Online-Kalibrierung in einem sich bewegenden System kein Objekt mit einer bekannten Textur und Ausdehnung, ein sogenanntes Target, zur Kalibrierung genutzt werden und deshalb müssen andere Korrespondenzen in den Bildpaaren genutzt werden. Um diese Korrespondenzen zu bestimmen, können verschiedene Verfahren eingesetzt werden. Dazu werden zwei Bilder der gleichen Szene verwendet. Mehrere Methoden, die für diese Arbeit in Frage kommen, werden in [Dang, 2007] beschrieben und verglichen. In dieser Arbeit wird eine Hard- und Softwareumgebung genutzt, in der eine Vielzahl von Bildverarbeitungsverfahren für Fahrerassistenzsysteme verfügbar sind. Alle diese Methoden sind ausführlich evaluiert worden. Dazu gehören die Fluss-Methode aus [Stein, 2004] und SGM-Methoden aus [Gehrig und Rabe, 2010]. Diese Methoden können mit zur Online-Kalibrierung genutzt werden.

In [Rabe, 2011, Kap. 5] werden eine Analyse der Fluss-Methode und ein Vergleich zwischen dem Fluss und weiteren Algorithmen zur Korrespondenzbestimmung vorgenommen. Die auf der verwendeten Hardware der Stereo-Kamera integrierte Fluss-Methode arbeitet auf ganzzahligen Pixelkoordinaten.

[Beauchemin und Barron, 1995; Zach et al., 2007; Brooks et al., 1998; Stein, 2004; Ho und Goecke, 2008] zeigen verschiedene Methoden, wie der Optische Fluss berechnet werden kann. [Zach et al., 2007; Stein, 2004; Ho und Goecke, 2008] gehen auf eine schnelle Berechnung des Flusses ein, indem in [Zach et al., 2007] eine totale Variation mit der L_1 -

Norm vorgenommen wird. [Stein, 2004; Ho und Goecke, 2008] beschreiben verschiedene Methoden, die Fluss-Berechnung durch unterschiedliche Transformationen zu beschleunigen. Die Methode nach [Stein, 2004] ist in vielen wissenschaftlichen Arbeiten evaluiert worden und erzeugt eine hohe Anzahl von Flussvektoren, die das Zeitintervall zwischen den Kalibrierungen verringern kann.

Der Fluss kann, wie in [Kanatani et al., 2000] beschrieben, genutzt werden, um die Fundamentalmatrix zu schätzen. Doch diese Methode zur Online-Kalibrierung benötigt viel Speicher und viel Rechenzeit und deshalb kann diese Methode nicht auf der gegebenen Hardware realisiert werden. Zudem kann die Berechnung der extrinsischen Parameter nicht von der Berechnung der intrinsischen Parameter getrennt werden. So müssen in einem weiteren Rechenschritt die intrinsischen und extrinsischen Parametern aus der Fundamentalmatrix separiert werden.

Sollten in späteren Arbeiten die intrinsischen zusätzlich zu den extrinsischen Parametern bestimmt werden und sich die Rechenzeit und der Speicher vergrößern, dann ist diese Methode eine weitere Möglichkeit, auf der gegebenen Hardware eine Online-Kalibrierung zu berechnen. Da auf der gegebenen Hardware der Fluss für die Realisierung einiger Assistenzsysteme aus Tab. 1.1 eingesetzt wird, wird durch die in dieser Arbeit erarbeiteten Methoden keine weitere zusätzliche Rechenzeit benötigt.

Ist die Genauigkeit bis auf einen Pixel genau gegeben, nennen wir dies eine pixelgenaue Darstellung. Die pixelgenaue Darstellung wird genutzt, um die Methode besser auf die gegebene und verwendete Hardware zu integrieren. Durch den Einsatz des Kanade-Lucas-Tomasi(KLT)-Tracker aus [Birchfield, 1997; Rabe, 2011], kann die Flussberechnung, die auf der Hardware realisiert ist, zur Subpixelgenauigkeit erweitert werden.

Auch eine Kreuzkorrelation in dem ausgewählten Bereich würde die subpixelgenaue Berechnung des Flusses ermöglichen, doch zeigt [Rabe, 2011], dass sich durch die Kreuzkorrelation ergibt, dass ein Subpixel die Tendenz hat, sich früher auf den Wert des ganzzahligen Pixels einzustellen – dies wird Pixellocking genannt – als beim KLT-Shift [Rabe, 2011, Kap. 5].

Das Semi-Global-Matching(SGM) wird in [Hirschmüller, 2011; Hermann et al., 2008; Hirschmüller, 2005] beschrieben. [Zhu et al., 2012; Haller und Nedeveschi, 2010; Banz et al., 2011; Gehrig und Rabe, 2010; Ernst und Hirschmüller, 2008] geben eine praktische Anwendung für das Schätzen von Disparitäten zur Bestimmung der Disparitätskarte. Es wird aber immer vorausgesetzt, dass das Stereo-Kamera-System initial kalibriert ist. Ist die Kalibrierung schlecht, so ist die Dichte der Disparitätskarte nicht mehr so hoch und nicht jedes Pixel hat eine Disparität, mit der die Distanzen in der Szene vermessen werden können. In [Rabe, 2011] wird eine Analyse über die SGM-Methode angegeben.

Eine weitere Methode zur Berechnung von Korrespondenzen in Stereobildern ist die Be-

rechnung der Kreuzkorrelation für jedes Pixel [Jähne, 2005]. Bei dieser Methode wird eine Tiefenkarte berechnet, die nicht sehr dicht sein muss, anders als bei der SGM-Methode. Zudem liefert diese Methode für jedes Pixel einen positiven Wert, wenn dieser unendlich weit, oder einen negativen Wert, wenn der Wert nicht zugeordnet werden kann. Dies ist beim SGM anders, weil hier eine Berechnung über ganze Bereiche im Bild angegeben werden kann und so in den Bereichen im Bild Annahmen gesetzt werden können, so dass nicht jedes Pixel einen exakten Wert der 3D-Fahrzeugumgebung wiedergibt. Dies ist z.B. bei Bäumen zu beobachten, indem beim SGM die Bereiche zwischen den Ästen falsch berechnet werden, wohingegen die Kreuzkorrelation auch hier für jedes Pixel einen realistischen Wert für die 3D-Rekonstruktion liefert. Da aber die Kreuzkorrelation eine sehr aufwendig zu berechnende Methode ist und ein frühes Pixellocking aufweist, wird diese Methode in dieser Arbeit nicht genutzt. In Abb. 2.1 ist links im Bild die Disparitätskarte mit der Kreuzkorrelationsmethode und rechts ist die Disparitätskarte mit der SGM-Methode berechnet worden.

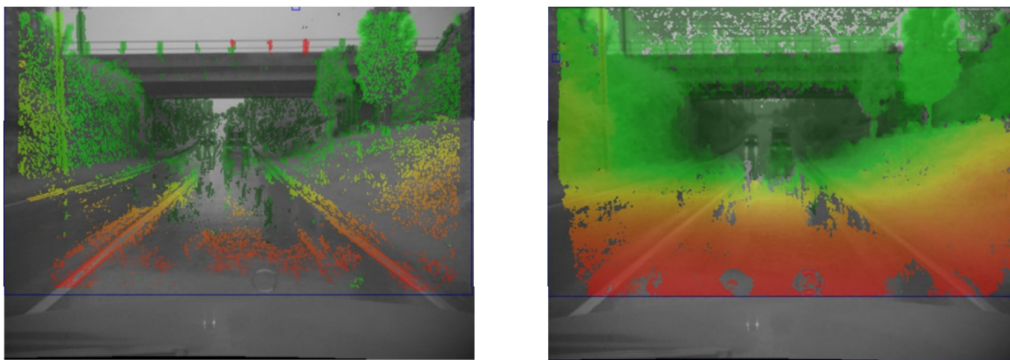


Abbildung 2.1.: Im linken Bild ist zu sehen, dass die Disparitätskarte mit der Kreuzkorrelationsmethode berechnet wurde. Im Gegensatz hierzu ist im rechten Bild die Disparitätskarte mit der SGM-Methode berechnet worden. Die Disparitäten, wie im linken Bild berechnet, sind sehr präzise, aber wie zu sehen, hat nicht jedes Pixel eine Disparität. Im rechten Bild ist fast für jedes Pixel eine Disparität bestimmt worden, aber die Methode verschmiert die Werte einer Disparität auf die Nachbarn, wodurch eine Ungenauigkeit entsteht (Bildquelle: [Hirschmüller, 2011]).

2.2. Ausreißerdetektion

Unter Ausreißerdetektion verstehen wir das Finden von Elementen, die nicht zu dem vorgegebenen Modell – hier der Kalibrierung eines Stereo-Kamera-Systems – gehören und so in der Schätzung zu großen Abweichungen führen können.

Durch das durch den Sensor bedingte Rauschen der Daten sowie die Asynchronität bei der Aufnahme mit beiden Kameras werden die Bilddaten verändert. Dies führt zu Ausreißern bei der Bestimmung der Korrespondenzen, die zu einer fehlerhaften Kalibrierung

führen würden. In diesem Abschnitt wird gezeigt, wie sich die Menge der Ausreißer in der Menge der Korrespondenzen bestimmen lässt.

2.2.1. Sampson-Distanz

Ein Fehlermaß bei der Berechnung der Online-Kalibrierung aus der Essential-Gleichung $x_l^T E x_r = 0$ ist die Sampson-Distanz [Hartley und Zisserman, 2004; Wöhler, 2009; Dang, 2007]. Dieses Maß misst den Abstand aller Korrespondenzen zur Epipolarlinie. Um den „mittleren Fehler“ aller Korrespondenzen zur Epipolarlinie zu bestimmen, wird der „Least Mean of Squares“ (LMeanS), wie in [Hartley und Zisserman, 2004; Wöhler, 2009] genutzt. Eine Verbesserung des LMeanS ist in [Xu und Zhang, 1996] beschrieben. In [Xu und Zhang, 1996] wird der „Least Median of Square“ (LMedS) evaluiert. Die LMedS-Methode ist robuster gegen die Ausreißer in der Berechnung als die LMeanS-Methode. Dies lässt sich in der Online-Kalibrierung in dieser Arbeit nutzen.

2.2.2. Die RANSAC-Methode

Der „Random Sample Consensus“ (RANSAC) [Fischler und Bolles, 1981] beschreibt eine Menge von Methoden in der Bildverarbeitung, die genutzt werden, um Ausreißer in einem gegebenen Modell zu finden. Um die RANSAC-Methode zu nutzen, muss die gegebene Aufgabe als Minimalproblem beschrieben werden. D.h. es wird die minimale Anzahl von Samples aus der Menge aller Samples zufällig gewählt, sodass die Aufgabe gelöst wird. Am Beispiel der Stereo-Kamera-Kalibrierung mit allen zwölf Parametern für eine stabile Lösung werden dann zwölf zufällig gewählte Korrespondenzen aus der Menge aller Korrespondenzen gewählt. Als eine stabile Lösung im Sprachgebrauch der Stereo-Kamera-Kalibrierung wird die Berechnung der intrinsischen und extrinsischen Parameter des Stereo-Kamera-Systems definiert. Dann wird eine Lösung für die beschriebene Aufgabe errechnet und mit den restlichen Korrespondenzen rektifiziert. Die Abweichung der restlichen Korrespondenzen vom Minimalproblem kann beispielsweise mit dem „Leased Mean Square“ errechnet werden. Als Nächstes wird wieder eine zufällige minimale Teilmenge aus der Menge aller Korrespondenzen gewählt und der Vorgang wiederholt. Dies geschieht so lange, bis die beste Lösung für das Problem gefunden wurde. [Torr und Murray, 1997; Ma et al., 2003; Hartley und Zisserman, 2004] zeigen, wie die RANSAC-Methode genutzt werden kann, um die Fundamentalmatrix zu schätzen. Dabei werden unterschiedliche Verfahren genutzt, um die Abweichung zwischen der aktuellen besten Fundamentalmatrix und der möglichen besten Fundamentalmatrix zu bestimmen.

Die Schätzung der Korrespondenzen ist ein iteratives Verfahren. Dazu werden die durch das Modell festgelegten Parameter bestimmt. Dies wird so lange durchgeführt, bis die beste Parametrisierung durch das Modell gegeben ist. Die beste Parametrisierung ist ge-

geben durch den Vergleich der aktuellen Parametrisierung zur nächsten Parametrisierung. Ist die nächste Parametrisierung besser, werden diese Parameter zur Parametrisierung des Modells genommen. Eine Parametrisierung wird als besser bezeichnet, wenn eine gegebene Fehlerfunktion einen besseren Wert hat. Als Fehlerfunktion wird in der Online-Kalibrierung der RMSE über die Distanz zwischen der Epipolarline und den Punkten (die Sampson-Distanz) genutzt.

Die Methoden, die im Rahmen dieser Arbeit entwickelt werden, schätzen die Rotation, die ein Teil der Essentialmatrix ist, so wie in Gl. (1.2) beschrieben.

In der Literatur wird die RANSAC-Methode häufig für die Schätzung der Fundamentalmatrix genutzt [Torr und Murray, 1997; Ma et al., 2003]. Um die Winkel (ϕ_1, ϕ_2, ϕ_3) der Rotationsmatrix der extrinsischen Kameraparameter der Essentialmatrix zu schätzen, wobei ϕ_1 die Änderung des Nick-, ϕ_2 die Änderung des Schiel- und ϕ_3 die Änderung des Rollwinkels sind, muss als erstes die Fundamentalmatrix in der RANSAC-Methode durch die Essentialmatrix ersetzt werden. Um, wie in der Literatur angegeben, die Fundamentalmatrix mit den zwölf Parametern zu schätzen, müssen mindestens zwölf Korrespondenzen bestimmt werden, die das Minimalproblem lösen.

Um die drei Parameter (ϕ_1, ϕ_2, ϕ_3) der Essentialmatrix zu bestimmen, werden wenigstens drei Korrespondenzen verwendet. Um diese drei Korrespondenzen zu bestimmen, müssen alle Korrespondenzen in Dreiergruppen zusammengefasst werden. Die Menge aller möglichen Teilmengen aus drei Korrespondenzen kann sehr groß sein.

Um nicht alle Teilmengen aus drei Korrespondenzen zu betrachten, kann eine Wahrscheinlichkeit herangezogen werden, die einen Wert bestimmt, der eine Aussage über die minimale Anzahl der Iterationen wiedergibt, die nötig sind, um eine Parametrisierung mit einer vorgegebenen Güte zu berechnen. Dies ist in Gl. (2.1) beschrieben.

$$1 - p = (1 - (1 - e)^c)^n, \quad (2.1)$$

mit n die Anzahl der Wiederholungen, die nötig sind, um eine Essentialmatrix zu berechnen, die von ausreißerfreien Korrespondenzen bestimmt wurde, p die Wahrscheinlichkeit, dass wenigstens eine Teilmenge ausreißerfrei ist, c ist die Anzahl der Korrespondenzen, die für die Berechnung des Modells notwendig sind und e die relative Anzahl von Ausreißern in der Menge aller Korrespondenzen. Die Parameter können empirisch oder mit dem Expectation-Maximization-Algorithmus bestimmt werden. Gl. (2.1) muss nun umgestellt werden, um die Anzahl der Wiederholungen zu berechnen. So folgt aus Gl. (2.1) die Gleichung

$$n = \frac{\log(1 - p)}{\log(1 - (1 - e)^c)}. \quad (2.2)$$

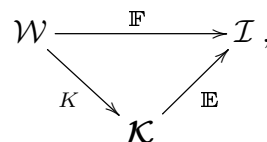
Durch die Bestimmung der Iterationen kann die Anzahl der Durchläufe stark gesenkt werden. Um dies zu zeigen, sei eine Menge von 1600 Korrespondenzen gegeben mit

20% Ausreißern und einer Wahrscheinlichkeit von 99%, dass eine ausreißerfreie Teilmenge gefunden wird. Dann werden 573 Iterationen benötigt, um eine Essentialmatrix zu bestimmen aus einer ausreißerfreien Menge von Korrespondenzen. Ohne diese Annahme müssten für die vollständige Durchsicht aller möglichen Kombinationen von Korrespondenzgruppen der Größe drei Korrespondenzen $\binom{1600}{3} = 681\,387\,200$ Iterationen gemacht werden, um die beste Essentialmatrix aus der Menge der gegebenen Korrespondenzen zu bestimmen.

Dabei wird die Anzahl der Ausreißer in der Menge aller Korrespondenzen empirisch, durch die „innere Genauigkeit“ oder dem Satz von Bayes, bestimmt.

2.2.3. Satz von Bayes in der Ausreißerdetektion

Ein weiteres Verfahren, um Ausreißer zu detektieren, ist die Betrachtung der Wahrscheinlichkeit, ob eine Korrespondenz zur Menge der Ausreißer gehört oder nicht. Um die Ausreißer zu filtern, wird der Satz von Bayes genutzt. Dazu wird betrachtet, welchen Benefit eine Korrespondenz zur Kalibrierung gibt. Die Berechnung der Wahrscheinlichkeiten ist ein Log-Likelihood-Ansatz [Bishop, 2006, S. 21]. Dabei sind zwei Klassen gegeben, die hier die Ausreißer und die Nicht-Ausreißer in der Menge der Korrespondenzen wiedergeben. [Brandt, 2006; Brandt und Heikkonen, 2000ab] zeigt die Anwendung des Satzes von Bayes auf die Menge der Korrespondenzen, um die Fundamentalmatrix zu schätzen. Diese Methode kann auf die Schätzung der Essentialmatrix angewendet werden. Im Fall der Essentialmatrixschätzung wird die Schätzung der Fundamentalmatrix durch die Schätzung der Essentialmatrix ersetzt. Dies ist möglich, da das Diagramm



mit K die Kameramatrix, \mathbb{E} die Essentialmatrix, \mathbb{F} die Fundamentalmatrix, \mathcal{W} das Weltkoordinatensystem, $\mathcal{I} \subset \mathbb{R}^2$ und \mathcal{K} das Kamerakoordinatensystem kommutativ ist [Wöhler, 2009; Hartley und Zisserman, 2004; Luong, 1992], d.h. $\mathbb{F} = \mathbb{E} \circ K$.

Die in [Brandt, 2006; Brandt und Heikkonen, 2000ab] beschriebene Methode wird auf affine Kamera-Systeme angewendet und wurde nicht mit einem Stereo-Kamera-System, wie es in dieser Arbeit verwendet wird, genutzt. Affine Kameras beschreiben einen Spezialfall der projektiven Geometrie, wobei auf acht von zwölf Freiheitsgraden reduziert wird. So behält das Kameramodell alle parallele Linien im Bild bei.

2.2.4. Klassische Maße aus der Photogrammetrie

In der Photogrammetrie ist die Methode der „inneren Genauigkeit“ für den Bündelausgleich [Luhmann, 2010; Triggs et al., 2000] bekannt. [Triggs et al., 2000] zeigt, wie die „innere

Genauigkeit" auf die Ausreißer in der Menge der Korrespondenzen wirkt. [Triggs et al., 2000] leitet Gl. (2.3) für die Anwendung im Bündelausgleich her.

$$\delta r_{ii}(W_i J_i Cov^{-1} J_i W_i + W_i) \delta r_{ii} = l \geq k, \quad (2.3)$$

mit δr_{ii} das Residuum der i -ten Korrespondenz, k ein gegebener Grenzwert, J_i die Jakobimatrix der Essentialmatrix, W_i das Gewicht, das der Korrespondenz eine Zuverlässigkeit zuordnet, und $Cov^{-1} = (J J^T)^{-1}$ die invertierte Kovarianzmatrix.

In Gl. (2.3) wird ein Residuum – hier durch die Sampson-Distanz berechnet – durch $J_i Cov J_i$ vom kartesischen Raum in den Raum der homogenen Transformationen projiziert, um so die erhaltene Krümmung der Zielfunktion gegen eine vorgegebene empirische Schwelle k zu testen. Diese Projektion erfolgt auf Basis einer linearen Approximation der inversen Zielfunktion. Die Herleitung der „inneren Genauigkeit" mit der Essentialgleichung ist analog zur Herleitung in [Triggs et al., 2000] beschrieben.

2.2.5. Prädiktionsband

Ein neuer Ansatz zur Ausreißerdetektion wird in [Martos et al., 2010] beschrieben. Die dort angegebene Methode schätzt die Kovarianz jeder Korrespondenz und bestimmt den Anteil der Korrespondenz zur Kalibrierung. Die Gleichung Gl. (2.4) wird aus der Student-T-Verteilung hergeleitet (siehe Kap. A.2).

$$\begin{aligned} \sigma(u, v) &= \frac{s}{2} \sqrt{(W_i J_i Cov(u, v) J_i W_i + W_i) \frac{RSS}{dof}} \\ &= \frac{s}{2} \sqrt{(J Cov(u, v) J + 1) \frac{RSS}{dof}}, \end{aligned} \quad (2.4)$$

mit (u, v) die Koordinate im linken Bild des Bildpaares, $RSS = \sum_i r_i$ das „Residual Sum of Square", dof der Freiheitsgrad und $s = cdf_{student-t}(1 - \frac{\alpha}{2}, dof)^{-1}$ die inverse Verteilungsfunktion der Student-T-Verteilung, mit α die Wahrscheinlichkeit der gesuchten Konfidenz. Diese Schätzung wird mittels der Prädiktionsbänder vorgenommen. Konfidenz- und Prädiktionsbänder sind in [Kardaun, 2005, S. 196] und [Martos et al., 2010] beschrieben. Eine Herleitung der Prädiktionsbänder ist in Kap. A.2 gegeben.

2.2.6. Robuste Fehlerfunktionen

In der Statistik werden robuste Fehlerfunktionen verwendet, um Ausreißer in einer gegebenen Menge zu finden und zu entfernen. In [Hartley und Zisserman, 2004, S. 616] werden verschiedene robuste Fehlerfunktionen angegeben, die sich zur Kalibrierung eignen.

Diese Funktionen zeichnen sich dadurch aus, dass sie in verschiedenen Anwendungen evaluiert wurden und eine Verbesserung in der Kalibrierung aufweisen [Hartley und Zisserman, 2004]. Zudem sind die Funktionen so angelegt, dass die Korrespondenzen schnell klassifiziert werden können. Um die empirische Schwelle für die robusten Fehlerfunktionen zu bestimmen, sind vorhergehende Untersuchungen notwendig.

2.3. Methoden zur Online-Kalibrierung

Sind die Korrespondenzen in den Bildpaaren bestimmt worden, werden diese genutzt, um die über die Zeit veränderten Kameraparameter zu bestimmen. Um diese Parameter zu schätzen, können verschiedene Methoden genutzt werden, die die Fundamentalmatrix der Epipolargeometrie schätzen. Die Fundamentalmatrix beschreibt die relative Position und Orientierung der beiden Kamerakoordinatensysteme zueinander [Wöhler, 2009; Hartley und Zisserman, 2004; Kanatani et al., 2000; Noury et al., 2007; Kanatani und Sugaya, 2009; Ma et al., 2003].

Sind die intrinsischen Parameter der Kameras bekannt, ist die geometrische Anordnung der Kameras durch die Essentialmatrix gegeben [Luong, 1992; Luong und Faugeras, 1994 1997]. Eine Erweiterung bilden die Methoden, die zu der Fundamental- oder Essentialmatrix noch die Linsenverzeichnung der Optik schätzen. Dies wird in [Barreto und Daniilidis, 2005] für die Fundamentalmatrix und in [Zhang, 1996; Graf, 2007] für die Essentialmatrix beschrieben. Im Rahmen dieser Arbeit werden die Verfahren ohne Linsenverzeichnung verwendet, da diese in der Initialkalibrierung einmal bestimmt wurde und sich über den Lebenszyklus der Kamera nicht ändert. Mit der Initialkalibrierung können die Parameter für die Linsenverzeichnung nach [Brown, 1966] bestimmt werden. Die Methode nach [Brown, 1966] stellt eine geometrische Berechnungsmethode dar, mit der die Linsenverzeichnung aus den digitalen Bilddaten herausgerechnet wird. Sind die Parameter für die geometrische Berechnungsmethode nach [Brown, 1966] bestimmt worden, können die Bilder als verzeichnungsfrei angenommen werden.

In [Hartley, 1997a] wird ein Verfahren zur Bestimmung der Fundamentalmatrix beschrieben, das die Fundamentalmatrix mit acht Korrespondenzen bestimmt. Die Essentialmatrix $\mathbb{E} := \mathbf{R}\mathbf{T}_\times$, mit \mathbf{R} die Rotationsmatrix und \mathbf{T}_\times die schief-symmetrische Matrix des Translationsvektors, kann durch weniger Korrespondenzen bestimmt werden. Der allgemeinste Ansatz zum Schätzen der Essentialmatrix in einer Online-Kalibrierung ist in [Torii et al., 2011; Åström et al., 1996] beschrieben. Hier werden sechs Korrespondenzen genutzt. In [Nistér und Schaffalitzky, 2004] wird die Essentialmatrix durch vier Korrespondenzen geschätzt. Dabei ist vorausgesetzt, dass die Kameras achsenparallel ausgerichtet sind, d. h. die z-Achsen beider Kameras sind parallel und zeigen vom Projektionsursprung durch den Kamerahauptpunkt in Richtung der inversen Projektion Abb. 2.2.

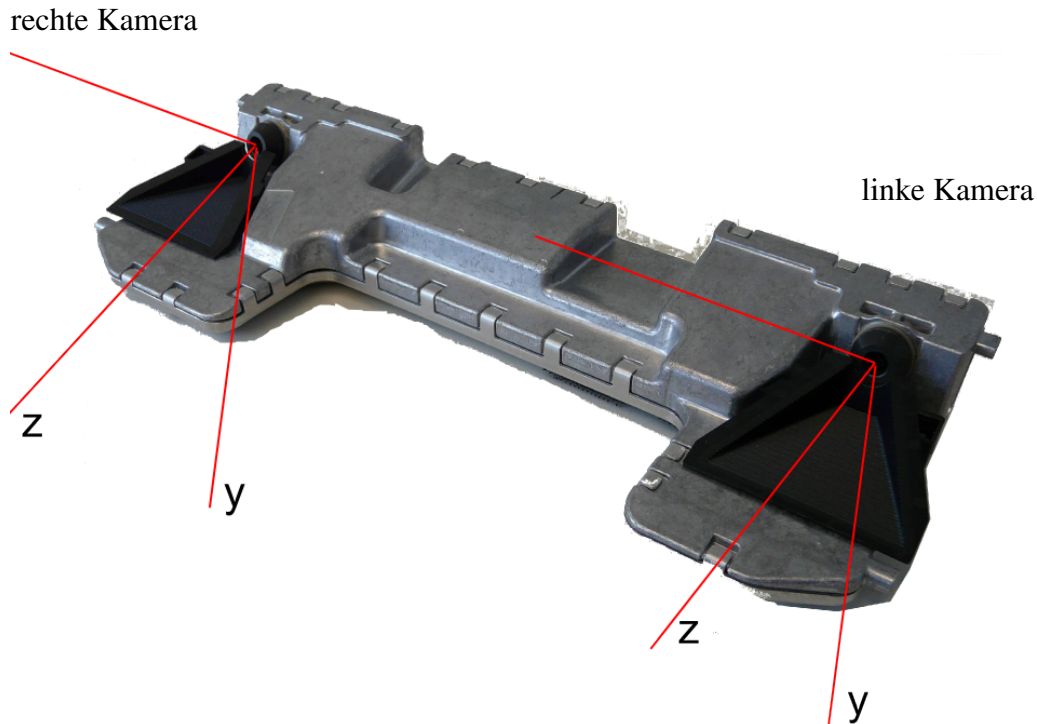


Abbildung 2.2.: Die beiden Kameras des Stereo-Kamera-Systems, das die Basis der experimentellen Untersuchungen im Rahmen dieser Arbeit ist, sind achsenparallel ausgerichtet. Die z-Achse zeigt in die Richtung der Abbildung.

Die vier Parameter beschreiben die Werte (ϕ_1, ϕ_2, ϕ_3) der Rotationsmatrix und den Parameter der Translation, der die Translation der beiden Kamerakoordinatensysteme beschreibt. Weitere Methoden zum Schätzen der Essentialmatrix werden in [Nister, 2004; Werner, 2003; Stewénus et al., 2006; Torr und Fitzgibbon, 2004] beschrieben. Diese Methoden bestimmen alle Parameter der Essentialmatrix bei beliebigem Systemaufbau, d.h. die Kameras müssen nicht achsenparallel angeordnet sein.

Die Sensordaten sind mit einem gaußschen Rauschen behaftet. Dieses Rauschen hat einen Einfluss auf die Korrespondenzen, die für die Online-Kalibrierung genutzt werden. Durch eine große Anzahl von Korrespondenzen, die ein kleines Signal-Rausch-Verhältnis haben, wird der Fehler der Online-Kalibrierung verkleinert. So ist in dieser Arbeit die Menge der Korrespondenzen sehr viel größer als die in den Methoden, die in [Dang et al., 2009; Dang und Hoffmann, 2004; Dang, 2007] angegeben werden.

2.3.1. Tensoren in der Stereo-Kamera-Kalibrierung

Der Trifokal-Tensor beschreibt die Orientierung und die Position von drei Kameras in einem System zueinander. Somit ist die Rolle der Fundamentalmatrix im Stereo-Kamera-System identisch mit der Rolle des Trifokal-Tensors im trinokularen Kamerasystem [Hartley und Zisserman, 2004; Ma et al., 2003]. [Shashua, 1997] gibt eine ausführliche Her-

leitung des Trifokal-Tensors. [Yu et al., 2006; Nistér, 2000; Torr und Zisserman, 1998; Scoleri et al., 2007; Förstner, 2000; Hartley, 1997b; Torr und Zisserman, 1997; Khropov et al., 2004] geben eine sehr ausführliche Darstellung über die Kalibrierung und Rekonstruktion von und durch den Trifokal-Tensor. [Thirthala und Pollefeys, 2005] nutzt den Trifokal-Tensor, um die Kameras zu kalibrieren und dabei die Linsenverzeichnung von Linsen mit großem Öffnungswinkel zu bestimmen. Der Trifokal-Tensor ist eine Methode, eine sehr genaue Kalibrierung zu erhalten. Da aber drei Kameras genutzt werden, ist die Berechnung des Trifokal-Tensors sehr aufwendig. Mit der Hinzunahme weiterer Kameras in einem Kameraverbund wird der Fehler der Online-Kalibrierung weiter verkleinert.

[Dang, 2007] stellt erstmals den Zusammenhang zwischen Trifokal-Tensor und der Stereo-Kamera-Kalibrierung her. Die in [Dang, 2007] entwickelte Kalibrieremethode ist für die Online-Kalibrierung in bewegten Fahrzeugen geeignet. Die Datenmenge, die zur Kalibrierung genutzt wird, ist sehr groß, da Tracks zu den Frames gespeichert werden, um die Korrespondenzen in allen Bildern des Tracks zu finden. Durch den verwendeten Kalman-Filter ist die Rechenzeit sehr hoch und durch die Nutzung von Tracks und dem Kalman-Filter ist die Kalibrierung sehr komplex. Die Berechnung der reprojizierten Elemente im Weltkoordinatensystem, die auch in [Dang, 2007] genutzt wird, wird auch im Bündelausgleich genutzt. In [Dang, 2007] wird gezeigt, dass es sich bei dieser Kalibrierung um die zuverlässigste Online-Kalibrierung für Stereo-Kamera-Systeme in bewegten Fahrzeugen handelt. Steht das Fahrzeug still, so wird keine Kalibrierung zur Verfügung gestellt.

In [Dang, 2007] wird ein Kalman-Filter genutzt, der über Tracks von acht bis zwölf Bildpaaren Korrespondenzen für eine neue Kalibrierung bestimmt. Dieses kann durch einen Levenberg-Marquardt-Algorithmus [Levenberg, 1944] beschleunigt werden.

Die in [Dang, 2007] beschriebene Methode kann zu einem Quadrifokal-Tensor erweitert werden. Dazu werden die zwei zeitlich aufeinanderfolgenden Bildpaare genutzt, um den Quadrifokal-Tensor zu berechnen. Durch die homogene Transformation, die aus der Fahrzeug-Odometrie bestimmt wird, ist die relative Position aller vier Kameras zueinander bekannt (vgl. Abb. 2.3) und mit der Formel aus [Kanatani und Sugaya, 2011; Hartley und Zisserman, 2004] kann der Quadrifokal-Tensor berechnet werden.

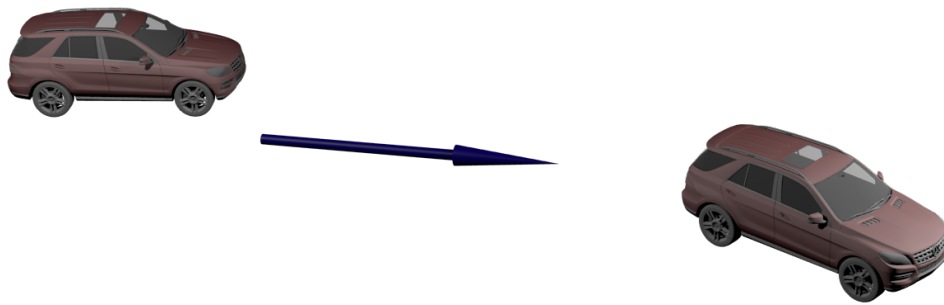


Abbildung 2.3.: Es existiert eine Transformation zwischen den beiden Fahrzeugpositionen. Diese Transformation kann aus den Daten des Fahrzeug-Odometrie bestimmt werden. Ist ein Kamera-System in dem Fahrzeug fest integriert, dann können die Koordinatensysteme der Kameras mit dieser Transformation zu einem neuen Mehrkamerasystem zusammengefasst werden.

Das Nutzen des Quadrifokal-Tensors zur Kalibrierung ist robuster als die Essential-Gleichung oder die in [Dang, 2007] beschriebene Methode. In [Hartley und Zisserman, 2004] wird auch der n-Tensor zur Kamerakalibrierung beschrieben und durch die sukzessive Nutzung zeitlich aufeinanderfolgender Bilder kann ein n-Tensor zur Kamerakalibrierung genutzt werden.

2.3.2. Bündelausgleich

Der Bündel(block)ausgleich ist eine weitere Methode, die zur Kalibrierung, sowohl online als auch offline, genutzt werden kann. Diese Methode ist ein Standardverfahren in der Photogrammetrie [Triggs et al., 2000; Ma et al., 2003; Kanatani, 2005; Kanatani und Sugaya, 2011; Eudes und Lhuillier, 2009; Luhmann, 2010]. Sie minimiert den Rückprojektionsfehler. Dieser Rückprojektionsfehler wird als „Mean Square Error“ (MSE) oder „Root Mean Square Error“ (RMSE) angegeben.

Dazu wird ein gegebener Punkt im Weltkoordinatensystem auf das Kameratarget (Bildfläche), auf der Basis der aktuellen Kameraparameter, projiziert und der Abstand zwischen den gemessenen und projizierten Punkten minimiert. Diese Methode kann, wenn die Projektionsmatrizen der Kameras bekannt sind, eine Rekonstruktion der Fahrzeugumgebung ermöglichen [Gargallo et al., 2007; Hung und Tang, 2006; Wong et al., 2004; Mouragnon et al., 2006; Jeong et al., 2010; Kanatani und Sugaya, 2011; Bastian und van den Hengel, 2003; Aravkin et al., 2012; Byröd und Åström, 2010; Lee und Yu, 2009; Engels et al., 2006; Eudes et al., 2010; Sibley et al., 2009; Kurz et al., 2011]. [Pollefeys, 1999] zeigt, wie mit dem Bündelausgleich Objekte rekonstruiert werden können. Dazu ist es wichtig, dass diese Objekte sich in der gleichen Szene befinden und von mehreren Kameras, die in einem Block zusammengefasst werden, erfasst werden. Zudem wird

gezeigt, dass die Kameras nicht fest verbunden sein müssen, da die Positionen der Kameras relativ zu einem Koordinatensystem mitgeschätzt werden können. [Hartley und Zisserman, 2004; Wöhler, 2009; Luhmann, 2010; Triggs et al., 2000] geben ein sehr allgemeines Verständnis und einen sehr weit gefassten Überblick über den Bündelausgleich. [Ma et al., 2003] gibt eine allgemeine Einführung zum Bündelausgleich mit dem Fokus auf Stereo-Kameras.

Die Fehlerfortpflanzung durch den Bündelblockausgleich wird in [Eudes und Lhuillier, 2009] beschrieben.

Da der Bündelausgleich eine quadratische Funktion minimiert, existiert ein eindeutiges Minimum. Dieses Minimum ist die globale Lösung. In der Praxis wird das globale Minimum gefunden, indem die Methode mit verschiedenen Anfangswerten gestartet wird. Dies ist so, da in der Praxis lokale Minima in der Nähe des globalen Minimums auftreten können. Das Berechnen des Minimums mit der Methode mit verschiedenen Anfangswerten führt dann zu einer erhöhten Berechnungszeit der exakten Parameter der Kalibrierung. Dennoch können mit dem Bündelausgleich alle internen und externen Parameter des Kameramodells und eines Kamerasystems bestimmt werden. Doch die Berechnung ist sehr aufwendig und komplex. Durch die hohen Anforderungen an die Rechenleistung und den großen Speicherbedarf kann die Methode nicht mit der gegebenen Hardware, die für den Einsatz im Fahrzeug geeignet ist, verwendet werden. Die Methode eignet sich aber für die 3D-Rekonstruktion sowie für die Initialkalibrierung.

Diese Methode ist die „allgemeinste Methode“, die in allen Bereichen der Rekonstruktion und Kalibrierung eingesetzt werden kann. Die Methode wird häufig für die Online- und Initialkalibrierung eingesetzt. Die Zeit zwischen zwei Kalibrierungen ist sehr lang und eignet sich deshalb nicht für eine Methode im Rahmen dieser Arbeit.

2.3.3. Essentialmannigfaltigkeit

Ein weiteres Verfahren zur Online-Kalibrierung basiert auf der Betrachtung des Fehlers bei der Schätzung der Essentialmatrix auf der Essentialmannigfaltigkeit, die eine riemannsche Mannigfaltigkeit ist [Rapp, 2013; Helmke et al., 2007; Soatto et al., 1994; Ma et al., 2003]. Die Masterarbeit von Herrn Matthias Rapp wurde von mir betreut und ist eine ausführliche Untersuchung der Arbeit von [Helmke et al., 2007], in der eine Online-Kalibrierung mit algebraischen Methoden, die auf der Essentialmannigfaltigkeit agieren, gemacht wird. Die Arbeiten von [Soatto et al., 1994; Ma et al., 2003] gehen nicht explizit auf die Eigenschaften der Essentialmannigfaltigkeit ein. Das in [Rapp, 2013; Helmke et al., 2007] beschriebene Maß betrachtet die Menge der eingehenden Korrespondenzen und deren Wirkung auf der Essential-Mannigfaltigkeit.

