

Thema der Dissertation

**Entwicklung und Validierung methodischer Konzepte einer
kamerabasierten Durchfahrtshöhenerkennung für Nutzfahrzeuge**

Dissertation zur Erlangung des akademischen Grades Doktoringenieur (Dr.-Ing.)

vorgelegt von

M.Sc. / Hänert, Stephan

geboren am: 25. November 1987 in: Querfurth, Deutschland

Gutachter:

Herr Prof. Dr. sc. techn. habil. Hans-Gerd Maas
Technische Universität Dresden

Herr Prof. Dr.-Ing. Markus Gerke
Technische Universität Braunschweig

Herr Prof. Dr.-Ing. Günther Prokop
Technische Universität Dresden

Dresden, der 03.03.2020

Disclaimer:

Veröffentlichungen über den Inhalt dieser Arbeit sind nur mit schriftlicher Genehmigung der Volkswagen AG zugelassen.

Die Ergebnisse, Meinungen und Schlüsse dieser Dissertation sind nicht notwendigerweise die der Volkswagen AG.

Danksagung

Diese Arbeit ist im Rahmen meiner Doktorandentätigkeit bei der Volkswagen AG innerhalb der Nutzfahrzeugentwicklung entstanden. Ich möchte mich an dieser Stelle bei allen bedanken, die mich während dieser Zeit unterstützt, fachlich begleitet und gefördert haben.

Ich danke Prof. Dr. sc. techn. habil. Hans-Gerd Maas vom Institut für Photogrammetrie und Fernerkundung der TU Dresden für sein Engagement und die investierte Zeit im Rahmen der wissenschaftlichen Betreuung meiner Arbeit. Besonders die fachlichen Diskussionen und das wertvolle Feedback waren mir eine große Unterstützung. Darüber hinaus bedanke ich mich bei meinem Zweitgutachter Prof. Dr.-Ing. Markus Gerke vom Institut für Geodäsie und Photogrammetrie der TU Braunschweig für den direkten und unkomplizierten Austausch sowie für die Zweitbegutachtung der Arbeit. Außerdem möchte ich mich bei meinem Drittgutachter Prof. Dr.-Ing. Günther Prokop vom Institut für Automobiltechnik der TU Dresden bedanken, dass er trotz vieler anderweitiger Verpflichtungen die Begutachtung der Arbeit übernommen hat.

Mein Dank gilt zudem den Kolleginnen und Kollegen in der Abteilung. Zunächst möchte ich mich bei Thomas Mächtig, Kai Grünitz, Heiko Hartmann und Markus Huhn für die Möglichkeit bedanken, dieses interessante Thema bearbeiten zu können. Ihr habt mir stets den Freiraum dafür gegeben und mich in schwierigen Situationen unterstützt und motiviert. Zudem danke ich Eugen Schneider, Simon Bork und Marwan Sabih für das Korrekturlesen meiner Arbeit sowie den zahlreichen fachlichen Gesprächen und Feedback zu meinem Promotionsvorhaben. Die konstruktiven Anregungen und das Hinterfragen jeglicher Aspekte haben wesentlich zum Abschluss dieser Arbeit beigetragen. Des Weiteren bedanke ich mich bei Dr. Dominik Vock für die wertvollen Anregungen zum Thema und der Unterstützung in zahlreichen Kalibrierprozessen.

Besonderer Dank gilt meiner Familie, vornehmlich meinen Eltern Bettina und Ronald Hänert, die mich auf meinem bisherigen Lebensweg und damit auch über den langen Entstehungsprozess dieser Arbeit jederzeit und in jeglicher Hinsicht über den Maßen unterstützt haben. Ihr seid die Besten!

Abschließend möchte ich mich bei meiner Freundin Audrey Stref-Lamberty für Ihre unendliche Geduld und Unterstützung vor allem an den arbeitsintensiven Wochenenden und Abenden gegen Ende der Dissertation bedanken.

Kurzfassung

Die vorliegende Arbeit beschäftigt sich mit der Konzeptionierung und Entwicklung eines neuartigen Fahrerassistenzsystems für Nutzfahrzeuge, welches die lichte Höhe von vor dem Fahrzeug befindlichen Hindernissen berechnet und über einen Abgleich mit der einstellbaren Fahrzeughöhe die Passierbarkeit bestimmt. Dabei werden die von einer Monokamera aufgenommenen Bildsequenzen genutzt, um durch indirekte und direkte Rekonstruktionsverfahren ein 3D-Abbild der Fahrumgebung zu erschaffen. Unter Hinzunahme einer Radodometrie-basierten Eigenbewegungsschätzung wird die erstellte 3D-Repräsentation skaliert und eine Prädiktion der longitudinalen und lateralen Fahrzeugbewegung ermittelt. Basierend auf dem vertikalen Höhenplan der Straßenoberfläche, welcher über die Aneinanderreihung mehrerer Ebenen modelliert wird, erfolgt die Klassifizierung des 3D-Raums in Fahruntergrund, Struktur und potentielle Hindernisse. Die innerhalb des Fahrschlauchs liegenden Hindernisse werden hinsichtlich ihrer Entfernung und Höhe bewertet. Ein daraus abgeleitetes Warnkonzept dient der optisch-akustischen Signalisierung des Hindernisses im Kombiinstrument des Fahrzeugs. Erfolgt keine entsprechende Reaktion durch den Fahrer, so wird bei kritischen Hindernishöhen eine Notbremsung durchgeführt.

Die geschätzte Eigenbewegung und berechneten Hindernisparameter werden mithilfe von Referenzsensorik bewertet. Dabei kommt eine dGPS-gestützte Inertialplattform sowie ein terrestrischer und mobiler Laserscanner zum Einsatz. Im Rahmen der Arbeit werden verschiedene Umgebungssituationen und Hindernistypen im urbanen und ländlichen Raum untersucht und Aussagen zur Genauigkeit und Zuverlässigkeit des Verfahrens getroffen.

Ein wesentlicher Einflussfaktor auf die Dichte und Genauigkeit der 3D-Rekonstruktion ist eine gleichmäßige Umgebungsbeleuchtung innerhalb der Bildsequenzaufnahme. Es wird in diesem Zusammenhang zwingend auf den Einsatz einer Automotive-tauglichen Kamera verwiesen. Die durch die Radodometrie bestimmte Eigenbewegung eignet sich im langsamen Geschwindigkeitsbereich zur Skalierung des 3D-Punktraums. Dieser wiederum sollte durch eine Kombination aus indirektem und direktem Punktrekonstruktionsverfahren erstellt werden. Der indirekte Anteil stützt dabei die Initialisierung des Verfahrens zum Start der Funktion und ermöglicht eine robuste Kameraschätzung. Das direkte Verfahren ermöglicht die Rekonstruktion einer hohen Anzahl an 3D-Punkten auf den Hindernisumrissen, welche zumeist die Unterkante beinhalten. Die Unterkante kann in einer Entfernung bis zu 20 m detektiert und verfolgt werden. Der größte Einflussfaktor auf die Genauigkeit der Berechnung der lichten Höhe von Hindernissen ist die Modellierung des Fahruntergrunds. Zur Reduktion von Ausreißern in der Höhenberechnung eignet sich die Stabilisierung des Verfahrens durch die Nutzung von zeitlich vorher zur Verfügung stehenden Berechnungen. Als weitere Maßnahme zur Stabilisierung wird zudem empfohlen die Hindernisausgabe an den Fahrer und den automatischen Notbremsassistenten mittels einer Hysterese zu stützen.

Das hier vorgestellte System eignet sich für Park- und Rangiervorgänge und ist als kostengünstiges Fahrerassistenzsystem interessant für Pkw mit Aufbauten und leichte Nutzfahrzeuge.

Abstract

The present work deals with the conception and development of a novel advanced driver assistance system for commercial vehicles, which estimates the clearance height of obstacles in front of the vehicle and determines the passability by comparison with the adjustable vehicle height. The image sequences captured by a mono camera are used to create a 3D representation of the driving environment using indirect and direct reconstruction methods. The 3D representation is scaled and a prediction of the longitudinal and lateral movement of the vehicle is determined with the aid of a wheel odometry-based estimation of the vehicle's own movement. Based on the vertical elevation plan of the road surface, which is modelled by attaching several surfaces together, the 3D space is classified into driving surface, structure and potential obstacles. The obstacles within the predicted driving tube are evaluated with regard to their distance and height. A warning concept derived from this serves to visually and acoustically signal the obstacle in the vehicle's instrument cluster. If the driver does not respond accordingly, emergency braking will be applied at critical obstacle heights.

The estimated vehicle movement and calculated obstacle parameters are evaluated with the aid of reference sensors. A dGPS-supported inertial measurement unit and a terrestrial as well as a mobile laser scanner are used. Within the scope of the work, different environmental situations and obstacle types in urban and rural areas are investigated and statements on the accuracy and reliability of the implemented function are made.

A major factor influencing the density and accuracy of 3D reconstruction is uniform ambient lighting within the image sequence. In this context, the use of an automotive camera is mandatory. The inherent motion determined by wheel odometry is suitable for scaling the 3D point space in the slow speed range. The 3D representation however, should be created by a combination of indirect and direct point reconstruction methods. The indirect part supports the initialization phase of the function and enables a robust camera estimation. The direct method enables the reconstruction of a large number of 3D points on the obstacle outlines, which usually contain the lower edge. The lower edge can be detected and tracked up to 20 m away. The biggest factor influencing the accuracy of the calculation of the clearance height of obstacles is the modelling of the driving surface. To reduce outliers in the height calculation, the method can be stabilized by using calculations from older time steps. As a further stabilization measure, it is also recommended to support the obstacle output to the driver and the automatic emergency brake assistant by means of hysteresis.

The system presented here is suitable for parking and maneuvering operations and is interesting as a cost-effective driver assistance system for cars with superstructures and light commercial vehicles.

Inhaltsverzeichnis

Danksagung

Kurzfassung

Abstract

1	Einleitung.....	1
1.1	Stand der Technik.....	2
1.1.1	Passive Systeme.....	2
1.1.2	Aktive Systeme.....	3
1.2	Funktion und System.....	7
1.2.1	Anforderungen.....	7
1.2.2	Systemüberblick	9
1.3	Zielsetzung	11
1.4	Gliederung der Arbeit.....	11
2	Grundlagen	13
2.1	Technische Grundlagen	13
2.1.1	Fahrzeug- und Referenzsensorik	13
2.1.2	Vernetzung und Systeme	20
2.1.3	Software und Entwicklungsumgebung	22
2.2	Kamerabasierte Rekonstruktionsverfahren.....	22
2.3	Mathematische und photogrammetrische Grundlagen	24
2.3.1	Koordinatensysteme	24
2.3.2	Projektive Geometrie.....	25
2.3.3	Geometrisches Kameramodell.....	27
2.3.4	Epipolargeometrie	29
2.4	Statistik.....	34
3	Kamerakalibrierung	37
3.1	Grundlagen	37
3.2	Bündelausgleichung	37
3.3	Kalibrierumgebung.....	38
3.4	Beurteilung	46
4	Fahrzeugeigenbewegung	49
4.1	Methoden zur Eigenbewegungsschätzung.....	49
4.2	Bewegungsmodell	51
4.3	Optimierung der Radodometrie	54
4.3.1	Optimierungstheorie	54

4.3.2	Implementierung	56
4.3.3	Auswertung	58
4.4	Skalierung der Szene	60
4.5	Fahrschlauchprädiktion	62
4.5.1	Bildbasierte Prädiktion	63
4.5.2	Eigenbewegungsbasierte Prädiktion	66
5	Merkmalsbasierte Rekonstruktionsverfahren	68
5.1	Grundlagen	68
5.1.1	Anforderungen an lokale Merkmale	68
5.1.2	Verfahrensschritte	69
5.1.3	Aufnahmekonfiguration	69
5.2	Punkte	72
5.2.1	Detektion	72
5.2.2	Deskription	76
5.2.3	Korrespondenzsuche	78
5.2.4	Rekonstruktion	85
5.3	Kanten und Linien	88
5.3.1	Kanten- und Liniendetektion	88
5.3.2	Korrespondenzsuche	89
5.3.3	Rekonstruktion	90
6	Direkte und hybride Rekonstruktionsverfahren	94
6.1	Verfahren	94
6.2	Direct Sparse Odometry	95
6.2.1	Grundlagen	95
6.2.2	Framework	97
7	Fahreruntergrund- und Hindernisdetektion	99
7.1	Vorverarbeitung	99
7.1.1	Rauschen und Ausreißer	99
7.1.2	Raumunterteilende Datenstrukturen	102
7.2	Fahreruntergrunddetektion	104
7.2.1	Grundlagen	105
7.2.2	Fahreruntergrundmodellierung	106
7.3	Struktur- und Hindernisdetektion	111
8	Evaluierung	114
8.1	Eigenbewegungsschätzung	114
8.2	Umgebungsrekonstruktion	116

8.2.1	Punkte	116
8.2.2	Linien.....	123
8.3	Fahruntergrundmodellierung.....	126
8.4	Hindernisdetektion	127
8.4.1	Grundlagen	128
8.4.2	Auswertung.....	130
8.5	Fazit	135
9	Warnstrategie.....	136
10	Zusammenfassung	138
11	Ausblick.....	142
	Literaturverzeichnis	144
	Abkürzungsverzeichnis	153
	Mathematische Notation.....	154
	Abbildungsverzeichnis	155
	Tabellenverzeichnis.....	160
Appendix A	Kamerakalibrierung.....	161
Appendix B	Fahruntergrundmodellierung.....	162
Appendix C	Hindernisdetektion.....	166

1 Einleitung

Der Weg zur Vision Zero, dem unfallfreien Fahren im Straßenverkehr, ist sehr lang und es gilt viele Hindernisse zu meistern. Jährlich finden auf den Straßen Deutschlands ca. 2.6 Mio. Unfälle statt; davon ca. 390000 mit Personenschaden und knapp 3200 mit Verkehrstoten. Die Zahlen sind durch Einführung von passiven und aktiven Sicherheitssystemen seit den 1950er Jahren erheblich gesunken, bleiben in den letzten Jahren jedoch nahezu konstant [1]. Denn obwohl moderne Assistenzsysteme den Markt stark durchdrungen haben und zum Teil gesetzlich vorgeschrieben sind, sind circa 90 % aller Unfälle auf ein Fehlverhalten des menschlichen Fahrers zurückzuführen.

Ein wichtiger Bestandteil zur Reduzierung der Unfälle soll das automatische Fahren einnehmen. Ein Ansatz, welcher Bedienungs- und Interpretationsfehler des Menschen auf ein Minimum reduzieren kann. Die Vorstufe zum hoch- und vollautomatischen Fahrsystemen, die ab ca. 2025 erwartet werden [2], sind teilautomatisierte Fahrerassistenzfunktionen, die bereits jetzt zum Einsatz kommen. Die Basis solcher Assistenzfunktionen liegt in der Erkennung des fahrzeugnahen Umfelds durch die fahrzeugeigene Sensorik. Es gilt das Umfeld möglichst detailliert und in Echtzeit zu erfassen, das „Gesehene“ zu interpretieren und anhand der vorliegenden Daten Systeme zu entwickeln, die dem Fahrer bei der Fahraufgabe unterstützen.

Nutzfahrzeuge, die sich durch ihr höheres zulässiges Gesamtgewicht, ihre größeren Abmaße und die Art ihrer Nutzung von Personenkraftwagen unterscheiden, ermöglichen durch den gezielten Einsatz in ländlichen und städtischen Gebieten eine Effizienzsteigerung im Personen- und Gütertransport. Die konstruktiven und nutzungsbedingten Besonderheiten müssen bei der Auslegung von Assistenzsystemen berücksichtigt werden, sodass eine einfache Adaption aus dem Pkw-Bereich oft nicht ausreicht. Leichte Nutzfahrzeuge eignen sich bspw. aufgrund ihrer Zuladung für Umzüge oder den Transport größerer Güter und werden für diese speziellen Zwecke häufig kurzzeitig von Fahrzeugverleihern vermietet. Dabei ist für unerfahrene Nutzer vor allem die Einschätzung der eigenen Fahrzeughöhe schwierig, sodass es gehäuft zu Unfällen kommt, wenn die lichte Höhe von Brücken und Einfahrten falsch interpretiert wird. Diese Art von Unfällen stellt eine direkte Gefahr für die Insassen als auch den nachfolgenden Verkehr dar. Doch auch ohne Personenschaden führt die Beschädigung der Dachstruktur, welche essentiell für die Steifigkeit und Stabilität der Karosserie ist, häufig zu einem Totalschaden des Fahrzeugs.



Abbildung 1-1 Beispiele für Kollisionen durch unzureichende Durchfahrthöhe [3, 4, 5]

Neben leichten Nutzfahrzeugen sind zum größten Teil Lastkraftwagen und Busse von diesen Unfällen betroffen. Laut [6] findet im Vereinigten Königreich (UK) im öffentlichen Straßenverkehrsnetz durchschnittlich alle 4.5 Std. eine Kollision aufgrund von zu geringer Durchfahrthöhe statt, wobei die Kosten für eine Brückenreparatur in diesen Fällen bei umgerechnet bis zu 160000 € liegen können. Selbst im Pkw-Segment geschehen solche Unfälle, wenn Dachaufbauten, wie z.B. Dachboxen oder Fahrradträger vergessen werden (s. Abbildung 1-1).

Durch den Einsatz von Schildern, Warnbaken oder speziellen Vermessungssystemen, die in der Nähe von Brücken und Unterführungen angebracht werden, soll die Zahl an Unfällen verringert werden. Diese Systeme setzen jedoch voraus, dass der Fahrer die Warnungen wahrnimmt und entsprechend reagiert. Im Gegensatz dazu soll innerhalb dieser Arbeit ein fahrzeugeigenes System zur Detektion von potentiellen Hindernissen mit eingeschränkter Durchfahrthöhe mittels monokularer Kamera entwickelt werden. Nach der Hindernisdetektion und Berechnung der lichten Höhe erfolgt der Abgleich mit der Fahrzeughöhe und ggf. eine Warnausgabe. Sofern der Fahrer diese ignoriert, reagiert das System im Notfall selbstständig mit einer Gefahrenbremsung.

1.1 Stand der Technik

Auf dem Markt gibt es eine Vielzahl von Systemen zur Warnung vor Brücken, Einfahrten und Durchfahrten mit beschränkter lichter Höhe. Die Systeme helfen dem Fahrer sich der Höhe seines Fahrzeuges bewusst zu werden und warnen ihn, bevor es zu einem teils schwerwiegenden Unfall kommt. Lange Ausfallzeiten oder sogar Totalschäden der Fahrzeuge sind häufig eine Konsequenz der Unfälle, sodass Fuhrparkunternehmer einen erheblichen Kostenaufwand und Zeitverlust damit tragen.

Die Systeme sind im öffentlichen Bereich auch bei Gemeinden und Ländern von Interesse, da Unfälle auch starke Beschädigungen an Brücken und Durchfahrten verursachen. Die Erhaltung und Instandsetzung von Brücken ist mit enormen Kosten verbunden und behindern oft für einen langen Zeitraum den Verkehr. Die Systeme, die im Folgenden vorgestellt werden, lassen sich in zwei Gruppen unterteilen.

1.1.1 Passive Systeme

Die passiven Systeme sind reine Anzeigen und warnen den Fahrer vor einer auf der Route befindlichen vertikalen Engstelle. Dazu zählen Verkehrszeichen mit Höhenangaben oder Warnbaken.



Abbildung 1-2 Beispiele passiver Warnsysteme. Links: Parkhauseinfahrt [7]; Mitte und Rechts: Brückenunterführung [8, 9]

Passive Systeme sind nicht in der Lage die Höhen des Fahrzeugs und der Durchfahrt zu bestimmen und zu vergleichen und dienen somit nur unzureichend der Unfallvermeidung. Als problematisch sind vor allem Schilder mit Höheninformationen anzusehen. Zum einen sind diese oft direkt am Hindernis angebracht, sodass sich die Erkennung aus großen Distanzen als problematisch erweist. Bei entsprechender Geschwindigkeit kann demnach auch ein Notbremsmanöver den Unfall nicht verhindern, sondern maximal dessen Stärke reduzieren kann. Zum anderen ist nicht gewährleistet, dass die angezeigte Höhe der Realität entspricht. Gerade in privaten Geländen oder in Parkhäusern sind diese nur eine Orientierung und es kann nicht gewährleistet werden, dass die angezeigte Höhe auch minimalen Höhe im gesamten Bereich des Geländes entspricht.

1.1.2 Aktive Systeme

Aktive Systeme sind in der Lage die Höhe des Fahrzeugs oder des Hindernisses zu bestimmen und situativ den Fahrer zu warnen. Sie können in zwei weitere Klassen unterteilt werden.

Fahrzeugdetektion

Zunächst werden Fahrzeugdetektionssysteme vorgestellt, die vor kritischen Engstellen die Höhe des Fahrzeuges bestimmen und den Fahrer warnen. Diese sind oft weit vor dem eigentlichen Hindernis positioniert, um dem Fahrer die Möglichkeit des Verlassens der Fahrspur und Wählen einer Ausweichroute zu ermöglichen. Sind infrastrukturell solche Bedingungen nicht gegeben, kann nur noch auf die unzureichende Höhe der Durchfahrt gewarnt und sich auf die Einleitung eines Bremsmanövers fokussiert werden.

Die Detektion der Fahrzeughöhe kann durch verschiedene Sensorsysteme erfolgen. Im einfachsten Fall wird vor dem eigentlichen Hindernis ein sogenanntes Warnportal installiert. Dies können je nach Ausführung eine beweglich aufgehängte Warnbake, Ketten oder reflektierende Stangen sein.

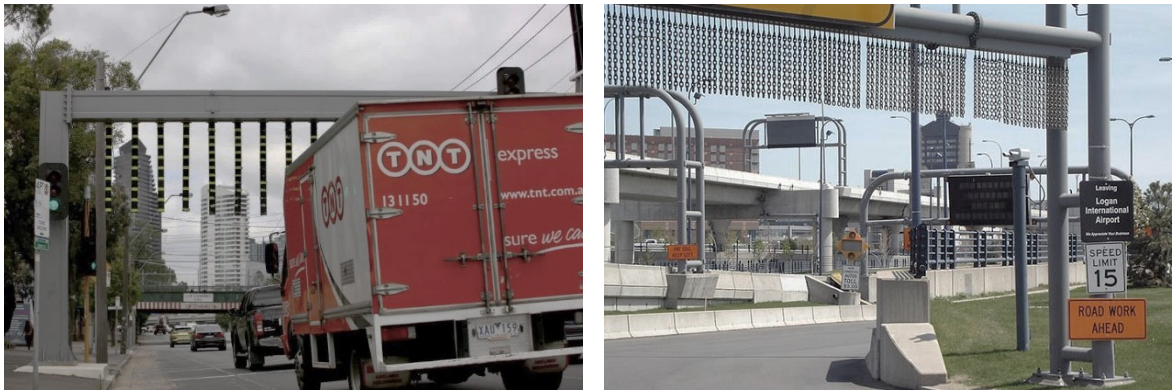


Abbildung 1-3 Beispiel Warnportale. Links: Reflektierende Stangen [10]; Rechts: Ausführung mit Ketten [11]

Neben den einfachen Warnportalen bieten viele Hersteller eine kontaktlose Variante zur Detektion der Fahrzeughöhe an. Diese reichen von Ein-Sensorsystemen bis hin zu Multi-Sensorsystemen, die Verbesserungen bei ungünstigen Wetterbedingungen und Beleuchtung bieten. *Trigg Industries International INC* vertreibt ein System, welches über im Boden eingelassene Induktionsschleifen das Fahrzeug detektiert [11]. Das auf Brückenhöhe angebrachte *Double Beam* sendet zwei Strahlen im sichtbaren und Infrarotbereich aus. Werden diese durchbrochen, so erfolgt die Warnung an den Fahrer über ein Warnschild. Ein ähnliches System wurde von [12] entwickelt, welches vorbeifahrende Fahrzeuge mit einer Geschwindigkeit bis zu 250 km/h detektiert.



Abbildung 1-4 Laserbasiertes System von [11]

Ein weiteres System wurde von [13] vorgestellt, wobei ein an der Brücke angebrachter LiDAR (Light Detection and Ranging) die ankommenden Fahrzeuge überwacht und bis zu 300 m Entfernung zunächst eine Klassifizierung vornimmt und anschließend deren Kritikalität bewertet.

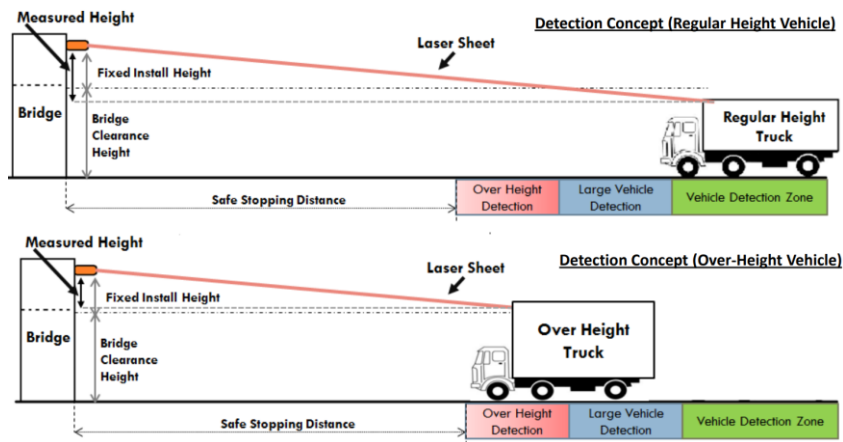


Abbildung 1-5 Schema: LaRa-OHVD [13]

Als Alternative zu Infrarot- bzw. LiDAR-basierten Systemen sind kamerabasierte Systeme oft kostengünstiger. Diese werden bereits vor allem zur Verkehrsüberwachung genutzt. Neben der Schätzung des Verkehrsaufkommens über Fahrzeugzählungen können indes Wetter- und Straßenkonditionen analysiert werden. Ein darüber hinaus gehender Algorithmus zur Ermittlung der Fahrzeughöhe kann in drei Bereiche unterteilt werden. Das sind neben der Kamerakalibrierung die Fahrzeugdetektion sowie -verfolgung. Die Kalibrierung ermöglicht dabei die Transformation zwischen Bild- und Weltkoordinatensystem und wird zur Bestimmung der Fahrzeugdimension genutzt.

Ein kamerabasiertes System wird in [14] vorgestellt. Die Kamera zeichnet dabei im rechten Winkel zur Fahrbahn die vorbeifahrenden Fahrzeuge auf. Durch die vorgenommene Kalibrierung kann der Bereich der kritischen Fahrzeughöhe im Bild markiert und ausgewertet werden. Mittels einer Kombination von SURF-Detektor (Speeded Up Robust Features) und Merkmalsverfolgung über optischen Fluss werden Fahrzeuge detektiert und anschließend von Rauschen und Fremdobjekten unterschieden. In [15] wird der auf das Hindernis zufahrende Verkehr ebenfalls mittels Monokamera beobachtet. Mittels Liniendetektion und Blob Tracking wird die Distanz zwischen oberen und unteren Pixelkoordinaten eines Fahrzeugs bestimmt und in 3D-Weltkoordinaten projiziert (s. Abbildung 1-6).



Abbildung 1-6 Höhenbestimmung nach [15] v.l.n.r.: Linienerkennung innerhalb einer Bounding Box, Auswahl der in Fahrtrichtung liegenden Linie mit kürzestem Abstand zur linken oberen Ecke der Bounding Box, Blob Detektion, Canny Edge Filter und Auswahl der untersten Linie, Abstandsberechnung zwischen oberer und unterer Linie

Die vorgestellten Systeme sind lokal begrenzt für nur eine Durchfahrt ausgelegt und mit Anschaffungskosten zwischen 5000 \$ und 25000 \$ sehr teuer. Allein kamerabasierte Systeme sind mit rund 700 \$ vergleichsweise günstig. Der Installations- und Wartungsaufwand ist abhängig von der gewählten Sensorkonfiguration [16]. Die Systeme zur Fahrzeugdetektion setzen voraus, dass der Fahrer ihre Warnausgaben wahrnimmt und berücksichtigt. Geschieht dies nicht, ist ein Unfall unvermeidbar.

Hindernisdetektion

Am Fahrzeug angebrachte Systeme zur Hindernisdetektion sind in der Lage die Höhe eines beliebigen potentiellen Hindernisses zu detektieren und können darüber hinaus auch automatische Bremsengriffe einleiten. Sie bieten einen zusätzlichen Schutz für den Fahrer. Diese Systeme sind universeller einsetzbar, erfordern jedoch einen hohen Entwicklungsaufwand, um die Kritikalität von Durchfahrten in verschiedenen Situationen zu erkennen.

Vor allem die Detektion des potentiellen Hindernisses in ausreichender Entfernung ist komplex. Neben der Vermeidung von Fehldetektionen ist die Bestimmung der tatsächlichen Durchfahrtshöhe unter Berücksichtigung des vertikalen Fahrbahnverlaufs am Kritischsten.

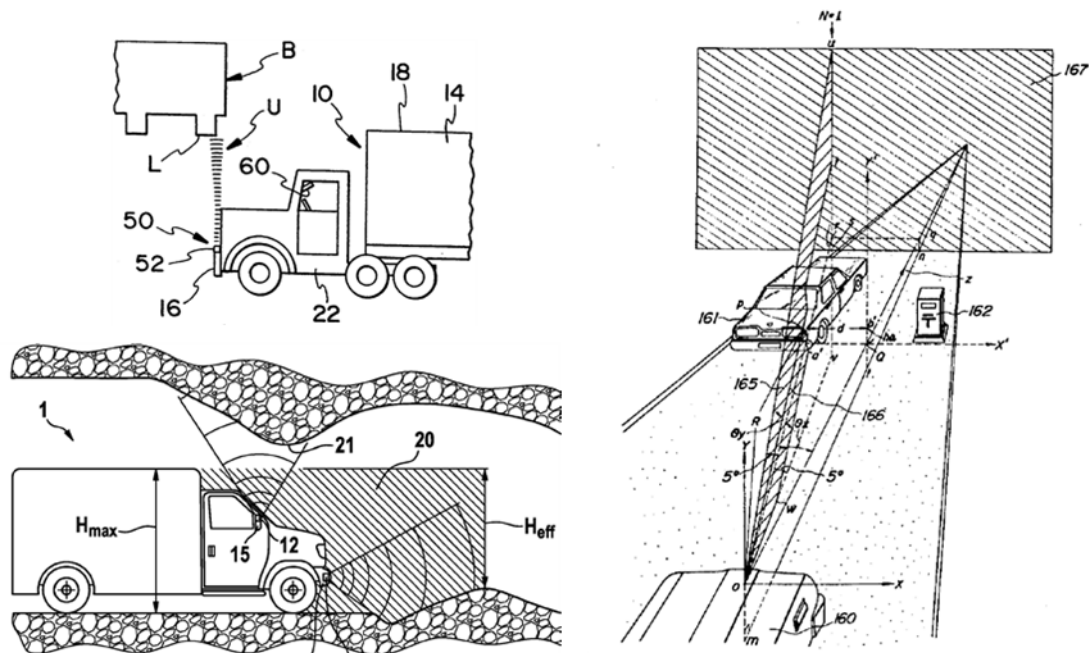


Abbildung 1-7 Systeme zur Berechnung der Durchfahrtshöhe. Links oben: Ultraschallbasiertes System [17]; Links unten: Zwei-Sensor System [18]; Rechts: LiDAR-basiertes System [19]

In [17] und [20] wird für die Berechnung der Durchfahrtshöhe ein am Fahrzeug vertikal angebrachter Ultraschallsensor genutzt. Dieser kann sich zum einen fest verbaut an der Front des Fahrzeugs befinden, oder sich variabel, bspw. an den Spiegelhalterungen befestigen lassen. Die beiden Systeme gleichen ihre Einbauhöhe mit der Fahrzeughöhe ab und warnen den Fahrer distanzabhängig optisch oder akustisch. Beide Systeme sind aufgrund der vertikalen Einbaulage ausschließlich für langsame Geschwindigkeiten geeignet. Eine Erweiterung wird in [21] vorgestellt, bei der der Sensor leicht in Fahrtrichtung gekippt wird, wodurch sich eine Frühwarnung erreichen lässt, sich der Rechenaufwand jedoch entscheidend vergrößert. Ein vollständig in Fahrtrichtung liegendes System wurde bereits in [19] vorgestellt. Dies ist auf die Nutzung mit einem LiDAR zugeschnitten und ermöglicht die Berechnung von Durchfahrtshöhen sowie -breiten. Eine Kombination aus Berechnung der Hindernishöhe mit Berücksichtigung des vertikalen Fahrbahnverlaufs wird in [18] beschrieben.

Dabei wird jedoch auf ein Zwei-Sensor System verwiesen. Ein System zur Durchfahrtshöhenerkennung mittels Stereokamera wurde in Zusammenarbeit von Bosch und Jaguar Land Rover entwickelt [22].

Das in dieser Arbeit vorgestellte System soll im Hinblick auf Kosten aus maximal einem Umfeldsensor bestehen und dennoch die Hindernishöhe in Abhängigkeit des Fahrbahnverlaufs ermitteln können. Wie eingangs erwähnt soll dies mit einer in Fahrtrichtung verbauten Monokamera geschehen. Existierende Verfahren zur monokamerabasierten Abstands- und Höhenberechnung sollen im Folgenden vorgestellt werden.

Monokamerabasierte Verfahren

Eine einfache Umsetzung der Funktion ist die Erweiterung der bestehenden Verkehrszeichenerkennung zur zusätzlichen Detektion von Verkehrszeichen mit Höheninformationen. Aktuell werden vor allem Geschwindigkeitsbegrenzungen und Überholverbote ausgewertet. Die Schilder sind je nach Land genormt, können in einer Datenbank zusammengefasst und per Template Matching ausgewertet werden. Diese Methode hat jedoch folgende Nachteile. Verkehrszeichen mit Höheninformationen sind oft direkt am Hindernis angebracht, sind aus großer Entfernung nur schwer zu detektieren und finden in privaten Geländen nur bedingt Anwendung. Darüber hinaus muss angezeigte Höhe nicht zwangsweise der genauen Durchfahrtshöhe entsprechen. Alternativ zur Verkehrszeichenerkennung sollen weitere kamerabasierte Verfahren vorgestellt werden, welche eine Detektion des Hindernisses, dessen Höhe und Entfernung ermöglicht.

In [23, 24, 25] werden Ansätze zur Bestimmung von Höheninformationen aus einem einzelnen Bild vorgestellt. Die gemeinsame Grundlage bildet das Finden eines vertikalen Fluchtpunkts und die Detektion der Fluchtlinie (oft Horizont) der Grundebene. Mithilfe von Referenzhöhen im Bild oder durch Kenntnis der Brennweite und Einbaulage der Kamera, kann die Höhe eines beliebigen Objekts im Bild bestimmt werden, welche sich unter der Manhattan-Welt-Annahme orthogonal zur Grundebene befindet. Die Verfahren erfordern ein Markieren des gewünschten Objekts im Bild. Insbesondere das Finden der exakten Hindernisunterkante und dessen Fußpunkt auf der Grundebene (Fahrbahn) bedingt einen komplexen automatisierten Algorithmus. Als problematisch wird zudem das Auffinden des Horizonts gesehen, da dies vorrangig auf einer Liniendetektion basiert, welche in unstrukturierten Gebieten fehlschlagen kann.

Im Automotive- und Robotik-Bereich wird zur Auswertung von Objektinformationen im Gegensatz zu den Einbildverfahren oft eine Bildfolge der im bewegten Fahrzeug befindlichen Kamera genutzt. So wird in [26] die Änderung der Skalierung über mehrere Kamerabilder zur Berechnung der Distanz eines vorausfahrenden Fahrzeugs genutzt. In [27] wird die Entfernung eines statischen Objekts in großen Entfernungen mittels optischen Fluss detektiert. In [28] wird ein System zur Berechnung von Entfernungen und Geschwindigkeiten auf Basis einer Nummernschilderkennung vorgestellt. All diese Systeme haben gemein, dass Einschränkungen bezüglich der Umgebungsparameter gemacht werden. Die Straßenoberfläche wird vereinfacht als eben angenommen. Besondere Informationen zu den Hindernissen, wie Größe, Form oder Farbe werden als bekannt angenommen, sodass Detektionsalgorithmen speziell auf diese Umgebungsparameter abgestimmt werden. Das Hindernis wird nach der Detektion über mehrere Bilder verfolgt und die Berechnung des Abstandes und weiterer Größen erfolgt über die Skalierungsänderung. Die Algorithmen sind für bestimmte Verkehrssituationen zugeschnitten, sodass das System im normalen Fahrbetrieb nur selten entsprechend genutzt werden kann.

Im vorliegenden Fall soll ein allgemeiner Ansatz gewählt werden, der es erlaubt beliebige Hindernisse mittels einer Bildfolge zu identifizieren sowie deren Abstand und Höhe im Sinne der Unterfahrbarkeit zu berechnen. Es sollen dafür keine Einschränkungen bezüglich der Hindernisparameter gemacht werden. Dazu werden die Informationen nicht allein aus dem 2D-Bild,

sondern innerhalb eines 3D-Raums des naheliegenden Umfelds des Fahrzeugs bezogen. Die Kamera wird zur Erstellung dieses 3D-Raums genutzt.

1.2 Funktion und System

Im Rahmen dieser Arbeit wird ein System zur Erkennung von Durchfahrtshöhen auf Basis des Volkswagen T6 entwickelt. Die Rahmenbedingungen sollen hier vorgestellt und erläutert werden.

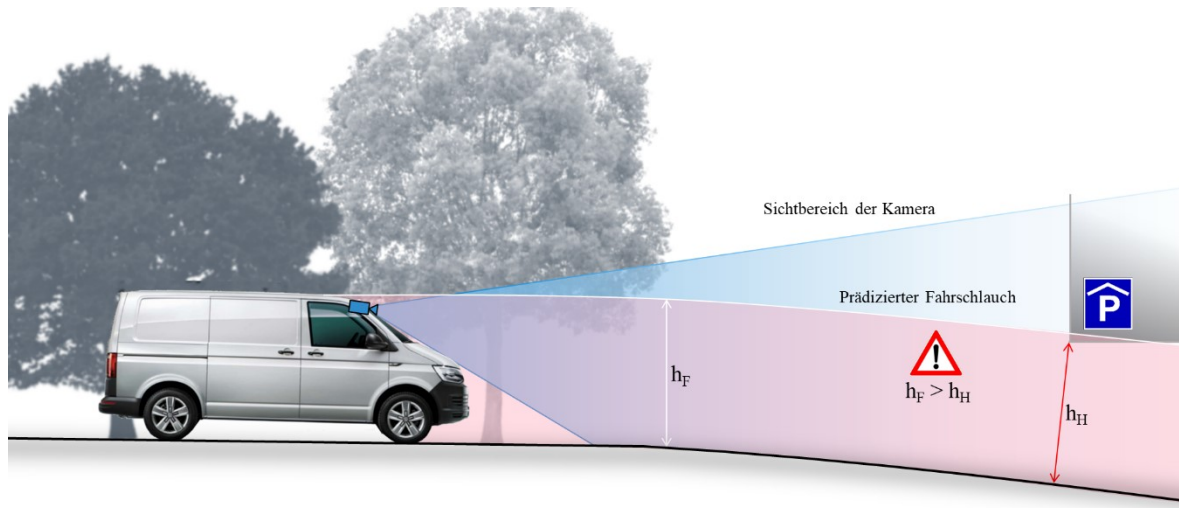


Abbildung 1-8 Schema: Kamerabasierte Durchfahrtshöhenerkennung

Der schematische Aufbau der Funktion ist in Abbildung 1-8 gezeigt. Dabei erkennt das Fahrzeug mittels Monokamera das vor sich liegende potentielle Hindernis; hier eine Parkhauseinfahrt. Unter Berücksichtigung der eigenen Fahrzeugbewegung, der Fahrzeughöhe und des Fahrbahnverlaufs wird eine Fahrschlauchprädiktion durchgeführt. Durch den Vergleich der Hindernis- und Fahrzeughöhe kann eine Aussage über die Passierbarkeit der Einfahrt getroffen werden.

1.2.1 Anforderungen

Leichte Nutzfahrzeuge werden im Markt mit einer hohen Spreizung an Ausstattungen und Varianten angeboten. Neben einfachen Transportern, die vor allem im gewerblichen Umfeld genutzt werden, werden auf gleicher Basis auch hochausgestattete Familien- und Freizeitmobile für den privaten Bereich entwickelt. Die Auswahl an Umfeldsensoren ist mit dem Hintergrund der Wirtschaftlichkeit eingeschränkt, sodass komplexe Fahrerassistenzsysteme im leichten Nutzfahrzeugbereich oft mit kostengünstigen und einfachen Sensoren realisiert werden müssen. Insofern wird innerhalb dieser Arbeit die bereits in allen modernen Fahrzeugen serienmäßig vorhandene Monokamera genutzt.

Die gesammelten Anforderungen an das System können Tabelle 1-1 entnommen werden. Diese legen einen Bereich fest, in dem die Funktion zuverlässig arbeiten soll. Diese Bereiche sind zum Teil fahrzeugbezogen und geben Aufschlüsse auf die notwendigen Fahrzeugdaten, die maximale Fahrzeugdynamik und die zu verwendende Hardware. Zum anderen wird die Fahrzeugumgebung beschrieben, in der die Funktion zum Einsatz kommt. Das sind bspw. spezielle Hinderniseigenschaften. Darüber hinaus werden Anforderungen an den Betrieb der Funktion und ihre Integration im Fahrzeug gegeben.

Tabelle 1-1 Anforderungen an das Gesamtsystem

Anforderung	Hintergrund
Verwendung von seriennaher videobasierter Sensorik	Dies ermöglicht eine kostengünstige Integration der Funktion in bereits bestehende Systeme.
Beliebiges statisches Hindernis	Mit der Funktion soll eine Vielzahl an möglichen Hindernissen detektiert werden. Diese sind in den meisten Fällen statisch (Tunneleinfahrt, Schild, Rohre, ...) oder bewegen sich nur minimal (Baumkronen bei leichtem Wind, ...). Dynamische Hindernisse werden nicht berücksichtigt.
Integration einer Eigenbewegungsprädiktion	Detektion des kritischsten Hindernisses zur Reduktion von falschen Warnausgaben.
Berücksichtigung des Verlaufs des Fahruntergrundes	Gewährleistung der exakten Hindernisberechnung in Kuppen und Wannen.
Signale der Fahrzeugvernetzung stehen zur Verfügung	Die fahrdynamischen Parameter des Fahrzeugs bieten die Grundlage für die Berechnung und Prädiktion der Eigenbewegung.
Funktion ohne satellitengestützte Navigationssoftware	Je nach Ausstattungsvariante des Fahrzeugs ist eine Satellitennavigation nicht zwangsweise verbaut.
Maximale Geschwindigkeit von 60 km/h	Die Funktion soll im urbanen Raum zur Anwendung kommen. Durchfahrten auf Überlandstraßen und Autobahnen werden nicht berücksichtigt.
Maximale Querbeschleunigung von 3.5 m/s ²	Die Funktion soll auch in Kurvenfahrten zur Verfügung stehen.
Hindernisdetektion nach Funktionsneustart innerhalb von 5 m	Ein schnelle Initialisierung ist erforderlich, um nach dem Funktionsstart bereits nah zum Fahrzeug liegende Hindernisse zu detektieren.
Höhengenauigkeit von 5 cm im Nahbereich	Bei zu großer aber auch zu kleiner Toleranz erfolgt keine Akzeptanz durch Fahrer. Dem System wird nicht vertraut.
Entfernungsgenauigkeit von max. 2 m bei Erstdetektion	Die Hindernisentfernung fließt in die Warnstrategie ein und hat direkten Einfluss auf die Auslösung des automatischen Notbremsassistenten.
Implementierung einer Warnstrategie	Das detektierte Hindernis muss dem Fahrer als Rückmeldung angezeigt und die mögliche Kritikalität bewertet werden. Erfolgt keine Reaktion durch den Fahrer, wird eine automatische Notbremsung durchgeführt.

Fortsetzung Tabelle 1-1 Anforderungen an das Gesamtsystem

Robustheit	Die Funktion soll bei einer Vielzahl von verschiedenen Hindernissen und Umwelteinflüssen einsetzbar sein.
Präzision	Das Hindernis muss mit einer hohen Genauigkeit detektiert werden, um den Fahrer optimal zu warnen und einen Unfall zu vermeiden.
Echtzeitfähigkeit	Das System muss echtzeitfähig im Fahrzeug integrierbar sein, sodass auf Änderungen im Verkehr rechtzeitig reagiert werden kann.
Langzeitstabilität	Die Funktion soll über die gesamte Dauer des Fahrzeugbetriebs zur Verfügung stehen.

1.2.2 Systemüberblick

Das in dieser Arbeit entwickelte System zur Durchfahrtshöhenerkennung soll hier kurz vorgestellt werden. Dabei wird auf die einzelnen Komponenten eingegangen und deren Hintergrund erläutert. Zunächst erfolgt die Unterteilung in das Fahrzeug- und das Referenzsystem. Das Fahrzeugsystem zielt auf die Integration der finalen Funktion zur Berechnung von Durchfahrtshöhen ab und soll so im Fahrzeug möglichst seriennah für den Endnutzer implementiert werden. Das Referenzsystem steht während der Entwicklungsphase zur Verfügung und dient als Grundlage für die Bewertung des Fahrzeugsystems.

Das Fahrzeugsystem kann in vier übergeordnete Bereiche unterteilt werden. Grundlegend für die Funktion ist die Nutzung der im Fahrzeug verbauten Frontkamera. Dies ist eine kalibrierte perspektivische Kamera, die hinter der Frontscheibe auf Höhe des Rückspiegels verbaut ist. Die Odometrie, d.h. die Eigenbewegung des Fahrzeugs, wird mittels fahrzeuginternen Signalen und Kamerainformationen berechnet, sodass jedem Bild der Bildfolge eine Pose (Position und Lage) zugewiesen werden kann. Im anschließenden Bereich der 3D-Rekonstruktion wird mithilfe dieser Bildfolge ein 3D-Abbild des nahen Fahrzeugumfelds geschaffen. Die verschiedenen Verfahren werden im Laufe der Arbeit vorgestellt. Das 3D-Umfeld lässt sich im nachfolgenden Schritt in verschiedene Klassen unterteilen. So erfolgt eine Fahruntergrunddetektion und -modellierung, um auch in Kuppen und Wannen Hindernishöhen berechnen zu können. Es erfolgt eine Klassierung in Struktur und potentielle Hindernisse. Hindernisse weisen eine lichte Höhe zum Fahruntergrund auf. Umgebungsstruktur, bspw. andere Fahrzeuge oder Hauswände, hat direkten Kontakt zum Fahruntergrund und wird in der Warnausgabe vernachlässigt. Zur Einschränkung des Hindernissuchraums wird eine Fahrschlauchprädiktion implementiert. Die Höhenberechnung, der Vergleich mit der Fahrzeughöhe und die Warnung erfolgt ausschließlich auf Hindernisse im prädizierten Fahrschlauch. Wird die Warnung auf ein potentiell zu niedriges Hindernis vom Fahrer ignoriert, so erfolgt der automatische Bremsengriff bis hin zu einer Notbremsung.

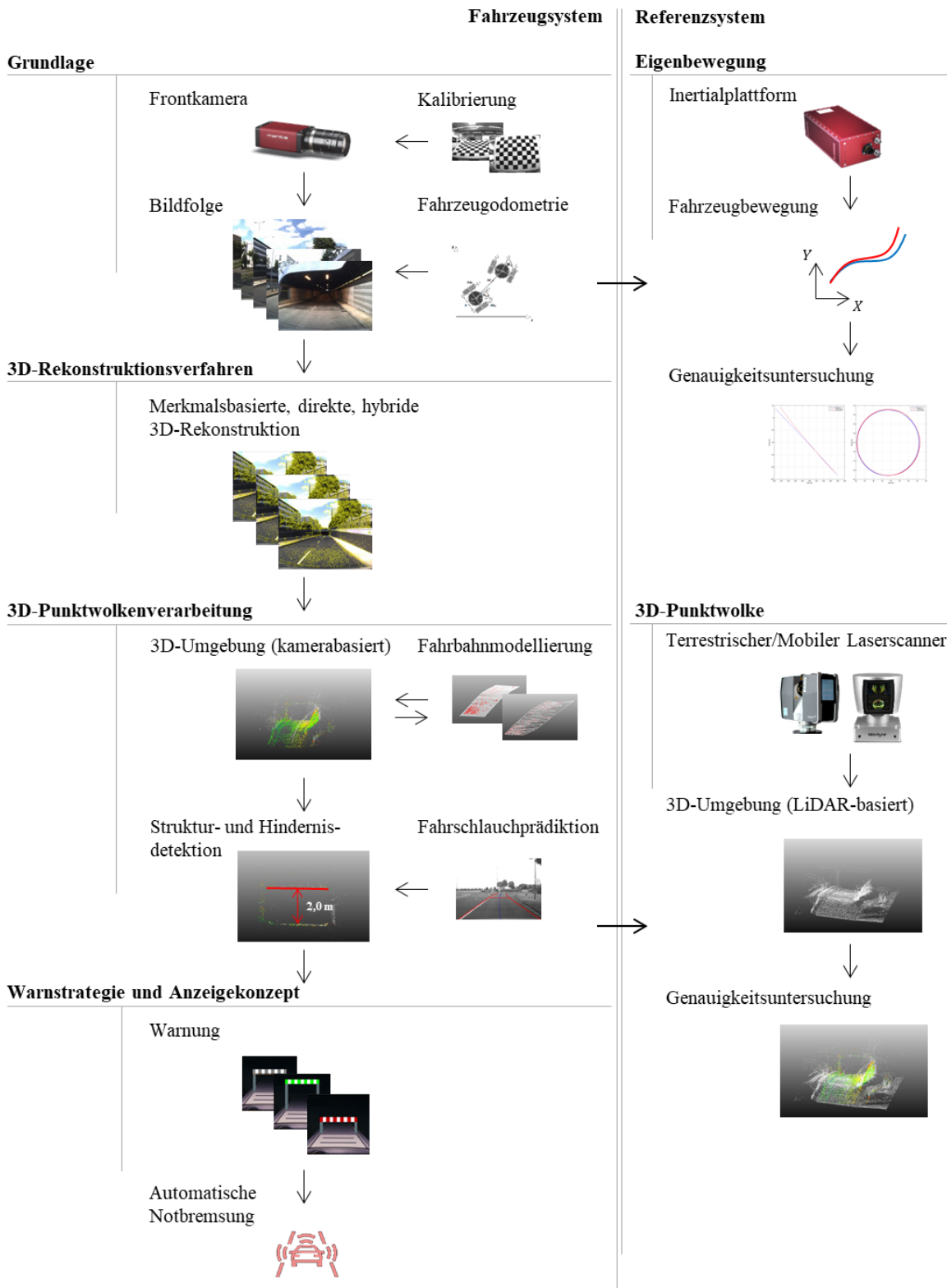


Abbildung 1-9 Durchfahrtshöhenerkennung – Systemüberblick

Das Referenzsystem lässt sich in zwei Bereiche unterteilen. Die Ermittlung der fahrzeugeigenen Odometrie geschieht über die Verwendung der Raddtickinformationen in Verbindung mit einem Bewegungsmodell. Mithilfe der Inertialplattform OXTS RT3002, welche die Fahrzeugbewegung hochgenau misst, soll die Radodometrie optimiert und mögliche Fehlerquellen identifiziert werden.

Zum Referenzieren der Hindernishöhen werden zwei separate LiDAR-Systeme eingesetzt. Zum einen wird der terrestrische Laserscanner Faro Focus3D X 130 HDR zur hochgenauen Vermessung spezieller Hindernisse genutzt. Zum anderen wird der mobile Laserscanner Velodyne HDL64e S3 auf dem Dach des Versuchsfahrzeuges befestigt und liefert während der Fahrt exakte 3D-Punkte, die mit der Punktwolke aus den Kameradaten verglichen werden können. Mithilfe der Laserscannerdaten können Analysen zur Genauigkeit der fahrzeugsystemseitigen Berechnung der lichten Höhe durchgeführt werden.

1.3 Zielsetzung

Das Ziel dieser Arbeit ist die Entwicklung einer Wirkkette zur Detektion von potentiellen Hindernissen mit eingeschränkter lichter Höhe. Das zu entwickelnde Low-Cost System soll allein auf der bereits in modernen Fahrzeugen integrierten Front-Monokamera und den vorhandenen Fahrdynamiksensoren basieren. In der vorliegenden Arbeit soll Schritt für Schritt die Wirkkette aufgebaut und bezüglich ihrer Robustheit und Genauigkeit untersucht werden. Unter Hinzunahme von Bildsequenzen aus typischen Fahrsituationen und ausgewählter Referenzsensorik werden entscheidende Fehlereinflussfaktoren innerhalb der Verarbeitungskette identifiziert und bewertet. Resultat der Untersuchung ist die Bestimmung der Genauigkeit der Höhenberechnung für eine Vielzahl unterschiedlicher Hindernisse und Fahrsituationen. Damit einher geht auch die Abschätzung, inwiefern ein kamerabasiertes System zur Detektion von Durchfahrtshindernissen im Fahrzeug integriert werden kann und welche Verfahrensgrenzen bestehen.

1.4 Gliederung der Arbeit

Zunächst wird in Kapitel 2 auf die technischen, mathematischen und photogrammetrischen Grundlagen eingegangen. Dabei werden die verwendete Fahrzeugsensorik zur Realisierung der Fahrerassistenzfunktion und die Referenzsensorik, die zur Evaluierung der Funktion dient, aufgeführt. Nach einer Diskussion über die verschiedenen Techniken der 3D-Rekonstruktion aus einer Bildfolge werden relevante mathematischen Hintergründe dazu näher erläutert.

In Kapitel 3 wird verstärkt auf die für die Funktion relevante Kamerakalibrierung eingegangen, verschiedene Vergleiche zwischen 2D- und 3D-Testfeldkalibrierungen aufgeführt, der Einfluss von wechselnden Umgebungstemperaturen untersucht und eine Abschätzung der Relevanz der Kalibrierung auf die Funktionsgenauigkeit gegeben.

Die Lokalisierung im Weltkoordinatensystem steht in Kapitel 4 im Vordergrund. Die fahrzeuginterne Sensorik wird hierbei evaluiert, deren Fehlereinfluss bestimmt und verschiedene Optimierungsstrategien erläutert. Ein wesentlicher Faktor ist die Fahrzeugodometrie nicht nur im Rahmen der Lokalisierung, sondern auch in der Skalierung des rekonstruierten 3D-Modells und der Entwicklung der Fahrschlauchprädiktion, auf die innerhalb der Arbeit eingegangen wird.

Kapitel 5 stellt verschiedene Ansätze vor, die es ermöglichen eine 3D-Szene aus einer Bildfolge zu erzeugen. Dabei werden Detektion, Verfolgung und Rekonstruktion indirekter Merkmale, wie Bildpunkte und Linien, untersucht.

In Kapitel 6 wird zudem auf die Integration einer direkten Methode eingegangen und deren Eignung in der vorliegenden Automotive-Anwendung geprüft.

Neben der Erzeugung der 3D-Punktwolke steht im weiteren Verlauf der Arbeit, in Kapitel 7, die Extraktion der potentiellen Hindernisse im Vordergrund. Dabei wird auf verschiedene Methoden

eingegangen die Punktwolke von fehlerhaften 3D-Punkten zu befreien und in einer geeigneten Struktur zu unterteilen. Die Extraktion des Fahruntergrundes und die Definition von Umgebungsstruktur und Hindernissen stehen dabei im Vordergrund.

Die Analyse der einzelnen Funktionsbausteine im Hinblick auf die Genauigkeit der Höhenhindernisse, die Zuverlässigkeit der Detektion und Stabilität der Gesamtfunktion wird in Kapitel 8 vorgenommen.

Ein Konzept zur Stabilisierung der Hindernisausgabe, der Anzeige der Hinderniskritikalität und die Entwicklung einer Warnstrategie wird in Kapitel 9 vorgestellt.

Zudem werden die Ergebnisse der hier vorgenommenen Untersuchung in Kapitel 10 zusammengefasst und mögliche Ansätze zur zukünftigen Optimierung und Erweiterung der Funktion in Kapitel 11 betrachtet.

2 Grundlagen

2.1 Technische Grundlagen

2.1.1 Fahrzeug- und Referenzsensorik

Innerhalb dieses Unterkapitels soll eine Bewertung der Sensorik in Bezug auf ihren Einsatz im Fahrzeug hinsichtlich der final definierten Funktion und deren Rahmenbedingungen aufgezeigt werden. Darüber hinaus wird die Referenzsensorik vorgestellt.

Nach [29] kann die Sensorik für Fahrerassistenzsysteme in zwei Kategorien unterteilt werden. Zum einen sind dies die Fahrdynamiksensoren, zum anderen die Umfeldsensoren. Um die Durchfahrtshöhenerkennung funktional optimal auf die vorhandenen Fahrzeugplattformen abzustimmen, sollen im Folgenden die benötigte Sensorik aus beiden Kategorien vorgestellt werden. Dabei steht im Fokus, eine möglichst genaue und zuverlässige Funktion zu schaffen, die auf einer geringen Anzahl von möglichst kostengünstigen Sensoren basiert.

Fahrdynamiksensoren

Im Rahmen der Fahrerassistenzsystementwicklung zählen in diese Kategorie Sensoren, mit denen die Bewegung des Fahrzeugs gemessen oder berechnet werden kann. Dies sind z.B. Sensoren zur Bestimmung der Raddrehzahl, des Lenkradwinkels, des Bremsdrucks, Drehraten- und Beschleunigungssensorik. Die Sensorwerte werden auf dem fahrzeugeigenem CAN (Controller Area Network) verschiedenen Steuergeräten zur Verfügung gestellt und können somit abgegriffen und für die Funktion genutzt werden.

In diesem Rahmen wird erneut darauf verwiesen, dass die Funktion unabhängig von einem satellitengestützten Ortungssystem arbeiten soll, sodass ausschließlich Fahrdynamiksensoren für die Eigenbewegungsbestimmung genutzt werden.

Umfeldsensoren

Diese Art von Sensoren liefert wichtige Informationen über das Fahrzeugumfeld und soll es ermöglichen kritische Hindernisse für die Funktion zu identifizieren. Ziel ist es die genutzte Sensorik so zu bestimmen, dass bereits vorhandene Sensoren von anderen Fahrerassistenzsystemen präferiert werden. Im Folgenden sollen die Basistechnologien vorgestellt und auf ihre Eignung zur Realisierung der Funktion bewertet werden.

Ultraschall

Ultraschallsensoren werden aktuell vorwiegend zur Erfassung von Hindernissen während des Einparkvorgangs genutzt. Sie werden an Front, Heck und an der Seite vor allem im unteren Bereich des Fahrzeugs angebracht. Sie basieren auf dem Puls-Echo Verfahren, wobei Schallimpulse ausgesandt und von in der Nähe befindlichen Hindernissen reflektiert werden. Die Sensoren empfangen die Reflektion und unter Annahme der Schallgeschwindigkeit kann über Laufzeitmessung der relative Abstand zum Objekt gemessen werden. Darüber hinaus werden mehrere Sensoren untereinander verschaltet, sodass auch indirekte Reflektionen detektiert werden können. Durch eine Reichweite von maximal 5 m und aufgrund ihrer langsamen Reaktionszeit eignen sich diese Sensoren ausschließlich für Anwendungen im Niedergeschwindigkeitsbereich bis 15 km/h.

RADAR

Das RADAR (Radio Detection and Ranging) wird überwiegend für Assistenzsysteme der Längsführung genutzt und bietet bspw. die Grundlage für die automatische Distanzregelung ACC (Adaptive Cruise Control) oder die Notbremsfunktion. Es arbeitet nach dem Prinzip des Doppler-Effekts, wobei eine elektromagnetische Welle bestimmter Frequenz ausgesendet wird, die sich nach der Reflektion an einem Objekt ändert. Der Frequenzunterschied zwischen ausgesendeter und empfangener Welle ist dabei ein Maß für die Relativgeschwindigkeit. Über die Laufzeit kann der Abstand berechnet werden. Durch Anordnung mehrerer Antennen in einer Halbmondform kann eine Winkelauflösung erreicht werden.

Das RADAR wird am unteren Teil der Fahrzeugfront angebracht und verfügt je nach Ausführung (Short-, Mid- oder Long-Range RADAR) über einen horizontalen Öffnungswinkel von bis zu 90°, vertikalen Öffnungswinkel von bis zu 10° und Reichweiten bis zu 250 m. Es eignet sich aufgrund seiner Mehrzielfähigkeit vor allem zur Detektion von beweglichen Objekten wie Fahrzeuge, Radfahrer und Fußgänger.

Videobasierte Sensorik

Die Frontkamera ist ein bildgebender Sensor, welcher in Fahrzeugmitte hinter der Frontscheibe angebracht und vor allem für die Funktionen der Spurhaltung und Verkehrszeichenerkennung genutzt wird. Darüber hinaus wird die Kamera zur Objektklassifizierung genutzt, wodurch zwischen Fahrzeugen, Fußgängern, Radfahrern, oder statischen Objekten wie Schildern, Fahrbahnmarkierungen und Gebäuden unterschieden werden kann. Diese Informationen stellen einen großen Mehrwert gegenüber anderer Sensoriken dar, wodurch Kameras eine einfache und gleichzeitig vielfältige Sensorbasis darstellen. Größter Nachteil ist jedoch die Abhängigkeit von Umwelteinflüssen. Während ungünstige Beleuchtungsverhältnisse wie Dunkelheit, Schlagschatten und Gegenlicht oft durch eine hohe Dynamik und intelligente Aufnahmestrategien ausgeglichen werden können, stellen Verschmutzungen, Regen und Nebel eine bedeutend größere Herausforderung dar.

Neben der meist perspektivischen Frontkamera gibt es weitere Ausführungen im Automotive-Bereich. Diese sollen kurz erläutert werden, stellen jedoch speziell im leichten Nutzfahrzeugsegment aufgrund ihrer hohen Systemkosten eine Ausnahme dar und werden deshalb nicht für die Funktion der Durchfahrtshöhenerkennung berücksichtigt. Fischaugenkameras werden zur Beobachtung des nahen Umfelds während des Einparkvorgangs genutzt. Aufgrund ihrer großen horizontalen und vertikalen Öffnungswinkel ergeben sich Überlappungen von Bildern mehrerer Kameras. Durch Stitching kann ein Bild der Vogelperspektive des Fahrzeugumfelds dem Fahrer angezeigt werden, in dem Distanzen zu umliegenden Objekten eingeschätzt werden können. Mehrkamera- bzw. Stereokamerasysteme kommen heute vor allem im hochpreisigen Segment zum Einsatz und ermöglichen die direkte Bestimmung von 3D-Koordinaten im Umfeld des Fahrzeugs. Sogenannte Time-of-Flight Kameras ermöglichen die Verknüpfung von Bildinformationen mit einem modulierten optischen Signal, wodurch sich zu jedem Pixel eine nach dem Laufzeitverfahren bestimmte 3D-Position berechnen lässt [29]. Infrarotkameras werden im Automotive-Bereich für Nachtsichtsysteme eingesetzt und in zwei Ausführungen integriert. Während das Nahinfrarot eine zusätzlich im infraroten Bereich liegende Lichtquelle benutzt und somit die Reflektion von Objekten ausgewertet werden, registriert das Ferninfrarot Wärmestrahlung und hebt somit vor allem Lebewesen wie Fußgänger oder Tiere hervor. Es eignet sich jedoch nur bedingt für die Detektion von statischen Objekten wie Einfahrten oder Unterführungen.

LiDAR

Das LiDAR ist ein optisches System, dessen Wirkprinzip dem des RADARs ähnelt. Anstatt der Nutzung von Mikrowellen werden sichtbare, ultraviolette oder infrarote Wellen ausgesendet. Sie werden meist mehrstrahlig ausgeführt und eignen sich zur hochgenauen Abstandmessung. Der Laserscanner nutzt das Prinzip des LiDARs. Dabei erfolgt jedoch eine Ablenkung der Strahlen mittels rotierendem Spiegel. Dies ermöglicht neben einer hohen Reichweite auch große horizontale Öffnungswinkel sowie verbesserte Winkelgenauigkeit. Laserscanner kommen heute vor allem in zwei Ausführungen im Fahrzeugbereich vor. Hochpreisige Fahrzeuge verfügen über mehrzeilige Laserscanner mit horizontalen Öffnungswinkel bis ca. 145° im Frontbereich. In der Entwicklung von automatischen Fahrzeugen werden vorrangig ein oder mehrere 360° Laserscanner im Dachbereich eingesetzt. Mechanisch scannenden Systeme sind jedoch vergleichsweise teuer und stör anfällig bei schlechter Witterung. Die Zukunft der Technologie kann im Einsatz von sog. Solid-State LiDARs liegen, welche es ermöglichen auch ohne rotierenden Spiegel ein 3D-Abbild der Umgebung zu erzeugen [30].

In der nachfolgenden Tabelle werden gängige Technologien und ihre Eignung für die hier zu entwickelnde Funktion gegenübergestellt. Die technischen Kenngrößen orientieren sich dabei an derzeit in Serie befindlichen Sensoren.

Tabelle 2-1 Vergleich der Umfoldsensorik

	Ultraschall	RADAR	Frontkamera	LiDAR
Messdistanz	5 m	160 m	50 m	150 m
Horizontaler Öffnungswinkel	90°	12° - 90°	55°	110° - 360°
Vertikaler Öffnungswinkel	90°	13°	45°	3.2° - 26.8°
Sensorausgabe	Laufzeiten der Reflektionen; vorverarbeitete Objektliste	Laufzeiten und Winkel der Reflektionen; vorverarbeitete Objektliste	RGB-Bildfolge	Vorverarbeitetes 3D-Umfeld des Fahrzeugs
Integrationsaufwand und Wirtschaftlichkeit	+	+	+	-

Ultraschallsensoren eignen sich aufgrund der geringen Messdistanz nur für Funktionen im niedrigen Geschwindigkeitsbereich und erfüllen damit nicht die Anforderungen an das Gesamtsystem, welches für den urbanen Raum bis 60 km/h ausgelegt werden soll. Das RADAR eignet sich durch Ausnutzung des Doppler-Effekts vorrangig zur Detektion von dynamischen Objekten wie Fahrzeugen oder Radfahrern. Darüber hinaus ist Detektionsstärke abhängig vom Objektmaterial und zu gering bei typischen Hindernissen, welche aus Kunststoff, Holz oder Stein bestehen können. Das LiDAR stellt aus messtechnischer Sicht den idealen Sensor zur Detektion von Höhenhindernissen dar. Wird der Nachteil des geringen vertikalen Öffnungswinkels durch eine Verortung in der Nähe des Fahrzeugdachs ausgeglichen, können hochgenaue 3D-Punktwolken direktmessend erzeugt und Höhenhindernisse extrahiert werden. Allein die hohen Systemkosten unterbinden derzeit den Einsatz eines LiDARs im leichten Nutzfahrzeug. Insofern wird in der vorliegenden Arbeit ein Algorithmus

entwickelt und bewertet, der die Bildfolge der perspektivischen Frontkamera nutzt, um eine 3D-Punktwolke des Umfelds zu rekonstruieren. Darin werden kritische Hindernisse identifiziert und dem Fahrer angezeigt. Die für die Funktionsentwicklung ausgewählte Industriekamera Allied Vision Manta G-125C ist an die technischen Spezifikationen einer Serienfahrzeugkamera angelehnt.

Tabelle 2-2 Kamerakonfiguration

	Eigenschaft
Modell	AVT Manta G-125C
Bildsensor	Sony ICX445 CCD 1/3 Zoll
Bildgröße [Pixel]	1292x964
Pixelgröße [μm]	3.75
Max. Bildwiederholfrequenz [fps]	30
Objektiv	KOWA LM4NCL
Brennweite [mm]	3.5
Öffnungswinkel	77° horizontal, 57.7° vertikal

Die vorrangig in der Industrie verwendete Allied Vision Manta G-125C besitzt einen 1/3 Zoll RGB-Sensor. Mittels GigaEthernet Datenschnittstelle werden die Bilder mit 30 Hz im Format 1292x964 Pixel aufgezeichnet. Die Kamera besitzt ein KOWA Objektiv mit einer festen Brennweite von 3.5 mm und ist hinter der Frontscheibe in Höhe des Rückspiegels angebracht.

Referenzsensorik

Neben der Fahrzeugsensorik, welche für die finale Funktion im echten Serienbetrieb genutzt wird, kommt im Entwicklungsprozess weitere Sensorik für das Referenzieren zum Einsatz. Es handelt sich dabei um hochgenaue Systeme, die es ermöglichen eine Bewertung der Serienfunktion bzgl. Güte und Robustheit zu geben. Grundprinzip der Funktion ist der *SfM* Ansatz (Structure from Motion), bei dem neben der Kartierung eine gleichzeitige Lokalisierung durchgeführt wird. Insofern werden die Referenzsysteme so gewählt, dass zum einen die Eigenbewegung des Fahrzeugs und die 3D-Punktwolke der Umgebung bewertet werden können.

Referenz für die Eigenbewegungsschätzung

Die im Fahrzeug genutzte Sensorik zum Referenzieren der Eigenbewegungsschätzung ist die OXTS RT3002. Sie kombiniert eine Inertialplattform (engl. Inertia Measurement Unit – IMU) und dGPS (differentielles GPS).



Abbildung 2-1 OXTS RT3002 [31]

Die IMU besitzt drei Beschleunigungs- und drei Drehratensensoren. Für die Positionierung im Weltkoordinatensystem (WKS) wird ein Zweiantennen-GPS genutzt, welches in der Lage ist nicht nur die Position und Geschwindigkeit des Fahrzeugs im WKS zu ermitteln, sondern auch dessen Ausrichtung. Durch die Verwendung eines GSM-Modems, welches Korrekturdaten über die Satellitenkonfiguration via Mobilfunk zur Verfügung stellt, wird die Genauigkeit der GPS Positionierung erhöht.

Um eine optimale Eigenbewegungsschätzung zu erhalten, werden die Daten der IMU und des dGPS mittels Kalman Filter fusioniert. Somit können systematische Fehler, wie z.B. der durch akkumulierte Fehler der Beschleunigungssensoren entstehende Drift in der Position des Fahrzeugs, korrigiert werden. Es lassen sich somit über lange Zeiträume sehr genaue Daten zur Eigenbewegung des Fahrzeugs berechnen.

Tabelle 2-3 Messgenauigkeit OXTS RT3002 [31]

	Messgenauigkeit
Position [m]	0.01
Geschwindigkeit [km/h]	0.05
Gierwinkel [°]	0.1
Nick-/Wankwinkel [°]	0.03
Beschleunigung [m/s ²]	0.01
Winkelraten [°/s]	0.01

Die RT3002 eignet sich aufgrund ihrer hohen Systemkosten allein zur Entwicklung und Bewertung von Fahrodynamikfunktionen, und wird so nicht in einem Serienfahrzeug verbaut. Eine kostengünstigere Alternative ist die Verwendung der bereits im ESC-Steuergerät berechneten Fahrodynamikenngrößen, die als Eingangsparameter für Bremsregelsysteme dienen.

Referenz für die 3D-Punktwolke

Die Ausgabe des kamerabasierten SfM-Verfahrens ist eine 3D-Punktwolke, welche zur Detektion der potentiellen Hindernisse genutzt wird. Die Genauigkeit der 3D-Punkte hat direkten Einfluss auf die Bestimmung der Hindernishöhe und ist somit entscheidend für die Realisierung und Integration der Gesamtfunktion. Zum Referenzieren der Punktwolke, im Besonderen der Hindernishöhen, werden innerhalb dieser Arbeit zwei verschiedene Laserscanner genutzt. Neben dem FARO Focus3D X 130 HDR, welcher zu den terrestrischen Laserscannern (TLS) gehört, wird der mobile Laserscanner (MLS) Velodyne HDL64e S3 verwendet. Beide Plattformen haben Vor- und Nachteile, die einen direkten Einfluss auf die jeweiligen Anwendungs- und Einsatzgebiete hat. Die Laserscanner werden im kommenden Absatz kurz vorgestellt und ihr Verwendungszweck als Referenzsensorik beschrieben.

Terrestrische bzw. statische Laserscanner werden zur Vermessung aus einem bzw. wenigen Standpunkten genutzt. Je nach Einstellung der Scanparameter können nahezu beliebig detaillierte Szenen in einem Zeitraum von wenigen Minuten bis Stunden erstellt werden. Durch das Zusammenführen mehrerer Scans aus unterschiedlichen Blickwinkeln können größere Bereiche abgedeckt und Verdeckungen durch Objekte vermieden werden. Zur Registrierung der Scans werden in der Szene positionierte Marker oder der in [32] beschriebene Iterative Closest Point (ICP) Algorithmus genutzt.