Die in [Helmke et al., 2007] angegebene Methode komprimiert die Menge aller Korrespondenzen in einer Matrix $\mathcal{M} \in \mathbb{R}^{9 \times 9}$. Diese Matrix \mathcal{M} hat all die Eigenschaften aller Korrespondenzen, die der Methode übergeben werden. Dies ist auch der Nachteil der Me-

thode. Die Ausreißer erhalten den gleichen Einfluss wie alle anderen Korrespondenzen. Sollten aber alle Korrespondenzen von Ausreißern befreit sein, dann ist diese Matrix \mathcal{M} ausreißerfrei und kann die Schätzung der Essentialmatrix beschleunigen. Zudem basieren alle Formeln zur Schätzung der Essentialmatrix auf algebraischen Methoden und sind somit sehr robust in ihrer Anwendung.

Da sich die Parameter der Essentialmatrix auf einer Mannigfaltigkeit anordnen lassen, gibt es eine Nachbarschaftsbeziehung zwischen den Parametern. So kann schnell eine Parametrisierung der Stereo-Kamera gefunden werden, da sich die Parameter der Stereo-Kamera nur gering ändern und sich diese „geringfügigen“ Änderungen in einer Nachbarschaft auf der Mannigfaltigkeit zur letzten Parametrisierung befinden müssen.

2.4. Triangulierung

Die Triangulierung wird genutzt, um die 3D-Rekonstruktion aus den Daten der Stereo-Kamera zu berechnen. Dabei wird jeder Bildpunkt $(u, v)_{\mathcal{I}}$ in einen Weltpunkt $(x, y, z)_{\mathcal{W}}$ projiziert. [Hartley, 1999; Kanatani, 2005; Hartley und Zisserman, 2004; Wöhler, 2009] geben allgemeine Ansätze zur Triangulierung.

[Kanatani und Niitsuma, 2011] rekonstruiert die Szene ohne die Verwendung des Reprojektionsfehlers wie in [Ma et al., 2003]. Die Methoden zur Triangulierung in [Ma et al., 2003; Kanatani und Niitsuma, 2011] beziehen sich auf Kameras in einem Stereo-Kamera-System in allgemeiner Lage. [Kanatani und Niitsuma, 2011] geht auf die geometrische Triangulierung ein und zeigt, dass bei achsenparalleler Kameraausrichtung die Triangulierung sich schnell und speichereffizient nutzen lässt. Dieses stellt einen Bezug zu [Hartley und Sturm, 1997; Lindstrom, 2010] her, die nur achsenparallele Kamerasysteme betrachten. Andere Verfahren zur Triangulierung sind beschrieben in [Torres-Mendez et al., 2004; Kazhdan et al., 2006; Hartley und Schaffalitzky, 2009; Khwaja und Goecke, 2008; Prasad und Fitzgibbon, 2006]. Diese Methoden sind für bestimmte Fälle entwickelt worden. Besonders die Berechnung des Grassmann-Tensors in [Hartley und Schaffalitzky, 2009] benötigt eine lange Rechenzeit. Dafür ist die Triangulierung mit dem Grassmann-Tensor numerisch stabiler.

Durch die in dieser Arbeit verwendete Kamerakonfiguration ist eine geometrische Methode zur effizienten Berechnung der 3D-Rekonstruktion nutzbar. Die in [Fusiello et al., 2000] beschriebene Methode nutzt Gl. (1.5) und wird auch im Rahmen dieser Arbeit verwendet.

3. Initialkalibrierung

Um Objekte in einem Bild zu vermessen, werden exakte Werte für die Kameraparameter genutzt. Diese Werte werden durch eine Kalibrierung der extrinsischen und intrinsischen Kameraparameter bestimmt.

Um ein Stereo-Kamera-System zu kalibrieren, werden zuerst die intrinsischen Parameter und die Linsenverzeichnung jeder Kamera des Kameraverbunds für sich kalibriert. Die Linsenverzeichnung wird durch die Bauweise von Kameras hervorgerufen, sie ist in [Brown, 1971] beschrieben und in Abb. 3.1 dargestellt.

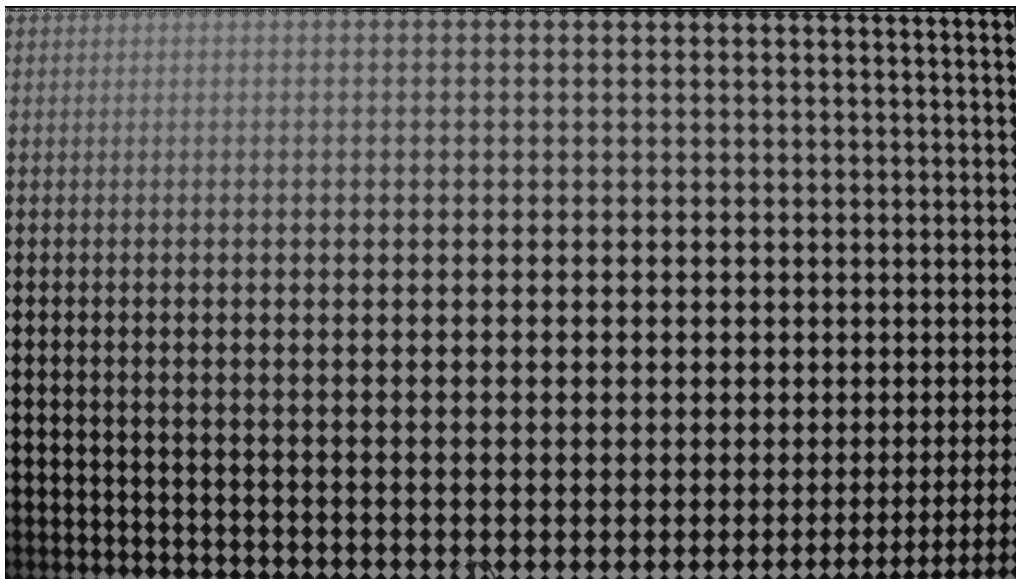


Abbildung 3.1.: Das Target repräsentiert ein Rautenmuster mit Geraden. Diese Geraden krümmen sich wegen der Linsenverzeichnung und machen so ein weiteres Messen schwer.

Danach werden die Orientierung und die Position der Kameras im Stereo-Kamera-System zueinander in der extrinsischen Kalibrierung bestimmt.

Dabei wird eine bekannte Geometrie im Computer gespeichert und auf einer ebenen Fläche, dem sogenannten Target, angebracht. Dieses Target wird in verschiedenen Distanzen, Positionen und Orientierungen mit der Kamera aufgenommen, um die Kamerakonstante, den Kamerahauptpunkt und das Aspectverhältnis zu approximieren. In Abb. 3.2 wird das Schachbrett in einer größeren Entfernung als in Abb. 3.3 repräsentiert. Die Repräsentation des Schachbrettes in unterschiedlichen Entfernungen wird genutzt, um si-

herzustellen, dass die Messungen von Objekten in unterschiedlichen Entfernungen zuverlässig sind.

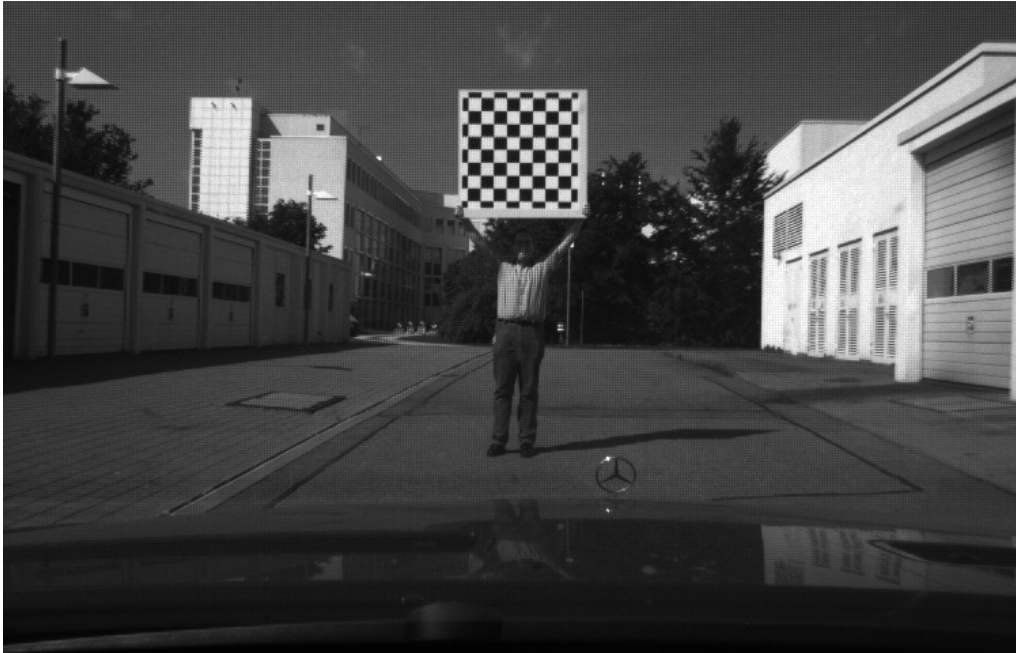


Abbildung 3.2.: Wenn eine Kamera auf ein Schachbrett ausgerichtet wird und ein weiteres Schachbrett zur intrinsischen Parameterbestimmung eingesetzt wird, kann die extrinsische Position sehr genau ermittelt werden. Durch die große Distanz zwischen dem Kamera-System und dem Target werden die Parameter in großen Distanzen exakt eingestellt.

Dann wird das im Computer hinterlegte Muster mit der Stelle des im Bild befindlichen Targets abgeglichen. Für diese Projektion werden die aktuellen Kameraparameter genutzt. Stimmen die Projektionen des Targets aus dem Bild und die computerinterne Repräsentation des Targets überein, dann ist die Kamera kalibriert. Andernfalls müssen die Parameter der Kamera so lange angepasst werden, bis das Modell im Computer mit der Geometrie auf dem aufgenommenen Target übereinstimmt. Als Fehler, der den Unterschied zwischen dem computerinternen Modell und der Projektion darstellt, wird der Rückprojektionsfehler genutzt. Dies entspricht dem Bündelausgleich (s. Kap. 2.3.2). Sind die intrinsischen Parameter der beiden Kameras im Stereo-Kamera-System bestimmt worden, werden die Position und die Orientierung der beiden Kameras zueinander bestimmt. Dazu wird das Target in beiden Bildern zur gleichen Zeit präsentiert.

Verschiedene Methoden zur Kalibrierung sind in [Heikkila und Silven, 1997; Weng et al., 1992; Brown, 1971; Bucher, 2000; Yang et al., 2008; Zhang, 1999; Tsai, 1987; Krüger und Wöhler, 2011; Krüger, 2007; Bouguet und Perona, 1999 1998; Mühlich und Aach, 2007] beschrieben. Da die Initialkalibrierung ein Target und eine bekannte Geometrie nutzt, kann eine solche Methode nicht für die Online-Kalibrierung herangezogen werden.



Abbildung 3.3.: Wenn eine Kamera auf ein Schachbrett ausgerichtet wird und ein weiteres Schachbrett zur intrinsischen Parameterbestimmung eingesetzt wird, kann die extrinsische Position sehr genau ermittelt werden. Durch die kurze Distanz zur Kamera kann die Kalibrierung im Nahbereich exakt bestimmt werden.

Eine sehr häufig verwendete Methode, die auch in dieser Arbeit zur Initialkalibrierung genutzt wird, wird in [Bouguet und Perona, 1998] vorgestellt. Die in [Bouguet und Perona, 1998] vorgestellte Methode nutzt ein Schachbrettmuster zur Kalibrierung. Ein solches Muster ist in Abb. 3.4 dargestellt. Durch die genaue Kenntnis der Geometrie des Schachbrettes, was die Anzahl der Quadrate pro Zeile sowie die Anzahl der Quadrate pro Spalte und die Kantenlänge eines Quadrates einschließt, können die Kameraparameter geschätzt werden.

Jeder Schnittpunkt der durch das Schachbrettmuster entstandenen horizontalen und vertikalen Linien kann in ein Koordinatensystem umgerechnet werden. Durch das Kreuzprodukt kann das 2D-Koordinatensystem entlang der horizontalen und vertikalen Linien zu einem 3D-Koordinatensystem erweitert werden. Ein solches 3D-Koordinatensystem ist in einem Schnittpunkt in Abb. 3.4 eingezeichnet.

Als Optimierungskriterium wird in dieser Arbeit der Bündelausgleich (siehe Kap. 2.3.2) eingesetzt. Der im Bündelausgleich genutzte Rückprojektionsfehler ist durch den RMSE gegeben. Um eine qualitative Abschätzung für den durch den Rückprojektionsfehler berechneten Wert zu haben, wird dieser empirisch für eine Vielzahl von Auswertungen bestimmt. Mit einem solchen empirischen Wert kann dann die Initialkalibrierung bewertet

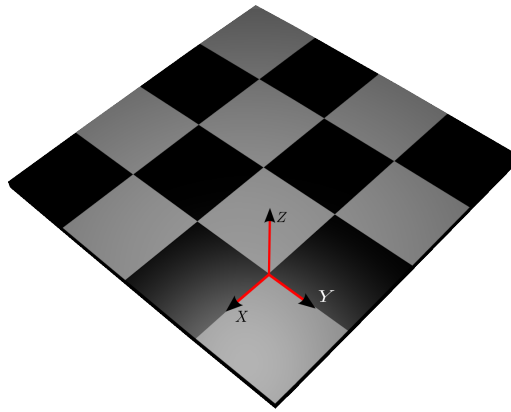


Abbildung 3.4.: In [Bouguet und Perona, 1998] wird ein Schachbrettmuster genutzt, um ein Kamerasystem zu kalibrieren. Die Anzahl der Quadrate in der horizontalen und vertikalen Richtung sowie die Kantenlänge eines Quadrates sind bekannt. Jeder Schnittpunkt der durch das Schachbrettmuster entstandenen horizontalen und vertikalen Linien kann in ein Koordinatensystem umgerechnet werden. Der Schnittpunkt ist die Koordinate, die zur Kalibrierung genutzt wird, d.h. die Abstände zwischen den Schnittpunkten des computer-internen Modells und den durch die Kamera aufgenommenen Schnittpunkte werden durch den Rückprojektionsfehler minimiert.

werden.

Mit den Werten aus der Initialkalibrierung werden im Rahmen von Kap. 5 die Werte aus der Online-Stereo-Kamerakalibrierung verglichen. Die dazu genutzte Funktion ist in Gl. (3.1) dargestellt.

$$\phi_j = \phi_{j_i} - \phi_{j_o} \text{ wobei } j = 1, 2, 3 \quad (3.1)$$

Der Wert ϕ_{j_o} ist dabei der, durch die Online-Stereo-Kamerakalibrierung , geschätzte Wert und der Wert ϕ_{j_i} ist der aus der Initialkalibrierung bestimmte Wert.

Mit einem externen Sensor kann die Qualität einer Online-Stereo-Kamerakalibrierung getestet werden. [Krüger et al., 2011] setzt den Velodyne-Laserscanner als externen Sensor ein, der die Genauigkeit der Initialkalibrierung überprüft. Der Velodyne-Laserscanner (siehe Abb. 3.5) zeichnet sich durch seine sehr hohe Genauigkeit aus und kann dadurch die Genauigkeit der Initialkalibrierung bewerten.

Ein Entfernungsfehler von $0,02 [m]$ auf $70 [m]$ bei einer Geschwindigkeit von $15 [\frac{U}{sec}]$ ist in der Spezifikation des Velodyne-Laserscanners angegeben. Dabei sendet der Velodyne einen Laserstrahl und misst die Zeit, die dieser Lichtstrahl benötigt, um wieder zurückzukommen. Aus dem Winkel, unter dem der Laserstrahl versendet wurde, und der Zeit, die der Laserstrahl benötigt hat, kann die 3D-Position eines Objekt berechnet werden. Die Rekonstruktion des Velodyne-Laserscanners wird dabei mit der 3D-Rekonstruktion der Stereo-Kamera verglichen.



Abbildung 3.5.: Der Velodyne-Laserscanner eignet sich, um die Kalibrierung von Kamerasystemen zu validieren, da der Velodyne-Laserscanner eine sehr hohe Entfernungsgenauigkeit von $0,02 [m]$ auf $70 [m]$ bei einer Umdrehungsgeschwindigkeit von $15 [\frac{U}{sec}]$ hat. Die Auflösung des Velodyne-Laserscanners hat eine horizontale Messdichte von ca. $2.000 [Pix] / 360 [^\circ]$ und eine vertikale Messdichte von $64 [Pix] / 26.8 [^\circ]$.

Die Stereo-Kamera und der Velodyne-Laserscanner können aufeinander kalibriert werden, um die 3D-Rekonstruktion beider Sensoren aufeinander abzugleichen. Wie die beiden Sensorsysteme aufeinander kalibriert werden, ist in Kap. A.3 beschrieben.

4. Genaue und robuste Online-Kalibrierung

Das kontinuierliche Überwachen und Anpassen der Kameraparameter über die Lebensdauer des Systems ist wichtig und wird von der Online-Kalibrierung sichergestellt. Diese Abweichungen von der Initialkalibrierung können durch thermische oder mechanische Einflüsse auf das Kamerasystem entstehen.

Die zwei in diesem Abschnitt beschriebenen Methoden bestimmen mit sehr wenigen Informationen aus dem Bild neue Parameter für eine Rektifizierung und die Methoden sind sehr robust und schnell anzuwenden. Um beide Methoden so robust zu realisieren, werden die Fahrzeug-Odometrie-Informationen, die im fahrenden Fahrzeug zur Verfügung stehen, genutzt.

Die erste Methode nutzt die Korrespondenzen und die Essential-Gleichung in einem Optimierungsschritt, um die Parameter in der Online-Kalibrierung neu zu schätzen.

Die zweite vorgestellte Methode nutzt geometrische Merkmale im Bild, die das Verwenden von Optimierungsverfahren, wie den Levenberg-Marquardt-Algorithmus, zur Berechnung der Winkel (ϕ_1, ϕ_2, ϕ_3) unnötig machen. So ist die Kalibrierung sehr robust und flexibel einsetzbar. Dies ist ein sehr großer Vorteil gegenüber den in der Literatur angegebenen Methoden zur Online-Kalibrierung, die einen Optimierungsschritt nutzen.

Die Online-Kalibrierung, die in dieser Arbeit beschrieben wird, ist ein Prozess mit mehreren Schritten, die in Abb. 4.1 dargestellt sind. Der erste Schritt ist die Rektifizierung der eingehenden Bildpaare. Dabei werden die Bilder so umgerechnet, dass korrespondierende Bildzeilen im rechten und linken Kamerabild zueinander passen. Um ein Bild zu rektifizieren muss eine Initialkalibrierung gemacht worden sein, da in die Rektifizierung die Parameter des Kamerasystems eingehen.

Der nächste Schritt der Online-Kalibrierung ist das Bestimmen von Korrespondenzen, um mit diesen die nächste Kalibrierung zu berechnen. Dazu können die Methoden aus Kap. 2.1 genutzt werden. Im Rahmen dieser Arbeit wurde die Methode nach [Stein, 2004] genutzt, da diese Methode wissenschaftlich untersucht ist und sich in unterschiedlichen Applikationen als robust erwiesen hat. Zudem ist diese Methode in dem verwendeten Framework implementiert und kann so verwendet werden. Durch die Fluss-Methode [Stein,

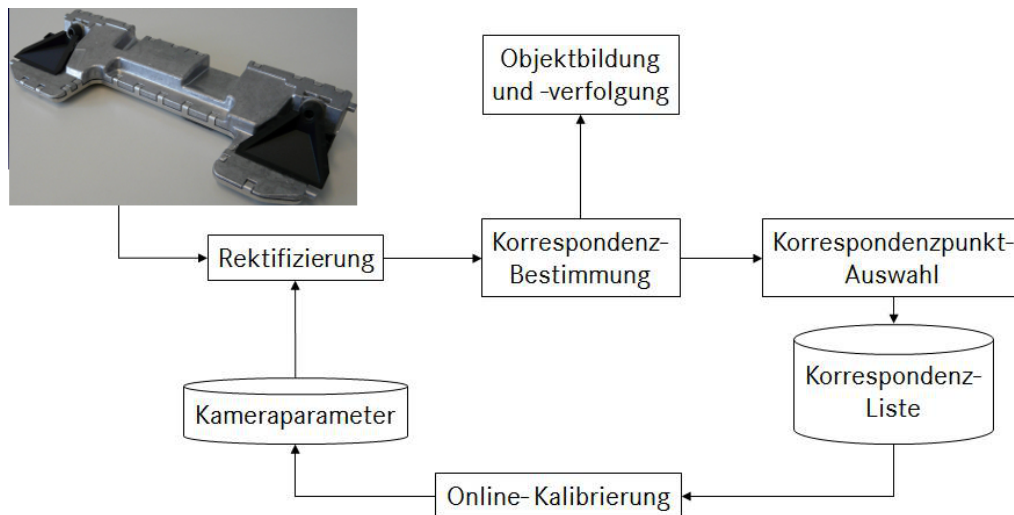


Abbildung 4.1.: Die Bilder werden in der Rektifizierung auf der Basis der Werte der Initialkalibrierung transformiert. Dann werden die Korrespondenzen in den rektifizierten Bildern unter Verwendung des Optischen-Flusses und der SGM-Algorithmus, bestimmt. Diese Liste wird dann in einer Kalibrierung genutzt, um die nächsten Werte der Rektifizierung zu berechnen.

2004] werden die Korrespondenzen in den rektifizierten und verzeichnungsfreien Bildern im Zeitschritt t_{n-1} und t_n bestimmt und durch die Semi-Global-Matching(SGM)-Methode werden die zu den Flussvektoren gehörenden Disparitäten bestimmt.

Mit diesen Korrespondenzen werden dann die Parameter für die nächste Rektifizierung bestimmt. Die Daten aus der Kalibrierung werden dann zur Rektifizierung der Bilder herangezogen. Im Rahmen dieser Arbeit werden die drei Rotationsparameter (ϕ_1, ϕ_2, ϕ_3) in der Essentialmatrix online geschätzt. Dabei ist die Schielwinkeländerung ϕ_2 der Winkel, der selbst bei geringen Abweichungen einen großen Fehler in der 3D-Rekonstruktion verursacht.

Abb. 4.2 zeigt den Teil der Kalibrierung des System-Diagramms aus Abb. 4.1, der in diesem Kapitel näher beschrieben wird. Die Korrespondenz-Bestimmung mit dem Fluss ist als gegeben vorausgesetzt. Da der Fluss durch die Fahrzeugbewegung signifikant beeinflusst wird und dadurch die Punkte im linken Bild den korrespondierenden Punkten im rechten Bild nicht immer richtig zugeordnet werden können, wird die Information über die Fahrtrichtung aus der Fahrzeug-Odometrie vom Control-Area-Network(CAN)-Bus des Fahrzeugs entnommen. Auf dem im Fahrzeug integrierten CAN sind die Informationen aller Sensoren in einem Fahrzeug kodiert, so auch die Information der Gierrate. Aus der Gierrate kann entnommen werden, ob das Fahrzeug im Augenblick geradeaus fährt.

Durch Einbeziehung der Daten aus dem CAN-Bus werden die Fehler, die durch Ungenauigkeiten des Flusses in die Kalibrierung mit eingebracht werden, verringert. Um eine geeignete Ausgangsposition für die Online-Kalibrierung zu sichern, ist die Kalibrierung

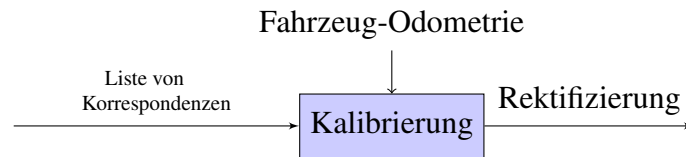


Abbildung 4.2.: Die Fahrzeug-Odometrie, d.h. die Fahrzeuggeschwindigkeit und der Gierwinkel, wird für die Kalibrierung der drei Rotationswinkeländerungen ϕ_1, ϕ_2, ϕ_3 genutzt.

der Stereo-Kamera, anders als bei den anderen Verfahren, in zwei Stufen unterteilt. Der erste Schritt schätzt den Roll- und Nickwinkel, wobei der Schielwinkel aus der Initialkalibrierung als fest angenommen wird. Nachdem durch die Schätzung von ϕ_1 und ϕ_3 die Epipolarlinie in beiden Bildern bis auf ein Pixel übereinstimmen, wird die zweite Stufe genutzt, um den Schielwinkel zu schätzen und so die Epipolarlinie in Deckung zu bringen. Die Grenze von einem Pixel wurde empirisch durch eine Vielzahl von Untersuchungen bestimmt.

Um eine Verteilung der Korrespondenzen gleichmäßig über das Bild zu erhalten, wird das Binning eingesetzt. Durch wissenschaftliche Untersuchungen im Rahmen dieser Arbeit wurde gezeigt, dass sich eine Gleichverteilung der Korrespondenzen über das gesamte Bild am besten für eine Kalibrierung eignet. Binning ist hier das Teilen eines Bildes in Bildsektoren, die eine maximale Anzahl von Korrespondenzen aufnehmen können. Durch eine Gleichverteilung der Korrespondenzen wird die Genauigkeit der Kalibrierung für den Schielwinkel verbessert [Feldmann et al., 2011]. In [Feldmann et al., 2011] wurde zudem gezeigt, dass die Verteilung der Korrespondenzen einen größeren Einfluss auf die Kalibrierung hat als die Wahl der robusten Fehlerfunktion, die eingesetzt wird [Hartley und Zisserman, 2004]. Näheres zum Binning ist in Kap. 4.3 zu finden.

Nach dem Bestimmen der Korrespondenzen wird die Rektifizierung nach [Fusiello et al., 2000] berechnet. Die Theorie zur Rektifizierung ist in [Hartley, 1999; Wöhler, 2009; Faugeras, 1993] beschrieben.

Methoden zur Schätzung des Schielwinkels unabhängig von Roll- und Nickwinkel für die Essentialmatrix werden in diesem Kapitel beschrieben. In den in Kap. 2 beschriebenen Methoden werden die Parameter der Essentialmatrix in einem Schritt geschätzt.

4.1. Online-Kalibrierung auf Basis von Entfernungsdaten

In diesem Abschnitt sind wir davon ausgegangen, dass die erste Stufe der Online-Kalibrierung die korrespondierenden Bildzeilen auf ein Pixel genau bestimmt hat. Um im zweiten Schritt der Kalibrierung die drei Winkel (ϕ_1, ϕ_2, ϕ_3) zu berechnen, werden in der hier be-

schriebenen Methode die Odometrie-Informationen genutzt, um den zurückgelegten Weg zwischen zwei Aufzeichnungen zu berechnen und durch die Odometrie-Information wird sichergestellt, dass das Fahrzeug geradeaus fährt. Die gerade Fahrt ist eine wichtige Voraussetzung an diese hier beschriebene Methode. Wird nun der Punkt der 3D-Rekonstruktion genutzt, der den Koordinaten der Korrespondenz im linken Bild entspricht, dann kann die Distanz aus der 3D-Rekonstruktion mit dem zurückgelegten Weg verrechnet werden. Der zurückgelegte Weg und die Distanz aus der 3D-Rekonstruktion müssen gleich sein, wenn die Stereo-Kamera kalibriert ist.

Wissenschaftliche Untersuchungen haben gezeigt, dass das Signal-zu-Rausch-Verhalten der Sensordaten bei einer Fahrzeug-Geschwindigkeit von ca. $8 \left[\frac{m}{s} \right]$ bis ca. $36 \left[\frac{m}{s} \right]$ genügend groß ist, dass die Kalibrierung einen geringen Unterschied zur Ground Truth zeigt. Außerhalb dieses Intervalls ist das Signal-zu-Rausch-Verhalten klein und die Kalibrierung wird ungenau. In dem Signal-zu-Rausch-Verhalten ist der optische Fluss, als auch das SGM-Verfahren zur Disparitätsbestimmung, schon mit eingeflossen.

Mit der Geschwindigkeitsangabe aus dem Steuergerät der Elektronischen Stabilitätsregelung (ESP), die im CAN-Bus abgefragt werden kann, und der Zeit, die zwischen den beiden Aufnahmen vergangen ist, kann die Distanz berechnet werden, die zwischen den beiden Aufnahmen zurückgelegt wurde. In $\mathcal{I}_{\tau-1}$ wird eine 3D-Rekonstruktion des Bereiches im Stereo-Kamera-Sichtbereich mittels Triangulierung aus den bestimmten Korrespondenzen berechnet. Dies wird dann im Bild \mathcal{I}_{τ} wiederholt. Die Differenz dieser beiden Triangulierungen muss der durch das Fahrzeug gefahrenen Strecke, bis auf einen kleinen Fehler, gleichen. Die Korrespondenzen, die dieser Bedingung genügen, werden für die Berechnung der neuen Parameter der Rektifizierung genutzt.

Sei \mathcal{D} die Distanz, die zwischen zwei Aufnahmen vom Fahrzeug zurückgelegt wurde. Dann folgt

$$\begin{aligned} \mathcal{D} &= z_{\tau-1} - z_{\tau} \\ &= \left(\frac{f \cdot b}{\hat{d}} \right)_{\tau-1} - \left(\frac{f \cdot b}{\hat{d}} \right)_{\tau} \\ &= \frac{f \cdot b}{\hat{d}_{\tau-1}} - \frac{f \cdot b}{\hat{d}_{\tau}}, \end{aligned} \tag{4.1}$$

mit f die Brennweite, b die Basisbreite des Stereo-Kamera-Systems, $\hat{d} = d + \Delta d$ die Disparität mit Fehler Δd . Abb. 4.3 stellt dies graphisch dar.

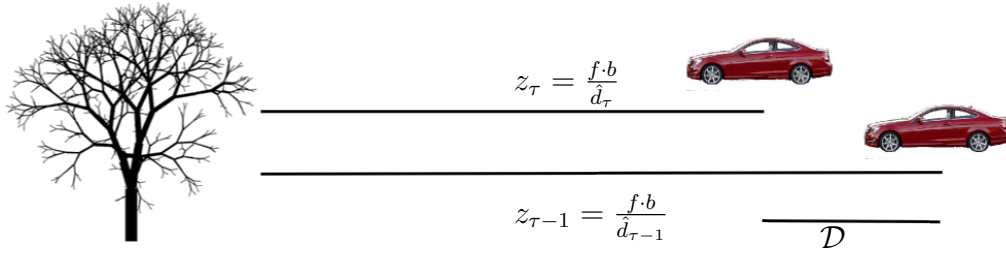


Abbildung 4.3.: Die Distanz zwischen dem Kamerasystem und einem unbewegten Objekt wird in der 3D-Rekonstruktion der Stereo-Kamera gemessen. Wird diese Distanz zwischen dem Objekt und dem Stereo-Kamera-System zum Zeitpunkt $\tau - 1$ und τ gemessen, kann die Differenz \mathcal{D} aus den beiden 3D-Rekonstruktionen bestimmt werden. Diese Distanz \mathcal{D} entspricht dem zurückgelegten Weg des Fahrzeugs. Dies ist in Gl. (4.1) beschrieben.

Sei der Fehler Δd in der 3D-Rekonstruktionen zum Zeitpunkt $\tau - 1$ und τ gleich, dann folgt mit Gl. (4.1)

$$\begin{aligned} \mathcal{D} &= z_{\tau-1} - z_\tau \\ &= \frac{\mathcal{F}}{\hat{d}_{\tau-1}} - \frac{\mathcal{F}}{\hat{d}_\tau} \end{aligned} \quad (4.2)$$

werden. Die Terme auf einen gemeinsamen Nenner gebracht und mit diesem Nenner multipliziert, folgt

$$\hat{d}_{\tau-1} \hat{d}_\tau \mathcal{D} = \mathcal{F} \underbrace{(\hat{d}_\tau - \hat{d}_{\tau-1})}_{=(d_\tau - d_{\tau-1})}. \quad (4.3)$$

Da \mathcal{D} ein Skalar ist, folgt mit einer Substitution von $\mathcal{F}(d_\tau - d_{\tau-1})\mathcal{D}^{-1} \equiv \hat{\mathcal{C}}$

$$(d_{\tau-1} + \Delta d)(d_\tau + \Delta d) = \hat{\mathcal{C}}. \quad (4.4)$$

Wird nun die linke Seite der Gleichung ausmultipliziert, folgt

$$\Delta d^2 + (d_{\tau-1} + d_\tau)\Delta d + d_{\tau-1}d_\tau = \hat{\mathcal{C}} \quad (4.5)$$

und wird diese Gleichung normalisiert und $\hat{\mathcal{C}} - d_{\tau-1}d_\tau = \mathcal{C}$, folgt

$$\Delta d^2 + (d_{\tau-1} + d_\tau)\Delta d - \mathcal{C} = 0. \quad (4.6)$$

Diese Gleichung kann dann mit der pq-Formel gelöst werden

$$\Rightarrow \Delta d_{1,2} = -\frac{d_{\tau-1} + d_\tau}{2} \pm \sqrt{\frac{(d_{\tau-1} + d_\tau)^2}{4} + \mathcal{C}}. \quad (4.7)$$

Die Gleichung Gl. (4.7) wird genutzt, um den Schielwinkel ϕ_2 einzustellen.

Durch die Funktion können nun die Korrespondenzen klassifiziert werden. Alle Korrespondenten, bei denen $\Delta d_{1,2}$ kleiner als ein Schwellwert σ ist, sind für eine Kalibrierung geeignet.

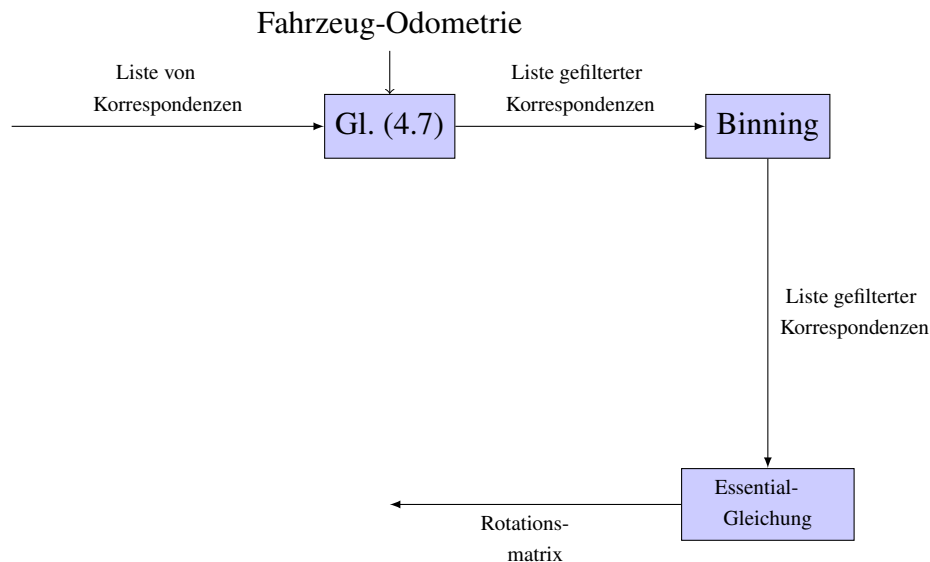


Abbildung 4.4.: Die Kalibrierung aus Abb. 4.1 wird in drei Schritte unterteilt. Durch Gl. (4.7) können die Korrespondenzen, die sich zur Kalibrierung eignen, von den Ausreißern getrennt werden. Nachdem die Korrespondenzen ausgesiebt wurden, werden die Korrespondenzen im Bild gesammelt, sodass sich eine Gleichverteilung ergibt. Dieses Anordnen nennen wir Binning. Die nun übrigen Korrespondenzen aus dem Binning werden dann für die Optimierung der Parameter für die Rektifizierung aus der Essential-Gleichung genutzt.

Abb. 4.4 ist eine Darstellung, wie Gl. (4.7) sich in das System der Online-Kalibrierung einbettet. Die Kalibrierung durchläuft drei Schritte, bis die Rektifizierung die Werte aus der Kalibrierung nutzen kann.

Im ersten Schritt werden alle Korrespondenzen durch die Gl. (4.7) von den Ausreißern getrennt. Im zweiten Schritt wird dann das Binning durchgeführt, sodass die Korrespondenzen auf dem Bild gleichverteilt sind. Genauer zum Binning ist in Kap. 4.3 beschrieben. Im letzten und dritten Schritt werden die im Binning bestimmten Korrespondenzen genutzt, um im Optimierungsschritt die neuen Parameter für die Rektifizierung aus der Essential-Gleichung zu schätzen.

4.2. Online-Kalibrierung mit Bezug auf die Fluchtpunkte in beiden Bildern

Die zuvor beschriebene Methode zur Berechnung der Veränderung des Schielwinkels ϕ_2 erfordert die Kenntnis der Geschwindigkeit und des Gierwinkels des Fahrzeugs, da die

Methoden zur Berechnung der Korrespondenzen durch Einbeziehen von Gierwinkel- und Geschwindigkeitsinformation genauer werden. Zusätzlich reagiert die Kalibrierung sensitiv auf die Verteilung der Korrespondenzen über die Bildfläche. Dies kann zu ungewollten Fehlern in der Kalibrierung führen und so falsche Werte an die Rektifizierung weiterleiten. Die nun dargestellte Methode ist die zweite Methode zur Online-Kalibrierung in dieser Arbeit. Sie ist nicht mehr von der Geschwindigkeit und der Verteilung der Korrespondenzen abhängig. Bei der zweiten Methode ist nur die Angabe „Geradeausfahrt“ aus den CAN-Daten erforderlich und nicht mehr wie in der ersten Methode „Geradeausfahrt“ und „Geschwindigkeit“.

Aus der projektiven Geometrie ist bekannt, dass sich Parallelen im Raum in einem Punkt im Bild – dem Fluchtpunkt – schneiden müssen. Der Fluchtpunkt hat eine Disparität von Null, d.h. der Fluchtpunkt ist unendlich weit entfernt. Deshalb müssen die Koordinaten der Fluchtpunkte im linken und rechten Bild identisch sein. Die horizontalen Koordinaten der Fluchtpunkte in beiden Bildern können dann die Abweichung des Schielwinkels des Stereo-Kamera-Systems angeben. Die hier beschriebene Methode bestimmt in beiden korrespondierenden Bildern den Fluchtpunkt und berechnet aus diesem die Abweichung des Schielwinkels vom Sollwinkel.

Zuerst ist zu zeigen, dass sich die Flussvektoren bei Geradeausfahrt des Fahrzeugs in einem Punkt im Bild schneiden, dem sogenannten Fluchtpunkt. Hierzu seien

$$p_0 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_0 \quad \text{und} \quad p_1 = \begin{pmatrix} x \\ y \\ z \end{pmatrix}_1 \quad (4.8)$$

zwei Punkte im Weltkoordinatensystem zum Zeitpunkt $t = 0$ und $t = 1$ und es sei $\tilde{\mathbf{d}} = (dx, dy, dz)^t = p_1 - p_0$ der Abstand zweier Punkte im Weltkoordinatensystem. Dann gilt

$$p_1 = p_0 - \tilde{\mathbf{d}}. \quad (4.9)$$

Aus der projektiven Geometrie, wie in [Wöhler, 2009; Hartley und Zisserman, 2004; Kanatani, 2005; Luhmann, 2010] beschrieben, folgt, dass der Punkt p_0 im linken Bild auf die Koordinaten $(u, v, d)_0$

$$\begin{pmatrix} u \\ v \\ d \end{pmatrix}_0 = \begin{pmatrix} \frac{f \cdot y_0}{x_0} \\ \frac{f \cdot z_0}{x_0} \\ \frac{F}{x_0} \end{pmatrix} \quad (4.10)$$

und dass p_1 auf die Koordinaten $(u, v, d)_1$

$$\begin{pmatrix} u \\ v \\ d \end{pmatrix}_1 = \begin{pmatrix} \frac{f \cdot y_1}{x_1} \\ \frac{f \cdot z_1}{x_1} \\ \frac{f}{x_1} \end{pmatrix} \quad (4.11)$$

abgebildet werden.

Dann kann der Flussvektor mit den Koordinaten d_u (für die horizontale Koordinate) und d_v (für die vertikale Koordiante), gegeben durch

$$\begin{pmatrix} d_u \\ d_v \end{pmatrix} = \begin{pmatrix} u_1 - u_0 \\ v_1 - v_0 \end{pmatrix} \quad (4.12)$$

berechnet werden. Die zugehörige Disparität d kann dann berechnet werden mit

$$d = d_1 - d_0. \quad (4.13)$$

Dann ist die Geradengleichung (g_u, g_v) für die horizontale Bildkoordinate u und die vertikale Bildkoordinate v mit $g_u = \frac{b(d_y f - d_x u)}{d_y x_0 - d_x y_0}$ und $g_v = \frac{b(d_z f - d_x v)}{d_z x_0 - d_x z_0}$ gegeben. Wenn g_u nach u aufgelöst wird, folgt

$$u = \frac{d_y f}{d_x} \quad (4.14)$$

und wenn g_v nach v aufgelöst wird, folgt

$$v = \frac{d_z f}{d_x}. \quad (4.15)$$

Setzen wir nun $(u, v) = (0, 0)_i$ mit $i = (1, 2)$, folgt, dass $d_x > 0$ sein muss, und da für die Kamerakonstante $f > 0$ gilt, folgt, dass für den Punkt $d_y = 0$ und $d_z = 0$ gelten muss. D.h. der Fluchtpunkt kann durch die Flussvektoren im Bild bestimmt werden, wenn sich der Weltpunkt p zum Zeitpunkt $t = 0$ und $t = 1$ auf der gleichen Position befindet und sich das Fahrzeug auf diesen Punkt geradeaus zubewegt.

Die zweite entwickelte Methode zur Online-Kalibrierung bestimmt damit im linken und im rechten Bild den Fluchtpunkt. Mit Gl. (4.14) und Gl. (4.15) folgt, dass die Fluchtpunkte existieren und durch den Schnittpunkt der Flussvektoren gefunden werden können. Mit der Gleichung Gl. (4.16) kann der Schnittpunkt aus der Menge der Korrespondenzen berechnet werden.

$$\begin{aligned} u_i &= \frac{1}{O} \left| \begin{array}{cc|c} l u_\tau & l v_\tau & (l u_\tau - r u_\tau) \\ r u_\tau & r v_\tau & \\ \hline l u_{\tau-1} & l v_{\tau-1} & (l u_{\tau-1} - r u_{\tau-1}) \\ r u_{\tau-1} & r v_{\tau-1} & \end{array} \right| \\ &= \frac{1}{O} [(l u_{\tau-1} - r u_{\tau-1})(l u_\tau v_\tau - r u_\tau l v_\tau) \end{aligned} \quad (4.16)$$

$$\begin{aligned}
& - ({}_l u_\tau - {}_r u_\tau)({}_l v_{\tau-1} - {}_r v_{\tau-1})] \quad (4.17) \\
\text{mit } O &= \begin{vmatrix} {}_l u_\tau - {}_r u_\tau & {}_l v_\tau - {}_r v_\tau \\ {}_l v_{\tau-1} - {}_r v_{\tau-1} & {}_l u_{\tau-1} - {}_r u_{\tau-1} \end{vmatrix} \\
&= ({}_l u_\tau - {}_r u_\tau) \cdot ({}_l v_{\tau-1} - {}_r v_{\tau-1}) \\
&- ({}_l v_\tau - {}_r v_\tau) \cdot ({}_l u_{\tau-1} - {}_r u_{\tau-1}) \\
&\neq 0
\end{aligned}$$

mit

$(u, v)_\tau^T$ der Bildpunkt $(u, v)^T$ links im linken Bild zum Zeitpunkt τ , $(u, v)_\tau^T$ der Punkt $(u, v)^T$ rechts im linken Bild zum Zeitpunkt τ und analog die Punkte im Bild zum Zeitpunkt $\tau - 1$. Dabei ist $(u, v)_\tau^T$ der linke Bildpunkt im linken Bild zum Zeitpunkt τ und $(u, v)_{\tau-1}^T$ ist der korrespondierende rechte Bildpunkt zum Zeitpunkt $\tau - 1$ im linken Bild. Der Bildpunkt $(u, v)_\tau^T$ und $(u, v)_{\tau-1}^T$ bilden einen Vektor und äquivalent die rechten korrespondierenden Bildpunkte $(u, v)_\tau^T$ und $(u, v)_{\tau-1}^T$ im linken Bild. Die Bildpunkte werden durch das addieren der Disparität ins rechte Bild translatiert. Sollte $O = 0$ gelten, dann sind die Vektoren parallel oder linear abhängig, und so existiert kein Schnittpunkt in einem Bild oder es existieren unendlich viele Schnittpunkte.

Um ein Stereo-Kamera-System zu kalibrieren, werden die Koordinaten der Fluchtpunkte in der horizontalen Ausrichtung im linken und rechten Bild verglichen. Sind diese Koordinaten gleich, ist das Stereo-Kamera-System kalibriert und weichen die Koordinaten voneinander ab, dann kann die Abweichung mit

$$\Delta U = u_l - u_r, \quad (4.18)$$

mit u_l die Koordinate des Fluchtpunktes im linken Bild und u_r die Koordinate des Fluchtpunktes im rechten Bild gemessen werden. ΔU nennen wir den Schielwinkelfehler. Der Schielwinkelfehler ist in Pixelkoordinaten angegeben und kann durch

$$\begin{aligned}
\tan(\beta) &= \frac{\Delta U}{f} \\
\beta &= \arctan\left(\frac{\Delta U}{f}\right) \quad (4.19)
\end{aligned}$$

in den Winkelraum für den Rotationsanteil der Essentialmatrix umgerechnet werden. Dabei ist f die Brennweite und ΔU die Größe des Schielwinkelfehlers in Pixeln. Abb. 4.5 zeigt den Zusammenhang zwischen dem Winkelraum und dem Pixelraum.

Mit dieser Methode ändert sich der Teil der Kalibrierung des Systemdiagramms Abb. 4.4 zu Abb. 4.6

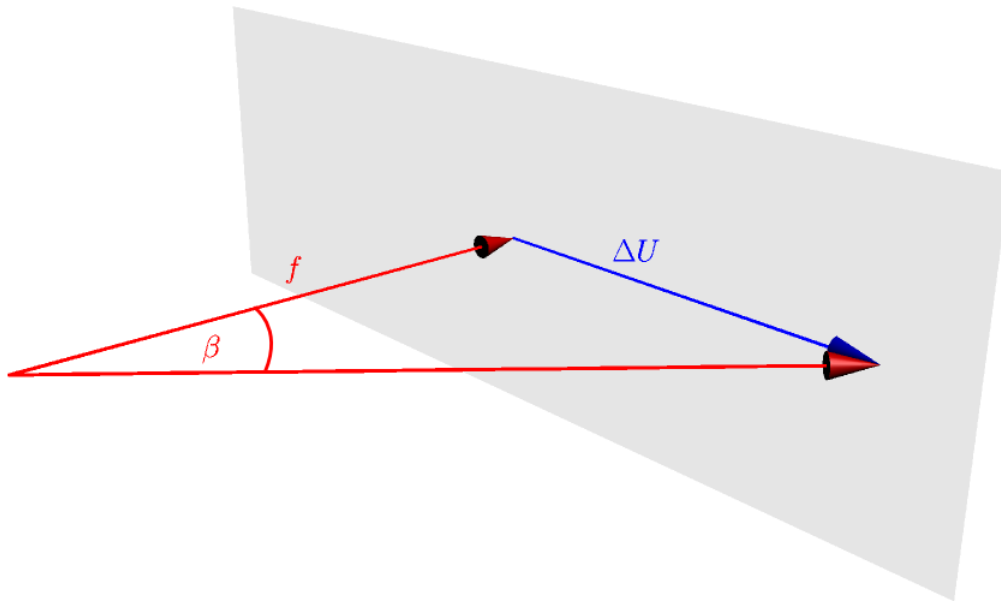


Abbildung 4.5.: Der Schielwinkelfehler ΔU kann in Pixelkoordinaten angegeben werden. Ist ΔU der Schielwinkelfehler in Pixeln, und f die Brennweite, dann ist der Schielwinkelfehler in Winkelkoordinaten Gl. (4.19).

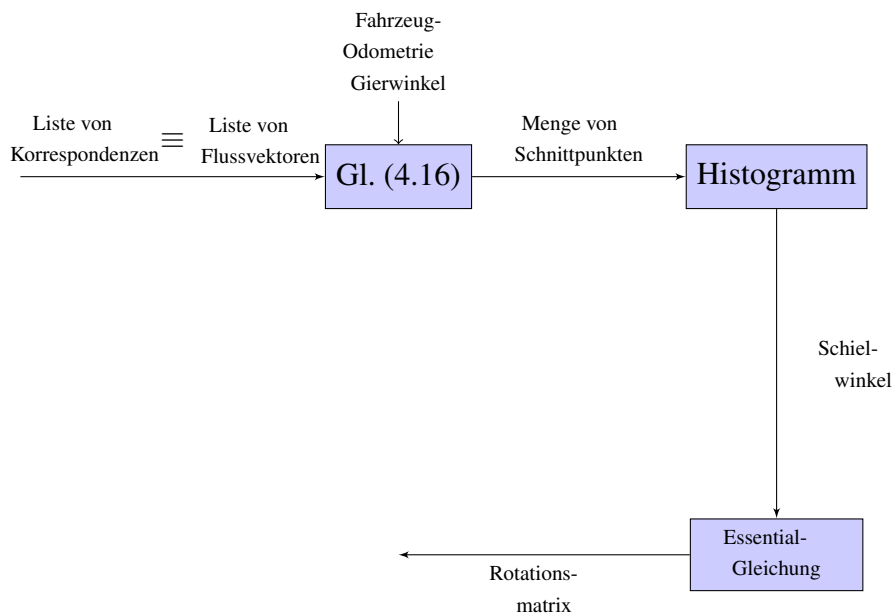


Abbildung 4.6.: Die Kalibrierung aus Abb. 4.4 wird in drei Schritte unterteilt. Zuerst wird bei gerader Fahrt der Fluchtpunkt berechnet. Dann ist der Schielwinkelfehler die Differenz der horizontalen Koordinate des Fluchtpunktes aus dem linken und rechten Bild. Dieser Fehler kann vom Pixelkoordinatensystem ins Bogenmaß umgerechnet werden. Durch die Essential-Gleichung wird die Rotationsmatrix der Essentialmatrix bestimmt. Die Rotationsmatrix wird dann in der Rektifizierung genutzt, um im nächsten Schritt die Bilder zu transformieren.

Das Bestimmen des Fluchtpunktes wurde in der Literatur zur Online-Kalibrierung noch nicht verwendet. Dies ist die erste Verwendung des Fluchtpunktes zur Online-Kalibrierung.

4.3. Analyse von Einflüssen

Sensordaten haben ein additives Rauschen. Das Rauschen der Daten hat einen Einfluss auf die Korrespondenzen und auf die in dieser Arbeit neu entwickelten Methoden. In [Dang, 2007] und [Rabe, 2011] werden die hier eingesetzten Standard-Methoden wie Fluss und Semi-Global-Matching auf ihre Genauigkeit und die Robustheit untersucht. Es wurden keine Analysen über den Einfluss von Rauschen auf die Sensordaten bzgl. Fehlern in den Messungen und Rekonstruktionen durchgeführt. In diesem Abschnitt wird gezeigt, wie sich das Rauschen auf die Komponenten der Korrespondenzen auswirkt und so auf die Winkel, die wir für die Online-Kalibrierung schätzen. Dieses Kapitel wird mit der Betrachtung der Verzeichnung durch die Windschutzscheibe abgeschlossen.

4.3.1. Einfluss der Komponenten einer Korrespondenz auf die Online-Kalibrierung

In diesem Abschnitt wird gezeigt, wie sich die geometrische Lage der Korrespondenzen auf die Online-Kalibrierung auswirkt. Dazu wird betrachtet, wie die Komponenten der Korrespondenzen (x_l, x_r) mit $x_l = (u, v, 1)_l^T$ und $x_r = (u, v, 1)_r^T$ sich durch die Essential-Gleichung transformieren. Aus dieser Transformation lässt sich dann eine geometrische Interpretation über den Einfluss der Komponenten der Korrespondenz auf die Kalibrierung entnehmen. Da im Rahmen dieser Arbeit die Rotationsparameter (ϕ_1, ϕ_2, ϕ_3) betrachtet werden, fokussieren wir uns hier auf den Einfluss dieser Komponenten der Korrespondenzen durch die Transformation, die durch die Rotation der Essential-Gleichung beschrieben wird. Aus Gl. (1.2) folgt

$$x_l^T \mathbb{E} x_r = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_l^T \mathbf{T}_\times \mathbf{R}(\mathbf{x}, \phi_1) \mathbf{R}(\mathbf{y}, \phi_2) \mathbf{R}(\mathbf{z}, \phi_3) \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_r, \quad (4.20)$$

mit $(u, v)^T \in [\pm\mathbb{R}]$ normalisiert. Sei $\mathbb{E}(\phi_1, \phi_2, \phi_3)$ der Einfluss der Komponenten der Korrespondenzen auf die Essential-Gleichung, dann ist $\mathbb{E}(\phi_1)$ der Einfluss von Winkel ϕ_1 , um die x-Achse auf die Essential-Gleichung, $\mathbb{E}(\phi_2)$ der Einfluss von Winkel ϕ_2 , um die y-Achse und $\mathbb{E}(\phi_3)$ der Einfluss des Winkels ϕ_3 , um die z-Achse, auf die Essential-Gleichung.

Um die Auswirkung der Komponenten der Korrespondenzen auf die Kalibrierung zu bestimmen, wird die Essential-Gleichung berechnet

$$\begin{aligned}
\mathbb{E}(\phi_1, \phi_2, \phi_3) &= \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_l^t \mathbf{T}_\times \mathbf{R}(\mathbf{x}, \phi_1) \mathbf{R}(\mathbf{y}, \phi_2) \mathbf{R}(\mathbf{z}, \phi_3) \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}_r \\
&= -t_1 \sin(\phi_1) \cos(\phi_2) - t_2 \sin(\phi_2) \\
&+ u_l(t_2 \cos(\phi_1) + t_3 \sin(\phi_1)) \cos(\phi_2) \\
&+ u_r(t_1(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) \\
&+ \sin(\phi_3) \cos(\phi_1)) - t_2 \cos(\phi_2) \cos(\phi_3) \\
&+ u_l(t_2(\sin(\phi_1) \sin(\phi_3) - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3)) \\
&- t_3(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) + \sin(\phi_3) \cos(\phi_1))) \\
&- v_l(t_1(\sin(\phi_1) \sin(\phi_3) - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3)) \\
&- t_3 \cos(\phi_2) \cos(\phi_3)) - v_l(t_1 \cos(\phi_1) \cos(\phi_2) \\
&- t_3 \sin(\phi_2)) - v_r(t_1(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \\
&- \cos(\phi_1) \cos(\phi_3)) - t_2 \sin(\phi_3) \cos(\phi_2) \\
&- u_l(t_2(\sin(\phi_1) \cos(\phi_3) + \sin(\phi_2) \sin(\phi_3) \cos(\phi_1)) \\
&+ t_3(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) - \cos(\phi_1) \cos(\phi_3))) \\
&+ v_l(t_1(\sin(\phi_1) \cos(\phi_3) \\
&+ \sin(\phi_2) \sin(\phi_3) \cos(\phi_1)) + t_3 \sin(\phi_3) \cos(\phi_2))). \tag{4.21}
\end{aligned}$$

Das Minimum der Funktion $\mathbb{E}(\phi_1, \phi_2, \phi_3)$ ist der optimale Wert der Winkel für die Kalibrierung. Um das Minimum der Funktion zu bestimmen, wird das Differential $\nabla \mathbb{E}(\phi_1, \phi_2, \phi_3)$ berechnet und Null gesetzt.

$$\nabla \mathbb{E}(\phi_1, \phi_2, \phi_3) = \left(\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_1}, \frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_2}, \frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_3} \right)^t \tag{4.22}$$

mit

$$\begin{aligned}
\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_1} &= t_1 v_l \sin(\phi_1) \cos(\phi_2) - t_1 \cos(\phi_1) \cos(\phi_2) \\
&- u_l(t_2 \sin(\phi_1) - t_3 \cos(\phi_1)) \cos(\phi_2) \\
&- u_r(t_1 v_l(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) + \sin(\phi_3) \cos(\phi_1)) \\
&+ t_1(\sin(\phi_1) \sin(\phi_3) - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3)) \\
&- u_l(t_2(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) + \sin(\phi_3) \cos(\phi_1)) \\
&+ t_3(\sin(\phi_1) \sin(\phi_3) - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3))) \\
&- v_r(-t_1 v_l(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) - \cos(\phi_1) \cos(\phi_3)) \tag{4.23}
\end{aligned}$$

$$\begin{aligned}
& + t_1(\sin(\phi_1) \cos(\phi_3) + \sin(\phi_2) \sin(\phi_3) \cos(\phi_1)) \\
& + u_l(t_2(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) - \cos(\phi_1) \cos(\phi_3)) \\
& - t_3(\sin(\phi_1) \cos(\phi_3) + \sin(\phi_2) \sin(\phi_3) \cos(\phi_1))) = 0 \\
\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_2} & = t_1 \sin(\phi_1) \sin(\phi_2) - t_2 \cos(\phi_2) - u_l(t_2 \cos(\phi_1) \\
& + t_3 \sin(\phi_1)) \sin(\phi_2) + u_r(t_1 \sin(\phi_1) \cos(\phi_2) \\
& + t_2 \sin(\phi_2) - u_l(t_2 \cos(\phi_1) + t_3 \sin(\phi_1)) \cos(\phi_2) \\
& + v_l(t_1 \cos(\phi_1) \cos(\phi_2) - t_3 \sin(\phi_2))) \cos(\phi_3) \\
& + v_l(t_1 \sin(\phi_2) \cos(\phi_1) + t_3 \cos(\phi_2)) \quad (4.24) \\
& - v_r(t_1 \sin(\phi_1) \cos(\phi_2) + t_2 \sin(\phi_2) - u_l(t_2 \cos(\phi_1) \\
& + t_3 \sin(\phi_1)) \cos(\phi_2) + v_l(t_1 \cos(\phi_1) \cos(\phi_2) \\
& - t_3 \sin(\phi_2))) \sin(\phi_3) = 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_3} & = -u_r(t_1(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) - \cos(\phi_1) \cos(\phi_3)) \\
& - t_2 \sin(\phi_3) \cos(\phi_2) - u_l(t_2(\sin(\phi_1) \cos(\phi_3) \\
& + \sin(\phi_2) \sin(\phi_3) \cos(\phi_1) + t_3(\sin(\phi_1) \sin(\phi_2) \sin(\phi_3) \\
& - \cos(\phi_1) \cos(\phi_3))) + v_l(t_1(\sin(\phi_1) \cos(\phi_3) \\
& + \sin(\phi_2) \sin(\phi_3) \cos(\phi_1) + t_3 \sin(\phi_3) \cos(\phi_2))) \quad (4.25) \\
& - v_r(t_1(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) + \sin(\phi_3) \cos(\phi_1)) \\
& - t_2 \cos(\phi_2) \cos(\phi_3) + u_l(t_2(\sin(\phi_1) \sin(\phi_3) \\
& - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3)) - t_3(\sin(\phi_1) \sin(\phi_2) \cos(\phi_3) \\
& + \sin(\phi_3) \cos(\phi_1))) - v_l(t_1(\sin(\phi_1) \sin(\phi_3) \\
& - \sin(\phi_2) \cos(\phi_1) \cos(\phi_3)) - t_3 \cos(\phi_2) \cos(\phi_3))) = 0
\end{aligned}$$

Es wird in dieser Berechnung angenommen, dass die Kameraparameter sich nur geringfügig ändern, d.h. α ist sehr klein. Dann kann die Approximation

$$\begin{pmatrix} \sin(\alpha) \\ \cos(\alpha) \end{pmatrix} \approx \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (4.26)$$

in der Berechnung genutzt werden. Weiter ist im Rahmen dieser Arbeit angenommen, dass die Translation zwischen beiden Kameras im Kamerasystem nur entlang der x-Achse ist. Dann folgt, dass der Translationsvektor $\mathbf{t} = (t_1, 0, 0)^T$ ist. Mit Gl. (4.26) können die Gleichungen Gl. (4.23) bis Gl. (4.25) zu

$$\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_1} \approx -t_1(v_l v_r + 1) \quad (4.27)$$

$$\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_2} \approx t_1 u_r v_l \quad (4.28)$$

$$\frac{\partial \mathbb{E}(\phi_1, \phi_2, \phi_3)}{\partial \phi_3} \approx t_1 u_r. \quad (4.29)$$

vereinfacht werden. Die Gleichungen Gl. (4.27) bis Gl. (4.28) sind ein Bewertungsmaß, um die Rangfolge der Korrespondenzen mit Hinblick auf den gesuchten Parameter zu bewerten. Sind die Korrespondenzen, für den betrachteten Winkel ϕ_1 , ϕ_2 oder ϕ_3 , durch die Gleichungen Gl. (4.27) bis Gl. (4.28) geordnet, dann sind die Korrespondenzen nach dem Einfluss auf die Kalibrierung des Winkels dem Rang nach sortiert.

Aus Gl. (4.27) folgt, dass $\mathbb{E}(\phi_1)$ minimal wird, wenn $v_l v_r = -1$ gilt, d.h. dass mindestens eine v-Koordinate der Korrespondenz $k := (u_l, v_l, u_r, v_r)^T$ gleich dem negativen Inversen der anderen sein muss.

$\mathbb{E}(\phi_2)$ wird minimal, wenn $u_r = 0$ oder $v_l = 0$ ist. Dies folgt aus Gl. (4.28).

Der Rollwinkel ϕ_3 aus dem \mathbb{E} wird minimal, wenn die Koordinate $u_r = 0$ ist. Dies folgt unmittelbar aus Gl. (4.29).

Dies zeigt, dass der Schielwinkel ϕ_2 und der Rollwinkel ϕ_3 beide von der u Komponente des rechten Bildes beeinflusst werden. Diese Zusammenhänge werden in Kap. 5 anhand realer Daten experimentell verifiziert.

4.3.2. Binning

Korrespondenzen sind häufig in Clustern im Bild angeordnet. Diese Cluster von Korrespondenzen können die Kalibrierung verschlechtern, da diese auf einem sehr eng begrenzten Bereich im Bild zu finden sind. Oft sind einige Bereiche im Bild sehr dünn mit Korrespondenzen besetzt. Der Himmel z.B. hat sehr wenige Korrespondenzen, wohingegen die Straßenoberfläche oder die Bordsteine viele Korrespondenzen erzeugen. Um eine möglichst gleiche Dichte von Korrespondenzen über das gesamte Bild bereitzustellen, wird das Binning eingesetzt.

Dazu wird, beispielsweise durch die Fluss-Methode von [Stein, 2004], festgestellt, wo die Korrespondenzen im Bild sind. Anschließend wird betrachtet, ob die Verteilung der Korrespondenzen im Bild einer vorgegebenen Modell-Verteilung entspricht (Kap. 4.3.3). Um eine Verteilung zu approximieren, werden Sektoren im Bild eingerichtet. Diese Sektoren sind so verteilt, dass nur eine vorgegebene Menge von Korrespondenzen in diesen Sektor aufgenommen werden kann. Jeder dieser Sektoren wird als ein Bin bezeichnet. Der Vorgang der Zuweisung von Korrespondenzen in den richtigen Bin wird Binning genannt. In Abb. 4.7 ist dies eine Gleichverteilung. Im Rahmen dieser Arbeit wurden verschiedene Verteilungen untersucht, die Auswertung dieser Verteilungen ist in Kap. 5.2 beschrieben. Kamerakalibrierungen werden mit Punkten mit unterschiedlichen Abständen zum Kamerasystem vorgenommen, damit die 3D-Rekonstruktion auch in verschiedenen Entfernungen zum Kamerasystem zuverlässig ist (siehe Kap. 3). Untersuchungen haben gezeigt,

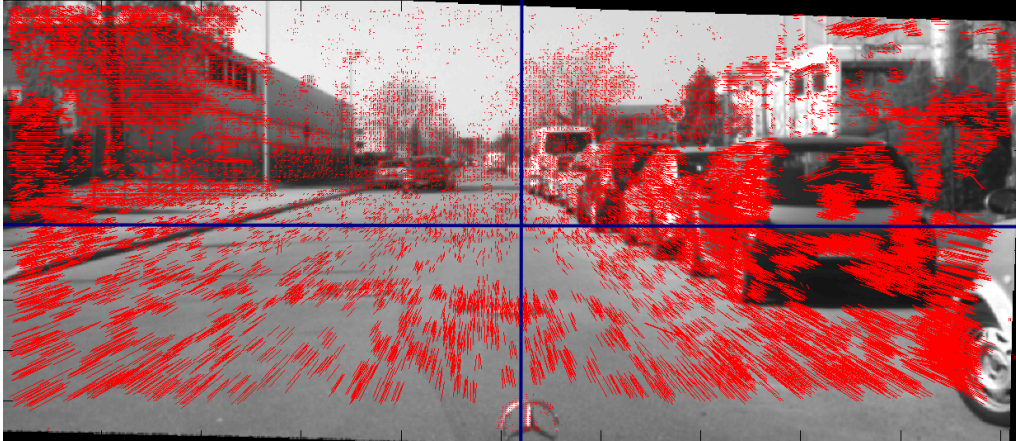


Abbildung 4.7.: Die Korrespondenzen können mit der Fluss-Methode in zeitlich aufeinanderfolgenden Bildern bestimmt werden. In jedem Bildsektor mit maximaler Größe werden dann die Korrespondenzen gesammelt. Ist diese maximale Anzahl von Korrespondenzen in den Bins erreicht, werden die Korrespondenzen für die Kalibrierung verwendet.

das Flussvektoren verschiedener Länge sich gut für eine Kalibrierung in verschiedenen Abständen eignen. Der Flussvektor verbindet zwei korrespondierende Punkte in zeitlicher Anordnung. Wenn nun ein Fahrzeug mit einem kleinen Gierwinkel gelenkt wird, ist die Länge der Flussvektoren ein Indikator für die Entfernung der Korrespondenz in der 3D-Rekonstruktion.

Bei kurzen Vektoren ist das Signal-zu-Rausch-Verhalten schlechter, bei langen Vektoren ist dieses Verhältnis besser. Deshalb werden viele lange Flussvektoren gesammelt und sehr kurze Flussvektoren nicht berücksichtigt. Daraus entstand eine Tiefenverteilung wie in Abb. 4.8 schematisch dargestellt.

4.3.3. Einfluss der Korrespondenzverteilung im Bild auf eine Kalibrierung

In der Literaturrecherche wurde keine Information zum Einfluss der Position der Verteilung der Korrespondenzen zur Kalibrierung gefunden. Deshalb wurde untersucht, welchen Einfluss die Verteilung der Beträge der Flussvektoren von Korrespondenzen über das Bild auf die Kalibrierung hat. Dabei wurden die Korrespondenzen ausgewählt, die der vorgegebenen Verteilung entsprachen. Dazu wurden die Verteilungen wie

- Gleichverteilung,
- Student-T-Verteilung,
- Normalverteilung,
- Inverse Normalverteilung,

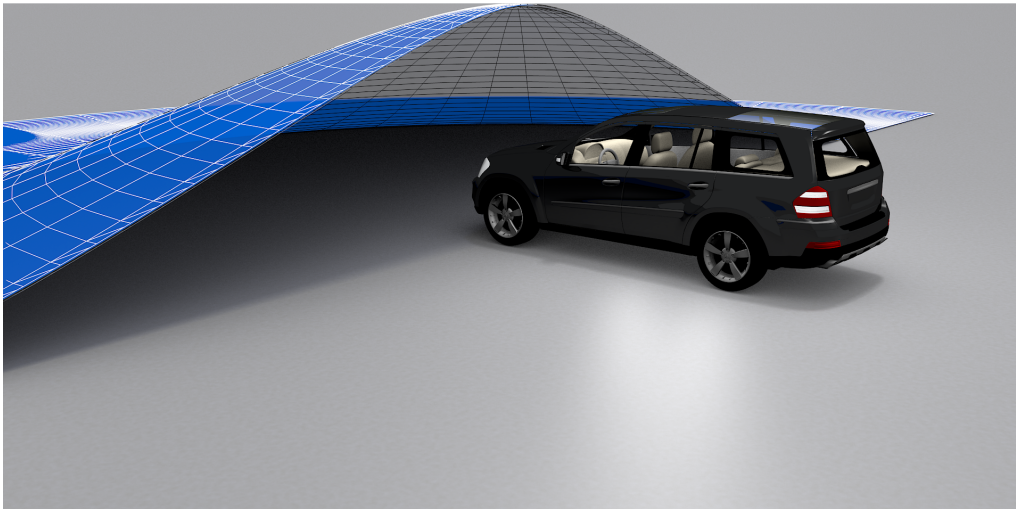


Abbildung 4.8.: Um das Kamerasystem in verschiedenen Entfernungen zu kalibrieren, sodass die Methoden zur 3D-Rekonstruktion zuverlässige Werte auch in weiten Distanzen liefern, müssen Korrespondenzen in verschiedenen Entfernungen betrachtet werden. Die Tiefeninformation aus den Flussvektoren wird durch deren Länge bestimmt. Kurze Vektoren repräsentieren Korrespondenzen in der Ferne und lange Vektoren repräsentieren nahe Korrespondenzen. Da die langen Vektoren ein besseres Signal-zu-Rausch-Verhalten haben, werden von diesen Vektoren mehr zur Auswertung gesammelt. Wird eine Gauss-Verteilung bezüglich der Vektorenlänge eingehalten, dann ist der Fehler durch das Rauschen der Daten gering gehalten. Dadurch ergibt sich eine Verteilung, die hier schematisch dargestellt ist. Viele Flussvektoren vor der Kamera und wenige Vektoren in einer großer Distanz halten den Fehler durch Messrauschen gering.

- Poisson-Verteilung und
- χ^2 -Verteilung

untersucht. Die Position einer Korrespondenz im Bild ist definiert durch die Koordinaten des Bildpunktes im linken Bild.

Korrespondenzen, die in einem Bild gefunden werden, sind mehr zufällig und gleichen keiner Verteilung. Dadurch muss über viele Bildpaare gesammelt werden. Doch selbst dann war es in den wissenschaftlichen, experimentellen Auswertungen dieser Methode bei der Mehrheit der untersuchten Verteilungen nicht möglich, eine Kalibrierung während der gesamten Sequenz zu erhalten. Somit gab es bei der Poisson-Verteilung und bei der χ^2 -Verteilung keine Kalibrierung, da das Kalibrieren mit den gesammelten Korrespondenzen innerhalb eines vorgegebenen Zeitfensters nicht möglich war.

4.3.4. Scheibenverzeichnung

In den Assistenzfunktionen, die mit einem Monokamerasystem realisiert wurden, wie die Verkehrszeichenerkennung, Fahrbahnoberflächenassistent oder Parkassistent, ist die

se Verzeichnung in die Berechnung nicht eingegangen, sodass diese Verzeichnung das erste Mal in der Nutzung des Stereo-Kamera-Systems in Erscheinung trat. In dem Bereich des Bildes, der von der Verkehrszeichenerkennung verarbeitet wird, ist die Verzeichnung durch die Windschutzscheibe sehr gering. Deshalb ist sie auch beim Stereo-Kamera-System zu vernachlässigen. Die Assistenzfunktionen, die im Nahbereich des Fahrzeugs Merkmale auf der Straße detektieren, haben ebenfalls diese Verzeichnung. Sie wird vernachlässigt, da bei einem Mono-Kamera-System bei diesen Anwendungen keine metrische Messung vorgenommen wird. Anders ist dies bei einem Stereo-Kamera-System, da dort mit dem Kamerasystem metrisch gemessen werden muss.

Durch das Auffinden von Korrespondenzen entlang einer Epipolarlinie müssen die Zeilen des linken und rechten Bildes übereinstimmen. Durch die Verzeichnung, verursacht durch die Windschutzscheibe, werden die Seiten des Bildes aber wie ein Gummiband verzogen und so weichen die Ränder von einem Rechteck ab. Dadurch sind die Zeilen des linken und rechten Bildes nicht durchgehend deckungsgleich über alle Bildzeilen angeordnet und so können die Korrespondenzen nicht exakt bestimmt werden, wodurch falsche Messungen erfolgen.

Wenn aber die Kameras eines Stereo-Kamera-Systems, die bereits im Fahrzeug montiert sind, kalibriert werden, dann ist ein Fehler in der 3D-Rekonstruktion nicht zu messen. Wird aber das Stereo-Kamera-System außerhalb des Fahrzeugs initial kalibriert und dann erst in das Fahrzeug eingebaut, erzeugt die Krümmung der Windschutzscheibe diese Verzeichnung. Diese durch die Windschutzscheibe verursachte Verzeichnung muss aus den Bildern herausgerechnet werden, sonst können die Daten aus den so aufgenommenen Bildern nicht mehr verwendet werden.

Da in der Literatur keine Lösung gefunden wurde, die diese Verzeichnung beseitigt, wird diese Herausforderung in diesem Abschnitt angegangen. In diesem Abschnitt wurden zwei Methoden entwickelt, die diese Aufgabe lösen.

Wissenschaftliche Untersuchungen haben gezeigt, dass diese Verzeichnung tatsächlich durch die Scheibenkrümmung erzeugt wird. Die hier entwickelten Methoden messen die durch die Scheibenkrümmung verursachten Verzeichnungen und können diese somit beseitigen.

Um die Verzeichnung zu messen, wurde das gleiche Target aus Schachbrettlinien im linken und rechten Bild aufgenommen. Die Koordinaten der Schnittpunkte aus den horizontalen und vertikalen Geraden wurden dann verglichen. Daraus ergab sich eine Abweichung, die als Windschutzscheibenverzeichnung beschrieben wurde.

Die erste entwickelte Methode bestimmt die Koeffizienten eines Polynoms und approximiert so die durch die Scheibenkrümmung verursachte Verzeichnung. Dieses Polynom wird in Gl. (4.30) beschrieben. Durch diese Approximation lässt sich dann die Verzeich-

nung auf dem Bild berechnen und damit die Rektifizierung wieder erfolgreich realisieren. Die zweite Methode ist eine theoretische Untersuchung mit differentialgeometrischen Methoden, mit der versucht wird, die Verzeichnung kontinuierlich und sehr genau zu berechnen. Da diese Methode nicht realisiert wurde, wird sie in diesem Abschnitt kurz beschrieben und in Kap. A.4 detaillierter erläutert.

4.3.4.1. Polynomialer Ansatz

Die Verzeichnung durch die Windschutzscheibe verzerrt das Bild, sodass die korrekte Bestimmung der Korrespondenzen schlecht oder gar nicht möglich ist. Es kann ein Polynom bestimmt werden, das die Verzeichnung durch die Windschutzscheibe approximiert. Dieses Polynom muss einmal nach dem Einbau des Kamerasystems ins Fahrzeug bestimmt werden und kann dann als fester Faktor vor die Rektifizierung in die Online-Kalibrierung integriert werden. Durch die Berechnung der Koeffizienten eines Polynoms lässt sich diese Verzerrung der Linien im Bild bestimmen, zurückrechnen und damit beseitigen. D.h. die Bildpunkte (u, v) können anhand des Polynoms an die eigentliche, erwartete Position im Bild verschoben werden. Nach der Verschiebung der Bildpunkte kann dann die Rektifizierung wiederhergestellt werden, sodass die Korrespondenzen wieder auf der Epipolarlinie gefunden werden.

Wie in Abb. 4.9 zu sehen, hat die durch die Windschutzscheibe verursachte Verzeichnung den größeren Fehler in der v -Koordinate entlang der Bilder, während entlang der u -Koordinate eine geringe Verzeichnung zu messen ist. Die Bildbreite ist die horizontale Koordinate des Diagramms und die Bildhöhe ist die vertikale Koordinate des Diagramms. Das Diagramm zeigt, wie an jeder Koordinate des Bildes die Abweichung durch die Scheibenverzeichnung ist. Die Abweichung ist von klein (blau) bis groß (rot) farblich gekennzeichnet.

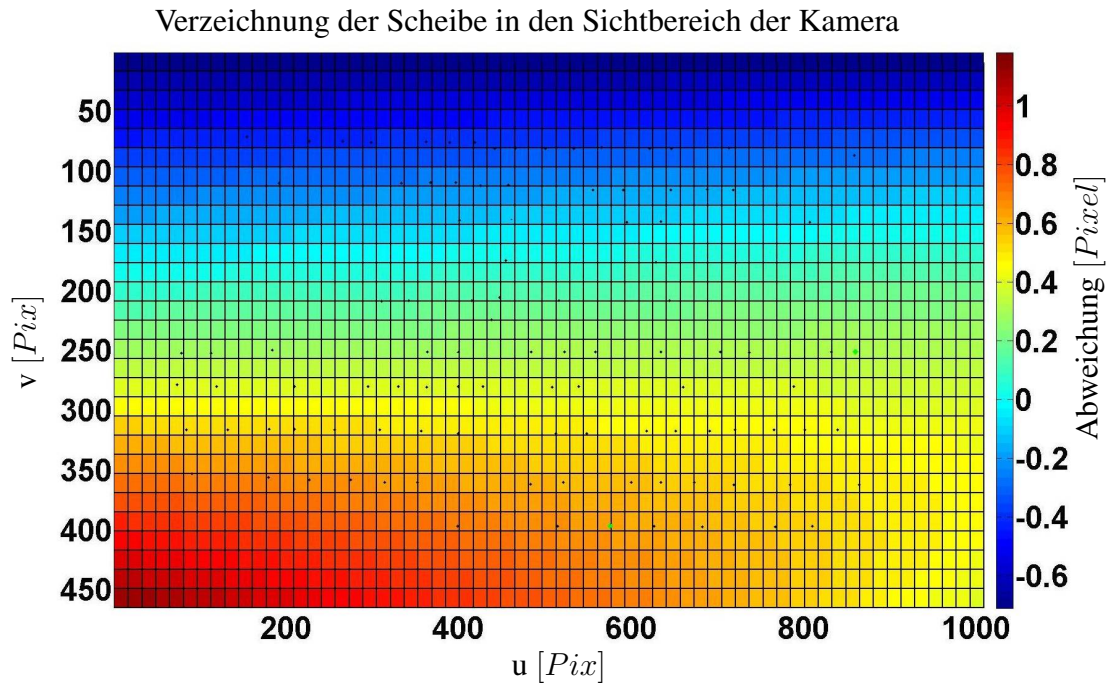


Abbildung 4.9.: Wenn die Stereo-Kamera außerhalb des Fahrzeugs kalibriert und dann in das Fahrzeug eingebaut wird, gibt es eine Verzeichnung, d.h. eine Verschiebung der Bildzeilen und Bildspalten. Diese Verzeichnung ist auf dem rektifizierten Bild nachweisbar und behindert das Finden von korrespondierenden Punkten im linken und rechten Bild. Der Faktor, um den sich der Bildpunkt verschiebt, ist hier in einer Farbskala kodiert.