Abbildung 2-2 Referenzsensorik Laserscanner. Links: Faro Focus3D X 130 HDR [33]; Rechts: Velodyne HDL64e S3 [34]

Beim Faro Focus3D X 130 HDR werden Entfernungen zu Objekten mittels Phasendifferenzmessung bestimmt. Das dauerstrichmodulierte System eignet sich vorrangig für Scans mit Entfernungen kleiner 130 m und zeichnet sich vor allem durch eine hohe Genauigkeit aus. Es wird im Rahmen der Entwicklung zur Erzeugung hochgenauer Punktwolken von Fahrbahnen und Hindernissen genutzt. Ein weiterer Laserscanner, der zum Einsatz kommt, ist der Velodyne HDL64e S3. Dieser auf dem Prinzip der Laufzeitmessung basierende MLS wird auf dem Dach des Versuchsfahrzeugs angebracht. Dadurch wird es ermöglicht auch über längere Fahrtstrecken eine Referenzpunktwolke zu erzeugen. In der angefügten Tabelle werden die beiden Laserscanner gegenübergestellt und verglichen.

Tabelle 2-4 Vergleich des TLS und MLS

	FARO Focus3D X 130 HDR	Velodyne HDL64e S3
Anwendungsgebiet	Hochgenaue Scans von kleinen Gebieten im Innen- und Außenbereich	Erstellung von 3D-Stadtkarten, Echtzeit-Umfeldererkennung im Bereich des automatischen Fahrens
Datenerfassung	Einfache Datenaufzeichnung an festen Standpunkten	Kontinuierliche Datenaufzeichnung entlang der Fahrzeugtrajektorie
Funktionsprinzip	Dauerstrichmodulation mit Phasendifferenzmessung	Pulsmodulation mit Laufzeitmessung
Entfernung	0.6 - 130 m mit senkrechtem Einfallswinkel auf 90 % reflektierender Oberfläche	0.9 - 120 m mit senkrechtem Einfallswinkel auf 80 % reflektierender Oberfläche
Öffnungswinkel	360° horizontal, 300° vertikal	360° horizontal, 26.8° vertikal
Genauigkeit	2 mm Entfernung 0.009° Winkel	15 mm Entfernung 0.09° Winkel
Messgeschwindigkeit	max. 0.976 Mio. Punkte/s	max. 2.2 Mio. Punkte/s
Punktdichte	Abnehmend mit größerer Entfernung zum Scannerstandpunkt	Homogen entlang der Fahrzeugtrajektorie bei konstanter Geschwindigkeit



Abbildung 2-3 Beispielhafter Vergleich von TLS (oben) und MLS (unten) anhand eines typischen Hindernisses mit Messpunkten Links, Mitte und Rechts

Abbildung 2-3 zeigt die Punktwolken der beiden Systeme. Das statische System besitzt eine hohe Punktdichte bei gleichzeitig geringer Punktstreuung auf dem Hindernis und dem Fahruntergrund. Die mit dem mobilen System berechnete Punktwolke weist während der Bewegung des Fahrzeugs fehlerhafte Punktzuordnungen im ICP-Algorithmus auf, was sich in der größeren Punktstreuung in X - und Y -Richtung widerspiegelt. Dies wird anhand der zwei Stützpfeiler und Querverstrebungen an der Carportfront sichtbar. Beide Systeme lassen jedoch die Messung der Objekthöhe innerhalb von 1 cm zu und werden im Verlauf der Arbeit zur Fehlerbestimmung herangezogen.

Tabelle 2-5 Vergleich der berechneten Hindernishöhen zwischen TLS und MLS

Hindernishöhe [m]	TLS	MLS	Differenz
Messpunkt Links	2.687 m	2.694 m	0.007 m
Messpunkt Mitte	2.684 m	2.687 m	0.003 m
Messpunkt Rechts	2.718 m	2.729 m	0.011 m

Für schwer einsehbare oder besonders hohe Hindernisse wird zudem ein Laserdistanzmessgerät mit einer Genauigkeit von 2 mm verwendet.

2.1.2 Vernetzung und Systeme

Wie bereits im vorherigen Kapitel erläutert besitzt das Versuchsfahrzeug, ein Volkswagen T6, bereits serienmäßige Fahrdynamiksensoren, welche zur Funktionsentwicklung genutzt werden. Sensorwerte und Informationen werden über das Fahrzeugnetzwerk unter verschiedenen Steuergeräten (SG) ausgetauscht und können der Funktion zur Verfügung gestellt werden. Der elektrische bzw. elektronische Grundaufbau des Fahrzeugs stützt sich auf die SG-Vernetzung mittels CAN (Controller Area Network).

Controller Area Network

CAN ist ein speziell für die Fahrzeugvernetzung designedes Bussystem, welches von Bosch entwickelt wurde [35]. Der Datenaustausch zwischen Steuergeräten erfolgt prioritätenbasiert mit einer Geschwindigkeit von 1 Mbit/s. Die einzelnen Signale der Sensoren und berechneten Werte der Steuergeräte werden zu Botschaften zusammengefasst. Die CAN-Botschaften werden mit einem Identifier versehen, der eine Unterscheidung ermöglicht und gleichzeitig die Priorität der Botschaft vorgibt. Ein modernes CAN-Bussystem wird so ausgelegt, dass eine maximale Buslast von 50 % nicht überschritten wird, um eine „weiche“ Echtzeitfähigkeit aller Signale zu gewährleisten [36]. Auf die Daten wird mit sog. CAN-Transceivern (Hardware-Interface) zugegriffen und anschließend an einen Automotive-Echtzeitrechner weitergeleitet. Das im T6 integrierte CAN-Bussystem verfügt über drei separate CANs, welche über ein Gateway Daten austauschen.

Antriebs-CAN

Der Antriebs-CAN arbeitet mit einer Baudrate von 500 kbit/s und ist für die grundlegenden Fahrzeugfunktionen zuständig. Wichtige Teilnehmer auf diesem CAN sind das Motor-SG und das ESC-SG.

Komfort-CAN

Dieser CAN steht Funktionen des Infotainments und Komfortanwendungen zur Verfügung. Da diese Funktionen nicht echtzeitkritisch sind, werden Botschaften auf diesem CAN mit einer geringeren Baudrate (125 kbit/s) versendet. Teilnehmer des CAN sind unter anderem das Radio- und Navigationssystem, das Klima-SG und die Steuergeräte für elektrische Schiebetüren.

Kombi-CAN

Der Kombi-CAN arbeitet ebenso wie der Komfort-CAN auf 125 kbit/s und wird für die Anzeigen im Kombiinstrument sowie für weitere Funktionen wie Fahrtenschreiber und Fernlichtassistent genutzt.

Für die Funktionsrealisierung sind vor allem die Daten des Antriebs-CAN des Serienfahrzeugs von Nutzen, da hier die Informationen zu den jeweiligen Radticks und zum Lenkwinkel bereitgestellt werden und das Steuergerät zur automatischen Notbremsung (ESC-SG) angeschlossen ist. Die Daten des Bussystems können mit CAN-Transceivern abgegriffen und wieder in das Netzwerk zurückgespeist werden.

Konzept

Das Vernetzungskonzept beschreibt die Anbindung der benötigten Hardwarekomponenten im Serienfahrzeug. Neben den fahrzeugeigenen Daten auf dem Antriebs-CAN wird zusätzlich ein Referenz-CAN für die Inertialplattform OXTS RT3002 und Ethernet-Verbindungen für die Übermittlung der Rohdaten der Monokamera (AVT Manta G-125C) sowie des Laserscanners (Velodyne HDL64e S3) implementiert.

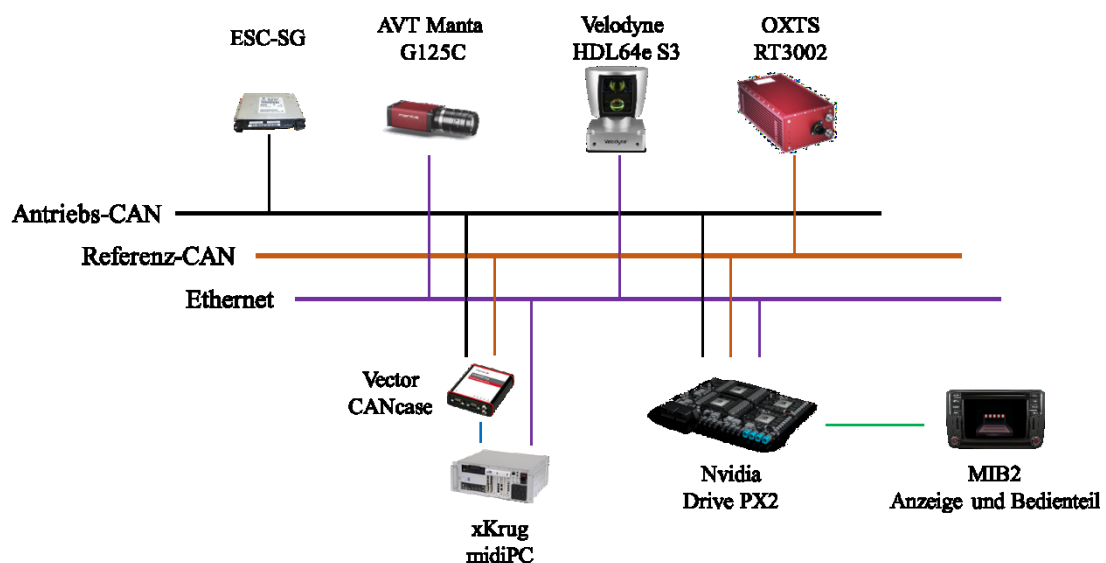


Abbildung 2-4 Vernetzungskonzept

Die Echtzeitfähigkeit des Systems ist ein entscheidender Faktor bei der Entwicklung der Assistenzfunktion. Das Delay zwischen Messung und Empfang der Messdaten darf die festgelegten Zeitkriterien nicht überschreiten. Innerhalb dieses Projekts werden dazu alle Sensoren an einen Echtzeitrechner (xKrug midiPC) angeschlossen. Dieser ist per USB mit einem Vector CANcase verbunden, welches die Daten der CAN-Netzwerke ausliest. Seriennahe Zielhardware für die Funktion ist das Nvidia Drive PX2 Entwicklungsboard. Dieses sendet per LVDS (Low-Voltage Differential Signaling) die Warnausgaben an das Anzeige- und Bedienelement MIB2 (Modularer Infotainmentbaukasten der 2. Generation). Etwaige Notbremsinformationen gelangen per Antriebs-CAN zum ESC-Steuergerät, welches eine automatische Bremsung einleitet.

Notbremsassistent

Der Notbremsassistent, oder auch AWW (Anhaltewegverkürzung), ist ein serienmäßiges RADAR-basiertes System, welches in kritischen Situationen die Aufmerksamkeit des Fahrers zurückgewinnen soll und im Notfall eine automatische Bremsung durchführt. Das System ist in vier Auslösestufen untergliedert, die Abbildung 2-5 entnommen werden können. In Phase 1 findet die optische und akustische Warnung des Fahrers sowie der Prefill der Bremsanlage statt. Erfolgt kein Eingreifen, so wird in Phase 2 ein Bremsruck ausgeführt. In Phase 3 beginnt die automatische Verzögerung des Fahrzeugs, wobei diese geschwindigkeitsabhängig ist. Sofern der Fahrer dennoch

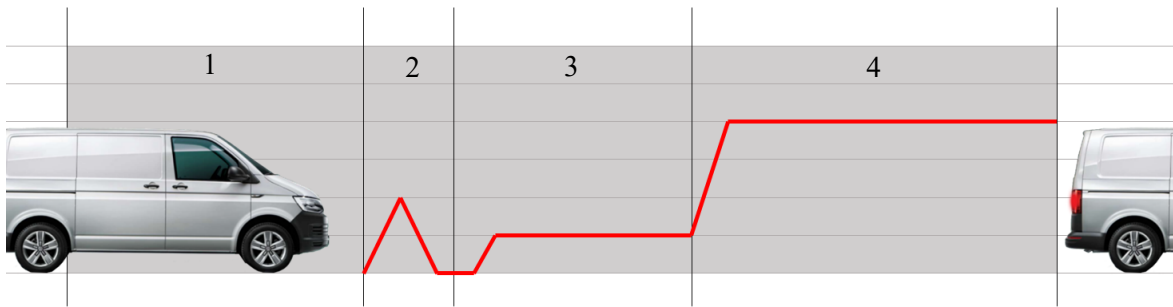


Abbildung 2-5 Schema: Notbremsassistent

nicht reagiert oder selbst zu schwach bremst, so geht das System in Phase 4 über, bei der der volle Bremsdruck abgerufen wird. Im System der Durchfahrthöhenerkennung werden die vom RADAR detektierten Objekte durch die Höhenhindernisse erweitert. Die Auslösung der einzelnen Phasen wird vor allem über die Berechnung der *TTC* (Time to Collision) mit

$$TTC = \frac{s_{Hindernis}}{v_{ego}} \quad \text{Gl. (2-1)}$$

bestimmt, wobei $s_{Hindernis}$ die Distanz zum Hindernis und v_{ego} die Fahrzeuggeschwindigkeit darstellen. Die Abstimmung des gewünschten Bremsverhaltens wird entscheidend über die Grenzwerte der *TTC* in den jeweiligen Phasen beeinflusst und so bestimmt, dass das Fahrzeug nach Durchführen einer Vollbremsung vor dem erkannten Hindernis zum Stehen kommt.

2.1.3 Software und Entwicklungsumgebung

Als simulative Entwicklungsumgebung wird ADTF (Automotive Data and Time-Triggered Framework) genutzt [37]. Das Tool dient der Vorentwicklung und Potentialabschätzung realer Assistenzsysteme. Diese erfüllen in frühen Entwicklungsstadien häufig noch keinen Sicherheitsanforderungen, sodass Tests in Realfahrzeugen nicht nur mit hohem Aufwand und Kosten verbunden sind, sondern auch Fahrer und Fahrumgebung gefährden können. ADTF ist modular aufgebaut und ermöglicht die Integration von vordefinierten aber auch die eigens in C++ programmierten Modulen. Die Daten können per Input- und Output-Pins an Module weitergegeben, aufgezeichnet und abgespielt werden. Ein wesentlicher Vorteil von ADTF ist die Verarbeitung der unterschiedlichen Daten unter einem gemeinsamen Zeitstempel, welche eine spätere Synchronisierung vereinfacht.

2.2 Kamerabasierte Rekonstruktionsverfahren

Grundlage für die Detektion von Hindernissen ist die in dieser Arbeit entwickelte Pipeline zur 3D-Rekonstruktion der Fahrzeugumgebung aus Bildinformationen. Neben einigen Analysen, die sich ausschließlich auf ein einziges Bild begrenzen [38, 39, 40], sind es vor allem die Zweibild- und Mehrbildauswertungen, die es ermöglichen eine hohe Detailtiefe und Genauigkeit innerhalb des 3D-Modells zu erzeugen. Bei diesen Verfahren wird das Objekt bzw. die Szene aus möglichst unterschiedlichen Blickrichtungen und Standpunkten aufgenommen und ein zusammenhängendes virtuelles Abbild erstellt. Die heutigen Verfahren zur Mehrbildrekonstruktion zeichnen sich durch einen hohen Automatisierungsgrad aus, wobei stets neue Verfahren und Pipelines für eine beständig wachsende Anzahl an Einsatzzwecken entwickelt werden. Die Anwendungsfelder reichen dabei von der bildbasierten Erzeugung genauer und realistischer Objektmodelle über die menschliche Bewegungsanalyse bis hin zum Einfügen von virtuellen Objekten in reale Szenen im Einsatzfeld der

Augmented Reality. Das Potential der kamerabasierten 3D-Rekonstruktion wird im Gebiet der Photogrammetrie sichtbar. Wesentlicher Vorteil des bildbasierten Ansatzes ist die Möglichkeit der schnellen und berührungslosen Vermessung von nahezu beliebig großen und komplexen Szenen und Objekten aus mehreren Dimensionen. Neben Anwendungen wie der Erstellung von digitalen Oberflächenmodellen aus Drohnen- oder Satellitenaufnahmen, der Generierung von 3D-Stadtmodellen, der Vermessung in der Land- und Forstwirtschaft, findet die Mehrbildauswertung immer mehr Einzug in Automotive-Anwendungen. Zu nennen sind dabei die Deformationsanalyse bei Crashtests, die bildbasierte Einmessung von Bauteilen im Produktionsprozess sowie die Überwachung der Werkzeugabnutzung. Ein weiteres Themengebiet ist die Entwicklung von modernen Fahrerassistenzsystemen. Unter der Nutzung von Ein- oder Mehrkameranensystemen wird ein 3D-Modell der Fahrzeugumgebung erstellt, welches zur Objektdetektion, Segmentierung oder Bewegungsanalyse herangezogen werden kann. Zur Rekonstruktion aus Bildern einer Monokamera, die auch hier verwendet wird, werden im Automotive-Bereich unterschiedliche Verfahren verwendet, die sich in wenigen Details unterscheiden.

Das Verfahren, das auch in der vorliegenden Arbeit Anwendung findet, ist *Structure from Motion* (SfM). Es beschreibt die Berechnung der Pose einer bewegten Kamera und die gleichzeitige Erschaffung eines 3D-Abbilds der Umgebung. Dies wird meist auf Grundlage von automatisch extrahierten Bildmerkmalen durchgeführt, welche über mehrere Bilder gematcht bzw. getrackt werden. Über eine nichtlineare Optimierung, der Bündelgleichung, werden durch Minimierung des Rückprojektionsfehlers die Qualität der 3D-Rekonstruktion, die Schätzung der Kamerabewegung und die Kalibrierung der Kamera verbessert. Frei verfügbare Tools, die eine solche automatisierte Pipeline zur Verfügung stellen, sind COLMAP [41] oder VisualSfM [42]. Die Entwicklung im Bereich SfM wurde in den letzten 30 Jahren so weit vorangetrieben, dass ausgewählte Algorithmen auch in kommerzieller Software Anwendung finden. Beispiele hierfür sind Agisoft Metashape oder Autodesk Recap. Diese Programme verfügen darüber hinaus über weitere Funktionen wie bspw. eine automatische Markerdetektion, Georeferenzierung und Kamerakalibrierung. Ein weiterer Ansatz, der viele Parallelen aufzeigt, ist *Simultaneous Localization and Mapping* (SLAM) Unter SLAM werden Verfahren zusammengefasst, die es einem Roboter, einer Drohne, einem Fahrzeug oder einfacherweise einer Kamera ermöglichen, sich in einer teils unbekanntenen Umgebung zu orientieren und gleichzeitig eine Karte des Terrains zu erstellen bzw. zu erweitern. Obwohl die Ansätze SfM und SLAM nahezu identisch sind, wird hier der Idee gefolgt, dass SLAM primär auf die Lokalisierung abzielendes Verfahren ist, wohingegen SfM größeren Fokus auf die genaue Rekonstruktion der Punktwolke legt. Darüber hinaus existiert der Ansatz des *Multiple View Stereo* (MVS). Darunter fallen Verfahren, die von bereits bekannten Kameraposen ausgehen und sich ausschließlich auf den Teil der Rekonstruktion fokussieren. MVS arbeitet mit einer Vielzahl an unterschiedlichen Bildern mit teils großen Standpunktvariationen und zielt auf eine möglichst dichte 3D-Rekonstruktion ab. Beispielhaft sind CMVS (Clustering View for Multi-view Stereo) und PMVS (Patch-based Multi-view Stereo) zu nennen [43, 44]. Die dichten photogrammetrischen Modelle eignen sich zur hochgenauen Vermessung von Objekten, erfordern jedoch eine sehr große Berechnungszeit und finden somit keine Anwendung innerhalb dieser Arbeit.

2.3 Mathematische und photogrammetrische Grundlagen

2.3.1 Koordinatensysteme

In der vorliegenden Arbeit werden vier verschiedene Koordinatensysteme verwendet, deren Eigenschaften im folgenden Absatz erläutert werden.

Für die absolute Positionierung und Ortung des Fahrzeugs wird das kartesische und feststehende Weltkoordinatensystem (WKS) verwendet.

Das Umfeldsensorik- oder Fahrzeugkoordinatensystem (USK) mit dem Index F ist ein sich mit dem Fahrzeug mitbewegendes Koordinatensystem, welches seinen Ursprung im Fußpunkt der Senkrechten durch die konstruktive Mitte der Fahrzeughinterachse auf der Fahrbahnoberfläche hat. Die einzelnen Sensorkoordinatensysteme im Fahrzeug werden relativ zum USK angegeben. Die Ausrichtung orientiert sich an der DIN 70000. Daraus geht ein rechtsdrehendes, orthogonales Koordinatensystem hervor, wobei die X -Achse in Fahrzeugrichtung, parallel der Längsachse des Fahrzeugs verläuft. Y zeigt nach links, Z nach oben.

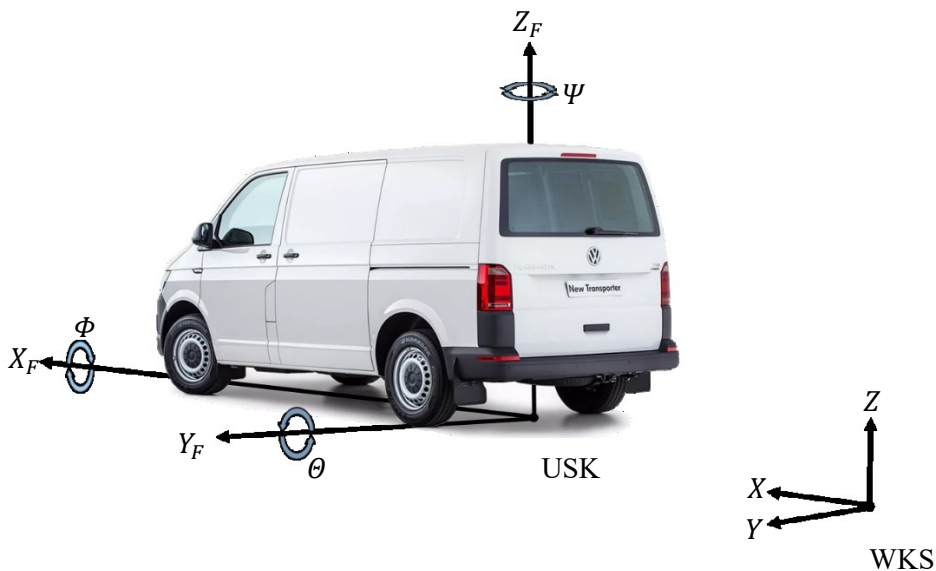


Abbildung 2-6 Definition von WKS und USK

Eine Drehung um X wird als Wanken bezeichnet und mit dem Winkel Φ beschrieben. Die Drehung um Y (Nicken) beschreibt der Winkel Θ . Das Gieren, welches eine Drehung um die Hochachse Z angibt, erhält das Formelzeichen Ψ . Die Reihenfolge der Rotationen, um eine gesamtheitliche Drehung durchzuführen, ist über die DIN 70000 definiert und lautet Ψ - Θ - Φ . Die Achsen des USK sind nicht karosseriefest und unabhängig von Nick- und Wankbewegungen.

Ein weiteres wichtiges Koordinatensystem ist das Kamerakoordinatensystem (KKS), dessen Ursprung im Projektionszentrum der Kamera liegt. Über die Kalibrierung der äußeren Orientierung wird die Kamerapose (Position und Lage) relativ zum USK bestimmt.

Das PKS (Pixelkoordinatensystem) hat seinen Ursprung in der oberen linken Ecke des Kamerabildes. Es werden darüber die Positionen einzelner Pixel auf dem Bild definiert. Die u -Achse zeigt dabei nach rechts, die v -Achse nach unten und die Einheit ist Pixel. Transformationen zwischen PKS und KKS werden über das geometrische Modell bestimmt.

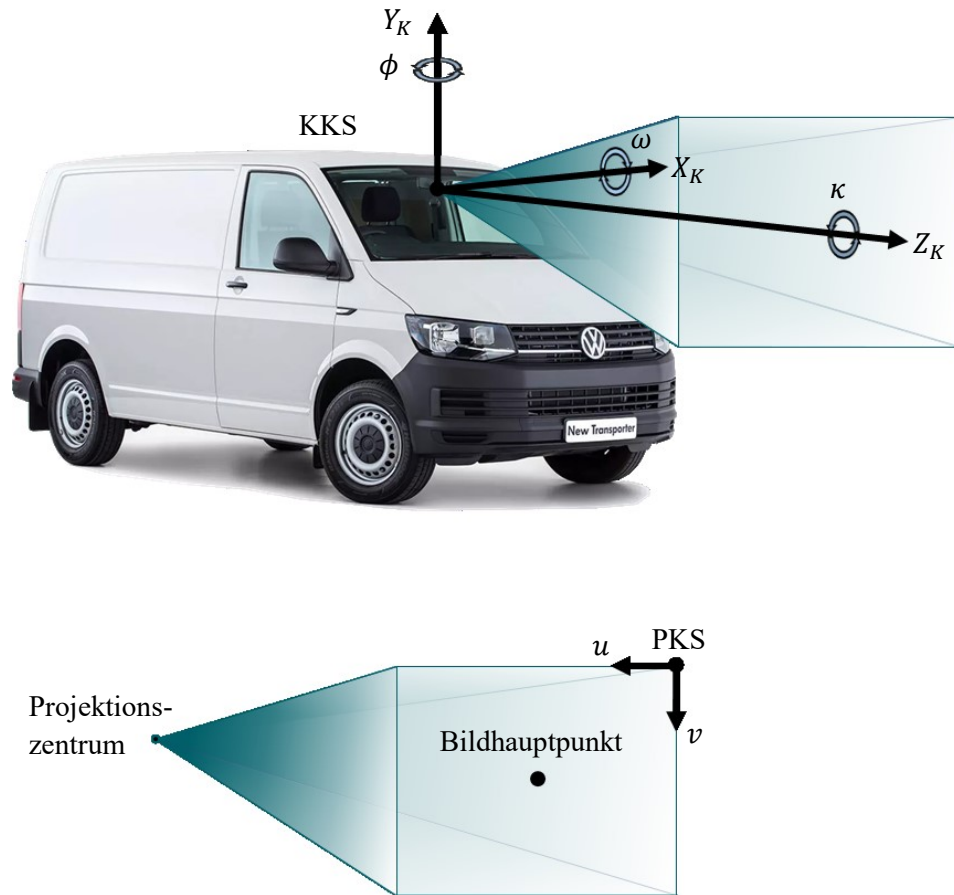


Abbildung 2-7 Definition von KKS und PKS

Die verwendeten Laserscanner zur Referenzmessung arbeiten nach ihrer Aufnahmekonfiguration in einem dreidimensionalen sphärischen Koordinatensystem. Die berechneten Koordinaten des terrestrischen Laserscanners werden unter Nutzung der mitgelieferten Software in ein dreidimensionales kartesisches Koordinatensystem umgewandelt. Die Objektkoordinaten des auf dem Dach befestigten mobilen Laserscanners erfordern zum Online Vergleich mit den Kameraobjektkoordinaten zudem eine Transformation in das mitbewegte Fahrzeug-USK.

2.3.2 Projektive Geometrie

Die projektive Geometrie im Raum \mathbb{P} wird im Rahmen dieser Arbeit für die Darstellung der perspektivischen Projektion eingeführt. Sie wird nach [45] als einfache Erweiterung der euklidischen Geometrie im Raum \mathbb{R} gesehen und ermöglicht die Formulierung des nichtlinearen Abbildungsprozess mittels linearer mathematischer Beziehungen unter der Verwendung homogener Koordinaten. Die im Folgenden beschriebenen Operationen werden dabei im \mathbb{P}^3 -Raum dargestellt.

2.3.2.1 Homogene Koordinaten

Homogene Koordinaten repräsentieren geometrische Elemente wie Punkte, Linien, Flächen und Transformationen im projektiven Raum. Die Koordinaten eines Raumpunktes werden dabei um den Skalierungsfaktor W erweitert.

$$\mathbf{X}_h = \begin{bmatrix} X_h \\ Y_h \\ Z_h \\ W \end{bmatrix} \quad \text{Gl. (2-2)}$$

Unter dem Sonderfall $W = 1$ ergeben sich die kartesischen Koordinaten des Punktes

$$\mathbf{X} = \begin{bmatrix} X_h/W \\ Y_h/W \\ Z_h/W \\ W/W \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \quad \text{Gl. (2-3)}$$

Über $W = 0$ können Punkte im Unendlichen beschrieben werden.

2.3.2.2 Projektive Transformationen

Mithilfe homogener Koordinaten können eine Vielzahl an Transformationen im projektiven Raum innerhalb eines linearen Gleichungssystems durchgeführt werden. Allgemein kann eine Transformation mit

$$\mathbf{X}^* = \mathbf{H} \cdot \mathbf{X} \quad \text{Gl. (2-4)}$$

dargestellt werden, wobei \mathbf{H} die Transformations- bzw. Abbildungsmatrix beschreibt.

$$\mathbf{H} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \quad \text{Gl. (2-5)}$$

Eine wichtige Transformation innerhalb dieser Arbeit ist die Ähnlichkeitstransformation. Sie setzt sich aus der Translation \mathbf{H}_T , Rotation \mathbf{H}_R und Skalierung \mathbf{H}_S mit dem Maßstab m zusammen und besitzt 7 Freiheitsgrade. Sie findet Anwendung im Rahmen der Punktwolkenrekonstruktion und -verarbeitung.

$$\mathbf{X}^* = \mathbf{H}_T \cdot \mathbf{H}_S \cdot \mathbf{H}_R \cdot \mathbf{X}$$

$$\begin{bmatrix} X^* \\ Y^* \\ Z^* \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & X_0 \\ 0 & 1 & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m & 0 & 0 & 0 \\ 0 & m & 0 & 0 \\ 0 & 0 & m & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \text{Gl. (2-6)}$$

Mehrere Transformationen können aneinander gekettet und zusammengefasst werden. Handelt es sich dabei um eine nicht singuläre Abbildungsmatrix, kann auch eine Rücktransformation erfolgen. In dieser Arbeit von großer Bedeutung ist die Abbildung eines 3D-Objektpunkts (\mathbb{P}^3) auf einen korrespondierenden 2D-Punkt in der Bildebene (\mathbb{P}^2). Diese Abbildung kann mithilfe der Gleichung

$$\mathbf{x}' = \mathbf{P}\mathbf{X} \quad \text{Gl. (2-7)}$$

berechnet werden, wobei \mathbf{P} eine 3x4-Matrix darstellt, die auch Kameramatrix genannt wird. Diese Matrix besitzt 11 Freiheitsgrade, die sich aus den Parametern der inneren und äußeren Orientierung der Kamera zusammensetzen.

2.3.3 Geometrisches Kameramodell

Das geometrische Kameramodell beschreibt den Abbildungszusammenhang zwischen 3D-Objektkoordinate $\mathbf{X} = (X, Y, Z)^T$ und 2D-Bildkoordinate $\mathbf{x}' = (x, y)^T$ unter Berücksichtigung der inneren und äußeren Orientierung der Kamera. Sofern das Kameramodell das reale Abbildungsverhalten nur bedingt repräsentiert, werden Korrekturterme mit Zusatzparametern eingeführt. Im hier vorliegenden Fall der hinter der Windschutzscheibe befindlichen Monokamera bietet sich das Modell der *Zentralperspektiven Kamera* an. Eine Adaption des Modells sollte stattfinden, sofern für die Realisierung der Funktion Fischaugenkameras verwendet werden. Aufgrund ihrer geringen Brennweite und Öffnungswinkeln von über 180° , wird eine Fischaugen-Abbildung nach [46] vorgeschlagen.

2.3.3.1 Zentralperspektive Kamera

Dieses Modell eignet sich für Kameras, die die 3D-Szene auf einer ebenen Fläche durch nahezu geradlinig verlaufende Strahlen abbilden. Beispielhaft ist dies in Abbildung 2-8 dargestellt. Dabei verläuft der vom Objektpunkt \mathbf{X} ausgehende Strahl zum Projektions- oder Kamerazentrum $\mathbf{O}' = (X_0, Y_0, Z_0)$ und schneidet die Bildebene I' im Punkt \mathbf{x}' . Der lotrechte Abstand zwischen Projektionszentrum und Bildhauptpunkt \mathbf{H}' wird als Kamerakonstante c bezeichnet. Zur mathematischen Vereinfachung wird die Positivlage der Bildebene angenommen, wodurch die Kamerakonstante negativ angesetzt ist. Bei unendlicher Fokussierung ist die Kamerakonstante annähernd der Brennweite f gleichzusetzen.

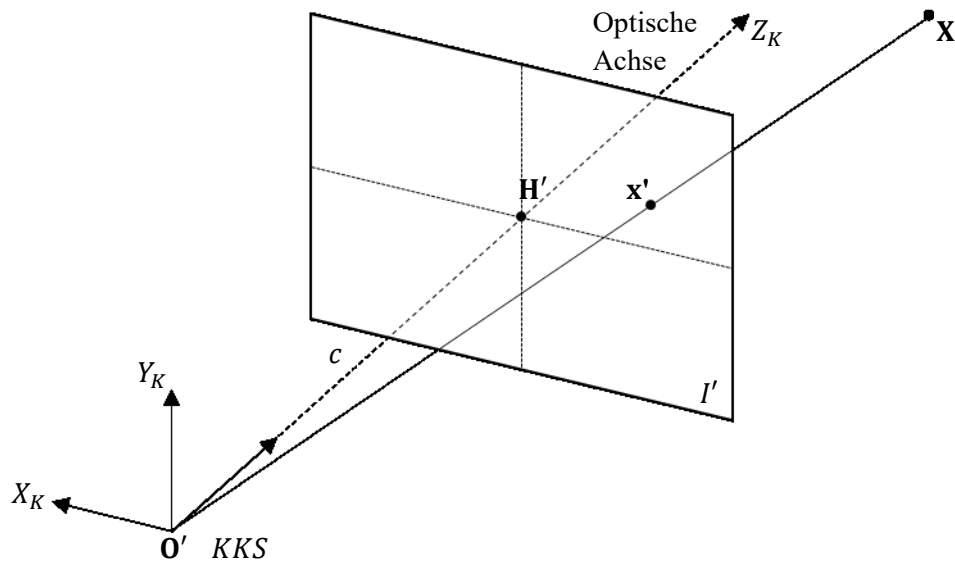


Abbildung 2-8 Prinzip der Zentralperspektiven Kamera

Gegeben sei der Punkt $\mathbf{X} = (X, Y, Z)^T$, welcher auf der Bildebene mit dem Abstand $Z = f$ in dem Punkt $\mathbf{x}' = (x, y)^T$ abgebildet werden soll. Der Richtungsvektor $(0,0,1)$ beschreibt die optische Achse und der Bildhauptpunkt wird mit $(0,0, f)$ angegeben. So lässt sich durch den Strahlensatz folgendes Verhältnis aufstellen:

$$\frac{x'}{X} = \frac{y'}{Y} = \frac{f}{Z} \quad \text{Gl. (2-8)}$$

Somit gilt für die Abbildung:

$$[X, Y, Z] \rightarrow \left[\frac{f}{Z}X, \frac{f}{Z}Y \right] \quad \text{Gl. (2-9)}$$

In homogenen Koordinaten kann dies als

$$\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \text{Gl. (2-10)}$$

formuliert werden (vgl. Gl. 2-7). Die Kameramatrix setzt sich dabei aus der Kalibriermatrix \mathbf{K} und der Einheitsprojektionsmatrix \mathbf{P}_p zusammen, welche wiederum die Einheitsmatrix und einen Nullvektor enthält $\mathbf{P}_p = [\mathbf{I}|\mathbf{0}]$.

$$\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}] = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{Gl. (2-11)}$$

Der Unterschied zwischen dem idealen zentralprojektiven Modell und der realen Abbildung kann über geometrische Größen beschrieben werden, die im Folgenden erläutert werden.

2.3.3.2 Innere Orientierung

Die optische Achse, die im idealen Modell direkt durch die beiden Halbierenden der Bildebene läuft, weicht in Realität davon ab, was über die Verschiebung des Hauptpunkts mit dem Offset x'_0 und y'_0 berücksichtigt wird. Verfügt der Bildsensor der realen Kamera nicht über exakt quadratische Pixel, so unterscheidet sich die Brennweite in horizontaler und vertikaler Richtung. Dies wird über die Parameter f_x und f_y berücksichtigt. Steht die optische Achse nicht senkrecht auf dem Bildsensor, so wird der Parameter s' eingefügt, welcher die Scherung beschreibt. Somit ergibt sich die innere Orientierung zu 5 Parametern, die in der Kalibriermatrix

$$\mathbf{K} = \begin{bmatrix} f_x & s' & x'_0 \\ 0 & f_y & y'_0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{Gl. (2-12)}$$

zusammenfasst werden.

In der realen Kamera kann es zudem zu Abbildungsfehler kommen, die durch Verzeichnungen des Objektivs hervorgerufen werden. Diese können über bildfehlerbeschreibende Funktionen berücksichtigt werden. Zunächst sei die radial-symmetrischen Verzeichnung erwähnt, die eine Ablenkung des Lichtstrahls lateral zur optischen Achse hervorruft. Diese kann nach [47] über das Polynom

$$\Delta r'_{rad} = k_0 r'^2 + k_1 r'^4 + k_2 r'^6 + k_3 r'^8 \quad \text{Gl. (2-13)}$$

beschrieben werden.

Die Ablenkung in die Richtungen der Bildebene ergibt sich zu:

$$\Delta x'_{rad} = x' \frac{\Delta r'_{rad}}{r'} \quad \Delta y'_{rad} = y' \frac{\Delta r'_{rad}}{r'} \quad \text{Gl. (2-14)}$$

Tangentiale oder asymmetrische Verzeichnungen, die durch eine Dezentrierung der Linsen hervorgerufen werden, können mit

$$\Delta x'_{tan} = p_1(r'^2 + 2x'^2) + 2p_2x'y' \quad \Delta y'_{tan} = p_2(r'^2 + 2y'^2) + 2p_1x'y' \quad \text{Gl. (2-15)}$$

formuliert werden. Affine und durch Scherung verursachte Verzeichnungen werden hier nicht berücksichtigt, da diese Parameter im Rahmen Bandende-Kalibrierung der Kamera ebenfalls nicht bestimmt werden. Die Zusatzparameter werden anschließend zu

$$\Delta x' = \Delta x'_{rad} + \Delta x'_{tan} \quad \Delta y' = \Delta y'_{rad} + \Delta y'_{tan} \quad \text{Gl. (2-16)}$$

zusammengefasst.

2.3.3.3 Äußere Orientierung

Zur Bestimmung von Objektkoordinaten, müssen zunächst die sechs Parameter der äußeren Orientierung der Kamera bestimmt werden. Sie beschreiben die Pose der Kamera im übergeordneten Koordinatensystem und wird über eine 3x3-Rotationsmatrix und einen 3x1-Translationsvektor angegeben. Dabei gilt:

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad \mathbf{t} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} \quad \text{Gl. (2-17)}$$

Die Kameramatrix kann damit zu

$$\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} = \mathbf{K}[\mathbf{R}|\mathbf{t}] \quad \text{Gl. (2-18)}$$

erweitert werden. Die insgesamt 11 Parameter der inneren und äußeren Orientierung sowie die Zusatzparameter der Abbildungsfehler werden innerhalb der Kamerakalibrierung bestimmt, die im Kapitel 3 näher erläutert wird. Die äußere Orientierung der Kamera innerhalb des USK wird einmalig bei der Kalibrierung bestimmt und dann als fest angenommen. Durch die Bewegung des Fahrzeugs ändert sich jedoch die äußeren Orientierung der Kamera im WKS, wodurch diese in jedem Zeitschritt berechnet werden muss (s. Abschnitt 2.3.4.4).

2.3.4 Epipolargeometrie

Die Epipolargeometrie ist ein mathematisches Modell zur Beschreibung der geometrischen Zusammenhänge zwischen Objektbildern, die von zwei verschiedenen Kameras aufgenommen wurden. Über diesen geometrischen Zusammenhang kann bestimmt werden wo sich der Bildpunkt \mathbf{x}' aus dem Bild I' der Kamera O' im Bild I'' einer anderen Kamera O'' befindet. Der Zusammenhang wird allein über die innere und äußere Orientierung der beiden Kameras bestimmt, nicht über die reale Struktur des Objekts. Der Objektpunkt und die beiden Kamerazentren spannen eine Ebene auf, welche die Bildebenen in den Epipolarlinien l' und l'' schneidet. Die Epipolargeometrie fordert nun, dass alle Objektpunkte \mathbf{X} , die auf dem Sichtstrahl \mathbf{r}' liegen und damit

das erste Bild in \mathbf{x}' schneiden, einen homologen Bildpunkt \mathbf{x}'' auf der Epipolarebene und damit auf der Epipolarlinie des zweiten Bildes erzeugt. Beispielhaft ist dies für die beiden Objektpunkt \mathbf{X}_1 und \mathbf{X}_2 in Abbildung 2-9 dargestellt.

Die geometrischen Eigenschaften der Epipolargeometrie werden im Verlauf der Arbeit genutzt, um:

1. Die Korrespondenzen zwischen zwei Bildern zu beurteilen.
2. Die Kamerabewegungen innerhalb der Bildfolge zu schätzen.
3. Die 3D-Rekonstruktion von homologen Punkten durchzuführen.

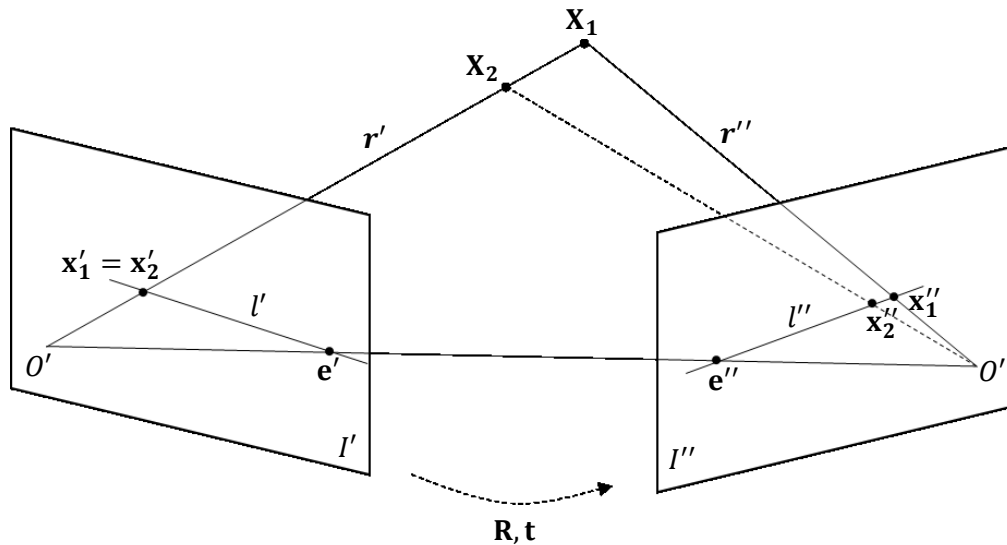


Abbildung 2-9 Korrespondenzen durch Epipolargeometrie

2.3.4.1 Fundamentalmatrix

Die Fundamentalmatrix ist die konkrete mathematische Umsetzung der geometrischen Zusammenhänge, die sich aus der Epipolargeometrie ergeben, wird mit dem Formelzeichen \mathbf{F} beschrieben und ist eine 3×3 Matrix. Sie enthält die gesamten Informationen über die Koplanarität von Projektionszentren, Bildpunkten und 3D-Punkt. Für Punktkorrespondenzen in zwei Bildern gilt die Epipolarbedingung:

$$\mathbf{x}''^T \mathbf{F} \mathbf{x}' = 0 \quad \text{Gl. (2-19)}$$

Durch \mathbf{F} kann zu jedem Bildpunkt \mathbf{x}' eine Epipolarlinie l'' bestimmt werden, auf welcher der korrespondierende Bildpunkt \mathbf{x}'' liegen muss.

$$l'' = \mathbf{F} \mathbf{x}' \quad \text{Gl. (2-20)}$$

Durch dieses Wissen kann die Suche nach möglichen Korrespondenzen vereinfacht werden, da sich der ursprünglich zweidimensionale in einen eindimensionalen Suchraum verringert. Zur eindeutigen Bestimmung der Fundamentalmatrix werden die Posen der beiden Kameras O' und O'' benötigt. Liegen diese nicht vor, so kann die Fundamentalmatrix über die Punktkorrespondenzen berechnet werden. Folgende Algorithmen bieten sich zur Berechnung der Fundamentalmatrix an.

8-Punkt Algorithmus

Dieser Algorithmus beschreibt die Bestimmung der Fundamentalmatrix mithilfe von acht Punkten nach [48]. Die Genauigkeit beschränkt sich dabei auf einen Skalierungsfaktor, was jedoch für die spätere Generierung der Epipolarlinien unerheblich ist. Durch Ausmultiplizieren von $\mathbf{x}''^T \mathbf{F} \mathbf{x}' = 0$ entsteht folgender Term

$$x''x'f_{11} + x''y'f_{12} + x''f_{13} + y''x'f_{21} + y''y'f_{22} + y''f_{23} + x'f_{31} + y'f_{32} + f_{33} = 0. \quad \text{Gl. (2-21)}$$

Für jedes Punktpaar ist somit eine Gleichung definiert, woraus sich das homogene lineare Gleichungssystem

$$\begin{bmatrix} x''_1x'_1 & x''_1y'_1 & x''_1 & y''_1x'_1 & y''_1y'_1 & y''_1 & x'_1 & y'_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x''_nx'_n & x''_ny'_n & x''_n & y''_nx'_n & y''_ny'_n & y''_n & x'_n & y'_n & 1 \end{bmatrix} \mathbf{f} = \mathbf{A}\mathbf{f} = 0 \quad \text{Gl. (2-22)}$$

ergibt. Das Gleichungssystem lässt sich mit 8 Punkten durch Singulärwertzerlegung lösen.

7-Punkt Algorithmus

Die Fundamentalmatrix kann auch mit nur sieben Punktkorrespondenzen berechnet werden. Unter der Berücksichtigung der Koplanaritätsbedingung

$$\det(\mathbf{F}) = 0 \quad \text{Gl. (2-23)}$$

kann über die Singulärwertzerlegung ein Lösungsraum mit zwei Vektoren bestimmt und die Fundamentalmatrix als gewichtete Linearkombination daraus formuliert werden.

2.3.4.2 Essentielle Matrix

Die essentielle Matrix \mathbf{E} ist eine spezielle Form der Fundamentalmatrix für normalisierte Bildkoordinaten [45]. Eingeführt von [48], beschreibt sie die geometrischen Zusammenhänge zweier kalibrierter Kameras. Die essentielle Matrix wird im vorliegendem Fall der Fundamentalmatrix vorgezogen, da im Fahrzeug mit einer kalibrierten Kamera gearbeitet wird und die Rechenzeit dadurch verringert werden kann. Die Definition der essentiellen Matrix lautet

$$\mathbf{x}''^T \mathbf{E} \mathbf{x}' = 0. \quad \text{Gl. (2-24)}$$

Bei \mathbf{x}' und \mathbf{x}'' handelt es sich dabei um normalisierte Punktkoordinaten. Ist die innere Orientierung der Kamera bekannt, ergeben sich zusätzliche Randbedingungen

$$\mathbf{E}\mathbf{E}^T\mathbf{E} - \frac{1}{2}\text{Spur}(\mathbf{E}\mathbf{E}^T)\mathbf{E} = 0, \quad \text{Gl. (2-25)}$$

wodurch sich die Anzahl der Freiheitsgrade damit auf 5 reduziert [49].

2.3.4.3 Berechnung von Inliern

Zur Minimierung von Ausreißern, die durch die Berechnung der essentiellen Matrix auf Basis von falschen Punktkorrespondenzen entstehen, werden iterative Lösungsansätze genutzt, um die tatsächlichen Korrespondenzen (Inlier) zwischen zwei Bildern zu bestimmen [45]. Häufig wird dafür das *RANSAC*-Verfahren (Random Sample Consensus) integriert, dessen Aufgabe darin besteht eine zufällige und minimale Auswahl von Punktkorrespondenzen zu bestimmen, mit denen sich die Parameter der essentiellen Matrix stabil berechnen lassen. Dabei werden folgende Schritte durchlaufen:

1. Herstellen von Punktkorrespondenzen \mathbf{x}' und \mathbf{x}''
2. Schätzung mittels RANSAC
 - a. Zufällige Auswahl von Korrespondenzen und Berechnung von \mathbf{E} mittels 5-Punkt Algorithmus.
 - b. Berechnung der Distanz d_{\perp} zwischen Korrespondenz \mathbf{x}'' und der zugehörigen Epipolarlinie l'' .
 - c. Berechnung der Inlier, für die $d_{\perp} < d_{max}$ gilt, wobei d_{max} einen eigens gewählten Schwellwert darstellt.
3. Wähle \mathbf{E} mit den meisten Korrespondenzen für die $d_{\perp} < d_{max}$ gilt.

Eine Erweiterung dieses Ansatzes mit automatischer Schwellwertbestimmung ist *AC-RANSAC* (A-Contrario RANSAC), welches von [50] vorgestellt wurde. Darüber hinaus existieren auch auf *Expectation Maximization* basierende Lösungen [51]. Diese Verfahren ermöglichen, dass der Abstand des Punkts zur Epipolarlinie zur Bewertung der Lösung herangezogen wird. Damit kann eine qualitativ höherwertige Lösung gewählt werden, welche nicht die maximale Anzahl an Inliern erhält.

2.3.4.4 Berechnung der Kameraorientierung

Die essentielle Matrix wurde über die Punktkorrespondenzen bestimmt, sodass nun über eine Singulärwertzerlegung von

$$\mathbf{E} = [\mathbf{t}]_{\times} \mathbf{R} = \mathbf{R} [\mathbf{R}^T \mathbf{t}]_{\times} \quad \text{Gl. (2-26)}$$

die relative Orientierung zweier Kameras bis auf einen Maßstab von \mathbf{t} bestimmt werden kann [48]. Das Gleichungssystem hat vier potentielle Lösungen, wobei diese gefunden werden muss, die eine Rekonstruktion von Objektpunkten im Sichtbereich beider Kameras ermöglicht. Der Maßstab kann über die Bekanntheit eines Objektpunktes und dem dazugehörigen Bildpunkt bestimmt werden.

Im hier entwickelten Verfahren wird der Maßstab bereits bei Start der Funktion innerhalb einer Initialisierungsphase bestimmt. Details dazu werden in Abschnitt 4.4 erläutert. Da innerhalb des Verfahrens davon ausgegangen werden kann, dass Objektpunkte in mehreren Bildern der Bildfolge sichtbar sind, kann die bereits 3D-rekonstruierte Szene zur Berechnung der absoluten Kameraorientierung herangezogen werden. Dafür wird die Lösung des räumlichen Rückwärtsschnitts berechnet, welcher mind. drei rekonstruierte Objektpunkte voraussetzt, die nicht auf gemeinsamen Gerade liegen dürfen.

2.3.4.5 Triangulation

Sind Punktkorrespondenzen und relative Orientierungen der kalibrierten Kameras bekannt, kann die Szenenrekonstruktion über den räumlichen Vorwärtsschnitt der Raumstrahlen \mathbf{r}' und \mathbf{r}'' (vgl. Abbildung 2-10) erfolgen. Dabei wird davon ausgegangen, dass die Bildkoordinaten und Orientierungen fehlerfrei sind und sich die beiden Raumstrahlen in exakt einem Punkt schneiden. Es gilt idealerweise:

$$\mathbf{x}' = \mathbf{P}'\mathbf{X} \text{ und } \mathbf{x}'' = \mathbf{P}''\mathbf{X}. \quad \text{Gl. (2-27)}$$

Die beiden Kameramatrizen besitzen unterschiedliche homogene Skalierungsfaktoren, die jeweils mithilfe des Kreuzproduktes entfernt werden. Für jeden Bildpunkt, gibt es drei Gleichungen von denen zwei unabhängig sind:

$$\mathbf{x}' \times \mathbf{P}'\mathbf{X} = \begin{bmatrix} x'(\mathbf{p}'^{3T}\mathbf{X}) - (\mathbf{p}'^{1T}\mathbf{X}) \\ y'(\mathbf{p}'^{3T}\mathbf{X}) - (\mathbf{p}'^{2T}\mathbf{X}) \\ x'(\mathbf{p}'^{2T}\mathbf{X}) - y'(\mathbf{p}'^{1T}\mathbf{X}) \end{bmatrix} = 0 \quad \text{Gl. (2-28)}$$

\mathbf{p}'^{iT} bezeichnet dabei die i -te Reihe der Kameramatrix \mathbf{P}' . Dies lässt sich in ein lineares Gleichungssystem der Form $\mathbf{A}\mathbf{X} = 0$ überführen

$$\mathbf{A}\mathbf{X} = \begin{bmatrix} x'\mathbf{p}'^{3T} - \mathbf{p}'^{1T} \\ y'\mathbf{p}'^{3T} - \mathbf{p}'^{2T} \\ x''\mathbf{p}''^{3T} - \mathbf{p}''^{1T} \\ y''\mathbf{p}''^{3T} - \mathbf{p}''^{2T} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = 0 \quad \text{Gl. (2-29)}$$

und anschließend lösen. Da die Merkmalspositionen und relativen Orientierungen jedoch fehlerbehaftet sind, folgen windschiefe Geraden daraus, die einen Punkt \mathbf{X}' und \mathbf{X}'' erzeugen. Die Projektion erfüllt demnach nicht die Epipolarbedingung und der Objektpunkt \mathbf{X} muss geschätzt werden. Im einfachsten Fall wird dabei der kürzeste Abstand des Objektpunkts zu beiden windschiefen Geraden bestimmt.

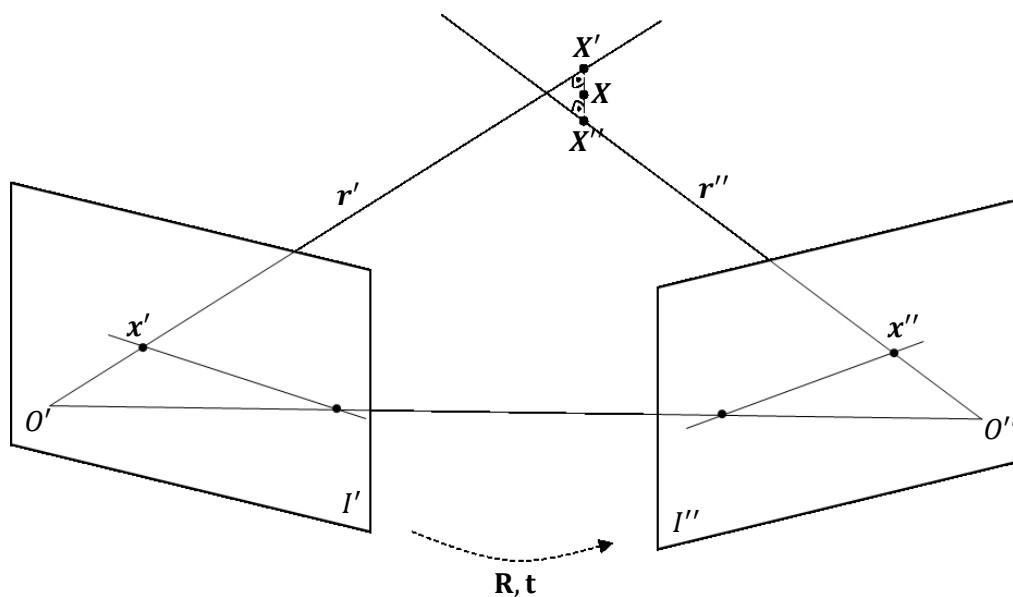


Abbildung 2-10 Räumlicher Vorwärtsschnitt

Eine weitere hier eingesetzte Variante ist die Minimierung des geometrischen Fehlers der rückprojizierten Bildpunkte $\hat{\mathbf{x}}'$ und $\hat{\mathbf{x}}''$. Der Fehler e berücksichtigt den euklidischen Abstand zwischen gemessenen und rückprojizierten Bildpunkten gemäß

$$e = d(\mathbf{x}', \hat{\mathbf{x}}')^2 + d(\mathbf{x}'', \hat{\mathbf{x}}'')^2. \quad \text{Gl. (2-30)}$$

Die Kostenfunktion kann mittels Levenberg-Marquardt Methode [52] minimiert werden. Ein polynomialer Ansatz zur Triangulation, welcher nicht iterativ verbesserte Bildkoordinaten anhand des Abstands zu den Epipolarlinien ermittelt, wird darüber hinaus in [45] vorgestellt und evaluiert.

2.4 Statistik

In diesem Abschnitt werden grundlegende Begriffe der in dieser Arbeit verwendeten statistischen Methoden zur Auswertung, Analyse und Weiterverarbeitung der gesammelten Daten erläutert. Es wird sich dabei an [53] und [49] orientiert.

Messabweichungen

Jede Messgröße (gemessene Zufallsgröße) hat zu dem Zeitpunkt ihrer Beobachtung einen bestimmten Wert. Messungen sind jedoch stets fehlerbehaftet, sodass sich eine Differenz der beobachteten Messgröße (Ist) von der ihr zugeordneten Bezugsgröße (Soll) ergibt, welche als Abweichung bzw. Fehler bezeichnet wird. Die Bezugsgröße wird auch als Referenzwert oder wahrer Wert bezeichnet.

Grobe Abweichungen

Grobe Fehler (Ausreißer) können in praktischen Anwendungen nur selten vermieden werden. Zurückgehend auf Fehlmessungen oder falsche Punktkorrespondenzen müssen Ausreißer im Prozess so früh wie möglich eliminiert werden, da sie einen direkten Einfluss auf die Qualität der Unbekannten (3D-Punkte) und somit auf das Gesamtergebnis haben. Darüber hinaus gibt es zufällige und systematische Fehler, welche nach ihrer Art der Entstehung unterschieden werden.

Zufällige Abweichungen

Das Ergebnis einer Messung liegt stets in einem gewissen Streubereich. Die Messungen unterliegen dabei Einflüssen, die eine beidseitige Schwankung um den wahren Wert hervorruft und als zufällige Abweichungen bezeichnet werden. Sie können bspw. von Ungenauigkeiten der Messmittel stammen und durch äußere Bedingungen hervorgerufen werden. Während der Ausgleichsrechnung können die zufälligen Fehler minimiert werden.

Systematische Abweichungen

Diese Art von Fehler führt zu einseitigen Abweichungen vom wahren Wert und stammen häufig von Vereinfachungen in den physikalischen oder mathematischen Modellen. Sie wirken sich somit auf alle Messwerte aus. Systematische Fehler lassen sich durch Kalibrierungen erfassen und werden durch Korrekturterme berücksichtigt. Unbekannte systematische Abweichungen addieren sich zu den zufälligen Abweichungen und bestimmen die resultierende Messunsicherheit.

Genauigkeit

Die Beurteilung der Akzeptanz und Nutzbarkeit des hier entwickelten Messsystems kann anhand seiner Genauigkeit erfolgen. Die innere Genauigkeit, welche bspw. die geschätzten Standardabweichungen der Kalibrierparameter, Bildkoordinaten, Kameraorientierungen oder 3D-Punkte angibt, lässt jedoch nur einen bedingten Rückschluss auf die Richtigkeit der rekonstruierten Umgebung zu. Die Bestimmung der äußeren Genauigkeit, welche über den Vergleich der ermittelten Größen mit unabhängigen Referenzdaten erfolgt, gibt dabei Aufschluss über die praktische Einsetzbarkeit des Systems zur Berechnung von Hindernishöhen und -abständen.

Zuverlässigkeit

Innerhalb einer Ausgleichung wird die Zuverlässigkeit zum einen über die Redundanzanteile der Beobachtungen und zum anderen über die Detektionsfähigkeit von groben Fehlern definiert. Darüber hinaus kann die Zuverlässigkeit des Systems im Sinne des Endnutzers analysiert werden, indem mit einer Vielzahl an Umgebungen und Sequenzen die Detektionsrate für ein potentielles Hindernis bestimmt wird.

Lage- und Streuungskennwerte

Arithmetischer Mittelwert

Der (arithmetische) Mittelwert \bar{x} beschreibt den Quotient aus Summe der Beobachtungswerte x_i und Anzahl der Beobachtungen n .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{Gl. (2-31)}$$

Spannweite

Die Spannweite (Spanne) R bezeichnet den maximalen Abstand der Beobachtungen innerhalb einer Messreihe. Da in dieser auch Ausreißer enthalten sind, gilt die Spannweite als nicht erwartungstreu.

$$R = X_{max} - X_{min} \quad \text{Gl. (2-32)}$$

Normalverteilung

Eines der wichtigsten Verteilungsmodelle der Statistik ist die Normalverteilung. Sie beruht auf der Aussage des zentralen Grenzwertsatzes, nach dem die Summe von n unabhängigen beliebigen Messungen mit endlicher Varianz annähernd normalverteilt ist. Die Dichtefunktion ist mit

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{Gl. (2-33)}$$

angegeben und wird dem Erwartungswert μ und der Varianz σ^2 definiert.

Standardabweichung

Die Standardabweichung ist eine Maß für die Streuung einer Zufallsgröße und wird auch als mittlerer Fehler bezeichnet. Bei bekanntem Erwartungswert berechnet sie sich zu

$$\sigma = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i \varepsilon_i}{n}} \quad \text{Gl. (2-34)}$$

mit $\varepsilon_i = x_i - \mu$. Der Erwartungswert μ enthält dabei den Mittelwert aller möglichen Messwerte und kann mit dem arithmetischem Mittelwert \bar{x} empirisch geschätzt werden. Für empirisch bestimmte Standardabweichungen wird hier das Formelzeichen s genutzt.

Beziehung zweier Elemente

Kovarianz

Innerhalb eines Zufallsvektors besitzt jedes Element eine Varianz σ^2 . Zwischen zwei Elementen x und y können weitere Information zur „gemeinsamen Streuung“ angegeben werden, die Kovarianz. Sie entspricht dem mittleren Produkt der Abweichungen der einzelnen Messungen von ihrem jeweiligen Mittelwert und wird mit

$$\sigma_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) = \frac{1}{n} \varepsilon_x^T \varepsilon_y \quad \text{Gl. (2-35)}$$

berechnet.

Korrelation

Die Dimensionen zweier Zufallszahlen beeinflussen der Wert der Kovarianz. Über eine Division über ihre jeweiligen Standardabweichungen können die Kovarianzen normiert werden, sodass sich der Korrelationskoeffizient zu

$$\rho_{xy} = \frac{\sigma_{xy}}{\sqrt{\sigma_x^2 \cdot \sigma_y^2}} \quad \text{Gl. (2-36)}$$

ergibt. Hohe Korrelationskoeffizienten weisen auf eine hohe Abhängigkeit der Parameter hin und sollten bspw. bei der Kamerakalibrierung vermieden werden.

3 Kamerakalibrierung

3.1 Grundlagen

Die Kamerakalibrierung dient der Ermittlung der inneren und äußeren Orientierung. Sie hat direkten Einfluss auf die Qualität der nachfolgenden Auswerteprozesse und ist somit von großer Bedeutung. In der Literatur finden sich verschiedene Ansätze zur Kalibrierung der inneren Kameraorientierung. Sie werden nach [54] in drei Kategorien eingeteilt.

Bei der *Laborkalibrierung* werden die physikalischen Beziehungen zwischen Sensorarray der aufnehmenden Kamera und dem Sichtstrahl optisch bestimmt, bspw. unter Nutzung eines Kollimators. Der Aufwand einer solchen Kalibrierung ist vergleichsweise groß und erfordert Laborequipment. In Praxis wird diese Methode nur begrenzt eingesetzt. Bei der *Testfeldkalibrierung* werden die Beziehungen zwischen Punkten und/oder Linien auf einem Testfeld und deren Abbildungen auf den Kamerabildern bestimmt, um die innere Orientierung zu berechnen. Das Testfeld kann dabei durch ein ebenes Schachbrett oder Punktarray oder bspw. eine 3D-Anordnung von bekannten Punkten repräsentiert werden. Dieses Verfahren wird sehr häufig eingesetzt und ist bereits heute Bestandteil der Bandende-Kalibrierung der Fahrzeugkamera. Das Verfahren der *Selbstkalibrierung* bestimmt die Parameter der inneren Orientierung durch Bündelausgleichung. Um die nötige Genauigkeit zu erlangen empfiehlt sich die Nutzung eines Testfelds oder künstlicher Marker. Entgegen der Testfeldkalibrierung müssen die Punkt- bzw. Markerpositionen jedoch nicht bekannt sein, da sie während der Bündelausgleichung mitgeschätzt werden. Solange in der Szene reichlich und gut identifizierbare Punkte bestimmt werden können, kann die Selbstkalibrierung der Fahrzeugkamera während der Fahrt geschehen.

Im vorliegenden Fall wird die Kamera mithilfe eines bekannten Testfelds kalibriert. Dies kann wiederum unter Nutzung von einem oder mehreren Bildern geschehen. Die *Einzelbildkalibrierung* stützt sich auf die Methode des räumlichen Rückwärtsschnitts, wobei in lineare und nichtlineare Ansätze für die Lösung der Kollinearitätsgleichungen unterschieden wird. Es sind dafür mindestens sechs räumlich verteilte 3D-Passpunkte erforderlich, die sich nicht auf einer Ebene befinden dürfen. Die *Mehrbildkalibrierung*, welche auch hier gewählt wird, bestimmt die Kameraparameter mithilfe mehrerer aus unterschiedlichen Richtungen aufgenommener Bilder des Testfelds. Dabei wird ebenfalls das Modell des räumlichen Rückwärtsschnitts herangezogen und die Parameter der inneren und äußeren Orientierung sowie die Testfeldkoordinaten über eine Bündelausgleichung geschätzt. Die Präzision der Bildpunktmessung, die räumliche Verteilung der Messpunkte und die Aufnahmeconfiguration beeinflussen dabei maßgeblich die Qualität der Kamerakalibrierung [55]. Ebene Testfelder mit Schachbrett- oder Punktmuster ermöglichen eine einfache Kalibrierung, liefern jedoch je nach Anwendungsfall unzureichend genaue Ergebnisse, sodass für die hier geforderte 3D-Messung der Durchfahrtshöhen während der Kamerakalibrierung zusätzlich ein 3D-Testfeld mit bekannten Markern zum Einsatz kommt.

3.2 Bündelausgleichung

Die Bündelausgleichung beschreibt den Schritt der iterativen Lösung der erweiterten Kollinearitätsgleichungen (Gl. (3-1) und Gl. (3-2)). Sie beschreiben den nichtlinearen Abbildungsprozess eines Objektpunktes in Bildkoordinaten unter Berücksichtigung der Kamerakalibrierung. Für den Parameter z' gilt dabei $z' = -c$.

$$x' = x'_0 + z' \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta x' \quad \text{Gl. (3-1)}$$

$$y' = y'_0 + z' \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} + \Delta y' \quad \text{Gl. (3-2)}$$

Zur iterativen Lösung des nichtlinearen Gleichungssystems werden Kleinste-Quadrate-Ansätze, wie das Gauss-Markov oder das Levenberg-Marquardt Modell genutzt [56, 57]. Neben der Ermittlung der inneren und äußeren Orientierung wird das Ausgleichsverfahren auch zur Bestimmung und Optimierung der 3D-Koordinaten der Bildpunkte verwendet. Die Bündelausgleichung bedingt eine initiale Schätzung der Kameraparameter und führt anschließend eine lokale Optimierung durch, die den Rückprojektionsfehler minimiert.

3.3 Kalibrierumgebung

Die Kamera wird im Versuchsfahrzeug hinter der Frontscheibe auf Höhe des Rückspiegels befestigt und anschließend die inneren und äußeren Kameraparameter bestimmt.

3.3.1.1 Äußere Orientierung

Die Bestimmung der relativen äußeren Orientierung im USK wird mithilfe eines 2D-Testfelds und Markern durchgeführt. Das Testfeld wird im Sichtbereich der Fahrzeugkamera positioniert und legt das Koordinatensystem fest. Marker werden auf dem Fahrzeug und auf dem Boden verteilt. Mithilfe des Maßstabs (angelehnt an den Tisch in Abbildung 3-1 rechts) wird ein metrisches Modell des Fahrzeugs erstellt, anschließend die Bodenebene, die Räder und Hinterachse extrahiert sowie der Ursprung des USK festgelegt. Wird das von der Fahrzeugkamera aufgenommene Bild des Testfelds in die Ausgleichung einbezogen, so kann deren Pose im USK bestimmt werden.



Abbildung 3-1 Bestimmung der äußeren Orientierung im USK mithilfe eines AICON 3D Systems

Die Parameter der äußeren Orientierung werden in Tabelle 3-1 dargestellt. Die innere Orientierung wird dabei als bekannt angenommen. Die Trennung von Kalibrierung der inneren und äußeren Orientierung ist explizit gewünscht, da dadurch Korrelationen aus 3D-Punktberechnungen und Größen der Kameraparameter vermieden werden.

Tabelle 3-1 Äußere Orientierung der Kamera im Fahrzeug-USK

	Orientierung
X [m]	2.591
Y [m]	-0.018
Z [m]	1.694
Φ [°]	1.364
Θ [°]	0.973
Ψ [°]	1.535

Es wird davon ausgegangen, dass keine weitere Rotation oder Translation der Kamera innerhalb des USK während der Fahrt vorliegt. Eine etwaige Verformung aufgrund von Wärmeeinstrahlung wird als vernachlässigbar gering eingestuft. Der Zugang zur Kamera durch den Fahrer wird über eine Abdeckung verhindert.

3.3.1.2 Innere Orientierung

Die innere Orientierung beschreibt den Unterschied zwischen dem idealen zentralprojektiven Modell der Lochkamera und dem realen Abbildungsmodell. Während der Kalibrierung mit 2D- und 3D-Testfeldern werden die Parameter der Brennweite (f_x, f_y) und des Bildhauptpunktes (x'_0, y'_0) ermittelt. Die Abbildungsfehler werden mit der Methode nach [47] modelliert, welche als Standard für perspektivische Kameras gilt. Dabei ergeben sich drei Parameter für die radiale und zwei Parameter für die tangentielle Verzeichnung (k_1, k_2, k_3, p_1, p_2).

2D-Testfeld

Es werden vier verschiedene 2D-Testfelder mit Schachbrettmuster verglichen, um den Einfluss des Testfelds auf die Kalibriergüte zu bestimmen. Die unterschiedlichen Kalibrierfälle können dabei Tabelle 3-2 entnommen werden, wobei dabei Schachbretter mit 35, 96, 234 bzw. 294 Punkten zum Einsatz kommen. Um optimale Ergebnisse zu erhalten werden mindestens 35 Bilder vom Testfeld aus möglichst verschiedenen Winkeln und Distanzen aufgenommen. Die anschließende Kalibrierung wird mit der Methode nach [58] durchgeführt.

Tabelle 3-2 Kalibrierfälle mit 2D-Testfeld

Name des Kalibrierfalls	Testfeldmuster	Anzahl der Punkte	Versuche
S35	Schachbrett	35	5
S96	Schachbrett	96	5
S234	Schachbrett	234	5
S294	Schachbrett	294	5

Zur Evaluierung der Ergebnisse wird der Rückprojektionsfehler herangezogen, der den Fehler zwischen dem detektierten 2D-Punkt und dem projizierten 3D-Punkt beschreibt. Anschließend werden die Rückprojektionsfehler aller verwendeter Bilder gemittelt und miteinander verglichen.

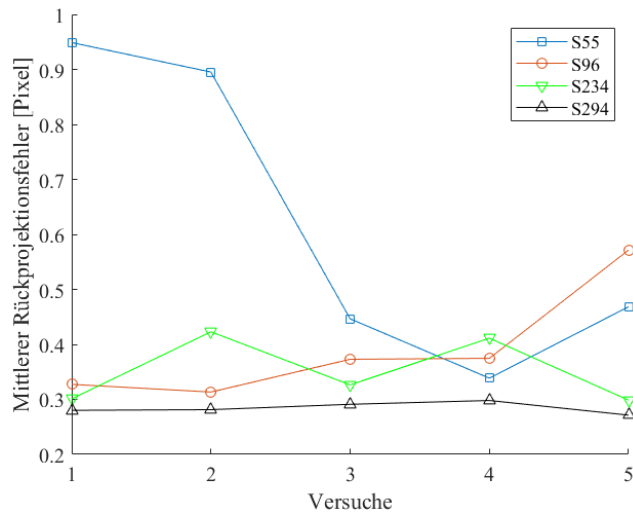


Abbildung 3-2 Gemittelter Rückprojektionsfehler bei unterschiedlichen Kalibrierfällen

Das Ergebnis zeigt, dass die Qualität der Kalibrierung mit wenigen Punkten bei mehrmaliger Kalibrierung mit abweichenden Aufnahme Konfigurationen keine stabilen Ergebnisse liefert. Die nachfolgende Tabelle fasst die Kalibrierergebnisse mit 2D-Testfeldern zusammen. Das Ergebnis zeigt, dass die Punktzahl entscheidenden Einfluss auf die Güte der Kalibrierung hat. So können beim Kalibrierfall S294 stabile Ergebnisse unter Berücksichtigung verschiedener Aufnahme Konfigurationen erzielt werden. Besonders wird dies in den Parametern der Brennweite und des Bildhauptpunktes sichtbar. Für die Parameter der Verzeichnung werden für alle Testfälle ähnliche Ergebnisse erzielt. Die Unterschiede sind nur marginal können nahezu vernachlässigt werden. Sofern kein 3D-Testfeld zur Verfügung steht, wird demnach empfohlen ein Schachbrett mit einer hohen Anzahl an Eckpunkten zu wählen.