Experimentelle Untersuchungen haben ergeben, dass sich die Verzeichnung durch das Polynom

$$(au + b)v^2 + (cu + d)v + e = \Delta v \quad (4.30)$$

approximieren lässt. Um die Koeffizienten (a, b, c, d, e) in Gl. (4.30) zu bestimmen, wird die RANSAC-Methode (siehe Kap. 2.2.2) genutzt. Jede Abweichung kann als ein Drei-Tupel mit den Komponenten $(u, v, \Delta v)$ dargestellt werden. Da diese Abweichungen sehr gering sein können, wurde für die wissenschaftliche Auswertung eine Vielzahl von Bildpaaren genutzt.

Als Hypothese wurde der kleinste Abstand zwischen der rechteckigen, erwarteten Bildebene und den Messpunkten genommen. Somit ist die Fitnessfunktion

$$r_i = \min_{a,b,c,d} \|\Delta v(u_i, v_i) - ((a \cdot u_i + b)v_i^2 + (c \cdot u_i + d)v_i + e)\|. \quad (4.31)$$

Dabei ist $\Delta v(u_i, v_i) - ((a \cdot u_i + b)v_i^2 + (c \cdot u_i + d)v_i + e)$ die Differenz zwischen der gemessenen Abweichung $\Delta v(u_i, v_i)$ an den Koordinaten $(u, v)_i$ und der Ebene aus Abb. 4.10 an der gleichen Position. Um die Ebene zu berechnen, werden mindestens fünf Messpunkte aus der Menge aller Messpunkte gewählt. Diese fünf Messpunkte werden dann genutzt, um die fünf Koeffizienten a, b, c, d, e in Gl. (4.30) zu schätzen. Dies wird solange wieder-

holt, bis das Optimum gefunden ist. Mit Gl. (2.1) kann das Finden des Optimums zeitlich optimiert werden. In Abb. 4.10 wird eine solche Fläche in der Punktmenge dargestellt.

Ausreißer Grafik

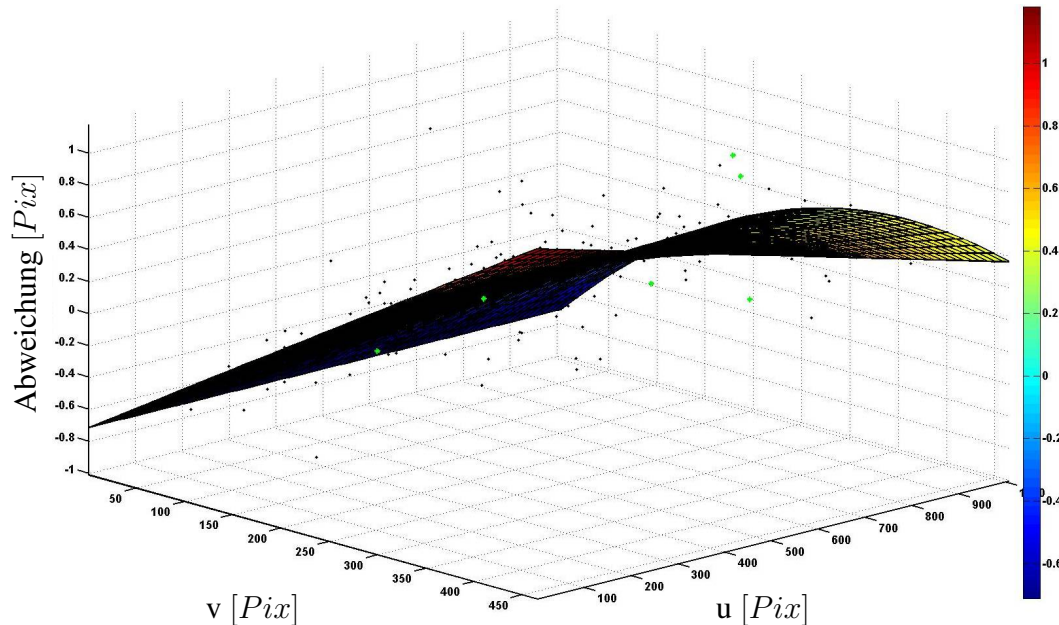


Abbildung 4.10.: Nachdem die Koeffizienten (a, b, c, d, e) durch die RANSAC-Methode bestimmt wurden, kann die Fläche, die durch das Polynom beschrieben wird, in die Menge der Messpunkte eingefügt werden.

Da sich die Windschutzscheibenkrümmung nicht ändert, wird diese Methode nach der Initialkalibrierung einmalig durchgeführt.

Um den Einfluss von Abweichungen durch die Windschutzscheibe in den Untersuchungen zu minimieren, wurde das Kamerasystem vor jeder Fahrt kalibriert, nachdem die Kamera in das Fahrzeug eingebaut wurde.

4.3.4.2. Differentialgeometrischer Ansatz

Die Verzeichnung durch die Windschutzscheibe macht es schwer, Korrespondenzen im Bild zu finden. Die Schwierigkeit folgt aus der Verschiebung der Zeilen, sodass sich die Korrespondenzen nicht in der gleichen Zeile befinden. Um die Rektifizierung im verzeichneten Bild wieder herzustellen, wurde im vorherigen Abschnitt ein Polynom genutzt. Die zuvor beschriebene Methode funktioniert sehr gut, was die wissenschaftlichen Untersuchungen verdeutlichen. Dennoch wurde eine weitere Methode erdacht, um die Verzeichnung durch die Windschutzscheibe in einem früheren Stadium der Entwicklung zu bestimmen und so das Bild immer richtig zu rektifizieren und auch bei einer leicht verschobenen Position des Stereo-Kamera-Systems hinter der Windschutzscheibe die Verzeichnung zu beseitigen.

Mit dieser Methode kann die Verzeichnung der Windschutzscheibe schon bei der Konstruktion der Scheibe bestimmt werden. Zusätzlich kann die Verzeichnung an jeder Stelle der Windschutzscheibe kontinuierlich berechnet werden. In dieser Methode wird angenommen, dass die Scheibe ein Segment eines Paraboloids ist. Dann gibt es einen kleinen Bereich auf diesem Paraboloid, durch den die Kameras des Stereo-Kamera-Systems hindurchsehen. Da die Krümmung und der Punkt der optischen Achse bekannt sind, kann der Fehler zwischen dem erwarteten und dem verzeichneten Bild berechnet werden. Dazu wird eine parallele Ebene zu der Bildebene der Kamera im Schnittpunkt zwischen optischer Achse und dem Paraboloiden gelegt. Die Distanz zwischen der Oberfläche des Paraboloiden und der zur Kamerabildebene verlaufenden Ebene ist dann die Verzeichnung, die durch die Windschutzscheibe gegeben wird. Durch diese Methode kann eine schnelle Berechnung durch neue Scheibenmodelle und durch andere Positionierungen bereitgestellt werden.

Die theoretischen Überlegungen sind in Kap. A.4 genauer beschrieben.

5. Experimentelle Untersuchungen

In den vorangegangenen Kapiteln wurde erklärt, welche Methoden im Rahmen dieser Arbeit entwickelt wurden, um eine Online-Kalibrierung einer Kamera, die in einem Fahrzeug hinter der Frontscheibe angebracht ist, zu ermöglichen.

In Kap. 1 wird beschrieben, warum solche Methoden zur Online-Kalibrierung essenziell sind, damit die Fahrerassistenzsysteme, die den Fahrer während der Fahrt unterstützen sollen, zuverlässig funktionieren. Bei optischen Messmitteln, wie dem eingesetzten Stereo-Kamera-System, verursachen selbst kleine Abweichungen von der Initialkalibrierung große Abweichungen in der Berechnung der Werte für die Fahrerassistenzsysteme. In Kap. 1.2 wird dies beschrieben.

Das für die Untersuchungen genutzte Stereo-Kamera-System ist in Abb. 5.1 dargestellt. Die Kameras haben beide eine Kamerakonstante von ca. $1024 [Pix]$. Die Basisbreite des Stereo-Kamera-Systems ist ca. $0.20 [m]$. Das Stereo-Kamera-System hat dann eine Stereo-Kamera-Konstante von $1024 [Pix] \cdot 0.2 [m] \approx 204.8 [Pix \cdot m]$. Bei dem Stereo-Kamera-System sind die optischen Achsen beider Kameras, die z-Achse des Kamerakoordinatensystems, parallel und die y-Achse zeigt nach unten. Diese beiden Achsen sind in einer Kamera in Abb. 5.1 rot eingezeichnet.

Zu Beginn der Untersuchungen in diesem Kapitel wird gezeigt, welchen Einfluss die einzelnen Komponenten einer Korrespondenz auf die Online-Kalibrierung haben. Dazu werden, wie in Kap. 4.3.1 gezeigt, die Korrespondenzen nach den Einflüssen auf den Winkel sortiert und ausgewertet.

Anschließend wird untersucht, welchen Einfluss die Position einer Korrespondenz im Bild auf die Online-Kalibrierung hat. Dazu werden die Korrespondenzen untersucht, ob diese zu der vorgegebenen Verteilung passen. Wenn die Korrespondenzen dem Kriterium genügen, dann wird die Korrespondenz beibehalten, ansonsten wird die Korrespondenz verworfen. Beschrieben ist das in Kap. 4.3.3.

Das Binning, wie in Kap. 4.3.2 beschrieben, hat sich als Verteilungsfunktion als sehr robust herausgestellt und half, die Online-Kalibrierung mit einer einfachen Methode signifikant zu verbessern. Um dies zu zeigen, wird das Binning im nächsten Abschnitt mit den robusten Fehlerfunktionen aus [Hartley und Zisserman, 2004, S. 636] verglichen.

Da sich die *innere Genauigkeit*, wie in Kap. 2.2.4 beschrieben, beim Bündelausgleich als sehr robust herausgestellt hat, wird diese im Rahmen dieser Arbeit als Nächstes untersucht.

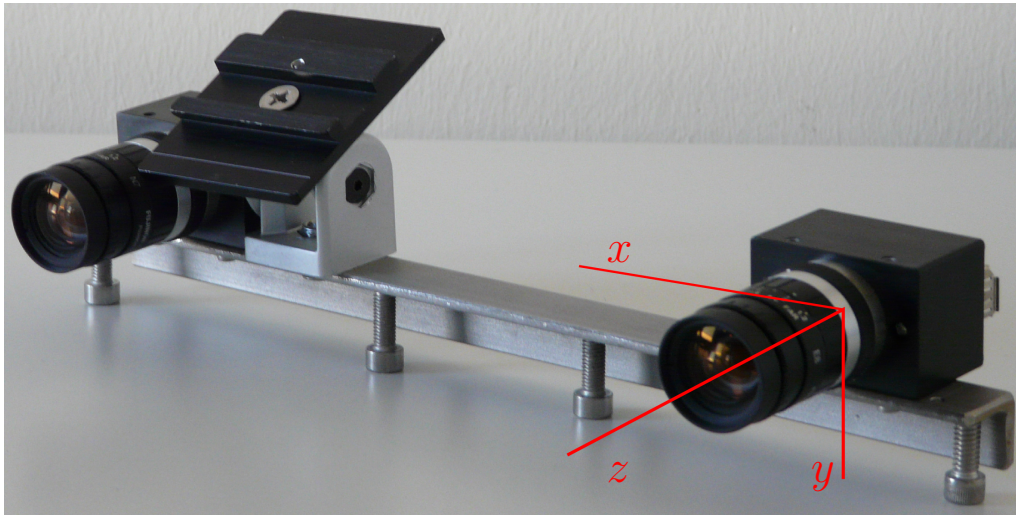


Abbildung 5.1.: Abgebildet ist der Aufbau des Versuchs-Stereo-Kamera-Systems, das für eine Vielzahl der Aufnahmen genutzt wurde. Die Kameras haben eine Kamerakonstante von $1024 [Pix]$ und einer Basisbreite von $0.20 [m]$. Dies entspricht einer Stereo-Kamera-Konstante von $1024 [Pix] \cdot 0.2 [m] = 204.8 [Pix \cdot m]$. Die Kameras sind so angeordnet, dass die optische Achse, die z -Achse des Kamerakoordinatensystemes, parallel zueinander ausgerichtet sind und die y -Achse nach unten zeigt. Die beiden Achsen sind im Bild rot eingezeichnet.

Im Anschluss der Untersuchung der *inneren Genauigkeit* wird untersucht, wie genau die ϕ_1, ϕ_2, ϕ_3 durch eine Online-Kalibrierung bestimmt werden müssen, damit eine Abweichung von nur wenigen Zentimetern in einer Entfernung von $50 [m]$ in der 3D-Rekonstruktion gegeben werden kann.

Abgeschlossen werden die Untersuchungen der Methoden aus dieser Arbeit mit den beiden Methoden, die zur Online-Kalibrierung genutzt werden können. Die zuerst beschriebene Methode nutzt die Informationen über die Entfernung zweier zeitlich aufeinanderfolgender Bilder von einem Realweltobjekt. Beschrieben ist diese Methode in Kap. 4.1. Die Methode zur Online-Kalibrierung, die dann beschrieben wird, bestimmt den Fluchtpunkt in den Bildern aus dem linken und rechten Bild. Beschrieben ist diese Methode in Kap. 4.2

In Tab. 5.1 ist ein kurzer Überblick gegeben, welche Untersuchung vorgenommen wird und es wird die Methode beschrieben, in welchem Kapitel die Untersuchung zu dieser Methode erfolgt.

Als Fehlermaß, um die Methoden zu bewerten, wird der Unterschied zwischen den Werten aus der Online-Kalibrierung und den Werten aus der Offline-Kalibrierung (siehe Kap. 3) genutzt. In einigen Fällen werden die Werte aus der Ground Truth auch als Initialkalibrierung im Sinne von Kap. 3 genutzt. Das Fehlermaß ist durch

Tabelle 5.1.: Im Rahmen dieser Arbeit wurden unterschiedliche Methoden und Funktionen zur Online-Kalibrierung entwickelt und untersucht. Zuerst wird untersucht, welchen Einfluss die vier Komponenten einer Korrespondenz auf die Online-Kalibrierung haben. Anschließend wird untersucht, wie sich die Kalibrierung durch die Position von Korrespondenzen im Bild verbessern lässt. Aus der Untersuchung der Positionierung der Korrespondenzen hat sich das Binning als besonders robust herausgestellt. Deshalb wird untersucht, wie sich das Binning auf bestehenden Kalibrierungen auswirkt. Anschließend wird die *innere Genauigkeit* betrachtet, die im Rahmen dieser Arbeit vom Bündelausgleich für die Essential-Gleichung hergeleitet wurde. Bevor die Methoden zur Online-Kalibrierung in einem Fahrzeug beschrieben werden, werden die Auswirkungen durch eine Abweichung auf die Disparität beschrieben.

	Methode/Daten	Beschrieben in Kap.	Evaluierung in Kap.
1.	Komponenten der Korrespondenzen	Kap. 4.3.1	Kap. 5.1
2.	Verteilung der Korrespondenzen	Kap. 4.3.3	Kap. 5.2
3.	Korrespondenzauswahl: Binning	Kap. 4.3.2	Kap. 5.3
4.	Korrespondenzauswahl: innere Genauigkeit	Kap. 2.2.4	Kap. 5.4
5.	Einfluss der Kalibrierung auf die Rekonstruktion	Kap. 1.2	Kap. 5.5
6.	Schielwinkelkalibrierung: Vergleich der zurückgelegten Entfernung	Kap. 4.1	Kap. 5.6.1
7.	Schielwinkelkalibrierung: Fluchtpunktmethode	Kap. 4.2	Kap. 5.6.2

$$\phi_j = \phi_{j_i} - \phi_{j_o} \text{ wobei } j = 1, 2, 3 \quad (5.1)$$

wobei sich i in Gl. (5.1) auf den Index der Ground Truth und o sich auf Online-Kalibrierung, bezieht, gegeben. Das Fehlermaß zeigt, wie und in welchem Ausmaß eine Dekalibrierung im Einsatz erfolgt ist und wie diese durch die Online-Kalibrierung zu korrigieren ist. Ist der Wert sehr groß, so werden Distanzen zu Objekten sehr ungenau geschätzt. Es werden daher hohe Anforderungen an die zur Korrektur eingesetzte Online-Kalibrierung gestellt. Die Messungen der ϕ_j unter realistischen Bedingungen zeigt an, ob eine Online-Kalibrierung überhaupt erforderlich ist.

Insgesamt wurden mit den verwendeten Stereo-Kamera-Systemen neun Stunden Video aufgezeichnet. Mit dem Stereo-Kamera-System aus Abb. 5.1 wurden ca. fünf Stunden für die Evaluierung der Methoden aufgezeichnet. Dies ergibt rund 300.000 Bildpaare. Die verbleibenden vier Stunden wurden mit einem Stereo-Kamera-System aufgezeichnet, das mit dem Stereo-Kamera-System aus Abb. 1.2 vergleichbar ist. Die Bildpaare wurden bei unterschiedlichen Wetterverhältnissen und zu unterschiedlichen Tageszeiten aufgezeichnet. Eine genauere Auflistung der Sequenzen ist in Tab. 5.2 gegeben.

Tabelle 5.2.: Verschiedene Sequenzen wurden zur Evaluierung der Methoden mit dem Stereo-Kamera-System aus Abb. 5.1 aufgezeichnet. Dabei wurden insgesamt 309.920 Bildpaare aufgezeichnet. Dies entspricht einer Gesamtzeit von ca. 5.4 [h].

	Bezeichnung	Wetter	# Bilder	Tageszeit
1.	Ringweg	verschneit	4.308	Vormittag
2.	Industriegebiet Böblingen	sonnig	14.416	Nachmittag
3.	Albrundfahrt	neblig	82.160	Vormittag
4.	Rundfahrt	sonnig	75.254	Vormittag
5.	Geislinger Steige	regnerisch	47.604	Nachmittag
6.	Kreisverkehr	bewölkt	9.602	Nachmittag
7.	Städte	bewölkt	76.576	Mittags

Die Methoden, die in [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] beschrieben sind, wurden im Rahmen dieser Arbeit implementiert, damit die Aussage durch die Analyse mit den Daten aus den Testsequenzen, die in Tab. 5.2 aufgezählt sind, vergleichbar sind mit den Auswertungen der Methoden, die im Rahmen dieser Arbeit entwickelt wurden.

5.1. Einfluss der Komponenten einer Korrespondenz auf die Kalibrierung

Die Essential-Gleichung Gl. (1.2) nutzt die normierten Korrespondenzen, um die Rotations- und Translationsparameter, die die Parameter der Essentialmatrix \mathbb{E} darstellen, zu schätzen. Vorhergehende wissenschaftliche Untersuchungen haben gezeigt, dass sich die Translationsparameter des Stereo-Kamera-Systems, wie es in Abb. 5.1 dargestellt ist, nur gering ändern, wohingegen die Rotationsparameter sich stark unter Temperatureinfluss verändern (siehe Kap. 1.1).

Werden nun die Rotationsparameter aus den Korrespondenzen berechnet, dann hat jede Komponente dieser Korrespondenz einen Einfluss auf diese Berechnung. Die Approximation Gl. (4.27) gibt an, welche Komponente einer Korrespondenz den Nickwinkel ϕ_1 beeinflusst, die Approximation Gl. (4.28) gibt an, welche Komponente den Schielwinkel ϕ_2 bei der Berechnung beeinflusst und die Approximation Gl. (4.29) zeigt den Zusammenhang der Komponenten der Korrespondenz und die Berechnung des Rollwinkels ϕ_3 . Zur Auswertung der Methode wurde die 1. Sequenz „Ringweg“ aus Tab. 5.2 genutzt. Damit die Menge der Korrespondenzen immer gleich groß ist, wird das Bild in vier gleich große Teile aufgeteilt, wie in Abb. 4.7 gezeigt und es werden pro Sektor 400 Korrespondenzen gesammelt.

Die Menge der Korrespondenzen wird so sortiert, dass die Änderungen Gl. (4.27) möglichst klein sind. D.h. die Korrespondenzen, die sich in der Menge ganz hinten befinden, haben nur wenig bis gar keinen Einfluss auf die Online-Kalibrierung des Nickwinkels. Identisch hierzu wird die Menge der Korrespondenzen nach der Approximation Gl. (4.28) sortiert und die Menge der Korrespondenzen wird anschließend nach der Approximation Gl. (4.29) sortiert. Anschließend wird die erste Menge genutzt, um den Nickwinkel ϕ_1 zu berechnen, mit der zweiten sortierten Menge wird der Schielwinkel ϕ_2 berechnet und der Rollwinkel ϕ_3 wird mit der dritten Menge berechnet. Dadurch sind die Korrespondenzen, die sich für die Kalibrierung am besten eignen, die ersten Elemente in den Mengen.

Wenn nun die Mengen sortiert sind, dann haben die Elemente am Ende jeder Menge keinen oder nur einen geringen Einfluss auf die Online-Kalibrierung und können vernachlässigt werden. Um dies zu zeigen, wird jede der Mengen, um $\frac{1}{4}$ ihrer Größe verkleinert, sodass jede der Mengen eine Größe von $\frac{3}{4}$ der Ausgangsgrößen an Korrespondenzen hat. Dieses sind die 75% der geeignetsten Korrespondenzen für die Kalibrierung.

Zum Vergleich der Methode werden die Kalibrierungen ohne jede Ausreißerdetektion und ohne Korrespondenzverarbeitung vorgenommen. Dann wurde auch noch die *innere Genauigkeit* genutzt, die in Kap. 2.2.4 beschrieben ist, um zu zeigen, ob Fehlerfunktionen sich durch diese Methode verbessern.

In der Abb. 5.2 wird gezeigt, wie sich der Nickwinkel ϕ_1 verhält. Dabei ist durch rote Linien der Wert ohne Einfluss auf die Kalibrierung angegeben, mit schwarzen Linien wird die Auswertung mit der *inneren Genauigkeit* verbunden, durch grüne Kreise werden die Werte markiert, die mit der sortierten und verkleinerten Menge errechnet wurden, und mit schwarzen Kreuzen werden die Werte gekennzeichnet, die zusammen mit der verringerten und sortierten Menge an Korrespondenzen und der *inneren Genauigkeit* errechnet wurden. In Abb. 5.2 ist auch zu sehen, dass keine Methode eine Verbesserung des Nickwinkels ϕ_1 gegenüber einer Berechnung ohne eine Änderung der Korrespondenzen hat. Sowohl Ausreißerdetektionen als auch die Sortierung der Menge verbessern das Ergebnis der Kalibrierung mit gegebenen Korrespondenzen.

Der Rollwinkel ϕ_3 und der Nickwinkel ϕ_1 sind bei der Auswertung sehr ähnlich. In Abb. 5.3 ist auch zu sehen, dass weder mit der *inneren Genauigkeit* noch mit der optimierten Menge an Korrespondenzen eine Verbesserung gegenüber einer Berechnung des Rollwinkels ϕ_3 zu sehen ist.

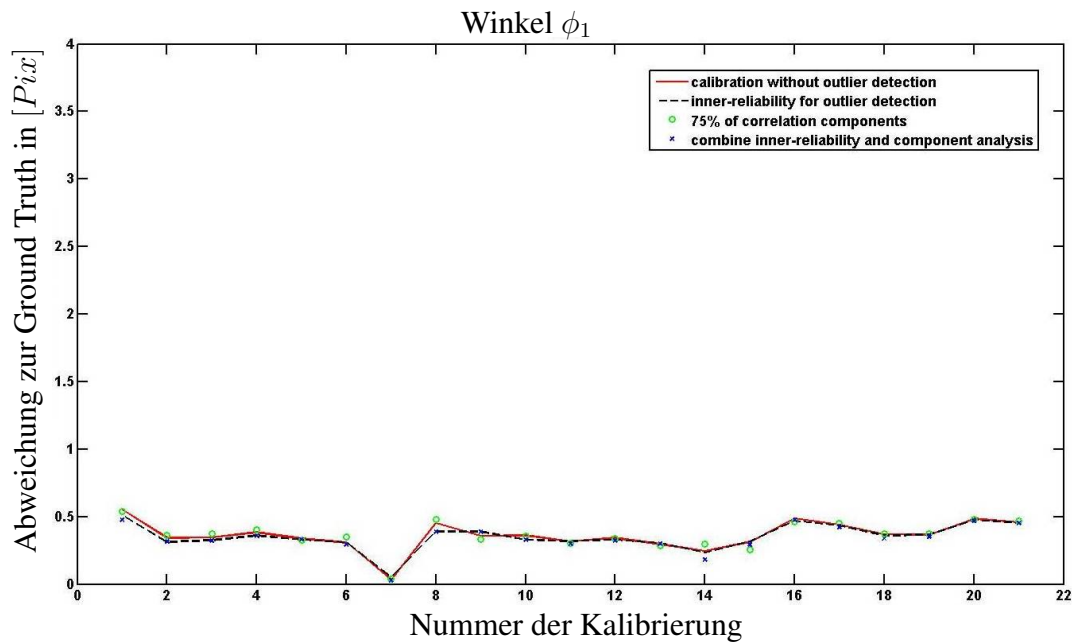


Abbildung 5.2.: Dargestellt ist der Vergleich der Werte aus der Online-Kalibrierung für den Nickwinkel ohne eine Ausreißerdetektion, was mit einer roten Linie eingezeichnet ist, mit der Ausreißerdetektion, hier die *innere Genauigkeit*, was durch eine schwarz gestrichelte Linie dargestellt ist, der optimierten Menge, die für den Versuch um $\frac{1}{4}$ verkleinert wurde, in der Darstellung mit grünen Kreisen eingezeichnet, und die Kombination der optimierten Menge zur Berechnung des Nickwinkels mit der *inneren Genauigkeit*, hier mit schwarzen Kreuzen eingezeichnet.

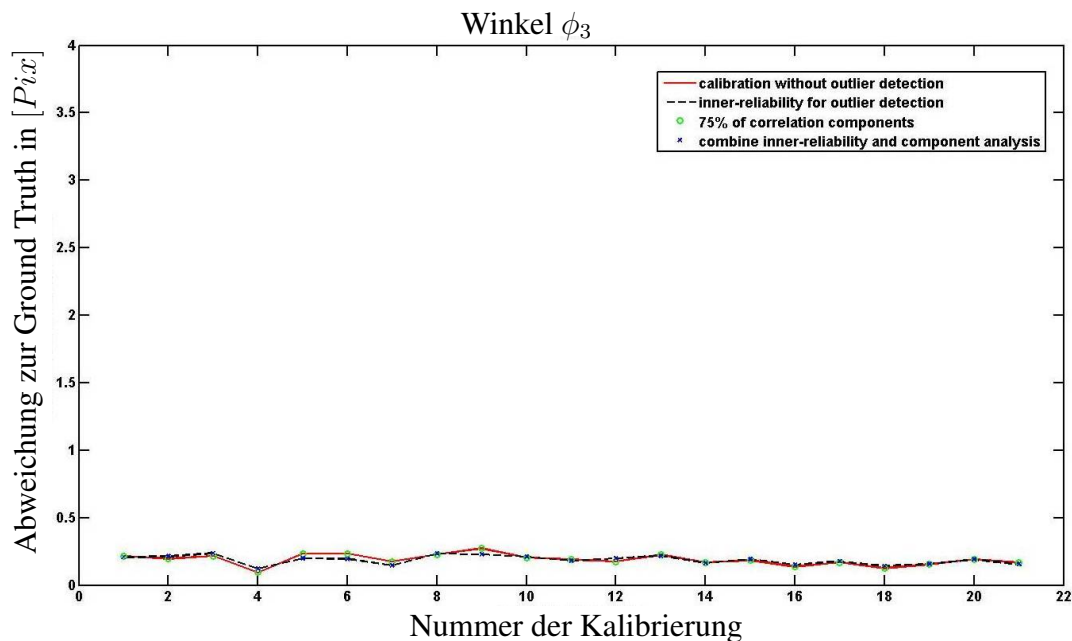


Abbildung 5.3.: Dargestellt ist der Vergleich der Werte aus der Online-Kalibrierung für den Rollwinkel ohne eine Ausreißerdetektion, was mit einer roten Linie eingezeichnet ist, mit der Ausreißerdetektion, hier die *innere Genauigkeit*, was durch eine schwarz gestrichelte Linie dargestellt ist, der optimierten Menge, die für den Versuch um $\frac{1}{4}$ verkleinert wurde, in der Darstellung mit grünen Kreisen eingezeichnet, und die Kombination der optimierten Menge zur Berechnung des Rollwinkels mit der *inneren Genauigkeit*, hier mit schwarzen Kreuzen eingezeichnet.

Anders als bei Nick- und Rollwinkel ϕ_1 und ϕ_3 gibt es beim Schielwinkel ϕ_2 signifikante Änderungen, wenn Methoden zur Ausreißerdetektion oder eine optimierte Korrespondenzmenge genutzt werden. Dies ist in Abb. 5.4 dargestellt. Es ist zu sehen, dass die verkleinerte Korrespondenzmenge zu wenige Korrespondenzen beinhaltet und sich dadurch die Schätzung des Schielwinkels der Online-Kalibrierung verschlechtert. Das Verwenden der *inneren Genauigkeit* (durch eine schwarze Linie dargestellt) bringt dagegen eine signifikante Verbesserung. Wenn nun die sortierte und um 25% verkleinerte Menge zusammen mit der *inneren Genauigkeit* genutzt wird, ist die Kombination beider Methoden gleich gut. Durch die Reduzierung der Komponenten, hier 400 Korrespondenzen, ist die Berechnung der Online-Kalibrierung beschleunigt worden, ohne eine Verschlechterung der Qualität der Online-Kalibrierung zu erzeugen. Die Berechnung der Online-Kalibrierung mit dieser reduzierten Mengen an Korrespondenzen war ca. 1.3-mal schneller als die Berechnung mit allen 1.600 Korrespondenzen.

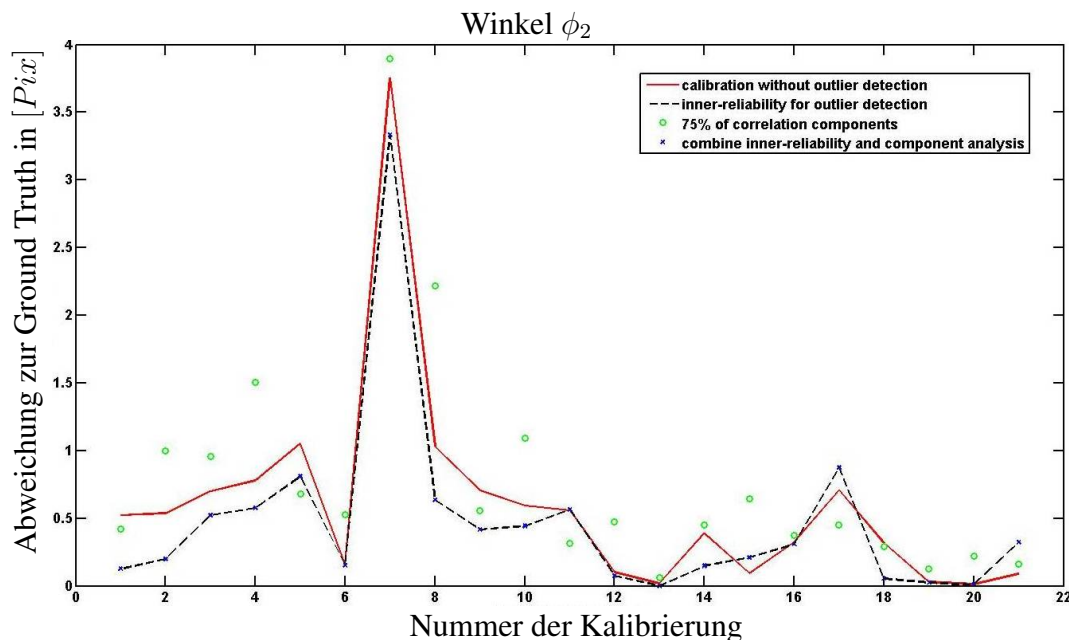


Abbildung 5.4.: Es ist zu sehen, dass die Online-Kalibrierung sich durch die *innere Genauigkeit* verbessern lässt. Mit der sortierten und verkleinerten Korrespondenzmenge ist die Geschwindigkeit der Online-Kalibrierung ca. 1.3-mal so schnell wie die Berechnung ohne eine sortierte und verkleinerte Korrespondenzmenge.

Die Online-Kalibrierung ist dann mit der robusten Fehlerfunktion so schnell wie die Berechnung der Online-Kalibrierung ohne eine Änderung der Menge der Korrespondenzmenge oder dem Nutzen der *inneren Genauigkeit* zur Ausreißerdetektion.

Zusammenfassung

Diese Untersuchung zeigt, dass sich durch eine Sortierung, nach den Approximationen Gl. (4.27) bis Gl. (4.29), die Menge der Korrespondenzen verringern lässt, ohne Verlus-

te bei der Qualität der Ergebnisse der Online-Kalibrierung. Beim Schielwinkel ist durch das starke Rauschen in den Daten die Hinzunahme einer robusten Fehlerfunktion, wie die *innere Genauigkeit*, essenziell, da sonst die Parameter bei einer kleineren Korrespondenzmenge schlechter geschätzt werden. Wichtigstes Ergebnis dieser Untersuchung ist die Verringerung der Verarbeitungszeit. Durch die Reduzierung der Anzahl der Korrespondenzen kann die Geschwindigkeit der Berechnung der Online-Kalibrierung mit der Ausreißerdetektionsmethode, wie der *innere Genauigkeit*, genutzt werden, ohne die Qualität der Online-Kalibrierung zu verschlechtern.

5.2. Verteilung der Korrespondenzen

In der Untersuchung Kap. 5.1 wurde genutzt, dass das Bild in vier Sektoren geteilt wird. In jedem Sektor wurden 400 Korrespondenzen gesammelt.

Um zu untersuchen, wie sich solche Verteilungen auf die Online-Kalibrierung auswirken, wurden verschiedene Verteilungen untersucht. In Kap. 4.3.3 ist dies beschrieben. Um die Verteilungen zu untersuchen, wurde die 1. Sequenz „Ringweg“ aus Tab. 5.2 genutzt. Der Versuchsaufbau wird in Tab. 5.3 dargestellt.

Tabelle 5.3.: Es werden verschiedene Verteilungen genutzt, um zu untersuchen, ob sich eine Online-Kalibrierung verbessert. Dazu wurden die Korrespondenzen im Bild untersucht, ob die Lage dieser Korrespondenzen einer vorgegebenen Verteilung entsprechen.

Bezeichnung der Verteilung	Parameter
χ^2 -Verteilung	Anzahl der Korrespondenzen
Poisson-Verteilung	Anzahl der Korrespondenzen
Gauss-Verteilung	Anzahl der Korrespondenzen pro Sektor
Inverse Gauss-Verteilung	Anzahl der Korrespondenzen pro Sektor
Gleichverteilung \equiv Binning	Anzahl der Sektoren

Um zu überprüfen, ob eine χ^2 -Verteilung oder eine Poisson-Verteilung der Korrespondenzen vorliegt, wurde der χ^2 -Test genutzt. Der χ^2 -Test nimmt eine Nullhypothese an, wie in [Klenke, 2006; Witten et al., 2011] beschrieben. Die hier betrachtete Nullhypothese ist, dass die Korrespondenzen im Bild in der Verteilung vorliegen. Sowohl für die χ^2 -Verteilung als auch für die Poisson-Verteilung konnte auf der ganzen Sequenz keine Online-Kalibrierung durchgeführt werden, da die Verteilung der Korrespondenzen nicht den geforderten Verteilungen entsprach.

Bei beiden Verteilungen wurden 2000 Korrespondenzen im Bild gesammelt. Bei allen neuen Korrespondenzen wurde nachgesehen, wie weit man von der geforderten Verteilung entfernt ist. Eine Heuristik hierfür ist durch den χ^2 -Test gegeben. Sollten sich die

im Bild befindenden Korrespondenzen nicht der χ^2 -Verteilung oder Poisson-Verteilung annähern, wird zufällig eine Anzahl von Korrespondenzen gelöscht. Die Anzahl der gelöschten Korrespondenzen entspricht $\frac{1}{5}$ der Anzahl der neuen Korrespondenzen. Bei einer Annäherung von 95% an die χ^2 -Verteilung oder Poisson-Verteilung wird eine Online-Kalibrierung durchgeführt.

Bei der Gauss-, der inversen Gauss- und bei der Normalverteilung wurde ein anderes Herangehen gewählt. Bei der Untersuchung dieser drei Verteilungen wurde untersucht, wie viele Kalibrierungen während der Laufzeit der Sequenz gemacht wurden. Zudem wurde untersucht, wie der Mittelwert und die Standardabweichung bei den gegebenen Verteilungen sind.

Gauss-Verteilung und inverse Gauss-Verteilung

Bei der Gauss-Verteilung wurden kreisförmige Sektoren, die von der Bildmitte ausgehen, eingerichtet. Jeder Sektor hat eine Schwelle, was die Anzahl zu findender Korrespondenzen betrifft. Die Schwellenwerte werden von der Mitte nach außen kleiner.

Um die Auswirkung dieser Verteilung zu untersuchen, werden im Bild gefundene Korrespondenzen den Sektoren so lange hinzugefügt, bis der Schwellenwert erreicht ist. Ist der Schwellenwert bei allen Sektoren erreicht, wird eine Online-Kalibrierung durchgeführt.



Abbildung 5.5.: Um die Gauss-Verteilung vorzugeben, wird das Bild in kreisförmige Abschnitte unterteilt. Diese Abschnitte oder Sektoren gehen ringförmig von der Mitte aus zum Rand. Jeder Sektor bekommt eine feste Schwelle, dabei ist die Schwelle der Korrespondenzen im innersten Kreis am größten und nimmt mit der Distanz zum Rande hin ab.

Es wurden verschiedene Einstellungen gewählt, um die Methode wissenschaftlich zu untersuchen. Die Auswertung ist in Tab. 5.4 abgebildet.

Tabelle 5.4.: **Gauss-Verteilung:** Es wird gezeigt, wie sich die Kalibrierung ändert, wenn sich die Verteilung der Korrespondenzen im Bild ändert. Es zeigt sich, dass bei der Zunahme von Korrespondenzen bei der Kalibrierung die Abweichungen der ϕ_{io} kleiner wird. In dieser Tabelle ist der Schwellenwert für den Bin angegeben, der die meisten Korrespondenzen beinhaltet. Dann die Anzahl der Sektoren, die für die Untersuchung genutzt werden. Abschließend sind die Daten aus der Auswertung mit der Anzahl der Kalibrierungen während der Sequenz angegeben.

Schwellen-Anzahl wert der Sektoren	Mittelwert für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Standardabweichung für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Anzahl der Kalibrie- rungen während der Sequenz	
400	2	(0.46, 1.0, 0.3)	(0.16, 0.88, 0.04)	16
500	2	(0.46, 1.0, 0.3)	(0.16, 0.87, 0.04)	16
600	2	(0.45, 0.98, 0.3)	(0.16, 0.7, 0.03)	15
400	3	(0.45, 0.99, 0.3)	(0.15, 0.7, 0.03)	16
500	3	(0.45, 0.97, 0.3)	(0.14, 0.7, 0.03)	16
600	3	(0.45, 0.96, 0.3)	(0.14, 0.69, 0.03)	15
400	4	(0.45, 0.96, 0.3)	(0.12, 0.7, 0.03)	15
500	4	(0.44, 0.96, 0.3)	(0.11, 0.69, 0.03)	15
600	4	(0.44, 0.96, 0.3)	(0.1, 0.69, 0.03)	14

Es ist zu erkennen, dass bei einer größeren Anzahl von Korrespondenzen der Mittelwert abnimmt und die Standardabweichung kleiner wird. Doch nimmt die Dauer zwischen den Kalibrierungen zu.

Bei der inversen Gauss-Verteilung werden, wie bei der Gauss-Verteilung, die Sektoren um den Mittelpunkt des Bildes angeordnet. Hier wird aber darauf geachtet, dass im innersten Kreis am wenigsten Korrespondenzen vorhanden sind und am Rand die meisten Korrespondenzen.

Wie auch zuvor wird die Auswertung in einer Tabelle angegeben. In Tab. 5.5 ist angegeben, wie der Schwellenwert für den größten Bin ist. Der größte Bin ist bei der inversen Gauss-Verteilung der Rand des Bildes. Dann ergeben sich der Mittelwert und die Standardabweichung für die drei Rotationswinkel (ϕ_1, ϕ_2, ϕ_3). Abschließend wird angegeben, wieviele Kalibrierungen während der Online-Kalibrierung durchgeführt wurden.

Der Vergleich zwischen der Gauss-Verteilung und der inversen Gauss-Verteilung zeigt, dass die inverse Gauss-Verteilung häufiger kalibriert. D.h. die Zeit zwischen den Kalibrierungen ist kleiner.

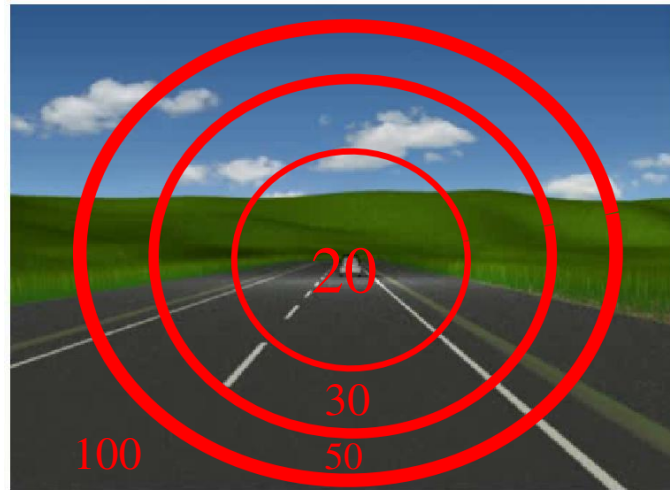


Abbildung 5.6.: Das Bild wird in Abschnitte geteilt. Dabei ist jeder Abschnitt kreisförmig. Jeder Abschnitt bekommt eine Schwelle für die maximale Anzahl von Korrespondenzen, die von außen nach innen zum innersten Kreis abnehmen.

Tabelle 5.5.: **Inverse Gauss-Verteilung:** In dieser Tabelle ist der Schwellenwert für den Bin angegeben, der die meisten Korrespondenzen beinhaltet. Dann die Anzahl der Sektoren, die im Bild eingerichtet sind. Abschließend sind die Daten aus der Auswertung mit der Anzahl der Kalibrierungen während der Sequenz angegeben.

Schwellen-Anzahl wert	der Sektoren	Mittelwert für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Standardabweichung für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Anzahl der Kalibrie- rungen während der Sequenz
400	2	(0.47, 0.97, 0.3)	(0.15, 0.88, 0.04)	18
500	2	(0.46, 0.95, 0.3)	(0.14, 0.87, 0.04)	18
600	2	(0.46, 0.94, 0.3)	(0.14, 0.7, 0.03)	17
400	3	(0.45, 0.96, 0.3)	(0.14, 0.7, 0.03)	18
500	3	(0.45, 0.95, 0.3)	(0.13, 0.7, 0.03)	17
600	3	(0.44, 0.95, 0.3)	(0.13, 0.69, 0.03)	16
400	4	(0.44, 0.96, 0.3)	(0.12, 0.7, 0.03)	16
500	4	(0.44, 0.94, 0.3)	(0.1, 0.68, 0.03)	16
600	4	(0.43, 0.94, 0.3)	(0.1, 0.68, 0.03)	16

Gleich-Verteilung

Um die Gleich-Verteilung zu untersuchen, wird ein Bild in n Sektoren geteilt. Jeder Sektor bekommt die gleiche Anzahl von Korrespondenzen. Diese Anzahl wird durch eine Schwelle fixiert. In Abb. 5.7 ist dies schematisch dargestellt.

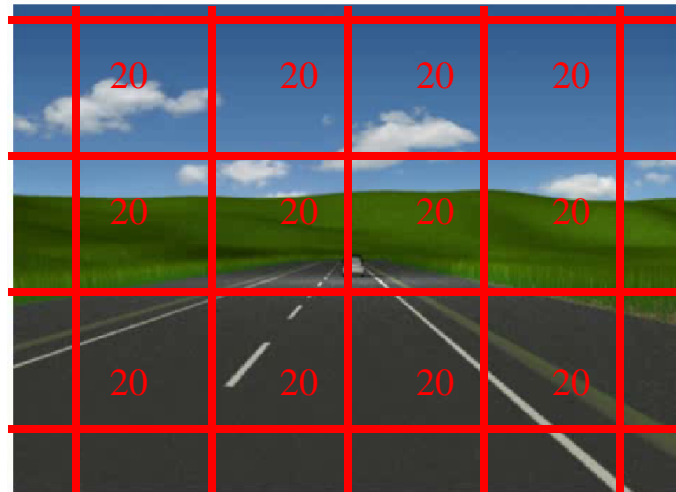


Abbildung 5.7.: Das Bild wird in Abschnitte geteilt. Dabei ist jeder Abschnitt rechteckig. Alle Abschnitte bekommen die gleiche Schwelle, für die feste Anzahl an Korrespondenzen.

In Tab. 5.6 wird dargestellt, wie sich die Auswertung der Online-Kalibrierung auf der gegebenen Sequenz verhält. Es ist aber zu sehen, dass während der ganzen Sequenz keine Kalibrierungen durchgeführt wurden, da die Sektoren zu kleinflächig werden, sodass nicht genug Korrespondenzen aus dem Bild genutzt werden können, um eine Kalibrierung zu starten.

Tabelle 5.6.: **Gleich-Verteilung:** Die 500 und 600 Korrespondenzen pro Sektor wurden nicht erreicht, daher ist keine Kalibrierung auf 1.8 [min] Fahrt durchgeführt worden.

Schwellen-Anzahl wert	Anzahl der Sektoren	Mittelwert für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Standardabweichung für (ϕ_1, ϕ_2, ϕ_3) [Pix]	Anzahl der Kalibrie- rungen während der Sequenz
400	4	(0.5, 0.97, 0.3)	(0.3, 0.82, 0.04)	21
500	4	(0.5, 0.96, 0.3)	(0.3, 0.77, 0.04)	19
600	4	(0.5, 0.96, 0.3)	(0.3, 0.73, 0.03)	18
400	16	(0.5, 0.8, 0.3)	(0.3, 0.46, 0.03)	2
500	16	-	-	-
600	16	-	-	-

Zusammenfassung

Es hat sich in dieser Untersuchung gezeigt, dass sich das Binning, wie in Kap. 4.3.2, als eine Approximation einer Gleich-Verteilung für eine robuste und genaue Online-Ka-

librierung auszeichnet, da zudem noch in einer kurzen Zeit eine geforderte Menge an Korrespondenzen bereit gestellt wird.

Andere Verteilungen, wie die χ^2 -Verteilung oder die Poisson-Verteilung, eignen sich nicht, da in der gesamten Sequenz nicht genügend Korrespondenzen zusammenkommen, um die Verteilung zu bilden.

Eine Verteilung der Korrespondenzen, wie z.B. die Gauss-Verteilung, ist geeignet für die Kalibrierung der Rotationsparameter, wenn wenige Ausreißer in der Stichprobe enthalten sind.

5.3. Korrespondenzauswahl mit Binning entspricht der Gleich-Verteilung

In den vorherigen Untersuchungen wurde gezeigt, dass sich die Online-Kalibrierung verbessern lässt, wenn Verteilungen der Korrespondenzen im Bild vorgenommen werden. In der Untersuchung zur Verteilung der Korrespondenzen im Bild ist zu sehen, wie sich das Binning als eine mögliche Realisierung der Gleich-Verteilung darstellt.

In dieser Untersuchung wird das Binning, das in Kap. 4.3.2 beschrieben ist, untersucht. Dazu wird die erste Sequenz aus Tab. 5.2 „Ringweg“ genutzt.

Bekannte und wissenschaftlich untersuchte robuste Fehlerfunktionen sind in [Hartley und Zisserman, 2004, S. 636] beschrieben. Es werden drei verschiedene Funktionen für die Untersuchung genutzt. Diese Funktionen sind in Tab. 5.7 dargestellt. Dabei wird die Lebesgue-Norm, die auch p -Norm genannt wird, für zwei unterschiedliche Konstanten p genutzt. Der Definitionsbereich der Parameter ist in Tab. 5.8 angegeben. Die Corrupted-Gaussian nutzt eine lineare Approximation in einer Gauss-Verteilung. Die Blake-Zisserman-Fehlerfunktion ist der Corrupted-Gaussian sehr ähnlich, hier wird aber ein fester Wert und keine Verteilung angenommen.

Tabelle 5.7.: Robuste Fehlerfunktionen werden eingesetzt, um Ausreißer in der Menge der Korrespondenzen zu finden. Durch die verschiedenen Funktionen werden die Ausreißer unterschiedlich bewertet.

	Robuste Fehlerfunktion	Gleichung
1.	Lebesgue-Norm	$\mathcal{L}^p(\delta r_i) = \int_{\mathbb{R}^n} (\delta r_i ^p dt)^{\frac{1}{p}}$
2.	Corrupted-Gaussian	$C(\delta r_i) = \sum_{i=1}^n -\ln(\alpha e^{-\delta r_i^2} + (1 - \alpha) e^{\frac{(-\delta r_i)^2}{\omega}})$
3.	Blake-Zisserman	$C(\delta r_i) = \sum_{i=1}^n -\ln(e^{-\delta r_i^2} + \xi)$.

Die erste Funktion in Tab. 5.7, die Lebesgue-Norm, wird genutzt, um Ausreißer auszusortieren. Dabei gibt der Wert p an, in wie weit der Ausreißer die Funktion beeinflusst. Die Ausreißer werden mit der p -Norm für $p = \infty$ sehr hoch bewertet und gehen in eine

Online-Kalibrierung mit dem höchsten Einfluss ein. Die Lebesgue-Norm für $p = \infty$ wird Infinity-Norm genannt.

Tabelle 5.8.: Die Variablen der Fehlerfunktionen haben unterschiedliche Definitionsbereiche. Diese werden in dieser Tabelle beschrieben. Die anderen Werte der betrachteten Fehlerfunktionen sind konstant.

Robuste Fehlerfunktion	Parameter	Definitionsbereich
Lebesgue-Norm	p	$p \in \mathbb{N}_1$
Corrupted-Gaussian	α	$\alpha \in]0, \dots, 1[$
Blake-Zisserman	ξ	$\xi \in \mathbb{N}$

Um den Einfluss des Binning auf die Online-Kalibrierung zu zeigen, werden die Untersuchungen zuerst mit den robusten Fehlerfunktionen allein durchgeführt. So lässt sich der Einfluss auf die Online-Kalibrierung durch das Binning verdeutlichen.

Lebesgue-Norm ($p = 1$)

Bei der Lebesgue-Norm, die in Tab. 5.7 beschrieben ist, wird $p \in \{1, 2\}$ untersucht.

Die Auswertung für die Lebesgue-Norm mit $p = 1$ ist in Abb. 5.8 dargestellt.

Die Skalen der drei Graphen in Abb. 5.8 und Abb. 5.9 zeigen, dass die Abweichung zwischen den Werten aus der Online-Kalibrierung und der Ground Truth unterschiedlich ist. So ist die Zahl für die Abweichung zwischen dem Nickwinkel ϕ_1 und dem Schielwinkel ϕ_2 um den Faktor zehn skaliert. D.h. die Abweichungen für den Schielwinkel sind zehnmal so groß wie die Abweichungen für den Nickwinkel. Ähnlich verhält es sich mit der Abweichung zwischen dem Roll- und Schielwinkel.

In der ersten Zeile ist die Abweichung des Nickwinkels ϕ_1 , in der Mitte die Abweichung des Schielwinkels ϕ_2 und unten ist die Abweichung des Rollwinkels ϕ_3 dargestellt. Dabei wird, wie in Gl. (5.1) beschrieben, die Abweichung zwischen der Ground Truth und der Online-Kalibrierung mit der Fehlerfunktion angegeben. Diese Abweichung ist auf den Ordinaten angegeben. Auf den Abszissen sind die Nummern der Kalibrierung abgetragen. In der Abb. 5.8 ist zu sehen, dass der Nickwinkel ϕ_1 einen Wertebereich zwischen $0 [Pix]$ und $0.5 [Pix]$ hat. Es gibt besondere Bereiche in der Abbildung, die besonders große Ausreißer über längere Abschnitte auf der Abszisse anzeigen. Diese liegen in den Bereichen der Online-Kalibrierung mit den Nummern $0 - 100$, $508 - 600$ und $1080 - 1433$. In diesem Abschnitt der Sequenz ist das Fahrzeug in eine Kurve gefahren.

Es ist zu erkennen, dass der Schielwinkel ϕ_2 die größten Abweichungen hat. Der Schiel- und Rollwinkel haben beide bei der Kalibrierung mit der Nummer 970 einen Ausreißer. Dieser Ausreißer entsteht dadurch, dass die Korrespondenzen nicht über das Bild verteilt sind, sondern dass die Korrespondenzen an nur einer Stelle im Bild auftauchen.

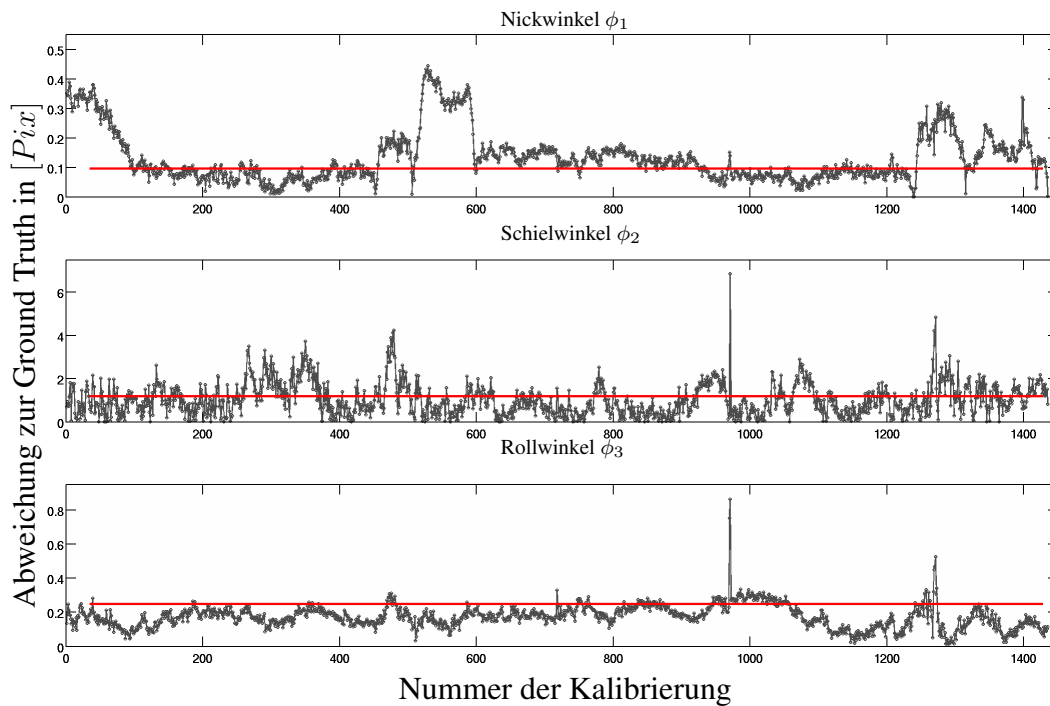


Abbildung 5.8.: **Lebesgue-Norm** ($p = 1$): Die drei Winkel (ϕ_1, ϕ_2, ϕ_3) sind hier abgebildet. Dabei sind die Residuen mit der Lebesgue-Norm mit $p = 1$ gewichtet worden. Ausreißer werden sehr gering gewichtet, sodass diese nur einen sehr geringen Anteil an der Online-Kalibrierung haben. Die Mittelwerte sind durch eine rote Linie in der Grafik markiert.

Es ist zu sehen, dass die Kurven der drei Graphen unterschiedlich sind. Die Abweichungen des Nickwinkels in Bezug auf den Schielwinkel sind zehnmal kleiner und die Abweichungen zwischen dem Rollwinkel in Bezug auf den Schielwinkel sind ähnlich.

Zu sehen ist, dass es viele Ausreißer gibt, die eine enorme Abweichung zur Ground Truth haben. Die größte Abweichung ist über sechs Pixel, was einen enormen Fehler in der 3D-Rekonstruktion hat.

In Abb. 5.9 ist zu sehen, wie sich der Verlauf der Online-Kalibrierung mit Binning ändert. Dies liegt am Sammeln der Korrespondenzen in Bins. In dieser Untersuchung wurden vier Bins eingerichtet, mit einem Schwellenwert von 400 Korrespondenzen pro Bin. Das wichtigste Ergebnis ist, dass Ausreißer über die gesamte Sequenz vermieden werden.

Lebesgue-Norm ($p = 2$)

In Abb. 5.10 wird die Lebesgue-Norm mit $p = 2$ auf der gleichen Sequenz wie die Lebesgue-Norm mit $p = 1$ angewendet.

In Abb. 5.11 ist die Lebesgue-Norm mit $p = 2$ mit dem Binning auf der gleichen Sequenz genutzt worden wie schon die Lebesgue-Norm mit $p = 2$ ohne Binning, wie sie in Abb. 5.10 abgebildet ist. Dabei ist zu sehen, dass durch das Binning diese inakzeptablen großen Abweichungen aus Abb. 5.10 unterdrückt werden und gar nicht mehr entstehen.

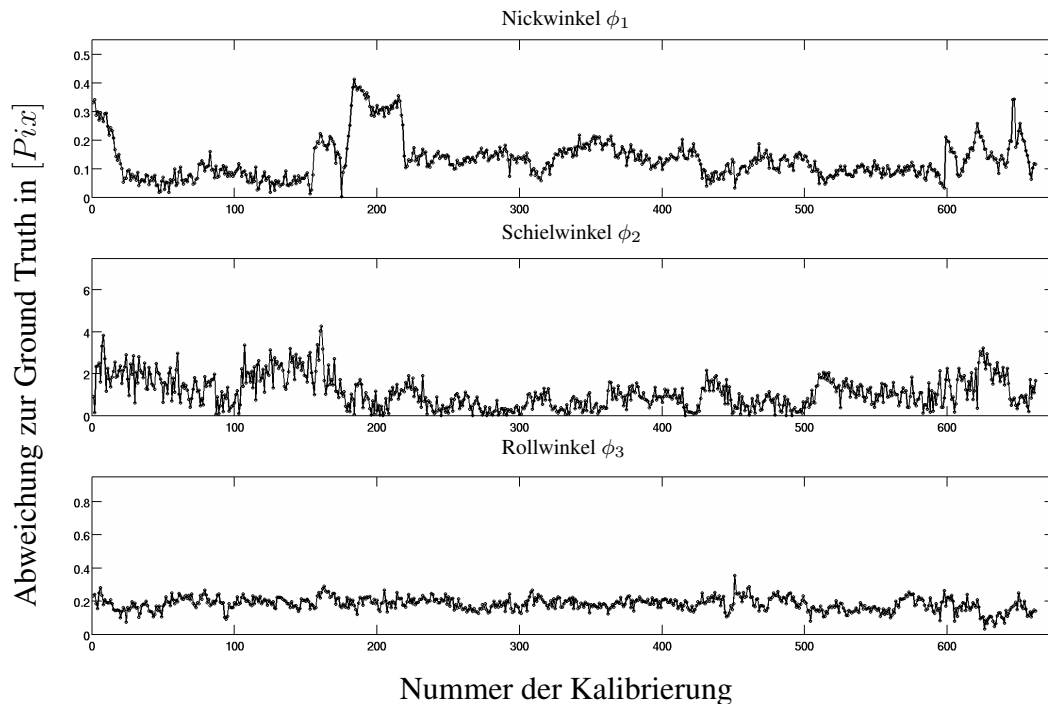


Abbildung 5.9.: **Lebesgue-Norm mit Binning** ($p = 1$): Der Verlauf der drei Winkel. Oben der Nickwinkel, in der Mitte der Schielwinkel und unten der Rollwinkel. Die Abzisse stellt die Nummer der Kalibrierung dar. Da bei dieser Untersuchung das Binning genutzt wurde, gibt es weniger Kalibrierungen.

Die Skalen der Graphen sind unterschiedlich. Während die Skalen des Nick- und Rollwinkels sehr nah beieinander liegen, ist die Abweichung im Graphen des Schielwinkels um den Faktor zehn größer als die der anderen Winkel. Daher muss der Schielwinkel stärker in den Fokus weiterer Untersuchungen eingebunden werden.

Anders als bei der Lebesgue-Norm ohne Binning, sind hier keine extremen Ausreißer zu sehen.

Corrupted-Gaussian

Die Corrupted-Gaussian wird in [Hartley und Zisserman, 2004] als eine Verbesserung der Lebesgue-Norm bezeichnet. Dabei ist $\alpha \in [0,1]$ und ω die angenommene Standardabweichung. Dabei wird der Parameter α genutzt, um einen glatten Übergang zwischen der Menge der Ausreißer hin zur Menge der zur Online-Kalibrierung genutzten Korrespondenzen zu haben.

Zur wissenschaftlichen Analyse dieser Fehlerfunktion wurde die gleiche Testsequenz wie bei der Lebesgue-Norm genutzt. Abb. 5.12 stellt das Ergebnis aus dieser Analyse dar. Es zeigt sich, dass auch hier die mittlere Abweichung $0,8 [Pix]$ beträgt. Dabei wurde eine Standardabweichung von $\omega = 1,02 [Pix]$ genutzt. Dieser Wert wurde genutzt, da bei der Untersuchung keine Fehlerfunktionen und Ausreißerdetektionen genutzt wurden und dies die Standardabweichung dieser Untersuchung ist.

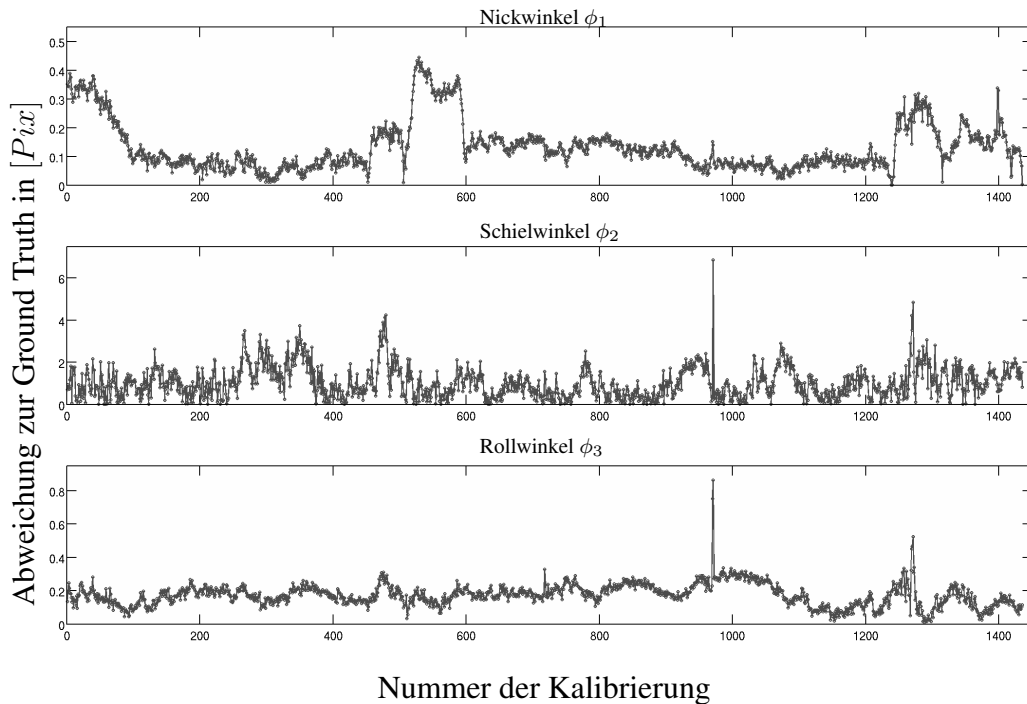


Abbildung 5.10.: **Lebesgue-Norm** ($p = 2$): Dargestellt ist der Verlauf der Abweichungen der drei Winkel (ϕ_1, ϕ_2, ϕ_3). Die Skalen sind dabei sehr unterschiedlich, da bei der Abweichung des Nickwinkels ϕ_1 und der Abweichung des Rollwinkels ϕ_3 sehr kleine Änderungen auftreten. Die Abweichung des Schielwinkels ϕ_2 ist in der Mitte abgetragen. Auf der Abzisse ist die Nummer der Kalibrierung abgebildet und auf der Ordinate ist die Abweichung des Winkels abgetragen.

Die Kurven des Nick- und Rollwinkels liegen nah beieinander, wobei die Abweichung des Schielwinkels zehnmal größer sind.

Es ist zu sehen, dass auch hier, wie schon bei der Lebesgue-Norm mit $p = 1$, enorme Ausreißer auftreten können, die in der 3D-Rekonstruktion einen enormen Fehler zur Folge haben.

In Abb. 5.13 ist zu sehen, wie sich das Binning auf das „Corrupted-Gaussian“ auswirkt. Auch hier sind die Ausreißer nicht mehr so stark und so zahlreich, was die Varianz stark verkleinert.

Blake-Zisserman

Die Blake-Zisserman-Fehlerfunktion, die in [Hartley und Zisserman, 2004] beschrieben ist, ist eine Alternative zur „Corrupted-Gaussian“-Fehlerfunktion. Es wird bei dieser Funktion angenommen, dass es zwischen der Menge der Ausreißer und den restlichen Korrespondenzen, anders als beim „Corrupted-Gaussian“, eine harte Schwelle gibt. Die Funktion ist in Tab. 5.7 und Tab. 5.8 angegeben. Dabei ist ξ der Faktor, der die Varianz der Ausreißer in der Menge angibt. Dieser Wert ist identisch mit dem Wert, der aus der Funktion $e^{\frac{(-\delta r_{ik})^2}{\omega}}$ in der „Corrupted-Gaussian“ berechnet wird.

Wie in Abb. 5.14 zu erkennen, stellt sich auch bei dieser robusten Fehlerfunktion das

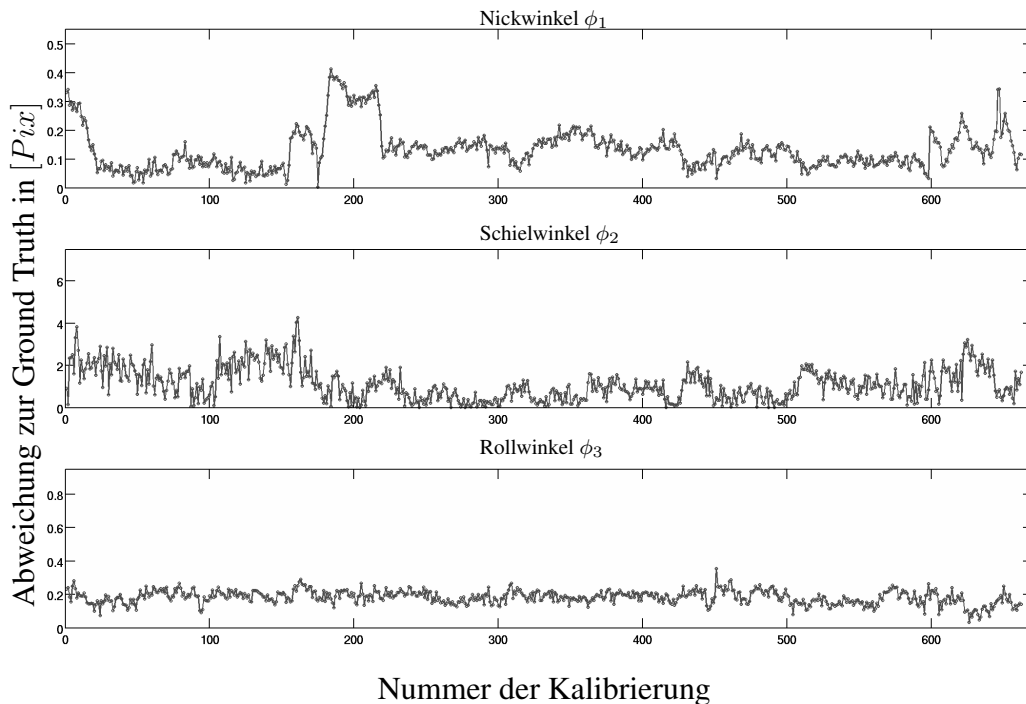


Abbildung 5.11.: **Lebesgue-Norm mit Binning** ($p = 2$): Dargestellt ist der Verlauf der Abweichungen der drei Winkel (ϕ_1, ϕ_2, ϕ_3). Die Skalen sind dabei sehr unterschiedlich, da bei der Abweichung des Nickwinkels ϕ_1 und der Abweichung des Rollwinkels ϕ_3 sehr kleine Änderungen vorliegen und die Schielwinkelabweichung ϕ_2 große Abweichungen zeigt.

Auf der Abzisse ist die Nummer der Kalibrierung abgebildet und auf der Ordinate die Abweichung des Winkels. Hier wurde das Binning mit hinzugenommen, deshalb ist die Anzahl der Kalibrierungen geringer.

Die Skalen der Graphen sind alle unterschiedlich. Dies ist bedingt durch die starken Abweichungen zwischen dem Nick- und Rollwinkel zum Schielwinkel, der um den Faktor zehn größer ist.

Anders als bei der Auswertung ohne Binning, ist hier zu erkennen, dass große Abweichungen, die durch die Ausreißer entstehen, ausbleiben.

gleiche Verhalten wie bei der zuvor untersuchten „Corrupted-Gaussian“-Funktionen, dabei sind die mittlere Abweichung und die Varianz sehr ähnlich.

Durch die Hinzunahme des Binnings wurden die Ausreißer auch hier minimiert und die Varianz verringert. Dies ist in Abb. 5.15 dargestellt.

Zusammenfassung

Es hat sich gezeigt, dass das Binning als Verteilung die Aufnahme großer Ausreißer während der Online-Kalibrierung verhindert. Die robusten Fehlerfunktionen aus der Literatur alleine konnten zur Online-Kalibrierung nicht zuverlässig genutzt werden. Eine Kalibrierung sollte zur Vermeidung der Ausreißer immer von einem Binning begleitet werden. Während der zusätzlichen Totzeit, die zum Füllen der Bins benötigt wird, ist kei-

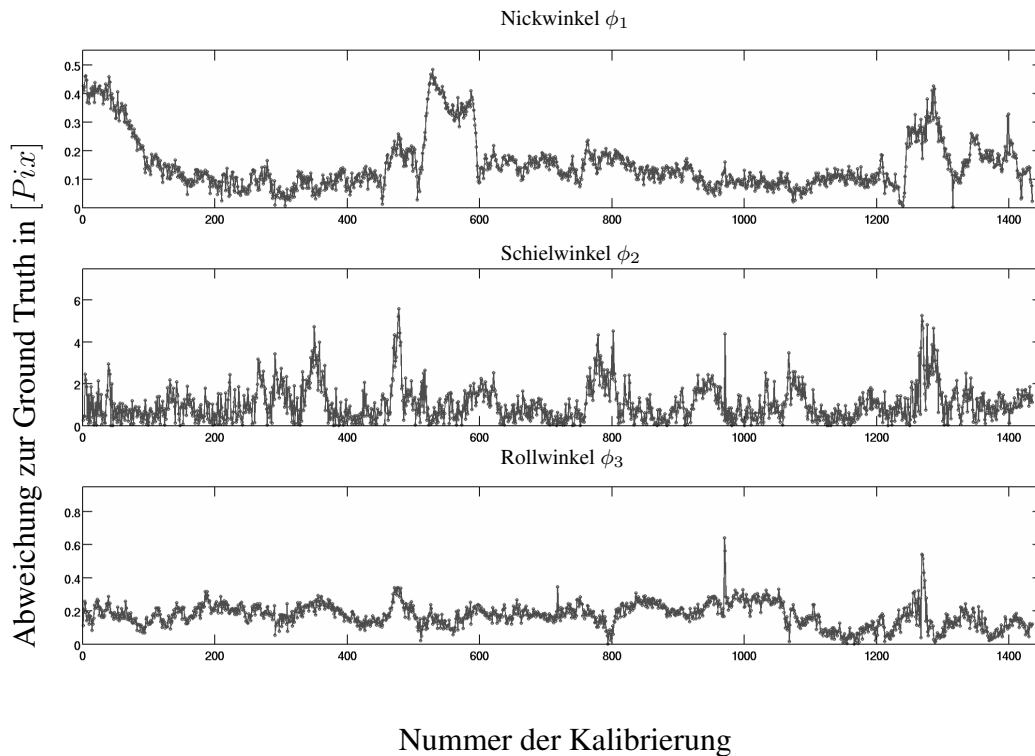


Abbildung 5.12.: **Corrupted-Gaussian:** Es ist wie bei den Untersuchungen zur Lebesgue-Norm zu sehen, dass die Ausreißer sehr häufig vorkommen und große Fehlerwerte haben können.

In den Skalen der Graphen ist zu erkennen, dass die Abweichung des Schielwinkels ϕ_2 bis zu zehnmal größer sein kann als die der anderen Winkel. Deshalb muss der Schielwinkel stärker in den Fokus zukünftiger Methoden zur Online-Kalibrierung rücken.

Es ist bei dieser Funktion zu sehen, dass sich viele große Abweichung der Ground Truth über die gesamte Zeit einstellen, die sich in der 3D-Rekonstruktion als enorme Fehlmesungen darstellen.

ne Schätzung der Kameraparameter möglich.

Es wird auch gezeigt, dass sich der Winkel ϕ_1 und der Winkel ϕ_3 durch die verwendeten Methoden nicht signifikant verbessern lassen.

5.4. Korrespondenzauswahl mit der *inneren Genauigkeit*

Wie in Kap. 2.2.4 beschrieben, ist die *innere Genauigkeit* ein Maß, das beim Bündelausgleich genutzt wird, um Ausreißer in der Menge der Korrespondenzen zu finden. Durch die Taylor-Reihenzerlegung kann die Gleichung, die den Bündelausgleich beschreibt, approximiert werden. So kann auch die Essential-Gleichung approximiert werden. Daraus folgt Gl. (2.3) nach dem dritten Glied der Reihe.

In Gl. (2.3) wird ein Gewicht W genutzt, um den Einfluss von Korrespondenzen zu

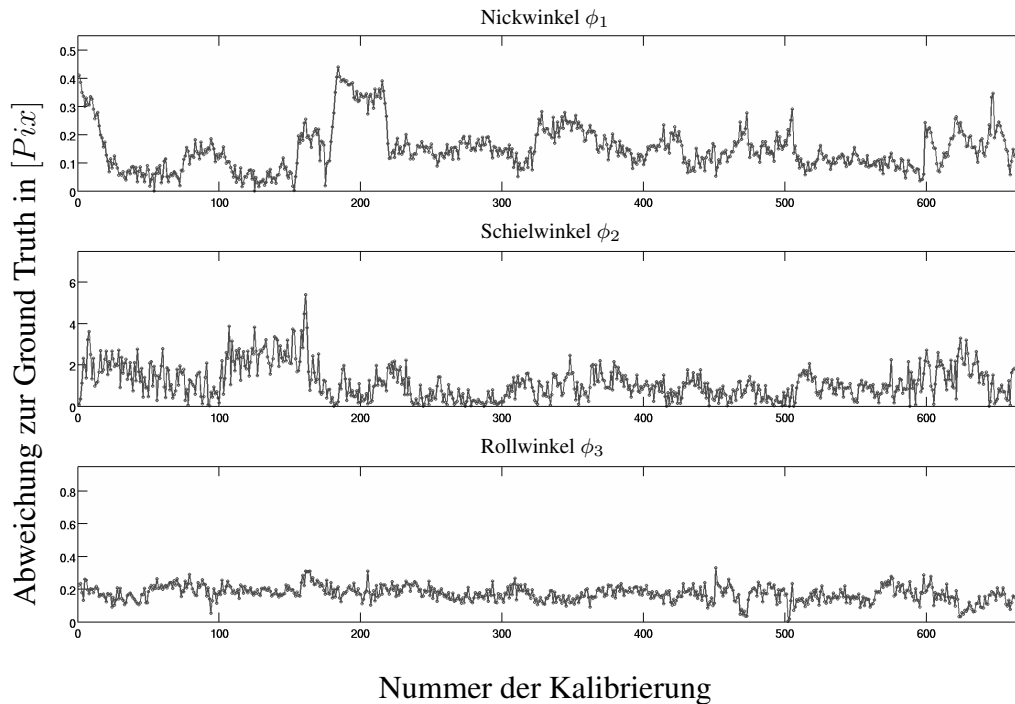


Abbildung 5.13.: **Corrupted-Gaussian mit Binning:** Jeder Punkt ist die Abweichung der Winkel nach der Online-Kalibrierung. Dabei zeigt der obere Graph die Abweichung des Nickwinkels ϕ_1 , der mittlere Graph die Abweichung des Schielwinkels ϕ_2 und der untere Graph stellt die Abweichung des Rollwinkels ϕ_3 dar.

Die Skalen der Graphen sind unterschiedlich, da der Schielwinkel eine bis zu zehnmal höhere Abweichung zur Ground Truth hat als die anderen beiden Winkel.