Tabelle 3-3 Parameter der Kalibrierung (2D- Testfeld)

	S35		S96		S234		S294	
	Mittelwert	Std. Abw.	Mittelwert	Std. Abw.	Mittelwert	Std. Abw.	Mittelwert	Std. Abw.
f_x [Pixel]	973.57	5.2999	974.29	2.392	972.92	3.053	975.01	0.2521
f_y [Pixel]	976.98	6.4162	974.35	1.836	971.07	2.724	974.54	0.5417
x_0 [Pixel]	638.99	14.352	648.15	2.177	644.29	2.928	644.79	1.908
y_0 [Pixel]	467.89	9.2995	482.55	4.785	479.09	3.956	481.78	0.3698
k_1 [1/mm]	-0.3705	0.0129	-0.3725	0.029	-0.3796	0.0035	-0.3747	0.0021
k_2 [1/mm ³]	0.1994	0.0287	0.2114	0.0108	0.2255	0.0101	0.2132	0.0069
k_3 [1/mm ⁵]	-0.0623	0.0328	-0.0738	0.0143	-0.0519	0.0524	-0.0761	0.0087
p_1 [1/mm]	0.0007	0.0018	-0.0001	0.0005	-0.0003	0.0006	-0.0002	0.0001
p_2 [1/mm ³]	0.0006	0.0006	-0.0001	0.0001	-0.0002	0.0004	-0.0001	0.0001
\bar{e} [Pixel]	0.6195	0.2811	0.3905	0.1057	0.3504	0.0586	0.2824	0.0097

3D-Testfeld

Die Bestimmung der Entfernung und Durchfahrthöhe von markanten Hindernissen wird im Rahmen dieser Arbeit in einem durch die Kamera rekonstruierten 3D-Punktraum durchgeführt. Es bietet sich demnach an die Art der Kalibrierung an den praktischen Einsatzzweck anzugleichen und ein 3D-Testfeld zu verwenden. Die räumliche Tiefe des Testfelds erlaubt eine bessere Bestimmbarkeit der Parameter und die Reduzierung der Korrelationen der Parameter untereinander [49]. Bei der Kalibrierung mit einem 2D-Testfeld wird dies nur bedingt durch die Aufnahme von Bildern in unterschiedlichen Distanzen erreicht.

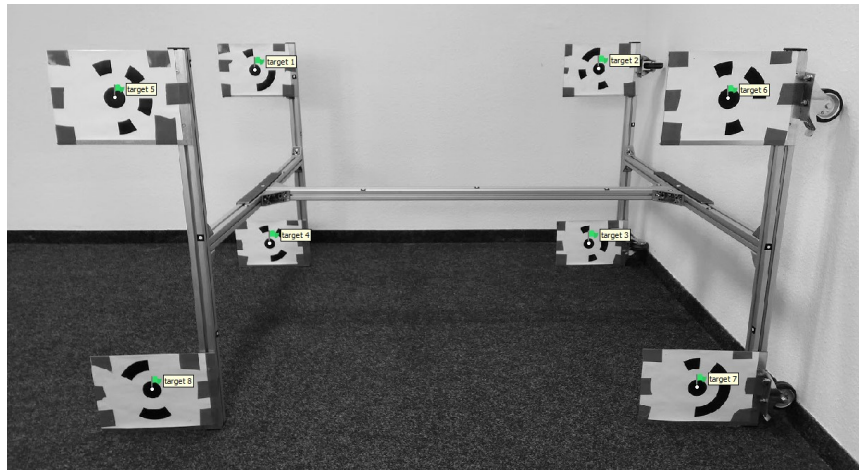


Abbildung 3-3 3D-Testfeld bestehend aus acht definierten Markerpositionen

Unter Berücksichtigung der Anfangsparameter von S294 werden mit dem 3D-Testfeld weitere Kalibrierungen durchgeführt. Die aufgeklebten Marker werden dabei von Agisoft Metashape erkannt und in der Bündelausgleichung berücksichtigt. Im ersten Fall werden die Marker ohne bekannte/vermessene Position im Objektkoordinatensystem verwendet. Aus Tabelle 3-4 wird ersichtlich, dass die Qualität in allen Parametern und dem Rückprojektionsfehler gesteigert werden kann. Die Standardabweichungen der Brennweite können um den Faktor 3.8 bzw. 8.2 und des Hauptpunktes um 12.4 bzw. 3.1 verringert werden. Eine Verbesserung der Verzeichnungsparameter ist aufgrund der geringen Punktzahl nicht zu erwarten. Der Rückprojektionsfehler ist um den Faktor 2.4 reduziert worden. Der Anstieg der Standardabweichung ist mit der Betrachtung von nur 8 Punkten zu erklären, da die fehlerhafte Bestimmung eines Punktes im Vergleich zu den 294 Punkten des Schachbretts bereits sehr stark das Ergebnis beeinflusst. Im zweiten Fall, d.h. mit Markerpositionen, die mittels eines Laserdistanzmessers vermessen wurden, lassen sich nur marginale Verbesserungen erreichen. Dies ist voraussichtlich auf die Vermessung mittels Laserdistanzmesser zurückzuführen, dessen Genauigkeit mit ± 2 mm angegeben ist. Die berechneten Markerkoordinaten konnten nach der Ausgleichung mit einem gemittelten Fehler von 7.24 mm angegeben werden.

Tabelle 3-4 Parameter der Kalibrierung (3D- Testfeld)

	2D-Schachbrett		3D-Marker (o. Pos.)		3D-Marker (m. Pos.)	
	Mittelwert	Std. Abw.	Mittelwert	Std. Abw.	Mittelwert	Std. Abw.
f_x [Pixel]	975.01	0.2521	974.02	0.059	973.98	0.059
f_y [Pixel]	974.54	0.5417	973.94	0.061	973.91	0.060
x_0 [Pixel]	644.79	1.908	647.39	0.113	647.39	0.152
y_0 [Pixel]	481.78	0.3698	486.75	0.086	486.48	0.117
k_1 [1/mm]	-0.3747	0.0021	-0.3764	0.0002	-0.3767	0.0002
k_2 [1/mm ³]	0.2132	0.0069	0.2114	0.0005	0.2120	0.0005
k_3 [1/mm ⁵]	-0.0761	0.0087	-0.0727	0.0001	-0.0731	0.0004
p_1 [1/mm]	-0.0002	0.0001	-0.0001	0.0001	-0.0001	0.0001
p_2 [1/mm ³]	-0.0001	0.0001	-0.0002	0.0001	-0.0001	0.0001
\bar{e} [Pixel]	0.2824	0.0097	0.1343	0.0341	0.1338	0.0348

Die Auswertung zeigt, dass mit einem 3D-Testfeld eine genauere Bestimmung der inneren Orientierung möglich ist. Vor allem die Methode ohne bekannte Position der 3D-Marker eignet sich für eine Kalibrierung der Fahrzeugkamera im täglichen Einsatz. Beispielweise können diese Marker innerhalb einer Garage angebracht und bei Einfahrt bzw. Ausfahrt automatisch detektiert und zur Nachkalibrierung genutzt werden. Bei der Anbringung der Marker ist auf die spätere Aufnahmeconfiguration zu achten, die aufgrund der fest verbauten Kamera und der geringen Rotationsfähigkeit des Fahrzeugs eingeschränkt sein kann.

Bei der Kalibrierung werden innerhalb der Bündelausgleichung die Beobachtungen im Bild bestmöglich auf den Objektpunkten abgebildet. Dabei kann es zu Überparametrisierungen und Korrelationen zwischen den Parametern der inneren sowie inneren und äußeren Orientierung kommen. Sofern der Berechnungsvorgang innerhalb eines mathematischen Systems vorgenommen wird, ist der Einfluss der Korrelationen zu vernachlässigen [49]. Da die Kalibrierparameter jedoch im späteren Verlauf der Arbeit für die Bestimmung der 3D-Punkte des Fahrumfelds genutzt wird, können die Korrelationen zu einer Verfälschung des funktionalen Modells führen. Geeignete Aufnahmeconfigurationen können Korrelationen jedoch während der Kalibrierung mindern. Durch gekantete Aufnahmen in verschiedenen Distanzen und Drehrichtungen werden unter Nutzung eines 2D-Testfelds stabile Parameter für die Brennweite und den Bildhauptpunkt berechnet. Vorteil des 3D-Testfelds ist das Vorhandensein einer zusätzlichen Maßstabsinformation in Aufnahmeichtung, welches eine Reduktion der Korrelationen ermöglicht. Verzeichnungsparameter können zuverlässig über eine formatfüllende Bildabdeckung bestimmt werden. Tabelle 3-5 zeigt, dass mithilfe des 3D-Testfelds die Korrelation der Brennweite und des Bildhauptpunktes reduziert werden konnte. Es ergeben sich hohe Abhängigkeiten zwischen der Brennweite und den Verzeichnungsparametern bei beiden Kalibrierungen, deren Auswirkung jedoch aufgrund der hohen Genauigkeit der Parameter als gering eingestuft wird. Die Verzeichnungsparameter k_1 und k_2 bzw. k_3 sind hochkorreliert. Die Untersuchung zeigt, dass auf die Modellierung von k_2 und k_3 ohne Qualitätsverlust verzichtet werden kann. Die Abhängigkeit der Dezentrierung der Linse und den tangentialen Verzeichnungsparametern ist in den Korrelationskoeffizienten von x_0 und p_1 bzw. y_0 und p_2 zu erkennen.

Kamerakalibrierung

Tabelle 3-5 Korrelationen der inneren Orientierung. Oben: Nach Kalibrierung mit 2D-Testfeld; Unten: Nach Kalibrierung mit 3D-Testfeld ohne bekannte Markerpositionen

	f	x_0	y_0	k_1	k_2	k_3	p_1	p_2
f	1	-0.03 -0.05	-0.42 0.07	-0.44 -0.34	0.26 0.29	-0.21 -0.25	-0.04 0.06	0.62 0.11
x_0		1	0.05 -0.04	0.02 0.04	-0.01 -0.02	0.01 0.01	-0.36 -0.65	0.04 0.03
y_0			1	0.09 -0.04	-0.04 0.02	0.04 -0.01	-0.02 0.00	-0.56 -0.59
k_1				1	-0.95 -0.97	0.87 0.91	0.00 -0.06	-0.13 0.00
k_2					1	-0.97 -0.98	-0.01 -0.06	0.05 0.01
k_3						1	-0.03 -0.04	-0.05 -0.02
p_1							1	-0.05 -0.01
p_2								1

Langzeitstabilität

Die Kamerakalibrierung wird am Ende der Fahrzeugproduktion vorgenommen und die Parameter im Kameraspeicher hinterlegt. Die Durchfahrtshöhenerkennung, die über den gesamten Zeitraum des Fahrzeugbetriebs zur Verfügung stehen soll, greift auf diese Kameraparameter zurück. Hier soll untersucht werden, inwiefern sich die Kameraparameter im Laufe der Zeit verändern. Dafür wurde die Kamerakalibrierung über mehrere Wochen wiederholt und die Parameteränderungen ausgewertet. Das Fahrzeug und damit die Kamera werden dabei unter realen Nutzungsbedingungen betrieben. Aufgrund der nur geringen Abweichung der Verzeichnungsparameter, die im 1/100- bzw. 1/1000-Bereich liegen, werden diese in der Übersicht nicht angezeigt.

Tabelle 3-6 Zeitliche Veränderung der Kalibrierparameter

	Tag 1	Tag 7	Tag 15	Tag 20
Δf_x [Pixel]	0.122	0.221	0.411	1.003
Δf_y [Pixel]	0.122	0.220	0.410	1.001
Δx_0 [Pixel]	0.578	0.492	0.513	5.801
Δy_0 [Pixel]	0.512	0.001	0.049	4.039

Wie in Tabelle 3-6 zu sehen, weisen die Kalibrierparameter bereits nach wenigen Nutzungstagen Veränderungen auf. Während sich die Brennweite in x und y im gleichen Maße ändert, verschiebt sich der Hauptpunkt unregelmäßig. Dies lässt auf eine Instabilität der Kamera schließen. Hervorgerufen von Erschütterungen während des Betriebs im Fahrzeug oder des Ein- und Ausbaus, wird dies während der Kalibrierung sichtbar. Bei der vorliegenden Kamera, der AVT Manta G-125C, primär als Industriekamera entwickelt, müssen die Aspekte der mechanischen Instabilität berücksichtigt werden und eine Selbstkalibrierung während der Fahrt erfolgen.

Temperatur

Neben den Erschütterungen, die während der Fahrt auf die Kamera einwirken, sind Temperaturunterschiede eine weitere Einflussquelle für Veränderungen der inneren Orientierung. Die Temperaturunterschiede, denen das Fahrzeug im Laufe des Nutzungszeitraums unterliegt sind enorm. Je nach Auslieferungsland können die Temperaturen im Fahrzeug stark schwanken. Ausgehend von einer Starttemperatur von 65 °C werden die Bildpunktabweichungen bis hin zu einer Temperatur von 29 °C untersucht. Dies entspricht ungefähr der Betriebstemperatur der Kamera im gewählten Umfeld.

Die Kamera wird dabei zur einfachen Handhabung aus dem Fahrzeug ausgebaut, mit einem Heißluftfön gleichmäßig erhitzt und die Temperatur am Gehäuse mittels Infrarot-Lasermessgerät bestimmt. Die unterschiedlichen Temperaturen am Kameragehäuse, der Linse und dem Sensor sollten idealerweise getrennt bestimmt werden, um den Einflussfaktor der unterschiedlichen Bauteile zu bestimmen. Da dies konstruktiv und messtechnisch hier nicht möglich ist, wird die Kamera als gesamte Einheit untersucht und die Temperatur möglichst langsam und gleichmäßig in der Erwärmungsphase erhöht. Um etwaige Einflüsse über den Luftstrom des Heißluftgeräts zu verhindern, wird die Messung während der Abkühlung durchgeführt.

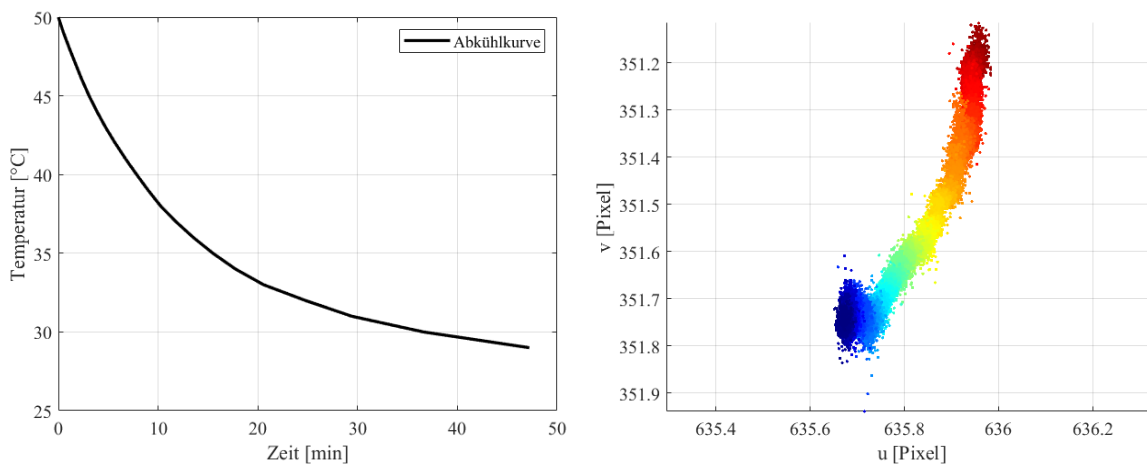


Abbildung 3-4 Temperatureinfluss auf die Punktdetektion. Links: Abkühlkurve; Rechts: Änderung der Punktkoordinate 146 (Rot - Starttemperatur, Blau - Endtemperatur)

Abbildung 3-4 zeigt, dass mit 29 °C fast die Betriebstemperatur der Kamera erreicht wurde und davon ausgehend keine starken Änderungen der Bildpunktkoordinaten zu erwarten sind. Im rechten Teil wird ersichtlich, dass sich der Bildpunkt vor allem in der y -Richtung ändert. Aus Abbildung 3-5 geht hervor, dass dies für alle Punktkoordinaten zutrifft. Es wird davon ausgegangen, dass die seitliche Sensorbefestigung eine Ausdehnung in horizontaler Richtung blockiert und sich der Sensor somit gesamtheitlich in vertikaler Richtung bewegt. Die geringen Auslenkungen in x -Richtung können von zusätzlichen Sensordeformationen und Linsenverformungen hervorgerufen werden. Die Auswertung mehrerer Abkühlkurven hat ergeben, dass das Ergebnis sehr stark von der gleichmäßigen Erwärmung der Kamera abhängig ist. Ein intensiver Temperaturtest in einer maßgefertigten Klimakammer wird in [59] durchgeführt, welche neben dem Aufwärmen und Abkühlen auch die gezielte Minderung der Luftfeuchte ermöglicht, um das Beschlagen der Linse während der Temperaturanpassung zu verhindern. Die Aufnahmen werden mit einem ähnlichen Kameramodell derselben Firma durchgeführt, welches hier verwendet wurde. Die Ergebnisse zeigen bereits einen Einfluss des Hochfahrens der Kamera nach einem Kaltstart. Unter einer Einspannung des Sensors und der Kameralinse an einem Werkstoff mit niedrigem Wärmeleitkoeffizienten konnten die Fehler in x - und

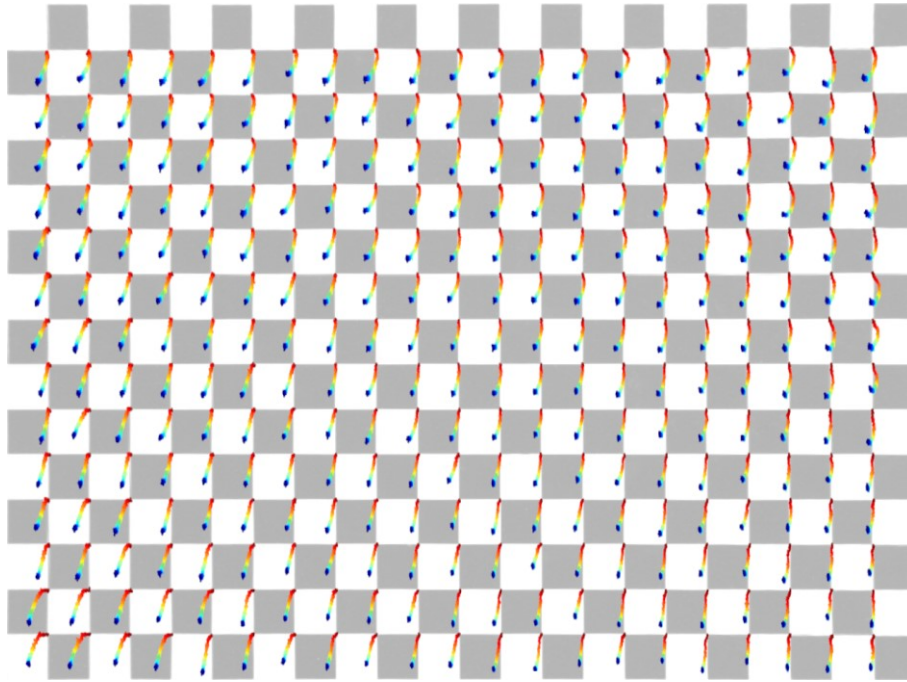


Abbildung 3-5 Temperatureinfluss auf die Punktdetektion. Zur besseren Visualisierung wird die Änderung der Bildkoordinaten mit dem Faktor 50 multipliziert und die Helligkeit des Schachbrettmusters erhöht.

y -Richtung um bis zu 48 % bzw. 85 % reduziert werden. Weiterhin geht aus [59] hervor, dass die Änderungen der Bildpunkte von Temperaturabweichung und nicht der absoluten Temperatur abhängig sind. Insofern sollte die Kamera nach dem Hochfahren und Erreichen einer eingeschwungenen Betriebstemperatur kalibriert und betrieben werden. Zudem wird festgehalten, dass der Effekt der Punktverschiebung beim Abkühlen und Erwärmen ein lineares Verhalten aufweist, welches nahezu identisch ist [59, 60]. Da die Bewegung im gesamten Sensorfeld nahezu identisch verläuft, kann der Temperatureffekt mithilfe einer Hauptpunktverschiebung ausgeglichen werden. Alternativ werden zusätzlich die Verzeichnungsparameter angepasst und als Lookup-Table in der Kamera hinterlegt.

Bei der vorliegenden Aufnahmeconfiguration verursacht die Temperaturänderung einen maximalen Fehler der detektierten Koordinate in x -Richtung von 0.68 Pixel und in y -Richtung von 1.06 Pixel. Dieser Einfluss wird allerdings nur sichtbar, wenn die Kamera mit einer großen Abweichung von ihrer normalen Betriebstemperatur betrieben wird. Während der Aufnahme der Bildfolge ist davon auszugehen, dass die Temperaturänderungen zwischen zwei Bildern bzw. der mittleren Tracklänge von 20 Bildern so gering sind, dass der Einfluss auf die Merkmalsverfolgung vernachlässigt werden kann. Nach [60] liegt die Änderung der Brennweite der untersuchten Digitalkamera bei ca. 0.1 % mit Temperaturänderung von 30 °C. Im vorliegenden Fall kann der Einfluss jedoch nicht exakt bestimmt werden. Es wird von einem maximal zugelassenen Gesamtfehler der Brennweite von 2 Pixel ausgegangen, der sich aus den Fehlern der initialen Kamerakalibrierung, der mechanischen Instabilität und den Temperaturschwankungen zusammensetzt.

3.4 Beurteilung

Die ermittelten Fehlergrößen, die sich während der Kamerakalibrierung und dem darauffolgenden Betrieb der Kamera im Fahrzeug ergeben werden herangezogen, um deren Einfluss auf die Höhenberechnung eines Hindernisses zu bewerten. Als entscheidender Faktor ist in diesem Zusammenhang die Temperaturänderung zu sehen, sofern davon ausgegangen wird, dass die Kalibrierung mittels 3D-Testfeld und während der eingeschwungenen Betriebstemperatur der Kamera vorgenommen wird, die im besten Fall mithilfe einer Selbstkalibrierung während der Fahrt oder über eine tägliche Kalibrierung unter Nutzung von Zielmarken verbessert wird.

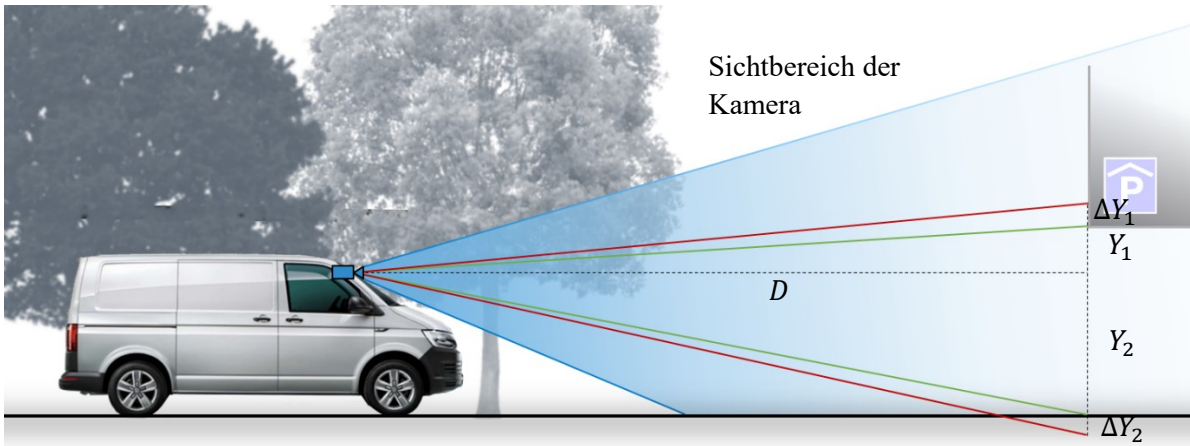


Abbildung 3-6 Simulationsmodell zur Ermittlung der Genauigkeit der bestimmten Hindernishöhe unter Einfluss der Kamerakalibrierung.

Das verwendete Modell zur Fehlerbestimmung kann Abbildung 3-6 entnommen werden. Es nutzt dabei die Kollinearitätsgleichungen (Gl. (3-1), Gl. (3-2)) zur Abbildung der beiden Höhen Y_1 und Y_2 :

$$Y_1 = \frac{D}{f} (y'_1 - y'_0) \quad \text{Gl. (3-3)}$$

$$Y_2 = \frac{D}{f} (y'_2 - y'_0). \quad \text{Gl. (3-4)}$$

Aufgrund des negativen Wertes des Terms $(y'_2 - y'_0)$ werden die bestimmten Höhen voneinander abgezogen, um die Gesamthöhe zu berechnen.

$$H = Y_1 - Y_2. \quad \text{Gl. (3-5)}$$

Weiterhin gilt:

$$H = \frac{D}{f} (y'_1 - y'_2) \quad \text{Gl. (3-6)}$$

Den Einfluss der Brennweite kann demnach mit

$$\frac{\partial H}{\partial f} = -\frac{1}{f^2} D(y'_1 - y'_2) \quad \text{Gl. (3-7)}$$

berechnet werden. Der letzte Term ist bei gleichbleibender Höhe und Brennweite immer konstant, sodass sich unter Berücksichtigung von Gl. (3-6) folgende Vereinfachung ergibt:

$$\frac{\partial H}{\partial f} = -\frac{1}{f^2} fH = -\frac{H}{f} \quad \text{Gl. (3-8)}$$

Damit wird ersichtlich, dass die Brennweite ausschließlich einen Einfluss auf bestimmte Hindernishöhen hat, die jedoch entfernungsunabhängig sind. Die Einbaulage der Kamera wird über die Parameter der äußeren Orientierung aus Abschnitt 3.3.1.1 berücksichtigt. Die kritische Hindernishöhe wird beispielhaft mit 2 m angenommen. Der Messfehler für die Brennweite wird mit 2 Pixel angenommen. Die Verschiebung des Hauptpunktes wird mit maximal 1 Pixel in x - und 2 Pixel in y -Richtung. Weitere Einflüsse durch die Verzeichnungsparameter können Appendix A entnommen werden (s. Abbildung A-1, Abbildung A-2).

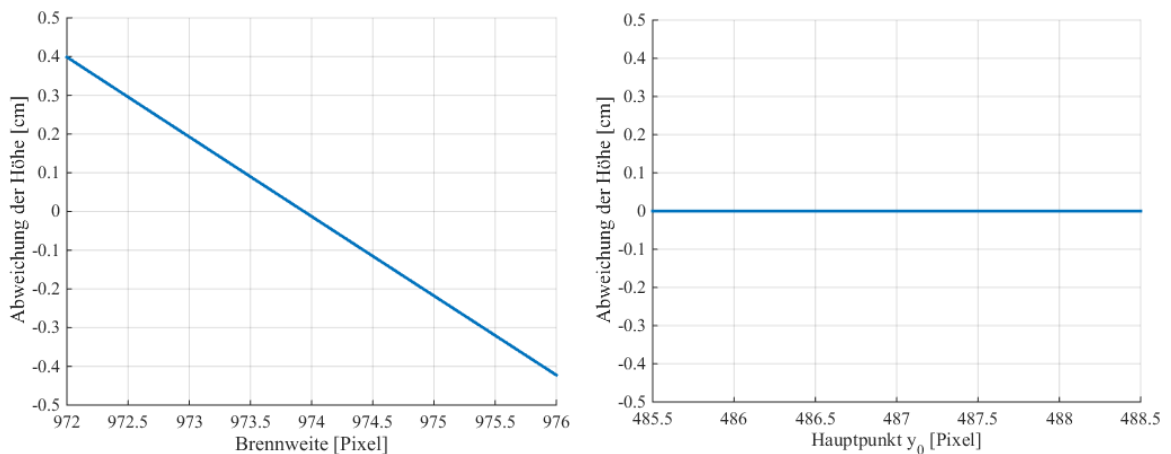


Abbildung 3-7 Einfluss der Brennweite (links) und des Bildhauptpunkts (rechts) auf die Höhenbestimmung für Hindernisse mit einer Höhe von 2 m.

Die in Abbildung 3-7 gezeigten Verläufe zeigen die Abweichung der Hindernishöhe, welche durch Parameterfehler der Brennweite und des Bildhauptpunkts hervorgerufen werden. Die Abweichung für ein 2 m hohes Hindernis durch die Änderung der Brennweite liegt bei ca. 0,2 cm/Pixel. Die Ungenauigkeit des Bildhauptpunktes, hier nur y_0 berücksichtigt, liegt in der Größenordnung von 10^{-3} cm und kann im weiteren Verlauf vernachlässigt werden. Dies ist auf den relativen Abstand der Bildkoordinaten der Hindernisses y_1 und die des Fahruntergrunds y_2 zurückzuführen, der für die Berechnung der Hindernishöhe herangezogen wird. Der Abstand bleibt bei Verschiebung des Hauptpunktes nahezu konstant, sodass sich kein bemerkbarer Fehler in der Höhenbestimmung ergibt.

Die Auswertung zeigt, dass die zuverlässige Bestimmung der inneren Orientierung von vielen Faktoren abhängt. Mithilfe eines 3D-Testfelds und einer geeigneten Aufnahmeconfiguration kann die Qualität der Parameter der Kalibriermatrix im Vergleich zum 2D-Schachbrett deutlich verbessert werden. Der Fehler der Brennweite konnten um ca. 77 % für f_x und 89 % für f_y reduziert werden. Der Fehler des Bildhauptpunkts wurde um 94 % für x_0 und 77 % für y_0 verringert. Die Verzeichnungsparameter werden bereits innerhalb der Kalibrierung mit dem 2D-Testfeld zuverlässig bestimmt. Die bestimmten Parameter beschreiben die innere Orientierung der Kamera zum Zeitpunkt der Testfeldkalibrierung. Im realen Fahrgeschehen wirken jedoch wechselnde äußere Bedingungen

auf die Kamera ein, sodass davon ausgegangen werden kann, dass die Parameter einer Änderung unterliegen. Durch Erschütterungen können bei mechanisch instabilen Kameras vor allem im Bereich der Brennweite und des Bildhauptpunktes Schwankungen vorliegen. Darüber hinaus werden durch unterschiedlichen Umgebungstemperaturen mechanische Ausdehnungen im Kameraaufbau hervorgerufen, die die Bildpunktbestimmung beeinflussen. Die Untersuchung zeigt, dass die Einflüsse auf die Bestimmung von Höhenhindernissen jedoch nahezu vernachlässigbar sind. Dennoch empfiehlt sich eine Selbstkalibrierung, bei der die Parameter während der Fahrt optimiert werden. Alternativ kann ein Verfahren verwendet werden, bei dem die Parameter über definierte Zielmarken innerhalb einer Garage bzw. eines Stellplatzes am Ende eines Tages erneut bestimmt werden. Im gleichen Prozess kann die externe Orientierung abgeglichen werden, die jedoch aufgrund der festen Verklebung und der schweren Erreichbarkeit der Kamera als nahezu konstant angesehen werden kann. Dabei ist zu berücksichtigen, dass die vollständige Pose der Kamera nur entscheidend ist, sofern eine Transformation der 3D-Punktwolke in das für alle Fahrerassistenzsysteme gemeinsame USK erfolgen soll. Wird das System eigenständig betrieben, so müssen allein die Einbauhöhe über Grund und die Entfernung zur Fahrzeugfront bestimmt werden.

4 Fahrzeugeigenbewegung

Die Schätzung der Fahrzeugeigenbewegung ist im vorliegenden Kontext essentiell für die Maßstabsberechnung während der 3D-Rekonstruktion des Fahrumfelds. Der Maßstab wird über die Fahrzeugbewegung bestimmt, indem unterschiedliche Kameraposen innerhalb der Bildfolge ausgewertet werden. In diesem Kapitel sollen die verschiedenen Möglichkeiten zur Berechnung der Eigenbewegung vorgestellt werden. Darüber hinaus werden Möglichkeiten zur Prädiktion dieser beschrieben. Der daraus entstehende prädizierte Fahrschlauch des Fahrzeugs wird im Rahmen der Hindernisdetektion zur Einschränkung des 3D-Raums und damit zur Reduzierung von Fehldetektionen genutzt.

4.1 Methoden zur Eigenbewegungsschätzung

Wichtig für eine genaue 3D-Rekonstruktion ist die Positions- und Richtungsbestimmung des Fahrzeugs, im Besonderen jedoch der Kamera. Über die nachfolgend gelisteten Parameter der externen Orientierung der Kamera, welche die Kamerapose im USK beschreiben, kann die Kamerabewegung im WKS (s. Abschnitt 2.3.1 und Abbildung 2-6) ermittelt werden.

$$X, Y, Z, \Phi, \theta, \Psi$$

Die äußere Orientierung der Kamera im USK wird dabei als fest angenommen, d.h. dass Fehler durch Karosserieverwindung oder eine Verformung der Kamerahalterung z.B. durch Wärme oder Vibrationen vernachlässigt werden.

Die relative Bewegung der Kamera im Weltkoordinatensystem ist dabei essentiell, um eine möglichst genaue Skalierung der 3D-Punktwolke zu erhalten. Um die Bewegung des Fahrzeugs zu bestimmen gibt es mehrere Ansätze, die anschließend vorgestellt werden.

Radodometrie

Die Radodometrie (auch Koppelnavigation) basiert auf Raddrehwinkel-Impulsmessungen durch aktive und passive Magnetfeldsensoren. Dabei wird der zurückgelegte Drehwinkel der einzelnen Räder um dessen Achse in radfesten Koordinaten berechnet. Durch den Radrollradius kann anschließend die Geschwindigkeit des Fahrzeugs an den jeweiligen Radaufstandspunkten berechnet werden. Zur Berechnung der Fahrzeugbewegung können die Drehwinkel und die Abrollumfänge aller vier Räder, der Lenkwinkel an der Vorderachse, die Spurbreite und der Radstand mit einbezogen werden. Die Berechnung erfolgt dann mithilfe eines physikalischen Ein- bzw. Zweipurmodells.

Die Genauigkeit der Eigenbewegungsschätzung ist abhängig von der endlichen Anzahl der Ticks pro Radumdrehung und der Genauigkeit der jeweiligen Parameter, die in die Modellberechnung mit eingehen. Darüber hinaus können umgebungsabhängige Störungen auftreten. Dabei handelt es sich um Schlupf, der durch einen niedrigen Reibwert, hohe Beschleunigungen oder durch die Fahrt durch Schlaglöcher hervorgerufen werden kann.

Durch die Radodometrie lässt sich eine Eigenbewegung schätzen, die eine hohe relative Genauigkeit aufweist, sofern hochdynamische Fahrmanöver ausgeschlossen sind. Die entstehenden Fehler werden jedoch aufsummiert, sodass eine Verzerrung der Koppelnavigation zur tatsächlichen Bewegung des Fahrzeugs entsteht, die mit zurückgelegter Strecke meist zunimmt. Insofern eignet sich dieses Modell zur absoluten Bewegungsschätzung nur für wenige hundert Meter.

Sofern die Radodometrie allein genutzt wird, können nur die Bewegungsparameter X, Y und Ψ berechnet werden, da über die Radticks keine Steigungen oder Gefälle in Längs- bzw. Querrichtung zum Fahrzeug erfasst werden. Dieses System muss demnach zusätzlich durch weitere Informationen gestützt werden, sofern eine genauere Fahrzeugbewegung gefordert ist.

GNSS gestützte Systeme

Bei globalen Navigationssatellitensystemen (GNSS) werden die Abstände der Phasenzentren zwischen Satelliten- und Empfängerantenne durch Laufzeitmessung des Signals erfasst [29]. Die Qualität steigt mit Anzahl der verfügbaren Satelliten und deren Ausrichtung zueinander. Darüber hinaus werden Satellitenuhrfehler und Ephemeridenfehler (Abweichungen zwischen realen und berechneten Satellitenumlaufbahnen) in der Nachricht versendet, um anschließende Korrekturen durchzuführen. Bekannte GNSS Systeme sind bspw. das vom Verteidigungsministerium der russischen Föderation betriebene GLONASS, das unabhängige europäische Satellitensystem Galileo oder das chinesische System Beidou. Vorrangig wird jedoch das ursprünglich für das amerikanische Militär entwickelte NavSTAR GPS verwendet.

Die Wirkweise aller Systeme sind ähnlich und die Positionsgenauigkeit, die in freien Gebieten und günstiger Satellitenkonstellation erreicht werden kann, beträgt dennoch nur wenige Meter. Um dies zu verbessern, kann dGPS (differentielles GPS) genutzt werden. Dabei werden bspw. Funkempfänger, die in Reichweite einer hochgenau vermessenen Basisstation stehen, genutzt (SATEL). Darüber hinaus gibt es Satellitenreferenzdienste, die über GNSS Referenzstationen die Satellitenbewegungen beobachtet und berechnete Korrekturdaten via Internet oder Mobilfunk aussenden (ASCOS). Durch diese Maßnahmen lassen sich Positionsgenauigkeiten von wenigen Zentimetern erreichen. Die GNSS Positionierung kennzeichnet eine langzeitstabile absolute Messgenauigkeit aus, die keine mit der Zeit anwachsenden Sensorfehler hat. Umgebungsbedingungen, wie Beugung und Reflektion von magnetischen Wellen können einen Mehrwegempfang verursachen, bzw. kann in stark bebauten Gebieten auch eine Abschattung des Signalempfangs erfolgen.

IMU basierte Systeme

Eine Inertialmesseinheit, dargestellt in Abbildung 2-1, beinhaltet Beschleunigungs- und Kreiselensoren, die die absoluten und dreidimensionalen Werte der Beschleunigung und Drehrate im Inertialsystem erfassen. Durch Integration der Drehraten erfolgt die Berechnung der Lage. Die Integration der Beschleunigung erlaubt die Berechnung der Geschwindigkeit und durch Integrieren erfolgt die Berechnung der Position. Erd- bzw. fahrzeugfeste Störgrößen wie die Erdbeschleunigung und -drehrate, Coriolisbeschleunigung und Transportrate werden dabei vom System kompensiert. Die Angabe der Fahrzeugpose wird gegenüber einem frei gewählten festen Weltkoordinatensystem angegeben.

Typische Fehler sind Messrauschen, Nullpunktfehler und Skalenfaktorfehler. Von Vorteil ist die stetige Verfügbarkeit der IMU, da sie von äußeren Faktoren unabhängig ist.

Lokalisierung mittels Umfeldsensorik

Alternativ dazu kann die Eigenbewegungsschätzung mittels Umfeldsensorik erfolgen. Neben Ultraschall-, RADAR- oder LiDAR Sensorik bieten sich vor allem Kameras als kostengünstige Alternative an. Dabei werden zunächst markante Merkmale innerhalb eines Bildes extrahiert. Anschließend werden diese Merkmale über verschiedene Techniken wie Merkmalspaarung oder Merkmalstracking über den optischen Fluss in den darauffolgenden Bildern verfolgt. Über die Bewegung der Merkmale innerhalb der Bildfolge werden 3D-Punkte des Umfelds bestimmt. Anschließend kann die Bewegung der Kamera bzw. des Fahrzeugs relativ zu den 3D-Punkten in einem lokalen Koordinatensystem berechnet werden. Dieser Vorgang ist Teil der SfM- und SLAM- Algorithmen. Eine verwandte Art ist die visuelle Odometrie, die den Fokus primär auf die Lokalisierung legt und mit weniger 3D-Punkten rechnet. Da Kameras die Odometrie nur indirekt berechnen, eignen sie sich ohne Referenzmaße im Bild nicht zur Skalierung der eigens aufgebauten rekonstruierten Szene.

Die hohe Spreizung an Ausstattungsvarianten in Nutzfahrzeugen bedingt ein einfaches System zur Eigenbewegungs-schätzung und Bestimmung der Skalierung, da nicht davon ausgegangen werden kann, dass stets hochwertige Sensorik zur Positions- und Lagebestimmung im Fahrzeug verbaut ist. Demnach bietet sich das System zur Berechnung der Radodometrie an. Zur Evaluierung dieses Systems wird ein Referenzsystem genutzt, welches die Ansätze der dGPS-gestützten Navigation und der IMU (vgl. Kapitel 2.1) verknüpft.

4.2 Bewegungsmodell

In diesem Kapitel werden die Grundlagen zum genutzten Bewegungsmodell vorgestellt. Ziel dabei ist es ein möglichst einfaches Bewegungsmodell zu nutzen, welches zum einen schnell berechnet werden kann und zum anderen möglichst der tatsächlichen Fahrzeugbewegung entspricht. Eine Auswertung zeigt dabei die auftretenden Fehler im Vergleich zur Referenz.

Die Fahrdynamik kann zunächst über das lineare Einspurmodell beschrieben werden. Dieses physikalische Modell wurde in [61] vorgestellt und eignet sich für Systeme der Fahrumgebungserfassung, welche nicht in dynamischen Grenzbereichen agieren [29, 62]. Bei diesem Modell werden beide Räder einer Fahrzeugachse in Achsmitte zusammengefasst und Nick-, Wank- sowie Hubbewegungen vernachlässigt. Es werden nur die Translation in der Ebene (X, Y) und die Rotation um die vertikale Achse (Ψ) beschrieben. Weitere Annahme ist die Vernachlässigung von angreifenden Reifenmomenten, Roll- und Luftwiderstand, sodass die Fahrdynamik auf die Kinematik reduziert werden kann. Die Fahrzeugmasse wird im Schwerpunkt zusammengefasst, welcher vereinfacht in den Ursprung des USK verschoben wird. Die sich ergebenden geometrischen Abhängigkeiten werden in Abbildung 4-1 dargestellt.

Die Rotation des Fahrzeugs bei Kurvenfahrt erfolgt auf imaginären Kreisbahn um den Kreismittelpunkt M in infinitesimal kleinen Zeitschritten $\Delta\tau$. M ist durch den Schnittpunkt der vertikal durch beide Räder verlaufenden Geraden definiert, wobei sich der Radius der Kreisbahn R_H durch den Radstand L und den Lenkwinkel ϕ , welcher sich aus der Mittelung aus ϕ_L und ϕ_R ergibt, berechnet wird.

$$R_H = \frac{L}{\tan \phi} \quad \text{Gl. (4-1)}$$

Die Änderung der Orientierung $\Delta\Psi$ [°] ergibt sich zu

$$\Delta\Psi = \psi \cdot \Delta\tau = \frac{v}{R_H} \cdot \Delta\tau = \frac{v}{L} \tan \phi \cdot \Delta\tau. \quad \text{Gl. (4-2)}$$

wobei durch ψ die Gierrate in [1/s] beschrieben wird. Die Geschwindigkeit v wird über die Mittelung der vorderen und hinteren Radticks und Abrollumfänge im Zeitintervall $\Delta\tau$ gegeben. Die Änderung der Positionen $\Delta X_F, \Delta Y_F$ [m] wird mit

$$\begin{aligned} \Delta X_F &= v \cos \Delta\Psi \cdot \Delta\tau, \\ \Delta Y_F &= v \sin \Delta\Psi \cdot \Delta\tau \end{aligned} \quad \text{Gl. (4-3)}$$

berechnet.

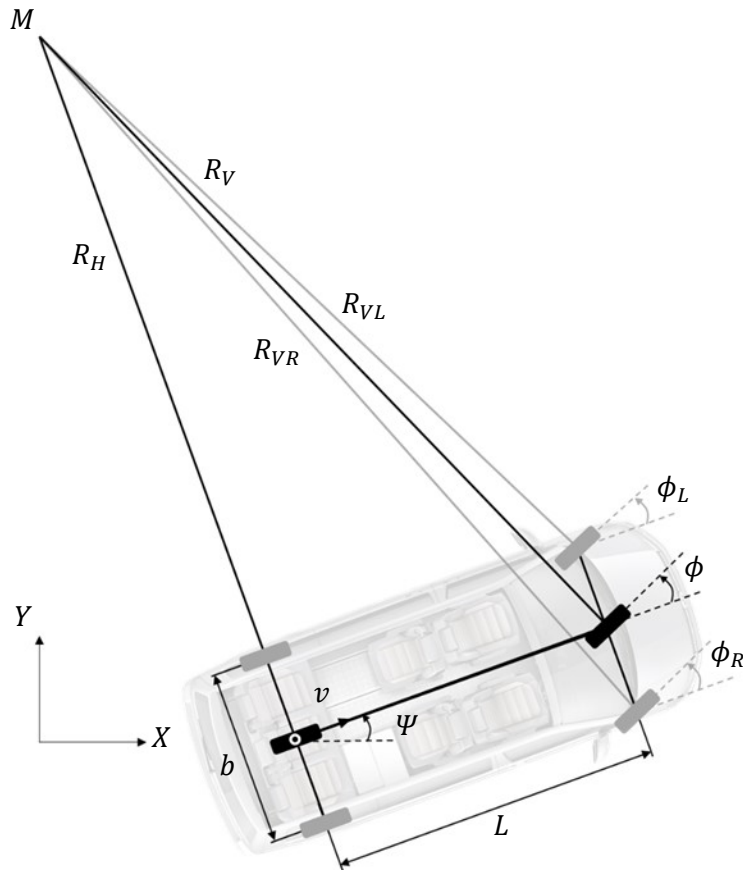


Abbildung 4-1 Schema. Kinematisches Einspurmodell

Die Genauigkeit des Ansatzes ist abhängig von den auf dem CAN bereitgestellten Daten zur Fahrzeuggeschwindigkeit und Lenkwinkel. Der in diesem Ansatz genutzte Lenkwinkel leitet sich aus dem Lenkradwinkel ab und erfordert Kenntnis der Übersetzung innerhalb des Lenkgestänges und die genaue Ausrichtung von Lenkrad und Vorderrädern. Die Ungenauigkeit dieser Parameter rufen einen Drift hervor, welcher die berechnete Position mit steigender Distanz von der tatsächlichen Position abweichen lässt. Beispielhaft sind die berechneten Fahrzeugpositionen in X/Y und Ψ in Abbildung 4-2 für eine Sequenz dargestellt.

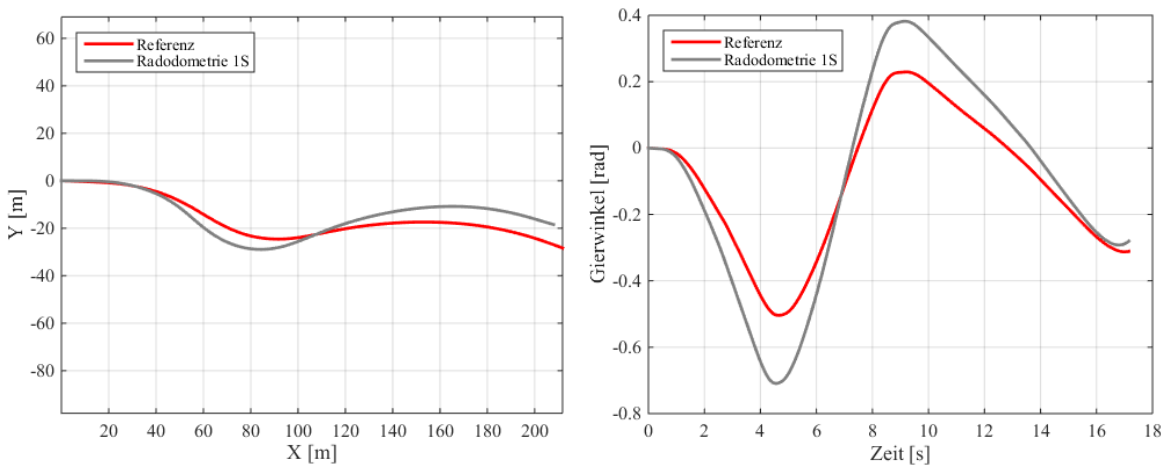


Abbildung 4-2 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S

Fahrzeugeigenbewegung

Die Bewegungsschätzung mittels Einspurmodell zeigt einen deutlichen Drift und lässt auf eine ungenaue Berechnung des Lenkwinkels schließen. Durch Nutzung der Radticks und der Spurbreite b können die Berechnungen der Fahrzeugbewegung redundant innerhalb eines Zweispurmodells ausgeführt und die Robustheit der Odometrie gesteigert werden. Die Fusion beider Ansätze erfolgt mittels Kalman Filter.

Zur Sensierung der Radumdrehung sind beim Volkswagen T6 an der Radnabe 48 magnetische Zahnradflanken angebracht. Während der Rotation werden jeweils die steigende und fallende Flanke mittels Hall-Element detektiert, sodass sich pro Radumdrehung eine Gesamtzahl von 96 Ticks ergeben. Die Wegstrecke eines Rades auf der Kreisbahn kann mit

$$\Delta X = \frac{N_{Tick}}{I_{Tick}} u \quad \text{Gl. (4-4)}$$

berechnet werden, wobei N_{Tick} die Anzahl der gemessenen Ticks, I_{Tick} die Gesamtanzahl der Ticks pro Radumdrehung und u den Abrollumfang des Rades beschreiben. Die lineare Bewegung des Fahrzeugs berechnet sich dabei aus der Mittelung von linken und rechten Hinterrad.

$$\Delta X = \frac{\Delta X_{HL} + \Delta X_{HR}}{2}. \quad \text{Gl. (4-5)}$$

Die Rotation um die vertikale Achse erfolgt über

$$\Delta \Psi = \arctan\left(\frac{\Delta X_{HR} - \Delta X_{HL}}{b}\right). \quad \text{Gl. (4-6)}$$

Unter Berücksichtigung der Zeitschritte ergibt sich das Gleichungssystem zu

$$\begin{bmatrix} X(t+1) \\ Y(t+1) \\ \Psi(t+1) \end{bmatrix} = \begin{bmatrix} X(t) \\ Y(t) \\ \Psi(t) \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta \Psi \end{bmatrix}, \quad \text{Gl. (4-7)}$$

mit dem die Fahrzeugpose bestimmt werden kann. Demzufolge kann in jedem Zeitschritt dem von der Kamera aufgenommenen Bild eine Pose zugewiesen werden, die es erlaubt einen Maßstab für die Rekonstruktion der Fahrzeugumgebung zu berechnen.

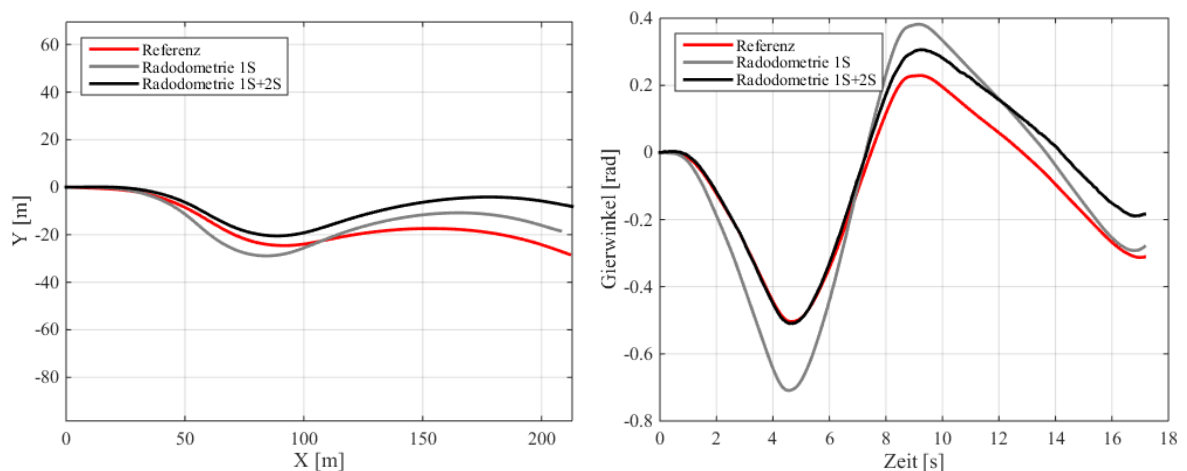


Abbildung 4-3 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S und kombinierter Ansatz 1S+2S

Abbildung 4-3 zeigt, dass durch die Nutzung der Radticks die Eigenbewegungsschätzung stabilisiert wurde. Die Bewegung gleicht zunehmend der Referenz, dennoch ist ein Drift ab ca. 50 m gefahrener Wegstrecke sichtbar. Die Berechnung der Radodometrie erfolgt hierbei noch unter Berücksichtigung der auf den Spezifikationen entnommenen Standardwerte für die Abrollumfänge, die Spurbreite und den Achsabstand, welche in folgendem Parametervektor zusammengefasst werden

$$\mathbf{z}_{sim} = \begin{bmatrix} u_{VL} \\ u_{VR} \\ u_{HL} \\ u_{HR} \\ b \\ L \end{bmatrix}. \quad \text{Gl. (4-8)}$$

Die absolute Eigenbewegung des Fahrzeugs wird im Rahmen der Fahrzeugfunktion zur Überwachung der visuellen Odometrie genutzt. Steigt der absolute Fehler zwischen visueller und Radodometrie, so wird das Verfahren und die Maßstabsberechnung reinitialisiert. Für die Berechnung der Skalierung ist hingegen die relative Pose von aufeinanderfolgenden Bildern entscheidend.

4.3 Optimierung der Radodometrie

Die Auswertung des letzten Abschnitts zeigen, dass ein Odometriefehler bei nicht vorgenommener Kalibrierung in der absoluten Fahrzeugposition vorliegt. Um diesen zu verringern, wird eine Anpassung der berechneten Fahrzeugbewegung an die Referenzbewegung mittels Parameteroptimierung vorgenommen. Im Serienfahrzeug ist bereits ein Algorithmus integriert, welcher Abweichungen der Abrollumfänge im Fahrbetrieb feststellt und ggf. Anpassungen vornimmt. Dies geschieht über die Beobachtung der Radticks während der langsamen Geradeausfahrt. Dennoch müssen geeignete Startparameter gefunden werden, damit dieser Prozess möglichst effizient konvergiert.

Ziel der hier vorgestellten Optimierung ist die Anpassung der theoretischen Größen der Spurbreite, des Radstands und der Abrollumfänge der vier Räder an die tatsächliche Größe. Bevor jedoch genauer auf die Größen eingegangen wird, soll eine Einführung in die Optimierungstheorie erfolgen.

4.3.1 Optimierungstheorie

Die im folgenden Abschnitt vorgestellten Grundlagen der Optimierungstheorie orientieren sich an [63].

Globales Optimum

Die Optimierung beschreibt einen Prozess, der auf die Lösung eines Problems abzielt. Dabei liegt eine Situation vor, bei der eine Entscheidung mit einer Wahl aus einer Menge verschiedener Alternative erfolgen soll. Die Alternativen unterscheiden sich dabei, dass Einflussgrößen auf die Lösung entweder minimiert oder maximiert werden. Aus mathematischer Sicht soll ein Wert x^* gefunden werden, für welchen die mathematische Funktion f den kleinstmöglichen bzw. größtmöglichen Wert einnimmt.

Fahrzeugeigenbewegung

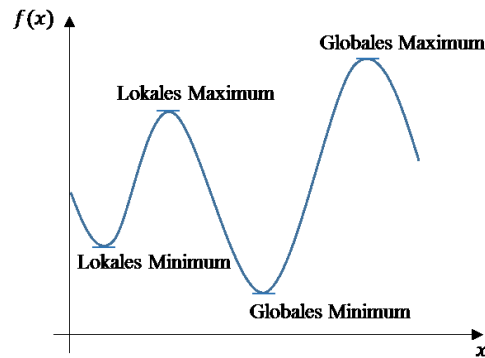


Abbildung 4-4 Extremwerte als Kandidaten für die Optimierung

Beschrieben wird das Finden einer Lösung für x^* , mit der Eigenschaft, dass keine weitere, bessere Lösung existiert. Dabei steht im Vordergrund die Lösungsfindung eines globalen Minimums bzw. Maximums, ohne dass die Suche nach dem Finden des lokalen Optimums beendet wird.

Problemraum

Am Anfang einer Optimierung steht die Definition des Problemraums, wobei dieser aus einer Menge \mathcal{X} aller möglichen Lösungen x^* besteht. Für ein Optimierungsproblem können dabei verschiedene Problemräume definiert werden. Innerhalb des Problemraums \mathcal{X} wird das Element x^* als Lösungskandidat bezeichnet. In der vorliegenden Arbeit wird ein evolutionärer Optimierungsalgorithmus verwendet, sodass die Begrifflichkeiten an die Natur angelehnt sind. Problemraum wird durch Phänom und ein Lösungskandidat durch Phänotyp beschrieben. Des Weiteren ist das Phänom eingeschränkt, was auf die Darstellungsprobleme eines Computers im Hinblick auf reelle Zahlen zurückzuführen ist.

Zielfunktion

Die Zielfunktion beschreibt das Maß für das Optimum. Parameter der Zielfunktion sind die Lösungskandidaten x^* aus dem Problemraum \mathcal{X} . Sie liefert einen Wert, welcher der Qualität der Lösung entspricht. Im vorliegenden Fall ist die Zielfunktion Gegenstand einer Minimierung, wobei der relative Fehler zwischen Radodometrie und dGPS/IMU basierter Odometrie minimiert werden soll. Die Zielfunktion muss nicht zwangsweise mathematisch definierbar sein.

Optimierungsalgorithmen

Optimierungsalgorithmen können in zwei Kategorien eingeteilt werden. Dabei ist die Rede von deterministischen und probabilistischen Algorithmen.

Bei einem deterministischen Optimierungsalgorithmus gibt es pro Ausführungsschritt nur maximal eine Möglichkeit zum weiteren Vorgehen. Kann keine Lösung berechnet werden, so wird der Algorithmus terminiert. Werden beispielweise die Lösungen x_1^* und x_2^* berechnet und es gilt $f(x_1^*) < f(x_2^*)$, dann wird entweder x_1^* oder x_2^* gewählt. Gleiche Eingangsdaten liefern gleiche Ergebnisse. Diese Art von Optimierungsalgorithmus wird eingesetzt, wenn Entscheidungen effizient anhand der zur Verfügung stehenden Daten getroffen werden können. Sie sind sinnvoll, wenn Nutzwert und Charakteristik der Lösungen einen klaren Zusammenhang besitzen. Bekannt sind bspw. Suchalgorithmen wie A*, Breadth-First Search oder Depth-First Search.

Dagegen eignen sich probabilistische Optimierungsalgorithmen bei sehr komplexen, dynamischen Problemen, bei denen die Beziehung der Lösbarkeit des Problems und der Lösungskandidaten nicht offensichtlich ist. Dabei basiert mindestens eine Instruktion des Algorithmus auf Zufallszahlen,

wodurch er nicht deterministisch, also probabilistisch/stochastisch wird. Durch die Möglichkeit der Auswahl mehrerer Lösungen wird bspw. ein Abbrechen des Algorithmus‘ in einem lokalen Minimum erschwert. Beispiele sind Hill Climbing, Tabu Search, evolutionäre Algorithmen oder Schwarmintelligenz. Die letzten beiden stellen dabei eigene Klassen mit verschiedenen Ausprägungen dar.

4.3.2 Implementierung

Zur Verbesserung der Radodometrie soll ein evolutionäres probabilistisches Optimierungsverfahren herangezogen werden. Ziel ist es den relativen für die Skalierung entscheidenden Fehler zwischen dGPS/IMU Odometrie als Referenz und Radodometrie zu verringern.

Folgende Schritte werden dabei durchlaufen:

- Definition des Problemraums \mathcal{X} inkl. Einschränkungen der Elemente und Rahmenbedingungen
- Definition der Zielfunktion f oder mehrerer Zielfunktionen $f_1 \dots f_n$
- Definition eines Komparators $c(x_1^*, x_2^*)$, welcher zwischen den Lösungskandidaten auf Basis der Zielfunktion auswählt
- Auswahl des Optimierungsalgorithmus‘
- Konfiguration des Algorithmus‘
- Optimierungsprozess

Problemraum

Die primären Einflussgrößen auf die Eigenbewegungsschätzung in der X/Y -Ebene sind die vier Abrollumfänge der Räder, die Spurbreite und der Radstand. Diese sind jedoch mit Messunsicherheiten behaftet, deren Rahmenbedingungen hier definiert werden.

Reifenumfänge

Die Darstellung der richtigen Geschwindigkeit im Kombiinstrument und die optimale Abstimmung der Bremsregelsysteme im Serienfahrzeug erfordern unter anderem die Kenntnis des Abrollumfanges der Reifen. Dieser ist von zahlreichen Faktoren wie Geschwindigkeit, Temperatur, Luftdruck, Abnutzung und Produktionsungenauigkeit abhängig. Der Umfang wird mit einer Toleranz von +1,5% und -4,5% versehen, da außerhalb dieses Bereichs eine Tachoangleichung erforderlich wäre.

Spurbreite

Die Spurbreite ist abhängig vom zugrundeliegenden Fahrwerk und ändert sich je nach Reifenbreite. Sie wird von Federwegänderungen beeinflusst, die hier jedoch vernachlässigt werden. Die Spurbreite wird mithilfe der Achsvermessung am Ende der Fahrzeugproduktion bestimmt, wobei die Genauigkeit bei ca. 1 mm liegt.

Achsabstand

Der Achsabstand ist abhängig von produktionstechnischen Genauigkeiten, die hier als kleiner 1 mm angenommen werden.

Der Parametervektor wird somit durch den Messunsicherheitsvektor

$$\mathbf{z}_{real} = \begin{bmatrix} s_{u_{VL}} \\ s_{u_{VR}} \\ s_{u_{HL}} \\ s_{u_{HR}} \\ s_b \\ s_L \end{bmatrix} \quad \text{Gl. (4-9)}$$

ersetzt.

Eine weitere Einschränkung des Problemraums wird durch Hinzufügen einer Auflösung der einzelnen Parameter erreicht. Der somit verkleinerte Suchraum bildet die Grundlage für die schnelle Suche nach einem globalen Optimum.

Zielfunktion

Nach der Definition des Problemraums wird nun die Zielfunktion aufgestellt. Die Eigenbewegungsschätzung ist letztlich auf sechs Fahrzeugparameter zurückzuführen. Diese sollen so parametrisiert werden, dass der relative Fehler der Eigenbewegungsschätzung zwischen Rad- und Referenzodometrie möglichst gering ist. Dazu werden die Abweichungen des Bewegungsvektors L und des resultierende Winkels in jedem Zeitschritt $n := \{1, N\}$ mittels

$$cost_n = |L| + \Delta\Psi \quad \text{Gl. (4-10)}$$

berechnet, wobei

$$|L| = \sqrt{(\Delta X)^2 + (\Delta Y)^2} \quad \text{Gl. (4-11)}$$

gilt. Anschließend werden die Kosten über die gesamte Sequenz zur folgenden Zielfunktion aufsummiert.

$$cost = \sum_{n=1}^N cost_n \quad \text{Gl. (4-12)}$$

Nun lässt sich die Komparatorfunktion aufstellen, mit der zwei Lösungskandidaten x_1^* und x_2^* verglichen werden. Im Sinne der Fehlerminimierung, wird die Komparatorfunktion den Lösungskandidaten bestimmen, dessen Zielfunktion die kleineren Fehler bzw. Kosten aufweist

$$c(x_1^*, x_2^*) = \begin{cases} x_1^*, & \text{falls } f_{sum}(x_1^*) \leq f_{sum}(x_2^*) \\ x_2^*, & \text{sonst.} \end{cases} \quad \text{Gl. (4-13)}$$

Mithilfe der Komparatorfunktion können nun die Lösungskandidaten x^* im Problemraum \mathcal{X} bezüglich ihrer Kriterien zur Fehlerminimierung innerhalb der Eigenbewegungsschätzung miteinander verglichen werden.

Optimierungsalgorithmus

Zur Optimierung der Eigenbewegungsschätzung wird ein evolutionärer Algorithmus verwendet, dessen Prinzip im Folgenden erläutert werden soll. Nach [63] basiert er auf Populationen und wird den metaheuristischen Optimierungsalgorithmen zugeordnet. Angelehnt an die biologische

Evolution finden sich Mechanismen wie Mutation, Kreuzung, natürliche Selektion oder Überleben des Stärkeren wieder.

Folgende Schritte werden beim evolutionären Optimierungsalgorithmus durchlaufen:

1. Lösungskandidaten x^* werden im Rahmen der evolutionären Optimierungstheorie als Individuen $\bar{\omega}$ bezeichnet. Zunächst wird die erste Generation $\gamma = 1$ der Population \mathcal{P} mit $n = |\mathcal{P}|$ Individuen $\bar{\omega}$ zufällig erzeugt. Dabei ist $\mathcal{P} \subset \mathcal{X}$.
2. Nach Erstellen der ersten Generation wird eine Schleife durchlaufen. Diese kann durch Finden des lokalen bzw. globalen Optimums oder durch Erreichen der maximalen Iterationsanzahl terminiert werden.
3. Während der Evaluierung werden für verschiedene Parametersätze, d.h. verschiedene Individuen $\bar{\omega}$, jeweils die Zielfunktionen $f(\bar{\omega})$ berechnet.
4. Die Fitnessfunktion $f_{fit}(\bar{\omega})$ bestimmt die Überlebensfähigkeit eines Individuums $\bar{\omega}$ innerhalb der Population \mathcal{P} . Hierbei handelt es sich im einfachsten Fall um die Anzahl der Individuen, deren Zielfunktionsresultate schlechter sind.
5. Während der Auslese werden die stärksten Individuen für die Reproduktion bestimmt. Die Ziel- bzw. Fitnessfunktion liefert hierbei das Auswahlkriterium. Die fittesten Individuen werden demnach als Elternmenge \mathcal{F} in der nächsten Generation γ_{i+1} berücksichtigt. Dabei gilt: $\mathcal{F} \subset \mathcal{P} \subset \mathcal{X}$.
6. Die Reproduktionsphase bildet durch Rekombination bzw. Kreuzung der Individuen $\bar{\omega} \in \mathcal{F}$ und anschließender Mutation Nachkommen.
7. Bei Erreichen eines der Terminierungskriterien, stoppt der Algorithmus an dieser Stelle und gibt das bisher stärkste Individuum $\bar{\omega} \in \mathcal{F}$ aller Generationen aus.

In der Optimierung werden verschiedene Sequenzen mit möglichst große Variation genutzt, um Optimierungen auf bestimmte Fahrsituationen zu vermeiden.

4.3.3 Auswertung

Die optimierten Ergebnisse der absoluten Position des Fahrzeugs in der X/Y-Ebene und der Verlauf des Gierwinkels können Abbildung 4-5 entnommen werden.

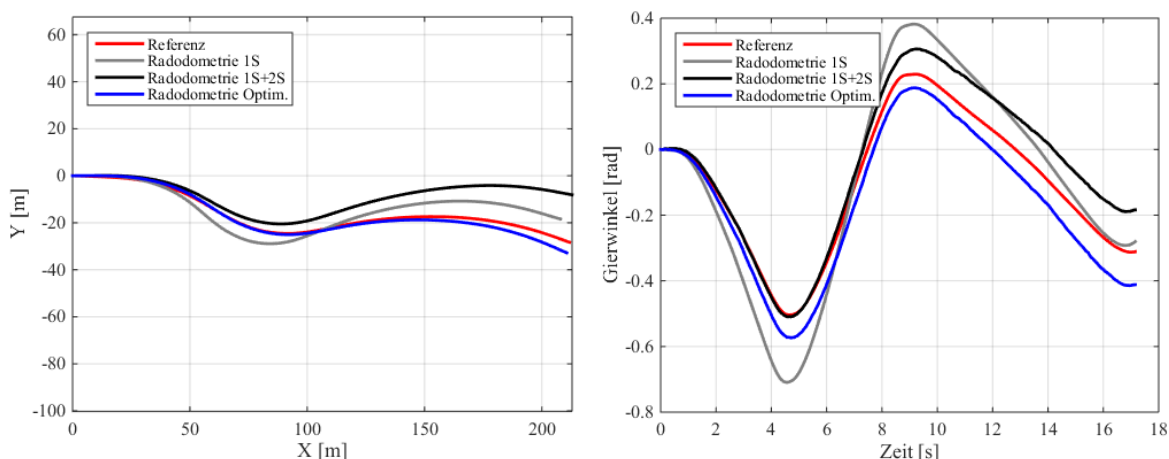


Abbildung 4-5 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S, kombinierter 1S+2S Ansatz und optimierter 1S+2S Ansatz

Obwohl durch das kombinierte Verfahren aus Einspur- und Zweispurmodell bereits die Berechnung der Odometrie in Kurven verbessert wurde, erreicht dieser Ansatz einen höheren absoluten Fehler

als das Einspurmodell allein. Die Verwendung von Radticks kompensiert demzufolge kleine Fehler im Lenkwinkel, verursacht jedoch bei Nutzung der Standardparameter für den Parametervektor aus Gl. (4-9) einen höheren Drift quer zur Fahrtrichtung. Erst durch eine Optimierung kann der absolute Fehler auf größere Distanz minimiert werden.

In Abbildung 4-6 ist die Verteilung des für die Maßstabsberechnung entscheidenden relativen Fehlers dargestellt. Auch hier zeigt sich, dass die Nutzung des kombinierten Ansatzes ohne Optimierung keine Vorteile gegenüber dem Einspurmodell liefert. Erst die Anpassung des Parametervektors ermöglicht die Minimierung des Gesamtfehlers des Bewegungsvektors, welcher sich im vorliegenden Beispiel auf maximal ca. 2.5 cm beläuft.

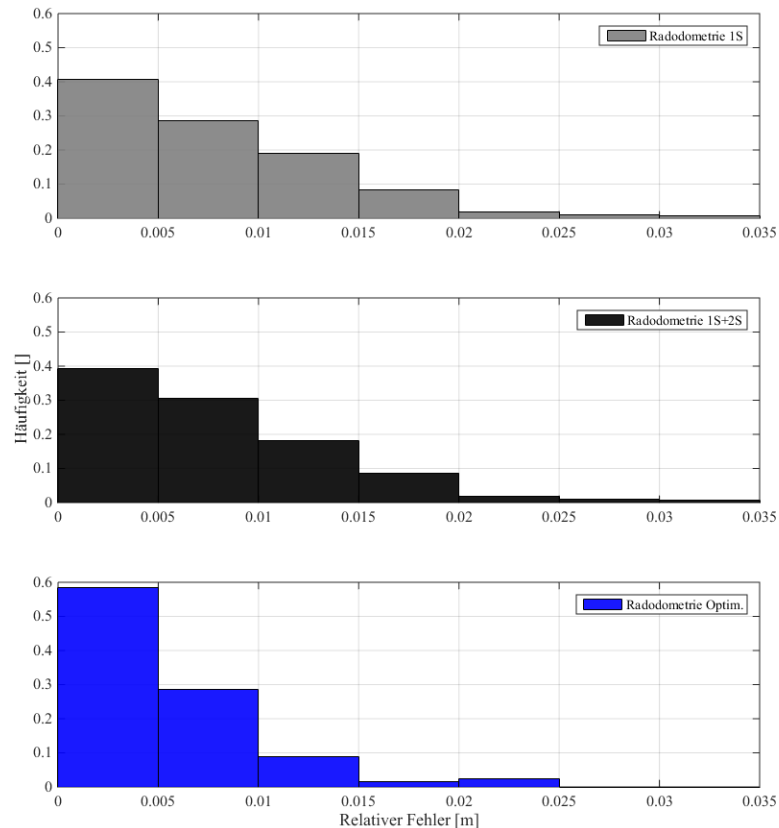


Abbildung 4-6 Relativer Fehler der Radodometrie. Einspurmodell 1S, kombinierter 1S+2S und optimierter 1S+2S Ansatz

Eine Analyse der einzelnen Komponenten zeigt, dass der Fehler vorwiegend durch die Bewegung quer zur Fahrtrichtung hervorgerufen wird. Durch die Untersuchung von neun Testsequenzen, die neben Geraden, Kurvenfahrten und Steigungen auch langsame und schnellere Abschnitte enthalten, ergeben sich für die optimierte Radodometrie folgende Fehler:

$$e_{\Delta X} = 0.02 \pm 0.5 \text{ cm}$$

$$e_{\Delta Y} = -0.2 \pm 0.8 \text{ cm}$$

$$e_{\Delta Z} = -0.3 \pm 0.7 \text{ cm}$$

$$e_G = -0.01 \pm 0.8 \text{ cm}$$

Die maximale Abweichung des relativen Bewegungsvektors wurde innerhalb aller Sequenzen mit 2.8 cm gemessen und wird für die anschließende Berechnung der Skalierung des 3D-Raums als Fehler berücksichtigt.