Das Verwenden von Binning zeigt auf der verwendeten Sequenz, dass sich die Ausreißer minimieren und sich so in der 3D-Rekonstruktion keine enormen Fehlmessungen einstellen.

bewerten. Da bei der Online-Kalibrierung nicht klar ist, welches Gewicht jeder Korrespondenz zugeordnet ist, wird das Gewicht für alle W gleich gesetzt. D.h. das Gewicht $W_i = 1 \forall i$. Der Schwellwert k , der die Ausreißer identifiziert, wurde empirisch bestimmt. Zur Vergleichbarkeit zu den vorigen Untersuchungen wird auch hier die erste Sequenz aus Tab. 5.2 “Ringweg” genutzt. In Tab. 5.9 sind die Auswertungen dargestellt. Dabei wurde für jede Kombination der Werte der Mittelwert nach der Kalibrierung bestimmt. Dieser Mittelwert ist in der Tabelle eingetragen. Die Variablen, die zyklisch geändert werden, waren die Größe des Bins und der Schwellwert k aus Gl. (2.3).

Tabelle 5.9.: Um zu untersuchen, welchen Einfluss die *innere Genauigkeit* hat, werden der Schwellwert und die Anzahl der Korrespondenzen im Bin geändert. Dabei ergeben sich unterschiedliche mittlere Abweichungen nach dem Durchlaufen einer Sequenz.

Dabei $k \in [0, 1[$ der Schwellwert, der angibt, wann eine Korrespondenz ein Ausreißer ist und wann nicht. Dabei gibt es keinen Schwellwert, wenn $k = 0$ gilt und für $k = 1$ sind alle Korrespondenzen Ausreißer.

Schwellenwert k	Binningparameter $\left[\frac{\text{Korrespondenzen}}{\text{Bin}} \right]$	Mittlere Ab- weichung der Winkel $(\phi_1, \phi_2, \phi_3) [Pix]$
0	100	(0.35, 1.90, 0.35)
0	300	(0.33, 0.98, 0.21)
0	400	(0.33, 0.95, 0.21)
0	500	(0.33, 0.94, 0.21)
0	600	(0.33, 0.91, 0.21)
0	700	(0.33, 0.89, 0.21)
0.1	100	(0.33, 1.0, 0.21)
0.1	300	(0.33, 0.87, 0.21)
0.1	400	(0.33, 0.83, 0.21)
0.1	500	(0.33, 0.82, 0.21)
0.1	600	(0.33, 0.79, 0.21)
0.1	700	(0.33, 0.77, 0.21)
0.2	100	(0.33, 0.99, 0.21)
0.2	300	(0.33, 0.86, 0.21)
0.2	400	(0.33, 0.84, 0.21)
0.2	500	(0.33, 0.81, 0.20)
0.2	600	(0.33, 0.78, 0.20)
0.2	700	(0.33, 0.76, 0.20)
0.3	100	(0.33, 0.98, 0.20)
0.3	300	(0.33, 0.85, 0.20)
0.3	400	(0.33, 0.83, 0.20)
0.3	500	(0.33, 0.79, 0.20)
0.3	600	(0.33, 0.77, 0.20)
0.3	700	(0.33, 0.75, 0.20)
0.4	100	(0.33, 0.97, 0.20)
0.4	300	(0.33, 0.84, 0.20)
0.4	400	(0.33, 0.82, 0.20)
0.4	500	(0.33, 0.78, 0.20)
0.4	600	(0.33, 0.75, 0.20)
0.4	700	(0.33, 0.73, 0.19)
0.5	100	(0.33, 0.95, 0.20)
0.5	300	(0.33, 0.83, 0.20)
0.5	400	(0.33, 0.83, 0.19)
0.5	500	(0.33, 0.77, 0.19)

0.5	600	(0.33, 0.74, 0.19)
0.5	700	(0.33, 0.72, 0.19)
0.6	100	(0.33, 0.94, 0.20)
0.6	300	(0.33, 0.87, 0.19)
0.6	400	(0.33, 0.87, 0.19)
0.6	500	(0.33, 0.76, 0.19)
0.6	600	(0.33, 0.73, 0.19)
0.6	700	(0.33, 0.71, 0.19)
0.7	100	(0.33, 0.94, 0.19)
0.7	300	(0.33, 0.87, 0.19)
0.7	400	(0.33, 0.87, 0.19)
0.7	500	(0.33, 0.75, 0.19)
0.7	600	(0.33, 0.75, 0.19)
0.7	700	(0.33, 0.74, 0.20)
0.8	100	(0.33, 0.95, 0.20)
0.8	300	(0.33, 0.87, 0.20)
0.8	400	(0.33, 0.86, 0.20)
0.8	500	(0.33, 0.77, 0.20)
0.8	600	(0.33, 0.76, 0.19)
0.8	700	(0.33, 0.76, 0.19)
0.9	100	(0.33, 1.0, 0.20)
0.9	300	(0.33, 0.9, 0.20)
0.9	400	(0.33, 0.84, 0.20)
0.9	500	(0.33, 0.78, 0.20)
0.9	600	(0.33, 0.78, 0.19)
0.9	700	(0.33, 0.76, 0.20)
1.0	100	(0.33, 1.0, 0.20)
1.0	300	(0.33, 0.9, 0.20)
1.0	400	(0.33, 0.84, 0.20)
1.0	500	(0.33, 0.78, 0.20)
1.0	600	(0.33, 0.78, 0.19)
1.0	700	(0.33, 0.76, 0.20)

Eine grafische Darstellung von Tab. 5.9 ist in Abb. 5.16 dargestellt. Es ist zu sehen, dass der Schwellenwert k einen geringeren Einfluss auf die Online-Kalibrierung hat als das Binning. Aus Abb. 5.16 ist auch zu entnehmen, dass, wenn keine Online-Kalibrierung während einer Fahrt vorgenommen wird, eine Abweichung von 1.9 [Pix] im Mittel über die Fahrt gegeben ist. Dies entspricht nicht der erwarteten Distanz von 102.40 [m], son-

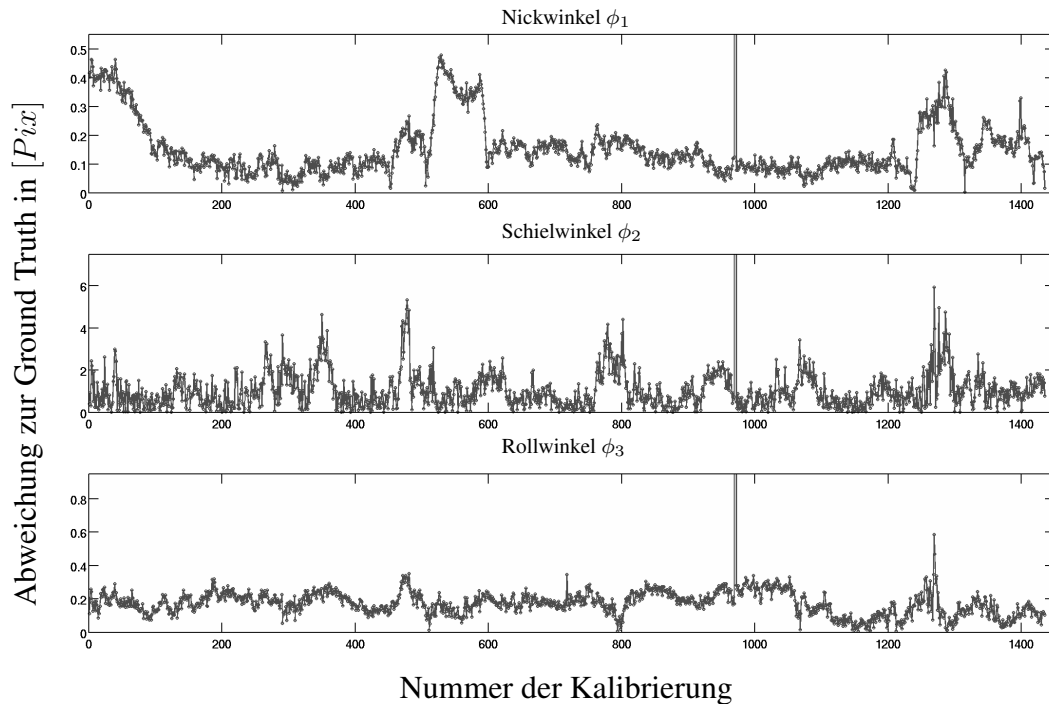


Abbildung 5.14.: **Blake-Zisserman:** Dabei zeigt der obere Graph die Abweichung des Nickwinkels ϕ_1 , der mittlere Graph die Abweichung des Schielwinkels ϕ_2 und der untere Graph stellt den Verlauf der Kalibrierungen über die Zeit der Abweichung des Rollwinkels ϕ_3 dar.

Die Abzisse stellt die Nummer der Kalibrierung in der Testsequenz dar. Die Ordinate zeigt die Abweichung der neuen Werte aus der Kalibrierung gegenüber der Ground Truth.

Auch hier ist die Skala der Ordinate bei den einzelnen Winkeln unterschiedlich. Dies ist notwendig, da der Schielwinkel um den Faktor zehn größer ist als die beiden anderen. D.h. es ist wichtig, dass der Fokus in der Entwicklung zukünftiger Methoden den Fokus auf die Kalibrierung des Schielwinkels legt.

Es ist zu erkennen, dass sich auch bei dieser Untersuchung starke Abweichungen einstellen und so große Fehlmessungen in der 3D-Rekonstruktion zu erkennen sind.

dern einer Rekonstruktion von 52.51 [m] oder 2048.00 [m]. Diese enorme Abweichung von der eigentlichen Distanz zeigt auch, dass eine Kalibrierung essenziell für das Betreiben eines Stereo-Kamera-Systems zum Vermessen der Fahrzeugumgebung und das Regeln von Aktoren im Fahrzeug ist.

Durch das Binning wird, wie in Kap. 5.2 gezeigt, der Einfluss von Ausreißern durch eine Gewichtung während einer Sequenz signifikant minimiert. Dieser Versuch zeigt, dass die *innere Genauigkeit*, wie hier verwendet, nur einen geringen Einfluss auf die Online-Kalibrierung hat.

In Abb. 5.17 ist ein Bild aus der Testsequenz zu sehen. Die grünen Linien sind die Korrespondenzen, die zur Kalibrierung genutzt werden, und die roten Linien sind die Ausreißer. Dabei ist die Einstellung für das Binning so, dass die Größe der einzelnen Bins auf 400 gesetzt ist, mit einem Schwellwert der *inneren Genauigkeit* von $k = 0.6$.

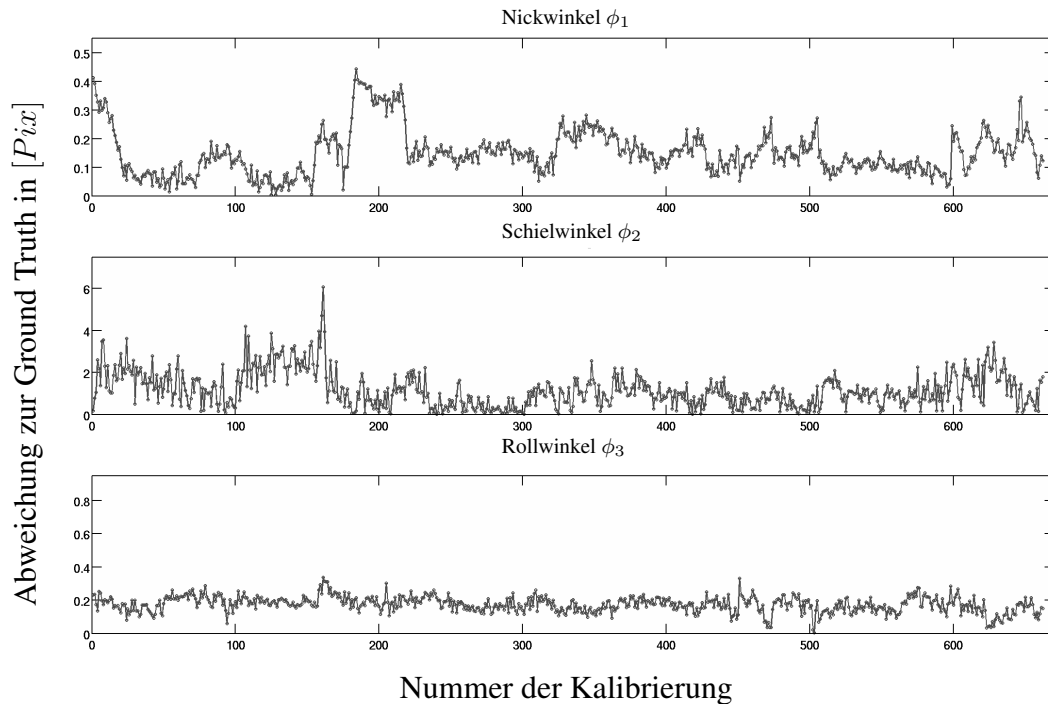


Abbildung 5.15.: **Blake-Zisserman mit Binning:** Dabei ist zeigt obere Graph den Nickwinkel ϕ_1 , der mittlere Graph den Schielwinkel ϕ_2 und der untere Graph stellt den Verlauf der Kalibrierungen über die Zeit des Rollwinkels ϕ_3 dar. Die Abzisse stellt die Nummer der Kalibrierung in der Sequenz dar. Die Ordinate zeigt die Abweichung der neuen Werte aus der Kalibrierung gegenüber der Ground Truth.

Die Ordinaten der Graphen sind unterschiedlich. Da die Abweichung des Schielwinkels um den Faktor zehn größer ist als die Abweichung der Winkel ϕ_1 und ϕ_3 , ist dies Wahl der Skala wichtig, da sonst die Abweichungen des Nick- und Rollwinkels sich nicht qualitativ darstellen lassen.

Durch das Binning wird, anders als bei der Untersuchung ohne Binning, den Ausreißern so entgegengewirkt, dass diese während der gesamten Sequenz nicht auftreten.

Zu sehen ist, dass sich die Zahl der durch das Binning gesammelten Korrespondenzen reduzieren lässt, da diese als Ausreißer detektiert werden. Das Zusammenspiel von *innerer Genauigkeit* und Binning verbessert die Online-Kalibrierung.

Zusammenfassung

Die *innere Genauigkeit* hebt sich nicht signifikant von anderen Fehlerfunktionen, die in Kap. 5.3 genutzt wurden, ab. Binning in Kombination mit der *inneren Genauigkeit* zeigt das gleiche Verhalten wie die Auswertung der robusten Fehlerfunktionen in Kap. 5.3.

5.5. Rauscheinfluss auf die Online-Kalibrierung

Um zu zeigen, welche Auswirkung das Rauschen der Daten aus einem Bildpaar auf die Kalibrierung hat, wird berechnet, wie es sich mit der 3D-Rekonstruktion verhält.

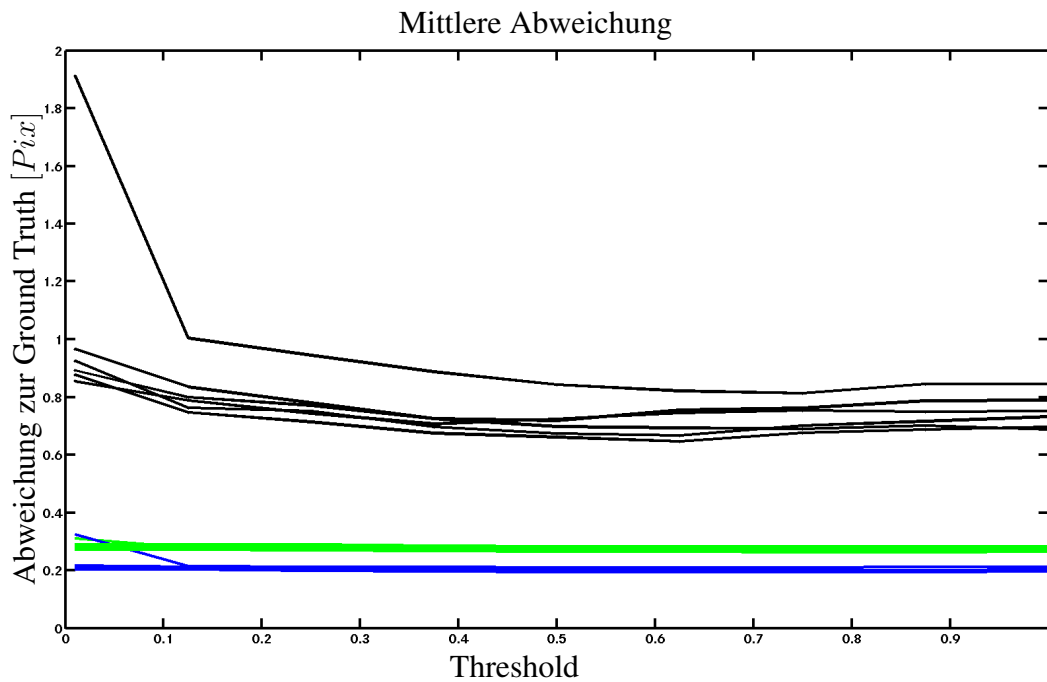


Abbildung 5.16.: Mit der Methode der *inneren Genauigkeit* können Ausreißer in der Menge der Korrespondenzen gefunden werden. Dabei wird ein Schwellwert k genutzt, der hier in der Abzisse abgetragen ist und angibt, wann eine Korrespondenz ein Ausreißer ist oder zur Kalibrierung genutzt wird. Die Abweichung des Nickwinkels (ϕ_1) wird hier in Grün dargestellt, die Schielwinkelabweichung (ϕ_2) in Schwarz und die Rollwinkelabweichung (ϕ_3) in Blau. Dabei ist jedes Datum der Mittelwert aus einem Durchlauf durch die Testsequenz. Die unterschiedlichen Linien zu einem Winkel sind die Verläufe bei Nutzen des Binning, um Korrespondenzen zu sammeln.

Die schwarzen Kurven zeigen, welchen Einfluss das Binning hat und welchen Einfluss der Schwellenwert k auf die Kalibrierung des Stereo-Kamera-Systems hat. Der Schielwinkel (ϕ_2) hat ohne eine Ausreißerdetektion eine mittlere Abweichung von mehr als $1.9 [Pix]$. Die sieben schwarzen Kurven, die die mittlere Abweichung über die Testsequenz des Schielwinkels ϕ_2 zur Ground Truth zeigen, stellen das Nutzen des Binning dar. Dabei werden verschiedene Schwellenwerte für die Anzahl der Korrespondenzen pro Bin genutzt. Dabei ist zu sehen, dass sich durch das Nutzen des Binnings die mittlere Abweichung des Schielwinkels ϕ_2 über die Sequenz um $1.0 [Pix]$ verbessert.

Die Verwendung der *inneren Genauigkeit* verbessert den Schielwinkel im Mittel über die Sequenz um $0.9 [Pix]$.

In Gl. (1.5) wird beschrieben, welchen Einfluss additives Rauschen auf die 3D-Rekonstruktion hat. Angenommen ist der Versuchsaufbau des Stereo-Kamera-Systems aus Abbildung 5.1. Die Kamerakonstanten beider Kameras sind $1024 [Pix]$ und eine Basisbreite von $0.2 [m]$. Das ergibt eine Stereo-Kamera-Konstante von $204.8 [Pix \cdot m]$. Um die Auswirkung von Rauschen zu bestimmen, wird eine Disparität von $2 [Pix]$ genommen, was einem Punkt in einer Entfernung von $102.40 [m]$ in der 3D-Rekonstruktion entspricht. Dabei ist zu berücksichtigen, dass Gl. (1.6) zur Berechnung genutzt wird.

Sei nun das Rauschen mit $\pm 0.8 [Pix]$ so groß, wie in der Literatur angegeben. Dann weicht der beobachtete Punkt von der eigentlichen Tiefe ab. Der Punkt in der 3D-Rekon-

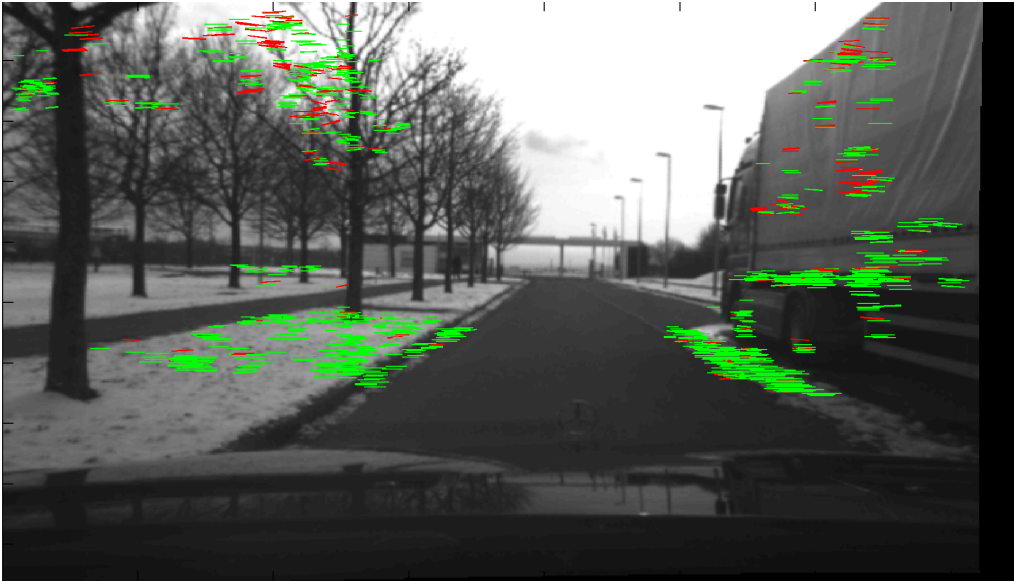


Abbildung 5.17.: Dieses Bild ist ein Bild aus der Testsequenz, um zu zeigen, wie die Ausreißer in der Menge der Korrespondenzen angeordnet sind. Die Außreißer, die mit der Methode der *inneren Genauigkeit* bestimmt wurden, sind als rote Linien markiert und die Korrespondenzen, die zur Kalibrierung genutzt werden, sind grün eingezeichnet.

struktion ist dann zwischen $\frac{204.8}{2.0+0.8} = 73.14 [m]$ und $\frac{204.8}{2.0-0.8} = 170.67 [m]$ entfernt.

Um nun darzustellen, wie genau in einem Bild gemessen werden muss, um einen Punkt in einer Distanz von $102.40 [m]$ zu berechnen, wird dieser Fehler zwischen der eigentlichen und der berechneten Position gesucht.

Für diese Untersuchung werden die Kameraparameter des Versuchsaufbau benutzt und eine Disparität von $2.0 [Pix]$ wird angenommen.

$$\begin{aligned}
 \hat{D} &= \frac{f \cdot b}{d} [m] \\
 &= \frac{204.8}{2.0} [m] \\
 \underbrace{\hat{D}=D \pm \Delta D}_{\Rightarrow} d &= \frac{204.8}{\hat{D}} [m]
 \end{aligned} \tag{5.2}$$

Sei nun eine 3D-Rekonstruktion mit einer Genauigkeit von $\pm 0.05 [m]$ in der Distanz von $102.40 [m]$ gefordert, dann folgt mit Gl. (5.2), dass die gemessene Disparität zwischen $\frac{204.8}{102.4+0.05} = 1.99902 [Pix]$ oder $\frac{204.8}{102.4-0.05} = 2.000977 [Pix]$ sein muss.

Es ist leicht zu erkennen, dass die Disparitäten nicht auf $\frac{9}{10000} [Pix]$ bestimmt werden können, wodurch eine solche Genauigkeiten in diesem Entfernungsbereich für diesen Versuchsaufbau nicht zu erreichen ist.

Eine Abweichung der Disparität von $0.2 [Pix]$ mit dem Versuchsaufbau wurde voraus-

gesetzt. Dies entspricht in der 3D-Rekonstruktion $\frac{204.8}{2.0+0.2} = 93.09 [m]$ bis $\frac{204.8}{2.0-0.2} = 113.78 [m]$ statt der $102.40 [m]$, die bei rauschfreien Daten möglich wären. Selbst mit diesen geringen Abweichungen von $0.2 [Pix]$ ist die 3D-Rekonstruktion auf $\pm 10 [m]$ genau.

Zusammenfassung

Diese Untersuchung zeigt, dass sich die Genauigkeit in der Ermittlung der Distanzen zu einem Objekt nur bis zu einem gewissen Maß erreichen lässt. Heutige Hardware und Methoden sind nicht in der Lage, die Genauigkeit auf wenige Zentimeter in der betrachteten Distanz von $102.40 [m]$ zu gewährleisten. Für kürzere Distanzen können höhere Genauigkeiten erreicht werden, da der Disparitätsfehler mit $\frac{1}{d}$ in den Distanzfehler eingeht.

5.6. Die Auswertung der Online-Kalibrierungsmethoden

Da die relative Ausrichtung der Kamerakoordinatensysteme zueinander achsenparallel ist, d.h. mit der z-Achse beider Kameras in Fahrtrichtung als gegeben vorausgesetzt wird, ist das System sehr fehleranfällig bezüglich kleiner Abweichungen im Schielwinkel, da bei kleinen Änderungen im Schielwinkel sich die Positionen der Epipole stark ändern [Kanatani, 2005].

Es hat sich in den vorherigen Untersuchungen gezeigt, dass der Nickwinkel ϕ_1 und der Rollwinkel ϕ_3 unabhängig von der Ausreißerbehandlung zu schätzen sind und um den Faktor zehn kleiner sind als der Schielwinkel ϕ_2 . Da aber in diesen beiden Stereo-Kamera-Parametern ϕ_1, ϕ_3 die Abweichung nur linear eingeht und in den Schielwinkel ϕ_2 mit $\frac{1}{d}$, wobei d die Disparität ist, ist für den Winkel ϕ_2 mit einer stärkeren Abweichung zu rechnen. In Gl. (1.6) ist dies hergeleitet und in Abb. 1.16 ist dies für den Schielwinkel (ϕ_2) abgebildet.

Deshalb wurden in Kap. 4 zwei Methoden zur Schielwinkelbestimmung beschrieben.

5.6.1. Online-Kalibrierung mit Bezug auf die Entfernungsdaten (Methode 1)

Die bisherigen Untersuchungen haben gezeigt, dass durch die Methoden der Ausreißererkennung und durch die Organisation der Korrespondenzmenge sich ein Stereo-Kamera-System nicht so kalibrieren lässt, dass die Abweichung im Mittel $0.2 [Pix]$ mit einer Standardabweichung von maximal $0.1 [Pix]$ eingehalten werden kann. Wird ein Punkt in der Distanz von $102.40 [m]$ betrachtet, dann wird bei einer Dekalibrierung des Schielwinkels ϕ_2 von $0.2 [Pix]$ der Punkt zwischen $113.78 [m]$ und $93.09 [m]$ rekonstruiert.

Abschn. 4.1 stellt eine der beiden Methoden zur Online-Kalibrierung vor, die im Rahmen dieser Arbeit entwickelt wurden. Mit dieser Methode kann ein Stereo-Kamera-System online kalibriert werden. Dazu werden die Informationen aus dem Fahrzeug genutzt, die Auskunft über die Fahrzeuggierate und die Geschwindigkeit des Fahrzeugs geben. Aus diesen Informationen kann dann errechnet werden, wie weit das Fahrzeug zwischen den Aufnahmen zweier aufeinanderfolgender Bildpaare gefahren ist. Daraus kann dann mit Gl. (4.2) die Entfernung bestimmt werden. Dann gibt es zwei Lösungen $\Delta d_{1,2}$ für den Disparitätsfehler.

Ist $\Delta d_{1,2} = 0$, so ist das Stereo-Kamera-System kalibriert und das Stereo-Kamera-System kann mit den vorherigen Werten rektifiziert werden. Ist das Stereo-Kamera-System nicht kalibriert, so gilt nach Gl. (4.7)

$$\Delta d_1 = \frac{d_\tau + d_{\tau-1}}{2} + \sqrt{\frac{(d_\tau + d_{\tau-1})^2}{4} + C} \quad (5.3)$$

oder

$$\Delta d_2 = \frac{d_\tau + d_{\tau-1}}{2} - \sqrt{\frac{(d_\tau + d_{\tau-1})^2}{4} + C}. \quad (5.4)$$

Die Abweichung des Schielwinkels ΔD wird aus den Fehlern der einzelnen Korrespondenzen bestimmt. Dazu wird der Mittelwert aus den Abweichungen der einzelnen Korrespondenzen berechnet. Da $\Delta d_1 > \Delta d_2$ gilt, folgt $\Delta D = \frac{1}{n} \sum_{i=0}^n \Delta_i d_1$ mit i der Index der Verschiebung der i -ten Korrespondenz.

Um das Rauschen in den Daten zu verkleinern, wurde das Binning, wie in Kap. 5.3 und Kap. 5.4 beschrieben, genutzt. Zur Auswertung wurde die Größe der Bins auf 400 Korrespondenzen festgelegt, wie in Kap. 5.4 bestimmt.

Zur Auswertung der Methode wurden mit diesem Kamerasystem fünf Stunden Videomaterial und Fahrzeuginformationen aufgezeichnet. D.h. auf allen sieben Sequenzen aus Tab. 5.2.

Die Methode wurde so implementiert, dass zuerst eine grobe Kalibrierung vorgenommen wird. Während dieser Phase der Kalibrierung wird vorausgesetzt, dass der Schielwinkel dem Winkel aus der Initialkalibrierung entspricht und es werden ausschließlich der Nick- und der Rollwinkel kalibriert, bis die Abweichung zweier aufeinanderfolgender Schätzungen von Nick- und Rollwinkel unter eine definierte Schwelle fallen.

Ist die Abweichung kleiner als ein Pixel, dann werden alle drei Winkel online kalibriert, sodass die gesamte Rotationsmatrix der Essentialmatrix bestimmt werden kann. Die Schielwinkelergebnisse werden mittels eines Medianfilters der Länge fünf gefiltert, um Rauschen zu unterdrücken. Um die Konvergenzfähigkeit der Methode zu zeigen, wird der

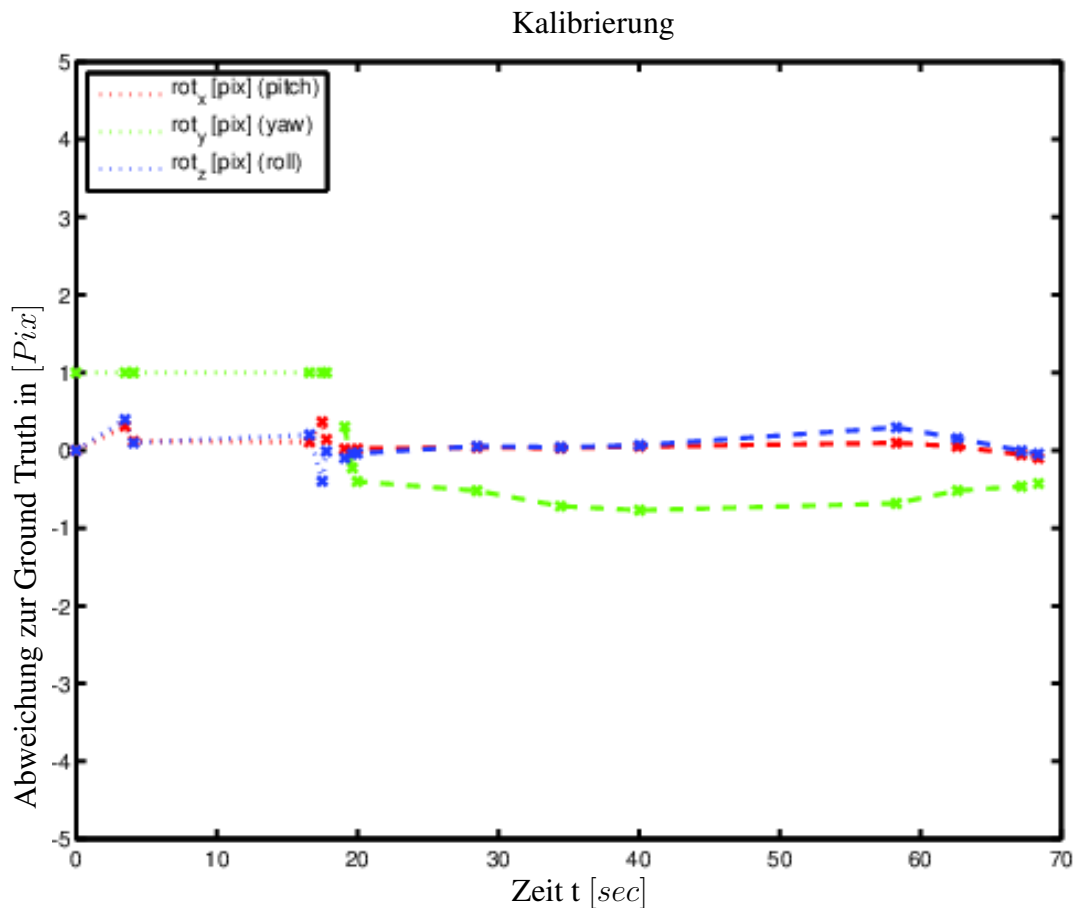


Abbildung 5.18.: **Methode 1 (ohne Medianfilter)**: In der Abbildung wird dargestellt, wie sich die Kalibrierung über die Zeit einstellt. Hier sind die zwei Phasen der Kalibrierung zu sehen. In den ersten 17 [sec] werden der Nick- und der Rollwinkel (ϕ_1, ϕ_3) so lange kalibriert, bis die Abweichung dieser beiden Winkel unter einem Pixel ist, dabei wird der Schielwinkel ϕ_2 künstlich auf den Wert von einem Pixel gesetzt, um die Konvergenz der Methode 1 in der zweiten Phase zu zeigen.

In der zweiten Phase wird zu den Winkeln ϕ_1 und ϕ_3 zusätzlich der Schielwinkel ϕ_2 kalibriert. Es zeigt sich, dass die Methode 1 konvergiert.

Durch die kurvige Straße wird die Abweichung vergrößert, weil bei einem Kurvenwechsel zwischen zwei Bildern der Flussvektor nicht zum Fahrzeugschielwinkel passt. Da aber eine Geradeausfahrt angenommen wird, wird mit den falschen Flussvektoren eine Online-Kalibrierung vorgenommen.

Ergebnis: Es zeigt sich, dass mit der Methode 1 eine mittlere Abweichung von 0.6 [Pix] auf der untersuchten Sequenz bei einer Standardabweichung von 0.3 [Pix] erreicht wird.

Startwert des in der Initialkalibrierung bestimmten Schielwinkels mit einem additivem Fehler von einem Pixel beaufschlagt.

Abb. 5.18 zeigt die Sequenz einer Kalibrierung. Die Sequenz „Albrundfahrt“ aus Tab. 5.2 stammt von einer Fahrt bei Nebel über eine kurvige Strecke auf der Alb. Dabei wird in Abb. 5.18 die Abweichung nach den in Gl. (5.1) beschriebenen Werten abgetragen. D.h. jeder Punkt entspricht der Abweichung zwischen der Online-Kalibrierung zur Initialkalibrierung. Durch die sehr kurvige Fahrt auf der schwäbischen Alb ist zu sehen, dass die Methode hochwertige und robuste Fahrzeug-Daten nutzt. Die Abweichungen von mehr

als $0.6 [Pix]$ des Schielwinkels sind auf die kurvenreiche Strecke zurückzuführen.

Zur Bestimmung, ob das Fahrzeug geradeaus fährt, was eine der Voraussetzungen der Methode ist, wird die Fahrzeuggierrate genutzt. Bei S-Kurven ist ein einmaliger Nullwert ein Zeichen für einen Wendepunkt, daher werden mehrere folgende Werte der Gierrate betrachtet, die ungefähr dem Nullwert entsprechen.

Die in diesem Abschnitt beschriebene Methode ohne Gierraten-Medianfilter wird in Abb. 5.18 dargestellt. Die Abweichung des Schielwinkels ϕ_2 von der Ground Truth ist dann $0.5 [Pix]$ mit einer Standardabweichung von $0.3 [Pix]$. Dieser Wert liegt unter dem Wert von $0.8 [Pix]$, die mit den Methoden aus [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] auf den Sequenzen erreicht wurden.

In einem Speicher können die Fahrzeuggierraten, die für die Kalibrierung gebraucht werden, abgelegt und abgeglichen werden, sodass gesichert ist, dass das Fahrzeug über mehrere Aufnahmen der Stereobilder geradeaus gefahren ist. Das Speichern und Abgleichen der Fahrzeuggierrate verbessert die Kalibrierung. Dies ist in Abb. 5.19 zu sehen. Wie auch in der Analyse zuvor werden die drei Abweichungen der Winkel ϕ_1 bis ϕ_3 über die Zeit dargestellt. Dabei wird darauf geachtet, dass das Fahrzeug auf fünf aufeinanderfolgenden Bildern geradeaus gefahren ist.

Wie zuvor stellt sich nach $17 [sec]$ die zweite Phase der Kalibrierung ein, bei der alle drei Winkel (ϕ_1, ϕ_2, ϕ_3) geschätzt werden. Bei der kurvigen Strecke ergibt sich, abweichend von den anderen Sequenzen, eine mittlere Abweichung des Schielwinkels ϕ_2 von $0.3 [Pix]$ bei einer Standardabweichung von $0.17 [Pix]$, bei einer Methode 1 ohne Medianfilter.

Mit dem eingesetzten Medianfilter und der Überprüfung der Gierrate sowie dem Filtern der Gierrate mit dem Medianfilter hat sich gezeigt, dass sich die mittlere Abweichung von $0.19 [Pix]$ mit einer Standardabweichung von $0.13 [Pix]$ relativ zur Ground Truth über alle Sequenzen nach im Mittel $20 [sec]$ eingestellt hat.

Dieser Wert ist um ein Viertel kleiner als alle aus der Literatur bekannten Werte zur Online-Kalibrierung in dynamischen Umgebungen. Dazu wurden die Methoden von [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] implementiert und auf den gleichen Sequenzen getestet, um die Vergleichbarkeit der Methoden zu zeigen.

Die Auswertung der Methoden aus dem Stand der Technik zur Schezung des Schielwinkels ist in Abb. 5.20.