4.4 Skalierung der Szene

Das Verfahren verfügt ausschließlich über die monokulare Kamera zur Erfassung und Rekonstruktion der Szene. Da die Kamerasequenz allein keine metrische Entfernung und Ausdehnung von erkannten Objekten liefert, muss eine Maßstabsberechnung durch zusätzliche Methoden bzw. Sensoren erfolgen. In [64] wird ein Algorithmus vorgestellt, welcher die absoluten Positions- und Geschwindigkeitsdaten der Inertialplattform durch die Hinzunahme von Bildinformationen mittels Spline Fitting und Kalman Filter verbessert und den Maßstab schätzt. Unter der Annahme der Bekanntheit des Fahruntergrunds wird in [65] ein Modell erstellt, welches den Drift im Maßstab kompensiert. In [66] wird eine Methode vorgestellt, bei der die Kameraeinbauhöhe und eine in die 3D-rekonstruierte Umgebung geschätzte und durch Fluchtpunkte verbesserte Bodenebene zur Berechnung des Maßstabs genutzt wird.

Im Gegensatz zu diesen Verfahren, werden hier die durch die Radodometrie bestimmte Fahrzeugposition in X und Y zur Maßstabsberechnung herangezogen. Dafür wird bei Start des Verfahrens zunächst eine Initialisierungsphase durchlaufen, bei der mittels Bündelausgleichung Kameraposen und 3D-Punkte optimiert und die Szene skaliert wird. Die Initialisierung erfolgt über eine definierte Basislänge B , welche sich aus einer Anzahl an Bewegungsvektoren L zusammensetzt (s. Abbildung 4-7). Während der obere Pfad durch die Radodometrie berechnet wird, zeigt der untere Pfad die äquivalente unskalierte Bewegung, welche durch ein Merkmalstracking und anschließender Positionsberechnung aus Kamerainformationen errechnet wird (visuelle Odometrie). Zur Vereinfachung werden die Zeitstempel der Kamerabilder und der Radodometrie vorher synchronisiert.

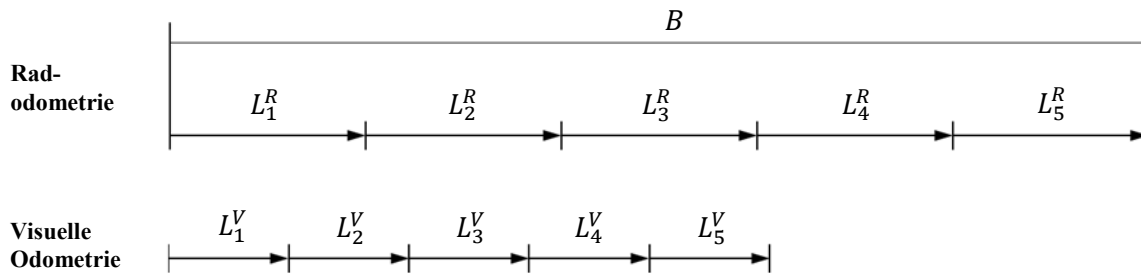


Abbildung 4-7 Bewegungsvektoren der Radodometrie und visuellen Odometrie

Die Berechnung des Maßstabs m erfolgt mithilfe Bewegungsvektoren und lässt sich in zwei verschiedenen Varianten berechnen.

$$m = \text{median} \left(\frac{L_1^R}{L_1^V}, \frac{L_2^R}{L_2^V}, \frac{L_3^R}{L_3^V}, \frac{L_4^R}{L_4^V}, \frac{L_5^R}{L_5^V}, \dots, \frac{L_n^R}{L_n^V} \right) \quad \text{Gl. (4-14)}$$

$$m = \text{mean} \left(\frac{L_1^R}{L_1^V}, \frac{L_2^R}{L_2^V}, \frac{L_3^R}{L_3^V}, \frac{L_4^R}{L_4^V}, \frac{L_5^R}{L_5^V}, \dots, \frac{L_n^R}{L_n^V} \right) \quad \text{Gl. (4-15)}$$

Welche Methode angewendet wird, hängt von der Genauigkeit der Radodometrie zusammen. Da die relative Bewegungsschätzung mittels Radodometrie mit einem maximalen Fehler des Bewegungsvektors von 2.8 cm sehr genau ist und keine Ausreißer detektiert wurden, eignet sich der Mittelwert zur Berechnung. Mithilfe der Standardabweichung der Bewegungsvektoren lässt sich die Standardabweichung des Mittelwertes formulieren (s. Gl. (4-16)), welche sich bei zunehmender Anzahl an Beobachtungen n verringert.

Fahrzeugeigenbewegung

$$s_m = \frac{s_L}{\sqrt{n}} = \frac{2.8 \text{ cm}}{\sqrt{n}}$$

Gl. (4-16)

Mittels einer Simulation wird der Objektdistanzfehler in Abhängigkeit der Basislänge aufgetragen. Somit lässt sich auf eine Mindestlänge B_{min} festlegen, die zur Initialisierung der Szene genutzt wird. Da dies direkt beim Losfahren geschieht, können geringe Geschwindigkeiten angenommen werden, wodurch sich kleine Bewegungsvektoren ergeben.

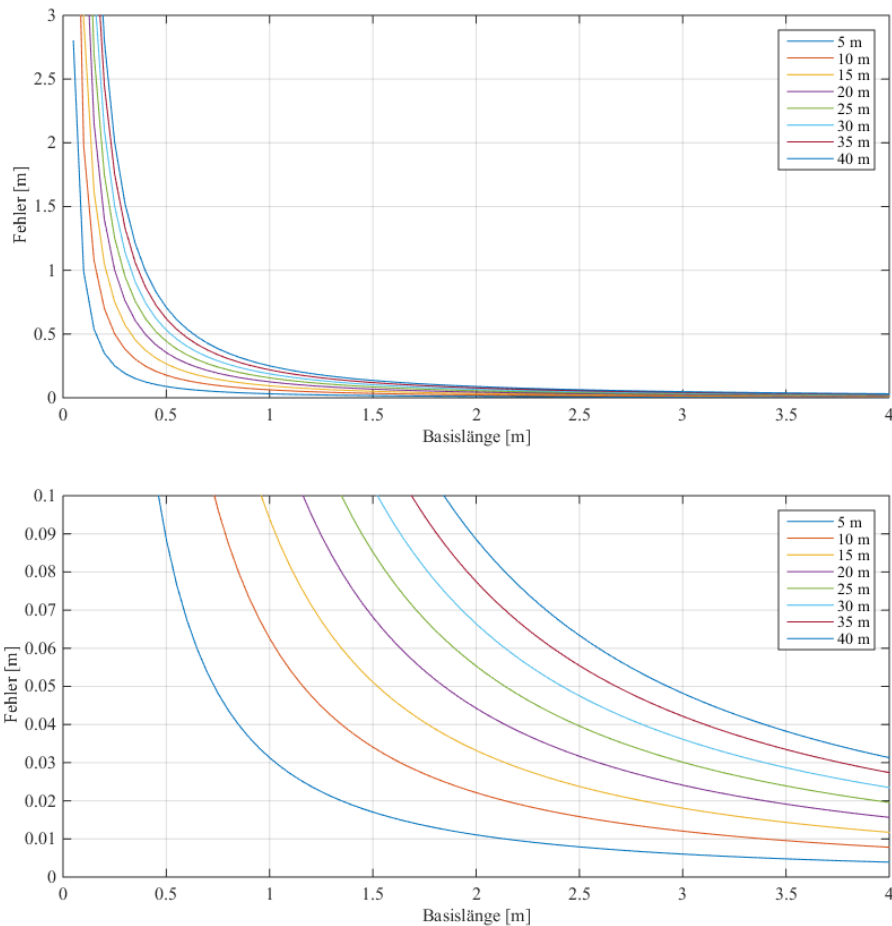


Abbildung 4-8 Auswirkung des Skalierungsfehlers auf die Objektdistanz. Oben: Fehler für ausgewählte Objektdistanzen von 5-40 m; Unten: Vergrößerter Ausschnitt bis zu einem maximalen Fehler von 0.1 m

Zu sehen ist, dass der Fehler eines Objektpunktes mit zunehmender Basislänge während der Initialisierung sinkt. Zudem ist der Fehler für Objekte in großen Distanzen erwartungsgemäß höher. Eine Übersicht wird in Tabelle 4-1 dargestellt, woraus hervorgeht, dass mit einer Basislänge von 4 m eine Genauigkeit von ca. 3.1 cm in einer Entfernung von 40 m zum Objekt erreicht wird. Es wird jedoch empfohlen zunächst über eine Basislänge von 2 m zu initialisieren, um sehr dichte Objekte direkt identifizieren zu können. Ist dies nicht der Fall, wird zur Steigerung der Genauigkeit die Szene mit einer größeren Basislänge erneut skaliert.

Tabelle 4-1 Maßstabsabhängiger Fehler [cm] bei unterschiedlichen Basislängen und festen Objektdistanzen

Objektdistanz [m]		5	10	15	20	25	30	35	40
Fehler [cm]	Basislänge 1 m	3.1	6.3	9.4	12.5	15.7	18.8	21.9	25.0
	Basislänge 2 m	1.1	2.2	3.3	4.4	5.5	6.6	7.7	8.9
	Basislänge 3 m	0.6	1.2	1.8	2.4	3.0	3.6	4.2	4.8
	Basislänge 4 m	0.4	0.8	1.2	1.6	2.0	2.3	2.7	3.1

Die Distanz D beschreibt hier den Abstand zwischen Kamera und Objekt. Gleiches gilt jedoch für beliebige Abstände im Punktraum, bspw. der Hindernishöhe, die sich aus Objekthöhe und Höhe des Fahruntergrunds ableitet. Ausgehend von einer für leichte Nutzfahrzeuge meist kritischen Hindernishöhe von 2 – 3 m, ergeben sich somit Fehler in der Höhenberechnung von unter 1 cm.

4.5 Fahrschlauchprädiktion

Die Reduzierung von fehlerhaften Warnungen ist ein Schlüsselement innerhalb dieser Arbeit. Neben der Rekonstruktion von möglichst genauen 3D-Punktwolken und der Identifikation möglicher Hindernisse, ist die Prädiktion der eigenen Fahrzeugbewegung von hoher Bedeutung. Die Warnstrategie sieht dabei vor, dass ausschließlich auf Hindernisse innerhalb des Fahrschlauchs gewarnt wird. Der Fahrschlauch beschreibt in diesem Zusammenhang die prädizierte Pose des Fahrzeugs im WKS mit Berücksichtigung dessen Höhe h und Breite b . Er umfasst die folgenden Kenngrößen.

$$X, Y, Z, \Phi, \Theta, \Psi, h, b$$

h und b sind dabei Fahrzeugkennwerte, die konstruktionsbedingt festgelegt sind. Die Fahrzeughöhe wird maßgeblich von den verschiedenen Fahrwerks- und Rad-Reifen-Kombinationen beeinflusst. Änderungen der Fahrzeughöhe aufgrund des Luftdrucks oder der Beladung werden im Zusammenhang dieser Arbeit nicht berücksichtigt.

Die Parameter Z , Φ und Θ werden durch den Verlauf des Fahruntergrundes bestimmt, dessen Extraktion und Modellierung in Kapitel 7.2 erläutert wird. Innerhalb dieses Kapitels wird auf die Prädiktion der Parameter X , Y und Ψ eingegangen, welche durch verschiedene Ansätze ermittelt werden können.

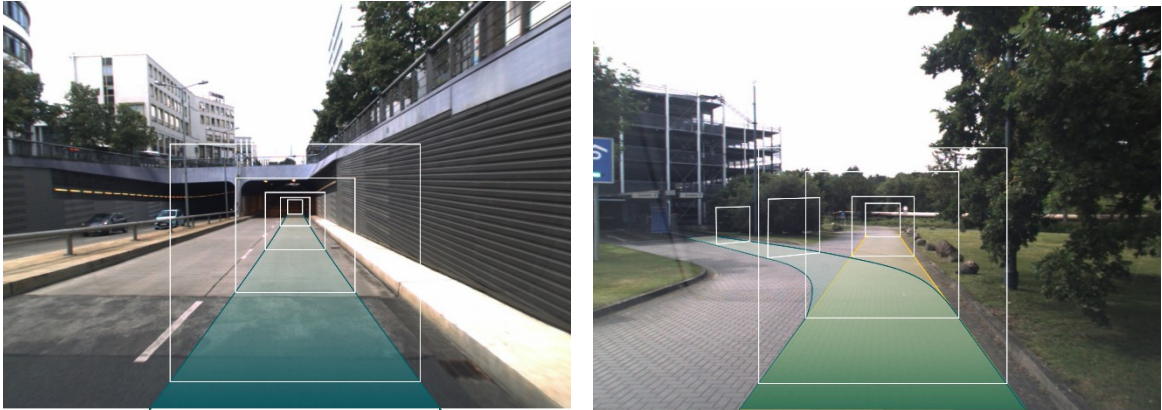


Abbildung 4-9 Fahrerschlauchprädiktion. Links: Prädiktion bei geradlinigen Verlauf durch Extrapolation der momentanen Eigenbewegung möglich; Rechts: Falsche Prädiktion durch Extrapolation (gelb) - Verwendung von Bildinformationen nötig (grün)

Es werden drei Ansätze zur Schätzung des Fahrerschlauchs vorgestellt, wobei diese in bildbasierte und eigenbewegungsbasierte Prädiktion unterteilt werden.

4.5.1 Bildbasierte Prädiktion

In diesem Unterkapitel sollen zwei bildbasierte Methoden zur Fahrerschlauchprädiktion vorgestellt werden. Im ersten Ansatz werden der Horizont und ein horizontaler Fluchtpunkt geschätzt, um die grobe Fahrtrichtung zu ermitteln.

Ein weiterer Ansatz ist die Erkennung von Fahrbahnmarkierungen und Fahrstreifen. Dies wird bereits in modernen Fahrerassistenzsystemen wie dem Spurverlassenswarner oder dem Spurhalteassistenten eingesetzt. Beide Systeme warnen den Fahrer bei Unachtsamkeit und mindern somit das Unfallrisiko im Straßenverkehr. Die meisten Systeme setzen dabei auf den Einsatz einer vergleichsweise kostengünstigen Frontkamera und Methoden der Bildverarbeitung.

Fluchtpunktschätzung

Die Existenz eines Fluchtpunkts ist eine charakteristische Eigenschaft der perspektivischen Abbildung. Geraden werden bei der projektiven Abbildung in die Bildebene auch als Geraden dargestellt. Parallele Geraden konvergieren dabei als Geradenbüschel im Fluchtpunkt v . Werden mehrere Fluchtpunkte von horizontalen Linien betrachtet, so liegen diese auf einer Geraden, dem Horizont [67].

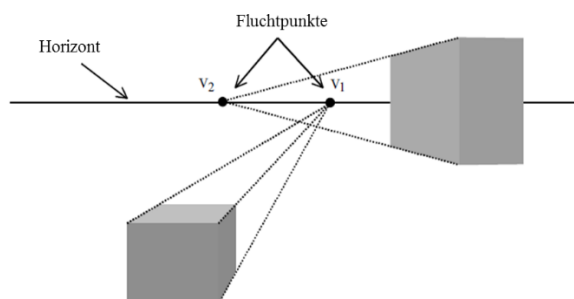


Abbildung 4-10 Horizont und Fluchtpunkte nach [68]

Fluchtpunkte werden im Automotive-Bereich vor allem bei der Umfeldwahrnehmung von automatischen Fahrzeugen verwendet. Die Navigation in unstrukturierten Gebieten, wie

Schotterstraßen, ist aufgrund fehlender Fahrbahnmarkierungen komplexer als in strukturierten städtischen Szenarien. Eine Fluchtpunkt- und Horizontschätzung kann in diesen Fällen bei der Identifikation der groben Fahrtrichtung dienen.

Typisch für die Fluchtpunkterkennung ist die Identifizierung von richtungsorientierten Elementen, die in Gruppen zusammengefasst und anschließend zur Fluchtpunktbestimmung genutzt werden. Diese Elemente können aus der Bildebene extrahierte Linien [69, 70, 71], orientierte Merkmale (z.B. durch Eckendetektor) [72, 73], oder sogar ganze Strukturen [74] sein. Die Verfolgung der Fluchtpunkte über mehrere Bilder wird mittels erweitertem Kalman Filter [71] oder Partikelfilter [75] realisiert.

Im vorliegenden Fall wird der Algorithmus [76] angepasst, um die Fluchtpunkte und den Horizont aus dem Bild zu extrahieren. Das Verfahren stützt sich zunächst auf eine Kantenerkennung mittels *Line Segment Detector* (LSD) [77]. Darüber hinaus werden über mehrere Bilder verfolgte Punktmerkmale genutzt, da diese Aufschluss über die Fahrzeugbewegung geben. Einzelne kurze Liniensegmente werden anschließend zu langen Linien verbunden und richtungsgebunden gruppiert. Die Fluchtpunktkandidaten werden berechnet und mittels Manhattan-Welt-Annahme drei rechtwinklig aufeinander stehende Fluchtpunkte gesucht. Zwei Fluchtpunkte befinden sich dabei auf Höhe der Bodenebene und definieren den Horizont. In Abbildung 4-11 werden in verschiedenen Sequenzen der in Fahrtrichtung zeigende Fluchtpunkt (in X-Richtung - rot) und der Horizont bestimmt. Der vertikale (in Z-Richtung - grün) und horizontale Fluchtpunkt (in Y-Richtung - blau) werden dabei nicht gezeigt.

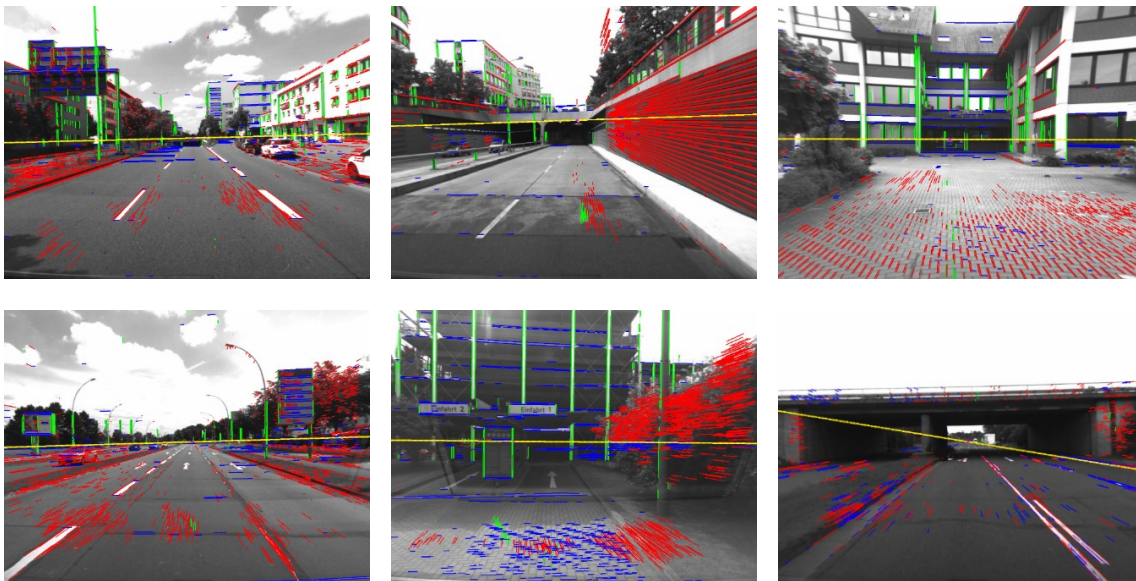


Abbildung 4-11 Fluchtpunkt- und Horizontdetektion

In allen Situationen hat der Algorithmus den in Fahrtrichtung liegenden Fluchtpunkt bestimmt, sodass auch in Kurvenfahrten (s. Unten Mitte) eine Prädiktion der Fahrzeugbewegung mittels Fluchtpunkt durchgeführt werden kann. Die Erkennung aller Fluchtpunkte und die Berechnung des Horizonts schlägt jedoch in Szenarien mit unzureichender Struktur (vgl. Unten rechts) fehl. Es werden mittels LSD zu wenig Linien für jede Richtung erkannt, was eine dahingehende Gruppierung erschwert. Weiterhin wird festgehalten, dass die Bestimmung des Horizonts nicht zwangsweise der Höhe des Fahruntergrunds entspricht (Oben Mitte, Oben rechts, Unten Mitte). Die in Kapitel 1.1.2 vorgestellten fluchtpunkt-basierten Verfahren zur Höhenberechnung von Hindernissen sind im vorliegenden Kontext demnach ungeeignet.

Freiflächen- und Fahrbahnschätzung

Im Rahmen automatischer Fahrfunktionen sind insbesondere die Kenntnis des eigenen Fahrspurverlaufs für die Querregelung notwendig. Bei komplexeren Manövern werden zudem die Nachbarfahrspuren beobachtet sowie statische und dynamische Objekte detektiert. Dies ermöglicht die Bestimmung des gesamten befahrbaren Bereichs, welcher auch Freifläche genannt wird. Sie kann unabhängig von vordefinierten Fahrspuren auch in unstrukturierten Geländen detektiert und bspw. durch bauliche Einschränkungen wie Bordsteine und stehende Fahrzeuge oder durch fahrdynamische Begrenzungen, wie maximaler Lenkeinschlag, verkleinert werden. Die Freiflächenerkennung wird im kamerabasierten Ansatz häufig durch eine Segmentierung direkt im Bild erreicht. Diese stützen sich zumeist auf neuronale Netze [78, 79].

Eine weitere Variante, die im Rahmen dieser Arbeit umgesetzt wurde, ist die Schätzung der Freifläche aus der 3D-Punktwolke der rekonstruierten Szene. Diese trennt Objekte vom befahrbaren Bereich. Der Ansatz bedingt eine möglichst dichte und genaue Punktwolke, damit auch kleinere Höhenunterschiede zwischen Fahrbahn und Bordsteinkanten erfasst werden können (s. Kapitel 7.2). Die Schätzung des Fahrbahnverlaufs über Fahrstreifenmarkierungen ist ein klassisches Thema des Computersehens im Automobilbereich und wurde im Laufe der Jahre durch unterschiedliche Verfahren umgesetzt. Prinzipiell lassen sich alle bildbasierten Verfahren in drei Abschnitte gliedern. Dabei werden zunächst die Fahrspurmarkierungen aus dem Bild extrahiert, der Fahrbahnverlauf approximiert und anschließend über mehrere Bilder stabilisiert. Eine Detektion nach [80] wird in Abbildung 4-12 gezeigt.

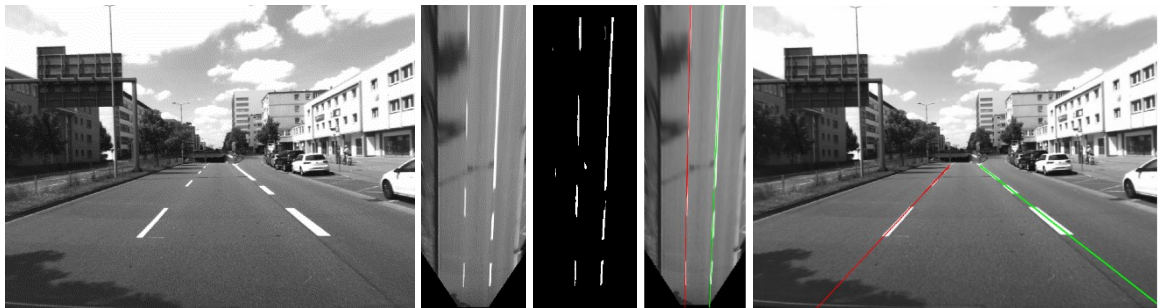


Abbildung 4-12 Detektion der Fahrbahnmarkierungen. v.l.n.r: Ausgangsbild; Transformation in die Vogelperspektive, Binärbild, Approximation der Linienverläufe; Rücktransformation der Linien in das Ausgangsbild

Zur Vereinfachung der Extraktion der Markierungen und Verhinderung von Fehldetektionen werden die Bilder vorbearbeitet. Neben einer Glättung zur Rauschunterdrückung werden zudem Region of Interests bestimmt. Diese können entweder als feste Pixelkoordinaten oder dynamisch, bspw. sich dem Horizont anpassend, implementiert werden. Darüber hinaus können auch Farbinformationen von RGB-Bildern ausgewertet werden. So können vergleichsweise einfach die Markierungen vom Asphalt unterschieden und auch Sondermarkierungen in Baustellen berücksichtigt werden [81]. In vielen Fällen erfolgt die anschließende Erstellung eines Kantenbild, in dem die Linien mittels Liniendetektoren oder morphologischen Operationen bestimmt werden [82]. In [83] wird hingegen die Extraktion von „Ridges“ forciert, welche die Mittellinie der Fahrbahnmarkierung darstellen. Die einzelnen Linien werden anschließend nach ihrer Richtung gruppiert, um bspw. rechte und linke Markierungen unterscheiden zu können. Zur Modellierung des Fahrstreifenverlaufs eignen sich verschiedene Ansätze. Während die Linienendpunkte genutzt werden können, um Geraden, Splines oder beliebige Polynome im Bild zu approximieren, versuchen einige Verfahren Verläufe zu schätzen, welche sich an Modellen orientieren, die im realen Straßenbau genutzt werden.



Abbildung 4-13 Detektion der Fahrbahnmarkierungen. Links: Fehlgeschlagene Modellierung; Rechts: Erfolgreiche Erkennung und Modellierung der Fahrbahnränder

In [84] wird die Annahme einer ebenen Grundfläche in Verbindung mit Kurven konstanter Krümmung getroffen. In [83] wird vorgeschlagen die konstante Krümmung durch einen längenabhängigen Term der Krümmung zu erweitern. Dies entspricht dem im Straßen- und Schienenbau verwendeten Klothoidenmodell [85]. Die Verfolgung und Stabilisierung des Modells wird innerhalb der Bildfolge bspw. über einen Kalman Filter oder RANSAC realisiert.

Die Liniendetektion und Extraktion der Fahrbahnmarkierungen ist stark von Umgebungseinflüssen geprägt. Überdeckungen durch vorausfahrende Fahrzeuge, abgenutzte Markierungen, verschiedene Beleuchtungssituation oder Regen verlangen einen robusten Algorithmus. Wie in Abbildung 4-13 zu sehen, werden auch in schwierigen Situationen Markierungen erkannt. Im linken Bild kommt es jedoch zu einer Fehldetektion, da sich die Fahrbahnbegrenzungen nur bedingt von der Fahrbahn unterscheiden und daraufhin die Modellierung fehlschlägt. Im rechten Bild werden die Bordsteine erkannt und modelliert, allerdings befindet sich der Carport außerhalb dieser Markierungen und wird daraufhin in der Hindernisdetektion missachtet. Um auch Situationen in der Fahrschlauchprädiktion zu berücksichtigen, die sich bildbasiert als problematisch erweisen, wird die eigenbewegungs-basierte Prädiktion eingeführt.

4.5.2 Eigenbewegungs-basierte Prädiktion

Dieser Ansatz beschreibt eine Prädiktion des Fahrschlauchs über die Extrapolation der momentanen Fahrzeugbewegung. Dabei werden die in Abschnitt 4.2 hergeleiteten Bewegungsgleichungen in die nahe Zukunft prädiziert. Um auch bei langsamen Geschwindigkeiten eine stabile Fahrschlauchprädiktion zu erhalten, wird vermehrt auf die Informationen des Lenkwinkels ϕ zurückgegriffen, welcher über den Lenkradwinkel und den Übersetzungsfaktor berechnet wird. Es wird angenommen, dass das Fahrzeug nach Durchfahren einer Kurve wieder zum Zustand $\phi = 0$ zurückkehrt. Demzufolge wird in jedem Zeitschritt der Lenkwinkel über die folgende Funktion verkleinert prädiziert.

$$\phi(t + 1) = \phi(t)e^{-a \cdot \Delta t} \quad \text{Gl. (4-17)}$$

Über den Faktor a kann das Abnehmen des Lenkwinkels parametrisiert werden. Mittels prädizierten Lenkwinkels können über Gl. (4-1) für jeden Zeitschritt ein Kurvenradius der Hinterräder bestimmt werden, welcher anschließend als Fahrschlauch zusammengefasst wird.

Fahrzeugeigenbewegung

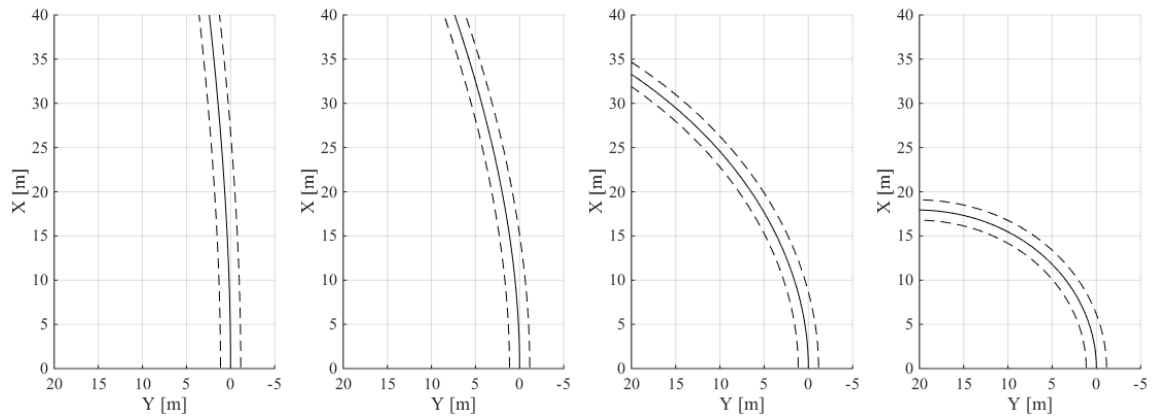


Abbildung 4-14 Eigenbewegungs-basierte Fahrschlauchprädiktion in der X/Y -Ebene. Die durchgezogene Linie beschreibt die Fahrzeugmitte. Die gestrichelten Linien stellen die Fahrzeugbreite dar. v.l.n.r.: Lenkradwinkel 10° , 30° , 90° , 180°

Abbildung 4-14 zeigt beispielhaft den prädizierten Fahrschlauch für vier ausgewählte Lenkradwinkel. Dabei ist darauf hinzuweisen, dass diese ausschließlich unter der Vernachlässigung von Reibwertverlusten gelten. Soll eine Prädiktion auch in hochdynamischen Situationen erfolgen, müssen weitere Informationen über die Umgebung und Fahrdynamik hinzugezogen werden.

5 Merkmalsbasierte Rekonstruktionsverfahren

5.1 Grundlagen

Bildmerkmale können in zwei Arten unterschieden werden. Globale Merkmale beschreiben das Bild gesamtlich, wobei deren Eigenschaften anhand aller Pixel bestimmt wird. Ein Bezug zur räumlichen Verteilung lässt dieser Ansatz nicht zu. Lokale Merkmale hingegen, wie bspw. einzelne Punkte, Linien, Regionen, Texturen oder Farben im Bild, sind entsprechend stark ausgeprägt und unterscheiden sich somit von ihrer benachbarten Umgebung. Lokale Merkmalsdetektoren sind oft für spezielle Anwendungen entwickelt und optimiert, wobei die Rahmenbedingungen der Anwendungen die Auswahl des Merkmalsdetektors beeinflussen.

5.1.1 Anforderungen an lokale Merkmale

Im vorliegenden Fall soll eine echtzeitfähige monokamerabasierte Funktion die Durchfahrtshöhe von vertikalen Engstellen bestimmen. Die Funktion erschafft ein dreidimensionales Abbild des Fahrzeugumfelds, welches sich durch eine große Anzahl von 3D-Elementen (Punkte/Linien) mit hoher Genauigkeit auszeichnet. Die Bildmerkmale müssen dabei perzeptiv bedeutsam, in verschiedenen Bildern extrahierbar sowie robust gegenüber Störeinflüssen sein.

Aufgrund der charakteristischen Bewegung des Fahrzeugs in Blickrichtung der Kamera und der hohen Bildabtastrate von 30 Bildern pro Sekunde, können Faktoren wie Robustheit gegenüber Rotation oder affinen Transformationen in der Auswahl des Merkmalsdetektors vernachlässigt werden. Die Invarianz der Merkmale hinsichtlich Skalierung und Standpunktänderungen, die vor allem in Szenen mit großer räumlicher Ausdehnung vorteilhaft ist, kann hier ebenfalls ignoriert werden. Als wesentlicher Vorteil des hier vorgestellten Verfahrens wird die geringe Bewegung zwischen den Bildern hervorgehoben, die es erlaubt einen Objektpunkt sehr häufig innerhalb der Sequenz zu detektieren und somit die Redundanz zu erhöhen. Da der rekonstruierte 3D-Raum zur metrischen Bestimmung der Durchfahrtshöhe genutzt wird, ist die Lokalisierungsgenauigkeit des Bildmerkmals entscheidend. Eine genaue Lokalisierung des Merkmals in mehreren Bildern kann die Anzahl an Inliern erhöhen und darüber hinaus die bildbasierte Berechnung der Fahrzeugbewegung verbessern. Die Identifizierung und Klassifizierung eines 3D-Objekts als Hindernis und die Bestimmung seiner Kritikalität ist vor allem von der Anzahl der von örtlich zusammenliegenden 3D-Punkten bzw. -Linien abhängig. Die Anzahl wird primär über die Eigenschaften des Detektors und die Wiederholgenauigkeit, d.h. die Anzahl korrekt zugeordneter Merkmale über mehrere Bilder, bestimmt. Zur verlässlichen Bestimmung der Bildorientierung ist eine Verteilung der Merkmale im gesamten Bild notwendig.

Im Sinne der Vielfalt der Szenarien im Straßenverkehr ist eine Beleuchtungsinvarianz ebenfalls von Bedeutung, da auch während kurzfristigen Änderung der Umgebungsbeleuchtung, bspw. bei der Einfahrt in ein Parkhaus oder durch Schattenwurf, ausreichend Merkmale gefunden und rekonstruiert werden müssen. Die Darstellung eines Fahrerassistenzsystems erfordert die Implementierung von echtzeitfähigen Algorithmen, sodass die Rechendauer bei der Bestimmung der Merkmale möglichst gering sein sollte. Zusammengefasst können die Anforderungen aufgelistet werden:

1. Hohe Anzahl an Merkmalen
2. Hohe Wiederholgenauigkeit
3. Hohe Lokalisierungsgenauigkeit
4. Formatfüllende Bildabdeckung
5. Invariant gegenüber Beleuchtungsänderungen
6. Hohe Berechnungsgeschwindigkeit

5.1.2 Verfahrensschritte

Das Verfahren zur Erstellung einer 3D-Punktwolke, aus einer Serie an Bildern wird als *Structure from Motion* bezeichnet und kann in vier Phasen unterteilt werden.

Detektion

Unter Detektion wird in diesem Zusammenhang das Finden von lokalen Bildmerkmalen beschrieben. Dabei handelt es sich um ausgewählte markante Bildpositionen, die über bestimmte Merkmalsdetektoren extrahiert werden. In den gängigen Verfahren werden dazu meist Punktmerkmale verwendet. Es können jedoch auch Linienmerkmale innerhalb der Pipeline genutzt werden.

Deskription

Während der Deskription gilt es die extrahierte Bildposition und dessen direktes Umfeld möglichst detailliert in einem unveränderlichen Eigenschaftsvektor beschreiben zu können, um das Wiederfinden des Merkmals in einem weiteren Bild zu ermöglichen.

Korrespondenzsuche

Diese Phase beschreibt Methoden die Deskriptoren unterschiedlicher Merkmale zu vergleichen, sodass einzelne Merkmale über die Bildfolge verfolgt werden können. Im Verlauf der Arbeit wird eine zweite Methode der Korrespondenzsuche vorgestellt, die ohne Merkmalsdeskription auskommt, das Tracking.

Rekonstruktion

Dieser Schritt beschreibt die Erstellung von 3D-Punkten auf Basis korrespondierender Bildmerkmale und orientiert sich an den Grundlagen aus Abschnitt 2.3.4.

5.1.3 Aufnahmekonfiguration

Aufgrund der primär in Blickrichtung der Kamera verlaufenden Fahrzeugbewegung entstehen hohe Überlappungen innerhalb der Bildfolge. Dies bietet die Möglichkeit ein hohes Maß an Redundanz während der Aufnahme der Straßenszene zu schaffen. Bereits detektierte Bildinhalte können über einen langen Zeitraum verfolgt, auf Neue kann schnell reagiert werden. Als problematisch kann jedoch die sich daraus ergebende schleifende Schnittgeometrie gesehen werden. Ausgehend von korrespondierenden Bildmerkmalen wird mithilfe der Epipolarometrie die Position des 3D-Punkts und die Kameraposen bestimmt. Da sich die Sichtstrahlen nur im idealen fehlerfreien Fall schneiden, kommt es je nach Genauigkeit in der Merkmalsdetektion bzw. -verfolgung zu Fehlern in der 3D-Punktberechnung. Der Einfluss der Störungen nimmt mit kleineren Schnittwinkeln zu. Schematisch kann dies Abbildung 5-1 entnommen werden.

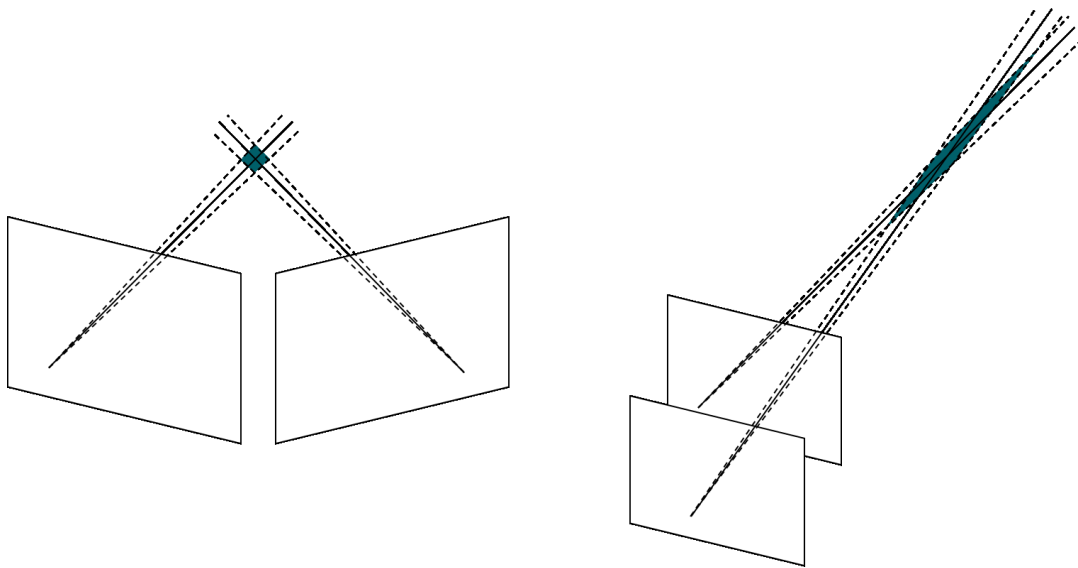


Abbildung 5-1 Ungenauigkeit aufgrund der Schnittgeometrie. Links: Kleiner markierter Bereich durch großen Schnittwinkel; Rechts: Aufnahmeconfiguration ähnlich einer Bildfolge im Fahrzeug. Große Ungenauigkeit vor allem in Kamerablickrichtung

Zur Evaluierung des Positionsfehlers des 3D-Punktes wird ein Modell aufgezeigt, welches den folgenden Annahmen unterliegt.

1. Es erfolgt die Rekonstruktion des 3D-Punktes aus zwei Bildern.
2. Die erste Kameraposition wird zu null gesetzt, die zweite bewegt sich entlang der Fahrtrichtung mit $[B, 0, 0]$. Rotationen werden vernachlässigt.
3. Die Kamera ist kalibriert, sodass systematische Fehler vernachlässigt werden können.
4. f_x und f_y werden mit f angenommen.
5. x_0 und y_0 sind null.

Unter diesen Annahmen werden die Kollinearitätsgleichungen aus Gl. (3-1) und Gl. (3-2) genutzt, um das funktionale Modell im Kamerakoordinatensystem aufzustellen.

$$x' = f \left(\frac{X}{Z} \right) \quad \text{Gl. (5-1)}$$

$$y' = f \left(\frac{Y}{Z} \right) \quad \text{Gl. (5-2)}$$

$$x'' = f \left(\frac{X}{Z - B} \right) \quad \text{Gl. (5-3)}$$

$$y'' = f \left(\frac{Y}{Z - B} \right) \quad \text{Gl. (5-4)}$$

Damit werden die geometrischen Zusammenhänge zwischen den Beobachtungen (den Bildkoordinaten) und den Unbekannten (den 3D-Punktpositionen) beschrieben.

Zunächst wird die Designmatrix \mathbf{A} der Form

$$\mathbf{A} = \begin{bmatrix} \frac{\partial x'}{\partial X} & \frac{\partial x'}{\partial Y} & \frac{\partial x'}{\partial Z} \\ \frac{\partial y'}{\partial X} & \frac{\partial y'}{\partial Y} & \frac{\partial y'}{\partial Z} \\ \frac{\partial x''}{\partial X} & \frac{\partial x''}{\partial Y} & \frac{\partial x''}{\partial Z} \\ \frac{\partial y''}{\partial X} & \frac{\partial y''}{\partial Y} & \frac{\partial y''}{\partial Z} \end{bmatrix} = \begin{bmatrix} \frac{f}{Z} & 0 & -\frac{fX}{Z^2} \\ 0 & \frac{f}{Z} & -\frac{fY}{Z^2} \\ \frac{f}{Z} & 0 & -\frac{fX}{(Z-B)^2} \\ 0 & \frac{f}{Z} & -\frac{fY}{(Z-B)^2} \end{bmatrix} \quad \text{Gl. (5-5)}$$

aufgestellt, welche die Ableitungen des funktionalen Modells nach den Unbekannten X , Y und Z enthält. Nun wird die Normalgleichungsmatrix \mathbf{N} mit

$$\mathbf{N} = \mathbf{A}^T \mathbf{P} \mathbf{A}, \quad \text{Gl. (5-6)}$$

aufgestellt. Diese enthält die Gewichtsmatrix \mathbf{P} , die unter Annahme gleichgenauer unkorrelierter Beobachtungen der Standardabweichung σ_0 zur Einheitsmatrix wird. Die Varianz-Kovarianz-Matrix der Unbekannten $\mathbf{K}_{\mathbf{XX}}$ berechnet sich zu

$$\mathbf{K}_{\mathbf{XX}} = \sigma_0^2 \mathbf{N}^{-1} = \sigma_0^2 (\mathbf{A}^T \mathbf{A})^{-1}. \quad \text{Gl. (5-7)}$$

Der Positionsfehler eines Objektpunktes vergrößert sich bei hohen Hindernisentfernungen, geringer Differenz zwischen Hindernis- und Einbauhöhe der Kamera sowie kleinen Bewegungsvektoren. Die Auswirkung auf den Fehler in den Komponenten Entfernung und Höhe der 3D-Punkte wird in Abbildung 5-2 dargestellt. Dabei wurde die Objektdistanz mit 40 m und der Bewegungsvektor in Fahrtrichtung mit 0.05 m angenommen. Ein wahrnehmbarer Entfernungsfehler wird innerhalb eines Radius' von 10 Pixel um den Bildhauptpunkt hervorgerufen. Im Fall von $X = 0$ und $Y = 0$ ergibt sich im Bildhauptpunkt ein theoretisch unendlich großer Fehler. Unter der Annahme, dass eine Genauigkeit von 0.05 m in der Objekthöhe gefordert wird und die Standardabweichung der Bildkoordinaten 1 Pixel beträgt, ergibt sich ein Entfernungsfehler von 4.56 m in der Nähe des Bildhauptpunktes. Der maximale Höhenfehler beträgt 0.001 m und wird am äußersten Rand des Bildes gemessen. Festzuhalten ist, dass sich der Fehler basierend auf der Aufnahmekonfiguration vor allem in Fahrtrichtung auswirkt, sodass dies bei der Unterteilung des Punktraumes (Abschnitt 7.1.2) in der Festlegung der Zellgröße berücksichtigt werden kann. Allerdings ist der Einfluss nur sehr nahe des Bildhauptpunktes wirklich bemerkbar und kann durch Forderung eines minimalen Winkels zur Punkttriangulation reduziert werden. Der Fehler der Höhenkomponente ist vernachlässigbar klein. Im hier vorgestellten Beispiel wird gezielt nur der Fehlereinfluss durch kleine Bewegungsvektoren und die Ungenauigkeit der Bildkoordinatenbestimmung eingegangen. Im realen Einsatz des Verfahrens kann jedoch von weiteren Einflussquellen ausgegangen werden. So sind die über die Punktkorrespondenzen bestimmten Kameraorientierungen aufgrund des Minimierungsverfahrens stets fehleranfällig. Zudem ist die inkrementelle Rekonstruktion von Bildsequenzen driftbehaftet, wodurch Fehler zwischen verfolgten Bildkoordinaten zweier Bilder zunächst gering sein können, sich jedoch im Laufe des Trackings entscheidend vergrößern. Darüber hinaus können auch fehlerhafte Tracks, Reflexionen oder bewegte Objekte die Genauigkeit der 3D-Punktbestimmung beeinflussen. In vorliegender Arbeit wird ein maximaler Entfernungsfehler von 2 m bei Erstdetektion des Hindernisses akzeptiert. So werden bereits früh einzelne 3D-Punkt Cluster als Hindernisse erkannt. Bei der weiteren Zufahrt sollte der Distanzfehler jedoch weiter abnehmen. Die Analyse dazu wird in Abschnitt 8.4.2.1 geführt.

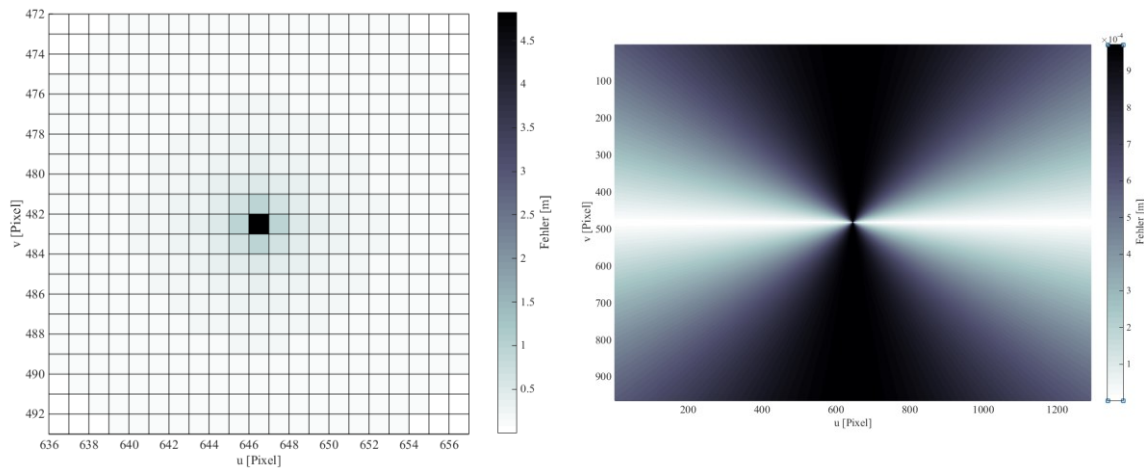


Abbildung 5-2 Aus den Ungenauigkeiten der Bildkoordinaten abgeleiteter Fehler (Hindernisentfernung 40 m) - Darstellung im Pixelkoordinatensystem. Links: Fehler in Fahrtrichtung mit einem Maximum von 4.56 m nahe des Bildhauptpunkts; Rechts: Fehler in vertikaler Richtung mit einem Maximum von 0.001 m am Bildrand.

5.2 Punkte

5.2.1 Detektion

Merkmalsdetektoren können klassisch in Eck- und Blobdetektoren unterteilt werden. Erstere fokussieren das Finden von Ecken, welche Schnittpunkte zweier oder mehrerer Linien darstellen. Sie ermöglichen eine schnelle Berechnung und weisen eine hohe Lokalisierungsgenauigkeit auf. Die über Eckendetektoren gefundenen Merkmale unterscheiden sich untereinander jedoch nur gering, was die Korrespondenzsuche in weiteren Bilder erschwert. Blobdetektoren zielen auf das Finden von Merkmalen auf verschiedenen Skalenebenen ab und sind deshalb häufig langsamer. Einzelne Merkmale unterscheiden sich jedoch stärker, wodurch auch Korrespondenzen über größere Blickwinkel- oder Skalenänderungen hergestellt werden können.

Je nach Anwendungsgebiet zeichnen sich Stärken und Schwächen eines Detektors ab. Im Rahmen dieser Arbeit sollen die gewählten Detektoren auf ihre Eignung zur Rekonstruktion des Fahrzeugumfelds bewertet werden. Ziel ist es eine möglichst dichte Punktwolke auf dem Hindernis und dem Fahruntergrund zu erhalten, die eine hohe Punktgenauigkeit aufweist. Zur Bewertung werden demnach eine Vielzahl an Testsequenzen ausgewählt, die unterschiedlichste Umgebungseinflüsse und Fahrzeugdynamiken vereinen. Die untersuchten Merkmalsdetektoren sollen im Folgenden vorgestellt werden.

SIFT

SIFT (Scale-Invariant Feature Transform) zählt zu einem der bekanntesten Merkmalsdetektoren und wurde von Lowe beschrieben [86, 87]. Zunächst wird das Bild auf verschiedene Größen skaliert und ein Gauß Filter angewendet, um unwichtige Details und Rauschen zu eliminieren. Anschließend werden über den Difference of Gaussian (DoG) Operator kontrastreiche Regionen im Bild hervorgehoben und jedes Pixels auf den unterschiedlichen Skalenebenen mit seiner Nachbarschaft auf die Existenz von lokalen Extrema untersucht. Durch eine gradientenbasierte Filterung können schwach texturierte Extrema entfernt und möglichst aussagekräftige Merkmale beibehalten werden. Im letzten Schritt werden über die verschiedenen Skalenebenen die lokalen Orientierungshistogramme aus der Nachbarschaft der Merkmale ermittelt. Dabei fließen die Gradientenstärke und -richtung in die finale Orientierung des Merkmals ein.

Das Verfahren ist skalen- und rotationsinvariant, jedoch aufgrund der Komplexität sehr rechenintensiv, wodurch es sich in Zusammenarbeit mit dem gleichnamigen Deskriptor vor allem für Offline Anwendungsfälle im Bereich der Objektdetektion [88] eignet. Die Güte der Merkmale ist sehr hoch, sodass hohe Genauigkeiten bei einer 3D-Rekonstruktion erreicht werden können.

SURF

SURF (Speeded-Up Robust Features) wurde von [89] vorgestellt und bedient sich vieler Grundelemente von SIFT. Anstelle von Skalenebenen arbeitet SURF jedoch auf dem Integralbild

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(x, y). \quad \text{Gl. (5-8)}$$

Zur Merkmalsdetektion wird das Fast Hessian Verfahren genutzt, wobei der DoG Operator durch einen Boxfilter ersetzt wird. Eine Skaleninvarianz wird vereinfacht über die Anpassung der Größe des Boxfilters erreicht. Dadurch werden unterschiedlich große Bereiche des Bildes abgedeckt und verschiedene Skalierungen simuliert.

Die Erkennung der Merkmale erfolgt über die Bestimmung der Determinante der Hesse Matrix jedes Pixels. Die Hesse Matrix enthält die zweiten partiellen Ableitungen der Bildintensitäten in x - und y -Richtung und kann durch die Verwendung von Haar Wavelets effizient approximiert werden. SIFT und SURF stellen jeweils Verfahren dar, welche auf eine Invarianz gegenüber Skalierung und Rotation abzielen. Der hohe Informationsgehalt und die Qualität der Merkmale vereinfacht die anschließende Korrespondenzsuche in der Bildfolge. Die Verfahren sind jedoch sehr rechenintensiv, wodurch sie sich nicht zwangsweise für einen Echtzeitbetrieb im Fahrzeug eignen.

Die nachfolgend vorgestellten Eckendetektoren arbeiten ausschließlich auf dem Originalbild. Dadurch wird eine schnellere Detektion möglich, die im vorliegenden Anwendungsfall entscheidend ist.

Harris

Der Harris Eckendetektor wurde von [90] entwickelt und basiert auf den Erkenntnissen von [91], der sich mit der Detektion von Ecken in Bildern beschäftigt. Ecken weisen dabei Intensitätsänderungen in x - und y -Richtung auf.

Der Algorithmus berechnet die Eigenwerte der lokalen Strukturmatrix, welche sämtlichen Ableitungen an einer Stelle (x, y) beschreibt.

Die bestimmten Eigenwerte λ_1 und λ_2 werden in der Auswertungsfunktion verknüpft, wobei Harris diese folglich definiert

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2. \quad \text{Gl. (5-9)}$$

Es werden damit folgende Differenzierungen erreicht:

- $|R|$ ist klein, wenn λ_1 und λ_2 klein \rightarrow flache Region
- $R < 0$, wenn $\lambda_1 \gg \lambda_2$ oder $\lambda_1 \ll \lambda_2 \rightarrow$ Kante
- R ist groß, wenn λ_1 und λ_2 groß und $\lambda_1 \approx \lambda_2 \rightarrow$ Ecke

Die Empfindlichkeit des Detektors wird dabei über die Variable k kontrolliert. Der Harris Operator ist einer der einfachsten und am häufigsten genutzten Eckendetektoren in der Bildverarbeitung.

Shi-Tomasi

Shi und Tomasi haben in ihrer Veröffentlichung [92] einen Merkmalsdetektor vorgestellt, der grundlegend auf dem Harris Eckendetektor basiert. Die Detektoren bestimmen die stärksten Eckpunkte, wobei sich deren Funktionen für diese Bewertung unterscheiden. Ähnlich zum Harris Eckendetektor werden die Eigenwerte der Strukturmatrix untersucht, um eine Ecke zu bestimmen. Es wird zudem eine Glättung vorgenommen, da die Strukturmatrix empfindlich gegenüber Rauschen ist. Shi und Tomasi führen folgende Auswertungsfunktion ein:

$$R = \min(\lambda_1, \lambda_2) \quad \text{Gl. (5-10)}$$

Es werden Ecken detektiert, sofern der Schwellwert λ_{min} überschritten wird. Der Ansatz von Shi-Tomasi zeichnet sich durch eine kürzere Berechnungszeit und stabilere Eckendetektion aus, wodurch ein anschließendes Tracking begünstigt wird. Er findet vorwiegend Anwendung im KLT Tracker.

FAST

Der FAST Detektor (Features from Accelerated Segment Test) wurde in [93] vorgestellt und zielt auf eine möglichst schnelle Merkmalsdetektion ab. Dabei wird es vermieden partielle Ableitungen zu bilden, sondern ein direkter Vergleich eines Pixels mit seinen Nachbarn durchgeführt.

Die Nachbarn umfassen dabei einen Kreis von 16 definierten Pixel um den Kandidaten, wobei alle Pixel innerhalb dieses Kreises als Patch bezeichnet werden. Wenn n benachbarte Pixel einen höheren Intensitätswert (+ Schwellwert) oder niedrigeren Intensitätswert (- Schwellwert) aufweisen, wird der Kandidat als Ecke definiert. Die Auswahl des Kandidaten, die Bestimmung des Schwellwertes und der 16 Nachbarpixel werden dabei mittels eines maschinellen Lernverfahrens durchgeführt. Um zu verhindern, dass benachbarte Pixel als separate Merkmale erkannt werden, wird ein Verfahren namens Non-maximum suppression angewendet. Dabei wird eine Auswertungsfunktion berechnet, die die Summe der absoluten Differenzen eines Kandidaten mit seinen 16 Nachbarn bestimmt. Liegen zwei Merkmalskandidaten sehr dicht beieinander, so wird der Kandidat verworfen, welcher einen geringeren Wert in der Auswertungsfunktion aufweist. Der Eckendetektor eignet sich aufgrund seiner hohen Geschwindigkeit vor allem für Echtzeitvideoanwendungen. Bereits in den von [93] durchgeführten Tests zeigt sich eine 18-fache Beschleunigung der Detektion gegenüber Harris und 45-fache Beschleunigung gegenüber SIFT. FAST arbeitet ausschließlich auf dem Originalbild, ist somit nicht skaleninvariant und aufgrund seines Designs anfällig für Rauschen. Da der Algorithmus schwellwertbasiert arbeitet, muss dieser auf ggf. auf wechselnde Umgebungssituationen angepasst werden.

Eine Bewertung der hier beschriebenen Detektoren wird in der Literatur bereits anhand von anwendungsorientierten Kriterien durchgeführt. Beispielsweise werden in [94] verschiedene Blob- und Eckendetektoren im Rahmen der Berechnung einer visuellen Odometrie bewertet. Wie im vorliegenden Projekt ist eine Monokamera dabei hinter der Windschutzscheibe angebracht durch eine primär vorwärts gerichtete Bewegung gekennzeichnet. Der KLT Tracker (Kanade-Lucas-Tomasi - s. Abschnitt 5.2.3) und SIFT liefern in Bezug auf den Translations- und Rotationsfehler in x und y den geringsten Fehler zur GPS-gestützten Referenzbewegung. Der kleinste Rückprojektionsfehler wird mit SIFT erreicht. Berechnungszeiten werden bei der Untersuchung jedoch vernachlässigt. In [95] wird ebenfalls Fokus auf Merkmalsdetektoren gelegt, die sich zur visuellen Odometrieberechnung eignen. Es werden insbesondere Robustheit, Genauigkeit und die Berechnungsgeschwindigkeit untersucht, wobei ein Algorithmus vorgeschlagen wird, der je nach Umgebungstextur den Merkmalsdetektor wechselt. Als optimal wurde dabei die Kombination aus Harris und FAST bestimmt. In [96] werden Detektoren bezüglich ihrer Eignung für visuellen SLAM

Merkmalsbasierte Rekonstruktionsverfahren

untersucht, wobei der Harris Eckendetektor die höchste Trackingrate besitzt. Anschließend folgen SIFT und SURF, welche in den bewerteten Sequenzen ähnliche Ergebnisse erzielen.

Einen Überblick über die gewählten Detektoren und ihre Eigenschaften wird in Tabelle 5-1 gegeben.

Tabelle 5-1 Merkmalsdetektoren und ihre Eigenschaften (Bewertung nach [97])

	Blob-detektor	Ecken-detektor	Rotations-invarianz	Skalen-invarianz	Affine Invarianz	Wiederhol-genauigkeit	Lokalisierungs-genauigkeit	Robustheit	Effizienz
SIFT	×		×	×	×	+++	++	+++	+
SURF	×		×	×	×	+++	++	++	++
Harris		×				+++	+++	++	++
Shi-Tomasi		×				+++	+++	++	++
FAST		×				++	++	++	++++

Primärer Fokus bei der Beurteilung der Merkmalsdetektoren liegt in diesem Abschnitt auf der Anzahl der Merkmale und der Berechnungsgeschwindigkeit. Die Merkmalsdetektion wurde dabei auf einer i7 CPU mit 2.8 GHz berechnet.

Tabelle 5-2 Vergleich der Detektoren

	SIFT	SURF	Harris	Shi-Tomasi	FAST
Merkmale pro Bild [n]	1668	3253	1231	2411	2534
Detektionszeit pro Bild [ms]	455.4	193.9	42.8	44.0	3.0
Detektionszeit pro Merkmal [ms]	0.273	0.059	0.034	0.018	0.0012

Unter Berücksichtigung von 72 repräsentativen Sequenzen wurden die durchschnittlichen Werte für die Merkmalsanzahl und Berechnungszeit, dargestellt in Tabelle 5-2, ermittelt. Die Sequenzen beinhalten eine möglichst große Variation an Umgebungsbeleuchtungen und -bedingungen. Dabei werden in verschiedenen Geschwindigkeiten geradlinige und kurvige Streckenabschnitte aufgezeichnet, welche über ebenen, steigenden und fallenden Fahruntergrund verfügen. Zudem sind große Hindernisse wie Brückenunterführungen und Tunnel sowie kleine Hindernisse wie freihängende Schilder und Warnbaken enthalten, die zur letztendlichen Höhenberechnung herangezogen werden.

Innerhalb der Detektion wird die maximale Merkmalsanzahl auf 3500 begrenzt. SIFT und SURF benötigen aufgrund der Detektion auf unterschiedlichen Skalenebenen am längsten, wobei mit SURF die meisten Merkmale im Bild erkannt werden. Die ähnlichen Laufzeiten von Harris und Shi-Tomasi sind durch ihren nahezu identischen Aufbau zu erklären. Sie unterscheiden sich ausschließlich in der Auswertungsfunktion. FAST ist aufgrund der laufzeitoptimierten Struktur der schnellste Detektor. Für eine genaue Berechnung der Kamerapose und der damit einhergehenden Fahrzeugbewegung wird eine möglichst hohe Verteilung der detektierten Merkmalspunkte im Bild gefordert. Diese wird im Rahmen der Analyse in Abschnitt 5.2.3.2 untersucht.

5.2.2 Deskription

Für die Merkmalsbeschreibung werden sogenannte Deskriptoren benutzt. Der Deskriptor, oder auch Eigenschaftsvektor, muss Informationen erhalten, die es ermöglichen, dasselbe Merkmal in einem weiteren Bild zu finden und anschließend zu paaren. Nach [98] kann zwischen gradientenbasierten und binärbasierten Deskriptoren unterschieden werden. Zu den gradientenbasierten oder auch floating point Deskriptoren gehören SIFT und SURF. BRIEF und FREAK sind binärbasierte Deskriptoren.

Inhalte eines solchen Eigenschaftsvektors sind z.B. die lokale Orientierung der Merkmalsumgebung oder ein Gradientenhistogramm. Bei Detektoren, die auf Skalenebenen aufsetzen, können zudem Informationen zu Pyramidenstufe, Texturmerkmale, Farbinformationen oder auch Frequenzbänder gespeichert werden. Ein Vektor kann wenige (bspw. 32) oder auch sehr viele Dimensionen besitzen (z.B. 128). Eine größere Dimension ermöglicht es mehr Hinweise auf die Besonderheit des Merkmals und der angrenzenden Region zu speichern, sodass sich Merkmale besser voneinander unterscheiden lassen. Allerdings geht damit eine Steigerung der Rechenzeit einher.

Auf die in dieser Arbeit ausgewählten Deskriptoren soll im Folgenden näher eingegangen werden.

SIFT

Der SIFT Deskriptor enthält für charakteristische Skalierungen lokale Bildinformationen in der Umgebung des Merkmals. Zunächst werden Lage und Richtung der Gradienten jedes Pixels in einer 16x16 Umgebung des Kandidaten untersucht, über eine Gauß-Verteilung gewichtet und anschließend in einer 4x4 Matrix zusammengefasst. Anschließend werden die 16 Unterregionen, die jeweils 8 Orientierungen erhalten, in einem 128 Einträge fassenden Vektor geschrieben, der den Merkmalsdeskriptor darstellt. Eine anschließende Normierung verhindert eine starke Abhängigkeit gegenüber der Beleuchtung. SIFT ist rotationsinvariant, sowie unabhängig von Skalierung und Beleuchtung. Der Vektor ist jedoch sehr aufwendig konstruiert, wodurch die Deskription und Korrespondenzsuche vergleichsweise langsam ist.

SURF

Beim SURF Deskriptor wird eine $20s \cdot 20s$ Nachbarschaft um den Kandidaten untersucht und in 4×4 Regionen unterteilt, wobei s die Skalenebene beschreibt.

Es wird ein Vektor

$$v = (\sum\{d_x\}, \sum\{d_y\}, \sum\{|d_x|\}, \sum\{|d_y|\}) \quad \text{Gl. (5-11)}$$

erstellt, welcher 64 Dimensionen besitzt. SURF ist invariant gegenüber Unschärfe und Rotation, jedoch nicht gegenüber affinen Transformationen sowie Beleuchtungsänderungen.

FREAK

FREAK (Fast Retina Keypoints) wurde in [99] vorgestellt und ist vom Aufbau der menschlichen Netzhaut inspiriert. Es wird ein netzhautähnlicher kreisrunder Bereich um das Merkmal angelegt und in mehreren symmetrischen Teilregionen die Intensitätswerte miteinander verglichen. Die Teilregionen werden mittels Gaußkernels geglättet. Der Deskriptor enthält die Intensitätsvergleiche von 512 vordefinierten Paaren. Der Deskriptor ist robust gegenüber aller Art von Deformationen, jedoch langsamer als andere Binärdeskriptoren [100].

BRIEF

BRIEF (Binary Robust Independent Elementary Features) wird in [101] vorgestellt. Es werden wie bei FREAK Intensitätswertpaare (128, 256 oder 512) miteinander verglichen, die sich immer an denselben Stellen in der Region um den Kandidaten befinden. Allerdings wird statt einer kreisförmigen Umgebung eine quadratische Region um den Kandidaten untersucht. Ist der Intensitätswert des ersten Pixels kleiner als des zweiten, so wird in den Deskriptorvektor eine 0 eingetragen; im umgekehrten Fall eine 1. Es wird ebenfalls wie bei FREAK ein Gauß Filter in der Umgebung des Kandidaten angewendet, um Einflüsse durch Rauschen zu minimieren. Das Verfahren arbeitet sehr schnell, ist invariant gegenüber Beleuchtungsänderungen, jedoch nicht invariant gegenüber Rotation und Skalierung.

In der Literatur, bspw. in [98], werden verschiedene gradientenbasierte und binärbasierte Merkmalsdeskriptoren bezüglich ihrer Rechenzeit und deren Inlieraten bewertet. Unter anderem werden dabei die Deskriptoren von SIFT, SURF, BRIEF und FREAK bei zufällig ausgewählten Bildern des *Oxford Buildings Datasets*¹ untersucht. Die Deskriptoren unterscheiden sich nur gering in deren Inlierate bei skalierten und rotierten Bildern sowie bei Standpunktvariationen. Der Einfluss ihrer Komplexität spiegelt sich eher in der Rechenzeit wider. FREAK ist dabei 12-mal schneller als SURF und 40-mal schneller als SIFT. Beim BRIEF Deskriptor sind die Faktoren 10 bzw. 30. Die Deskriptoren wurden dabei mithilfe des euklidischen Abstands (floating point basiert) und Hammingdistanz (binärbasiert) verglichen. In [102] werden Deskriptoren im Rahmen des Visual SLAM Ansatzes verglichen. Diese werden jedoch nicht nach Inlieren bewertet, sondern die Veränderung des Deskriptors eines Merkmals über aufeinanderfolgende Bilder beurteilt. Als gemeinsamer Merkmalsdetektor wird Harris genutzt, wobei ausschließlich Sequenzen untersucht wurden, welche in geschlossenen Räumen aufgenommen wurden. SURF zeichnet sich dabei in der gewählten Umgebung vor allem durch einen hohen Unterscheidungsgrad zwischen verschiedenen Merkmalen aus, wodurch die Korrespondenzsuche über mehrere Bilder vereinfacht wird. Die Deskriptoren werden bei der kamerabasierten Erzeugung von 3D-Punktwolken im Automotive-Bereich häufig durch geeignete Tracker (s. Abschnitt 5.2.3.2) ersetzt. Beim Einsatz von Trackingverfahren werden in einem Bild zunächst Merkmale mittels Detektor bestimmt. Anschließend werden unter Annahme von kleinen Bewegungen die Merkmalspositionen in den nächsten Bildern geschätzt. Kleine Positionsfehler akkumulieren sich somit über mehrere Bilder. Unter der Verwendung von Deskriptoren müssen in jedem Bild Merkmale detektiert werden, was jedoch eine unabhängige Korrespondenzsuche ermöglicht.

Innerhalb dieser Arbeit werden die Deskriptoren zunächst allein anhand ihrer Geschwindigkeit bewertet. Die Mittelung der Ergebnisse aus den 72 Testsequenzen wird in Tabelle 5-3 aufgezeigt. Eine gezielte Untersuchung der Detektor-Deskriptor-Paarung wird im nachfolgenden Abschnitt behandelt.

¹ <http://www.robots.ox.ac.uk/~vgg/data/oxbuildings/>

Tabelle 5-3 Berechnungszeiten der Deskriptoren [ms]

	SIFT	SURF	FREAK	BRIEF
Deskriptionszeit pro Bild [ms]	785.1	68.0	51.6	12.3
Deskriptionszeit pro Merkmal [ms]	0.3537	0.0306	0.0232	0.0055

SIFT weist aufgrund der Komplexität die größte Berechnungsdauer auf. Der verkleinerte Deskriptor von SURF ermöglicht im vorliegenden Fall eine Beschleunigung der Deskription vom Faktor 12. FREAK ist etwas schneller, wohingegen BRIEF die mit Abstand geringste Berechnungszeit aufweist. Für echtzeitfähige Anwendungen eignet sich bei bestehender Hardware allein der BRIEF-Deskriptor. Es wird hierbei jedoch auf die GPU-Varianten von SIFT und SURF verwiesen, die es ermöglichen die Berechnungszeiten für die Detektion und Deskription um bis zu 85 % zu reduzieren [103]. Dieser Wert wird jedoch erst bei einer möglichst großen Anzahl an Merkmalen erreicht. Bei nur geringem Berechnungsaufwand für die Grafikkarte steigt der prozentuale Anteil der Berechnungszeit für das Lesen und Schreiben der Informationen im Speicher.

5.2.3 Korrespondenzsuche

Im folgenden Abschnitt soll kurz auf unterschiedlichen Verfahren einer Korrespondenzsuche eingegangen werden. In der Arbeit werden dafür zwei grundlegende Varianten auf deren Eignung für die Integration im Fahrzeug untersucht.

5.2.3.1 Deskriptormatching

Die im letzten Abschnitt vorgestellten Deskriptoren bzw. Eigenschaftsvektoren werden im Rahmen der Bildauswertung für jedes Merkmal erstellt und können mittels Matchingverfahren auf ihre Ähnlichkeit überprüft werden. Grundsätzlich wird beim Matching die Distanz der Deskriptoren der einzelnen Merkmale miteinander verglichen. Die Merkmale, deren Deskriptoren den kleinsten Unterschied zueinander aufweisen, werden gepaart. Der Vergleich ist somit die Suche nach dem nächsten Nachbarn, die je nach Deskriptorgröße im n -dimensionalen Raum durchgeführt wird.

Die Distanzmessungen können z.B. bei gradientenbasierten Deskriptoren mittels euklidischem Abstand oder Mahalanobisdistanz durchgeführt werden. In [87] wird ein *Ratio Test* vorgeschlagen, bei dem das Verhältnis zwischen besten und zweitbestem Match ermittelt wird. Ein hoher Wert deutet auf eine große Ähnlichkeit mit einer weiteren Punktkorrespondenz, sodass keine Eindeutigkeit nachgewiesen und der Match verworfen wird. Binärbasierte Deskriptoren können mittels Hamming-Distanz verglichen werden, welche die Unterschiede in einem String-Vergleich bewertet.

Das Deskriptormatching ist ein sehr rechenaufwendiger Schritt, der durch verschiedene Techniken, wie Hashing, Clustering, Kd-Bäume oder GPU-Implementierung beschleunigt werden kann. Durch *Guided Matching* kann zudem der geometrische Suchraum für Punktkorrespondenzen im Bild verkleinert werden. Dabei können Informationen der Radodometrie und die daraus resultierende Kamerabewegung innerhalb der Bildfolge hinzugezogen werden.

Die in den vorigen Abschnitten vorgestellten Detektoren und Deskriptoren werden nun gepaart und über die Gesamtanzahl der Sequenzen die gemeinsamen Korrespondenzen von jeweils zwei aufeinander folgenden Bildern unter Einbeziehung der geometrisch nächsten Nachbarn bestimmt. Mittels RANSAC wird die essentielle Matrix eines Bildpaares bestimmt und alle Korrespondenzen, die einen Rückprojektionsfehler größer 0.8 Pixel aufweisen eliminiert. Die daraus ermittelten

Tabelle 5-4 Gemeinsame (oben) und geometrische (unten) Korrespondenzen zwischen einem Bildpaar [%]

		Detektor				
		SIFT	SURF	Harris	Shi-Tomasi	FAST
Deskriptor	SIFT	73.05	78.23	75.69	49.51	86.72
		59.10	42.23	52.82	31.95	45.35
	SURF	48.88	75.72	35.53	19.49	44.79
		38.06	53.64	22.41	10.55	30.48
	FREAK	65.18	72.62	66.83	50.53	64.04
		53.02	48.84	49.26	35.25	45.55
	BRIEF	78.82	75.41	81.89	66.75	83.10
		53.76	48.24	52.07	39.07	47.19

geometrischen Korrespondenzen (Inlier) können anschließend in die Rekonstruktion übergeben werden. Tabelle 5-4 zeigt eine Übersicht über die gemeinsamen und geometrischen Korrespondenzen der Detektor-Deskriptor-Paarungen. Aus der Tabelle wird ersichtlich welche Detektor-Deskriptor-Paarung sich im Rahmen einer Automotive-Anwendung mit in Fahrtrichtung gerichteter Kamera eignet. Es wird gezeigt, dass keine Kombination der Erwartung einer hohen Inlierrate gerecht wird. Obwohl sich die Bildinhalte infolge der geringen Bewegungen innerhalb der Sequenz nur wenig unterscheiden, wird im Vergleich eine maximale Inlierrate von ca. 60 % erreicht. Die Blobdetektoren liefern dabei die besten Ergebnisse mit ihren eigenen Deskriptoren (SIFT-SIFT, SURF-SURF). FREAK und BRIEF liefern für alle Detektoren ähnliche Ergebnisse, wobei BRIEF aufgrund der bis zu 4-fach kürzeren Laufzeit bevorzugt wird. Unter den Eckendetektoren werden die höchsten Matchingraten mit Harris erreicht.