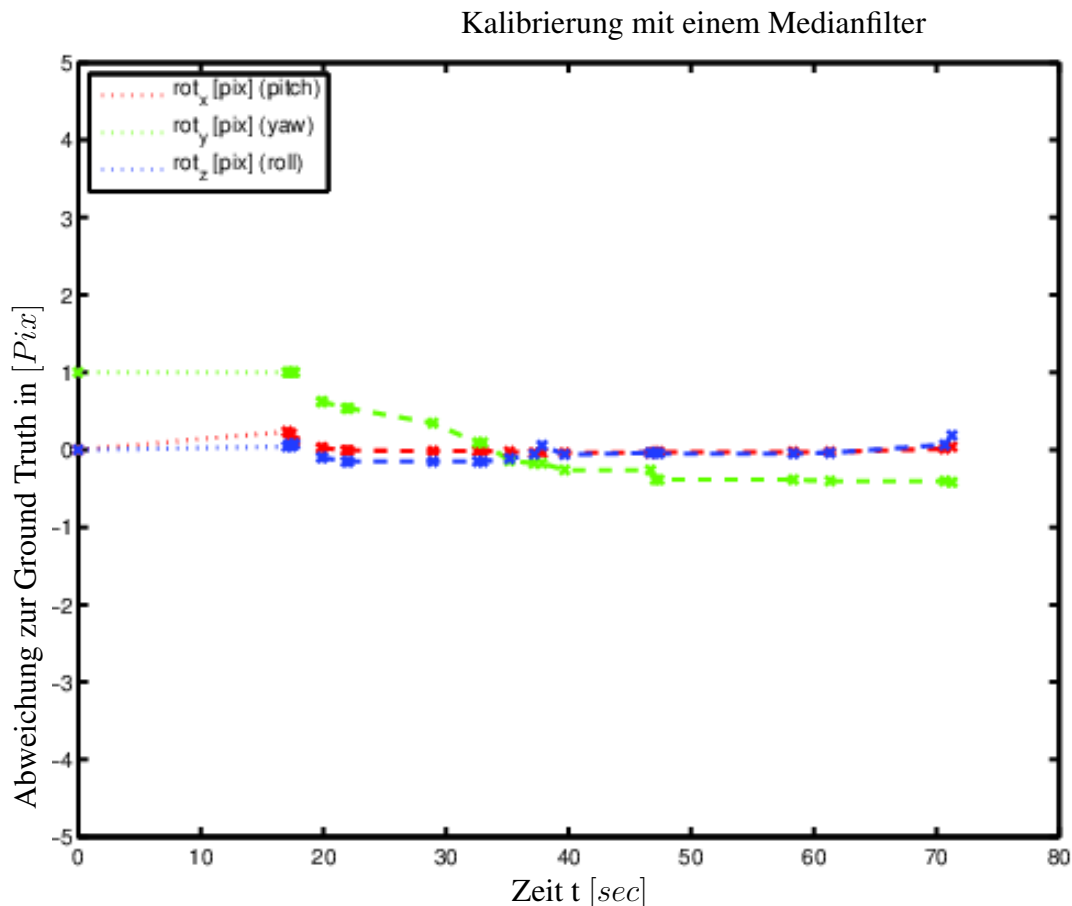


Abbildung 5.19.: **Methode 1 (mit Medianfilter)**: Dies ist das beste Ergebnis der ersten Methode. In der Abbildung wird dargestellt, wie sich die Kalibrierung über die Zeit einstellt, wenn ein Medianfilter der Länge fünf genutzt wird. Hier sind die zwei Phasen der Kalibrierung zu sehen. In den ersten 17 [sec] wird die Grobkalibrierung genutzt, die den Nick- und den Rollwinkel (ϕ_1, ϕ_3) einstellt. Dabei wird der Schielwinkel ϕ_2 künstlich auf den Wert von einem Pixel eingestellt. Dies soll zeigen, dass in der zweiten Phase der Kalibrierung die Methode 1 konvergiert. Wenn der Roll- und Nickwinkel bis auf ein Pixel genau eingestellt sind, werden alle drei Winkel (ϕ_1, ϕ_2, ϕ_3) eingestellt. In dieser Abbildung ist der Schielwinkel auf ein Pixel gesetzt worden, um die Konvergenz des Winkels durch die Kalibrieremethode 1 in der zweiten Phase der Kalibrierung zu zeigen.

Ergebnis: Es stellt sich eine mittlere Abweichung von 0.3 [Pix] bei einer Standardabweichung von 0.17 [Pix] ein.

Dabei wurden Abweichungen von mehr als 0.8 [Pix] erreicht, wobei die hier vorgestellte Methode 1 eine Standardabweichung von weniger als 0.6 [Pix] hat.

5.6.2. Online-Kalibrierung mit Bezug auf die Fluchtpunkte in beiden Bildern (Methode 2)

Aus der Literatur zur projektiven Geometrie ist bekannt, dass parallele Linien in einem Bild sich in einem Punkt, dem Fluchtpunkt, treffen. Durch den optischen Fluss, der auf

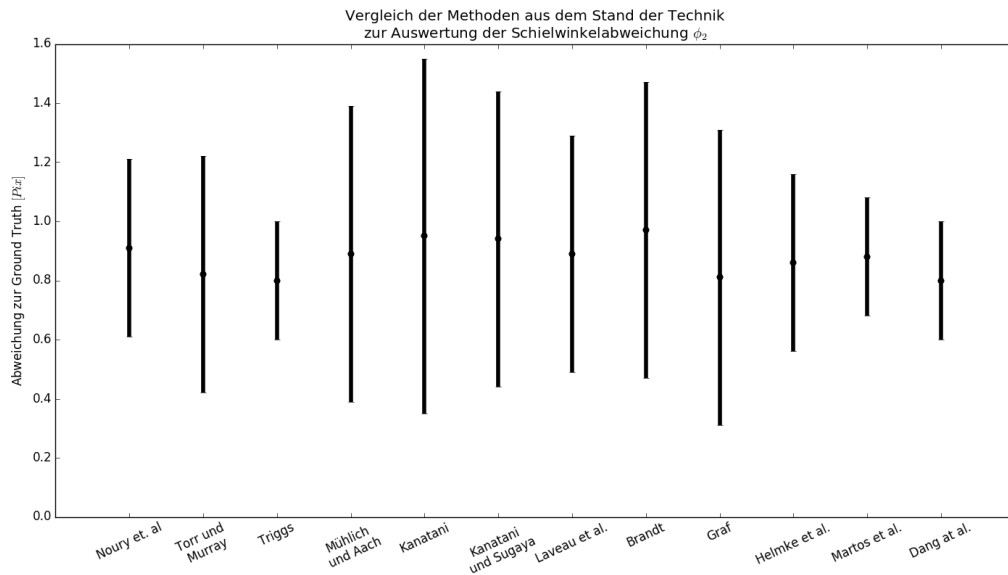


Abbildung 5.20.: Die Methoden aus dem Stand der Technik sind aufgezeichnet. Dabei ist zu sehen, dass die Abweichung zur Ground Truth im Mittel zwischen $0.8 [Pix]$ und $1.0 [Pix]$ liegt.

zwei zeitlich aufeinanderfolgenden Bildern angewendet wird, ergeben sich bei Geradeausfahrt die gesuchten Linien. In Abb. 5.21 ist dies zu sehen.

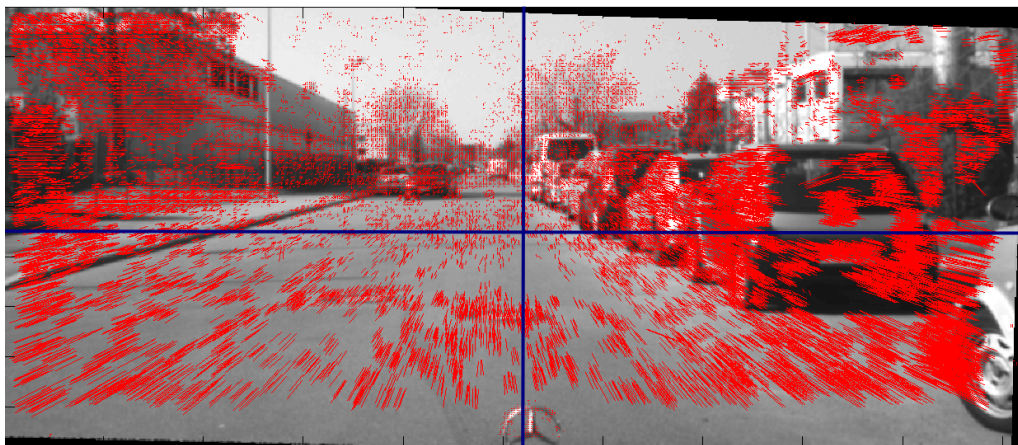


Abbildung 5.21.: Alle paarweisen disjunkten Flussvektoren aus zwei zeitlich aufeinanderfolgenden Bildern bei einer Geradeausfahrt haben einen Schnittpunkt im Bild. Dieser Punkt wird als Fluchtpunkt bezeichnet. Für eine Online-Kalibrierung eines Stereo-Kamera-Systems muss dieser Fluchtpunkt in beiden Bildern bestimmt werden.

Um den Fluchtpunkt aus der Menge aller Korrespondenzen in einem Bildpaar einer Kamera des Stereo-Kamera-Systems zu berechnen, wird in dieser Umsetzung die Gleichung Gl. (5.5) genutzt.

$$u_i = \frac{1}{O} \left| \begin{array}{cc|c} i u_\tau & i v_\tau & (i u_\tau - r u_\tau) \\ r u_\tau & r v_\tau & \\ \hline i u_{\tau-1} & i v_{\tau-1} & (i u_{\tau-1} - r u_{\tau-1}) \\ r u_{\tau-1} & r v_{\tau-1} & \end{array} \right| \quad (5.5)$$

$$= \frac{1}{O} [(i u_{\tau-1} - r u_{\tau-1})(i u_\tau v_\tau - r u_\tau i v_\tau) - (i u_\tau - r u_\tau)(i u_{\tau-1} v_{\tau-1} - r u_{\tau-1} i v_{\tau-1})] \quad (5.6)$$

$$\begin{aligned} \text{mit } O &= \begin{vmatrix} i u_\tau - r u_\tau & i v_\tau - r v_\tau \\ i v_{\tau-1} - r v_{\tau-1} & i u_{\tau-1} - r u_{\tau-1} \end{vmatrix} \\ &= (i u_\tau - r u_\tau) \cdot (i v_{\tau-1} - r v_{\tau-1}) \\ &\quad - (i v_\tau - r v_\tau) \cdot (i u_{\tau-1} - r u_{\tau-1}) \\ &\neq 0 \end{aligned}$$

mit $i(u, v)_\tau^T$ der Punkt $(u, v)^T$ im linken Bild zum Zeitpunkt τ , $r(u, v)_\tau^T$ der Punkt $(u, v)^T$ im rechten Bild zum Zeitpunkt τ und analog die Punkte im linken und rechten Bild zum Zeitpunkt $\tau-1$. Dabei ist $i(u, v)_\tau^T$ der Bildpunkt im linken Bild zum Zeitpunkt τ , $i(u, v)_{\tau-1}^T$ ist der korrespondierende Bildpunkt zum Zeitpunkt $\tau-1$, der räumlich korrespondierende Bildpunkt im rechten Bild zu diesem Bildpunkt ist $r(u, v)_{\tau-1}^T$ und der zeitlich korrespondierende Bildpunkt im rechten Bild ist $r(u, v)_\tau^T$. Sollte $O = 0$ gelten, dann sind die Vektoren parallel oder linear abhängig und so existiert kein Schnittpunkt in einem Bild oder unendlich viele Schnittpunkte.

Um nun in der Menge aller durch Gl. (5.5) berechneten Schnittpunkte den Fluchtpunkt zu berechnen, wird ein Histogramm genutzt. Dabei wird für die horizontale Koordinate des Bildes jedes Pixel als Bin im Histogramm eingeteilt, d.h. wenn das Bild 1024 [Pix] für die horizontale Koordinate hat, dann gibt es 1024 Bins für das Histogramm. Um die Schnittpunkte ins Histogramm einzutragen, werden die Schnittpunkte, die außerhalb der Bildfläche sind, vernachlässigt. Die restlichen Schnittpunkte werden in das Histogramm eingetragen. Nach wenigen Bildern ergibt sich das Histogramm aus Abb. 5.22.

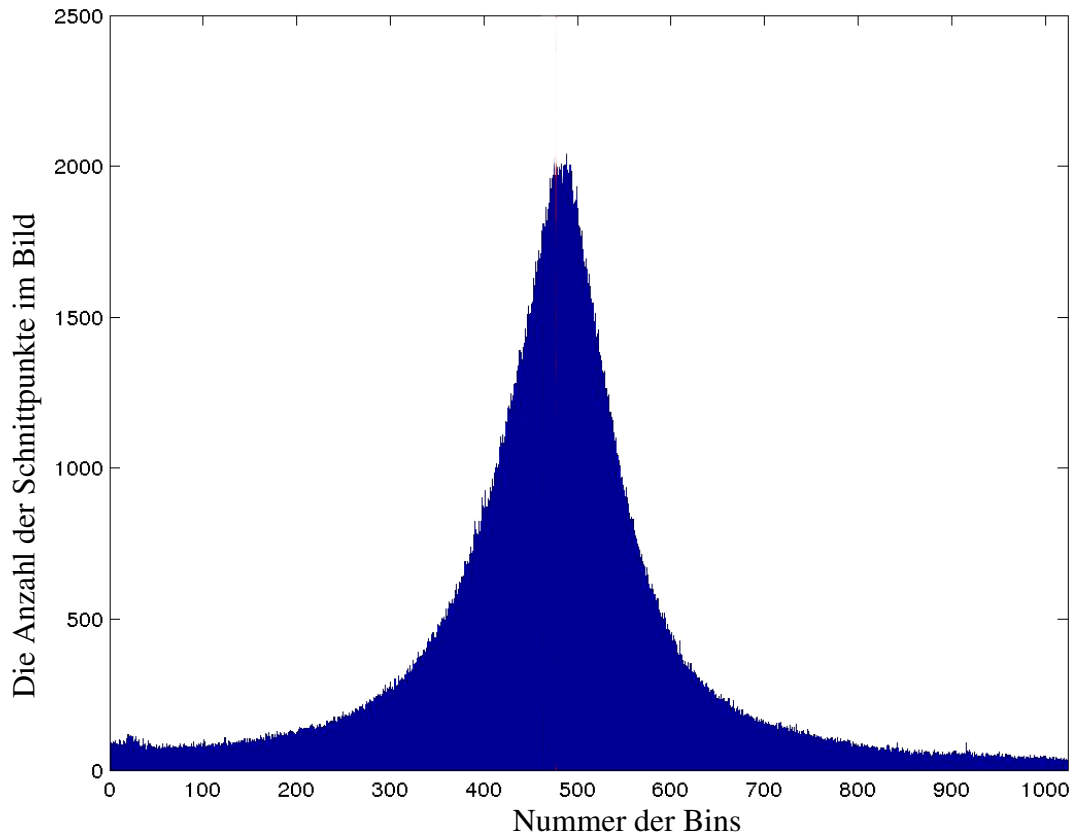


Abbildung 5.22.: **Methode 2:** Die Schnittpunkte der Flussvektoren werden in einem Histogramm eingetragen. Der Peak ist die horizontale Koordinate des Fluchtpunktes. Diese Auswertung läuft über viele Folgebilder, bis das Histogramm ein vorgegebenen Schwellenwert für die maximalen Elemente des höchsten Bins erreicht.

Das Maximum oder der Peak des Histogramms ist die horizontale Koordinate des Fluchtpunktes. Das Maximum kann durch das Fitten einer Laplace-Verteilung oder durch das Fitten einer Gauss-Verteilung berechnet werden. Die beiden Verteilungen sind sehr ähnlich und liefern in der Implementierung und auf den Testdaten gleiche Werte. Dabei wird hier aber die Menge von 2σ im Histogramm für die Gauss- oder Laplace-Verteilung genutzt, da sonst die Fehler durch Ausreißer einen großen Einfluss haben. Die Laplace-Verteilung wurde in der MatLab-Implementierung genutzt, in der C++-Implementierung wurde die Gauss-Verteilung eingesetzt. Die Laplace-Verteilung und die Gauss-Verteilung stehen durch die generalisierte Gauss-Verteilung in Beziehung. Um die Dekalibrierung zu bestimmen, werden die Fluchtpunkte in beiden Bildern der Stereo-Kamera bestimmt. Nachdem die Koordinate des Fluchtpunktes in beiden Bildern bestimmt ist, wird Gl. (4.19) genutzt, um die Größe ΔU des Schielwinkelfehlers zu ermitteln und somit die Dekalibrierung auszugleichen.

Das Stereo-Kamera-System von Abb. 5.1 wurde für die Evaluierung genutzt. Die in diesem Abschnitt beschriebene Implementierung der Methode zur Berechnung des Schielwinkels nutzt die erste, zweite, fünfte und siebte Sequenz aus Tab. 5.2. Da auch, wie bei

der Methode zuvor, der schnelle Wechsel zwischen Kurvenfahrten ein Problem erzeugt, wird auch hier darauf geachtet, dass die Gierrate durch einen Medianfilter von 0.1 Sekunden eine Geradeausfahrt zusichert.

In Abb. 5.23 ist die Auswertung einer kurzen Sequenz zu sehen. Es ist zu sehen, dass mit dieser Methode eine mittlere Abweichung ϕ_2 von 0.12 [Pix] mit einer Standardabweichung von 0.05 [Pix] erreicht wurde. Wird nun ein Punkt mit der exakten Entfernung von 102.40 [m] zum Stereo-Kamera-System mit dieser Abweichung berechnet, dann wird der Punkt in der Entfernung zwischen 96.60 [m] und 108.94 [m] rekonstruiert.

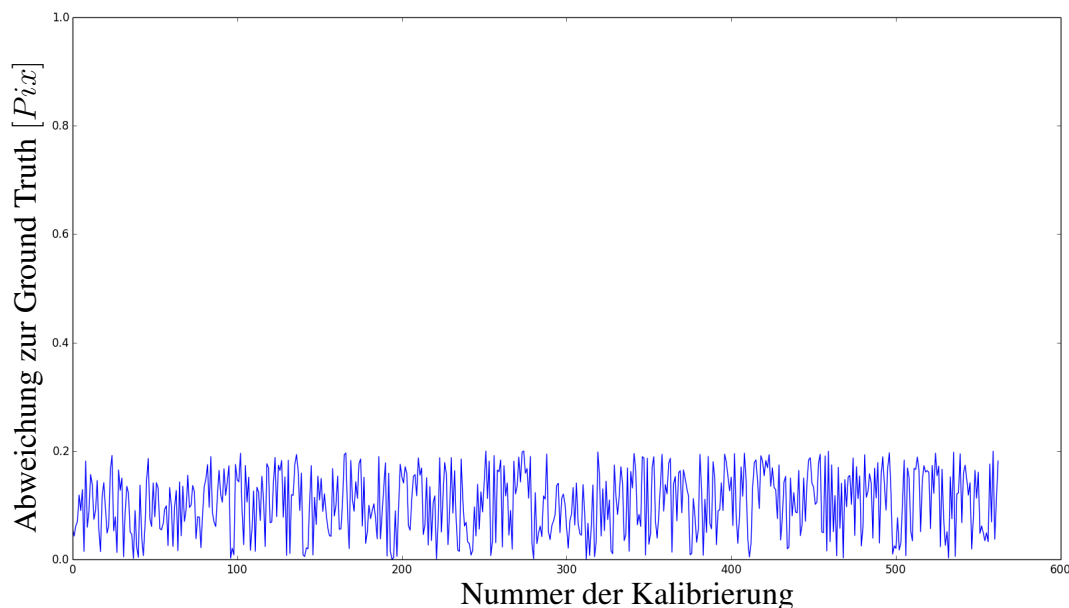


Abbildung 5.23.: **Methode 2:** Die Auswertung zeigt, dass sich die Kalibrierung des Schielwinkels gegenüber den in der Literatur genutzten Methoden deutlich verbessert. Die mittlere Abweichung ist 0.12 [Pix] über alle Sequenzen bei einer Standardabweichung von 0.05 [Pix].

Ergebnis: Über die Testsequenz stellt sich eine mittlere Abweichung von 0.12 [Pix] mit einer Standardabweichung von 0.05 [Pix] ein.

Die Methoden, die in [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] beschrieben sind, wurden nachimplementiert, um die Vergleichbarkeit der Methoden auf den gleichen Sequenzen darzustellen. Auf den gleichen Sequenzen haben diese Methoden eine Abweichung von mindestens 0.8 [Pix] wobei die hier beschriebene Methode mindestens viermal besser ist als die Methoden aus der oben genannten Literatur, da die Abweichung im Mittel 0.12 [Pix] groß ist und maximale Abweichung 0.2 [Pix] ist.

5.6.3. Zusammenfassung

Beide Methoden eignen sich zur Berechnung des Schielwinkels in einem Fahrzeug. Sie lassen sich aber nicht direkt vergleichen, da die Methode 1 zur Berechnung der Schielwinkelabweichung mit den Fahrzeugdaten sehr schnell ein Ergebnis berechnet, mit einem nur geringen Genauigkeitsvorteil zu den herkömmlichen Methoden. Die Methode 2, zur Berechnung auf Basis der Fluchtpunktschätzung, hat eine sehr hohe Genauigkeit, benötigt aber mehr Zeit für den Durchlauf einer Kalibrierung.

Die beiden Methoden ergänzen sich hervorragend. Die erste Methode liefert schnell eine gute Abschätzung, während die zweite Methode nach längerer Einschwingzeit ein hochgenaues Ergebnis liefert.

Die Methoden, die in [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] beschrieben sind, wurden im Rahmen dieser Arbeit nachimplementiert und in einer aufwendigen Auswertung gegen die Methoden, die in Abschn. 5.6.1 und in Abschn. 5.6.2 beschrieben sind, verglichen.

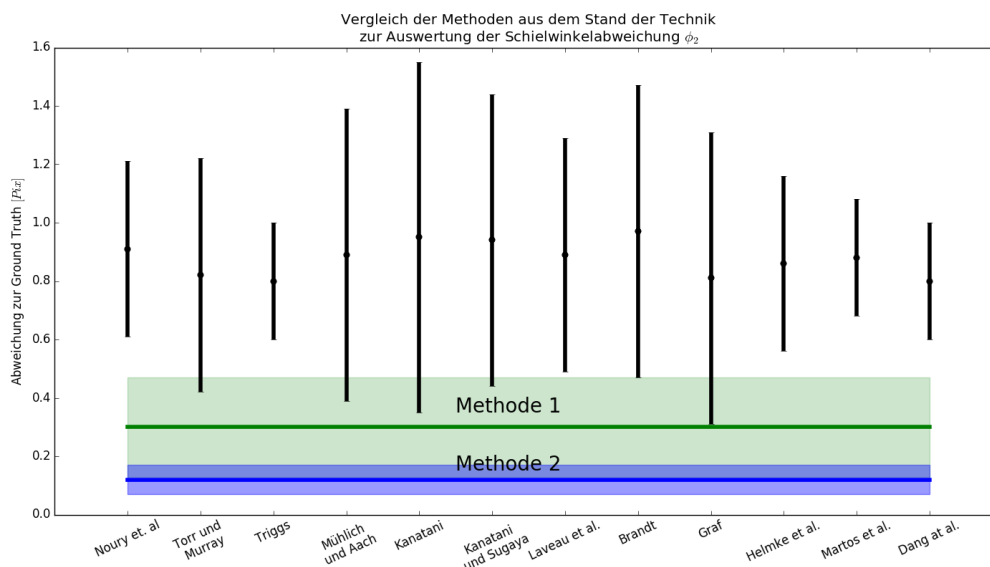


Abbildung 5.24.: Die Methoden aus dem Stand der Technik, die in [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000; Dang et al., 2009] beschrieben sind, wurden in einer aufwendigen Nachsimulation implementiert und auf den Sequenzen: 2,3,5 und 6 aus der Tab. 5.2 ausgewertet. Der Vergleich zu den in dieser Arbeit entwickelten Methode ist eingezeichnet. Dabei ist die Methode 1 grün eingezeichnet und die Methode 2 ist mit blau eingezeichnet.

Diese Methoden aus dem Stand der Technik wurden mit den Sequenzen: 2,3,5 und 6 aus der Tab. 5.2 ausgewertet. Die Auswertung der Schielwinkelabweichung und der Vergleich zu den Methoden, die zur Schielwinkelschätzung in dieser Arbeit beschrieben sind, ist in Abb. 5.24 gezeigt. Dabei ist zu sehen, dass die Methode 1 eine Abweichung zur Ground Truth von weniger als $0.6 [Pix]$ auf diesen Sequenzen hat und die Methoden aus dem Stand der Technik im Mittel nicht besser als $0.8 [Pix]$ gegenüber der Ground Truth sind.

6. Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wurden Methoden zur Ausreißerdetektion untersucht und verbessert. So wurde die klassische Methode aus der Photogrammetrie (*innere Genauigkeit*) genutzt, um Ausreißer in der Menge der Korrespondenzen zu bestimmen. Im Rahmen dieser Arbeit wurde für die Gleichung zur *inneren Genauigkeit* Gl. (2.3) das Gewicht $W_i = 1$ genutzt, um alle Korrespondenzen gleich zu gewichten. Das Gewicht kann in weiteren Arbeiten durch den EM-Algorithmus (Expectation-Maximization-Algorithmus), der auch in den Arbeiten von [Brandt und Heikkonen, 2000a] zur Ausreißerdetektion genutzt wird, geschätzt werden. So können Ausreißer in der Menge der Korrespondenzen besser bestimmt werden. Das kann zu besseren Kalibrierergebnissen führen.

Die Komponenten von Korrespondenzen geben eine Information über den Einfluss der Korrespondenz auf die Kalibrierung. Dies wurde in Kap. 4.3.1 hergeleitet und in Kap. 5.1 nachgewiesen. Es wurde gezeigt, dass, wenn die Rangordnung aus Gl. (4.27) bis Gl. (4.29) benutzt wird, die Menge der Korrespondenzen für jeden einzelnen Winkel optimiert werden und so die Kalibrierung schneller ermittelt werden kann, da die Korrespondenzen mit dem größten Einfluss für den gesuchten Parameter in der Menge aller Korrespondenzen gefunden werden und exklusiv genutzt werden können.

Die zwei Methoden zur Schätzung des Schielwinkels eignen sich beide zur Kalibrierung. Dabei ist die erste Methode in Kap. 4.1 und die zweite Methode in Kap. 4.2 hergeleitet. Die Methoden aus [Brandt, 2006; Graf, 2007; Helmke et al., 2007; Kanatani et al., 2000; Kanatani und Sugaya, 2009; Laveau et al., 1994; Mühlich und Aach, 2007; Martos et al., 2010; Noury et al., 2007; Torr und Murray, 1997; Triggs et al., 2000] wurden auf den gleichen Sequenzen mit den gleichen Stereo-Kamera-Systemen bewertet. Die besten Methoden aus der Literatur haben eine Abweichung von mindestens $0.8 [Pix]$ und die Methode 1 hat eine maximale Abweichung von $0.6 [Pix]$, was eine Verbesserung von $0.2 [Pix]$ bedeutet. Im Mittel ist die Methode 1 viermal besser als die in der Literatur beschriebenen Methoden.

Die zweite Methode wurde ebenfalls mit den Methoden aus der Literatur, die schon für die Auswertung der ersten Methode genutzt wurden, verglichen. Dabei wurden die selben Sequenzen für die Versuche genutzt. Dabei hat sich herausgestellt, dass die Methode 2 im Mittel achtmal genauer ist als die in der Literatur beschriebenen Methoden. Dabei hat

die Methode 2 eine maximale Abweichung von $0.2 [Pix]$ und im Mittel eine Abweichung von $0.12 [Pix]$. Die Abweichungen der Methoden aus der Literatur sind genau so wie in der Methode 1.

Dies zeigt, dass die hier entwickelten Methoden zur Kalibrierung, besonders die zwei Methoden zur Schätzung des Schielwinkels, sich sehr gut eignen, um ein Stereo-Kamera-System zu kalibrieren.

In dieser Arbeit wurde das erste Mal gezeigt, dass die Windschutzscheibe, die zwischen der Kamera und der Umwelt ist, einen entscheidenden Anteil an der Kalibrierung trägt. Selbst wenn das Kamera-System initial mit einem kleinen Fehler – in dieser Arbeit haben wir den RMSE zur Angabe des Fehlers genutzt – kalibriert wurde, ist durch die Verzeichnung eine Online-Kalibrierung sehr schwer und nicht optimal. In dieser Arbeit ist eine Methode dargestellt, die sowohl online als auch offline die Verzeichnung des Bildes durch die Windschutzscheibe bestimmen und rückrechnen kann. Dabei werden aber lokale Unregelmäßigkeiten, die sich wie Mikrolinsen auswirken, nicht betrachtet. Dieses kann zu lokalen Verzeichnungen im Bild führen und so kleine, aber durchaus erhebliche Verschlechterungen bei der Berechnung zur Folge haben. Da die Stereo-Kamera im Rahmen dieser Arbeit immer mit der Windschutzscheibe zusammen kalibriert wurde, sind die Bilder verzeichnungsfrei. Die Verzeichnung der Windschutzscheibe tritt auf, wenn das Stereo-Kamera-System zuerst kalibriert wird und dann ins Fahrzeug hinter die Windschutzscheibe integriert wird. Dieses Vorgehen wird im praktischen Einsatz durchgeführt, sodass die Untersuchungen hierfür erforderlich waren.

Es wurde gezeigt, dass sich ein Velodyne-Laserscanner und ein Stereo-Kamera-System so kalibrieren lassen, dass die Messungen in einem gemeinsamen Koordinatensystem erfolgen können. So kann der Velodyne-Laserscanner genutzt werden, um die Kalibrierung des Stereo-Kamera-Systems in absoluten Entfernungsdaten zu überprüfen. Dabei werden die 3D-Rekonstruktionen aus dem Velodyne-Laserscanner und dem Stereo-Kamera-System verglichen. Die Daten aus dem Velodyne-Laserscanner werden mit der Eigenbewegung des Fahrzeugs mitberücksichtigt, sodass eine 3D-Szene zum Zeitpunkt der Bildaufnahme zum Abgleich genutzt werden kann. Da die Daten aus dem Velodyne-Laserscanner immer als ein Block kommen, sind die Datenpakete verschoben, wenn sich das Fahrzeug bewegt. Wenn zu dem Datenpaket die Fahrzeug-Odometrie gespeichert wird, dann kann eine vollständige 3D-Rekonstruktion berechnet werden.

Im Rahmen dieser Arbeit wurden zwei Methoden zur Schätzung des Schielwinkels entwickelt. Methode 1 nutzt Methoden, die auf der gegebenen Hardware und dem Framework schon implementiert und integriert sind.

Es ist sinnvoll, in weiterführenden Arbeiten zu zeigen, wie die intrinsischen Parameter sich bei einer Dekalibrierung verhalten, um aufzuzeigen, ob die Hinzunahme neuer Parameter die Ergebnisse verbessern kann.

A. Anhang

A.1. Die erforderliche Rechenleistung auf dem DSP zur Online-Kalibrierung

Die Online-Kalibrierung der Stereo-Kamera wird direkt auf dem Digitalen Signalprozessor (DSP) im Stereo-Kamera-System vorgenommen. Die Rechenzeit, die auf dem DSP zur Verfügung steht, ist äquivalent zu einem DSP mit einer Taktfrequenz von 15 MHz. Um sicherzustellen, dass die Stereo-Kamera kalibriert ist, muss eine Kalibrierung in einer Zeitspanne von maximal einer Minute bereit stehen. Dies ist eine Forderung an das System, da der thermische Prozess die Kalibrierung der Kamera stark verändert. Da die Größe des Speichers auf der Hardware bei der Entwicklung der Stereo-Kamera festgelegt werden muss, wurde der Speicher für die Online-Kalibrierung auf ein Mindestmaß, in diesem Fall 2 kB, beschränkt. Daten, die den anderen Programmen zur Verfügung stehen, können mitgenutzt werden.

A.2. Herleitung des Prädiktionsintervalls

Das Konfidenzintervall hängt stark mit dem Prädiktionsintervall zusammen. Um die Zuverlässigkeit einer Aussage zu bewerten, wird das Konfidenzintervall genutzt. Um eine Wahrscheinlichkeit über die Aussage eines zukünftigen Ereignisses zu treffen, werden Prädiktionsintervalle genutzt. Es gilt, dass das Prädiktionsintervall bei gleichem Perzentil größer ist als das Konfidenzintervall [Kardaun, 2005, S. 196].

Sei eine Student-T-Verteilung gegeben. Dabei seien x_1, \dots, x_n Ereignisse mit einem Erwartungswert μ . Der Mittelwert ist gegeben durch

$$\bar{x} = \frac{\sum_i^n x_i}{n} \quad (\text{A.1})$$

und der Varianz

$$\sigma^2 = \frac{\sum_i^n (x_i - \bar{x})^2}{n - 1}, \quad (\text{A.2})$$

dann ergibt sich der Wert des betrachteten Quantils, hier der Median, der auch "t-Wert" genannt wird, aus

$$t = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}. \quad (\text{A.3})$$

Seien nun der Mittelwert und die Varianz einer gegebenen Normalverteilung $\mathcal{N}(\mu, \sigma^2)$ unbekannt. Dann ergibt sich aus der Student-T-Verteilung

$$t^{n-1} \sim \frac{x_{n+1} - \bar{x}_n}{\sigma \sqrt{1 + \frac{1}{n}}}. \quad (\text{A.4})$$

für den ‘t-Wert’. Wird nun Gl. (A.4) nach x_{n+1} aufgelöst, dann folgt

$$x^{n+1} = \bar{x}_n + \sigma \sqrt{1 + \frac{1}{n}} \cdot t^{n-1}. \quad (\text{A.5})$$

Dann ergibt sich die Wahrscheinlichkeit

$$Pr \left(\bar{x}_n - t_a \sigma \sqrt{1 + \frac{1}{n}} \leq x^{n+1} \leq \bar{x}_n + t_a \sigma \sqrt{1 + \frac{1}{n}} \right) = p, \quad (\text{A.6})$$

mit t_a das $100(\frac{1+p}{2})$ -te Perzentil der Student-T-Verteilung mit $n - 1$ Freiheitsgraden.

A.3. Kalibrierung des Velodyne-Laserscanners und des Stereo-Kamera-Systems

Der Sensor dient dazu, die 3D-Rekonstruktionen der Fahrzeugumgebung des Velodyne-Laserscanners und des Stereo-Kamera-Systems zu vergleichen. Dazu werden die Koordinatenursprünge beider Sensoren auf ein einheitliches Koordinatensystem umgerechnet. Um diese Berechnung zu vereinfachen, werden die Sensorsysteme auf einen gemeinsamen Ursprung kalibriert.

Dabei wird eine Kalibrierung des Stereo-Kamera-Systems und des Velodyne-Laserscanners auf eine gemeinsame Basis kalibriert, um die 3D-Rekonstruktionen zu vergleichen. Dazu wird das Target mit der bekannten Geometrie aus der Initialkalibrierung durch ein für den Velodyne-Laserscanner erkennbares Target erweitert. Dann ist dieses Target in der Velodyne-Laserscanner-3D-Rekonstruktion und der 3D-Rekonstruktion aus der Stereo-Kamera zu detektieren. Im Rahmen dieser Arbeit wurde ein rundes Objekt, wie in Abb. A.1, gewählt.



Abbildung A.1.: Das Kamera-Target ist so angebracht, dass die Mitte des Schachbrettes mit der Mitte des kreisförmigen Targets übereinstimmt. Um dann das Koordinatensystem des Stereo-Kamera-Systems und das Koordinatensystem des Velodyne-Laserscanners abzustimmen, wird der Mittelpunkt des runden Targets in der Rekonstruktion des Velodyne-Laserscanners bestimmt.

Dabei wird die Mitte des Targets in die Mitte des runden Objektes positioniert. Nachdem die Stereo-Kamera initial kalibriert ist, kann eine weitere Kalibrierung vorgenommen werden, sodass die Orientierung und Positionierung des Velodyne-Laserscanners bzgl. des Stereo-Kamera-Systems bestimmt werden kann. Dazu wird in der 3D-Rekonstruktion des Velodyne-Laserscanners (siehe Abb. A.2) der Mittelpunkt des runden Objektes durch die Funktion Gl. (A.7) bestimmt. Um die Koordinaten des Kreismittelpunkts $(a, b, c)^T$ zu finden, werden drei Punkte auf dem Rand des Kreises im 2D-Bild $P_{1..3} = (x, y)_{1..3}^T$ ausgewählt.

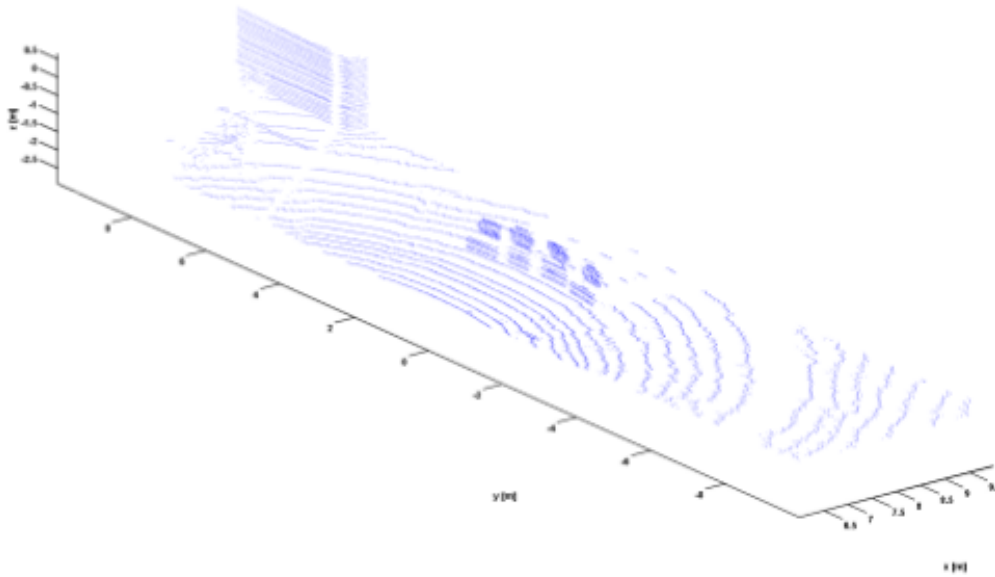


Abbildung A.2.: Der Velodyne-Laserscanner liefert eine 3D-Punkt wolke der Umgebung um das Fahrzeug.

$$\begin{pmatrix} a & b & c \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 1 \\ -x_1 & -x_2 & -x_3 \\ -y_1 & -y_2 & -y_3 \end{pmatrix} = \begin{pmatrix} a & -bx_1 & -cy_1 \\ a & -bx_2 & -cy_2 \\ a & -bx_3 & -cy_3 \end{pmatrix} = - \begin{pmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ x_3^2 + y_3^2 \end{pmatrix} \quad (\text{A.7})$$

mit $P_n(x_n, y_n)_{1, \dots, 3}^T$ den Punkten auf dem Rand des Kreises im Bild, dann ist $\hat{P} = (a, b, c)^T$ der Mittelpunkt des Kreises.

Dieser Punkt stimmt mit dem Mittelpunkt des Targets überein und kann mit dem Bündelausgleich in ein Basissystem umgerechnet werden.

A.4. Differentialgeometrischer Ansatz für die Scheibenverzeichnung

Die nun vorgestellte Methode zur Bestimmung der Verzeichnung ist universell einsetzbar, da zur Berechnung der Verzeichnung der Punkt auf der Windschutzscheibe und die optischen Achsen bekannt sein müssen. Bei der in Kap. 4.3.4 vorgestellten Methode mussten eine Vielzahl von Messungen mit verschiedenen Windschutzscheiben und verschiedenen Stereo-Kamera-Systemen vorgenommen werden. An dem Punkt, an dem die optische Achse die Windschutzscheibe schneidet, wird die Bildebene projiziert. Durch die Krümmung der Windschutzscheibe entlang der Hauptkrümmungsrichtungen kann dann die Verzeichnung abgelesen werden, sodass die Verzeichnung in jedem Punkt des Bildes zurückgerechnet werden kann und so die Rektifizierung wieder hergestellt wird. Ist die

projizierte Bildebene im Schnittpunkt von optischer Achse und Windschutzscheibe, dann ist die Verzeichnung die Projektion der Distanz jedes Bildpunktes zwischen einem Punkt auf dem Bild und dem korrespondierenden Punkt auf der Windschutzscheibe.