Eine hohe Inlierrate ermöglicht es eine große Anzahl an 3D-Punkten zu erzeugen. Die Qualität des photogrammetrischen Modells wird jedoch entscheidend über die Genauigkeit der Merkmale bestimmt. Als Indikator kann hierbei der gemittelte Rückprojektionsfehler dienen, welcher für alle Punktkorrespondenzen der Testsequenzen in Tabelle 5-5 gezeigt wird. Hier wird jedoch darauf verwiesen, dass die Bewegung von Merkmalen nahe des Kamerazentrums sehr gering ist. Aufgrund der schleifenden Schnittgeometrie ergeben sich ungenaue 3D-Punkte in Kamerablickrichtung, die dennoch einen geringen Rückprojektionsfehler aufweisen. Diese Merkmale werden mithilfe eines Mindestschnittwinkels eliminiert.

Tabelle 5-5 Mittelwerte der Rückprojektionsfehler [Pixel]

		Detektor				
		SIFT	SURF	Harris	Shi-Tomasi	FAST
Deskriptor	SIFT	0.397	0.452	0.473	0.414	0.541
	SURF	0.408	0.440	0.404	0.410	0.473
	FREAK	0.436	0.442	0.462	0.450	0.498
	BRIEF	0.478	0.450	0.491	0.467	0.544

Neben der höchsten Inlierrate hat SIFT-SIFT auch den geringsten Rückprojektionsfehler. Die Eckendetektoren besitzen den kleinsten Rückprojektionsfehler unter Verwendung des SURF-Deskriptors. FAST ist der vergleichsweise ungenaueste Detektor und sollte nur bei Anwendungen genutzt werden, die begrenzte Rechenressourcen zur Verfügung stellen.

Als weiteres Qualitätskriterium wird die Tracklänge untersucht. Dies ist ein Maß, welches die mittlere Bildanzahl angibt, über die ein Merkmal verfolgt werden kann. Eine hohe Tracklänge ermöglicht es einen Objektpunkt aus möglichst vielen Kameraposen zu betrachten und ihn somit mehrmals zu triangulieren bzw. zu optimieren. Wie in Tabelle 5-6 zu sehen, werden die mit SIFT extrahierten Merkmale am längsten verfolgt. Nur sehr kurze Tracks werden unter der Verwendung des FAST-Detektors erreicht. Dies ist auf die geringe Stabilität gegenüber Beleuchtungsänderungen zurückzuführen.

Tabelle 5-6 Mittelwerte der Tracklänge [Bilder]

		Detektor				
		SIFT	SURF	Harris	Shi-Tomasi	FAST
Deskriptor	SIFT	18.81	10.62	10.69	9.97	6.94
	SURF	19.13	11.61	12.04	10.72	6.40
	FREAK	18.54	10.85	10.78	9.88	6.96
	BRIEF	18.65	10.49	10.03	9.61	7.00

Die Untersuchung hat ergeben, dass die besten Ergebnisse hinsichtlich Inlierrate, Qualität und Stabilität mit Blobdetektoren erreicht werden. Die Nutzung der Skalenpyramiden bietet hier den entscheidenden Vorteil eine hohe Wiederholgenauigkeit zu erreichen. Den geringsten Rückprojektionsfehler und die größte Tracklänge wird bei den Eckendetektoren durch Harris erreicht. Allerdings geht dies mit einer geringen Merkmalsanzahl einher. Shi-Tomasi hat die geringsten Inlierraten, erzeugt jedoch qualitativ ähnliche Ergebnisse wie Harris. FAST ist mit Abstand der schnellste Detektor. Allerdings ist die Wiederholgenauigkeit und Qualität sehr gering. Kritisch bei allen Detektoren ist die Parametrierung des Schwellwertes zu sehen, welche über die Akzeptanz eines potentiellen Merkmals entscheidet. Die Vielzahl der Situationen und Umgebungsbedingungen im Straßenverkehr machen eine eindeutige Wahl des Parameters schwer, sodass ggf. Schwellwerte adaptiv angepasst werden müssen. Eine Möglichkeit ist dabei das stetige Herabsetzen des Schwellwertes während der Detektion im Bild, bis eine Mindestanzahl an Merkmalen bestimmt wurde.

Ein weiterer Ansatz, der es erlaubt die Merkmale über mehrere Bilder zu verfolgen, ist das Tracking.

5.2.3.2 Merkmalstracking

Die Korrespondenzsuche und -verfolgung (Tracking) kann auch auf Basis des optischen Flusses durchgeführt werden. Darunter versteht man in der Bildverarbeitung die Schätzung der Bewegungsrichtung und Geschwindigkeit eines Bildpunktes innerhalb einer Bildfolge. Das aufgebaute Flussfeld zwischen zwei Bildern enthält Vektoren, deren Bildpunkte denselben Objektpunkt abbildet. Der Vektor

$$v(x, t) = \begin{bmatrix} u(x, t) \\ v(x, t) \end{bmatrix} \quad \text{Gl. (5-12)}$$

kann für jeden Bildpunkt x unter der Annahme gleichbleibender Intensität des Objektpunktes auf der Bildebene $I(x, y, t)$ in der Bildsequenz bestimmt werden.

Somit folgt aus

$$\frac{dI}{dt} = \frac{\delta I}{\delta t} + \frac{\delta I}{\delta x} \frac{dx}{dt} + \frac{\delta I}{\delta y} \frac{dy}{dt} \quad \text{Gl. (5-13)}$$

die Bestimmungsgleichung der Geschwindigkeiten zu

$$\frac{\delta I}{\delta t} + \frac{\delta I}{\delta x} \frac{dx}{dt} + \frac{\delta I}{\delta y} \frac{dy}{dt} = 0. \quad \text{Gl. (5-14)}$$

Die Richtung des Gradienten und die zeitliche Dimension beschränken hierbei die Lösung, welche jedoch nicht eindeutig ist (s. *Aperturproblem*). Zudem sind in homogenen Bereichen die Ableitungen der Intensität I nach x und y null, wodurch kein optischer Fluss berechnet werden kann. Eine eindeutige Bestimmung des optischen Flusses ermöglicht unter anderem der Algorithmus nach Lucas und Kanade, welcher in der vorliegenden Arbeit implementiert wurde.

Methode nach Lucas und Kanade

In [104] wird davon ausgegangen, dass ein konstanter Fluss in der $m \times m$ -Region um einen zentralen Bildpunkt herrscht, wobei $1 \dots n = m^2$ gilt. Damit lässt sich das überbestimmte lineare Gleichungssystem

$$A\mathbf{v}_z + \mathbf{b} = 0 \quad \text{Gl. (5-15)}$$

oder

$$\begin{bmatrix} I_{x_1} & I_{y_1} \\ I_{x_2} & I_{y_2} \\ \vdots & \vdots \\ I_{x_n} & I_{y_n} \end{bmatrix} \mathbf{v}_z = \begin{bmatrix} -I_{t_1} \\ -I_{t_2} \\ \vdots \\ -I_{t_n} \end{bmatrix} \quad \text{Gl. (5-16)}$$

aufstellen, welches sich nach der Methode der kleinsten Quadrate berechnen lässt. I_x , I_y und I_t beschreiben dabei die Ableitungen der Intensität nach dem Ort bzw. der Zeit. Die Erweiterung der Methode führt eine Gewichtsfunktion $W(\mathbf{x})$ ein, um dem zentralen Bildpunkt mehr Einfluss zu verleihen. In homogenen Bildregionen wie der Fahrbahn kann aufgrund des fehlenden Intensitätsgradienten dennoch keine Lösung gefunden werden (s. Abbildung 5-3).

Weitere Verfahren werden in [105, 106] und vorgestellt. In ersterem wird eine Glattheitsfunktion eingeführt, welche voraussetzt, dass sich das Flussfeld um einen zentralen Bildpunkt nur geringfügig ändert. Der zweite Ansatz ist nicht von den Gradienten der Intensität abhängig, um ein Flussfeld zu berechnen. Er versucht die Lösung für ein approximiertes Polynom der Intensitätswerte in einer Bildregion zu finden. Beide Methoden sind in der Lage für jeden Bildpunkt den optischen Fluss zu berechnen, jedoch ist die Lösung in homogenen Regionen null.

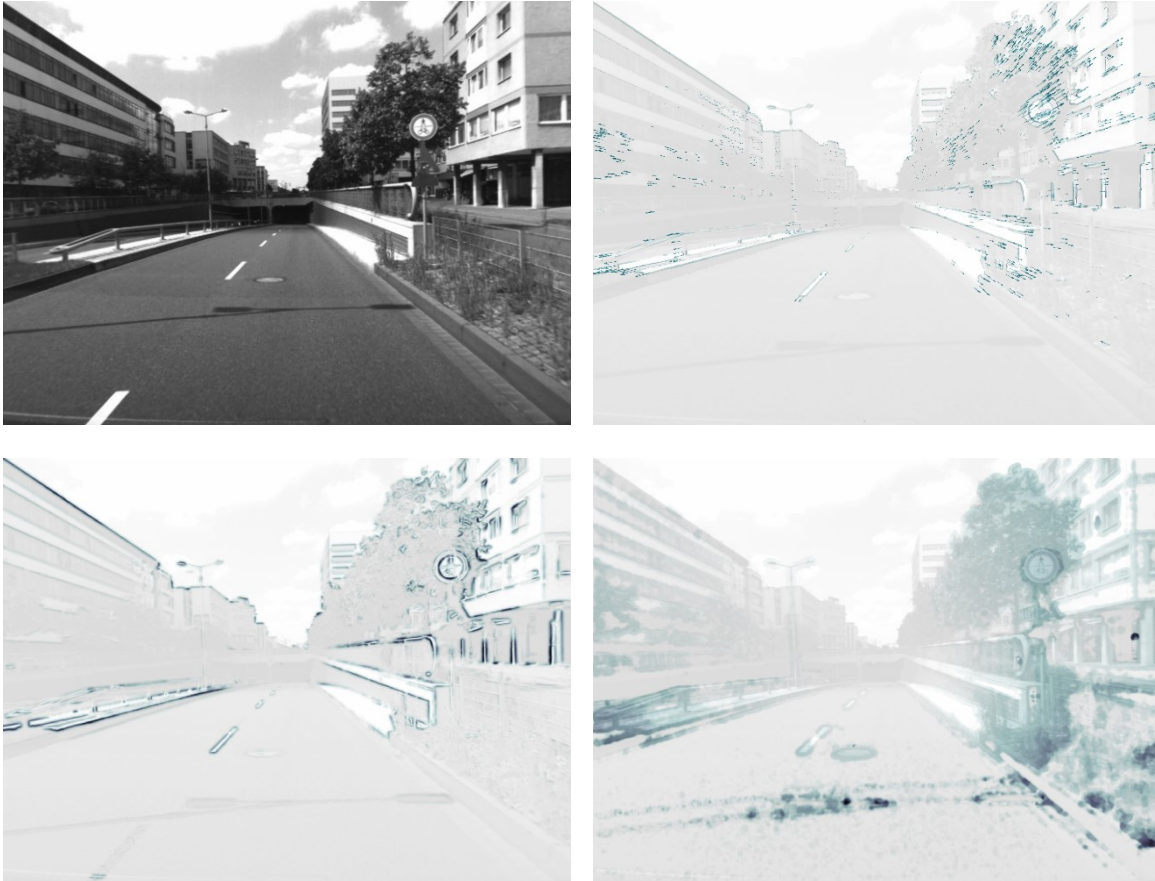


Abbildung 5-3 Optischer Fluss. Oben: Ausgangsequenz (links) und Methode nach Lucas und Kanade (rechts); Unten: Methode nach Horn und Schunck (links) sowie nach Farneback (rechts)

Das vor allem in der Videobearbeitung eingesetzte Verfahren nach Lucas und Kanade soll auch hier genutzt werden, um Merkmalskorrespondenzen über die Bildfolge aufzustellen. Dazu werden mit den in 5.2.1 beschriebenen Detektoren zunächst Merkmale in einem Bild detektiert, per KLT Tracker die Bewegung des Merkmals zum nächsten Bild ermittelt und die relative Orientierung der Bilder berechnet. Merkmale, die keine gemeinsame Korrespondenz aufweisen, werden gelöscht. In jedem Bild werden zudem in geringer Anzahl weitere Merkmale detektiert, sodass auf neue Bildinhalte reagiert werden kann. Die Schnittwinkel zwischen Kamerazentren und 3D-Punkten werden vor allem durch die Basisbreite beeinflusst. Die Erzeugung langer Tracks ist demzufolge ein Mittel, mit dem die Qualität des 3D-Punkts gesteigert werden kann.

Tabelle 5-7 Vergleich der Merkmalsdetektoren unter Nutzung des KLT Trackers

	Detektor				
	SIFT	SURF	Harris	Shi-Tomasi	FAST
Merkmale pro Bild [n]	1797	2012	1670	1956	1988
Gemeinsame Korrespondenzen [%]	98.59	98.17	98.67	98.47	98.65
Geometrische Korrespondenzen [%]	93.29	92.58	93.63	92.60	93.45
Rückprojektionsfehler [Pixel]	0.283	0.287	0.282	0.299	0.306
Tracklänge [Bilder]	22.22	20.33	22.02	20.01	23.19
Detektion/Tracking pro Bild [ms]	88.84	53.38	34.14	36.69	30.17
Detektion/Tracking pro Merkmal [ms]	0.049	0.027	0.020	0.019	0.015

In Tabelle 5-7 werden die Ergebnisse mit dem Lucas-Kanade Tracker dargestellt. Im Vergleich zur Korrespondenzsuche über Deskriptoren werden hierbei im Mittel weniger Merkmale pro Bild genutzt, da nicht in jedem Bild versucht wird die Maximalanzahl von 3500 Merkmalen zu detektieren. Allerdings wird die geringere Anzahl effizienter verfolgt. Dabei werden bis zu 2-3 mal höhere Inlierraten und Tracklängen erreicht und der mittlere Rückprojektionsfehler im Schnitt um 0.16 Pixel verringert. Wesentlicher Vorteil ist die Verkürzung der Berechnungszeit aufgrund der reduzierten Merkmalsuche, des Wegfallen des Deskriptionsschritts und der effizienteren Korrespondenzsuche. Beispielhaft sind in Abbildung 5-4 die verfolgten Merkmalspunkte der Detektoren dargestellt. Die Sequenz ist durch ein Hindernis gekennzeichnet, welches sich vom Hintergrund abhebt und eine klare Unterkante aufweist. Obwohl der Fahruntergrund über eine ausreichende Strukturierung verfügt, werden mit Harris und FAST nur wenige Merkmale auf der Straße verfolgt. Shi-Tomasi eignet sich hier als bester Detektor, da damit nicht nur die Markierungen, sondern auch Teile der Teernähte und der Schattenwurf unter dem Hindernis erkannt wird. Mit den Blobdetektoren SIFT und SURF werden viele Merkmale in vermeintlich uninteressanten Regionen gefunden. Beispielsweise werden mit SURF viele Merkmale auf dem Hindernis gefunden, jedoch nicht an dessen Unterkante bzw. den Deckenleuchten im Tunnel. Dafür eignen sich eher die Eckendetektoren, wobei auch diese im direkten Bereich der Unterkante aufgrund ihrer Auswertungsfunktion Merkmale verwerfen. Die mit Shi-Tomasi extrahierten Merkmale befinden sich in allen Bildbereichen, auf dem Hindernis sowie der Fahrbahn. Somit eignet sich dieser Detektor in Verbindung mit dem Lucas-Kanade Tracker für den vorliegenden Anwendungsfall am besten. Alternativ können auch verschiedene Detektoren im Verlauf der Sequenz kombiniert werden, um eine möglichst hohe Stabilität bei variierenden Umgebungsbedingungen zu gewährleisten.

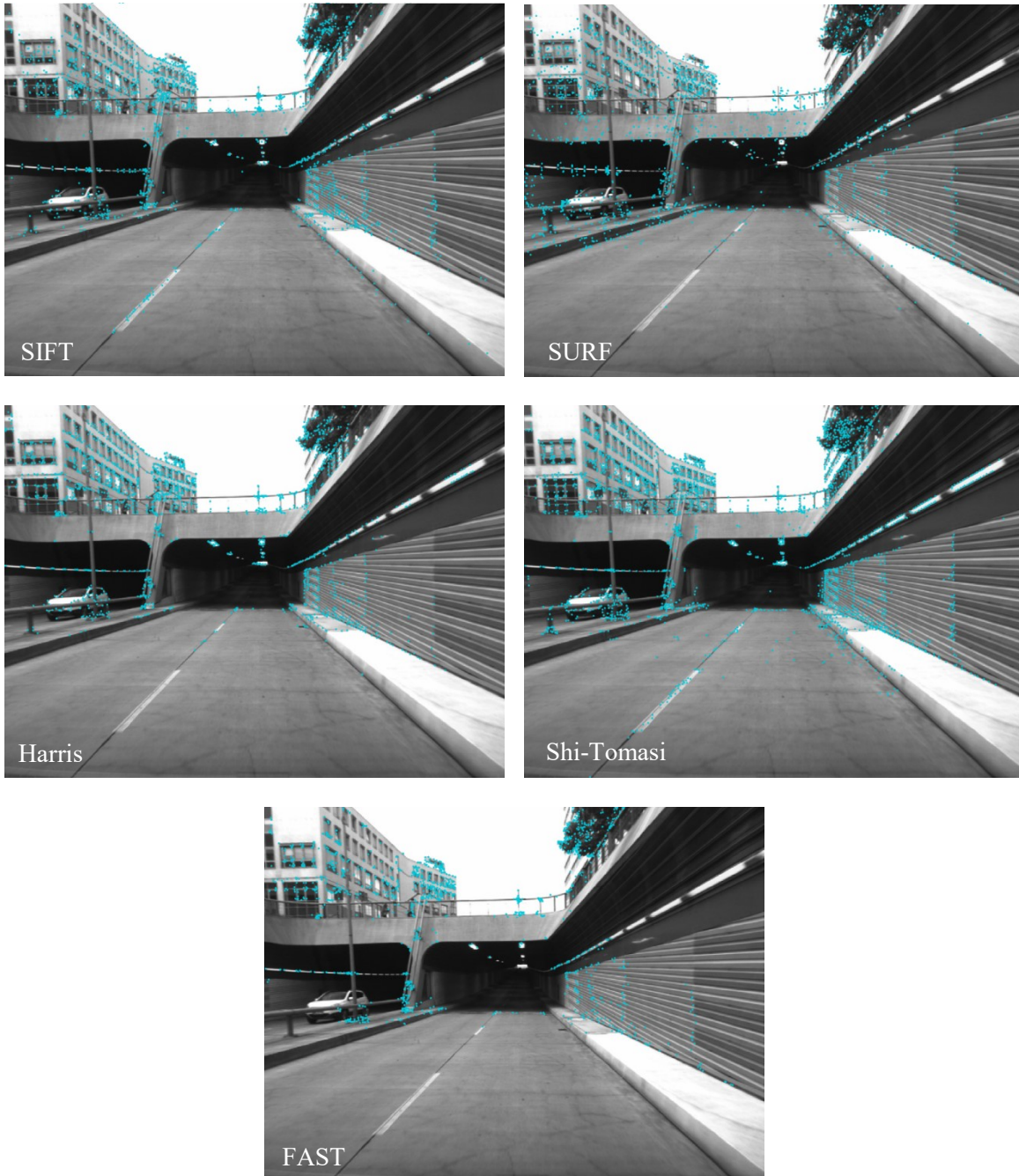


Abbildung 5-4 Gegenüberstellung der durch die Detektoren extrahierten Merkmale

Da beim normalen Betrieb im Straßenverkehr die Beleuchtung der größte Einflussfaktor ist und das Tracking nur von den gefundenen Pixelpositionen abhängig ist (keine Skalierung, Richtung, Deskriptoren), kann auf einen beleuchtungsabhängigen Schwellwert der Shi-Tomasi Auswertungsfunktion zurückgegriffen werden. Alternativ dazu wird hier der Schwellwert stark gesenkt und die Merkmale mithilfe einer Mindestdistanz im Bild verteilt. Somit wird auch in niedrig belichteten Sequenzen eine ausreichende Merkmalsdetektion gewährleistet. Beispielhaft ist dies in Abbildung 5-5 dargestellt.

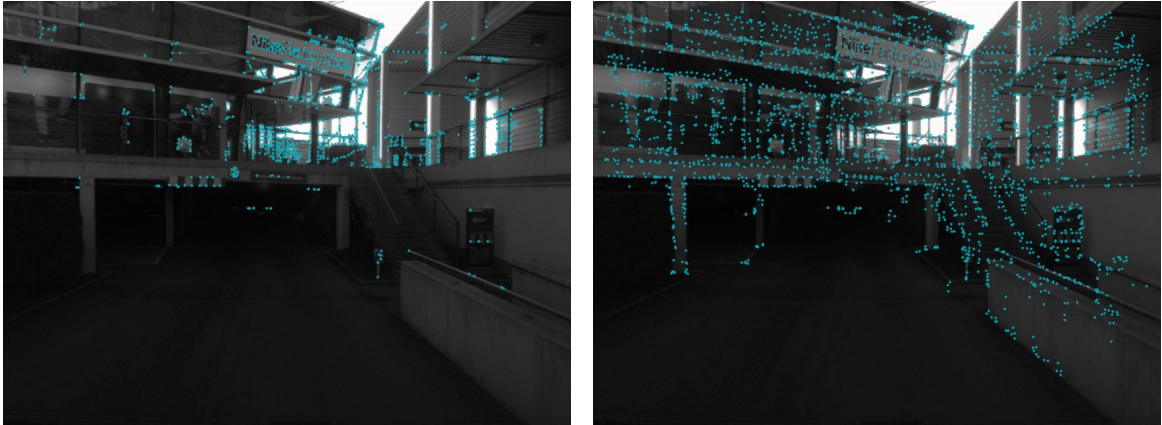


Abbildung 5-5 Detektierte Merkmale in schwach beleuchteten Sequenzen. Links: Initiale Parametrierung; Rechts: Niedriger Schwellwert und Mindestdistanz

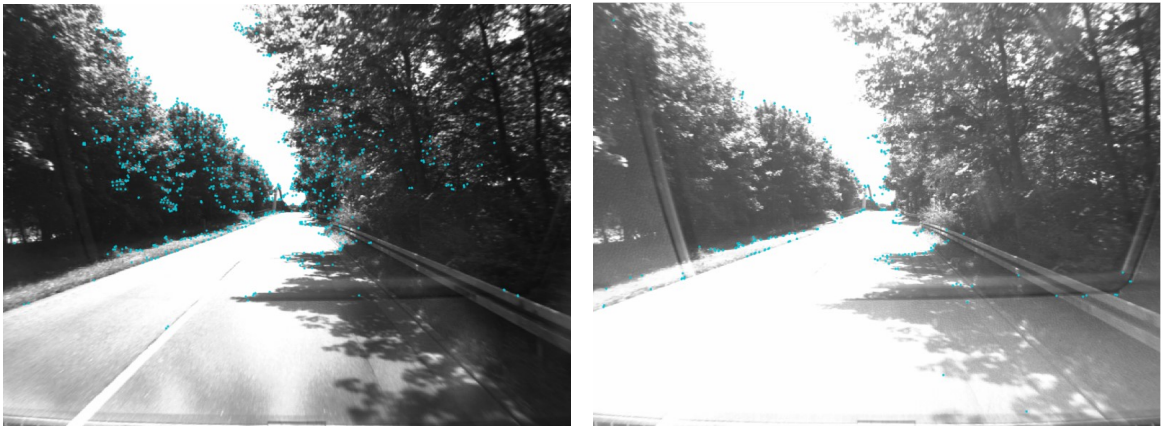


Abbildung 5-6 Verfolgte Merkmale während der Beleuchtungsänderung in zwei aufeinanderfolgenden Bildern

Als problematisch sind die in Abbildung 5-6 beispielhaft gezeigten Situationen der abrupten Beleuchtungsänderung in aufeinanderfolgenden Bildern. Obwohl eine Großzahl an Merkmalen detektiert werden, werden aufgrund der Annahme konstanter Intensitätsgradienten im KLT Tracker nur wenige Merkmale verfolgt, sodass keine umfassende Rekonstruktion erfolgen kann.

5.2.4 Rekonstruktion

Im Framework werden die Bildmerkmale mittels Shi-Tomasi Detektor extrahiert und über den KLT Tracker verfolgt. Es wird dafür die OpenCV Implementierung beider Algorithmen verwendet. Die Rekonstruktion erfolgt mit COLMAP [41], einem frei verfügbarem Tool, dem die verzeichnungsfreien Bilder, die Kalibriermatrix, die extrahierten Merkmale und Punktkorrespondenzen sowie die Bewegung der Radodometrie über eine definierte Schnittstelle zur Verfügung gestellt werden. Da auf Basis einer Bildfolge gearbeitet wird, die sich über die Zeit vergrößert, wird eine inkrementelle Rekonstruktion durchgeführt. Das Verfahren wird in Initialisierung und Hauptschleife unterteilt, welche im Folgenden vorgestellt werden.

5.2.4.1 Initialisierung

In der Initialisierungsphase, die bei Funktionsstart stattfindet, werden mittels Shi-Tomasi Detektor Merkmale im ersten Bild extrahiert und anschließend per KLT Tracker verfolgt. Über die Berechnung der essentiellen Matrix und Singulärwertzerlegung werden initiale Parameter für die Kameraorientierungen berechnet. Anschließend erfolgt die Erstellung des Modells unter Berücksichtigung des über die Radodometrie bestimmten Maßstabs (s. Abschnitt 4.4). Mittels Bündelausgleichung werden die 3D-Punkte und Kameraorientierungen optimiert, sodass am Ende der Initialisierungsphase eine metrische 3D-Rekonstruktion der Szene zur Verfügung steht. Dabei werden alle 3D-Punkte mit einem Rückprojektionsfehler größer 0.8 Pixel und mit einem Schnittwinkel kleiner 1° eliminiert. Beispielhaft ist eine initiale Punktwolke in Abbildung 5-7 dargestellt, bei der eine Zufahrt auf eine Brückenunterführung zu sehen ist. Die Initialisierung erfolgt hierbei über neun Bilder. Die Brückenunterkante ist bereits nach der Initialisierung zu erkennen und kann zur Berechnung der Hindernishöhe herangezogen werden. Der Fahruntergrund wird nur durch die Fahrbahnmarkierungen und in Nähe der Bordsteine repräsentiert.

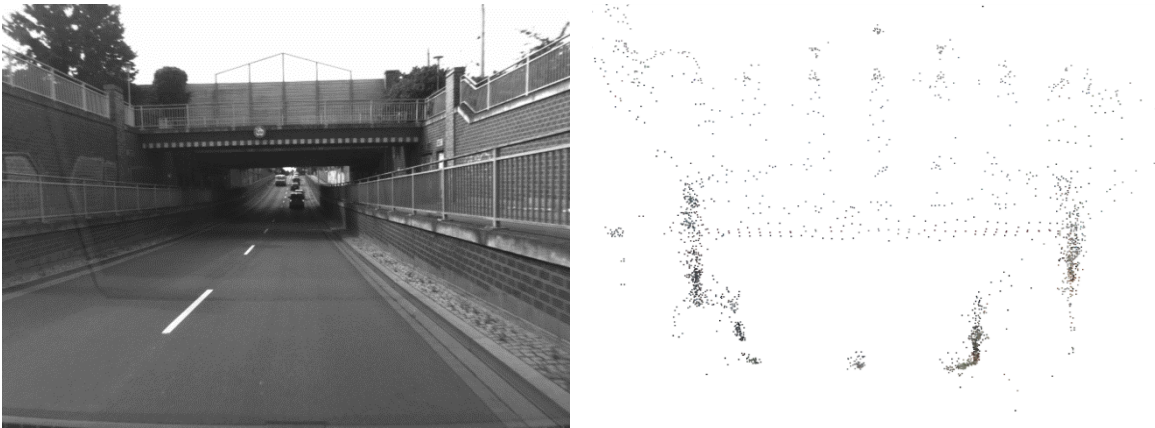


Abbildung 5-7 Zufahrt auf Brückenunterführung. Links: Bild der Sequenz; Rechts: Punktwolke der Frontansicht der Brückenunterführung nach der Initialisierungsphase (9 Bilder – Abstand zur Brücke ~32 m).

5.2.4.2 Hauptschleife

Ist die Initialisierung abgeschlossen erfolgt der Übergang in die Hauptschleife. Im Gegensatz zur vorherigen Herangehensweise, werden hierbei die initialen Kameraorientierungen nicht über die essentielle Matrix bestimmt, sondern über den räumlichen Rückwärtsschnitt. Dabei werden die vorhandenen 3D-Punkte in das neue Bild projiziert und in einer anschließenden Optimierung der Rückprojektionsfehler die Kamerapose bestimmt. Über die neuen 2D-2D Merkmalskorrespondenzen aus dem KLT Tracker können weitere 3D-Punkte trianguliert werden. Die zur Berechnung der initialen Kamerapose verwendeten 3D-Punkte können unverändert gelassen oder anhand der Informationen des aktuellen Bildes erneut trianguliert und verfeinert werden.

Das inkrementelle Rekonstruktionsverfahren hat die Eigenschaft, dass bei zunehmender Sequenzlänge ein Drift in den Kameraorientierungen zu beobachten ist [54]. Um dem entgegenzuwirken werden globale und lokale (inkrementelle) Bündelausgleichungen genutzt. Die globale Bündelausgleichung wird unter Berücksichtigung aller Kamerabilder jeweils bei Verdoppelung der Größe des rekonstruierten Modells durchgeführt, sodass zu Beginn des Verfahrens die Szene stabilisiert wird. Dementsprechend werden mit steigender Anzahl an Bildern weniger Ausgleichungen vorgenommen. Um die akkumulierten Fehler dabei zu reduzieren, wird zudem eine lokale Bündelausgleichung über die 15 letzten Bildern durchlaufen. Im realen Fahrzeug werden, im Hinblick auf die Rechenzeit, weit hinter dem Fahrzeug befindliche Kameras und

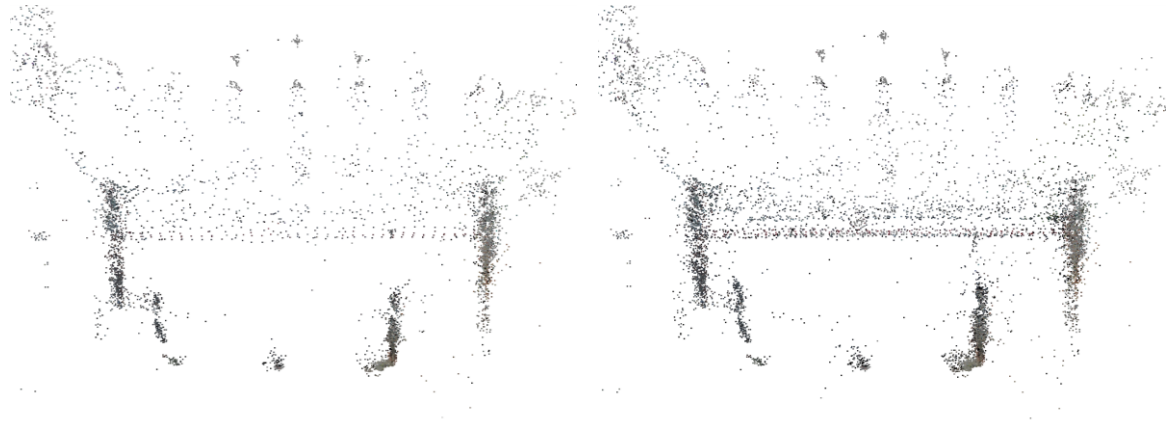


Abbildung 5-8 3D-Punktwolken der Zufahrt auf die Brückenunterführung. Links: Punktwolke nach 29 Bildern; Rechts: Punktwolke nach 49 Bildern

3D-Punkte verworfen, sodass die Szene eine nahezu konstante Größe aufweist und ausschließlich mit einer lokalen Ausgleichung gearbeitet wird. Abbildung 5-8 zeigt die rekonstruierten Punktwolke jeweils nach 29 bzw. 49 Bildern. Zu sehen ist, dass die Punktdichte am Hindernis und dessen Unterkante zunimmt, während die Fahrbahn weiterhin ausschließlich am Rand und in den Bereichen der Markierungen rekonstruiert wird. Auch in weiteren Szenarien wird deutlich, dass wenige Merkmale in homogenen Bereichen detektiert und verfolgt werden. In Abbildung 5-9 wird ersichtlich, dass der Fahruntergrund in Szenen mit geringer Umgebungsbeleuchtung nicht erkannt wird. Gleiches gilt für Hindernisse, deren Unterkante keine Markierungen aufweist und sich dadurch nicht entscheidend vom Hintergrund abhebt.

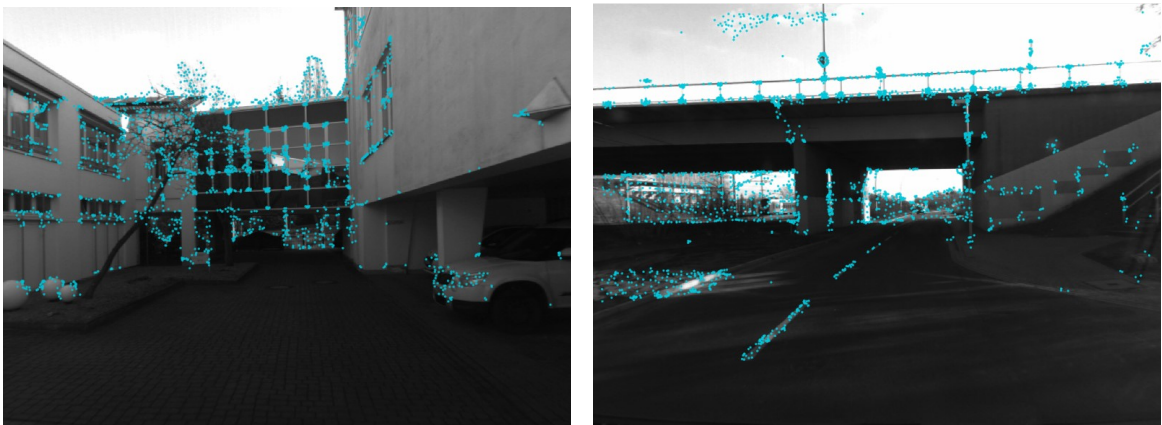


Abbildung 5-9 Merkmalsverfolgung bei geringer Umgebungsbeleuchtung; Links: Keine Detektion in dunklen homogenen Bereichen; Rechts: Unterkante der Brücke wird nicht identifiziert

Wie in den Abschnitten 5.2.1 und 5.2.3.2 beschrieben, zielt der ausgewählte Detektor auf die Identifizierung von Eckenmerkmalen im Bild ab. Auch mit einer Mindestdistanz zur besseren Verteilung der Merkmale im Bild erweisen sich Szenarien mit homogenen Bildbereichen, wie bspw. auf der Fahrbahnoberfläche, als kritisch. Zudem werden prinzipbedingt keine Merkmale an langen Kanten gefunden, welche häufig die geringste lichte Höhe eines Hindernisses darstellen. Als Alternative zu Punktmerkmalen werden im nächsten Abschnitt Kantensegmente näher beleuchtet, die während der Fahrt identifiziert, verfolgt und rekonstruiert werden. Darüber hinaus werden in Kapitel 6 direkte Verfahren vorgestellt, die merkmalsunabhängig, d.h. ausschließlich auf Bildintensitäten basierend, eine Rekonstruktion vornehmen.

5.3 Kanten und Linien

Bereits Anfang der 1980er Jahre wurden Linien im Kontext des 3D-Sehens verwendet. Sie wurden in zahlreichen Anwendungen zur Korrespondenzsuche zwischen Karten und Luftbildaufnahmen [107], zwischen Stereobildpaaren [108], zur Verfolgung [109] und Navigation [110] genutzt. Erst mit dem Einzug von punktbasierten Rekonstruktionsmethoden auf Basis von effizient und robust detektier- und beschreibbaren Merkmalen war es möglich große realistische Datensätze als 3D-Punktwolke darzustellen. Die Nutzung von Linien war indes auf kleine und beschränkte Umgebungen limitiert und mit einem erhöhten Rechenaufwand verbunden [111]. Die Entwicklung hat in den letzten Jahren jedoch weitere linienbasierte Verfahren hervorgebracht, die ein größeres Anwendungsspektrum mit hoher Robustheit ermöglicht [112, 113, 114]. Ähnlich wie bei den punktbasierten Verfahren können die linienbasierten Systeme in Kategorien SfM, SLAM und MVS eingeteilt werden. Die Linienrekonstruktion stellt im vorliegenden Fall nur eine Erweiterung der bereits durchlaufenen SfM-Pipeline dar. Dies bedeutet, dass die detektierten und gematchten Linien nur zur Rekonstruktion im 3D-Raum und nicht zur Berechnung der Kameraposition genutzt werden. Insofern handelt es sich hierbei um ein Linien-MVS Verfahren, welches integriert und auf die Eignung zur verbesserten Fahrbahn- und Hindernisdetektion bewertet wird.

5.3.1 Kanten- und Liniendetektion

Die Kantendetektion ist Teil der Bildverarbeitung und dient in erster Linie der Segmentierung der Bildelemente. Dabei sollen flächige und homogene Bereiche voneinander getrennt werden. Oft geschieht dies aufgrund der Berechnung des Grauwertgradienten zwischen einem Pixel und dessen direkter Umgebung. Gängige Operatoren, die zur Faltung angewendet werden, sind bspw. *Sobel* [115] oder *Canny* [116]. Ergebnis ist ein Kantenbild, wie in Abbildung 5-10 oben dargestellt. Der Informationsgehalt dieser Bilder ist enorm, sie eignen sich jedoch nicht für eine 3D-Rekonstruktion. Dafür sollen nur bestimmte Kantensegmente genutzt werden, die zuverlässig über eine Serie von Bildern verfolgt werden können. Mit dem Begriff Kantensegment wird in diesem Kontext eine 2D-Linie beschrieben, welche einen definierten Anfangs- und Endpunkt im Bild besitzt. Grundsätzlich kann ein beliebiger Detektor für Linien genutzt werden, welcher jedoch möglichst robust gegenüber Bildrauschen sein sollte. Ballard [117] befasste sich bspw. mit der Detektion von Linien, bei der ein Kantenbild mittels *Hough* Transformation in einen binären Parameterraum überführt werden. Innerhalb dieses Parameterraums werden mittels lokalen Maxima anschließend Linien detektiert. Eine Erweiterung dazu liefert [118], welche es ermöglicht Linien nach ihrer Länge zu filtern und einzelne kleine Segmente zu einer gemeinsamen Linie zusammenzufassen. Detektoren basierend auf der *Hough* Transformation eignen sich grundsätzlich zur Detektion von wichtigen Linien, jedoch ist ihre Genauigkeit vergleichsweise gering. Darüber hinaus sind sie sehr anfällig gegenüber Rauschen benötigen aufgrund ihrer Komplexität einen hohen Rechenaufwand [112].

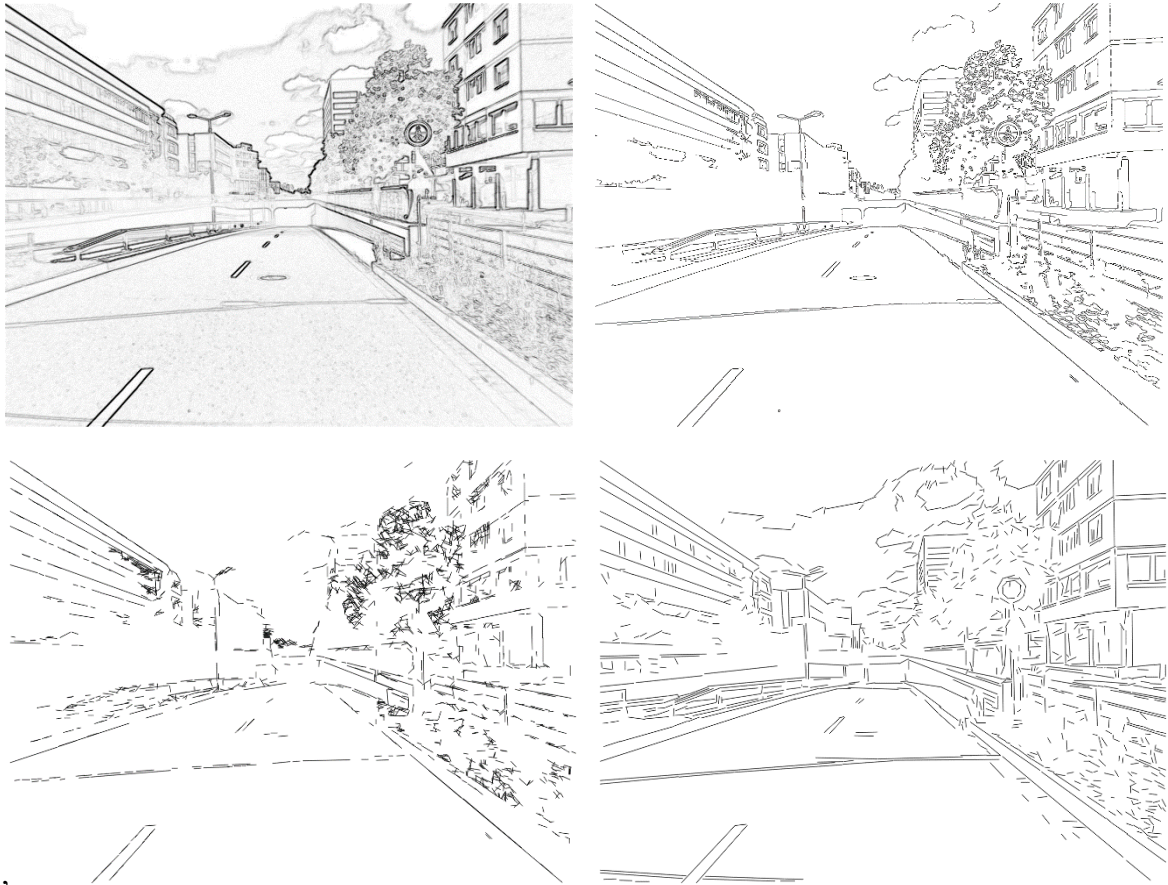


Abbildung 5-10 Kantenbilder. Oben links: Sobel; Oben rechts: Canny; Unten links: Probabilistische Hough Linien; Unten rechts: Line Segment Detector

Neuere Detektoren für Linien sind bspw. *LSD Line Segment Detector* [77] oder *EDLines* [119], welche Kantenpixel mit gleicher Orientierung iterativ zusammenfassen, diese mittels *A Contrario* Modell analysieren und in der Lage sind subpixelgenaue Linien zu erzeugen. Der Algorithmus ist adaptiv, sodass keine festen Schwellwerte oder Parameter für die Detektion eingestellt werden müssen. Damit eignen sich diese Algorithmen für dynamische Szenarien im Straßenverkehr und bieten die Grundlage für eine möglichst effiziente und genaue 3D-Rekonstruktion von Linien. Im vorliegenden Fall wird der *LSD* genutzt. Exemplarisch sind in Abbildung 5-10 unten Linien gezeigt, welche mittels Hough Lines nach [118] und *LSD* detektiert wurden.

5.3.2 Korrespondenzsuche

Die Korrespondenzsuche beschreibt den Prozess des paarweisen Zuordnens von Linien unterschiedlicher Bilder. Es bieten sich dabei mehrere Verfahren an, welche in zwei Klassen unterteilt werden können. Zum einen können Linien aufgrund ihrer paarweisen optischen und geometrischen Eigenschaften direkt miteinander gematcht werden. Ein zweiter Ansatz ist das Matchen aller Linien in einem Durchgang unter Einbeziehung globaler Eigenschaften, wie relative Endpunktpositionen, Winkel oder Schnittpunkte. In [120] wird ein Verfahren zum Linienmatching aus Triplets beschrieben. Die Bekanntheit der Kameraposen und des Trifokaltensors wird dabei für die intensitätsbasierte Korrespondenzsuche vorausgesetzt. Neben Deskriptoren, die Informationen aus Histogrammen nutzen [121], wurden patchbasierte Deskriptoren entwickelt [122, 123], welche an die Grundidee von Punktdeskriptoren wie SIFT und SURF angelehnt sind. Neuere Methoden kombinieren Ansätze, welche zumeist optische und geometrische Einschränkungen berücksichtigen.

Sie sind generalisiert entworfen und müssen möglichst robust und genau sein, um eine Vielzahl an korrekten Matches erzeugen, auch wenn nur zwei Bildern vorliegen.

Wesentlicher Unterschied zu oben genannten Verfahren ist die in dieser Arbeit vorliegende Informationsdichte, die aus der Bildfolge stammt. Unter dieser Voraussetzung können bereits berechnete Kameraposen und Bilder genutzt werden, um die Korrespondenzsuche und die Qualität von neu rekonstruierten Linien zu bewerten. In [112] wird ein Verfahren für Bildfolgen entwickelt, welches sich grundlegend auf die Einschränkungen stützt, die sich aus der Epipolargeometrie ergeben. Rechenaufwendige Deskriptoren werden vermieden. Zunächst werden alle Linien l'_k mit LSD in einem Bild I' detektiert. Anschließend werden zu jeder Linie l'_k die Epipolarlinien ihrer Endpunkte p'_k und q'_k im Nachbarbild I'' bestimmt (vgl. Abbildung 5-11).

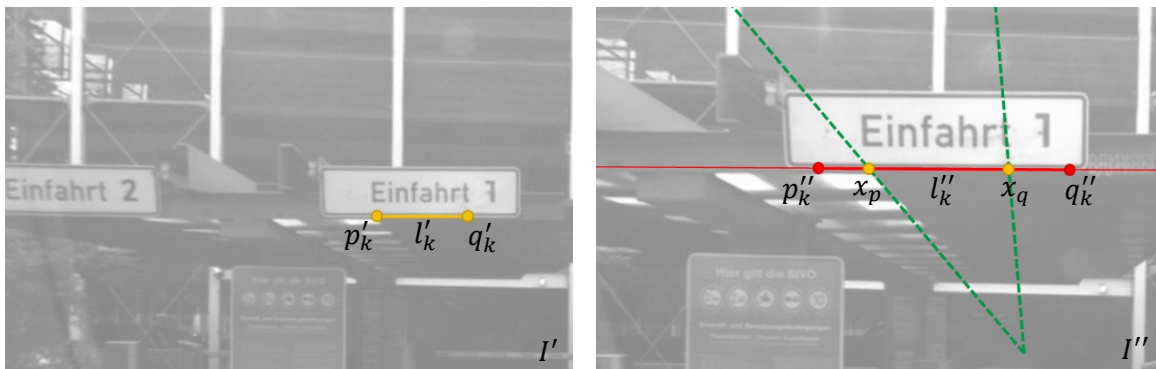


Abbildung 5-11 Erstellen einer Linienkorrespondenz mit Berücksichtigung der Epipolargeometrie nach [112]. Links: Detektierte Linie l'_k im Bild I' mit den Endpunkten p'_k und q'_k ; Rechts: Detektierte und ins Unendliche verlängerte Linie l'_k im Bild I'' . Schnittpunkte x_p und x_q der Epipolarlinien (grün) mit der unendlichen Linie, sowie die Endpunkte p''_k und q''_k .

Die dort detektierten Linien l''_k werden ins Unendliche verlängert und mit den Epipolarlinien (grün), welche mit den berechneten Kameraposen der SfM-Punktpipeline ermittelt werden, geschnitten. Die Schnittmenge zwischen l'_k und l''_k wird über das Verhältnis des euklidischen Abstands der Punktpaare $d(x_p, x_q)$ und $d(p''_k, q''_k)$ bestimmt. Überschreitet die Schnittmenge zweier Linien einen bestimmten Schwellwert, so wird von einer Korrespondenz ausgegangen. Dies ergibt eine Vielzahl an möglichen Korrespondenzen, die aufgrund ihrer Schnittmenge zunächst auf n Kandidaten reduziert werden können. Die finale Korrespondenz wird über die Bewertung der Ähnlichkeit der 2D- und hypothetischen 3D-Position der Linienpaare bestimmt. Dabei fließt die Kenntnis von mind. vier Nachbarbildern ein. Im nächsten Abschnitt wird darauf vertieft eingegangen.

5.3.3 Rekonstruktion

Die Rekonstruktion der in zwei Bildern identifizierten homologen Linien l' und l'' erfolgt über die Ermittlung der Schnittgeraden L der beiden projizierten Ebenen π' und π'' .

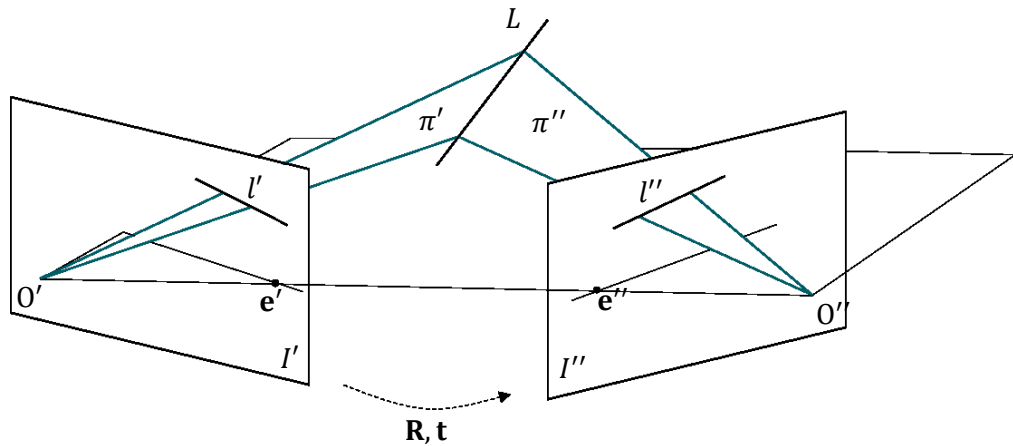


Abbildung 5-12 Linienrekonstruktion

Die Ebenen sind über $\pi' = \mathbf{P}'^T \mathbf{I}'$ und $\pi'' = \mathbf{P}''^T \mathbf{I}''$ definiert. Die Schnittgerade L ergibt sich durch

$$\mathbf{L} = \begin{bmatrix} \mathbf{I}'^T \mathbf{P}' \\ \mathbf{I}''^T \mathbf{P}'' \end{bmatrix}. \quad \text{Gl. (5-17)}$$

Die Rekonstruktion der 3D-Linie aus zwei Bildern erzeugt eine unendliche Gerade. Um endliche Linien zu erzeugen, müssen Korrespondenzen aus mindestens drei Bildern vorliegen. Bei der Triangulierung der Schnittgeraden aus mehreren Bildern werden die einzelnen Ebenen π^n in einer Matrix \mathbf{A} der Größe $n \times 4$ zusammengefasst, wobei jede Reihe durch eine Ebene repräsentiert wird. Mittels Singulärwertzerlegung $\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{V}^T$ werden die beiden größten Singulärwerte von \mathbf{V} berechnet, die den Rang-2 von \mathbf{A} am besten approximieren und zur Definition der 3D-Linie gewählt. Eine weitere Variante ist die Bestimmung der idealen 3D-Linie über die Minimierung des Rückprojektionsfehlers in jedem Bild [45]. Im vorliegenden Fall wird das Triangulierungsverfahren von [124] eingesetzt. Dabei werden 3D-Hypothesen (hier beispielhaft $H' \rightarrow''$ und $H'' \rightarrow'$) trianguliert, die auf einer gemeinsamen 3D-Linie L liegen und deren Endpunkte über die Endpunkte der 2D-Linien definiert wird (vgl. Abbildung 5-13).

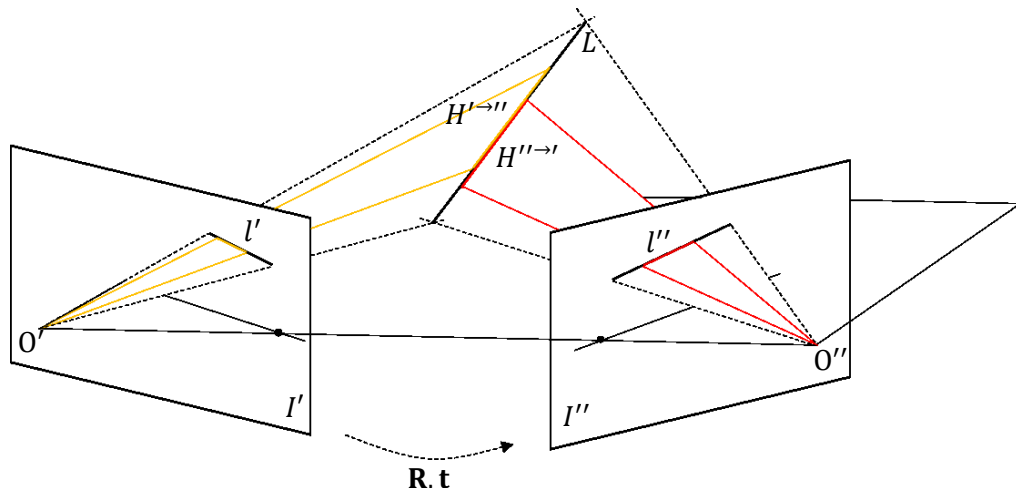


Abbildung 5-13 3D-Hypothesen $H' \rightarrow''$ (gelb) und $H'' \rightarrow'$ (rot) und die gemeinsame 3D-Linie L . Entnommen aus [124].

Diese Art der Korrespondenzsuche erzeugt jedoch eine große Anzahl an Hypothesen. Die korrekte Korrespondenz und 3D-Linie wird mithilfe von bis zu m Nachbarbildern bestimmt. Dabei werden weitere Hypothesen bis hin zu $H' \rightarrow^m$ aufgestellt und die Position und Lage der möglichen 3D-Linie

bewertet. Dies erfolgt innerhalb des Clusterings der 3D-Linien. Da der Maßstab bekannt ist, kann eine maximale Abweichung der 3D-Linien im Raum (hier: 5 cm) angegeben und alle Hypothesen verworfen werden, die außerhalb dieses Bereichs liegen. Die 3D-Linien müssen Hypothesen aus mindestens vier Bildern aufweisen. Anschließend erfolgt die Berechnung einer Matrix, welche die Hypothesen, deren Endpunkte und die 2D Residuen enthält. Über eine Singulärwertzerlegung und Berechnung des Schwerpunkts der Endpunkte wird eine resultierende 3D-Linie bestimmt. Zuletzt werden alle 2D-Segmente auf diese Linie projiziert und 3D-Segmente berechnet, die mindestens drei Hypothesen enthalten. Detaillierte Informationen lassen sich in [111, 112, 124] finden.

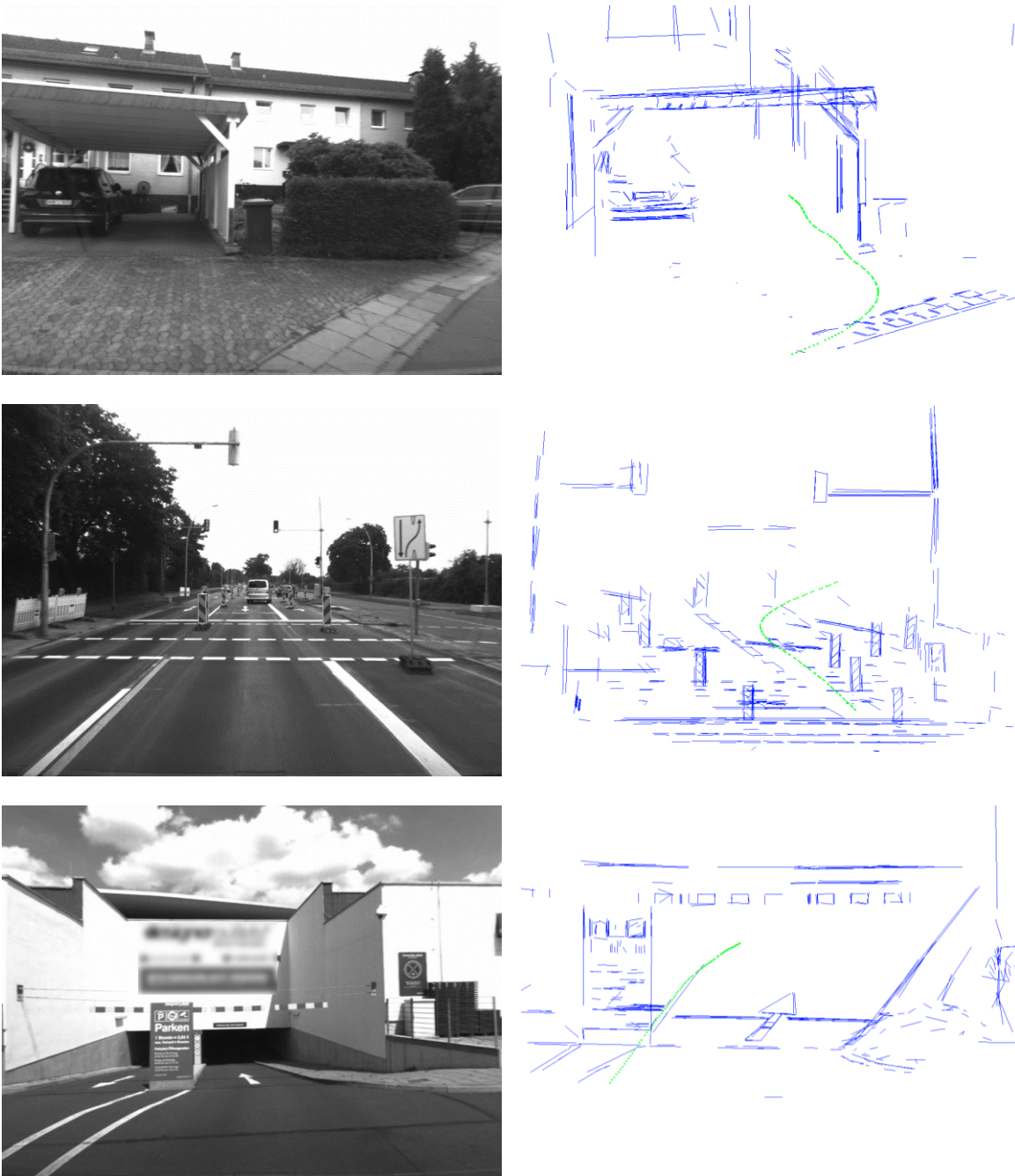


Abbildung 5-14 Rekonstruktion der 3D-Linien (Blau) mit Fahrzeugtrajektorie (Grün).

Abbildung 5-14 zeigt die Linienrekonstruktion in ausgewählten Sequenzen. Es wird sichtbar, dass sich das Verfahren sowohl für die Generierung von 3D-Linien auf Hindernissen als auch der Fahrbahnoberfläche eignet. Kurze Linien, welche bspw. im oberen Teil des Bildes im Bereich des Pflasters auftreten, werden aufgrund ihrer Länge im Bild gefiltert. Dies gilt ebenso für Fahrbahnmarkierungen. Entsprechend werden diese zum größten Teils quer zur Fahrzeugtrajektorie rekonstruiert. Die stückweise Unterteilung von Warnbaken können dennoch rekonstruiert werden, sofern sich das Fahrzeug mit ausreichend langsamer Geschwindigkeit auf das Hindernis zubewegt (s. Unten). Aber auch hier wird sichtbar, dass aufgrund des geringen Kontrasts zum Hintergrund weiße Abschnitte der Warnbake nicht erkannt werden. Des Weiteren werden an einer Kante oft mehrere Linien rekonstruiert. Unter Reduzierung der zugelassenen räumlichen Entfernung der Hypothesen kann dies vermieden werden. Damit geht jedoch eine kleinere Linienanzahl einher, die es nicht ermöglicht eine Hindernisdetektion zu realisieren. Eine Evaluierung der 3D-Linien wird in Kapitel 8 vorgenommen.

Alternativ zur vollen Rekonstruktion von 3D-Linien, können 2D-Linien bereits als Indiz für eine Hinderniskante innerhalb des Bildes betrachtet werden. Sofern die Ausrichtung der Linien nahezu horizontal verläuft, wird um die gefundenen Linien eine Region of Interest (ROI) bestimmt. Innerhalb dieser ROI wird der Schwellwert für die Auswertungsfunktion der Punktdetektion verringert, sodass mehr Punkte in diesem Bereich akzeptiert werden. Gegebenenfalls muss dafür der Mindestpunktabstand reduziert und das Trackingfenster verkleinert werden, sodass die größere Punktzahl verlässlich verfolgt wird.

6 Direkte und hybride Rekonstruktionsverfahren

Indirekte Rekonstruktionsverfahren setzen im ersten Schritt auf die Extraktion und Verfolgung von diskreten Bildmerkmalen. Unter der Annahme von fehlerbehafteten Messwerten wird im zweiten Schritt die 3D-Rekonstruktion ausgeführt, bei der das geometrische Modell und die Kamerabewegung geschätzt wird. Die intermediäre Repräsentation der Bildinhalte über Merkmale, d.h. Punkte oder auch Linien, wird auch als *indirekte* Methode bezeichnet. Alternativ dazu wurden vor allem in der letzten Dekade *direkte* Ansätze vorgestellt, die diese Darstellung überspringen und unmittelbar mit den eigentlichen Sensorinformationen arbeiten, d.h. den Intensitätswerten der Pixel. Hierbei wird im Gegensatz zur Minimierung des geometrischen Rückprojektionsfehlers mittels Bündelausgleichung der sogenannte photometrische Fehler minimiert. Darüber hinaus existieren *hybride* Verfahren, welche indirekte und direkte Ansätze kombinieren. Auf eine Auswahl wird hier näher eingegangen.

6.1 Verfahren

Direkte Verfahren können zunächst in dichte und semi-dichte Verfahren unterteilt werden. Während bei dichten Verfahren die Informationen aller Pixel evaluiert werden, benutzen semi-dichte Varianten ausschließlich die Pixel, deren Helligkeitsgradient bedeutend ist. Für die Durchfahrthöhenerkennung wird ausschließlich eine dünn besetzte Punktwolke benötigt, sodass hierbei auf eine Vorstellung von dichten Verfahren verzichtet wird. Ein hybrides Verfahren stellt SVO (Semi-direct Visual Odometry) dar [125]. Es vereint den merkmalsbasierten und direkten Ansatz, wobei Merkmale mit dem FAST Detektor extrahiert und deren direkte Nachbarschaft abgespeichert werden. Punkttiefen werden über Optimierung des photometrischen Fehlers zwischen der Merkmalsnachbarschaft und dem aktuellen Bild berechnet. Die Berechnung der Kamerabewegung erfolgt über den direkten Ansatz. SVO ist ein sehr effizientes Tool, welches in Echtzeit auf einer CPU läuft und später durch die Integration von Linien erweitert wurde. Ein vollständig direktes Verfahren ist LSD-SLAM (Large-Scale Direct Monocular SLAM) [126]. Es bedient sich der *Keyframe* Formulierung, welche pixelweise nach Intensitätsgradienten untersucht werden. Für jeden Kandidaten erfolgt die Bestimmung der inversen Tiefe über einen Extended Kalman Filter unter Annahme kleiner Basisbreiten zwischen dem Keyframe und den nachfolgenden Bildern. Das Verfahren beruht auf einer photometrischen Kalibrierung, welche unter verschiedenen Belichtungszeiten durchgeführt werden muss. Erfolgt diese nicht, so kommt es häufig zu Abrissen im Tracking bei variierender Umgebungsbeleuchtung. LSD-SLAM kann in Echtzeit auf einer CPU berechnet werden. Es wird z.B. für Augmented Reality Anwendungen auf Smartphones integriert [127].

Das in dieser Arbeit integrierte und untersuchte Verfahren ist DSO (Direct Sparse Odometry). Es wurde in [128] vorgestellt und kann als eine Erweiterung von LSD-SLAM angesehen werden.

6.2 Direct Sparse Odometry

DSO unterscheidet sich zur dem bereits vorgestellten indirekten Verfahren (SfM) in mehreren Bereichen. Die wesentlichen Besonderheiten sollen hierbei erläutert werden.

6.2.1 Grundlagen

Bildpunktauswahl

Anstatt einen der in Kapitel 5.2.1 vorgestellten Merkmalsdetektoren zu nutzen, werden bei DSO alle Pixel eines Keyframes auf ihren Intensitätsgradienten untersucht. Dabei wird das Bild in 32x32 Blöcke unterteilt und in jedem Block das Mittel aller Intensitätsgradienten als Schwellwert bestimmt. Die Bildpunkte mit den höchsten Gradienten über diesem Schwellwert werden für die anschließende Verfolgung abgespeichert. Wird eine Mindestanzahl an Bildpunkten nicht erreicht, so erfolgt eine Anpassung der Blockgröße. Vorteil ist, dass nicht nur Eckenmerkmale, sondern auch gesamte Kanten und Konturen detektiert werden. Die Unterteilung in Blöcke und die Nutzung des Mittelwerts der Gradienten ermöglicht die Detektion von Merkmalen im gesamten Bild, auch auf wenig strukturierten Bereichen, wie der Fahrbahn oder einfarbigen Wänden.

Neben den eigentlichen Bildpunkten wird zudem jeweils eine definierte Nachbarschaft von acht weiteren Pixel abgespeichert, die nach [128] den empirisch ermittelten besten Kompromiss zwischen Geschwindigkeit, ausreichender Information für das Tracking und Umgang mit Bewegungsunschärfe darstellt.

Keyframe Repräsentation

Bei dem klassischen indirekten Verfahren sind die Berechnung der Kamerapose und die Erzeugung der 3D-Punktswolke eng miteinander verknüpft. Beide Schritte werden mit jedem neuen aufgenommenen Bild durchlaufen. Diese zeitintensive Berechnung kann durch die Nutzung der Keyframe Repräsentation optimiert werden. Während die Lokalisierung auch hier in jedem Bild stattfindet, wird die Extraktion neuer Bildpunkte und die Rekonstruktion ausschließlich während eines Keyframes durchgeführt. Die Optimierung wird ähnlich zum indirekten Verfahren über den gesamten Zeitraum durchgeführt, wobei in DSO 3D-Punkte nicht durch X, Y, Z -Koordinaten, sondern durch ihre inverse Tiefe innerhalb eines Referenzbildes beschrieben werden.

Inverse Tiefe

Erstmals in [129] vorgestellt, eröffnet die Nutzung der inversen Tiefe die Position und Genauigkeit eines 3D-Punkts effizient bewerten zu können. Sie findet häufig Anwendung in aktuellen filterbasierten SLAM Verfahren, da ein 3D-Punkt erst nach Erreichen einer Normalverteilung in Filteralgorithmen wie dem Extended Kalman Filter für eine Prädiktion verwendet werden kann. Die inverse Tiefe bietet eine mathematische Grundlage, die es ermöglicht einen 3D-Punkt direkt nach der ersten Triangulierung normalverteilt darzustellen und somit lange Initialisierungsphasen für neue Punkte zu vermeiden. Ein 3D-Punkt wird hierbei mit dem 6-dimensionalen Vektor

$$\mathbf{X} = [X_K \ Y_K \ Z_K \ \phi \ \omega \ \rho] \quad \text{Gl. (6-1)}$$

beschrieben, wobei X_K, Y_K und Z_K die Position der Kamerazentrums und ϕ, ω Azimuth und Elevation im Kamerakoordinatensystem darstellen. ρ ist die inverse Tiefe, welche die Inverse der Distanz zwischen Objektpunkt und des Kamerazentrums darstellt.

Die euklidische Position des 3D-Punkts wird über

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_F \\ Y_F \\ Z_F \end{bmatrix} + \frac{1}{\rho} \mathbf{m}(\phi, \omega) \quad \text{Gl. (6-2)}$$

mit dem Richtungsvektor

$$\mathbf{m} = [\cos\omega \cdot \sin\phi, -\sin\omega, \cos\omega \cdot \cos\phi]^T \quad \text{Gl. (6-3)}$$

berechnet. Anschließend erfolgt die Transformation der 3D-Punkte in das USK.

Photometrischer Fehler

Im Gegensatz zur indirekten Variante, bei der der geometrische Rückprojektionsfehler minimiert wird, erfolgt die Optimierung bei direkten Verfahren über die Minimierung des photometrischen Fehlers. Parallelen können dabei zur Bestimmung des optischen Flusses gesehen werden. Beim optischen Fluss werden die Intensitätsgradienten genutzt, um das Flussfeld zu bestimmen. Bei den direkten Methoden wird unter Minimierung der Abweichung der Intensitätsgradienten zweier Bilder die Geometrie der Szene bestimmt. Der Fehler eines einzelnen Bildpunktes berechnet sich zu:

$$E_{x^j} = \sum_{\mathbf{x} \in N_{\mathbf{x}}} w_{\mathbf{x}} \left\| \left(I^j(\hat{\mathbf{x}}) - b^j \right) - \frac{t^j e^{a^j}}{t^i e^{a^i}} \left(I^i(\mathbf{x}) - b^i \right) \right\|_{\gamma} \quad \text{Gl. (6-4)}$$

$N_{\mathbf{x}}$ ist dabei die Nachbarschaft von \mathbf{x} , t^i und t^j die Belichtungszeiten der Bilder I^i und I^j . $\|\cdot\|_{\gamma}$ beschreibt die Huber Norm, welche kleine Fehler quadratisch und große Fehler linear bestraft. $\hat{\mathbf{x}}$ ist der projizierte Punkt im Bild I^j mit der inversen Tiefe $d_{\mathbf{x}}$.

$$\hat{\mathbf{x}} = \Pi_c(\mathbf{R}\Pi_c^{-1}(\mathbf{x}, d_{\mathbf{x}}) + \mathbf{t}) \quad \text{Gl. (6-5)}$$

Π_c beschreibt die Projektion unter Nutzung der Kameraparameter. Da keine aufwendige photometrische Kalibrierung während der Produktion durchgeführt wird, kommt die hinterlegte Helligkeitsfunktion zum Tragen, welche über $e^{-a^i}(I^i - b^i)$ verschiedene Belichtungszeiten berücksichtigt. Die Parameter werden dabei während der Verarbeitung geschätzt. Neben der Huber Norm wird ein zusätzlicher Wichtungsfaktor $w_{\mathbf{x}}$ eingeführt, welcher Bildpunkte mit starkem Gradienten geringer wichtet. Der gesamte photometrische Fehler wird mit

$$E_{photo} = \sum_{i \in M} \sum_{\mathbf{x} \in X_i} \sum_{j \in obs(\mathbf{x})} E_{x^j} \quad \text{Gl. (6-6)}$$

formuliert und beinhaltet alle Bilder M , alle beobachteten Bildkoordinaten X_i im Bild i und alle Bilder $obs(\mathbf{x})$, in denen \mathbf{x} sichtbar ist.

6.2.2 Framework

Das DSO Framework wird in zwei wesentliche Instanzen, das Front End und das Back End, unterteilt. Das Front End ist unter anderem für das Finden von relevanten Bildpunkten in den Keyframes verantwortlich. Während der Initialisierung werden diese Punkte aus bereits einer Kamerapose mit einer inversen Tiefe $\rho = 1$ und hoher Varianz rekonstruiert. Des Weiteren wird im Front End das Tracking der Bildpunkte unter Minimierung des photometrischen Fehlers durch Suche entlang der Epipolarlinie durchgeführt. Die Kamerapose des neuen Bildes wird dabei ermittelt und die inverse Tiefe und Varianz aller 3D-Punkte aktualisiert (s. Abbildung 6-1).

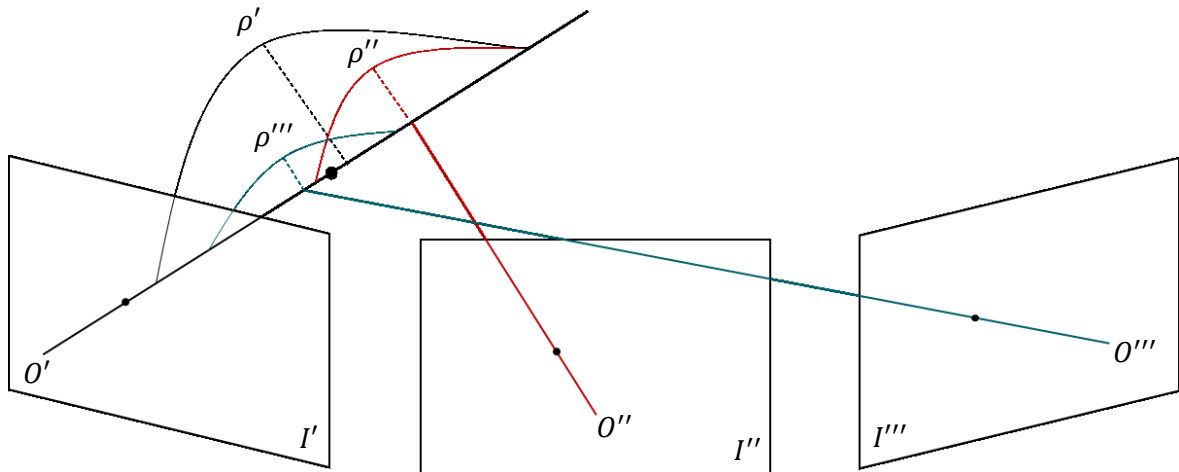


Abbildung 6-1 Initialisierung eines 3D-Punktes mit $\rho = 1$ und hoher Varianz im Bild I' . Anschließende Optimierung der inversen Tiefe und Varianz durch die nachfolgenden Bilder I'' und I''' .

Ist ein 3D-Modell vorhanden, so werden die neu extrahierten Bildpunkte (Kandidaten) eines Keyframes nicht direkt rekonstruiert/aktiviert. Zunächst werden aktive 3D-Punkte in das Bild projiziert und dann Kandidaten aktiviert, die den Abstand zu bisherigen 3D-Punkten maximieren. Somit wird eine möglichst ideale Verteilung der 3D-Punkte im Raum geschaffen. Sollte das Tracking fehlschlagen und der Fehler der geschätzten Kamerapose eines neuen Bildes mehr als doppelt so hoch wie der des vorherigen Bildes sein, so wird die Kamerapose verworfen und weitere Kombinationen aus verschiedenen Translationen und Rotationen geprüft. Unter Annahme von keiner, halber, einfacher bzw. doppelter Translation und 27 beliebigen Rotationen in verschiedenen Richtungen wird die Pose ausgewählt, die den kleinsten Fehler erzeugt. Somit kann auch bei Fehlschlagen des Trackings meist eine Relokalisierung stattfinden. Das Front End legt zudem fest, ob ein neues Bild als Keyframe eingestuft wird. Einfluss darauf nehmen die Änderung der Translation und des relativen Helligkeitswerts zum letzten Bild sowie die Änderung des Sichtbereichs zum letzten Keyframe. Im Mittel wird in den hier untersuchten Sequenzen jedes fünfte Bild als Keyframe deklariert. Das Back End von DSO übernimmt die Aufgabe der Marginalisierung und der lokalen Optimierung der Kameraposen und aller aktiven 3D-Punkte. Im Back End wird eine Maximalanzahl von sieben Keyframes aktiv gehalten, über die eine Optimierung nach Gauß-Newton stattfindet. Bei Hinzukommen eines neuen Keyframes, wird ein alter marginalisiert, wobei darauf geachtet wird, dass die Verteilung der Keyframes innerhalb der Bildfolge homogen ist. DSO verzichtet dabei gänzlich auf eine globale Optimierung.

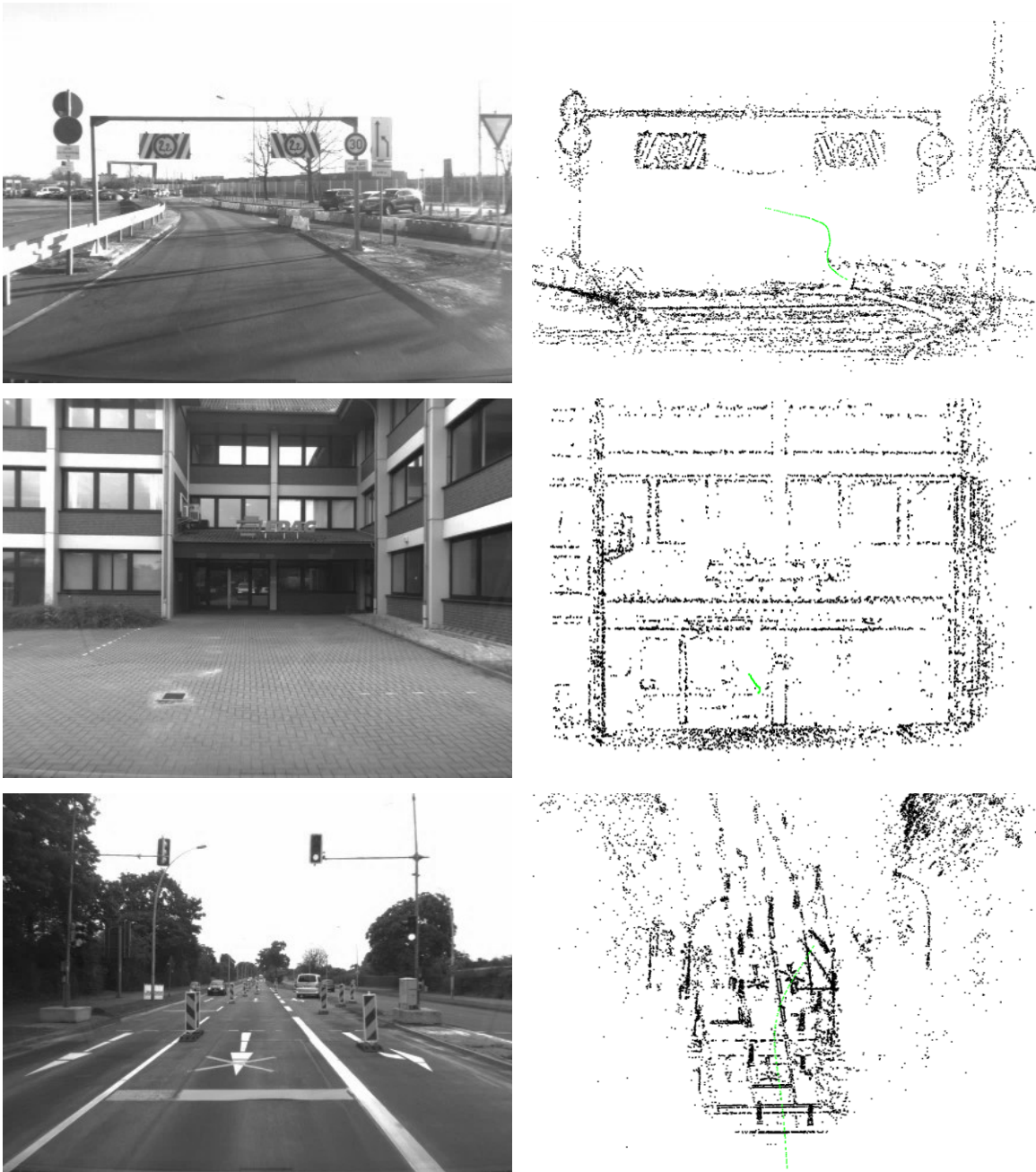


Abbildung 6-2 Direktes Verfahren. 3D-Rekonstruktion mit Fahrzeugtrajektorie (Grün)

In Abbildung 6-2 werden die Punktwolken drei verschiedener Sequenzen, die mit DSO rekonstruiert wurden, dargestellt. Fahruntergrund als auch Umgebungsstruktur werden mit einer hohen Punktdichte vorrangig entlang langer Gradienten sehr präzise rekonstruiert. Die Detektion der wahren Hindernisunterkante wird somit erleichtert. Aus der unteren Sequenz geht jedoch auch eine hohe Anzahl an einzelnen Punkten hervor, die im Rahmen der Weiterverarbeitung entfernt werden müssen. Es handelt sich hierbei um 3D-Punkte, die initialisiert wurden und deren Optimierung nicht konvergiert ist. Es wird vermutet, dass dies durch kurze Abstände der Keyframes oder bspw. durch Okklusion verlorene Tracks hervorgerufen wird. Im Gegensatz zum indirekten Verfahren wird bei DSO während der Rekonstruktion keine Einschränkung hinsichtlich minimalem Schnittwinkel getroffen. Somit werden tendenziell mehr Ausreißer in großen Distanzen rekonstruiert.

7 Fahrergrund- und Hindernisdetektion

In diesem Kapitel werden die Verfahrensschritte dargestellt, um potentielle nicht unterfahrbare Hindernisse innerhalb der Punktwolke zu identifizieren. Zunächst werden dafür in der Vorverarbeitung die Punktwolke unterteilt und Ausreißer eliminiert. Anschließend erfolgt eine Klassifizierung in Fahrergrund, Objekte/Struktur und potentielle Hindernisse (s. Abbildung 7-1).

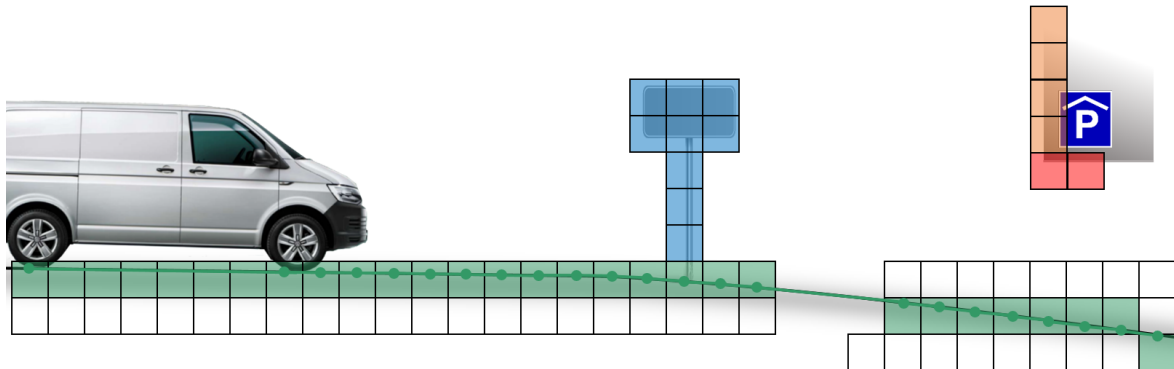


Abbildung 7-1 Unterteilung und Klassifizierung des 3D-Punktraums. Grün: Fahrergrund; Blau: Struktur/Objekte; Orange: Potentielle Hindernisse; Rot: Hindernis im Fahrschlauch

Die folgenden Abschnitte beschreiben diesen Prozess und gehen auf Eigenschaften und Besonderheiten der Verfahren ein.

7.1 Vorverarbeitung

Die Vorverarbeitung nimmt einen hohen Stellenwert bei der Analyse von Punktwolken ein. Das aus dem SfM- und DSO-Verfahren stammende photogrammetrische Modell besteht oft aus einer sehr hohen Anzahl an 3D-Punkten und muss für die Identifizierung von Hindernissen zunächst bearbeitet werden. Trotz der Optimierungen in den vorangehenden Schritten der Kamerakalibrierung, Odometrieberechnung und den Rekonstruktionsverfahren, ist die Punktwolke mit Rauschen und Ausreißern behaftet, die im Hinblick auf eine möglichst eindeutige und genaue Hindernisdetektion entfernt werden müssen. Des Weiteren sind die 3D-Punkte in der Regel ungeordnet abgespeichert, was eine Auswertung sehr zeitaufwendig und ressourcenintensiv macht. Zur Vereinfachung der Komplexität können räumlich nah zusammenliegende Punkte in Zellen zusammengefasst werden, was eine Weiterverarbeitung beschleunigt.

7.1.1 Rauschen und Ausreißer

Schon während des SfM-Verfahrens wird eine Verbesserung der Punktqualität angestrebt, sodass bereits qualitativ schlechte Punkte gefiltert und Punktpositionen durch eine Bündelausgleichung optimiert werden können. Dennoch können während der Bildanalyse und 3D-Rekonstruktion Artefakte auftreten, die sich zu diesem Zeitpunkt nur bedingt verhindern lassen. Ursache für solche Artefakte können bspw. dynamische Objekte sein, die sich während der Bildaufnahme mit einer ähnlichen Relativgeschwindigkeit zum Fahrzeug bewegen wie statische Objekte. Die erkannten Merkmale auf diesen Objekten werden nicht zwangsweise durch die Analyse des optischen Flusses eliminiert und somit im weiteren Verlauf trianguliert. Typische Fehlerquellen sind zudem Reflexionen, welche je nach Objektfläche stärker auftreten können. Kritisch ist hier vor allem der Fahrergrund, bei dem durch Fahrrinnen und andere Vertiefungen ein Wasserabfluss behindert wird. Wassertropfen oder Verschmutzungen auf der Frontscheibe können nicht nur die



Abbildung 7-2 Fehler in der 3D-Punktvolke. Grün: Ausreißer bzw. Detektionsfehler; Blau: isolierte Cluster
Rot: nicht isoliertes Rauschen

Merkmalsdetektion beeinflussen, sie reduzieren auch den sichtbaren Bereich und beeinflussen somit die Schätzung der Kamerabewegung und die Triangulation von 3D-Punkten. Damit die Hindernishöhe auch in schwierigen Situationen robust berechnet werden kann, erfolgt die Eliminierung von fehlerhaften Punkten. Die Fehler in der 3D-Punktvolke werden, angelehnt an [130], in drei verschiedene Kategorien eingeteilt, welche in Abbildung 7-2 beispielhaft an der 3D-Punktvolke einer Brückenunterführung gezeigt werden.

Ausreißer

Diese Fehlerklasse ist vermeintlich am einfachsten zu detektieren, da es sich um vereinzelte isolierte 3D-Punkte handelt. Ursache eines solchen 3D-Punkts ist häufig die Rekonstruktion von falschen Punktkorrespondenzen. Auch die Rekonstruktion von geometrisch korrespondierenden Bildmerkmalen, die ausschließlich über zwei Bilder verfolgt werden, kann eine Ursache sein. Eine Vielzahl an Ausreißern kann jedoch durch die Angabe einer Mindesttracklänge von $n \geq 3$ bereits frühzeitig vermieden werden. Die charakteristische Fahrzeugbewegung in Richtung der Kamerahauptachse und die daraus resultierenden kleinen Schnittwinkel fördern die Erzeugung von Ausreißern. Durch Festlegung eines Mindestschnittwinkels wird auch dieser Effekt bereits während der Rekonstruktion minimiert. Auch die Reduzierung des maximalen Rückprojektionsfehlers während der Optimierung reduziert Ausreißer im photogrammetrischen Modell auf ein Mindestmaß.

Da das rekonstruierte Modell jedoch auch nach Optimierung der Rekonstruktionspipeline nicht gesamtheitlich ausreißerfrei vorliegt, wird mithilfe einer raumunterteilenden Struktur eine Vielzahl an Ausreißern eliminiert. Die in Abschnitt 7.1.2 vorgestellten Algorithmen dienen dabei der Strukturierung der Punktvolke sowie Vereinfachung und Beschleunigung der Punktvolkenverarbeitung. Durch Angabe einer Mindestanzahl an 3D-Punkten pro Zelle können Ausreißer direkt bei der Unterteilung eliminiert werden. Im vorliegenden Fall werden für das indirekte und direkte Verfahren 2°Punkte pro Zelle gefordert, sofern die Zellgröße als Würfel mit einer Kantenlänge von 0.25 m angenommen wird. Die Parametrierung erfolgt unter Berücksichtigung der Ersterkennung für Hindernisse. Während eine niedrige Punktzahl pro Zelle die Erstdetektion fördert, werden allerdings auch nah beieinanderliegende Ausreißer zunächst in den Zellstrukturen berücksichtigt. Diese werden jedoch anschließend während der Eliminierung der isolierten Cluster entfernt.

Isolierte Cluster

Eine in der Literatur nur begrenzt behandelte Fehlerkategorie sind isolierte Cluster. Sie sind beispielweise auf Anomalien wie Reflektionen und bewegte Objekte zurückzuführen. Da diese häufig im Straßenverkehr vorkommen, ist diese Kategorie als kritisch zu bewerten. Isolierte Cluster können die Berechnung des potentiellen Hindernisses beeinflussen und Fehldetektionen hervorrufen. Reflektionen und bewegte Objekte können jedoch schon während des Trackings identifiziert werden. Die Bewegungsmuster der einzelnen Merkmale unterscheiden sich je nach Entfernung stark von statischen Objekten. Des Weiteren können bestimmte Bildregionen, die bewegliche Objekte enthalten, mittels neuronaler Netze segmentiert und für die Merkmalsdetektion anschließend maskiert werden.



Abbildung 7-3 Isoliertes Cluster aufgrund des bewegten Objekts. Links: Originalbild mit Fokus auf das Fahrzeug auf der gegenüberliegenden Fahrspur; Rechts: Rekonstruktion von Fahrzeugpunkten auf der eigenen Fahrspur

Abbildung 7-3 zeigt beispielhaft den Einfluss rekonstruierter Tracks von bewegten Objekten. Ohne vorherige Entfernung dieser, werden 3D-Punkte häufig im falschen Abstand zur Kamera trianguliert, sodass das Fahrzeug, welches sich, wie links zu sehen, auf der gegenüberliegenden Spur befindet auf die eigene Fahrbahn geschätzt wird (rechts im Bild). Treten solche Fehler in Höhe des Fahrzeugdachs auf, so kommt es zwangsweise zu Fehldetektionen. Die Eliminierung der isolierten Cluster wird innerhalb der Punktwolke durch eine lokale Dichteberechnung durchgeführt. Über die Mindestanzahl und Abstand der Zellen zueinander ist diese parametrierbar. Die Parametrierung der Dichteberechnung erfolgt anhand von Sequenzen mit vergleichsweise kleinen bzw. isolierten Hindernissen, wie Verkehrsschilder oder Warnbaken, die oberhalb der Straße angebracht sind.

Nicht isoliertes Rauschen

Das nicht isolierte Rauschen ist im vorliegenden Fall die kritischste aller Fehlerkategorien. Hervorgerufen durch Ungenauigkeiten in der Punktdetektion und -verfolgung, kleinen Schnittwinkeln und fehlerbehafteten Kameraorientierungen, ist das Rauschen ein primärer Einflussfaktor für die korrekte Bestimmung der Hindernishöhe. In der Literatur werden verschiedene Ansätze zur Reduzierung des Rauschens innerhalb von Punktwolken vorgestellt. Neben einer Flächenberechnung und dem anschließend schwellwertbasierten Entfernen von Punkten mit einem zu großen euklidischen Abstand [130] wird ebenso eine Neuberechnung der Punktwolke und Verschieben der Punkte an deren wahrscheinlichste Position basierend auf einer flächenbasierten Wahrscheinlichkeitsfunktion vorgestellt [131]. Beide Verfahren sind zeitintensiv und anwendbar, sofern eine dichte Punktwolke zur Verfügung steht, die sich zur lokalen Flächenberechnung eignet. Im vorliegenden Fall müssen je nach Situation und Punktdichte die Schwellwerte für die Eliminierung angepasst werden. Wird mit festen Schwellwerten gearbeitet, so werden bspw. in

dichten Bereichen der Punktwolke das Rauschen unterdrückt, während in dünn besetzten Bereichen vermehrt Punkte gelöscht werden. Die Analyse zeigt, dass die Reduktion des Rauschens bei schwer erkennbaren Hindernissen die Erstreduktion senkt, sodass im Sinne der möglichst frühen Warnung auf eine Eliminierung der Punkte verzichtet wird. Alternativ kann über qualitativen Eigenschaften der 3D-Punkte eine Wichtung während der Höhenberechnung erfolgen, sodass Punkte, die eine höhere Tracklänge oder bessere Schnittgeometrie aufweisen, verstärkt berücksichtigt werden.

7.1.2 Raumunterteilende Datenstrukturen

Die 3D-Punkte liegen auch noch nach dem Entfernen von Ausreißern und Rauschen ungeordnet als Liste vor. Sie besitzen weder Informationen über direkte Nachbarn noch über zusammenhängende Objekte, sodass für eine spätere Auswertung alle Punktabstände miteinander verglichen werden müssten. Um die Identifizierung von möglichen Hindernissen zu beschleunigen, werden hier mögliche Datenstrukturen vorgestellt, die Punkte räumlich unterteilen. So wird nicht nur der Zugriff auf einen 3D-Punkt beschleunigt, sondern auch die Möglichkeit geboten die Punkte anhand ihrer Ausrichtung in verschiedene Objektklassen zu gliedern.

Die vorgestellten Datenstrukturen werden hinsichtlich Speicheranforderungen, Berechnungszeit und den Umgang mit neuen Daten bewertet. Alle Strukturen basieren auf der vorigen Definition der maximalen 3D-Raumgröße. Diese wird durch einen Quader mit folgenden Ausdehnungen beschrieben:

$$\begin{aligned} X &= 0 \text{ m} : 40 \text{ m} \\ Y &= \min(FS) - 2 \text{ m} : \max(FS) + 2 \text{ m} \\ Z &= -5 \text{ m} : 5 \text{ m} \end{aligned}$$

In Y -Richtung werden die Grenzwerte der Fahrschlauchprädiktion (FS) berücksichtigt. 3D-Punkte außerhalb dieser definierten Grenzen werden entfernt. Eine weitere Option ist die Begrenzung der maximalen Ausdehnung auf die tatsächliche Größe des geschätzten Fahrschlauchs. Dadurch kann der Speicherbedarf und die Rechenzeit gesenkt werden. Da die Fahrschlauchprädiktion in Kurven und auf große Distanzen jedoch fehleranfällig ist, findet diese Variante hier keine Anwendung.

Gitter

Eine Vereinfachung der Komplexität kann zunächst mit einem K -dimensionalen Gitter (hier: $K = 3$) mit festgelegter Zellgröße erreicht werden. Die Zellgröße kann dabei unabhängig für jede Dimension definiert werden. Ausgehend vom Ursprung des USK wird die Struktur in alle Richtungen aufgebaut und die 3D-Punkte einzelnen Zellen zugeordnet. Die Zellen können unterschiedlich dicht besetzt sein, wobei eine Zelle nur dann genutzt wird, wenn eine Mindestanzahl an Punkten überschritten wird. Innerhalb der Zellen wird auf die 3D-Punkte mit Pointern verwiesen. Sollen weitere Punkte aus einer zeitlich späteren Rekonstruktion in das Gitter eingetragen werden, so wird die Zelle bestimmt und ein neuer Pointer angelegt. Das Löschen eines Punktes erfolgt dementsprechend über das Löschen des Pointers. Der Speicherbedarf einer solchen Struktur ist hoch, da leere Zellen nicht gelöscht werden.

Kd-Baum

Die Punktmenge wird hierbei durch eine achsenorientierte Ebene geteilt, wobei die Ebenenrichtung durch die Dimension mit der größten Zellausdehnung bestimmt wird. Die Teilung erfolgt unter Nutzung des Medians der Punktpositionen bis hin zum Erreichen eines Grenzwerts für die Punktzahl einer Zelle. Kd-Bäume eignen sich für die schnelle Identifizierung von Punkten innerhalb einer großen Datenmenge. Problematisch ist jedoch der Umgang mit variablen bzw.

dynamischen Daten. Beim Hinzufügen von neuen 3D-Punkten im Laufe der Sequenz sollte der Baum, um eine optimale Verteilung anzustreben, neu aufgebaut werden. Durch das Filtern von gleichartigen Punkten, Anordnung der oberen Ebenen als Gitter und die Festlegung einer Mindestanzahl an Punkten pro Blatt kann der Baum jedoch balanciert werden [132].

Octree

Die Unterteilung des Punktraumes erfolgt beim Octree-Verfahren in acht gleich große Quader. Diese Unterteilung wird rekursiv fortgesetzt bis ein festgelegter Grenzwert für die Punktdichte unterschritten wird. Wesentliche Vorteile dieser Struktur sind die variable Zellengröße, die sich an der Punktdichte orientiert und die Eigenschaft, dass nicht befüllte Zellen gelöscht werden können. Die Reaktion auf neu hinzukommende Daten ist jedoch vergleichsweise komplex. Beim Hinzufügen von neuen Punkten muss zunächst die übergeordnete Zelle detektiert werden. Danach erfolgt die Überprüfung, ob der vorgegebene Grenzwert noch eingehalten wird. Wird die Punktdichte überschritten, so erfolgt eine erneute Teilung in acht kleinere Zellen. Im umgekehrten Fall des Löschens von Punkten wird dieses Kriterium ebenfalls überprüft und ggf. acht kleinere Zellen zu einer Größeren zusammengefasst.

Im vorliegenden Fall werden durch die Rekonstruktion nur wenige tausend Punkte erzeugt, deren Position sich in jedem Zeitschritt ändert. Die variablen Zellgrößen, die beim Kd-Baum und Octree entstehen, sind bei den verschiedenen und schnell ändernden Szenen nicht zielführend, da stets eine Adaption der Datenstruktur vorgenommen werden muss. Die Wahl fällt in diesem Fall auf die Gitterstruktur. Sie benötigt den größten Speicheraufwand, jedoch ist das Verfolgen der Punkte über verschiedene Zeitpunkte und Positionen unkompliziert und kann über die Lokalisierungsdaten der Eigenbewegungsschätzung gestützt werden. In Abbildung 7-4 wird die Zufahrt auf freihängende Schilder zu ausgewählten Zeitpunkten dargestellt. Abbildung 7-5 wiederum zeigt beispielhaft die Unterteilung der originalen 3D-Punktwolke dieser Szene. Die minimale Punktzahl pro Zelle liegt hier bei zwei. Es ist zu erkennen, dass Fahrbahn und Umgebungsstruktur weiterhin repräsentiert werden. Gleichzeitig erfolgt die Eliminierung von Ausreißern und isolierten Clustern.



Abbildung 7-4 Zufahrt auf freihängende Warnschilder zu verschiedenen Zeitpunkten. Links: $I=30$; Mitte: $I=60$; Rechts: $I=90$

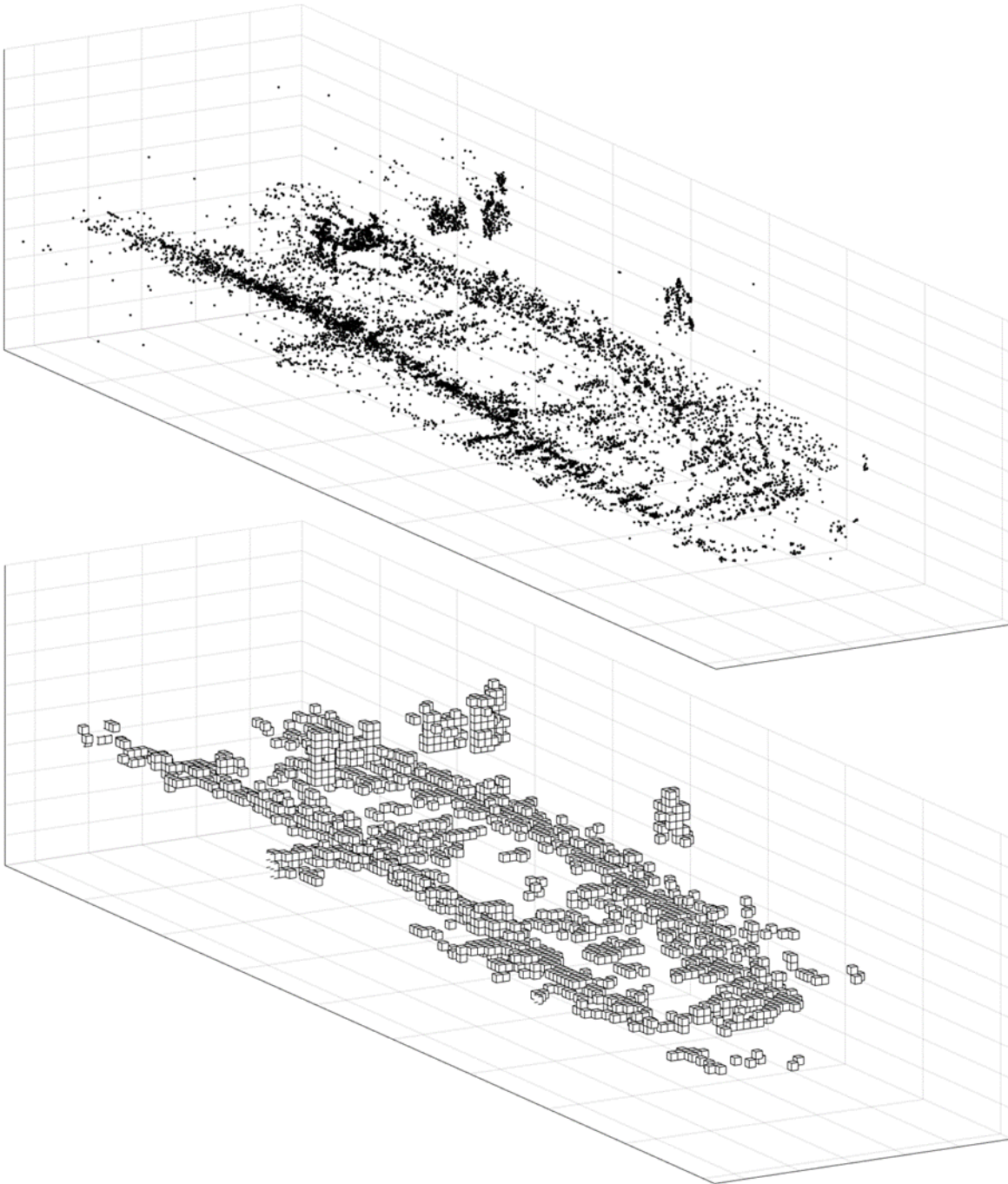


Abbildung 7-5 Unterteilung und Filterung von 3D-Punkten. Oben: Originale 3D-Punktwolke (SfM); Unten: Raumunterteilende Zellstruktur mit Würfeln der Kantenlänge von 0.25 m

7.2 Fahruntergrunddetektion

Die exakte Bestimmung der Durchfahrtshöhe bedingt das Referenzieren des Hindernisses auf den darunterliegenden Fahruntergrund. Zur Repräsentation des Fahruntergrunds, bzw. des befahrbaren Bereichs, wurde bereits in den 1980er Jahren bei der Exploration der Umgebung mit Robotern ein Belegungsgitter eingeführt [133]. In aktuellen Anwendungen sind diese Belegungsgitter stets die gängigste Methode zur Darstellung der befahrbaren Fläche [134, 135]. Nachteil der 2D-Gitterrepräsentation ist die Vernachlässigung des vertikalen Verlaufs der Fahrbahn.

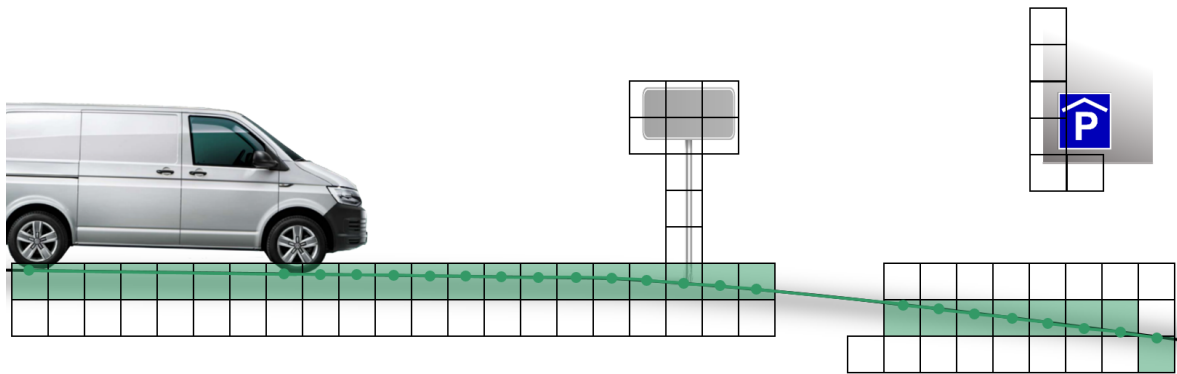


Abbildung 7-6 Berechnung der Fahrergrundhöhe und Klassifizierung der Bodenzellen (grün).

Eine vergleichsweise einfache Schätzung des Fahrergrunds als Ebene, wie sie in [136] unter Hinzunahme der Kamerahöhe berechnet wird, ist im vorliegenden Anwendungsfall nicht zielführend, da Kuppen und Wannen die Höhenberechnung des Hindernisses innerhalb der 3D-Punktwolke beeinflussen. Insofern wird hier ein Verfahren zur Modellierung des Höhenverlaufs des Fahrergrundes vorgestellt, welches es ermöglicht innerhalb der Zellstruktur Bodenzellen zu klassifizieren (s. Abbildung 7-6). Die Rekonstruktion von 3D-Punkten auf der Fahrhahnoberfläche kann je nach Struktur und Entfernung sehr dünn ausfallen, sodass der Algorithmus diese Freibereiche ebenfalls im Höhenplan interpolieren muss.

7.2.1 Grundlagen

7.2.1.1 Höhenplan

Der vertikale Verlauf einer Straße wird im sogenannten Höhenplan bestimmt. Zum Zeitpunkt der Straßenplanung wird unter Beachtung von baulichen, geologischen und topographischen Randbedingungen eine Biegelinie in das betreffende Gelände gelegt. Unter Einhaltung von Grenzwerten werden dann die Längsneigungen, Kuppen- und Wannenausrundungen bestimmt. Je nach Entwurfsklasse, Autobahn oder Landstraße, befinden sich die Längsneigungen der Fahrhah bei 1 - 8 %. Querneigungen dürfen im Bereich von 2.5 - 10 % liegen, um einerseits Wasserabfluss zu gewährleisten, jedoch im Winter ein Abrutschen aufgrund von Glätte zu verhindern [137]. Diese Grenzwerte fließen grundlegend in die Fahrergrundrekonstruktion ein, um zu verhindern, dass fehlerhafte Modelle erzeugt werden. Da die Funktion jedoch auch abseits von normiert angelegten Straßen genutzt werden soll, bspw. in Privatgeländen, sind die Grenzwerte nicht immer anwendbar. Bei der Bodenrekonstruktion werden demnach doppelt so große Grenzwerte zugelassen.

7.2.1.2 Normalenvektoren

Zur Modellierung des Fahrergrunds wird die durch das jeweilige Rekonstruktionsverfahren erzeugte 3D-Punktwolke herangezogen. Ziel ist es die 3D-Punkte herauszufiltern, die den befahrbaren Bereich darstellen. Um die Kandidaten zu ermitteln wird über jede Zelle des erzeugten Gitters iteriert und unter Berücksichtigung der 3D-Punkte der Nachbarzelle der jeweilige Normalenvektor \mathbf{n} berechnet. Die Nachbarn werden über einen definierten Kugelradius identifiziert und dienen der Bestimmung einer lokalen Fläche (Gl. (7-1)), die direkte Informationen über den Normalenvektor liefert. Sollten im ersten Durchlauf keine Nachbarn gefunden werden, so wird der Radius mit dem Faktor $2^{1/3}$ multipliziert, um das Suchvolumen zu vergrößern.

$$ax + by + cz + d = 0 \rightarrow \mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \text{Gl. (7-1)}$$

Die Normalenvektoren werden anschließend auf Basis ihrer Ausrichtung (z.B. $(c \gg a) \cap (c \gg b)$) und Winkel zueinander gefiltert. Ähnliche Eigenschaften wie der Fahruntergrund weisen jedoch auch weitere ebene Elemente, wie Fahrzeugdächer oder Tunneldecken, auf. Während der Modellierung des Fahruntergrundes müssen diese Elemente demzufolge eliminiert werden. Beispielhaft ist dies in Abbildung 7-7 aufgezeigt.

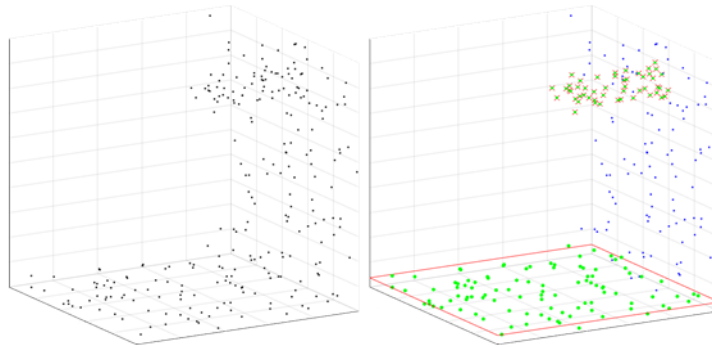


Abbildung 7-7 Eliminierung falscher Fahruntergrundkandidaten.

Es wird die originale und bereits klassifizierte Punktwolke einer Szene dargestellt, in der der modellierte Fahruntergrund, auf der rechten Seite die Wand und oben die Decke einer Tunnelleinfahrt zu sehen sind. Die potentiellen Bodenkandidaten werden in Grün markiert, die zur Umgebungsstruktur gehörenden 3D-Punkte in Blau. Die geschätzte Fahrbahnhöhe, hier durch das rote Rechteck dargestellt, erlaubt die Trennung der richtigen und falschen Kandidaten, welche mit einem roten Kreuz markiert werden.

7.2.2 Fahruntergrundmodellierung

Da der Fahruntergrund nicht als ebene Fläche angesehen werden kann, muss die Bestimmung der Untergrundhöhe lokal ablaufen. Über die Verbauhöhe der Kamera können dabei zunächst die 3D-Punkte identifiziert werden, welche direkt vor dem Fahrzeug auf Höhe des Untergrunds bei $Z \approx 0$ liegen. Ausgehend von dieser Basishöhe werden die Nachbarzellen in Fahrtrichtung exploriert und durch ein geeignetes Verfahren modelliert. Das Modell soll auch in Regionen mit wenigen 3D-Punkten eine adäquate Repräsentation des Untergrundes ermöglichen. Es bieten sich dafür unterschiedliche Realisierungsmöglichkeiten an, auf die kurz eingegangen werden soll.

7.2.2.1 Delaunay Triangulation

Ein häufig angewendetes Verfahren ist die Triangulierung der Punkte zu einer Netzstruktur. Beispielhaft ist dabei die Delaunay Triangulation zu erwähnen, die nicht nur bei der Modellierung von realen 3D-Modellen zum Einsatz kommt, sondern auch sehr häufig im Rahmen der Entwicklung künstlicher 3D-Welten bei Computerspielen Einsatz findet. Charakteristisch ist die Approximation einer Oberfläche durch Dreiecke. Durch Vergrößerung der maximalen Kantenlänge können auch Untergrundregionen modelliert werden, in denen nur wenige Punkte rekonstruiert wurden.

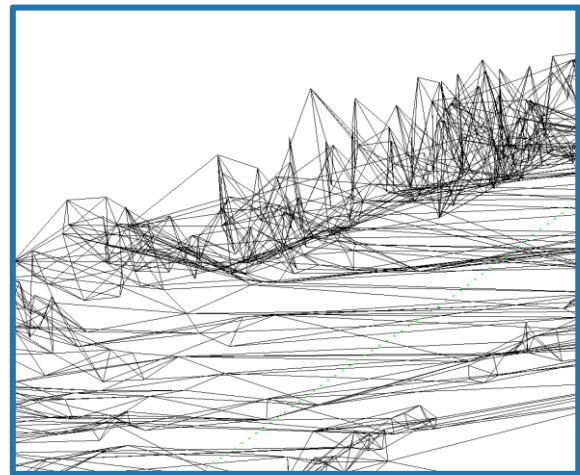
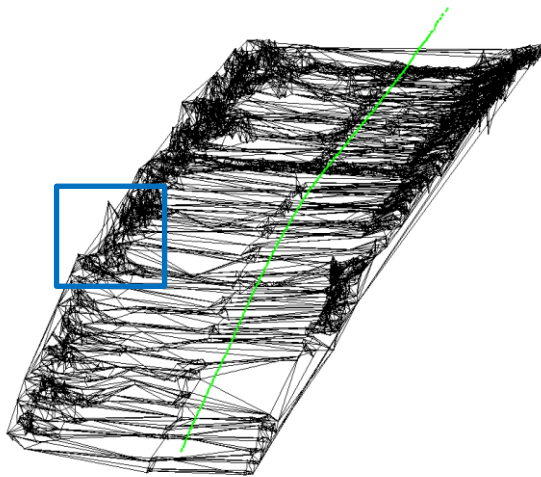
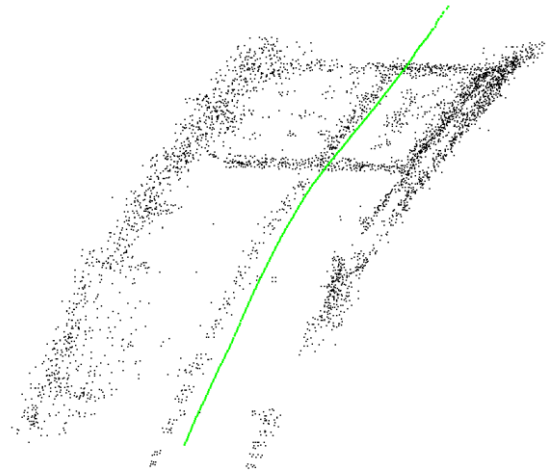


Abbildung 7-8 Delaunay Triangulation des Fahrergrund. Oben links: Aufnahme der Umgebung; Oben rechts: Identifizierte Kandidaten des Bodens und gefahrene Trajektorie; Unten links: Delaunay Triangulation; Unten rechts: Detailaufnahme des Bordsteins

Die in Abbildung 7-8 gezeigten extrahierten Fahrbahnpunkte werden mittels Delaunay Triangulation zu einem Netz verbunden, welches unten im Bild dargestellt wird. Es wird ersichtlich, dass Bordsteinkanten auch Bestandteil der Fahrergrundkandidaten sein können und somit zwangsweise in die Bodenschätzung mit eingehen. Wird keine zusätzliche Filterung angewendet, so folgt daraus eine ungenaue Modellierung der Fahrbahnrandbereiche. Um eine möglichst gesamtheitliche Rekonstruktion der Fahrbahn zu erhalten, wird während der Delaunay Triangulation die Erzeugung möglichst langer Kanten präferiert. Dies fördert jedoch die Einbeziehung von weit entfernten Punkten, wie etwa nicht entfernte Ausreißer. Beispielhaft ist dies in Abbildung 7-9 dargestellt. Da besonders in großen Entfernungen die Punktungenauigkeiten sehr hoch sind und Ausreißer auftreten können, wird innerhalb dieser Arbeit auf den Ansatz der Delaunay Triangulation verzichtet.

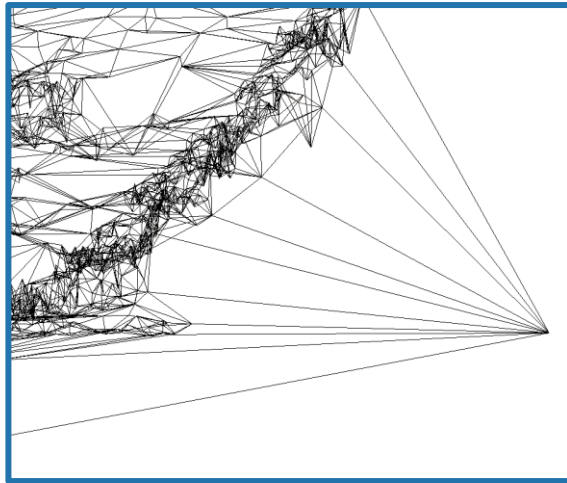
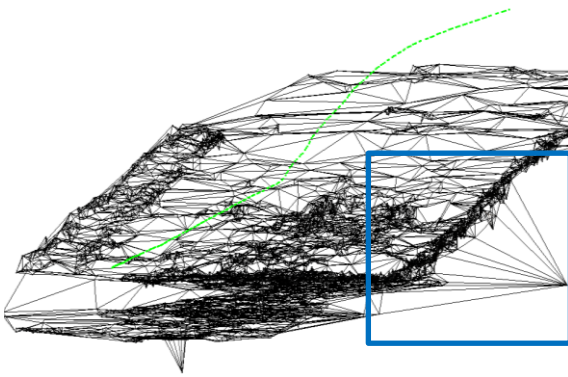
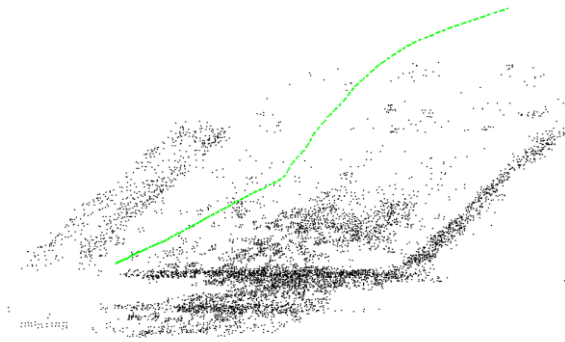


Abbildung 7-9 Delaunay Triangulation des Fahruntergrunds. Oben links: Aufnahme der Umgebung; Oben rechts: Identifizierte Kandidaten des Bodens und gefahrene Trajektorie; Unten links: Delaunay Triangulation; Unten rechts: Detailaufnahme des Ausreißers

7.2.2.2 B-Splines

Eine weitere Möglichkeit, die in [138] vorgestellt wurde, ist die Beschreibung der Fahrbahnoberfläche mittels B-Splines. Dabei wurde auf Basis der v -Disparität zweier Bilder einer Stereokamera das mathematische Modell der Oberfläche approximiert und mittels Kalman Filter stabilisiert. Vorteil der Darstellung über B-Splines ist die algebraische Abhängigkeit der Distanz zur Fahrbahnhöhe. Wird ein Hindernis in einer bestimmten Distanz wahrgenommen, so kann der Z -Wert der Fahrbahn durch Einsetzen in die Gleichung des Splines bestimmt und die relative Höhe zum Hindernis berechnet werden. Die Splines-Darstellung beruht auf der Basis von Stützpunkten, die in definierten Abständen entlang der Fahrzeuglängsachse innerhalb des USK verteilt werden. In unbefestigten Geländen, kann sich dies jedoch nachteilig auswirken, da auf starke Veränderungen nicht eingegangen wird. Eine geschwindigkeitsabhängige Positionierung der Stützpunkte oder das Einfügen zusätzlicher Punkte in Regionen großer Steigungsänderungen können diesen Effekt reduzieren. Sofern als Stützstellen die bereits vorgestellten Zellen genutzt werden, ergibt sich durch die der Modellierung nachgelagerten Berechnung des B-Splines ein zusätzlicher Zeitaufwand, wodurch diese Variante hier nicht zum Einsatz kommt.

7.2.2.3 Stückweise Ebenen

In der vorliegenden Arbeit wird die Methode der stückweise aneinandergesetzten Ebenen verwendet. Die Stützstellen der Ebenen sind abhängig von der Zellgröße und befinden sich jeweils in Zellmitte. Ausgehend von den 3D-Punkten direkt unter dem Fahrzeug, die die Aufstandspunkte der Reifen beinhalten, wird zunächst eine Basisebene bestimmt, die sich aus der kalibrierten Einbauhöhe der Kamera im Fahrzeug ableitet. Anschließend werden ausgehend von dieser Grundebene alle Zellen in X -Richtung exploriert, deren Normalenvektoren vornehmlich in Z -Richtung ausgerichtet sind. Dazu werden jeweils die 3D-Punkte innerhalb der benachbarten Zellen der Höhe nach aufgelistet und eine gemittelte Basishöhe des unteren Quartils berechnet. Dies soll den Effekt der gehäuft am Fahrbahnrand rekonstruierten 3D-Punkte kompensieren, welche durch Bordsteine oder Bepflanzung eine größere Höhe als die eigentliche Fahrbahn besitzen, jedoch während der Berechnung der Normalenvektoren nicht entfernt wurden. Die Höhen aller 3D-Punktkandidaten werden mit dieser Basishöhe verglichen und die Kandidaten entfernt, die eine Steigungsänderung der Fahrbahn von größer 20° hervorrufen würden. Die validen Punktkandidaten werden durch ihre inverse Distanz zur aktuellen Stützstelle gewichtet. Ein Punkt mit doppelter Distanz im Vergleich zu einem anderem, wird entsprechend nur halb so stark gewichtet. Um eine Übergewichtung zu vermeiden, werden die Gewichte über einem Grenzwert abgeschnitten. Dabei hat sich eine Mindestdistanz von 1 m (4 Zellen) als zielführend erwiesen. Anschließend wird über die validen und gewichteten 3D-Punkte die lokale Höhe des Zellverbunds bestimmt. Sollten Freibereiche innerhalb der Zellstruktur vorhanden sein, werden diese linear aus den Höhen der vorherigen und nachfolgenden Stützstellen im Anschluss interpoliert. Eine komplexere Darstellung der Bodenhöhe kann mithilfe der Exploration in X - und Y -Richtung geschehen. Die Untersuchung zeigt jedoch, dass nahe des prädierten Fahrschlauchs keine deutliche Verbesserung der Untergrundmodellierung erkennbar ist. Somit wird sich im Hinblick auf die Rechenzeit für die eindimensionale Höhenberechnung entschieden.

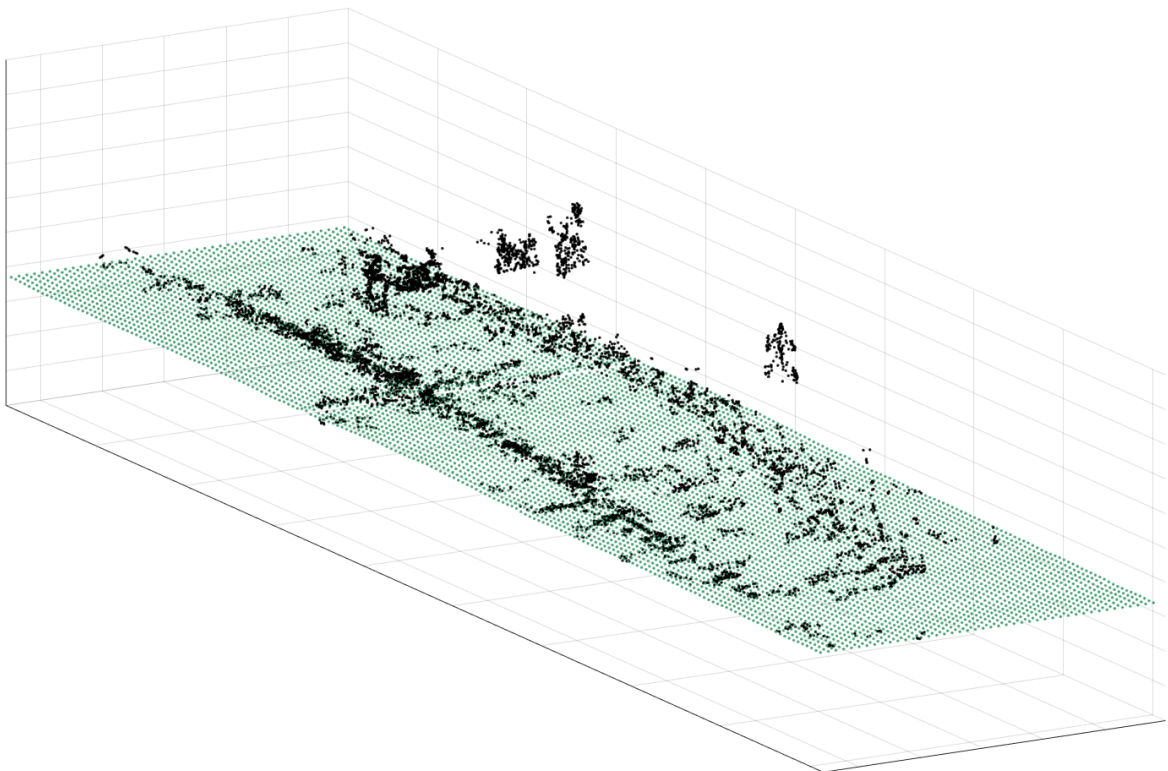


Abbildung 7-10 Extrahierter Fahrergrund. Schwarz: Gefilterte Punktwolke; Grün: Geschätzter Fahrergrund an jeder Stützstelle

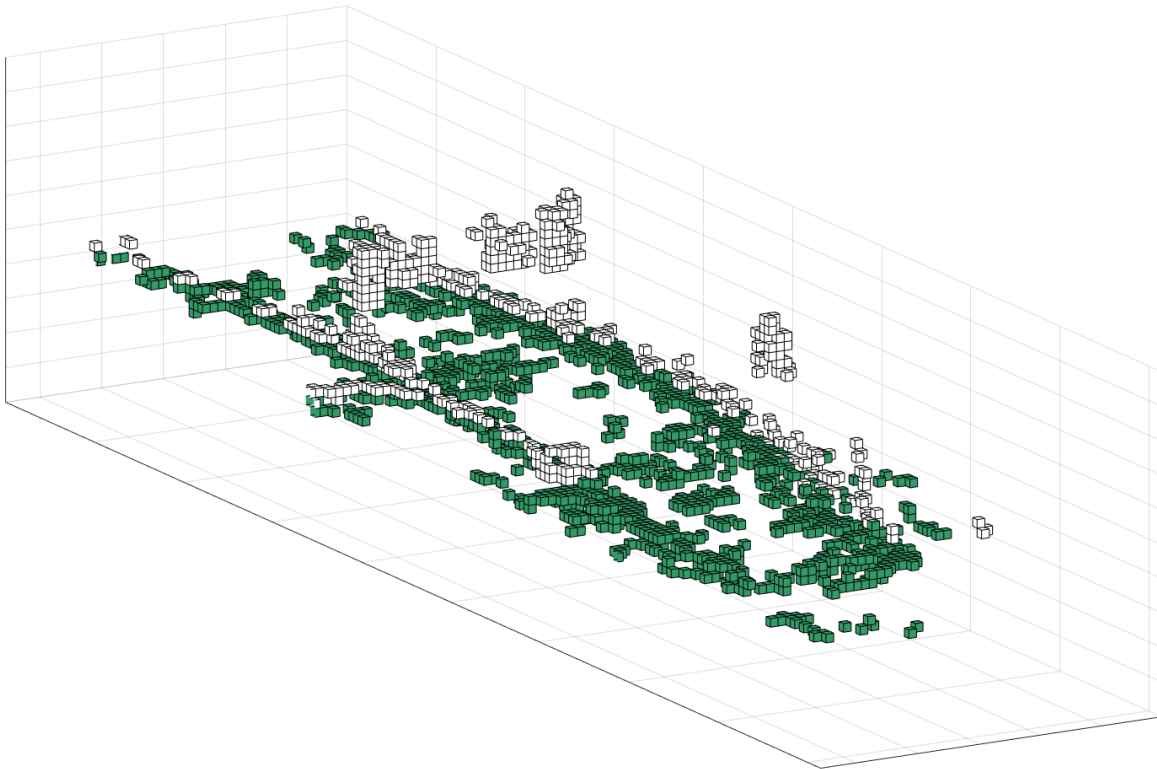


Abbildung 7-11 Klassifizierte Zellstruktur. Weiß: Unklassifiziert; Grün: Zellen des Fahrergrundes

In Abbildung 7-10 ist beispielhaft die Fahrergrundmodellierung für eine Sequenz gezeigt. Wie erwartet verläuft das Modell in diesem einfacheren Beispiel nahezu eben über die gesamte Breite der Punktwolke. Nach der Bestimmung des Höhenplans werden alle Zellen, die zum Fahrergrund gehören markiert (s. Abbildung 7-11). In Abbildung 7-12 werden weitere Sequenzen gezeigt, deren Punktwolke, Modellierung und Klassifizierung in Abbildung 7-13 dargestellt sind. Bei der Kurvenfahrt ist anzumerken, dass die maximale Prädiktion des Fahrschlauchs verkürzt werden sollte, da es sonst zu fehlerhaften Modellierungen kommt. Kritisch sind hierbei die seitlichen Begrenzungen, die anteilig in den Höhenplan eingehen. Zur Modellierung des Fahrergrundes der Auffahrt musste die maximal zulässige Steigung auf 45° erhöht werden, um den Höhenplan abzubilden zu können. Empfehlenswert ist eine geschwindigkeitsabhängige Beschränkung der maximalen Steigung, sodass auch in unbewegten Gelände eine Modellierung des Fahrergrundes stattfinden kann.



Abbildung 7-12 Ausgewählte Sequenzen zur Fahrbahnmodellierung. Links: Kurvenfahrt; Rechts: Auffahrt zu einem Parkplatz

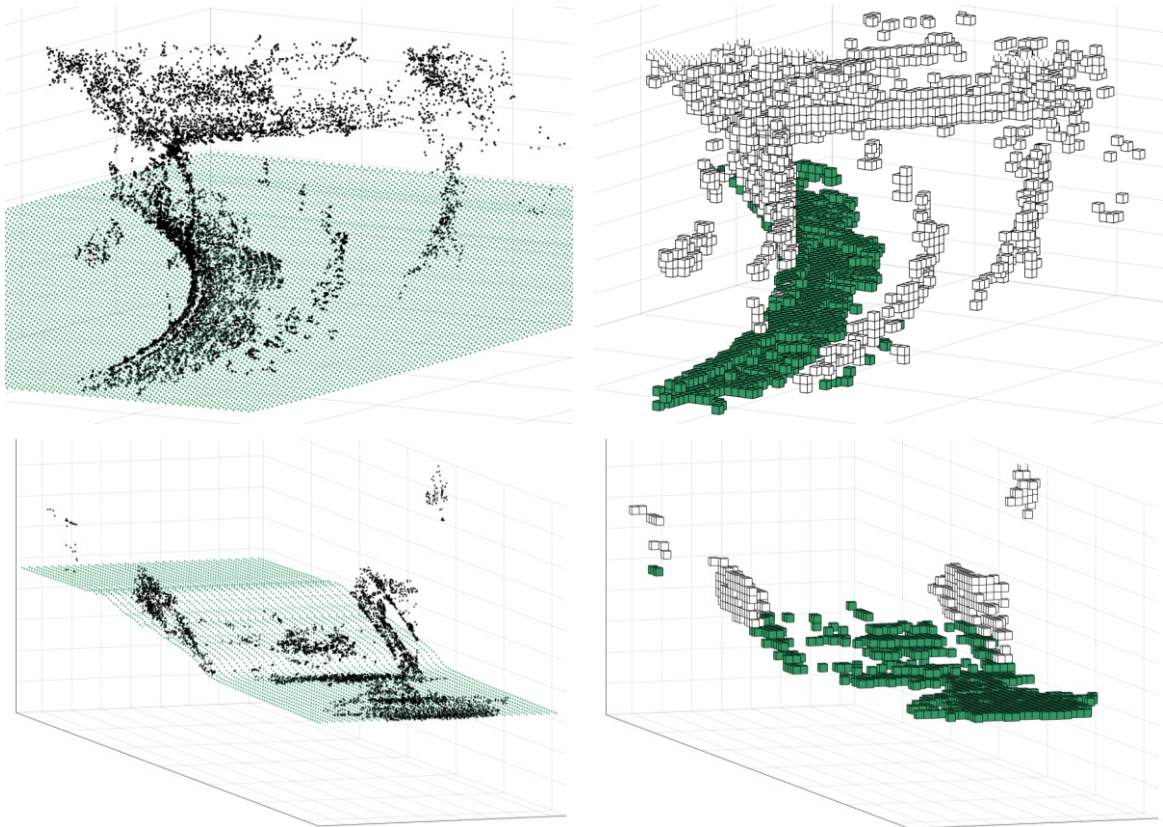


Abbildung 7-13 Fahrergrundmodellierung und Klassifizierung ausgewählter Sequenzen. Oben: Kurvenfahrt; Unten: Auffahrt zu einem Parkplatz

7.3 Struktur- und Hindernisdetektion

Nach der Modellierung der Fahrbahn und dem Extrahieren des Fahrergrunds werden darauf basierend die Objekte und Strukturen der Umgebung identifiziert. Eine Objektklassifikation wird im Rahmen der Bildverarbeitung häufig durch den Einsatz von neuronalen Netzen bereits direkt im Bild durchgeführt. Diese sind auf verschiedene Arten von Objekten trainiert und werden in Klassen unterteilt. Laut Anforderungen aus Abschnitt 1.2.1 sollen die hier berücksichtigten Hindernisse unabhängig von einer vorbestimmten Klasse sein, sodass ein allgemeinerer Ansatz gewählt wird, der auf dem bereits definierten 3D-Gitter beruht. Als Umgebungsstruktur werden in diesem Zusammenhang Zellen beschrieben, die einen direkten oder indirekten (durch ihre Nachbarn) Kontakt zur extrahierten Fahrbahnoberfläche besitzen. Potentielle Hindernisse sind die Zellen, die keinen Kontakt und einen Mindestabstand zur Fahrbahnoberfläche besitzen. Dieser ist hier auf 1.5 m festgelegt. Im Testfahrzeug liegt dies auf Höhe des Brustkorbs des Fahrers. Es wird davon ausgegangen, dass der Fahrer niedrigere Hindernisse eigenständig als nicht unterfahrbar wahrnimmt. In der Fahrzeugfunktion ist zudem die Beschränkung der maximalen Höhe über die Angabe der Fahrzeughöhe (inkl. Aufbau) möglich, die den Hindernisraum weiter einschränkt. In der vorliegenden Arbeit wird diese Höhe jedoch auf 5 m begrenzt, um auch Hindernisse in großen Höhen entsprechend bewerten zu können. Abbildung 7-14 zeigt die final klassifizierte Zellstruktur mit dem Fahrergrund, der Umgebungsstruktur, potentiellen Hindernissen und den Hinderniszellen, die innerhalb des prädizierten Fahrschlauchs liegen. Diese Zellen werden später zur Berechnung der Durchfahrtshöhe herangezogen. Die Wichtigkeit des prädizierten Fahrschlauchs zur Vermeidung von Fehldetektionen wird hierbei ersichtlich. Zum einen kann der genaue Auftreffpunkt am Hindernis bestimmt werden. Zum anderen werden damit etwaige Fehldetektionen vermieden.

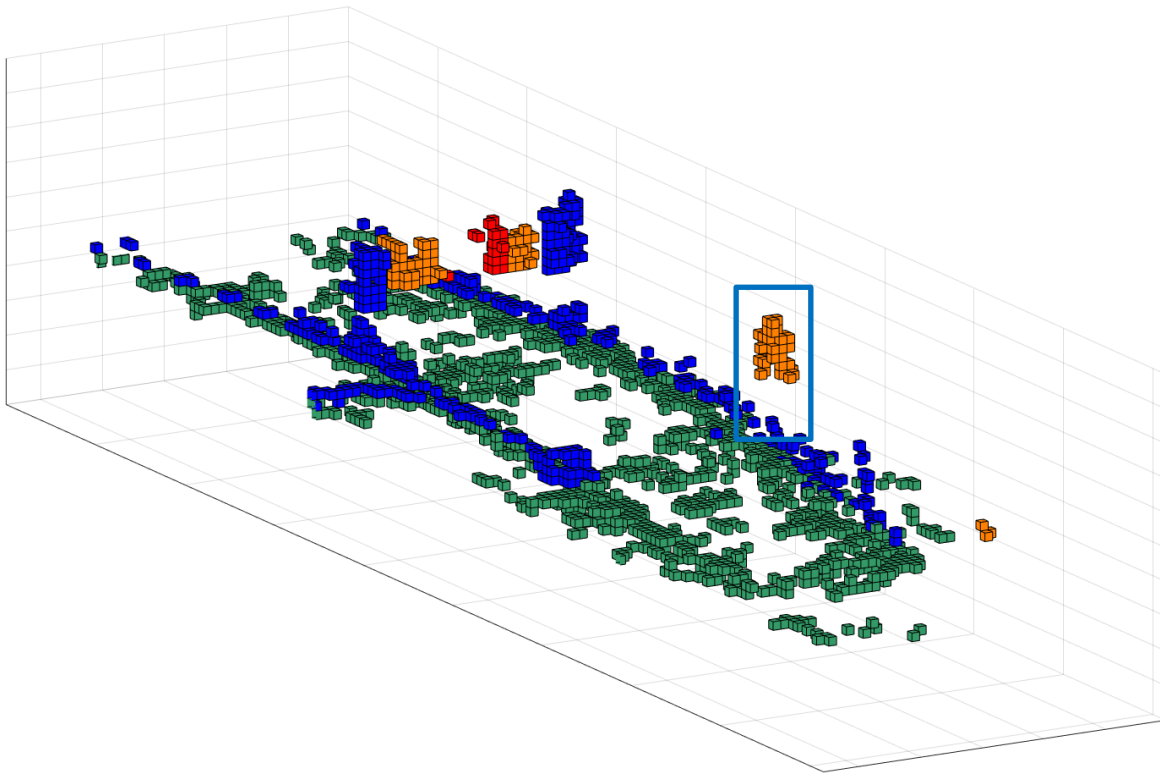


Abbildung 7-14 Klassifizierte Zellstruktur. Grün: Fahruntergrund; Blau: Struktur; Orange: Potentielles Hindernis; Rot: Identifizierte Hinderniszellen innerhalb des prädizierten Fahrschlauchs; Blauer Rahmen: Rekonstruiertes Verkehrsschild ohne Rohrpfosten

Beispielsweise wird in der oben gezeigten Abbildung 7-14 ein am rechten Fahrbahnrand befindliches Verkehrsschild (blauer Rahmen) als mögliches Hindernis fehlklassifiziert. Als problematisch stellt sich hierbei die Detektion und Rekonstruktion des Rohrpfostens heraus, sodass das Schild freihängend im 3D-Raum dargestellt wird. Durch den Fahrschlauch wird dieses vermeintliche Hindernis jedoch nicht in der Höhenberechnung berücksichtigt. Die klassifizierte 3D-Strukturen weiterer Sequenzen sind Abbildung 7-15 zu entnehmen. Die beiden oberen sind die bekannten Sequenzen mit Kurvenfahrt und Auffahrt dargestellt, welche bereits in Abschnitt 7.2.2.3 genutzt wurden. Die dritte Sequenz zeigt die Einfahrt in einen Tunnel. Die untere Sequenz zeigt die Zufahrt auf ein Parkhaus mit freihängendem Schild und Warnbake im Vordergrund sowie der tatsächlichen Einfahrt im Hintergrund. Beide Hindernisse werden vom Algorithmus detektiert.

Fahrergrund- und Hindernisdetektion

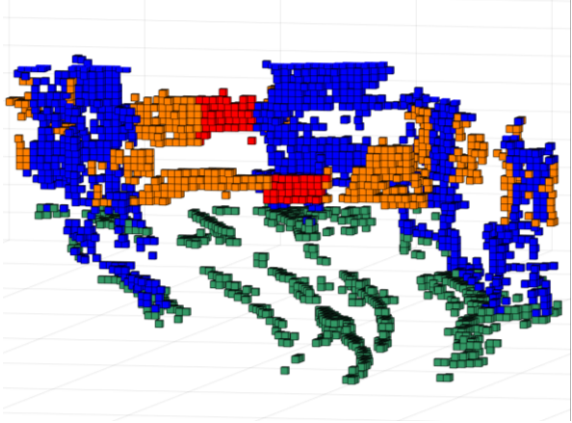
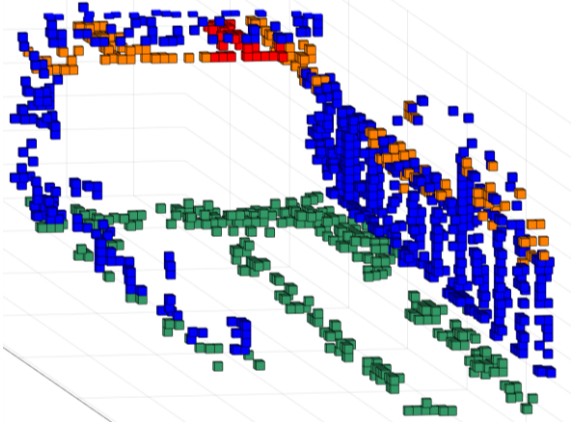
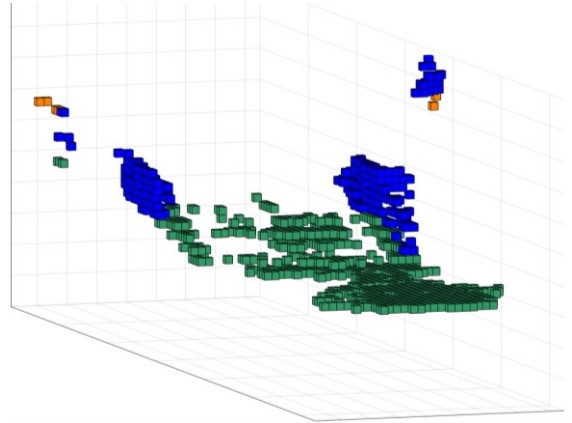
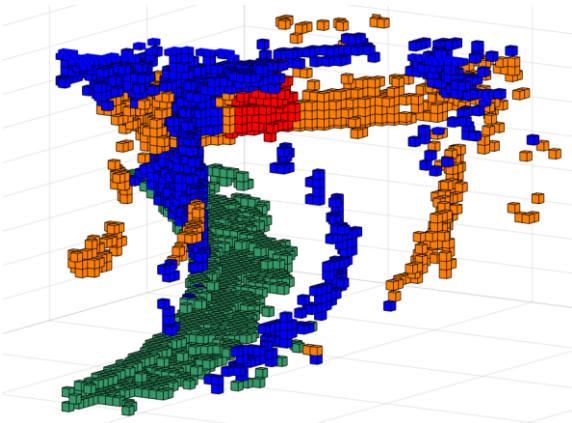


Abbildung 7-15 Klassifizierte 3D-Gitterstruktur ausgewählter Sequenzen.

8 Evaluierung

Innerhalb dieses Kapitels sollen die vorgestellten indirekten Verfahren (SfM, Line3D++) und das direkte Verfahren (DSO) auf Basis der 72 Sequenzen miteinander verglichen werden. Neben einem kurzen Vergleich zwischen Radodometrie und visueller Odometrie wird hier der Fokus auf die Bestimmung der äußeren Genauigkeit und der Zuverlässigkeit der Verfahren in den jeweiligen Verkehrssituationen gelegt.

8.1 Eigenbewegungsschätzung

Nach der Initialisierungsphase erfolgt der Übergang in die Hauptschleife, bei der die Kameraorientierungen über die vorhandenen 3D-Punkte und den räumlichen Rückwärtsschnitt berechnet werden (s. Abschnitt 5.2.45.2.4.2). Unter Berücksichtigung mehrerer Bilder und der Berechnung ihrer Position und Orientierung innerhalb der rekonstruierten 3D-Punktwolke, kann ebenfalls auf die Bewegung des Fahrzeugs geschlossen werden. Dieses als *visuelle Odometrie* bezeichnete Verfahren bestimmt im Vergleich zur Radodometrie die zusätzlichen Parameter der vertikalen Fahrzeugbewegung Z sowie die Rotationen um X und Y . Im folgenden Abschnitt werden die visuelle Odometrie aus dem indirekten und direkten Punktverfahren mit der Rad- und Referenzodometrie verglichen.

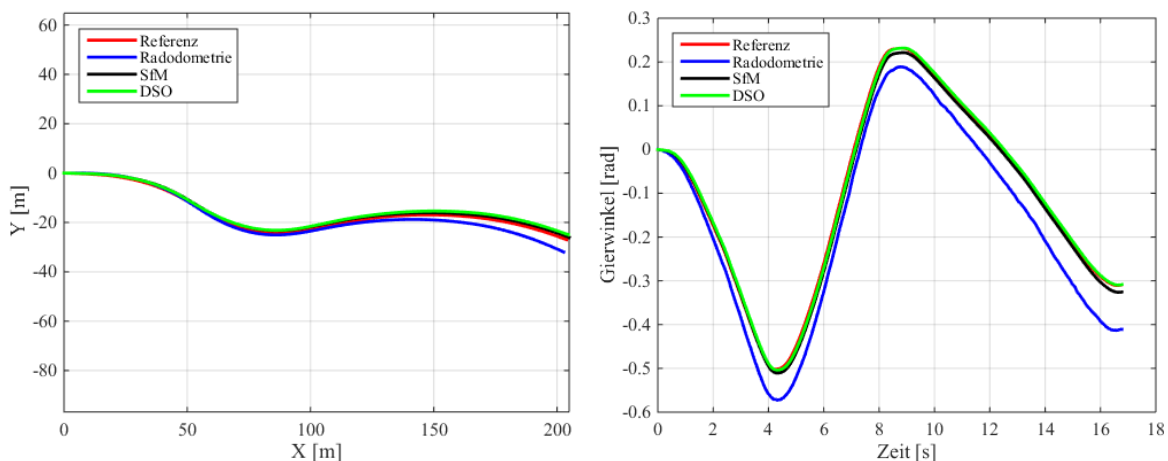


Abbildung 8-1 Vergleich der absoluten Position und Rotation von Rad- und visueller Odometrie (SfM, DSO) sowie der Referenz (IMU)

Wie Abbildung 8-1 zu entnehmen, kann durch beide kamerabasierte Verfahren der absolute Fehler im Vergleich zur Radodometrie in Translation und Rotation verringert werden. Zudem verfügen die visuellen Verfahren über einen geringeren Drift und eignen sich somit zur Bestimmung der absoluten Position in einem übergeordneten Weltkoordinatensystem. Es wird darauf verwiesen, dass die Ergebnisse der visuellen Verfahren maßgeblich von der Güte der Radodometrie abhängig sind, da diese zur initialen Berechnung des Skalierungsfaktors herangezogen wird. Eine präzise absolute Odometrie ermöglicht die Kartierung von Hindernissen in einem übergeordneten Koordinatensystem, dem primären Ziel bekannter SLAM-Ansätze. Für die hier vorliegende Funktion ist dies jedoch nicht notwendig, da die Positionen der Hindernisse im mitbewegten USK definiert werden.