Die Oberfläche der Windschutzscheibe wird approximiert als Ellipsoid. Ein solches Ellipsoid ist Abb. A.3 beispielhaft dargestellt

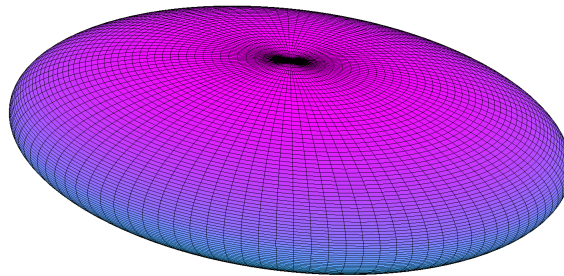


Abbildung A.3.: Die Windschutzscheibe ist ein kleiner Teil eines Paraboloiden. Hier ist ein Paraboloid dargestellt. Ein Punkt auf der Oberfläche kann durch die Methoden der Differentialgeometrie genau bestimmt werden und ausgehend von diesem Punkt kann die Krümmung des Paraboloiden bestimmt werden.

Um einen Punkt auf der Oberfläche zu bestimmen, muss die Funktion, die den Paraboloid beschreibt, in der Parameterdarstellung beschrieben werden [do Carmo, 1976]. Die Parameterdarstellung eines Paraboloides ist durch

$$\Omega(u, v) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \cdot \sin(u) \cdot \cos(v) \\ b \cdot \sin(u) \cdot \sin(v) \\ c \cdot \cos(u) \end{pmatrix} \text{ mit}$$

$$D = \left\{ \begin{pmatrix} k \\ l \end{pmatrix} : -2\pi \leq u \leq 2\pi, -2\pi \leq v \leq 2\pi \right\} \quad (\text{A.8})$$

gegeben. Mit Gl. (A.8) kann dann die Tangentialfläche bestimmt werden. Die Basisvektoren der Tangentialfläche an einem beliebigen Punkt auf der Oberfläche des Paraboloiden ist dann

$$\begin{aligned}\Omega_u(u, v) &\equiv \frac{\partial \Omega(u, v)}{\partial u} = \begin{pmatrix} a \cdot \cos(u) \cdot \cos(v) \\ b \cdot \cos(u) \cdot \sin(v) \\ -c \cdot \sin(u) \end{pmatrix} \\ \Omega_v(u, v) &\equiv \frac{\partial \Omega(u, v)}{\partial v} = \begin{pmatrix} -a \cdot \sin(u) \cdot \sin(v) \\ b \cdot \sin(u) \cdot \cos(v) \\ 0 \end{pmatrix}\end{aligned}\quad (\text{A.9})$$

wobei Ω ein Zeichen für eine Fläche darstellt. Im Rahmen dieser Arbeit ist mit Ω die gesamte Windschutzscheibe beschrieben. $\Omega(u, v)$ ist dann der Punkt an der 2D-Position (u, v) auf der Windschutzscheibe. Mit der ersten Fundamentalform der Differentialgeometrie, wie in [do Carmo, 1976] beschrieben, kann der Inhalt der Fläche und die Länge einer jeden gegebenen Bahn auf dem Flächensegment berechnet werden. Die erste Fundamentalform kann mit der folgenden Matrixdarstellung beschrieben werden [do Carmo, 1976], wobei die Blockmatrizen E, F, G eingefügt werden. Jeder dieser Matrizen ist die Ableitung in eine Richtung auf der Fläche im Punkt (u, v) :

$$(g_{ij}) = \begin{pmatrix} E & F \\ F & G \end{pmatrix}\quad (\text{A.10})$$

mit

$$E = \Omega_v(u, v) \cdot \Omega_v(u, v) = a^2(\cos(u) \cdot \cos(v))^2 + b^2(\cos(u) \cdot \sin(v))^2 + c^2(\sin(v))^2\quad (\text{A.11})$$

$$F = \Omega_u(u, v) \cdot \Omega_v(u, v) = b^2((\cos(u))^2 \cdot \sin^2(v)) - a^2((\cos(v))^2 \cdot \cos(u) \cdot \sin(u))\quad (\text{A.12})$$

und

$$G = \Omega_u(u, v) \cdot \Omega_u(u, v).\quad (\text{A.13})$$

Somit ist die erste Fundamentalform

$$ds^2 = E dv^2 - 2F dudv + G du^2.\quad (\text{A.14})$$

Ein Wegstück eines beliebigen Weges auf der Windschutzscheibe ist dann

$$ds = \sqrt{ds^2}\quad (\text{A.15})$$

und die Länge eines beliebigen Weges ist dann

$$l = \int_{\phi} ds, \quad (\text{A.16})$$

wobei ϕ der Weg ist. Es ist leicht zu sehen, dass die erste Fundamentalform eine quadratische positiv definite Bilinearform ist [do Carmo, 1976; Lang, 1999; Struik, 1961]. Mit der ersten Fundamentalform lässt sich jetzt die innere Geometrie einer Fläche beschreiben. Die innere Geometrie einer Fläche beschreibt Länge und Flächeninhalte einer gegebenen Fläche $\Omega' \subseteq \Omega$.

Mit der zweiten Fundamentalform der Differentialgeometrie wird die Orientierung der Fläche im umgebenden Raum beschrieben. Sie wird hier außerdem genutzt, um die Krümmung der Windschutzscheibe zu bestimmen. Die zweite Fundamentalform hat die Matrixdarstellung

$$(h_{ij}) = \begin{pmatrix} L & M \\ M & N \end{pmatrix}, \quad (\text{A.17})$$

mit L, M, N die Blockmatrizen der zweiten Ableitungen der Parameterdarstellung. L, M, N werden im Folgenden bestimmt. Die Koeffizienten von (h_{ij}) sind abhängig von der normierten äußeren Normale am Punkt (u, v) auf dem Paraboloiden und sind definiert als

$$\mathcal{N}(u, v) = \frac{\Omega_u(u, v) \times \Omega_v(u, v)}{\|\Omega_u(u, v) \times \Omega_v(u, v)\|}. \quad (\text{A.18})$$

Eine Veranschaulichung bietet Abb. A.4.

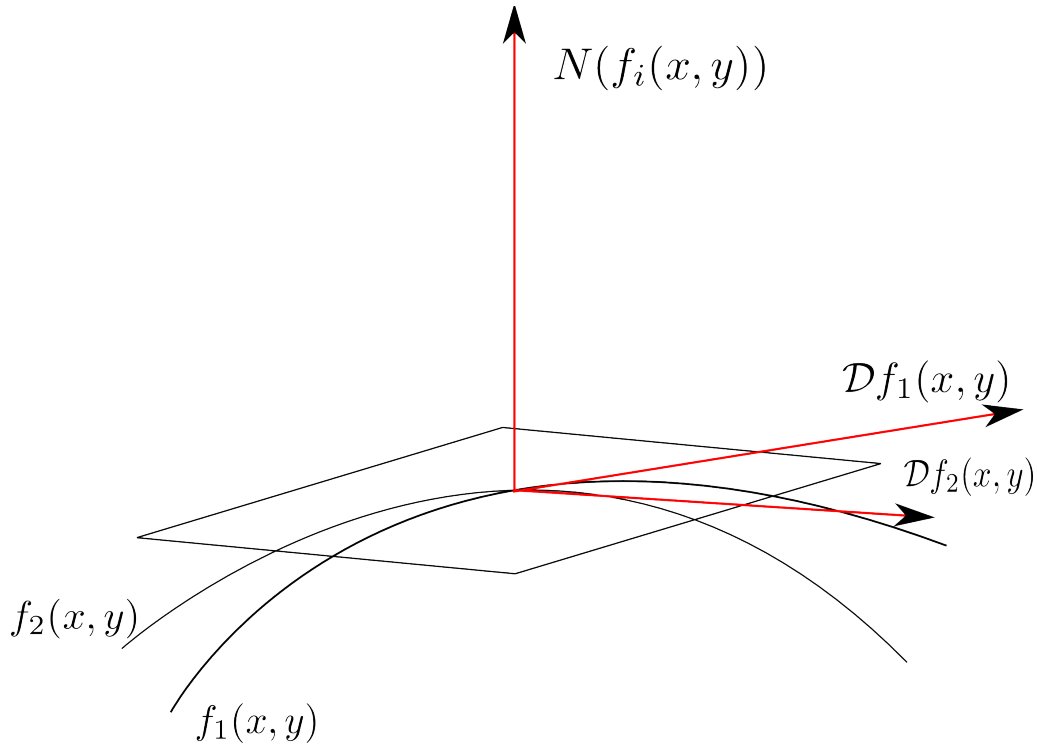


Abbildung A.4.: Die Normale auf einem Segment eines Paraboloiden mit gegebener Tangentialfläche am Punkt (x, y) .

Mit Gl. (A.18) können die Parameter von Gl. (A.17) bestimmt werden. Wenn nun

$$\begin{pmatrix} L \\ M \\ N \end{pmatrix} = \mathcal{N}(u, v) \cdot \begin{pmatrix} \Omega_{uu}(u, v) \\ \Omega_{uv}(u, v) \\ \Omega_{vv}(u, v) \end{pmatrix}, \quad (\text{A.19})$$

dann ist

$$d\sigma^2 = Ldv^2 + 2Mdudv + Ndu^2 \quad (\text{A.20})$$

eine Differentialform, mit der man leicht sehen kann, dass durch die zweite Fundamentform eine quadratische Form ist [do Carmo, 1976; Lang, 1999; Struik, 1961]. Die Determinante $(h_{ij}) = LN - M^2$ gibt Auskunft über die Krümmung der Fläche. Da wir hier ein Paraboloid betrachten, folgt, dass die Determinante immer größer als null ist. Um nun die Krümmung der Windschutzscheibe angeben zu können, betrachten wir die Weingarten-Abbildung [do Carmo, 1976; do Carmo und Flaherty, 1992; Dubrovin et al., 1985; Willmore, 1997]. Zuvor legen wir die Schreibweise $(g^{ij}) \equiv (g_{ij})^{-1}$ fest. Dann ist die Weingarten-Abbildung

$$\mathbb{W}(\Omega_u(u, v), \Omega_v(u, v)) = \begin{pmatrix} h_1^1(u, v) & h_2^1(u, v) \\ h_1^2(u, v) & h_2^2(u, v) \end{pmatrix} \quad (\text{A.21})$$

mit

$$\begin{aligned} h_{iu} &= g_{ij} h_u^j \\ h_u^i &= g^{ij} h_u^j. \end{aligned} \quad (\text{A.22})$$

Nun sind die zwei Hauptkrümmungsrichtungen κ_1, κ_2 die Eigenwerte der Weingartenabbildung. Die Hauptkrümmungsrichtungen sind die Basisvektoren, aus denen sich durch Linearkombinationen alle Krümmungen der Fläche berechnen lassen. Die beiden Basisvektoren sind orthogonal zueinander [do Carmo, 1976]. Zur Veranschaulichung der Hauptkrümmungsrichtungen dient Abb. A.5.

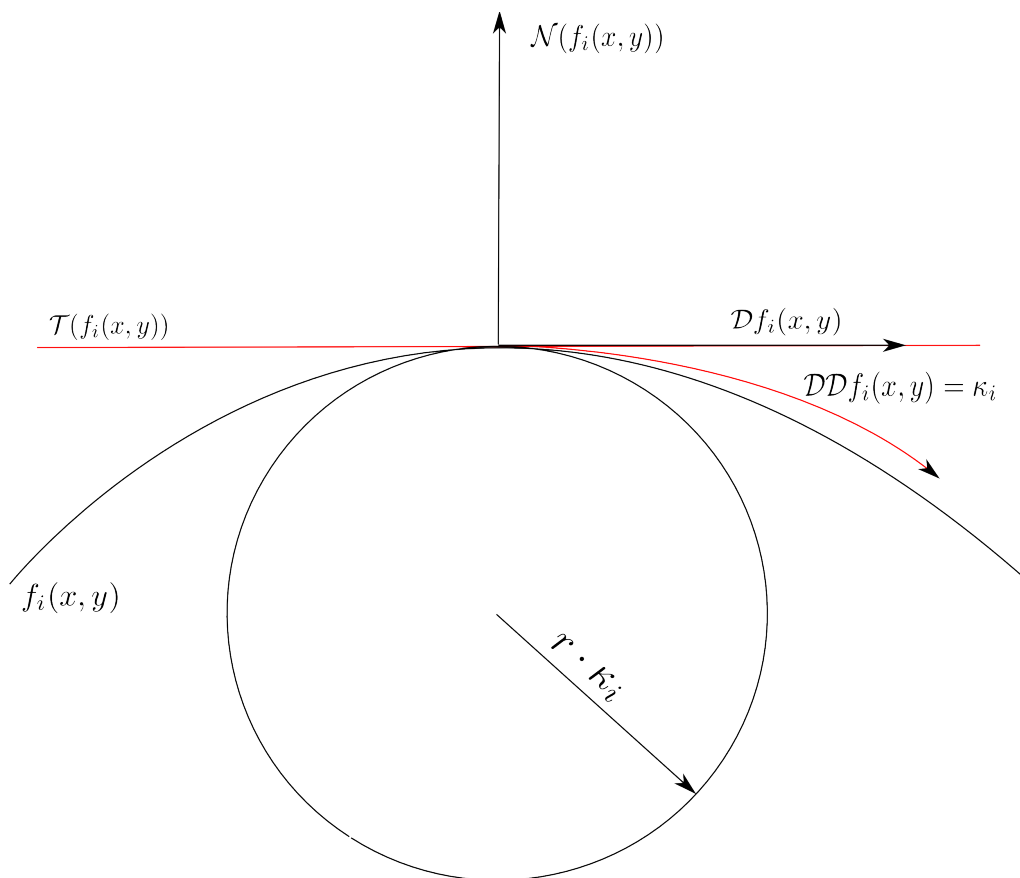


Abbildung A.5.: Die Hauptkrümmungsrichtungen sind die „Basisvektoren“, aus denen sich jede Krümmung auf dem Paraboloid berechnen lässt.

Um die Verzeichnung durch die Windschutzscheibenkrümmung bestimmen zu können, werden nun die Wege entlang der Hauptkrümmungsrichtungen berechnet. Der Punkt auf der Windschutzscheibe wird identifiziert durch die optische Achse, die ausgehend vom Bildhauptpunkt durch die Linse die Windschutzscheibe schneidet. Dies ist in Abb. A.6 dargestellt.

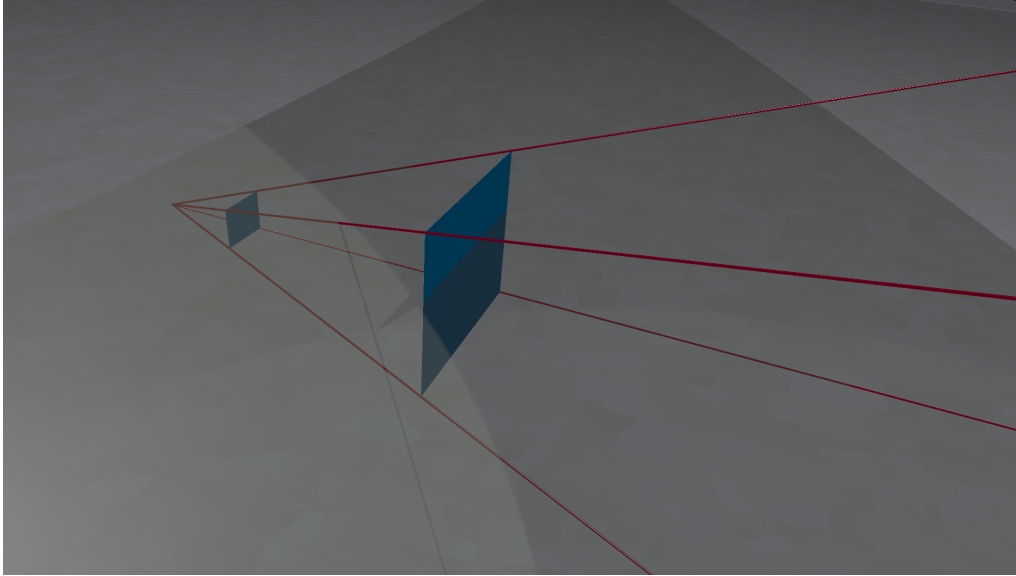


Abbildung A.6.: Die Bildfläche am Schnittpunkt der optischen Achse mit der Windschutzscheibe.

Ist nun eine Fläche entlang der Hauptkrümmungsrichtung berechnet worden, die die Windschutzscheibe in dem durch die optische Achse identifizierten Punkt berührt und die parallel zur Bildfläche ausgerichtet ist, dann ist die Verzeichnung gleich der Projektion zwischen der Fläche und der Windschutzscheibe.

Nun ändert sich dadurch die projektive Abbildung in drei Teile

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{cases} \hat{X} \in \Omega(u, v) = [R|t]X_W & \text{falls } X \notin \Omega(u, v) \\ (u', v')^T \in \Omega_{u,v}(u, v) = \min(\mathcal{K}\hat{X}), & \text{falls } X \in \Omega(u, v) \\ (u, v)^T \in \mathcal{I} = [R \in \mathbb{R}^2|t](u', v') & \text{falls } (u', v')^T \in \Omega_{u,v}(u, v) \end{cases} \quad (\text{A.23})$$

Dabei wird der 3D-Weltpunkt auf die Windschutzscheibe transformiert. Im nächsten Schritt wird der Punkt von der Windschutzscheibe auf die Tangentialfläche projiziert und abschließend wird dieser Punkt von der Tangentialfläche auf den Imager transformiert.

Dieser theoretische Ansatz ist ein Grundstein weiterer wissenschaftlicher Untersuchungen.

Literaturverzeichnis

- Aravkin, A., Styer, M., Moratto, Z., Nefian, A., und Broxton, M. (2012). Student's t robust bundle adjustment algorithm. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 1757–1760.
- Åström, K., Cipolla, R., und Giblin, P. (1996). Generalised epipolar constraints. In Buxton, B. und Cipolla, R., editors, *Computer Vision - ECCV 1996*, volume 1065 of *Lecture Notes in Computer Science*, pages 95–108. Springer Berlin Heidelberg.
- Banz, C., Blume, H., und Pirsch, P. (2011). Real-time semi-global matching disparity estimation on the gpu. In *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*, pages 514–521.
- Barreto, J. P. und Daniilidis, K. (2005). Fundamental Matrix for Cameras with Radial Distortion. In *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 - Volume 01*, ICCV '05, pages 625–632, Washington, DC, USA. IEEE Computer Society.
- Bastian, J. W. und van den Hengel, A. (2003). Computing Image-Based Reprojection Error on Graphics Hardware. In *DICTA*, pages 663–672.
- Beauchemin, S. S. und Barron, J. L. (1995). The computation of optical flow. *ACM Comput. Surv.*, 27(3):433–466, doi:10.1145/212094.212141.
- Birchfield, S. (1997). Derivation of Kanade-Lucas-Tomasi Tracking Equation.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Bouguet, J. und Perona, P. (1998). Camera Calibration from Points and Lines in Dual-Space Geometry. Technical report, California Institute of Technology.
- Bouguet, J. und Perona, P. (1999). 3d photography using shadows in dual-space geometry. *International Journal of Computer Vision*, 35:129–149.
- Brandt, S. und Heikkonen, J. (2000a). A Bayesian weighting principle for the fundamental matrix estimation. *Pattern Recognition Letters*, 21(12):1081 – 1092, doi:10.1016/S0167-8655(00)00071-4.

- Dubrovín, B., Fomenko, A., und Novikov, S. (1985). *Modern Geometry. Methods and Applications: Part 2: The Geometry and Topology of Manifolds*. Graduate Texts in Mathematics. Springer-Verlag.
- Engels, C., Stewénius, H., und Nistér, D. (2006). Bundle Adjustment Rules. In *Photogrammetric Computer Vision (PCV)*. ISPRS.
- Ernst, I. und Hirschmüller, H. (2008). Mutual Information Based Semi-Global Stereo Matching on the GPU. In *Proceedings of the 4th International Symposium on Advances in Visual Computing, ISVC '08*, pages 228–239, Berlin, Heidelberg. Springer-Verlag.
- Eudes, A. und Lhuillier, M. (2009). Error propagations for local bundle adjustment. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2411–2418.
- Eudes, A., Lhuillier, M., Naudet-Collette, S., und Dhome, M. (2010). Fast Odometry Integration in Local Bundle Adjustment-Based Visual SLAM. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 290–293.
- Faugeras, O. (1993). *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, Cambridge, MA, USA.
- Feldmann, A., Krüger, L., und Kummert, F. (2011). An Evaluation on Estimators for Stochastic and Heuristic Based Cost Functions Using the Epipolar-Constraint. In Gagalowicz, A. und Philips, W., editors, *MIRAGE*, volume 6930 of *Lecture Notes in Computer Science*, pages 40–50. Springer.
- Fischler, M. A. und Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, 24(6):381–395, doi:10.1145/358669.358692.
- Förstner, W. (2000). On Weighting and Choosing Constraints for Optimally Reconstructing the Geometry of Image Triplets. In *Proceedings of the 6th European Conference on Computer Vision-Part II, ECCV '00*, pages 669–684, London, UK. Springer-Verlag.
- Franke, U., Rabe, C., und Gehrig, S. K. (2007). Kollisionsvermeidung durch raumzeitliche Bildanalyse (Collision Avoidance based on Space-Time Image Analysis). *IT - Information Technology*, 49(1):25–32.
- Fusiello, A., Trucco, E., und Verri, A. (2000). A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12:16–22. 10.1007/s001380050120.
- Gargallo, P., Prados, E., und Sturm, P. (2007). Minimizing the Reprojection Error in Surface Reconstruction from Images. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8.
-

- Gehrig, S. und Rabe, C. (2010). Real-time Semi-Global Matching on the CPU. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*, pages 85–92.
- Graf, S. (2007). *Kamerakalibrierung mit radialer Verzeichnung - die radiale essentielle Matrix*. PhD thesis, Universität Passau, Innstrasse 29, 94032 Passau.
- Haller, I. und Nedeveschi, S. (2010). GPU optimization of the SGM stereo algorithm. In *Intelligent Computer Communication and Processing (ICCP), 2010 IEEE International Conference on*, pages 197–202.
- Hartley, R. (1997a). In defense of the eight-point algorithm. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(6):580–593, doi:10.1109/34.601246.
- Hartley, R. und Schaffalitzky, F. (2009). Reconstruction from Projections Using Grassmann Tensors. *International Journal of Computer Vision*, 83(3):274–293, doi:10.1007/s11263-009-0225-1.
- Hartley, R. I. (1997b). Lines and Points in Three Views and the Trifocal Tensor. *Int. J. Comput. Vision*, 22:125–140, doi:http://dx.doi.org/10.1023/A:1007936012022.
- Hartley, R. I. (1999). Theory and Practice of Projective Rectification. *Int. J. Comput. Vision*, 35(2):115–127, doi:10.1023/A:1008115206617.
- Hartley, R. I. und Sturm, P. F. (1997). Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157.
- Hartley, R. I. und Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition.
- Heikkila, J. und Silven, O. (1997). A Four-step Camera Calibration Procedure with Implicit Image Correction. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), CVPR '97*, pages 1106–, Washington, DC, USA. IEEE Computer Society.
- Helmke, U., Hüper, K., Lee, P. Y., und Moore, J. B. (2007). Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold. *International Journal of Computer Vision*, 74(2):117–136.
- Hermann, S., Klette, R., und Destefanis, E. (2008). Inclusion of a Second-Order Prior into Semi-Global Matching. In *Proceedings of the 3rd Pacific Rim Symposium on Advances in Image and Video Technology, PSIVT '09*, pages 633–644, Berlin, Heidelberg. Springer-Verlag.
-

- Hirschmüller, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 807–814 vol. 2.
- Hirschmüller, H. (2011). Semi-Global Matching-Motivation, Developments and Applications. *Invited Paper at the Photogrammetric Week*.
- Hirschmüller, H. (2008). Stereo Processing by Semiglobal Matching and Mutual Information. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(2):328–341, doi:10.1109/TPAMI.2007.1166.
- Ho, H. T. und Goecke, R. (2008). Optical flow estimation using fourier mellin transform. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8.
- Hung, Y. S. und Tang, W. K. (2006). Projective Reconstruction from Multiple Views with Minimization of 2D Reprojection Error. *Int. J. Comput. Vision*, 66(3):305–317, doi:10.1007/s11263-005-3675-0.
- Jähne, B. (2005). *Digitale Bildverarbeitung*. Springer.
- Jeong, Y., Nister, D., Steedly, D., Szeliski, R., und Kweon, I.-S. (2010). Pushing the envelope of modern methods for bundle adjustment. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1474–1481.
- Jordan, M. (2012). Die neue S-Klasse kommt mit neuen Sensoren - DISTRONIC PLUS nun mit Lenkassistent. <http://blog.mercedes-benz-passion.com/2012/11/die-neue-s-klasse-kommt-mit-neuen-sensoren/>. [Online; accessed 19-March-2013].
- Kanatani, K. (2005). *Statistical Optimization For Geometric Computation: Theory And Practice*. Dover Books on Mathematics Series. Dover Publ.
- Kanatani, K. und Niitsuma, H. (2011). Optimal Two-View Planar Scene Triangulation. In Kimmel, R., Klette, R., und Sugimoto, A., editors, *Computer Vision - ACCV 2010*, volume 6493 of *Lecture Notes in Computer Science*, pages 242–253. Springer Berlin Heidelberg.
- Kanatani, K., Shimizu, Y., Ohta, N., Brooks, M. J., Chojnacki, W., und van den Hengel, A. (2000). Fundamental matrix from optical flow: optimal computation and reliability evaluation. *Journal of Electronic Imaging*, 9(2):194–202, doi:10.1117/1.482739.
- Kanatani, K. und Sugaya, Y. (2009). Compact Fundamental Matrix Computation. In Wada, T., Huang, F., und Lin, S., editors, *Advances in Image and Video Technology*,
-

- volume 5414 of *Lecture Notes in Computer Science*, pages 179–190. Springer Berlin Heidelberg.
- Kanatani, K. und Sugaya, Y. (2011). Implementation and Evaluation of Bundle Adjustment for 3-D Reconstruction. In *Proceedings of the 17th Symposium on Sensing via Imaging Information (SSII 2011)*, pages IS4–02–1 – IS4–02–8, Yokohama, Japan.
- Kardaun, O. J. W. F. (2005). *Classical Methods of Statistics*. Springer.
- Kazhdan, M., Bolitho, M., und Hoppe, H. (2006). Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing, SGP '06*, pages 61–70, Aire-la-Ville, Switzerland, Switzerland. Eurographics Association.
- Khropov, A., Shokurov, A., Lempitskiy, V., und Ivanov, D. (2004). Reconstruction of projective and metric cameras for image triplets. *Graphicon conference*.
- Khwaja, A. und Goecke, R. (2008). Image Reconstruction from Contrast Information. In *Digital Image Computing: Techniques and Applications (DICTA), 2008*, pages 226–233.
- Klenke, A. (2006). *Wahrscheinlichkeitstheorie*. Springer.
- Krüger, L. (2007). *Model based object classification and localisation in multiocular images*. PhD thesis, Bielefeld University.
- Krüger, L., Emmert, V., Feldmann, A., und Lindner, F. (2011). Evaluating the Accuracy of Camera Calibration for Driver Assistance Systems. In *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage*. Wichmann.
- Krüger, L. und Wöhler, C. (2011). Accurate chequerboard corner localisation for camera calibration. *Pattern Recognition Letters*, 32(10):1428–1435.
- Kurz, C., Thormählen, T., und Seidel, H.-P. (2011). Bundle adjustment for stereoscopic 3d. In *Proceedings of the 5th international conference on Computer vision/computer graphics collaboration techniques, MIRAGE'11*, pages 1–12, Berlin, Heidelberg. Springer-Verlag.
- Lang, S. (1999). *Fundamentals Differential*. Graduate Texts in Mathematics Series. Springer Verlag.
- Laveau, S., Faugeras, O., Faugeras, O., Robotique, P., und Robotvis, P. (1994). 3-D Scene Representation as a Collection of Images and Fundamental Matrices. Technical report.
- Lee, W. H. und Yu, K. (2009). Bundle Block Adjustment with 3D Natural Cubic Splines. *Sensors*, 9(12):9629–9665, doi:10.3390/s91209629.
-

- Levenberg, K. (1944). A method for the solution of certain problems in least squares. *Quart. Applied Math.*, 2:164–168.
- Lindner, F. (2010). Ground Truthing von monokularen Entfernungsschätzungen mittels eines Mehrzeilen-Laserscanners. In Luhmann, T., editor, *Photogrammetrie, Laserscanning, Optische 3D-Messtechnik - Beiträge der Oldenburger 3D-Tage*, pages 210–220, Oldenburg, Germany. Wichmann - VDE Verlag.
- Lindner, F. (2012). *Adaptive traffic sign recognition*. PhD thesis, Bielefeld University.
- Lindstrom, P. (2010). Triangulation made easy. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1554–1561.
- Loose, H. (2013). *Dreidimensionale Straßenmodelle für Fahrerassistenzsysteme auf Landstraßen*. KIT Scientific Publishing.
- Luhmann, T. (2010). *Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen*. Wichmann Herbert.
- Luong, Q.-T. (1992). *Matrice fondamentale et auto-calibration en vision par ordinateur*. PhD thesis, Universite de Paris-Sud, Orsay.
- Luong, Q.-T. und Faugeras, O. D. (1994). A Stability Analysis of the Fundamental Matrix. In Eklundh, J.-O., editor, *ECCV (1)*, volume 800 of *Lecture Notes in Computer Science*, pages 577–588. Springer.
- Luong, Q.-T. und Faugeras, O. D. (1997). Self-Calibration of a Moving Camera from PointCorrespondences and Fundamental Matrices. *Int. J. Comput. Vision*, 22(3):261–289, doi:<http://dx.doi.org/10.1023/A:1007982716991>.
- Ma, Y., Soatto, S., Kosecká, J., und Sastry, S. (2003). *An Invitation to 3-D Vision: From Images to Geometric Models*. Interdisciplinary applied mathematics: Imaging, vision, and graphics. Springer.
- Martos, A., Krüger, L., und Wöhler, C. (2010). Towards Real Time Camera Self Calibration: Significance and Active Selection. In *Proc. Fourth International Symposium on 3D Data Processing, Visualization and Transmission*.
- Meiners, J. (2010). Glattegebügelt. <http://www.sueddeutsche.de/auto/mercedes-magic-body-control-bodenwellen-waren-gestern-1.1004365-8>. [Online; accessed 19-March-2013].
- Mouragnon, E., Lhuillier, M., Dhome, M., Dekeyser, F., und Sayd, P. (2006). 3d reconstruction of complex structures with bundle adjustment: an incremental approach. In *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*, pages 3055–3061.
-

- Müller, T., Rabe, C., und Franke, U. (2011). Dense6D - Position und Bewegung robust für jeden Bildpunkt. In *Proceedings of the 7th Workshop Fahrerassistenzsysteme*, Walting, Germany.
- Mühlich, M. und Aach, T. (2007). High Accuracy Feature Detection for Camera Calibration: A Multi-Steerable Approach. In *IN PROC. DAGM07: 29TH ANNUAL SYMPOSIUM OF THE GERMAN ASSOCIATION FOR PATTERN RECOGNITION*. Springer.
- Nistér, D. (2000). Reconstruction from Uncalibrated Sequences with a Hierarchy of Trifocal Tensors. In *Proceedings of the 6th European Conference on Computer Vision-Part I, ECCV '00*, pages 649–663, London, UK, UK. Springer-Verlag.
- Nister, D. (2004). An efficient solution to the five-point relative pose problem. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(6):756–770, doi:10.1109/TPAMI.2004.17.
- Nistér, D. und Schaffalitzky, F. (2004). What do Four Points in Two Calibrated Images Tell Us About the Epipoles. In *Proceedings of the 8th European Conference on Computer Vision*, pages 41–57.
- Noury, N., Sur, F., und Berger, M. O. (2007). Fundamental Matrix Estimation Without Prior Match. In *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, volume 1, pages I – 513–I – 516.
- Paragios, N., Chen, Y., und Faugeras, O. (2005). *Handbook of Mathematical Models in Computer Vision*. Springer.
- Pollefeys, M. (1999). *Self-calibration and metric 3D reconstruction from uncalibrated image sequences*. PhD thesis, K.U.Leuven.
- Prasad, M. und Fitzgibbon, A. (2006). Single View Reconstruction of Curved Surfaces. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, CVPR '06*, pages 1345–1354, Washington, DC, USA. IEEE Computer Society.
- Rabe, C. (2011). *Detection of Moving Objects by Spatio-Temporal Motion Analysis*. PhD thesis, University of Kiel, Kiel, Germany.
- Rapp, M. (2013). Stereokamerakalibrierung mittels Gauß-Newton-Verfahren auf einer Riemannschen Mannigfaltigkeit. Diplomarbeit, Universität Ulm.
- Scoleri, T., Chojnacki, W., und Brooks, M. (2007). A Decoupled Algorithm for Vision Parameter Estimation with Application to the Trifocal Tensor. In *Digital Image Computing Techniques and Applications, 9th Biennial Conference of the Australian Pattern Recognition Society on*, pages 138–143.
-

- Shashua, A. (1997). Trilinear tensor: The fundamental construct of multiple-view geometry and its applications. In Sommer, G. und Koenderink, J., editors, *Algebraic Frames for the Perception-Action Cycle*, volume 1315 of *Lecture Notes in Computer Science*, pages 190–206. Springer Berlin Heidelberg.
- Sibley, G., Mei, C., Reid, I., und Newman, P. (2009). Adaptive Relative Bundle Adjustment. In *Robotics Science and Systems (RSS)*, Seattle, USA.
- Soatto, S., Frezza, R., und Perona, P. (1994). Motion estimation on the essential manifold. In *Proceedings of the 3^d European Conference on Computer Vision, J.-O. Eklundh (Ed.), LNCS-Series Vol. 800-801, Springer-Verlag*, page 61.
- Stein, F. (2004). Efficient Computation of Optical Flow Using the Census Transform. In Rasmussen, C. E., Bülthoff, H. H., Schölkopf, B., und Giese, M. A., editors, *Proceedings of the 26th DAGM Symposium*, volume 3175 of *Lecture Notes in Computer Science*, pages 79–86. Springer.
- Stewénius, H., Engels, C., und Nistér, D. (2006). Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 60(4):284 – 294, doi:10.1016/j.isprsjprs.2006.03.005.
- Struik, D. (1961). *Lectures on Classical Differential Geometry: Second Edition*. Addison-Wesley series in mathematics. Dover Publ.
- Thirthala, S. und Pollefeys, M. (2005). The Radial Trifocal Tensor: A Tool for Calibrating the Radial Distortion of Wide-Angle Cameras. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01, CVPR '05*, pages 321–328, Washington, DC, USA. IEEE Computer Society.
- Torii, A., Kukulova, Z., Bujnak, M., und Pajdla, T. (2011). The Six Point Algorithm Revisited. In Koch, R. und Huang, F., editors, *Computer Vision - ACCV 2010 Workshops*, volume 6469 of *Lecture Notes in Computer Science*, pages 184–193. Springer Berlin Heidelberg.
- Torr, P. und Zisserman, A. (1997). Robust Parameterization and Computation of the Trifocal Tensor. *Image and Vision Computing*, 15:591–605.
- Torr, P. und Zisserman, A. (1998). Concerning Bayesian motion segmentation, model averaging, matching and the trifocal tensor. In Burkhardt, H. und Neumann, B., editors, *Computer Vision - ECCV'98*, volume 1406 of *Lecture Notes in Computer Science*, pages 511–527. Springer Berlin Heidelberg.
-

- Torr, P. H. S. und Fitzgibbon, A. (2004). Invariant fitting of two view geometry. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(5):648–650, doi:10.1109/TPAMI.2004.1273967.
- Torr, P. H. S. und Murray, D. W. (1997). The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision*, 24:271–300.
- Torres-Mendez, L., Dudek, G., und Di Marco, P. (2004). Inter-image statistics for scene reconstruction. In *Computer and Robot Vision, 2004. Proceedings. First Canadian Conference on*, pages 432–439.
- Triggs, B., Mclauchlan, P. F., Hartley, R. I., und Fitzgibbon, A. W. (2000). Bundle Adjustment – A Modern Synthesis. *Lecture Notes in Computer Science*, 1883:298+.
- Tsai, R. (1987). 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):323–344, doi:10.1109/JRA.1987.1087109.
- Weng, J., Cohen, P., und Herniou, M. (1992). Camera Calibration with Distortion Models and Accuracy Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:965–980, doi:http://doi.ieeecomputersociety.org/10.1109/34.159901.
- Werner, T. (2003). Constraint on five points in two images. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–203–8 vol.2.
- Willmore, T. (1997). *Riemannian Geometry*. Oxford Science Publications. Clarendon Press.
- Witten, I., Frank, E., und Hall, M. (2011). *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science.
- Wöhler, C. (2009). *3D Computer Vision: Efficient Methods and Applications*. Springer Verlag Berlin Heidelberg.
- Wong, K. H., Ming, M., und Chang, Y. (2004). 3D Model Reconstruction by Constrained Bundle Adjustment. *Pattern Recognition, International Conference on*, 3:902–905, doi:http://doi.ieeecomputersociety.org/10.1109/ICPR.2004.1334674.
- Xu, G. und Zhang, Z. (1996). *Epipolar Geometry in Stereo, Motion, and Object Recognition: A Unified Approach*. Kluwer Academic Publishers, Norwell, MA, USA.
-

- Yang, Z., Chen, F., Zhao, J., und Zhao, H. (2008). A novel camera calibration method based on genetic algorithm. In *Industrial Electronics and Applications, 2008. ICIEA 2008. 3rd IEEE Conference on*, pages 2222–2227.
- Yu, Y.-K., Wong, K. H., Chang, M.-Y., und Or, S. H. (2006). Recursive Camera-Motion Estimation With the Trifocal Tensor. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(5):1081–1090, doi:10.1109/TSMCB.2006.874133.
- Zach, C., Pock, T., und Bischof, H. (2007). A duality based approach for realtime tv-11 optical flow. In *Proceedings of the 29th DAGM conference on Pattern recognition*, pages 214–223, Berlin, Heidelberg. Springer-Verlag.
- Zhang, Z. (1996). On the epipolar geometry between two images with lens distortion. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 1, pages 407–411 vol.1.
- Zhang, Z. (1999). Flexible camera calibration by viewing a plane from unknown orientations. In *in ICCV*, pages 666–673.
- Zhu, K., Butenuth, M., und d’Angelo, P. (2012). Comparison of Dense Stereo Using CUDA. In Kutulakos, K., editor, *Trends and Topics in Computer Vision*, volume 6554 of *Lecture Notes in Computer Science*, pages 398–410. Springer Berlin Heidelberg.
-