In Abbildung 8-2 werden die relativen Fehler im Vergleich zur Referenz dargestellt. Dabei werden die Dimensionen getrennt aufgezeigt, um auf etwaige Stärken und Schwächen der Verfahren hinzuweisen. Die Nick- und Wankwinkel werden in diesem Zusammenhang vernachlässigt, da diese nicht für die Funktion entscheidend und vernachlässigbar klein sind.

Evaluierung

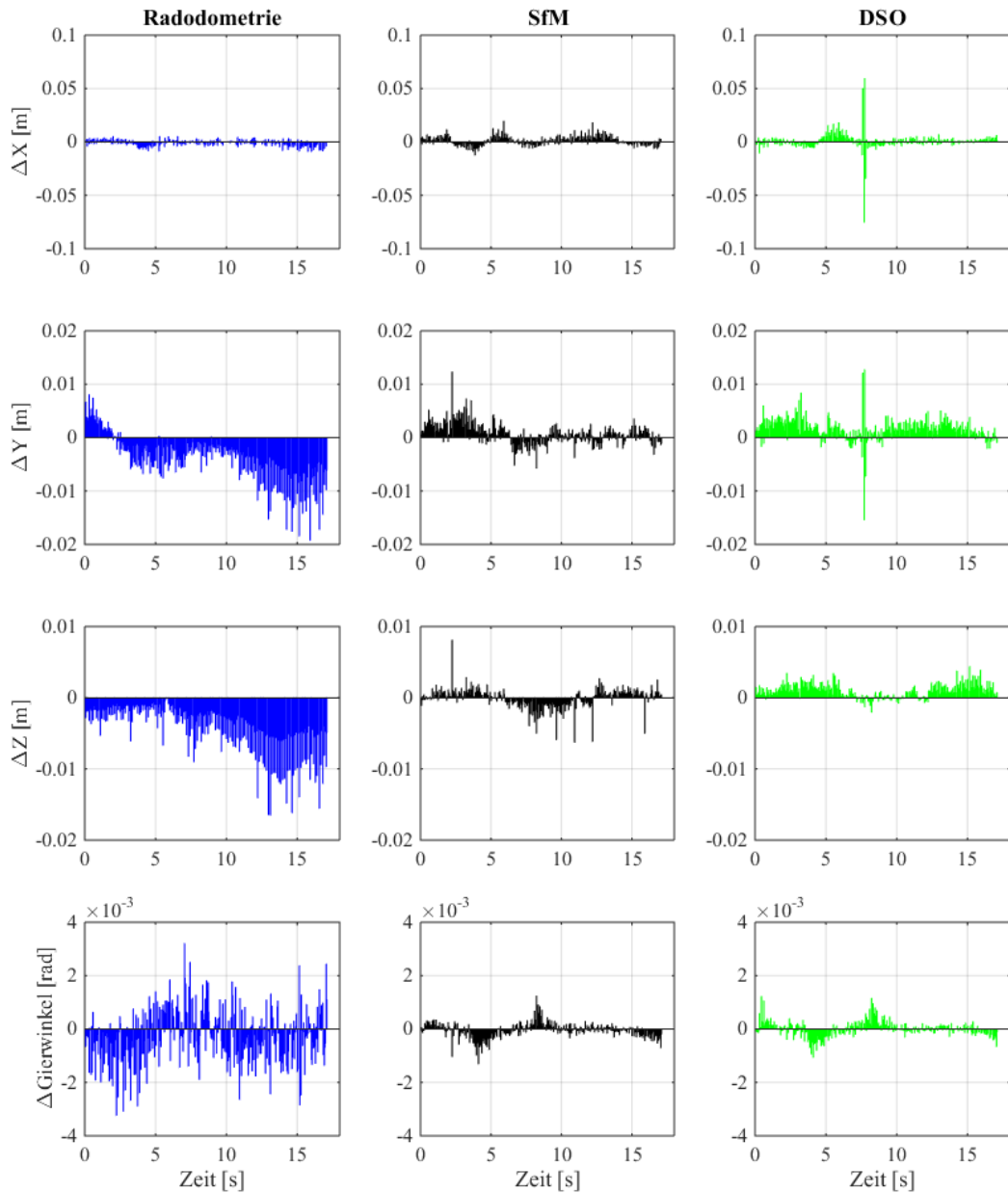


Abbildung 8-2 Relativer Fehler der einzelnen Dimensionen von Radodometrie und visueller Odometrie (SfM, DSO)

Die Radodometrie zeigt erwartungsgemäß einen Drift in Y - und Z -Richtung. Im Vergleich zu beiden visuellen Verfahren sind die Schwankungen der Gierwinkelberechnung stärker. Das indirekte SfM- und das direkte DSO-Verfahren reduzieren den Drift, können jedoch Ausreißer enthalten. In diesen Fällen wird empfohlen die berechnete Kamerapose zu verwerfen und diese somit nicht in die Optimierung der Kameraschätzung bzw. der 3D-Punkte einzubeziehen. Die Ausreißer treten vor allem bei wechselnden Bildinhalten in Kurven oder starken bei Beleuchtungsschwankungen auf.

8.2 Umgebungsrekonstruktion

Die Rekonstruktion der Umgebung wird während der Fahrt mittels indirektem und direktem Ansatz durchgeführt. Prinzipiell unterscheiden sich die beiden Verfahren in der Detektion der Bildmerkmale und der Bestimmung sowie Optimierung der 3D-Punkte. Darüber hinaus wird unter Annahme bekannter Kameraposen eine Linientriangulierung durchgeführt. Im nachfolgenden Abschnitt werden die drei Methoden miteinander verglichen. Dabei wird die Eignung der jeweiligen Verfahren zur Umgebungsrekonstruktion, im Speziellen die Rekonstruktion von Fahrbahn und Hindernissen, bewertet.

8.2.1 Punkte

Innerhalb dieses Abschnitts werden zunächst das indirekte und direkte Verfahren zur Punktreakonstruktion miteinander verglichen. Die Untersuchung erfolgt dabei anhand aller Sequenzen im 2D- und 3D-Raum. Dabei ist darauf zu verweisen, dass mit dem direkten DSO-Verfahren nur 51 der insgesamt 72 Sequenzen rekonstruiert werden konnten. Die Ursachen sollen hier erläutert werden.

8.2.1.1 Allgemein

Bildpunktdetektion und -verteilung

Die Detektion und Verfolgung von Bildpunkten ist entscheidend für die spätere Rekonstruktion einer dichten und genauen Punktwolke des Fahrumfelds. Wie bereits in Abschnitt 5.2.4.2 erläutert, ist die Detektion von Punkten an langen Kanten mittels Shi-Tomasi problematisch, da die Auswertungsfunktion diese Merkmale unterdrückt. Direkte Verfahren, die ohne diskrete Merkmale arbeiten, können Bildpunkte anhand ihrer Intensität und 3D-Repräsentation verfolgen. Dies ermöglicht auch Umrisse von Objekten als Ganzes zu rekonstruieren. Beispielfhaft ist dies in Abbildung 8-3 dargestellt, bei dem die verfolgten Bildpunkte der beiden Verfahren gegenübergestellt werden. Im Bild oben ist zu erkennen, dass mittels Shi-Tomasi kein Merkmal an der Unterkante der Brücke detektiert wurde. Das Hindernis wird in diesem Fall übersehen, bzw. eine falsche Höhe berechnet, die sich am Brückengeländer orientiert. Im Gegensatz dazu ist im rechten Teil des Bildes die Detektion durch DSO zu sehen, bei der die Unterkante zusammenhängend erkannt wird. In der unteren Sequenz wird eine hochstrukturierte Szene gezeigt, bei der beide Verfahren eine Vielzahl an Bildpunkten bereits aus größerer Distanz zum Hindernis (hier ca. 41 m) verfolgen. Es zeigt ebenso, dass DSO im Nahbereich und Regionen mit Bildunschärfe eine höhere Anzahl an Bildpunkten verfolgen und anschließend rekonstruieren kann. Allerdings trifft dies nicht auf alle Fahrsituationen zu. Während beim indirekten Verfahren vor allem die Bewegungsunschärfe problematisch ist, sind Beleuchtungsänderungen die Haupteinflussquelle für eine fehlerhafte Rekonstruktion beim direkten Verfahren. Mithilfe einer intelligenten Belichtungssteuerung oder durch Hinzunahme der Daten des serienmäßig verbauten Regen-Licht-Sensors kann jedoch die Bildhelligkeit möglichst konstant gehalten werden.

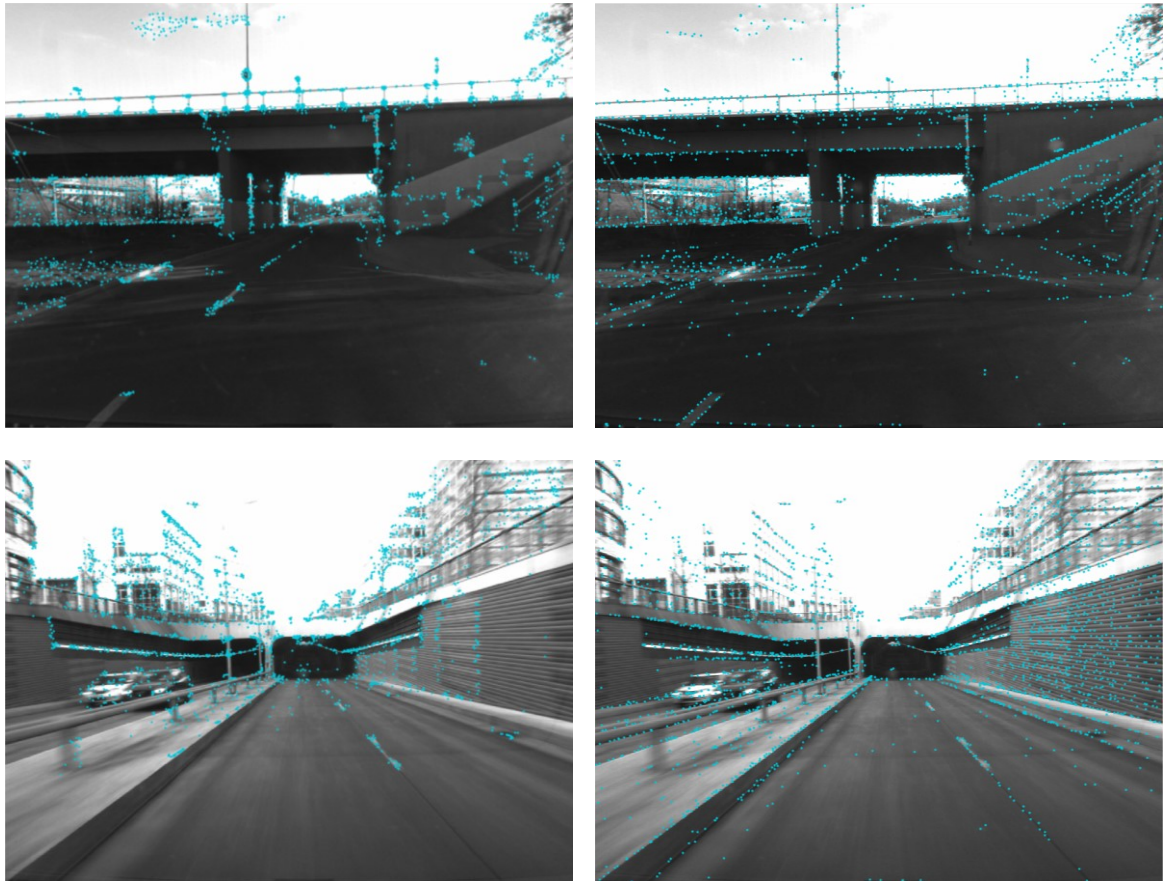


Abbildung 8-3 Vergleich der verfolgten Bildpunkte. Links: Indirektes SfM-Verfahren mit Shi-Tomasi Detektor; Rechts: Direktes DSO-Verfahren

Umwelteinflüsse

Im Verlauf der Auswertung zeigt sich, dass mit dem direkten Verfahren einige Sequenzen nicht initialisiert bzw. rekonstruiert werden können. Mit dem merkmalsbasierten Ansatz ist dies jedoch ausnahmslos möglich. Als problematisch stellen sich diese Sequenzen heraus, die während der Initialisierung eine starke Rotation aufweisen, bspw. bei Kurvenfahrten. Obwohl eine Initialisierung meist möglich ist, geht der Ansatz von einer primär lateralen Bewegung aus, wodurch der Kamerapfad fehlerhaft bestimmt und die Punktwolke verzerrt rekonstruiert wird. Dies ist auf die Bewegungsschätzung von DSO zurückzuführen, welche bei wenigen vorhandenen 3D-Punkten und großen Fehlern der geschätzten Kamerapose vermehrt die bisherige Kamerabewegung prädiziert, um die aktuelle Bewegung zu schätzen. Genauso wie bei DSO wird im Falle des SfM-Verfahrens per räumlichen Rückwärtsschnitt die Pose der Kamera ermittelt, allerdings kann diese über die KLT Tracks verifiziert werden. Während des Durchlaufens der Hauptschleife der Verfahren werden Rotationen sehr genau identifiziert. In diesen Fällen werden Instabilitäten vor allem durch starke und schwache Umgebungsbeleuchtung (bspw. Gegenlicht, Schatten) sowie schnellen Beleuchtungsänderungen (s. Abbildung 8-4 links) hervorgerufen. In Sequenzen mit Beleuchtungsänderungen, bei Regen, der Reinigung durch die Scheibenwischer und Reflexionen, treten wechselhafte Bildinhalte in kurzen Zeitabständen auf. Diese haben zur Folge, dass das DSO Front End die Erzeugung von neuen Keyframes fordert. Der Abstand zwischen Keyframes sinkt in diesem Fall im Mittel von fünf auf zwei Bilder. Die Erzeugung von neuen 3D-Punkten wird hierbei gefördert, allerdings werden die bereits vorhandenen 3D-Punkte nur bedingt in ihrer Position optimiert. Alte Keyframes werden zudem verworfen und können nicht zur Stabilisierung beitragen. Insofern kommt es hier vermehrt zum Abbrechen der Rekonstruktion und Reinitialisierung.



Abbildung 8-4 Fehlgeschlagene Initialisierung/Rekonstruktion mit dem direktem Ansatz. Links: Helle Umgebungsbeleuchtung und starke Beleuchtungsänderung; Mitte: Regen, Scheibenwischer, Reflektionen auf der nassen Fahrbahn; Rechts: Keine Objekte im Bildvordergrund

Ein Einfluss dieser Umgebungsbedingungen ist auch beim SfM-Verfahren zu beobachten. Es erfolgt eine Reduzierung der verfolgten Merkmale und rekonstruierten 3D-Punkte. Allerdings wird die Stabilität der Bewegungsschätzung und Rekonstruktion nur marginal beeinflusst. Im rechten Teil der Abbildung ist eine Situation zu sehen, in der nur wenige bis keine Objekte im Bildvordergrund auftauchen. Obwohl die Bildinhalte sich nur wenig ändern und eine stabile Beleuchtung vorherrscht, kann mit DSO keine stabile Rekonstruktion erreicht werden. Problematisch sind hierbei die weit entfernten 3D-Punkte, deren inverse Tiefe nicht konvergiert. In diesem Fall wird eine Reinitialisierung eingeleitet und eine neue Punktwolke aufgebaut. Dies wiederholt sich entsprechend oft, sodass keine zusammenhängende Szene erstellt wird. Das SfM-Verfahren ermöglicht in diesen Situationen eine durchgehende Rekonstruktion.

Für Sequenzen, die während der Dämmerung oder in der Nacht aufgenommen wurden, konnten keine Punktwolken erstellt werden. Es wird hierbei auf die eingesetzte Kamera verwiesen, welche nicht für Automotive-Anwendungen konzipiert wurde.

3D-Punktraum

Aus Tabelle 8-1 können die durchschnittlichen Punktzahlen der untersuchten Sequenzen entnommen werden. Mit dem direkten Verfahren werden demnach ca. dreimal mehr Punkte rekonstruiert als mit dem indirekten Ansatz. Dies ist unter anderem auf die Vorfilter (min. Tracklänge und min. Schnittwinkel) beim indirekten Verfahren zurückzuführen. Es werden weniger Punkte in großer Entfernung erzeugt, die eine mindere Qualität aufweisen. Beim direkten Verfahren werden diese Punkte jedoch mit einer inversen Tiefe von $\rho = 1$ initialisiert und anschließend optimiert. Allerdings ist die Entfernung bei Punkten in der Bildmitte so hoch, dass diese während der Reduzierung der Punktraumgröße (vgl. Abschnitt 7.1.2) entfernt werden (vgl. 52.65 - 63.58 %). Durch die in Abschnitt 7.1.1 erwähnte Reduzierung von Rauschen und Ausreißern werden zudem bei beiden Verfahren ca. 3.3 - 3.5 % der Punkte eliminiert, sodass sich eine Anzahl an gefilterten Punkten für die anschließende Fahruntergrund- und Hindernisdetektion ergibt, welche nur noch einen Faktor von ca. 2.2 zwischen indirekt und direktem Ansatz besitzt. Während der Klassifizierung wird deutlich, dass beide Verfahren eine ähnliche Punktverteilung aufweisen. Die Punkte verteilen sich gleichmäßig auf Fahruntergrund und Struktur. Ca. 10 % des Punktraums werden als potentielle Hindernisse, die keinen Bodenkontakt besitzen, klassifiziert. Ca. 15 % dieser Hindernispunkte (1.37 % und 1.74 % der gefilterten Punkte) befinden sich dabei innerhalb des prädierten Fahrschlauchs und werden als identifizierte Hindernispunkte im Hinblick auf Abstand und Höhe bewertet. Beide Verfahren verfügen über eine ähnliche Punktverteilung innerhalb der Klassen und eignen sich grundsätzlich für die Funktion der Durchfahrtshöhenerkennung, die eine genaue Identifizierung des Fahruntergrundes und der potentiellen Objekte bedingt.

Tabelle 8-1 Gemittelte Anzahl an 3D-Punkten beim indirekten und direkten Verfahren in den untersuchten Sequenzen

	Indirektes Verfahren	Direktes Verfahren
Rekonstruierte Punkte	10402	31736
Filterung		
Punkte außerhalb des 3D-Raums [%]	52.65	63.58
Rauschen und Ausreißer [%]	3.51	3.37
Gefilterte Punkte	4561	10487
Fahreruntergrund [%]	46.52	43.75
Struktur [%]	44.82	45.97
Potentielle Hindernisse [%]	8.66	10.28
Hindernispunkte im Fahrschlauch [%]	1.37	1.74

Das direkte Verfahren rekonstruiert mehr Punkte in großen Distanzen, die jedoch anschließend gefiltert werden. Aufgrund der geringen Dichte eignen sich diese Punkte nicht für eine Auswertung. Aus der Auswertung geht hervor, dass mit dem indirekten Verfahren eine zuverlässigere Rekonstruktion möglich ist als mit dem direkten Ansatz. Durch die Bewertung der Merkmalsverfolgung, d.h. der Analyse der Tracklänge und des Schnittwinkels sowie der Filterung im 3D-Raum über den maximalen Rückprojektionsfehler, werden ausschließlich 3D-Punkte mit hoher Qualität rekonstruiert. Die 3D-Szene mit DSO ist hingegen dichter und Objektumrisse werden vollständig erkannt. Problematisch sind hier jedoch Beleuchtungsschwankungen, Regen und Reflexionen. Es empfiehlt sich für einen Einsatz im Fahrzeug die Kombination beider Verfahren, wobei die Bewegungsschätzung über den indirekten Ansatz erfolgt und zusätzliche 3D-Punkte über das photometrische Verfahren erzeugt werden. Als Limitierung kann hier die verwendete Kamera gesehen werden, deren Belichtungssteuerung keine geeigneten Aufnahmen in Dämmerungs- und Nachtsituationen zulässt.

8.2.1.2 Fahrbahnrekonstruktion

Essentiell für die Berechnung der Hindernishöhe ist zunächst die Bestimmung des vertikalen Fahrbahnverlaufs. Da dieser nicht durch eine statische Ebene approximiert wird, bedarf es einer Vielzahl an Punkten auf dem Untergrund, um eine Modellierung zu ermöglichen. Zur Bestimmung der maximalen Detektionsreichweite werden die Punktwolken und die dazu gehörigen Trajektorien



Abbildung 8-5 Testsequenzen zur Fahrbahnrekonstruktion. Links: Asphaltierte Straße mit Markierungen, Bordsteinkanten und Schattenwurf; Mitte: Gepflasterte Fahrbahn; Rechts: Geringe Umgebungsbeleuchtung

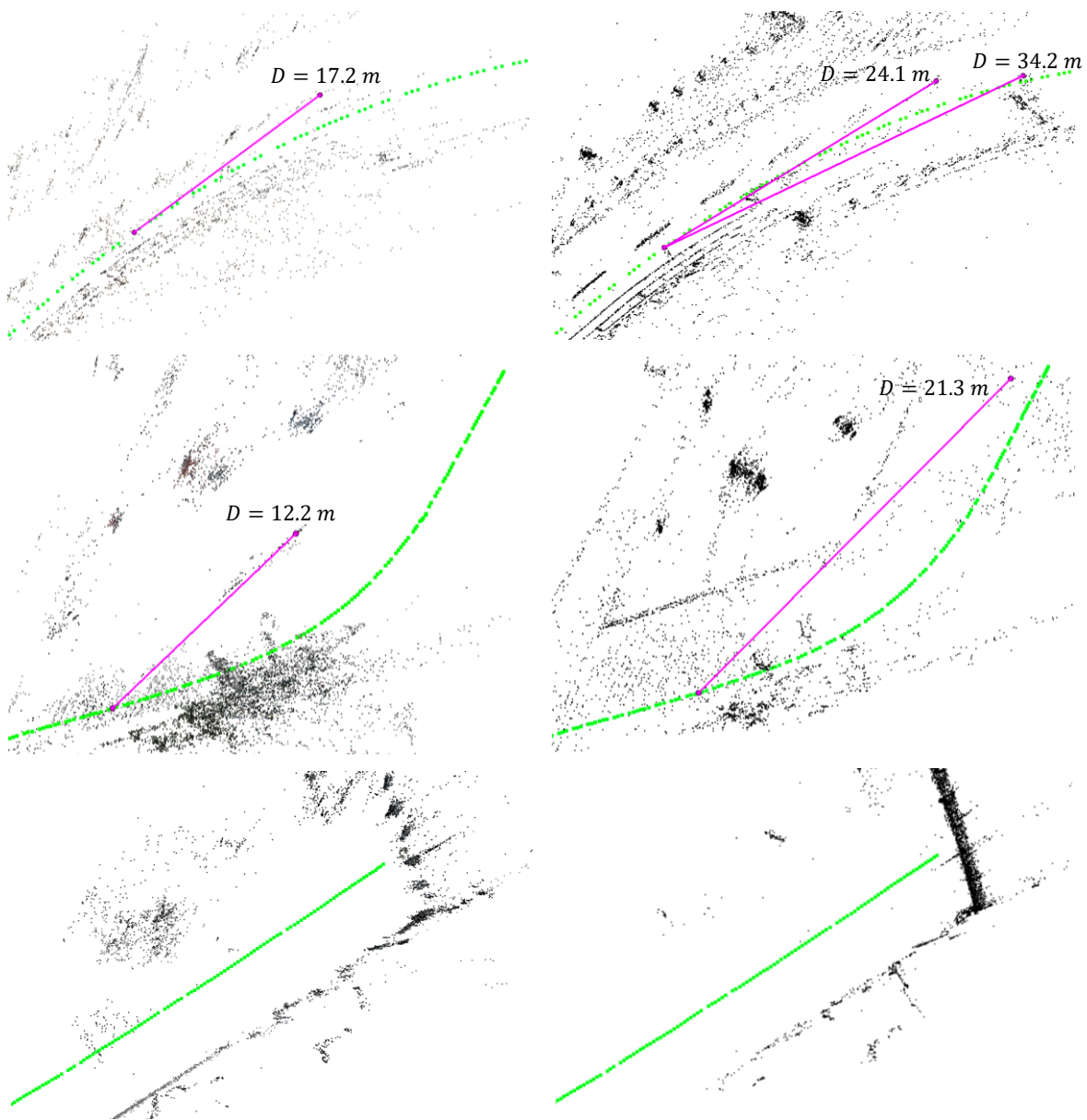


Abbildung 8-6 Fahrbahnrekonstruktion (Draufsicht) und maximaler Abstand D zu Bodenpunkten bei $I = 100$. Links: SfM; Rechts: DSO; Oben: Asphaltierte Straße mit Markierungen, Bordsteinkanten und Schattenwurf; Mitte: Gepflasterte Fahrbahn; Unten: Geringe Umgebungsbeleuchtung

händisch untersucht. Die Sequenzen unterscheiden sich in der Detailreife der Fahrbahn und der Umgebungsbeleuchtung. Exemplarisch werden hier drei Sequenzen gezeigt. Es kann davon ausgegangen werden, dass die Funktion zum Großteil auf öffentlichen Straßen und bebauten Privatgeländen zur Anwendung kommt. Dort sind vornehmlich Asphalt- bzw. gepflasterte Untergründe vorzufinden, auf denen je nach Situation Markierungen und Schattenwurf oder Bordsteinkanten am Fahrbahnrand enthalten sein können (s. Abbildung 8-5). Nach Start der Rekonstruktion mit SfM und DSO werden jeweils 100 Bilder der Sequenz durchlaufen und anschließend der detektierbare Untergrundbereich ermittelt. Nach dieser Anzahl von Bildern ist von einer initialisierten 3D-Szene auszugehen, die sich durch einen hohen Detailreichtum auszeichnet und sich zur Bodenrekonstruktion eignen sollte.

Oben in Abbildung 8-6 wird ersichtlich, dass die Fahrbahnmarkierungen mit dem SfM-Verfahren (links) vor allem in den Bereichen der Ecken rekonstruiert werden, was auf den Detektor zurückzuführen ist. Die Bordsteinkanten werden in der Distanz detektiert, lassen jedoch aufgrund der geringen Punktzahl keine eindeutige Schätzung der Fahrbahn zu. Die maximale Detektionsreichweite liegt hier bei ca. 17 m. Im Gegensatz dazu werden durch das indirekte Verfahren (rechts) mehr Punkte auf dem Untergrund und an den Bordsteinkanten rekonstruiert. Die Fahrbahnmarkierungen und der Schatten der Leitplanke werden im Nahbereich sehr detailliert und gesamtheitlich dargestellt. Die Streuung der Punkte in Fahrtrichtung wird bei der Darstellung der weiter entfernten Markierungen sichtbar. Es ergibt sich eine Detektionsreichweite zwischen ca. 24 m und 34 m. Im mittleren Teil des Bildes wird die Rekonstruktion einer gepflasterten Fahrbahn gezeigt, bei der die Detektionsreichweite auf bis zu 12 m bzw. 21 m sinkt. Die detailreiche Struktur des Fahruntergrunds erschwert das Tracking der Bildpunkte, sodass vor allem beim SfM-Verfahren wenige Merkmale stabil verfolgt werden können. DSO detektiert in erster Linie die Schattenkante unter der Einfahrt und die Bordsteinkante. Im unteren Teil des Bildes wird eine gering ausgeleuchtete Szene gezeigt. Ohne die Anpassung der Belichtungssteuerung der Kamera oder das Anschalten der Frontscheinwerfer können in diesen Situationen keine Bodenpunkte rekonstruiert werden. In solchen Fällen muss auf eine Grundebenschätzung zurückgegriffen werden, welche durch die vier Aufstandspunkte des Fahrzeugs und der kalibrierten Einbauhöhe der Kamera definiert wird.

8.2.1.3 Struktur- und Hindernisrekonstruktion

Die Präzision und Genauigkeit der 3D-Punkte ist entscheidend für die Berechnung der Hindernishöhe, während die Dichte der Punktwolke sich in erster Linie auf die Ersterkennung des Hindernisses auswirkt. Mit DSO werden vornehmlich Grauwertunterschiede an Linien verfolgt, welche häufig die Unterkanten von Hindernissen repräsentieren. SfM basiert auf den Eckpunkten des Shi-Tomasi Detektors, sodass Hindernisse häufig flächig rekonstruiert werden und die eigentliche Unterkante unterrepräsentiert dargestellt wird. Es ist davon auszugehen, dass dadurch die Entfernung zur Erstdetektion sowie Genauigkeit der Hindernishöhe sinkt. In Abbildung 8-7 sind beispielhafte Hindernisse aufgelistet, deren Rekonstruktionen in Abbildung 8-8 gezeigt werden.



Abbildung 8-7 Ausgewählte Hindernisse. Links: Einfahrt zu einem Parkdeck; Mitte: Tunnelleinfahrt; Rechts: Freihängende Warnschilder und -baken

In der ersten Sequenz wird ein Parkdeck gezeigt, welches mit niedriger Geschwindigkeit angefahren wird. Aus den Punktwolken wird ersichtlich, dass mit beiden Verfahren eine Vielzahl an 3D-Punkten auf dem Gelände und dem Fahrzeug erzeugt werden. Die eigentliche Unterkante wird allein durch DSO detektiert. In der zweiten Sequenz ist die Einfahrt in einen Tunnel dargestellt, welcher vor allem ein Hindernis für Lastkraftwagen und Sondertransporte darstellt. Diese inner- wie außerstädtisch vorkommenden Hindernisse werden häufig mit einer hohen Geschwindigkeit angefahren, sodass hierbei eine frühzeitige Detektion entscheidend ist. Allerdings sind sie schon aus großer Distanz zu erkennen, sodass die Verfolgung und Rekonstruktion über einen längeren Zeitraum erfolgen kann.

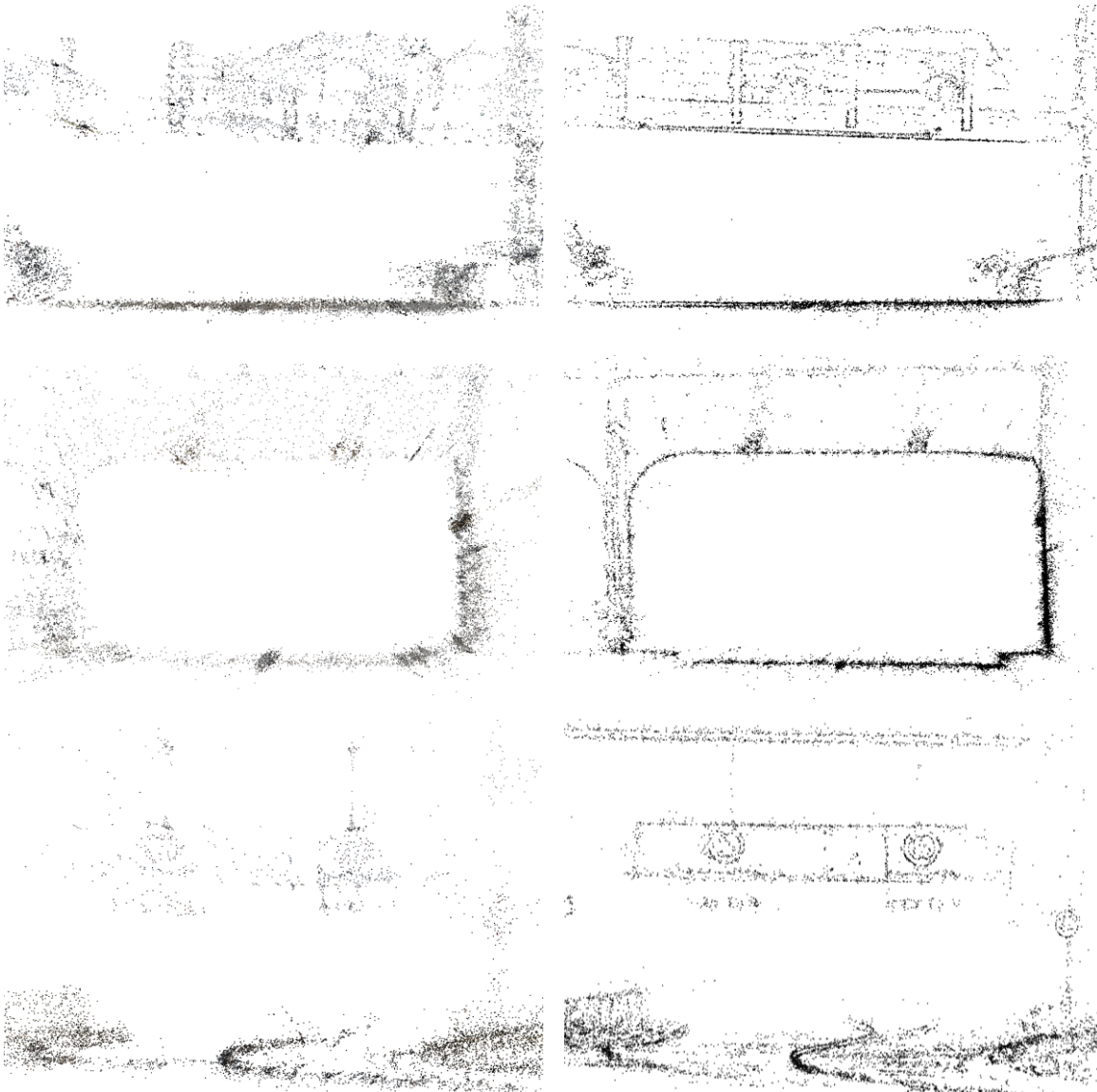


Abbildung 8-8 3D-Punktwolke der Hindernisse. Links: SfM; Rechts: DSO; Oben: Einfahrt zu einem Parkdeck; Mitte: Tunnelleinfahrt; Unten: Freihängende Warnschilder und -baken

Auch hier werden mit DSO vornehmlich Kanten nachgebildet, während das SfM-Verfahren mehr Punkte auf den Betonwänden rekonstruiert. Eine besondere Herausforderung ist die Detektion von freihängenden Schildern und Warnbaken, die in der dritten Sequenz dargestellt werden. Diese Hindernisse sind vergleichsweise klein und ihre Kennzeichnung ist innerhalb eines Graustufenbilds nur bedingt vom Hintergrund zu unterscheiden. Sie sind schon weit vor Parkhäusern aufgehängt, um etwaige Unfälle mit dem eigentlichen Hindernis zu vermeiden. Allerdings handelt es sich hierbei oft um Metallkonstruktionen, die selbst einen Schaden hervorrufen können. Diese Hindernisse sollten langsam angefahren werden, um die Fahrzeughöhe abgleichen zu können. Dementsprechend ist die Zeitspanne zur Detektion und Verfolgung ausreichend groß.

8.2.2 Linien

Die Linienrekonstruktion nach [111] wird in diesem Kapitel auf ausgewählten Sequenzen angewendet. Der Fokus liegt dabei vor allem auf der Rekonstruktion von Linien auf den Hindernisunterkanten, wodurch die Hindernisdetektion und die Berechnung der tatsächlichen Durchfahrtshöhe optimiert werden soll. Ein weiterer Aspekt ist die Rekonstruktion von Linien auf der Fahrbahn. Vorteil der Linien ist, dass über deren 3D-Richtungsvektor ihre Zugehörigkeit zur Fahrbahn verifiziert werden kann. Somit kann im weiteren Verlauf die Fahrbahnoberfläche innerhalb der rekonstruierten 3D-Punktwolke verlässlicher extrahiert und modelliert werden. Zur Vermeidung von fehlerhaften Korrespondenzen ist die Mindestlänge der zu detektierenden 2D-Kantensegmente im Bild parametrierbar. Sie ist so eingestellt, dass quer zur Fahrbahn liegende Hindernisse, deren Breite mindestens der halben Fahrzeugbreite entsprechen, bereits in einer Distanz von 40 m detektiert werden. Im Bild entspricht dies einer Länge von ca. 26 Pixel.

8.2.2.1 Fahrbahnrekonstruktion

Die angegebene Mindestlänge der Liniensegmente im Bild ermöglicht die Rekonstruktion von langen 3D-Linien in großer Entfernung, welche die Erstdetektion von Hindernissen verbessern sollen. Fehleranfällige kurze Linien sollen dabei vermieden werden. Dies wirkt sich jedoch negativ auf die Rekonstruktion von Fahrbahnmarkierungen aus, welche entlang der Kamerahauptachse ausgerichtet sind (s. Abbildung 8-9 oben).

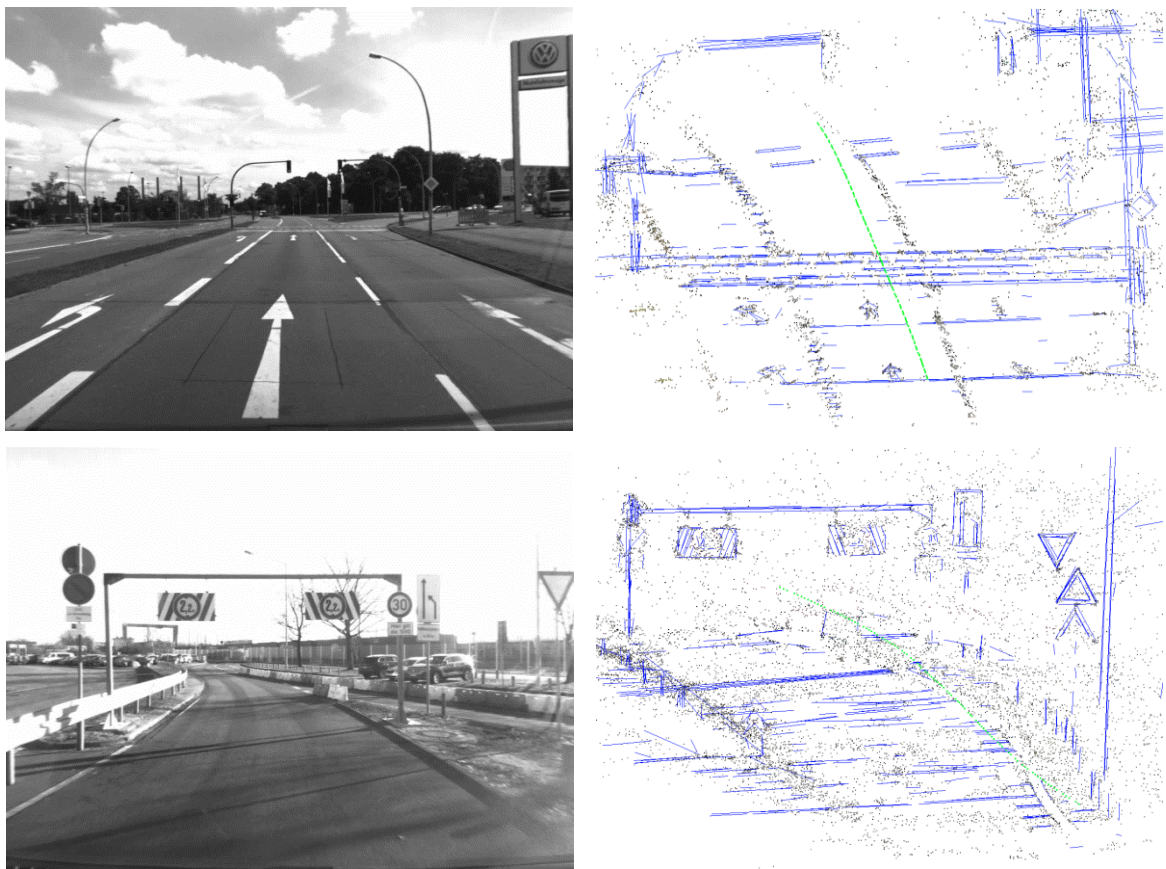


Abbildung 8-9 Linienbasierte Rekonstruktion des Fahreruntergrunds. Oben: Vornehmlich quer zur Fahrbahn rekonstruierte Linien; Unten: Rekonstruktion von Schlagschatten

Laut [139] weisen Fahrbahnmarkierungen innerorts eine Länge von 3 m auf. In der Nähe von Knotenpunkten kann die Länge auch 0.5 - 1 m betragen. Unter Berücksichtigung der in Kapitel 3 bestimmten inneren und äußeren Orientierung der im Fahrzeug verbauten Kamera, ergibt sich ein maximaler Abstand von ca. 12 m für die Erstdetektion auf der Fahrbahnoberfläche für 3 m-Markierungen (5 m für 1 m-Markierungen). Die Auswertung der Sequenzen zeigt, dass die Verfolgung der Linien und die Verifizierung zu einer gemeinsamen 3D-Linie häufig erst ab einem Abstand von 4 m zum Fahrzeug durchgeführt werden kann. Dementsprechend können diese Linien nur bei langsamer Fahrt im Nahbereich des Fahrzeugs zur Fahrbahnmodellierung herangezogen werden. Bei quer zur Fahrbahn liegenden Markierungen und Schlagschatten, welche durch flache Sonneneinstrahlung hervorgerufen werden, kann die Rekonstruktion jedoch bis zu einer maximalen Entfernung von ca. 17 m erfolgen.

8.2.2.2 Struktur- und Hindernisrekonstruktion

Im Bild gefundene horizontale Linien beschreiben einen Großteil der typischen Durchfahrtshindernisse, die im Straßenverkehr vertreten sind. Typische Objekte sind Parkhauseinfahrten, Brückenunterführungen und Schilder, die mit ihren Unterkanten häufig die maximale lichte Höhe repräsentieren. In Abbildung 8-10 werden ausgewählte Sequenzen, deren Punktwolken und die dazu mit *Line3D++* rekonstruierten Linien gezeigt. Da Unterkanten im Bild zumeist homogen dargestellt werden, wird das Finden und Verfolgen dieser mittels merkmalsbasiertem Verfahren erschwert. Das linienbasierte Verfahren ermöglicht auch die Erkennung und Rekonstruktion dieser Hindernisse, wie im oberen Teil des Bildes verdeutlicht wird. Durch den Tiefenfehler, der während der Rekonstruktion der Punkte auftritt, werden mehrere nah beieinanderliegende Hindernisse in der 3D-Punktwolke häufig zusammengefasst. Die Linien unterstützen bei der korrekten Identifizierung des kritischsten Hindernisses im 3D-Raum. Dies wird in Bildmitte verdeutlicht, bei der zwei versetzte Unterkanten gefunden werden. Im unteren Teil des Bildes wird die Rekonstruktion eines Portals mit freihängenden Schildern gezeigt. Das linke Schild wird dabei gesamtheitlich an allen Außenkanten erkannt, während im rechten Schild aufgrund der Kombination von rotweißer Schraffierung und geringem Kontrast zum Hintergrund keine zusammenhängende Unterkante detektiert wird.

Die Analyse der Sequenzen zeigt, dass die Rekonstruktion bei großen Kantensegmenten, bspw. bei Brücken oder Tunnelleinfahrten in einer Distanz von bis zu 19 m stattfinden kann. Bei kleineren Hindernissen wie Schildern oder Warnbaken liegt die Detektionsreichweite bei ca. 8 m. Eine vertiefte Betrachtung wird in Abschnitt 8.4 vorgenommen.

Evaluierung

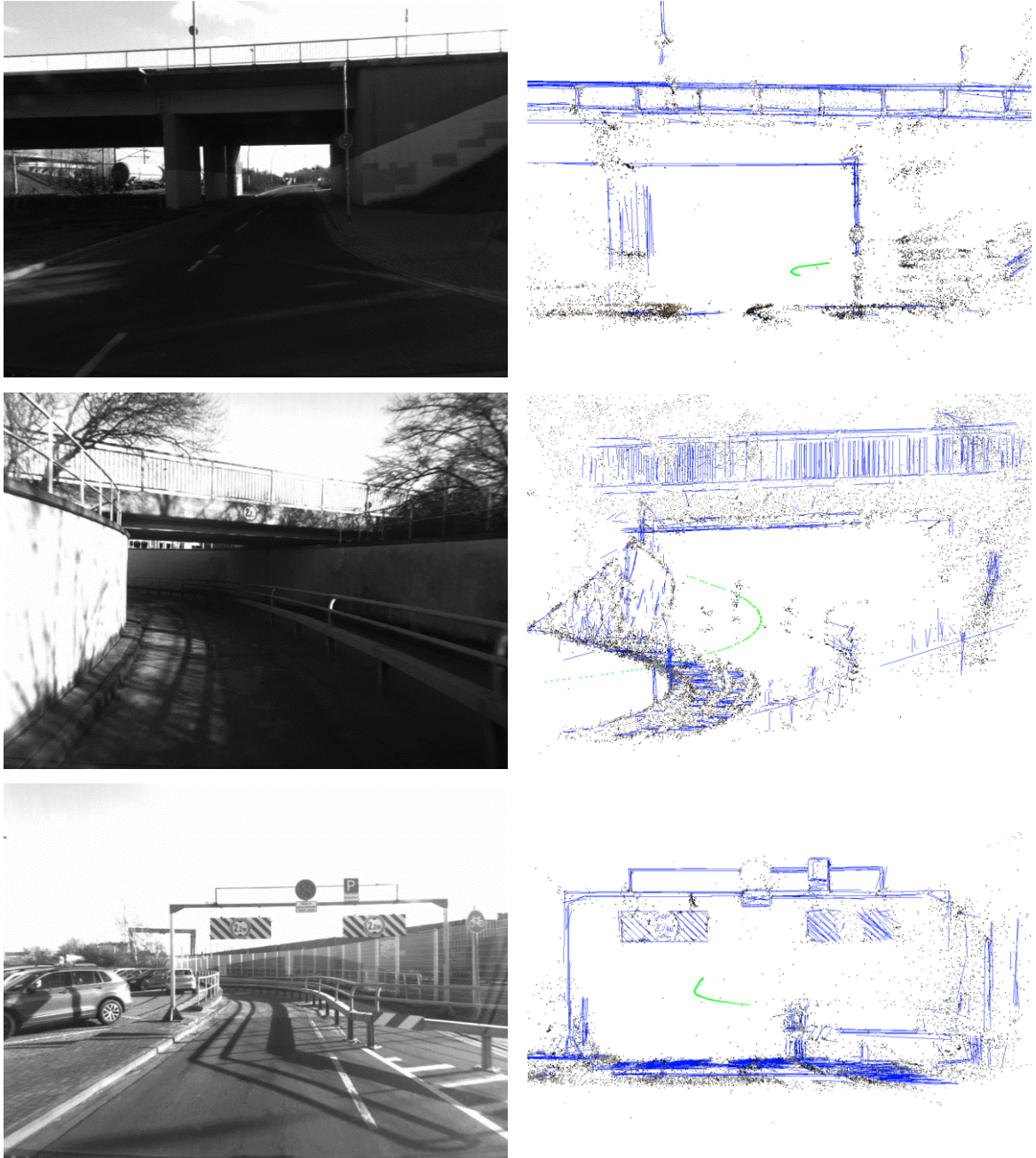


Abbildung 8-10 Linienbasierte Rekonstruktion von Hindernissen. Oben: Rekonstruktion der wahren Brückenunterkante, Fehlschlagen der Detektion mittels SfM-Verfahren; Mitte: Identifizierung von zwei räumlich getrennten Unterkanten; Unten: Gesamtheitliche Rekonstruktion des Portals durch Linien

8.3 Fahruntergrundmodellierung

Die Fahrbahnrekonstruktion ist entscheidend für die Berechnung der korrekten Hindernishöhe. Vor allem in großen Entfernungen ist die Approximation des Fahruntergrundes durch eine einfache Ebene fehleranfällig, sodass hierbei eine Modellierung kurzer Teilebenen stattfindet, die es ermöglicht sich dem vertikalen Fahrbahnverlauf anzupassen. Aus der Analyse der Punktwolken geht hervor, dass die Detektionsreichweite im Idealfall bei ca. 30 m liegt; selten jedoch darüber. Je nach Struktur der Fahrbahnoberfläche und Umgebungsbedingungen reduziert sich dieser Wert jedoch auf ca. knapp 12 m für das SfM-Verfahren bzw. 21 m für das DSO-Verfahren. Mithilfe quer zur Fahrbahn ausgerichteten Linien wurde eine maximale Reichweite von 17 m geschätzt. Die Fahruntergrundmodellierung nutzt diese Informationen, um in Abhängigkeit der Stützstellen der Gitterstruktur für jede Zelle in Fahrtrichtung eine Untergrundhöhe zu bestimmen. Die Genauigkeit der Modellierung über alle Sequenzen wird in Abbildung 8-11 dargestellt und in Tabelle 8-2 zusammengefasst². Die Parameter der Modellierung werden in beiden Punkträumen (SfM, DSO) konstant gehalten. Darüber hinaus ist das Modellierungsverfahren deterministisch, sodass der Fehler des Höhenplans vor allem auf die Repräsentation der Umgebung im 3D-Raum zurückzuführen ist. Als Referenz werden die Positionsdaten der Inertialplattform RT3002 herangezogen. Über die Länge einer Sequenz wird die vertikale Trajektorie des Fahrzeugs ausgewertet und die Bodenhöhe zu jedem Zeitschritt über die kalibrierte Kamerahöhe bestimmt.

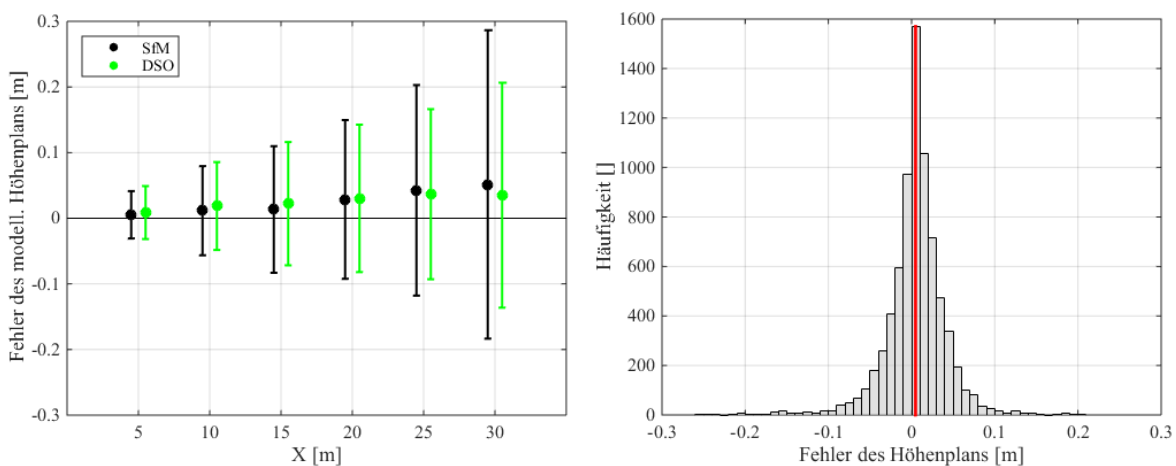


Abbildung 8-11 Fehler der Fahrbahnmodellierung aller Sequenzen. Links: Mittelwert und Standardabweichung des Höhenfehlers in der SfM- bzw. DSO-Punktwolke; Rechts: Exemplarische Verteilung bei 5 m (SfM)

Tabelle 8-2 Übersicht der Fehler der Fahrbahnmodellierung aller Sequenzen [cm]

	5 m	10 m	15 m	20 m	25 m	30 m
SfM	0.5 ± 3.6	1.1 ± 6.8	1.3 ± 9.6	2.8 ± 12.1	4.2 ± 16.0	5.1 ± 23.5
DSO	0.8 ± 4.0	1.9 ± 6.7	2.2 ± 9.4	3.0 ± 11.2	3.6 ± 13.0	3.6 ± 17.1

Die Analyse zeigt, dass die Fehler in den ausgewählten Abständen einer Normalverteilung unterliegen (s. Abbildung 8-11 rechts). Es werden Mittelwert und Standardabweichung der Modellierung miteinander verglichen. Ausgehend von 5.1 cm (SfM) und 3.6 cm (DSO) in einem Abstand von 30 m Entfernung, nimmt der Mittelwert der Modellierung nahezu linear auf 0.5 cm bzw. 0.8 cm im Nahbereich ab. Der Höhenplan wird bei beiden Rekonstruktionserfahren in allen

² Für eine detaillierte Untersuchung von ausgewählten Sequenzen wird auf Appendix 0 verwiesen.

Distanzen leicht überschätzt. Dies leitet sich vornehmlich aus den Bebauungen am Fahrbahnrand ab, da Bordsteinkanten, Schilder- und Warnbakenbefestigungen anteilig in die Berechnung der Normalenvektoren mit einbezogen werden. Um dem entgegenzuwirken kann der untersuchte 3D-Raum in Y weiter begrenzt werden, was jedoch zu einer dünnen Punktwolke am Boden führt, welche sich ggf. nicht für eine Modellierung eignet. Die größten Fehler ergeben erwartungsgemäß in weit entfernten Abständen zum Fahrzeug und nehmen ebenfalls linear mit sinkender Distanz ab. Aus den individuellen Sequenzen in 0 B geht hervor, dass Fahrbahnmarkierungen und Schattenwurf in beiden Verfahren die Genauigkeit um ca. 25 % steigern, sofern von einer nahezu ebenen und geraden Fahrbahn ausgegangen wird. Als problematisch haben sich jedoch folgende Szenarien herausgestellt. Bei Zufahrt auf eine Kurve wird zunächst, solange kein Einlenken durch den Fahrer stattfindet, der Fahrschlauch als Gerade approximiert. Dies führt in großen Entfernungen dazu, dass die Schätzung über die eigentliche Fahrbahn hinausführt, dortige Bodenobjekte miteinbezogen werden und damit die Untergrundhöhe in der Modellierung überschätzt wird. Des Weiteren führen Steigungen und Gefälle oberhalb der definierten 20° zu Fehlinterpretationen und hohen Fehlern bereits im Nahbereich. In Abschnitten in denen keine 3D-Punkte rekonstruiert werden, sollte auf die hier vorgestellte Bodenschätzung verzichtet werden. Dies geschieht vornehmlich in dunklen Sequenzen, bei großen Schattenwürfen, bspw. beim Einfahren in einen Tunnel. Die Approximation über eine Grundebene, die sich aus der Kamerahöhe ergibt, führt hierbei ebenfalls zur einer fehlerbehafteten Untergrundschätzung, welche jedoch stabiler ist. In Entfernungen größer 30 m sollte gänzlich auf eine Schätzung des Fahruntergrunds bzw. Extrapolation der vorherigen Untergrundhöhen verzichtet werden.

Zusammenfassend wird festgehalten, dass die Modellierung des Fahruntergrunds maßgeblich von der zur Verfügung stehenden Punktzahl beeinflusst wird. Während mit DSO im entfernten Bereich (>20 m) die genaueren Ergebnisse erreicht werden, kann beim SfM-Verfahren nur im Nahbereich eine zuverlässige Modellierung erfolgen. Es wird vorgeschlagen die maximale Entfernung für die Untergrundmodellierung an die Warnstrategie anzulehnen, sodass nur jeweils bis zur maximal eingestellten Vorwarnzeit prädiiziert und modelliert wird. Dabei sollte zudem die eine Stabilisierung des Fahruntergrunds unter Berücksichtigung zeitlich vorausgehender Modellierungen erfolgen. Dabei können die alten Stützpunkte entsprechend der Fahrzeugtrajektorie verschoben und als Startwerte für eine Optimierung der Untergrundhöhe angesehen werden.

8.4 Hindernisdetektion

Nach der Fahruntergrundmodellierung erfolgt die Bestimmung der Struktur und der potentiellen Hindernisse, wie in Abschnitt 7.3 beschrieben. Über die Prädiktion der Fahrzeugbewegung wird ein dreidimensionaler Fahrschlauch geschätzt, dessen Bewegung in Y über den Lenkwinkel und in Z über den modellierten Fahruntergrund bestimmt wird. Die Zellen potentieller Objekte, die sich innerhalb des Fahrschlauchs befinden, werden im Zuge der Hindernisdetektion ausgewertet. Dies beinhaltet die Bestimmung der minimalen Durchfahrtshöhe eines Hindernisses bei gleichzeitiger Berechnung des Abstands zum Fahrzeug. Die beiden Werte bilden die Basis für die nachgelagerte Warnstrategie, die von einer Anzeige der Kritikalität des Hindernisses im Kombiinstrument bis hin zur automatischen Notbremsung reicht. Abbildung 8-12 zeigt exemplarisch die gefilterten Punktwolken des SfM- und Linienvorgangs (oben) sowie die äquivalente DSO-Punktwolke (unten). Neben der Fahrzeugtrajektorie ist die auf den Fahrschlauch begrenzte geschätzte Bodenebene zu sehen. In Rot sind die identifizierten Hindernispunkte gekennzeichnet, die neben den Linien zur Höhenbestimmung ausgewertet werden.

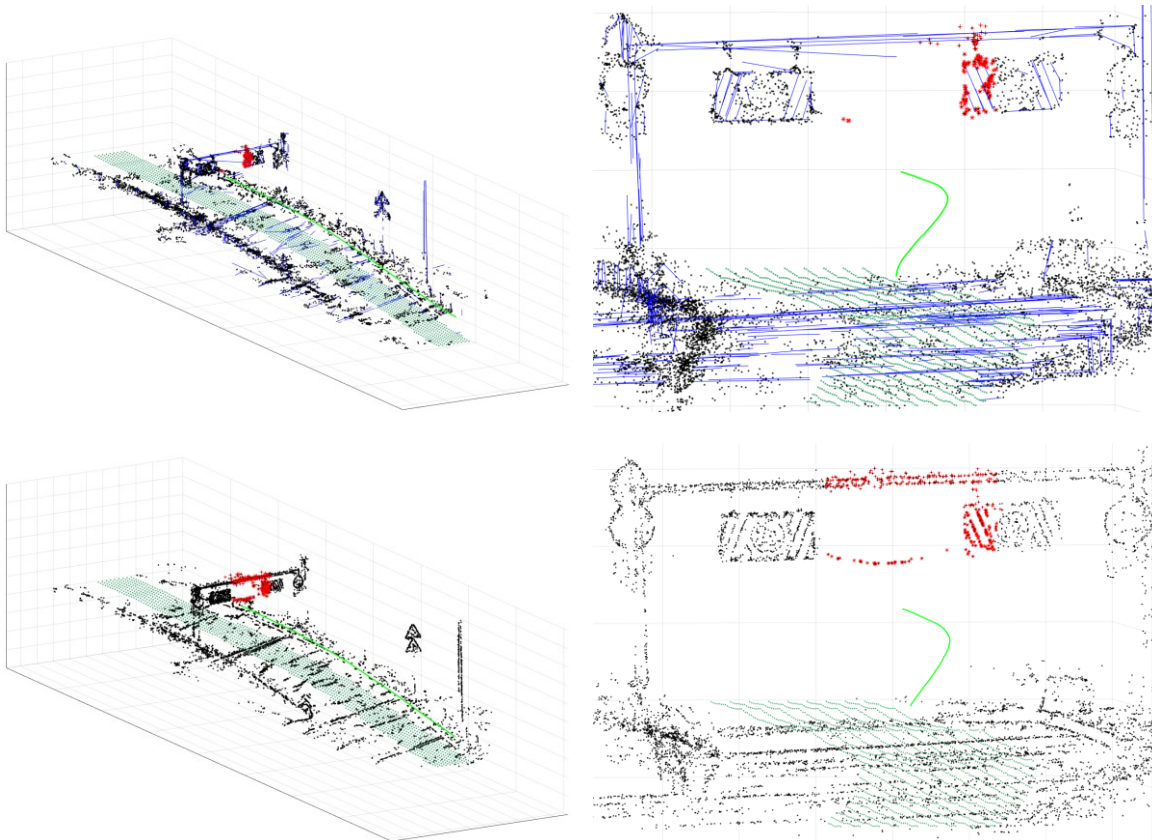


Abbildung 8-12 Gefilterter 3D-Punkt- und Linienraum mit Trajektorie (hellgrün), Fahruntergrund (dunkelgrün) und identifizierten Hindernispunkten innerhalb des Fahrschlauchs. Oben: SfM und Linien; Unten: DSO

8.4.1 Grundlagen

Grundlage für die Auswertung der Höhenberechnung sind die 3D-Punkte und 3D-Linien, die durch die vorgestellten Verfahren erzeugt werden. Im Falle der indirekten Verfahren wird innerhalb einer Sequenz nach der Initialisierung alle fünf Bilder eine Punktwolke extrahiert, gefiltert, klassifiziert und die Hindernisdetektion durchlaufen. Beim direkten Verfahren wird jeweils nach der Erstellung eines neuen Keyframes die Punktwolke extrahiert und ausgewertet. Dies führt dazu, dass sich mit zunehmender Bildanzahl der 3D-Raum allmählich aufbaut, immer mehr Punkte bzw. Linien auf dem Hindernis rekonstruiert werden und die minimale lichte Höhe sicherer bestimmt werden kann. Nachdem in den vorherigen Kapiteln die innere Genauigkeit für den Distanz- und Höhenfehler ermittelt wurde, soll hier die äußere Genauigkeit anhand von Referenzmaßen bestimmt werden.

Bestimmung der Referenzgrößen

Für die Auswertung der Detektionseigenschaften der implementierten Funktion werden für die Hindernishöhe und -entfernung Referenzdaten benötigt. Zur Höhenbestimmung werden dafür die Punktwolken des terrestrischen und mobilen Laserscanners herangezogen. Bei direkt erreichbaren Hindernissen wird hingegen die Messung mit einem Laserdistanzmessgerät mit der Genauigkeit von 2 mm durchgeführt. Zur Bestimmung der Entfernung wird der 3D-Raum des letzten Bildes einer Sequenz untersucht und per Hand die Lage des Hindernisses ermittelt (s. Abbildung 8-13).

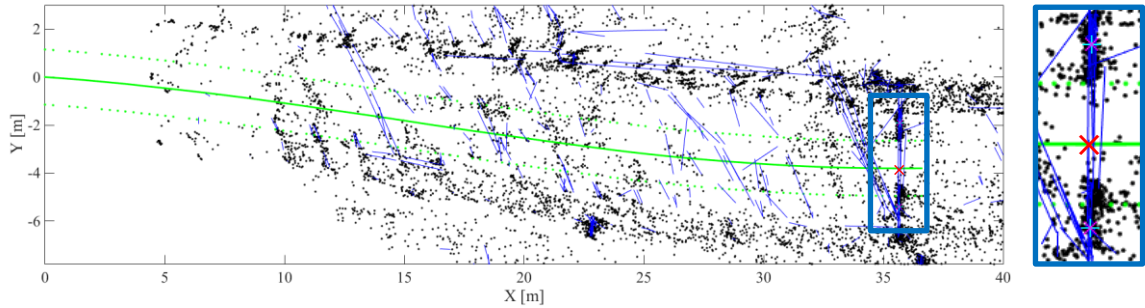


Abbildung 8-13 Manuelle Auswahl des Hindernisses. Links: Draufsicht einer SfM-Punktwolke, Trajektorie (grün), Fahrschlauch (grün - gepunktet) und detektierte Linien (blau); Rechts: Vergrößerter Ausschnitt des Hindernisses mit manuell gesetzten Marken (cyan-violett) und berechnetem Schnittpunkt mit der Trajektorie (rot)

Dafür werden auf dem Hindernis zwei Punkte markiert, deren resultierende Linie mit der Fahrzeugtrajektorie geschnitten wird. Während der Analyse wird in jedem Zeitschritt der kürzeste Abstand zwischen diesem Schnittpunkt und der Fahrzeugposition als Hindernisentfernung definiert und die Streuung der identifizierten Hindernispunkte dazu berechnet.

Berechnung der Hindernishöhe

Die identifizierten Hindernispunkte sind, wie aus Abbildung 8-12 zu entnehmen, auf der gesamten Höhe des prädizierten Fahrschlauchs verteilt. Im Realfall würde dieser über die Fahrzeughöhe begrenzt werden und eine Vielzahl der detektierten Punkte bereits ausschließen. Um jedoch auch sehr hohe Hindernisse evaluieren zu können wird der vertikale Fahrschlauch nicht begrenzt und eine Strategie zur Punktauswahl implementiert. Zunächst wird die mittlere Höhe der untersten Hinderniszelle berechnet und anschließend abhängig von der Hindernisentfernung die Punkte in einer Zone ausgewählt, deren Ausdehnung an den erwarteten bzw. zugelassenen Höhen- und Distanzfehler angelehnt ist. Der Höhenfehler ergibt sich anteilig aus dem in Abschnitt 8.3 bestimmten Fehler der Fahruntergrundmodellierung s_B . Dieser wird als Richtlinie für die Punktgenauigkeit des gesamten 3D-Raums angesehen, sodass sich zusätzlich ein Fehler der Hindernisunterkante s_H im gleichen Maße ergibt. Da diese Höhenfehler im skalierten 3D-Punktraum bestimmt wurden, kann davon ausgegangen werden, dass die Unsicherheiten von Brennweite und Maßstab bereits enthalten sind. Die Formel zu Berechnung der resultierenden Hindernishöhe H ergibt sich zu

$$H = H_H - H_B, \quad \text{Gl. (8-1)}$$

wobei H_H die Höhe des 3D-Punktes auf Objektunterkante und H_B die Höhe des lotrechten Bodenpunktes beschreiben. Über eine Fehlerfortpflanzung können die einzelnen Höhenfehler zu resultierenden inneren Genauigkeit für unterschiedliche Distanzen zusammengefasst werden.

$$s_G = \sqrt{\left(\frac{\partial H}{\partial H_H}\right)^2 s_H^2 + \left(\frac{\partial H}{\partial H_B}\right)^2 s_B^2} \quad \text{Gl. (8-2)}$$

Über $s_H = s_B = s_{Obj}$ ergibt sich somit

$$s_G = \sqrt{2} s_{Obj}. \quad \text{Gl. (8-3)}$$

In Tabelle 8-3 ist die resultierende innere Genauigkeit des Verfahrens zur Höhenbestimmung eines Hindernisses berechnet.

Tabelle 8-3 Resultierender Höhenfehler des Hindernisses in ausgewählten Distanzen zur Bestimmung der Ausdehnung der Evaluierungszone

		Höhenfehler [cm]					
		Hindernis s_H		Fahreruntergrund s_B		Gesamt s_G	
		SfM	DSO	SfM	DSO	SfM	DSO
Hindernisdistanz [m]	5	3.6	4.0	3.6	4.0	5.1	5.7
	10	6.8	6.7	6.8	6.7	9.6	9.5
	15	9.6	9.4	9.6	9.4	13.6	13.3
	20	12.1	11.2	12.1	11.2	17.1	15.9
	25	16.0	13.0	16.0	13.0	22.6	18.4
	30	23.5	17.1	23.5	17.1	33.2	24.2

Die zu evaluierenden Linien liegen in der bestimmten Zone für die SfM-Punktwolke, wobei hier die Einschränkung gilt, dass der Winkel einer Linie in der Y-Z-Ebene 20° nicht überschreiten darf. Damit sollen nur nahezu waagerechte Linien untersucht werden, die eine mögliche Hindernisunterkante darstellen. Für die Distanz wird ein Fehler von maximal 2 m gefordert, welcher sich beim Annähern verringern sollte. Nah hintereinanderliegende Hindernisse werden somit in großen Distanzen zusammengefasst und erst im Nahbereich bei größerer Punktzahl separiert. Nach der Bestimmung der zu evaluierenden Punkte und Linien wird die Hindernishöhe berechnet. Die 3D-Punkte sind dabei auf dem gesamten Hindernis verteilt, sodass eine Bildung des Mittelwertes die lichte Höhe überschätzen würde. In diesem Sinne wird in der dynamischen Zone der Mittelwert nur mit den Punkten gebildet, welche sich innerhalb der unteren 5 cm befinden. Da die Punktwolkenfilterung grobe Ausreißer zu diesem Moment bereits eliminiert hat, kann davon ausgegangen werden, dass dies der Unterkante des Hindernisses entspricht. Die aus den Linien berechnete Hindernishöhe wird über den mittleren Schwerpunkt aller Linien in der Detektionszone bestimmt.

8.4.2 Auswertung

In diesem Abschnitt wird die Genauigkeit der Entfernungs- und Höhenberechnung der punkt- und linienbasierten Verfahren anhand von Sequenzen mit potentiellen Hindernissen bewertet. Die Hindernisse werden dazu in zwei Kategorien unterteilt. Zum einen werden große hohe Hindernisse untersucht, welche sich deutlich vom Hintergrund unterscheiden und sich durch eine gute Sichtbarkeit bereits in großen Entfernungen auszeichnen. Des Weiteren werden kleine Hindernisse analysiert, welche häufig vor Parkhäusern und Tiefgaragen auf die maximal zulässige Höhe verweisen. In Abbildung 8-14 sind ausgewählte Beispielsequenzen dazu aufgezeigt. Hierbei wird auf Appendix C verwiesen, in dem die einzelnen Sequenzen getrennt voneinander ausgewertet werden.

Evaluierung



Abbildung 8-14 Ausgewählte Sequenzen mit Durchfahrtshindernissen. Oben: große gut sichtbare Hindernisse; Unten: kleine freihängende Hindernisse

Durch die Punkthäufigkeiten beider Kategorien, dargestellt in Abbildung 8-15, lässt sich eine Aussage über die Entfernung der Erstdetektion eines Hindernisses treffen. So werden große Hindernisse gleichermaßen mit dem indirekten und direkten Verfahren bis zu einer Entfernung von 30 m detektiert. Allerdings sind dabei nur vereinzelt Punkte auf dem Hindernis rekonstruiert, deren Genauigkeit im weiteren Verlauf überprüft wird. Linien werden, wie auch der Untersuchung aus Abschnitt 8.2.2 zu entnehmen, erst im Nahbereich genau genug rekonstruiert, um in die Hindernisdetektion und der damit eingehenden Berechnung der lichten Höhe einfließen zu können. Kleine Hindernissen, die sich nur bedingt vom Hintergrund differenzieren, werden erst im absoluten Nahbereich von 5 m vollständig rekonstruiert. Ausgehend davon, dass diese Hindernisse vorwiegend mit geringer Geschwindigkeit angefahren werden, kann im Ernstfall trotzdem eine Warnung und Notbremsung erfolgen. Es sind jedoch vereinzelt hohe Erstdetektionen mit beiden Punktverfahren möglich. Die rotweißen Warnbaken ermöglichen eine Linienrekonstruktion ebenfalls nur im absoluten Nahbereich, da zu kurze Linien bereits während der Detektion eliminiert werden.

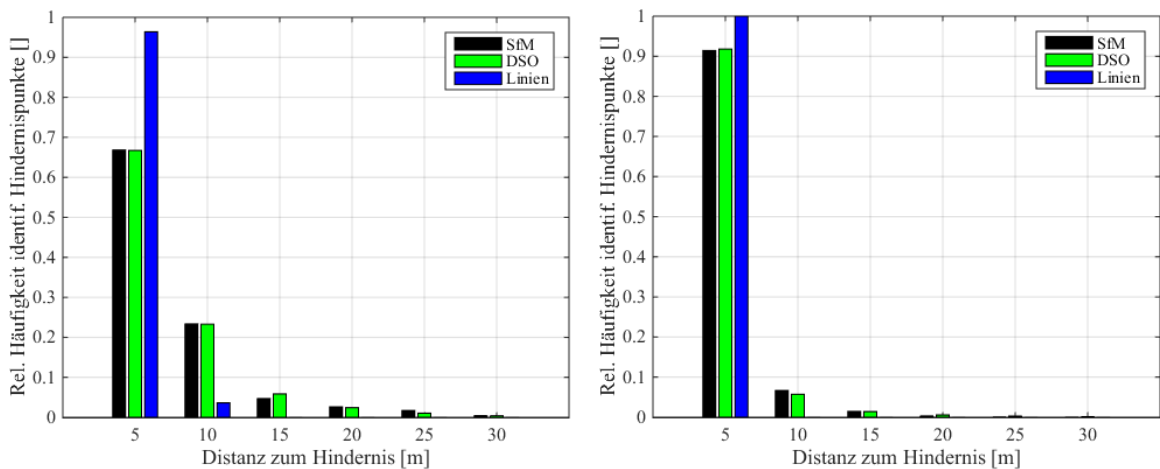


Abbildung 8-15 Punkthäufigkeiten in Abhängigkeit von der Distanz zum Hindernis. Links: große Hindernisse; Rechts: kleine Hindernisse

8.4.2.1 Hindernisentfernung

Die korrekte Berechnung der Hindernisentfernung ist ausschlaggebend für die Bestimmung der Time to Collision TTC (s. Gl. (2-1)), welche innerhalb der Warnstrategie die einzelnen Stufen der automatischen Notbremsung einleitet. Je früher ein Hindernis und dessen Kritikalität erkannt wird, desto höher kann zudem die maximale Geschwindigkeit der Funktion ausgelegt werden. Einhergehend mit der maximalen Untergrundmodellierung werden hier Hindernisse bis zu einem Abstand von 30 m untersucht. In Ausnahmefällen werden Hindernisse jedoch auch in größeren Entfernungen detektiert. In Abbildung 8-16 sind hierfür die Mittelwerte und Standardabweichungen der Hindernisentfernungen für große Hindernisse (links) und kleine Hindernisse (rechts) im erwähnten Bereich aufgetragen. Bei großen Objekten wie Brücken oder Tunneleinfahrten sind die berechneten Distanzen unter Nutzung der SfM-Punktwolke im Mittel sehr genau. Die Standardabweichungen liegen in großen Distanzen im Bereich von 60 - 80 cm. Bei der Zufahrt auf ein Hindernis ergeben sich jedoch mit längerer Beobachtung von Objektpunkten und der Annahme, dass sich das Hindernis im Bild mit abnehmender Distanz dem oberen Bildrand annähert, eine bessere Aufnahmegeometrie mit größeren Schnittwinkel. Das führt dazu, dass ab einer Distanz von ca. 10 m höhere Genauigkeiten (30 - 45 cm) erreicht werden können. Gleiches gilt für das DSO-Verfahren, wobei sich hier bereits in großen Entfernungen kleinere Fehler ergeben. Auffällig ist, dass DSO alle Hindernisse in ihrer Entfernung um ca. 20 - 25 cm überschätzt. Bei hohen Hindernissen wird durch DSO neben der Unterkante auch die Deckenstruktur der Unterführung rekonstruiert, wie z.B. Verstrebrungen, Metallträger oder angebrachte Beleuchtung. Diese wird ebenfalls als kritisch identifiziert und beeinflusst somit die Distanzberechnung. Das linienbasierte Verfahren ist in der Lage die Richtung der Linien in der Y-Z-Ebene richtig zu rekonstruieren, sodass die horizontale Ausrichtung der Linien meist mit der Unterkante übereinstimmt. Die Tiefe der zwei Linienendpunkte zu schätzen ist jedoch erheblich schwieriger, sodass sich hohe Fehler von 75 - 80 cm für die Distanzberechnung ergeben.

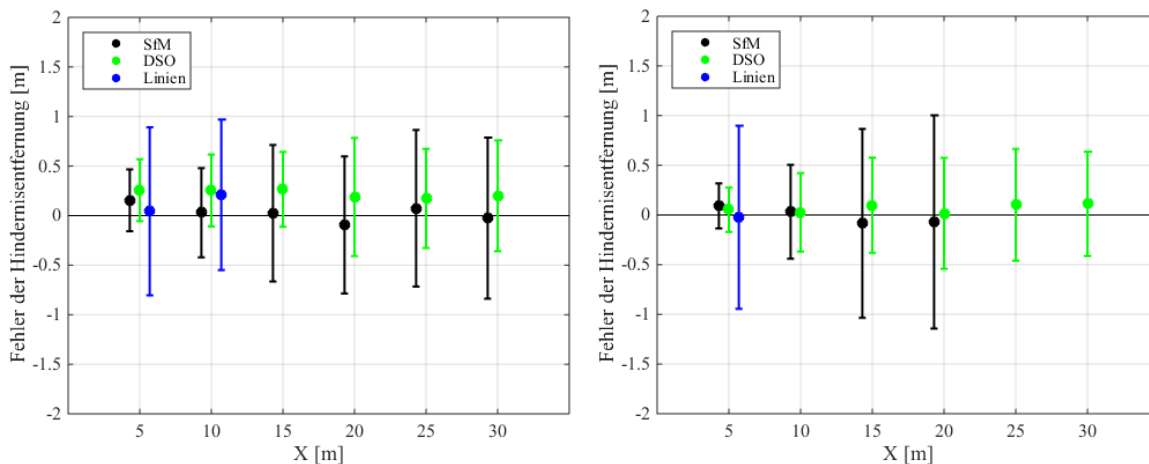


Abbildung 8-16 Analyse der Hindernisdetektion in Abhängigkeit von der Distanz zum Hindernis. Mittelwert und Standardabweichung der Distanzberechnung. Links: große Hindernisse; Rechts: kleine Hindernisse

Bei kleineren Hindernissen werden mit dem SfM-Verfahren nur sehr wenige Punkte rekonstruiert, welche stark schwanken. Auch hier kann erst ab 10 m eine genauere Distanzangabe mit einem Fehler zwischen 20 - 40 cm getätigt werden. Die Fehler beim DSO-Verfahren sind in großen Distanzen um ca. 50 % kleiner als beim SfM-Verfahren. Zudem werden die Hindernisse im Mittel nicht mehr überschätzt. Für Linien ergeben sich im Nahbereich für alle Hindernisgrößen ähnliche Fehler.

Zusammenfassend wird festgehalten, dass die maximalen Fehler beim SfM-Verfahren stark den zugrundeliegenden Hindernissen unterliegen. Bei kleinerem Hindernis und weniger verfolgten Punkten ergeben sich höhere Fehler. Erst im Nahbereich werden bei langsamer Fahrt ähnliche Genauigkeiten wie mit dem DSO-Verfahren erreicht. Dieses überschätzt bei hohen Hindernissen die Distanz. Der Standardabweichung ist jedoch nahezu unabhängig von der Hindernisgröße. Linien eignen sich bei dieser Aufnahmekonfiguration nicht zur Berechnung der Hindernisentfernung. Im Sinne der Warnstrategie wird empfohlen zur Berechnung der *TTC* die ermittelte Distanz um 1 m zu verkürzen. Damit kann davon ausgegangen werden, dass das Fahrzeug während der Notbremsung vor dem Hindernis zum Stehen kommt, sofern ideale Wetterbedingungen herrschen. Bei Regen sollte die Distanz aufgrund des niedrigen Reibwerts weiter verkürzt werden.

8.4.2.2 Hindernishöhe

Entscheidend für die Ermittlung der Passierbarkeit des Hindernisses ist die Berechnung der lichten Höhe, wobei die dabei erreichte Genauigkeit wiederum grundlegend für die nachgelagerte Warnstrategie ist. In diesem Abschnitt wird erneut die Trennung zwischen großen und kleinen Hindernissen vorgenommen, deren Auswertungen in Abbildung 8-17 dargestellt sind.

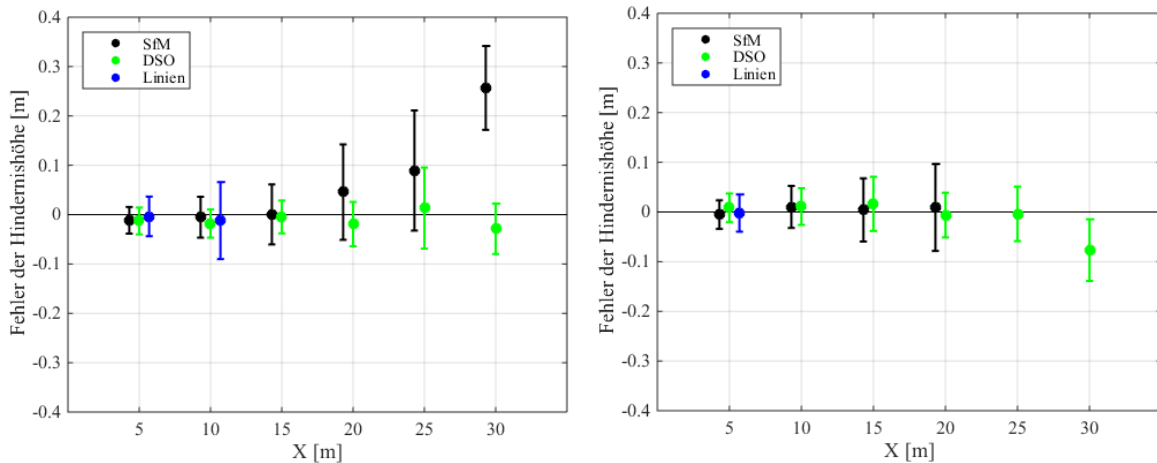


Abbildung 8-17 Analyse der Hindernisdetektion in Abhängigkeit von der Distanz zum Hindernis. Mittelwert und Standardabweichung der Höhenberechnung. Links: große Hindernisse; Rechts: kleine Hindernisse

Im linken Teil der Abbildung ist zu sehen, dass die Hindernishöhe innerhalb der SfM-Punktvolke in Entfernungen größer als 15 m tendenziell überschätzt wird. Dies ist entgegengesetzt der Vermutung, dass die Hindernisse aufgrund des ebenfalls überschätzten Höhenplans zu niedrig berechnet werden. Die Analyse zeigt, dass die SfM-3D-Punkte in dieser Entfernung auf dem Hindernis, nicht jedoch der Unterkante rekonstruiert werden. Erst bei Annäherung an das Hindernis erhöht sich die Punktdichte entscheidend für die Detektion der lichten Höhe. Diese wird in einem Abstand von 5 bzw. 10 m mit einem Fehler von 2,7 bzw. 4,1 cm berechnet. Ähnliche Werte mit dem DSO-Verfahren erreicht, allerdings liegt hier der wesentliche Vorteil auf der genaueren Detektion von Hindernissen in größeren Entfernungen. Mit DSO werden bereits ab 30 m eine ausreichende Anzahl an Punkten auf der Unterkante verfolgt. Die im Mittel zu niedrig geschätzte Höhe ist hierbei vermutlich auf den Offset des Höhenplans zurückzuführen. Die Standardabweichungen unter Nutzung der DSO-Punktvolke sind in Abständen über 20 m ca. 40 % kleiner als bei der SfM-Punktvolke. Die 3D-Linien liegen im Mittel auf der Unterkante des Hindernisses. Allerdings wird eine einzige Unterkante häufig durch mehrere kurze Linien repräsentiert, die sich in ihrer Höhe und Entfernung unterscheiden können. Dies kann, wie bereits in Abschnitt 5.3.3 geschildert, durch eine räumliche Filterung umgangen werden, allerdings sinkt dabei deutlich die Anzahl der identifizierten

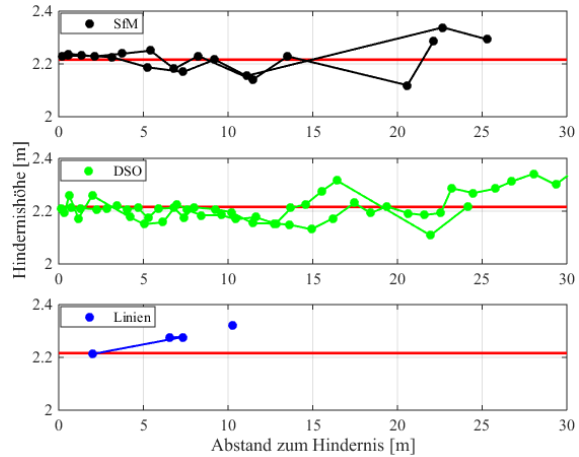


Abbildung 8-18 Zweifache Zufahrt auf eine freihängende Warnbake mit einer Höhe von 2.216 m. Zunächst erfolgt die Erkennung der größeren Schildunterkante. Bei Zufahrt auf das Hindernis schwankt die Höhe zwischen der Ober- und Unterkante der Warnbake. Die Hindernisdetektion mittels Linien schlägt fehl.

Linien. Bei kleineren Hindernissen wie Warnbaken zeigt sich, dass der Shi-Tomasi-Detektor im SfM-Verfahren erwartungsgemäß Merkmale an den Ecken der rotweißen Markierungen extrahiert und somit die Unterkante ab ca. 20 m Entfernung detektiert, in wenigen Ausnahmen auch darüber. Die Oberkante der häufig nur wenigen Zentimeter hohen Hindernisse wird im 3D-Raum jedoch ähnlich repräsentiert, sodass es häufig zu Schwankungen der lichten Höhe kommt (s. Abbildung 8-18). Im Nahbereich kann die Unterkante stabil verfolgt und der Fehler von anfangs 8.8 cm auf 2.9 cm reduziert werden. Ähnlich wie bei großen Hindernissen, wird innerhalb der DSO-Punktwolke der Höhe in Distanzen um die 30 m unterschätzt. Bei Annäherung an das Hindernis wird die Höhe im Mittel nur ca. 0.8 cm überschätzt und eine Standardabweichung von 2.9 cm erreicht. Linien, die ausschließlich innerhalb von 5 m geeignete Ergebnisse liefern, weisen einen Fehler von ca. 3.8 cm auf.

In Appendix C werden gesonderte Auswertungen einzelner Hindernisse gezeigt. Dabei wird die Zufahrt auf die Hindernisse unter verschiedenen Umgebungsbedingungen wiederholt, um somit eine Aussage über die Zuverlässigkeit der Höhendetektion zu treffen (vgl. Abbildung 8-18). Aus diesen Sequenzen geht hervor, dass die beiden punktbasierten Verfahren, vor allem jedoch DSO, am zuverlässigsten die Hindernisse detektieren. Sofern es zu einer Erstdetektion gekommen ist, wird das Hindernis in nahezu allen Fällen bis zum Unterfahren des Hindernisses verfolgt. Allerdings können lokale Schwankungen des Fahruntergrunds auftreten, welche zu Unterbrechungen während der Verfolgung zu führen. Da während der 3D-Raumunterteilung und der anschließenden Modellierung des Höhenplans keine Stabilisierung des Verfahrens durch Hinzunahme alter Zeitschritte und Modellierungen erfolgt, können vereinzelt Zellen auftreten, die den Höhenplan stark lokal beeinflussen. Besonders tritt dies bei Hindernissen wie Brückenunterführungen, Carports und Tunnel auf, bei denen in der Nähe der Einfahrt Seitenwände oder Geländer rekonstruiert werden, die anteilig in die Bodenschätzung eingehen und somit direkt am Hindernis einen Höhengsprung des Untergrundmodells verursachen. Die Begrenzung der auszuwertenden 3D-Punkte und Linien innerhalb des prädierten Fahrschlauchs, sofern in ausreichender Anzahl vorhanden, kann dem entgegenwirken. Die Auswertungen zeigen, dass mit punktbasierten Verfahren nahezu jedes Hindernis in einem Abstand von 15 - 20 m konstant verfolgt werden kann. Durch eine Kombination beider Verfahren und einer Stabilisierung der Funktion durch Hinzunahme älterer Zeitstempel wird davon ausgegangen, dass auch Abstände bis 25 - 30 m realisierbar sind. Das linienbasierte Verfahren eignet sich hingegen weder für eine frühe Erkennung des Hindernisses, noch zur Verfolgung.

Ausschließlich bei einer langen geradlinigen Zufahrt auf große Hindernisse, die sich stark vom Hintergrund unterscheiden, ist mit Linien eine stabile Hindernisdetektion möglich (s. Abbildung C-1). Dennoch sind bei allen Hindernistypen und Umgebungssituationen die punktbasierten Verfahren dem linienbasierten Verfahren zu bevorzugen.

8.5 Fazit

Wie bereits in Abschnitt 8.2.1 erwähnt, können mit dem indirekten Verfahren alle der insgesamt 72 Sequenzen und mit dem direkten Verfahren nur 51 Sequenzen rekonstruiert werden. Hauptursache sind dabei starke wechselnde Beleuchtungsänderungen, die eine bereits begonnene Rekonstruktion abbrechen lassen oder eine Relokalisierung nicht zulassen. Des Weiteren werden mit DSO keine Sequenzen initialisiert, die zu Beginn Rotationen aufweisen. Sofern von einer stabilen Umgebungsbeleuchtung ausgegangen werden kann, so ist die mit dem DSO-Verfahren rekonstruierte Punktwolke am geeignetsten zur Detektion von Durchfahrtshindernissen. Nach der Eliminierung von Ausreißern und Rauschen wird mit dem DSO-Verfahren ca. die doppelte Punktzahl im Vergleich zum SfM-Verfahren rekonstruiert. Die Punktverteilung auf Fahruntergrund und Umgebungsobjekten ist nahezu identisch, allerdings werden unter Nutzung des direkten Ansatzes Umrisse von Hindernissen und Kanten verfolgt, die im 3D-Raum eine hohe Punktdichte aufweisen und zumeist die niedrigsten Objekthöhe repräsentieren. Dies führt zu kleineren Höhenfehlern in nahezu allen Hindernisabständen. Zudem werden Hindernisse früher als solche identifiziert und effektiver verfolgt. Größter Einflussfaktor auf die Hindernisgenauigkeit ist neben der Detektion der eigentlichen Unterkante die Modellierung des Höhenplans, der aufgrund von Randbebauungen tendenziell überschätzt wird. Als zielführend wird die Kombination beider Verfahren gesehen, die die Stabilität der Rekonstruktion und Kameraschätzung von SfM mit der Punktzahl und -verteilung von DSO vereint. Das hier untersuchte linienbasierte Verfahren eignet sich ausschließlich, um bereits im Bild potentielle horizontale Hinderniskanten zu detektieren und die Punktverfolgung in einem ausgewählten Bereich darum zu konzentrieren. Auf eine Rekonstruktion der Linien kann im Rahmen dieser Anwendung verzichtet werden, da sich keine Vorteile in der frühzeitigen Identifizierung des Hindernisses oder Berechnung der lichten Höhe ergeben.

Unter der Annahme einer frühestmöglichen Detektion bei ca. 20 m Entfernung und der nur begrenzten Stabilität sowie Robustheit des Verfahrens, wird ein mögliches Einsatzszenario der Funktion allein im Geschwindigkeitsbereich der Park- und Rangierfunktionen gesehen. Für eine abschließende Betrachtung der Einsatzfähigkeit der Funktion im realen Straßenbetrieb werden weitere Untersuchungen und Experimente vorausgesetzt, welche die Detektionsgüte und Zuverlässigkeit final bewerten lassen.

9 Warnstrategie

Innerhalb dieses Kapitel wird eine potentielle Warnstrategie vorgestellt, die die Ergebnisse der Evaluierung berücksichtigt. Es wurde im letzten Kapitel gezeigt, dass die berechneten Hindernishöhen und -entfernungen stark schwanken können. Um die minimal für eine Notbremsung benötigte *TTC* einhalten zu können, ist es ratsam die Hindernisausgaben zu stabilisieren. Dies ermöglicht eine eindeutige Signalisierung der Kritikalität über eine Mensch-Maschine-Schnittstelle im Fahrzeug und den rechtzeitigen Eingriff bei nicht unterfahrbaren Hindernissen. Die Funktion der Durchfahrthöhenerkennung wird in die Klasse der Fahrerassistenzsysteme für frontalen Kollisionsschutz eingeteilt, welche sich nach [29] in drei verschiedenen Umsetzungsweisen unterscheidet.

Präventive Assistenz

Dabei wird eine Strategie der Unfallvermeidung beschrieben, welche die Wahrscheinlichkeit in kritische Situationen zu geraten verringert und damit die latente Gefahr herabsetzt. Dies kann durch Erhöhung des fahrdynamischen Handlungsspielraums oder durch eine Steigerung der Fahrerfähigkeit (z.B. Müdigkeitserkennung) erreicht werden.

Reaktionsunterstützung

Diese Strategie zur Unfallvermeidung ermöglicht es dem Fahrer in kritischen Situation rechtzeitig und richtig reagieren. Für Szenarien im Längsverkehr ist dabei die Unterstützung beim Verzögern im Fokus. Klassische Systeme sind bspw. das Antiblockiersystem ABS und das Stabilisierungsprogramm ESP.

Notmanöver

Bleibt eine rechtzeitige oder richtige Fahrerreaktion aus, so ermöglichen Notmanöver innerhalb der letzten Sekunde(n) vor einem möglichen Unfall einen „harten“ Eingriff. Dieser kann der Unfallvermeidung mindestens aber der Reduzierung der Unfallschwere dienen. Diese Art der Umsetzung eignet sich im Rahmen der hier vorgestellten Funktion. Beispielhaft sind mögliche Warnhinweise zur Darstellung im Kombiinstrument in Abbildung 9-1 gezeigt. Die Piktogramme sollen den Fahrer frühestmöglich auf ein potentielles Hindernis aufmerksam machen.

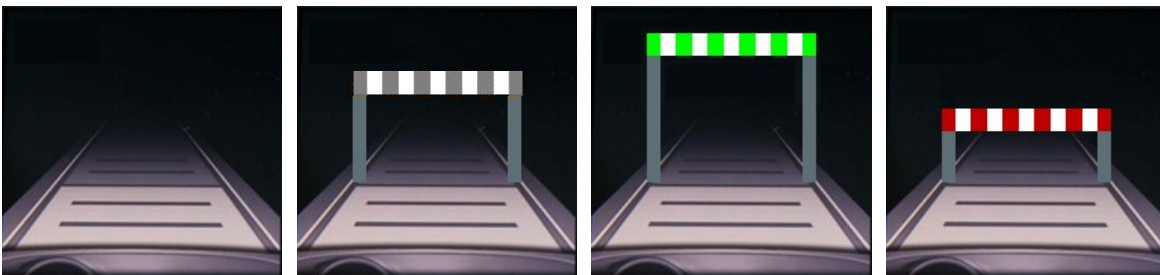


Abbildung 9-1 Warnhinweise zur Anzeige im Kombiinstrument. Ohne Balken: Kein Hindernis detektiert; Grauer Balken: Hindernis wurde detektiert und Höhe noch nicht zuverlässig bestimmt; Grüner Balken: Hindernis ist unterfahrbar; Roter Balken: Hindernis ist nicht unterfahrbar

Die Idee ist es mit einem zunächst graufarbenen Hindernisbalken ein detektiertes Hindernis zu symbolisieren, dessen Höhe bei großer Entfernung jedoch noch nicht zuverlässig bestimmbar ist. Anschließend soll mithilfe der in Abbildung 9-2 beschriebenen Hysterese die Unterscheidung in

Warnstrategie

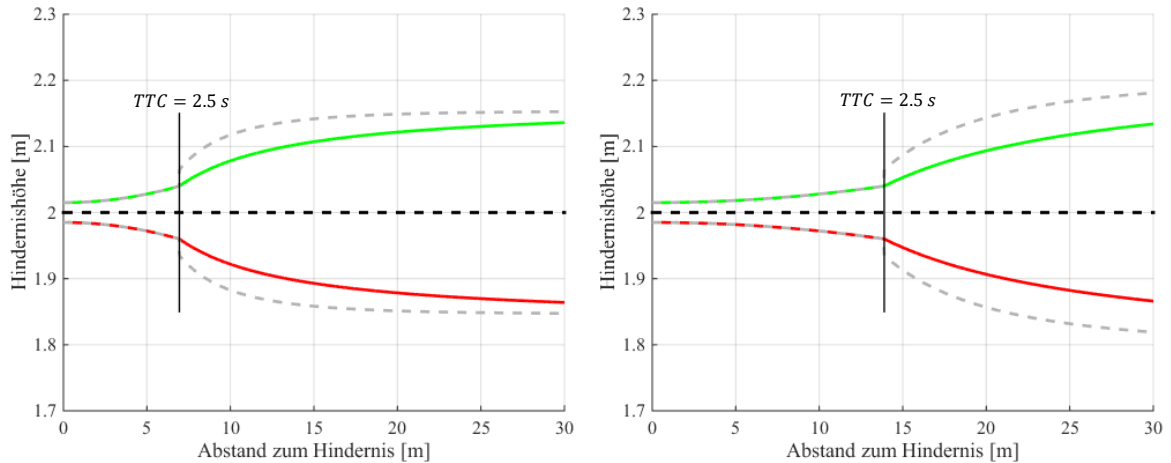


Abbildung 9-2 Hysterese zur Stabilisierung der Warnausgabe. Die Hindernishöhe wird mit 2 m und die TTC mit 2.5 s angenommen. In Grau wird die Unsicherheit dargestellt, Grün beschreibt das Limit für unterfahrbare Hindernisse - Rot für nicht unterfahrbare Hindernisse. Die Fahrzeughöhe ist in schwarz aufgetragen. Die Fahrzeuggeschwindigkeit beträgt links 10 km/h und rechts 20 km/h.

unterfahrbares (grün) und nicht unterfahrbares (rot) Hindernis getroffen werden. Die Hysterese dient der Stabilisierung der Hindernisausgabe und ihre Form hängt in erster Linie von den erreichbaren Genauigkeiten der Hindernisdetektion und der gefahrenen Geschwindigkeit ab. Sie ist so ausgelegt, dass die in früheren Zeitschritten detektierten Hindernishöhen zur Beurteilung der Kritikalität herangezogen werden und bei Erreichen der maximalen TTC von 2.5s eine Entscheidung über die Unterfahrbbarkeit des Hindernisses getroffen werden kann. Die Hysterese ist wie folgt zu lesen.

Bei Erstdetektion eines Hindernisses wird Anzeige direkt aus der Höhe abgeleitet. Ist es höher als der Verlauf der grünen Linie wird unterfahrbar ausgegeben. Ist es niedriger als der rote Verlauf, so erfolgt entsprechend die Ausgabe als nicht unterfahrbar. Der Bereich dazwischen gilt als nicht genau bestimmbar, sodass die graue Anzeige eingeblendet wird. Sofern dies geschieht, müssen zunächst die grauen Grenzwerte über- bzw. unterschritten werden, um auf die grüne bzw. rote Anzeige zu wechseln. Die Hysterese ist so ausgelegt, dass im Nahbereich eine Anzeige für unterfahrbar oder nicht unterfahrbar forciert wird. Die Standardabweichungen der Hindernisdetektion aus Abschnitt 8.4.2.2 zeigen jedoch, dass auch bei geringem Abstand zum Hindernis Restunsicherheiten bestehen. Die Entscheidung unterfahrbar/nicht unterfahrbar bei Hindernissen genau auf Fahrzeughöhe ist demzufolge nicht eindeutig festzulegen. Es wird empfohlen die automatische Bremsung auf dieses Hindernis zu initiieren.

10 Zusammenfassung

Moderne Fahrerassistenzsysteme durchdringen immer tiefgreifender den Markt und sind ein essentieller Bestandteil neuer Fahrzeuge. Neben Komfortfunktionen, die das Fahren erleichtern und das Fahrvergnügen steigern, sind es vor allem die Sicherheitsfunktionen, deren Wichtigkeit immens gestiegen ist. So wird das Vorhandensein eines automatischen Notbremsassistenten und eines Spurassistenten bereits heute in der Sicherheitsbewertung des Euro NCAP berücksichtigt. Viele dieser Systeme wurden ursprünglich für Lkw entwickelt und finden nun vorwiegend Anwendung in Pkw. Leichte Nutzfahrzeuge, die vorrangig von Speditionsunternehmen, Auslieferungsbetrieben und Fahrzeugvermietern betrieben werden, nutzen oft die adaptierten Assistenzsysteme aus dem Pkw-Bereich. Aufgrund ihrer Abmessungen und Tonnagen ergeben sich jedoch weitere Anwendungen, die spezielle Assistenzsysteme erfordern.

Im Rahmen dieser Arbeit wird ein neuartiges Fahrerassistenzsystem für (leichte) Nutzfahrzeuge entwickelt, welches mit einer möglichst kostengünstigen Hardware, die Durchfahrthöhe von potentiellen Hindernissen wie Parkhauseinfahrten, Tunnel und Schilderbrücken vermisst. Das System führt dabei den Abgleich zwischen Hindernis- und Fahrzeughöhe durch und warnt den Fahrer bzw. führt eine automatische Notbremsung aus, sofern eine Unfallgefahr besteht. Es besteht zum einen aus einer kostengünstigen Monokamera, die als Umfeldsensorik im Rahmen weiterer Assistenzsysteme bereits in den meisten Serienfahrzeugen hinter der Frontscheibe verbaut ist. Des Weiteren wird auf die Daten der Radtacksensoren und des Lenkwinkelgebers zurückgegriffen, um die Eigenbewegung des Fahrzeugs zu bestimmen. Die von der Kamera aufgenommene Bildsequenz wird genutzt, um mit ausgewählten 3D-Rekonstruktionsverfahren ein virtuelles Abbild der Fahrzeugumgebung zu erstellen, in dem nicht unterfahrbare Hindernisse identifiziert werden. Die Skalierung des erstellten 3D-Raums wird über die geschätzte Eigenbewegung vorgenommen. Das Auftreten von Kuppen und Wannsen wird über eine Schätzung des Fahruntergrunds modelliert, die Auswahl des kritischen Hindernisses erfolgt über eine Fahrschlauchprädiktion. Die Genauigkeit und Zuverlässigkeit der Hindernisdetektion ist von zahlreichen Faktoren abhängig, die im Laufe der Arbeit vorgestellt werden.

Grundlegend für die Funktion ist die verbaute Kamera. Im vorliegenden Fall wird sich für eine Industriekamera entschieden, deren Spezifikationen in etwa einer aktuellen Serienfahrzeugkamera entsprechen. Die Kalibrierung, d.h. die Bestimmung der externen und internen Orientierung steht zunächst im Fokus. Die externen Parameter werden mithilfe eines AICON 3D Systems vermessen und im Verlauf der Entwicklung als konstant angenommen. Als wesentliche Kenngrößen werden die Kamerahöhe und die Position in Fahrzeuginnenrichtung definiert, welche der späteren Bestimmung des Fahruntergrunds bzw. der Entfernungsberechnung zu etwaigen Hindernissen dienen. Im Hinblick auf die genaue Bestimmung der inneren Orientierung wird die 3D-Testfeldkalibrierung einer 2D-Testfeldkalibrierung vorgezogen. Die Verzeichnungsparameter werden in beiden Fällen nahezu identisch bestimmt, während sich der Fehler der Brennweiten (f_x, f_y) und des Bildhauptpunktes (x_0, y_0) hierbei um ca. 75/89 % bzw. 76/94 % verringert. Darüber hinaus werden hohe Korrelationen der Parameter untereinander vermieden. Die mechanische Instabilität der Kamera und die Umgebungseinflüsse führen jedoch bereits nach einem Tag im Fahrzeugbetrieb zu einer starken Veränderung der Parameter, sodass eine Selbstkalibrierung während der Fahrt empfohlen wird. Thermische Einflüsse spiegeln sich vor allem in einem nahezu gleichmäßigen Drift in allen Bildbereichen wider. Da die Höhenbestimmung der Hindernisse jedoch auf die relativen Abstände von zwei vertikal zu einander ausgerichteten Bildpunkten zurückzuführen ist, sind die Einflüsse des Drifts und der Verschiebung des Bildhauptpunktes auf den Höhenfehler vernachlässigbar gering. Der Einfluss der Brennweite wird entfernungsunabhängig mit ca. 0.2 cm/Pixel angegeben.

Im Rahmen der Untersuchung der Eigenbewegungsschätzung wird der entfernungsabhängige 3D-Punktfehler aufgrund der Maßstabsberechnung ermittelt. Die Analyse zeigt, dass die Nutzung des Zweispurmodells dem Einspurmodell bei der Bewegungsschätzung vorzuziehen ist. Weitere Verbesserungen können durch die Kalibrierung bzw. Parameteroptimierung der fahrdynamischen Einflussgrößen (Radabrollumfänge, Spurbreite, Radstand) erreicht werden, sodass sich der maximale relative Fehler des Bewegungsvektors zwischen zwei Bildern im vorliegenden Fall auf 2.8 cm bemisst. Die Referenzdaten werden hier einer dGPS-gestützten Inertialplattform OXTS RT3002 entnommen. Es gilt dabei die Annahme, dass die Funktion nicht im fahrdynamisch extremen Bereich zum Einsatz kommt und Fahrbahnunebenheiten sowie Schlupf vernachlässigt werden. Unter Annahme der Rekonstruktion einer ersten unskalierten 3D-Punktewolke nach Funktionsstart, ergibt sich nach einer Initialisierungsphase von 4 m, in der die Skalierung erfolgt, ein beispielhafter Höhenfehler von 0.15 cm bei einer Hindernishöhe von 2 m. Um etwaige nah zum Fahrzeug befindliche Hindernisse frühzeitig zu erkennen, kann die erste Skalierung bereits nach 2 m gefahrener Strecke erfolgen. In den vorliegenden Testsequenzen wurde die Skalierung jeweils einmalig am Anfang der Fahrt bestimmt und anschließend als konstant angenommen. Eine erneute Skalierung kann durchgeführt werden, sofern die Unterschiede von Radodometrie und visuellen Odometrie einen Schwellwert übersteigen. Eine häufige Maßstabsänderung kann jedoch die Stabilität des Verfahrens beeinträchtigen. Die Daten der Eigenbewegungsschätzung werden neben der Skalierung auch für die Prädiktion des horizontalen Fahrerschlauchs verwendet.

Basis für die Hindernisdetektion ist die Rekonstruktion des 3D-Abbilds der Fahrzeugumgebung. In der Arbeit wird zunächst ein indirektes, auf extrahierten Bildmerkmalen basierendes Verfahren namens *Structure from Motion (SfM)* vorgestellt. Die Untersuchung umfasst dabei die Bestimmung der Qualität und Berechnungsgeschwindigkeit verschiedener Merkmalsdetektoren und -deskriptoren. Da mit dem klassischen Deskriptormatching keine zufriedenstellende Anzahl an geometrischen Korrespondenzen (Inlier) erreicht worden ist und innerhalb der Sequenz von einer hohen Bildüberlappung ausgegangen werden kann, wird der bei aktuellen SLAM-Verfahren häufig genutzte KLT Tracker zur Merkmalsverfolgung untersucht. Aus den untersuchten Detektoren wird Shi-Tomasi für das weitere Verfahren herangezogen, da sich damit der vergleichsweise beste Kompromiss aus hoher mittlerer Merkmalsanzahl (1956 Merkmale), Matchingrate (92.6 %) und Tracklänge (20 Bilder) ergibt. Bei gleichzeitig geringem Rückprojektionsfehler (0.299 Pixel) und Rechenaufwand (36.69 ms/Bild) ist vor allem die gleichmäßige Merkmalsverteilung im Bild entscheidend für die Wahl des Detektors. So werden die Umrisse von Hindernissen als auch von Fahrbahnmarkierungen und Bordsteinkanten effektiv verfolgt und rekonstruiert. Neben dem merkmalsbasierten Verfahren wird ein weiteres direktes Verfahren vorgestellt. Bei *Direct Sparse Odometry (DSO)* werden einzelne Bildpunkte direkt anhand ihres Gradienten identifiziert und während der Rekonstruktion der 3D-Punkte über die Minimierung des photometrischen Fehlers verfolgt. Die Bildpunktauswahl wird im gesamten Bild durchgeführt, wobei der Ansatz vornehmlich Kanten detektiert. Zudem wird ein linienbasiertes Verfahren implementiert (*Line3D++*), welches über bekannte Kameraorientierungen die vom *Line Segment Detector* extrahierten Linien im Bild zu 3D-Linien rekonstruiert.

Ausgehend von den 3D-Rekonstruktionen der Verfahren, erfolgt im Anschluss die Begrenzung des 3D-Raums, Eliminierung von Rauschen und Ausreißern sowie die Unterteilung in eine Gitterstruktur. Dabei wird festgehalten, dass DSO eine hohe Punktzahl in großer Entfernung zur Kamera erstellt, die sich aufgrund der starken Verteilung nicht zur Hindernisdetektion eignen. Es wird davon ausgegangen, dass diese Punkte hohe Ungenauigkeiten aufgrund von fehlender Filterung während der Rekonstruktion aufweisen. Beim SfM-Verfahren werden potentiell ungenaue Punkte bereits während der Rekonstruktion aufgrund unzureichender Tracklänge und Schnittwinkel eliminiert.

Unter Berücksichtigung der Berechnung von Punktnormalen und der kalibrierten Kamerahöhe wird aus den verbleibenden Punkten der Fahruntergrund mit stückweise aneinandergereihten Flächen modelliert. Grenzwerte für maximale Steigungen werden den Entwurfsklassen für Landstraßen entnommen, sind jedoch für Privatgelände mit komplexeren Höhenplan geschwindigkeitsbasierend zu adaptieren. Die maximal zu detektierenden Entfernungen liegen bei beiden Punktverfahren bei 30 m. Beim SfM-Verfahren werden Genauigkeiten von ca. 3.8 cm im Nahbereich und 23.5 cm in großer Entfernung erreicht. Beim DSO-Verfahren liegen die Werte im Bereich von 4 - 17.1 cm. In beiden Punktwolken wird der Höhenplan tendenziell überschätzt, was auf Randbebauungen wie Schilderbefestigungen und Bordsteinkanten zurückzuführen ist. Die rekonstruierten 3D-Linien eignen sich nicht zur alleinigen Modellierung des Fahruntergrunds. Ausgehend vom bestimmten Höhenplan werden die Zellen der Gitterstruktur klassifiziert. Zellen, die dabei den Fahruntergrund schneiden werden als dieser klassifiziert. Umgebungsstruktur, wie bspw. Wände oder Fahrzeuge, haben direkten Kontakt zum Fahruntergrund und werden während der Hindernisdetektion ignoriert. Gitterzellen, die keinen indirekten oder direkten Bodenkontakt besitzen und eine Mindesthöhe erreichen, werden als potentielle Hindernisse klassifiziert. Kritische Hindernispunkte und -linien werden aus der Schnittmenge dieser Zellen und dem prädierten Fahrschlauch identifiziert.

Zur Berechnung der Hindernishöhe erfolgt zunächst eine Suche nach der untersten kritischen Hinderniszelle. Ausgehend von der bestimmten inneren Genauigkeit des Verfahrens, die den Fehler der Höhenberechnung hinsichtlich Kamerakalibrierung, Maßstabsberechnung, Fahruntergrundmodellierung und der daraus abgeleiteten Genauigkeit des 3D-Raums beinhaltet, werden entfernungsabhängige Detektionszonen definiert. Aus diesen Zonen werden die Hindernispunkte extrahiert, deren Höhe berechnet und die externe Genauigkeit bestimmt. Bei den Punktverfahren erweist sich der Mittelwert der identifizierten Punkte der unteren 5 cm, beim Linienverfahren der Mittelwert der Schwerpunkte aller Linien mit horizontaler Ausrichtung, als zielführend. Die untersuchten Hindernisse werden hinsichtlich ihrer Größe und Sichtbarkeit getrennt, um auf Unterschiede der Hindernisdetektion zwischen großen Hindernissen wie Brückenunterführungen oder Tunnelleinfahrten und kleinen Hindernissen wie freihängende Schilder oder Warnbaken hinzuweisen. Beim SfM-Verfahren werden große Hindernisse zunächst nicht an der Unterkante, sondern auf der Struktur darüber detektiert, sodass bei Distanzen größer als 15 m die Hindernishöhe im Mittel überschätzt wird. Erst im Nahbereich wird die Detektion der Unterkante zuverlässig und es werden Genauigkeiten bis zu 3 cm erreicht. Bei kleinen Hindernissen erfolgt die Detektion bis zu einer maximalen Entfernung von 20 m bei nahezu identischer Standardabweichung der Höhe. Die Entfernung der Hindernisse wird im Mittel sehr genau gemessen. Allerdings reduziert sich die Standardabweichung erst im Bereich unter 10 m Hindernisentfernung von vorher 60 - 80 cm auf 30 - 45 cm. Beim direkten Verfahren werden bei Hindernistypen in nahezu identischen Fehler in der Abstands- und Höhenberechnung erkannt. Hohe Hindernisse werden aufgrund ihrer Deckenkonstruktionen tendenziell zu weit entfernt geschätzt. Die mittleren Höhen werden aufgrund des Höhenplans eher unterschätzt. Die Standardabweichungen für die Entfernungen liegen im Bereich von 20 - 40 cm und für die Höhe zwischen 3 cm im Nahbereich und 6 cm auf größerer Distanzen. Aufgrund einer maximalen Erstdetektion von 10 m und Entfernungs- bzw. Höhenfehlern von 75 - 80 cm bzw. 5 - 8 cm, eignet sich das Linienverfahren ausschließlich zur Detektion von Hindernisunterkanten im Bild, sodass die Extraktion und Verfolgung von Punkten in diesen Bereichen verstärkt werden kann. Eine Kombination aus indirektem und direktem Verfahren wird zusammenfassend als zielführend für eine spätere Integration im Fahrzeug gesehen.

Im Hinblick auf die anfangs gestellten Anforderungen der Funktion der Durchfahrtshöhenerkennung kann, unter Verwendung der hier vorgestellten Verfahren, nicht die maximale Geschwindigkeit von 60 km/h realisiert werden. Die aktuelle Einsatzmöglichkeit der Funktion im realen Fahrzeug wird

Zusammenfassung

bei Geschwindigkeiten im Park- und Rangierbereich gesehen, wobei hier eine finale Bewertung der Detektionshäufigkeit und Zuverlässigkeit mithilfe von typischen Hindernissen und Szenarien, wie bspw. Parkhauseinfahrten, Carports oder Tiefgaragen, aussteht. Dennoch zeichnet sich die Funktion bereits jetzt durch eine kurze Initialisierungsphase von mind. 2 m und einer erreichten Höhengenaugigkeit von 3 cm im Nahbereich aus.

11 Ausblick

Die hier vorgestellten Konzepte zur Funktionsrealisierung einer Durchfahrtshöhenerkennung wurde mit Berücksichtigung aktueller Serienfahrzeugkameras anhand von Testsequenzen entwickelt und validiert. Dabei zeigt sich, dass ein günstiges Monokamerasystem prinzipiell geeignet ist, um im niedrigen Geschwindigkeitsbereich kritische Hindernisse zu identifizieren und deren Höhe so zu bestimmen, dass eine nachgelagerte Warnstrategie den Fahrer rechtzeitig informieren und ggf. eine automatische Notbremsung durchführen kann. Hier soll dennoch ein Überblick gegeben werden, welche Aspekte der Funktion mit Hinblick auf einen möglichen Serieneinsatz und die Erhöhung der Funktionsverfügbarkeit näher betrachtet werden sollten.

Zunächst wird hierbei auf den Einsatz einer monokularen Kamera verwiesen, die vor allem auch unter ungünstigen Beleuchtungsbedingungen, d.h. in der Dämmerung und Nacht sowie schnellen Beleuchtungsänderungen, verwertbare Bilder aufzeichnet. Des Weiteren ergibt sich mit höheren Öffnungswinkeln die Möglichkeit der Verfolgung von Hindernissen bis zu einem Abstand von wenigen Zentimetern. Dies gilt vor allem bei Fahrzeugen, die durch etwaige Dachaufbauten einen großen Abstand zwischen Fahrzeug- und Einbauhöhe der Kamera besitzen und daraufhin das Hindernis sich im Nahbereich aus dem Sichtbereich der Kamera bewegt. Darüber hinaus stabilisieren die verfolgten Merkmale in den äußeren Bildbereichen die Schätzung der Kamerapose.

In der entwickelten Funktion wird die durch die Radodometrie geschätzte Bewegung zur Skalierung der 3D-rekonstruierten Szene herangezogen. Es wird dabei davon ausgegangen, dass die externe Orientierung der Kamera im universellen Koordinatensystem konstant ist. Für Fahrzeuge mit abgeschirmter Kamerahalterung und zusätzlich steifem Fahrwerk ist dies zulässig, allerdings können je nach Beladung oder Fahrsituation Nick- und Wankbewegungen auftreten, die über die Radodometrie nicht abgebildet werden und somit die Maßstabsberechnung beeinflussen. Im Speziellen ist dies bei der Adaption der Funktion für Lkw zu untersuchen, deren Rotationen um die Fahrzeuglängs- und Querachse aufgrund des gefederten und gedämpften Führerhauses deutlich größer ausfallen können. Obwohl der Einfluss auf die Maßstabsberechnung als marginal eingeschätzt wird, können sich diese zusätzlichen Kamerabewegungen negativ auf das Tracking der Bildmerkmale auswirken, sodass eine höhere Anzahl an 3D-Punkte verworfen werden.

Im Hinblick auf die Genauigkeit und Zuverlässigkeit der Funktion ist eine möglichst dichte Rekonstruktion der Fahrzeugumgebung entscheidend. Dies kann zunächst unter Berücksichtigung der Rechenzeit über die Erhöhung der Merkmalsanzahl im Bild erreicht werden. Des Weiteren können unter Berücksichtigung von pixelgenau klassifizierten Bildern (s. Abbildung 11-1), die durch ein neuronales Netz der Funktion zur Verfügung gestellt werden, bereits vor der Rekonstruktion interessante Bereiche im Bild identifiziert werden. Mit dem Ziel die Hindernisse so dicht wie möglich zu rekonstruieren, werden diese Bereiche während der Merkmalsdetektion vornehmlich untersucht. Dies gilt z.B. für Bereiche mit dem Label Schilder, Gebäude und Natur (Bäume, Äste). Wiederum werden uninteressante Bereiche mit den Labels Fahrzeuge oder Fußgänger für die Merkmalsdetektion ausgeblendet. Diese können jedoch zur Analyse des zur Verfügung stehenden Sichtbereichs dienen und bspw. um die Funktion zu deaktivieren oder Hindernisausgaben zu unterdrücken. Fehlerhafte Warnungen, wie sie bei einer Fahrt hinter einem Lkw auftreten, können so vermieden werden. Die Label für den befahrbaren Bereich und Fahrbahnmarkierungen werden im Zuge der Fahrschlauchprädiktion verwertet. Im Allgemeinen kann festgehalten werden, dass sich mit den Auswertungen neuronaler Netze bereits sehr frühzeitig eine verbesserte Szeneninterpretation erreichen lässt, die das Verfahren robuster gegenüber Fehldetektionen macht.

Ausblick

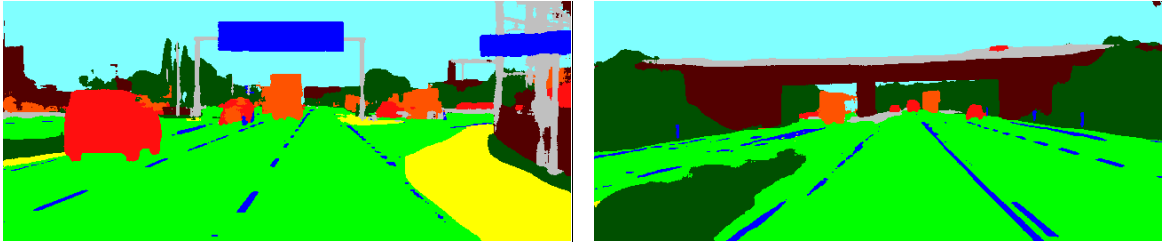


Abbildung 11-1 Pixelgenau klassifizierte Bilder aus einem neuronalen Netz.

Die Funktion wurde mit der Einschränkung auf die Nutzung eines Low-Cost Kamerasystems entwickelt, welches den Einsatz in aktuellen Fahrzeugen ermöglicht. Prinzipiell können die aufwendigen Rekonstruktionsverfahren vermieden werden, sofern eine Stereokamera mit fester Basisbreite zum Einsatz kommt. Eine vertikale Ausrichtung der beiden Kameras ist hierbei denkbar, jedoch hinsichtlich der Verwendbarkeit für weitere Fahrerassistenzsysteme zu bewerten. Darüber hinaus wird auf direktmessende Systeme verwiesen. Dabei bietet sich der Einsatz eines Solid State LiDARs an, der aufgrund des wegfallenden rotierenden Spiegels nicht nur kompakter, sondern auch kostengünstiger gegenüber aktueller Automotive-LiDAR-Hardware ist. Diese Systeme können unabhängig von der Umgebungsbeleuchtung 3D-Rekonstruktionen des Fahrzeugumfelds erschaffen und steigern somit die Funktionsverfügbarkeit. Im Sinne einer immer tiefgreifenderen Vernetzung der Fahrzeuge und Verkehrsinfrastruktur, können Hindernisse zudem ihre durch ein Fahrzeug oder händisch vermessene Höhe über den Kommunikationsstandard IEEE 802.11p oder per Mobilfunknetz an weitere Verkehrsteilnehmer übertragen werden.

Literaturverzeichnis

- [1] Statistisches Bundesamt (Destatis), „Polizeilich erfasste Unfälle - Unfälle und Verunglückte im Straßenverkehr,“ 2018. [Online]. Available: <https://www.destatis.de/DE/ZahlenFakten/Wirtschaftsbereiche/TransportVerkehr/Verkehrsunfaelle/Tabellen/UnfaelleVerunglueckte.html>. [Zugriff am 9 Januar 2019].
- [2] Verband der Automobilindustrie - VDA, „Automatisiertes Fahren,“ 2018. [Online]. Available: <https://www.vda.de/de/themen/innovation-und-technik/automatisiertes-fahren/automatisiertes-fahren.html>. [Zugriff am 9 Januar 2019].
- [3] J. Henn, „11foot8.com,“ 2009. [Online]. Available: <http://11foot8.com/faq/>. [Zugriff am 9 Januar 2019].
- [4] M. Damen, „Pakketbezorger rijdt zich klem onder viaduct Buikslotermeerplein,“ 2016.
- [5] Strava, 2015. [Online]. Available: <https://www.strava.com/athletes/5884091?hl=zh-CN>. [Zugriff am 9 Januar 2019].
- [6] B. Nguyen und I. Brilakis, „Understanding the problem of bridge and tunnel strikes caused by over-height vehicles,“ *Transportation Research Procedia* 14, pp. 3915-3924, 2016.
- [7] mux.de, „mux.de - Parkhaus Riem Arcaden,“ 2014. [Online]. Available: <https://www.mux.de/Parkhaus-Riem-Arcaden>. [Zugriff am 7 Januar 2019].
- [8] Osnabrücker Zeitung, „Endstation Eisenbahnbrücke: Fahrer vergisst Höhe seines Busses,“ 7 September 2010. [Online]. Available: <https://www.noz.de/lokales/emsbueren/artikel/319290/endstation-eisenbahnbruecke-fahrer-vergisst-hohe-seines-busses#gallery-jump-to>. [Zugriff am 7 Januar 2018].
- [9] B. Wittke, „Traktor fährt gegen Oranienburger Brücke,“ 17 September 2015. [Online]. Available: <http://www.maz-online.de/Lokales/Oberhavel/Traktor-faehrt-gegen-Oranienburger-Bruecke>. [Zugriff am 7 Januar 2018].
- [10] D. Bonica, „ABS News - Warning gantry on approach to Montague St bridge,“ 2018. [Online]. Available: <https://www.abc.net.au/news/2018-12-17/a-truck-approaches-a-low-bridge-height-warning-gantry-1/10626582>. [Zugriff am 23 Dezember 2018].
- [11] Trigg Industries International INC, „Over Height Vehicle Detection and Warning Systems (OHVDS),“ 2015.
- [12] CEOS, „Over-Height Vehicle Detection - TIRTL OHVD,“ 2017. [Online]. Available: <http://www.ceos.com.au/index.php/solutions/over-height-vehicle-detection>. [Zugriff am 23 Dezember 2018].
- [13] A. Singhal, „LaRa-OHVD: An Innovative Over-Height Vehicle Detection System to Protect our Bridges to Prosperity,“ 2015.
- [14] B. Nguyen und I. Brilakis, „Real-time validation of vision-based over-height vehicle detection system,“ *Advanced Engineering Informatics*, Bd. 38, pp. 67-80, 2018.

- [15] F. Dai, M. Park, M. Sandidge und I. Brilakis, „A vision-based method for on-road truck height measurement in proactive prevention of collision with overpasses and tunnels,“ *Automation and construction*, Bd. 50, pp. 29-39, 2015.
- [16] M. Maghiar, M. Jackson und G. Maldonado, „Warning systems evaluation for overhead clearance detection,“ *No. FHWAGA-16-1521 Georgia Southern University*, 2017.
- [17] P. P. Arvin, „Truck clearance anti-collision device“. United States Patent US 5,389,912, 1995.
- [18] S. Walenta, „Verfahren und Vorrichtung zur Erkennung der Durchfahrtsmöglichkeit eines Fahrzeugs“. Europa Patent EP 2,897,116 A1, 2013.
- [19] H. Endo, „Obstacle detection system for use in vehicles“. United States Patent US 4,477,184, 1984.
- [20] F. Nugent, „Overhead obstacle avoidance system“. United States Patent US 8,207,836 B2, 2012.
- [21] J. Kole, „Vehicle overhead clearance detection system and method of operation“. United States Patent US 8,354,920 B2, 2013.
- [22] C. Paukert, „How future Jaguars, Land Rovers will help drivers avoid a close shave,“ 2016. [Online]. Available: <https://www.cnet.com/roadshow/news/jaguar-land-rover-overhead-clearance-assist-smart-bike-kayak-rack/>. [Zugriff am 4 Januar 2019].
- [23] A. Criminisi, I. Reid und A. Zisserman, „Single view metrology,“ *International Journal of Computer Vision*, Bd. 40, Nr. 2, pp. 123-148, 2000.
- [24] M. Momeni-k, S. C. Diamantas, F. Ruggiero und B. Siciliano, „Height estimation from a single camera view,“ *VISAPP*, Bd. 1, pp. 358-364, 2012.
- [25] F. A. Andaló, G. Taubin und S. Goldenstein, „Efficient height measurements in single images based on the detection of vanishing points,“ *Computer Vision and Image Understanding*, pp. 51-60, 2015.
- [26] G. P. Stein, O. Mano und A. Shashua, „Vision-based ACC with single camera: bounds on range and rate accuracy,“ *IEEE Proceedings: Intelligent vehicles symposium*, pp. 120-125, 2003.
- [27] A. Wedel, U. Franke, J. Klapstein, T. Brox und D. Cremers, „Realtime depth estimation and obstacle detection from monocular video,“ *Joint Pattern Recognition Symposium*, pp. 475-484, 2006.
- [28] M. I. Arenado, J. M. P. Oria, C. Torre-Ferrero und L. A. Rentería, „Monovision-based vehicle detection, distance and relative speed measurements in urban traffic,“ *IET Intelligent Transport Systems*, Bd. 8, Nr. 8, pp. 655-664, 2014.
- [29] H. Winner, S. Hakuli, F. Lotz und C. Singer, *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, 3 Hrsg., Berlin Heidelberg New York: Springer- Verlag, 2015.

- [30] N. Mokey, „A self-driving car in every driveway? Solid-state lidar is the key,“ 15 März 2018. [Online]. Available: <https://www.digitaltrends.com/cars/solid-state-lidar-for-self-driving-cars/>. [Zugriff am 18 Dezember 2018].
- [31] Oxford Technical Solutions, „RT GNSS-aided inertial measurement systems - User manual,“ 2018.
- [32] S. Rusinkiewicz und M. Levoy, „Efficient variants of the ICP algorithm,“ *IEEE Third International Conference on 3-D Digital Imaging and Modeling*, 2001.
- [33] FARO Technologies, Inc., „Faro Laser Scanner Focus3D X 130 HDR,“ 2016.
- [34] Velodyne LiDAR, Inc., „HDL-64e S3 High Definition LiDAR Sensor,“ 2013.
- [35] U. Kiencke, S. Dais und M. Litschel, „Automotive Serial Controller Area Network,“ *SAE Technical Paper (No. 860391)*, 1986.
- [36] A. Albert und W. Gerth, „Evaluation and Comparison of the Real-Time Performance of CAN and TTCAN,“ *9th CAN in Automation Conference, iCC, Munich*, 2003.
- [37] Elektrobit, „EB Assist ADTF,“ 2019. [Online]. Available: <https://www.elektrobit.com/products/automated-driving/eb-assist/adtf/>. [Zugriff am 15 Januar 2019].
- [38] D. Hoeim, A. Efros und M. Geometri, „context from a single image,“ *IEEE International Conference on Computer Vision (ICCV)*, Bd. 1, pp. 654-661, 2005.
- [39] A. Saxena, S. Chung und A. Ng, „Learning depth from single monocular images,“ *Advances in Neural Information Processing Systems*, Bd. 18, p. 1161, 2006.
- [40] F. Han und S. Zhu, „Bottom-up/top-down image parsing by attribute graph grammar,“ *IEEE International Conference on Computer Vision (ICCV)*, Bd. 2, pp. 1778-1785, 2005.
- [41] J. L. Schönberger und J.-M. Frahm, „Structure-from-Motion Revisited,“ *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [42] W. Changchang, „Towards linear-time incremental structure from motion,“ *2013 International Conference on 3D Vision - 3DV 2013*, pp. 127-134, 2013.
- [43] Y. Furukawa, B. Curless, S. M. Seitz und R. Szeliski, „Towards Internet-scale multi-view stereo,“ *CVPR10*, 2010.
- [44] Y. Furukawa und J. Ponce, „Accurate, dense and robust multiview stereopsis,“ *IEEE Trans Pattern Analysis and Machine Intelligence (PAMI)*, Bd. 32, Nr. 8, pp. 1362-1376, 2010.
- [45] R. Hartley und A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge: Cambridge University Press, 2004.
- [46] S. Abraham und W. Förstner, „Fisheye-stereo calibration and epipolar rectification,“ *ISPRS Journal of Photogrammetry & Remote Sensing*, Bd. 59, pp. 278-288, 2005.

- [47] D. C. Brown, „Close-Range Camera Calibration,“ *Photogramm. Eng* 37.8, pp. 855-866, 1971.
- [48] H. C. Longuet-Higgins, A Computer Algorithm for Reconstructing a Scene from Two Projections, Los Altos, CA: Kaufmann, 1981, pp. 61-62.
- [49] T. Luhmann, Nahbereichsphotogrammetrie, Berlin: Wichmann, 2010.
- [50] P. Moulon, P. Monasse und R. Marlet, „Adaptive Structure from Motion with a Contrario Model Estimation,“ in *2012 11th Asian Conference on Computer Vision (ACCV)*, Daejeon, Korea, Springer, 2012, pp. 257-270.
- [51] J. u. M. H. Bartelsen, „Orientation of image sequences acquired from UAVs and with GPS cameras,“ *Surv. Land Inf. Sci.*, pp. 151-159, 2010.
- [52] J. J. Moré, „The Levenberg-Marquardt algorithm: implementation and theory,“ *Numerical analysis*, pp. 106-116, 1978.
- [53] G. Navratil und M. Staudinger, „Die Anwendung statistischer Methoden in Vermessungswesen und GIS,“ 2006.
- [54] W. Förstner und B. P. Wrobel, *Photogrammetric Computer Vision. Statistics, Geometry, Orientation and Reconstruction*, 2016.
- [55] T. Luhmann und H.-G. Maas, „Industriephoto­grammetrie,“ in *Photogrammetrie und Fernerkundung - Handbuch der Geodäsie*, Springer-Verlag, 2017, p. 125.
- [56] K. Levenberg, „A method for the solution of certain non-linear problems in least squares,“ *Quarterly of applied mathematics*, Bd. 2, Nr. 2, pp. 164-168, 1944.
- [57] D. W. Marquardt, „An algorithm for least-squares estimation of nonlinear parameters,“ *Journal of the society for Industrial and Applied Mathematics*, Bd. 11, Nr. 2, pp. 431-441, 1963.
- [58] Z. Zhang, „A flexible new technique for camera calibration,“ *IEEE Transactions on pattern analysis and machine intelligence*, Bd. 22, 2000.
- [59] M. Adamczyk, P. Liberadzki und R. Sitnik, „Temperature Compensation Method for Digital Cameras in 2D and 3D Measurement Applications,“ *Sensors*, Bd. 18, Nr. 11, 2018.
- [60] M. J. Smith und E. Cope, „The effects of temperature variation on single-lens-reflex digital camera calibration parameters,“ *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Bd. 38, 2010.
- [61] P. Riekert und T. E. Schunck, „Zur Fahrmechanik des gummibereiften Kraftfahrzeugs,“ *Ingenieur-Archiv*, pp. 210-224, 1940.
- [62] D. Ammon, Modellbildung und Systementwicklung in der Fahrzeugdynamik, Stuttgart: Teubner, 1997.
- [63] T. Weise, „Global optimization algorithms - theory and application,“ *Self-published*, Bd. 2, 2009.

- [64] G. Nützi, S. Weiss, D. Scaramuzza und R. Siegwart, „Fusion of IMU and Vision for Absolute Scale Estimation in Monocular SLAM,“ *Journal of Intelligent and Robotic Systems*, Bd. 61, pp. 287-299, 2011.
- [65] B. M. Kitt, J. Rehder, A. D. Chambers, M. Schonbein, H. Lategahn und S. Singh, „Monocular visual odometry using a planar road model to solve scale ambiguity,“ 2011.
- [66] J. Grater, T. Schwarze und M. Lauer, „Robust scale estimation for monocular visual odometry using structure from motion and vanishing points,“ *IEEE Intelligent Vehicles Symposium*, pp. 475-480, 2015.
- [67] F. Ackermann, „Photogrammetrie: Grundlagen, Verfahren, Anwendungen,“ Bd. 7, p. 27, 1976.
- [68] O. Schreer, in *Stereoanalyse und Bildsynthese*, Berlin, Springer-Verlag, 2005, p. 11.
- [69] R. T. Collins und R. S. Weiss, „Vanishing point calculation as a statistical inference on the unit sphere,“ *ICCV*, 1990.
- [70] A. Almansa, A. Desolneux und S. Vamech, „Vanishing point detection without any priori information,“ *TPAMI*, Bd. 25, Nr. 4, pp. 502-507, 2003.
- [71] Y. W. Seo, „Detection and Tracking the Vanishing Point on the Horizon,“ *CME-RI-TR-14-07*, 2014.
- [72] O. Barinova, E. Lempitsky, E. Tretjak und P. Kohli, „Geometric image parsing in man-made environments,“ *ECCV*, 2010.
- [73] J. M. Coughlan und A. L. Yuille, „Manhattan world: Orientation and outlier detection by Bayesian inference,“ *Neural Computation*, Bd. 15, Nr. 5, pp. 1063-1088, 2003.
- [74] A. Vedaldi und A. Zisserman, „Self-similar sketch,“ *ECCV*, 2012.
- [75] P. Moghadam und J. F. Dong, „Road Direction Detection Based on Vanishing-Point Tracking,“ *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1553-1560, 2012.
- [76] J. Lezama, R. Grompone von Gioi, G. Randall und J. M. Morel, „Finding Vanishing Points via Point Alignments in Image Primal and Dual Domains,“ *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 509-515, 2014.
- [77] R. G. Von Gioi, J. Jakubowicz, J. M. Morel und G. Randall, „LSD: a line segment detector,“ *Image Processing On Line*, Bd. 2, pp. 35-55, 2012.
- [78] J. Pazhayampallil, „Free Space Detection with Deep Nets for Autonomous Driving,“ 2015.
- [79] J. Jao, S. Ramalingam, Y. Taguchi, Y. Miki und R. Urtasun, „Estimating Drivable Collision-Free Space from Monocular Video,“ *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 420-427, 2015.
- [80] The Mathworks Inc, „MATLAB and Autonomous Driving Toolbox Release 2018a,“ *Lane Detection*, 2018.

- [81] N. Arshad, K. Moon, S. Park und J. Kim, „Lane detection with moving vehicle using color information,“ *In World Congress on Engineering and Computer Science*, 2011.
- [82] T. Aung und M. H. Zaw, „Video Based Lane Departure Warning System using Hough Transform,“ *International Conference on Advances in Engineering and Technology*, pp. 29-30, 2014.
- [83] A. López, J. Serrat, C. Canero, F. Lumbreras und T. Graf, „Robust lane markings detection and road geometry computation,“ *International Journal of Automotive Technology*, Bd. 11, Nr. 3, pp. 395-407, 2010.
- [84] A. Guiducci, „Parametric model of the perspective projection of a road with application to lane keeping and 3D road reconstruction,“ *Computer Vision and Image Understanding*, Bd. 73, Nr. 3, pp. 414-427, 1999.
- [85] G. Wolf, A. Bracher und B. Bösl, „Straßenplanung,“ Bd. 8, 2013.
- [86] D. G. Lowe, „Object recognition from local scale-invariant features,“ *Proceedings of the seventh IEEE international conference on computer vision*, Bd. 2, pp. 1150-1157, 1999.
- [87] D. G. Lowe, „Distinctive images features from scale-invariant keypoints,“ *International journal of computer vision*, Bd. 60, Nr. 2, pp. 91-110, 2004.
- [88] E. Karami, M. Shehata und A. Smith, „Image Identification Using SIFT Algorithm: Performance Analysis Against Different Image Deformations,“ *Proceedings of the 2015 Newfoundland Electrical and Computer Engineering Conference*, 2015.
- [89] H. Bay, T. Tuytelaars und L. Van Gool, „Surf: Speeded up robust features,“ *Computer Vision-ECCV*, pp. 404-417, 2006.
- [90] C. Harris und M. Stephens, „A combined corner and edge detector,“ *Alvey vision conference*, Bd. 15, Nr. 50, pp. 10-5244, 1988.
- [91] H. P. Moravec, „Obstacle avoidance and navigation in the real world by a seeing robot rover,“ *STANFORD UNIV CA DEPT OF COMPUTER SCIENCE*, 1980.
- [92] J. Shi und C. Tomasi, „Good features to track,“ *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 593-600, 1994.
- [93] E. Rosten und T. Drummond, „Machine learning for hig-speed corner detection,“ *9th European conference on computer vision (ECCV)*, pp. 430-443, 2006.
- [94] N. Govender, „Evaluation of Feature Detection Algorithms for Structure from Motion,“ 2009.
- [95] K. Otsu, M. Otsuki, G. Ishigami und T. Kubota, „An Examination of Feature Detection for Real-Time Visual Odometry in untextured natural Terrain,“ *Robot Intelligence Technology and Applications 2012*, pp. 405-414, 2013.
- [96] Ó. M. Mozos, M. Ballesta und O. Reinoso, „Interest Point Detectors for visual SLAM,“ *Conference of the Spanish Association for Artificial Intelligence*, pp. 170-179, 2007.

- [97] F. Fraundorfer und D. Scaramuzza, „Visual Odometry: Part II: Matching, Robustness, Optimization and Applications,“ *IEEE Robotics & Automation Magazine*, Bd. 19, Nr. 2, pp. 78-90, 2012.
- [98] A. Canclini, M. Cesana, A. Redondi, M. Tagliasacchi, J. Anscenso und R. Cilla, „Evaluation of low-complexity visual feature detectors and descriptors,“ *IEEE 18th International Conference on Digital Signal Processing (DSP)*, pp. 1-7, 2013.
- [99] A. Alahi, R. Ortiz und P. Vandergheynst, „Freak: Fast retina keypoint,“ *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 510-517, 2012.
- [100] M. Riegler, „A study and comparison of feature matching,“ p. 16, 2015.
- [101] M. Calonder, V. Lepetit, C. Strecha und P. Fua, „Brief: Binary robust independent elementary features,“ *European conference on computer vision*, pp. 778-792, 2010.
- [102] M. Ballesta, A. Gil, O. Martinez Mosos und O. Reinoso, „Local descriptors for visual SLAM,“ 2007.
- [103] H. Fassold und J. Rosner, „A real-time GPU implementation of the SIFT algorithm for large-scale video analysis tasks,“ *Real-Time Image and Video Processing*, 2015.
- [104] B. D. Lucas und T. Kanade, „An iterative image registration technique with an application to stereo vision,“ pp. 674-679, 1981.
- [105] B. K. Horn und B. G. Schunck, „Determining optical flow,“ *Artificial Intelligence*, Bd. 17, Nr. 1-3, pp. 185-203, 1981.
- [106] G. Farneäck, „Two-frame motion estimation based on polynomial expansion,“ *Scandinavian conference on Image analysis*, pp. 363-370, 2003.
- [107] G. Medioni und R. Nevatia, „Matching Images Using Linear Features,“ *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Bd. 6, pp. 675-685, 1984.
- [108] G. Medioni und R. Nevatia, „Segment-based stereo matching,“ *Computer Vision, Graphics, and Image Processing*, Bd. 31, Nr. 1, pp. 2-18, 1985.
- [109] R. Deriche und O. Faugeras, „Tracking Line Segments,“ *European Conference on Computer Vision (ECCV)*, pp. 259-268, 1990.
- [110] N. Ayache und O. Faugeras, „Building, Registrating, and Fusing Noisy Visual Maps,“ *The International Journal of Robotics Research*, Bd. 7, Nr. 6, pp. 45-65, 1988.
- [111] M. Hofer, „Building with Lines: Efficient 3D Scene Abstraction for the Built Environment,“ 2016.
- [112] M. Hofer, M. Maurer und H. Bischof, „Line3D: Efficient 3D Scene Abstraction for the Built Environment,“ *German Conference on Computer Vision and Pattern Recognition (GCPR)*, pp. 237-248, 2015.

- [113] B. Micusik und H. Wildenauer, „Structure from Motion with Line Segments Under Relaxed Endpoint Constraints,“ *International Conference on 3D Vision (3DV)*, Bd. 1, pp. 13-19, 2017.
- [114] L. Zhang, C. Xu, K. M. Lee und R. Koch, „Robust and Efficient Pose Estimation from Line Correspondences,“ *Asian Conference on Computer Vision (ACCV)*, pp. 217-230, 2012.
- [115] I. Sobel und G. Feldman, „A 3x3 Isotropic Gradient Operator for Image Processing,“ *Stanford Artificial Intelligence Project (SAIL)*, 1968.
- [116] J. Canny, „A Computational Approach to Edge Detection,“ *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Bd. 8, Nr. 6, pp. 679-698, 1986.
- [117] D. H. Ballard, „Generalizing Hough Transform to Detect Arbitrary Shapes,“ Bd. 13, Nr. 2, pp. 111-122, 1981.
- [118] J. Matas, C. Galambos und J. Kittler, „Robust Detection of Lines Using the Progressive Probabilistic Hough Transform,“ *Computer Vision and Image Understanding (CVIU)*, Bd. 78, Nr. 1, pp. 119-137, 2000.
- [119] C. Akinlar und C. Topal, „EDLines: Real-Time Line Segment Detection by Edge Drawing,“ *18th IEEE International Conference on Image Processing (ICIP)*, pp. 2837-2840, 2011.
- [120] C. Schmid und A. Zisserman, „Automatic Line Matching across Views,“ *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 666-671, 1997.
- [121] H. Bay, V. Ferrari und L. van Gool, „Wide-Baseline Stereo Matching with Line Segments,“ *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Bd. 1, pp. 329-336, 2005.
- [122] B. Khaleghi, M. Baklouti und F. O. Karray, „SILT: Scale-Invariant Line Transform,“ *IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA)*, pp. 78-83, 2009.
- [123] Z. Wang, F. Wu und Z. Hu, „MSLD: A Robust Descriptor for Line Matching,“ *Pattern Recognition*, Bd. 42, Nr. 5, pp. 941-953, 2009.
- [124] M. Hofer, M. Donoser und H. Bischof, „Semi-Global 3D Line Modeling For Incremental Structure-from-Motion,“ *British Machine Vision Conference (BMVC)*, 2014.
- [125] C. Forster, M. Pizzoli und D. Scaramuzza, „SVO: Fast Semi-Direct Monocular Visual Odometry,“ *IEEE International Conference on Robotics and Automation (ICRA)*, 2014.
- [126] J. Engel, T. Schöps und D. Cremers, „LSD-SLAM: Large-Scale Direct Monocular SLAM,“ *European Conference on Computer Vision (ECCV)*, 2014.
- [127] T. Schöps, J. Engel und D. Cremers, „Semi-Dense Visual Odometry for AR on a Smartphone,“ *International Symposium on Mixed and Augmented Reality (ISMAR)*, 2014.
- [128] J. Engel, V. Koltun und D. Cremers, „Direct Sparse Odometry,“ 2016.

- [129] J. Civera, A. J. Davison und J. M. Martínez Montiel, „Inverse Depth Parametrization for Monocular SLAM,“ *IEEE TRANSACTIONS ON ROBOTICS*, Bd. 24, Nr. 5, pp. 932-945, 2008.
- [130] X. Ning, F. Li, G. Tian und Y. Wang, „An Efficient Outlier Removal Method for Scattered Point Cloud Data,“ *PloS one*, Bd. 13, Nr. 8, 2018.
- [131] O. Schall, A. Belyaev und H.-P. Seidel, „Robust filtering of noisy scattered point data,“ *IEEE Proceedings Eurographics/IEEE VBTC Symposium Point-Based Graphics*, pp. 71-144, 2005.
- [132] S. Hempel, „Geometrierekonstruktion von LIDAR-Punktwolken zur Echtzeitverarbeitung für autonome Fahrzeuge,“ 2009.
- [133] A. Elfes, „Sonar-Based Real World Mapping and Navigation,“ 1987.
- [134] H. Badino, U. Franke und R. Mester, „Free Space Computation Using Stochastic Occupancy Grids and Dynamic Programming,“ *Proceedings of the International Conference for Computer Vision*, 2007.
- [135] B. Leibe, O. H. Jafari und D. Mitzel, „Real-Time RGB-D Based People Detection and Tracking for mobile Robots and Head-Worn Cameras,“ 2014.
- [136] M. Bansal, S. H. Jung, B. Matei, J. Eledath und H. Sawheney, „A real-time pedestrian detection system based on structure and appearance classification,“ *2010 IEEE International Conference on Robotics and Automation*, pp. 903-909, 2010.
- [137] T. Richter, *Planung von Autobahnen und Landstraßen*, 2016.
- [138] A. Wedel, H. Badino, C. Rabe, H. Loose, U. Franke und D. Cremers, „B-Spline Modeling of Road Surfaces With an Application to Free-Space Estimation,“ *IEEE Transactions of Intelligent Transportation Systems*, Bd. 10, Nr. 4, pp. 572-583, 2009.
- [139] Deutscher Verkehrssicherheitsrat, „Leifaden Fahrbahnmarkierungen,“ *Schriftenreihe Verkehrssicherheit*, Bd. 17, 2014.

Abkürzungsverzeichnis

2D, 3D:	Zweidimensional, Dreidimensional
ABS:	Antiblockiersystem
ADTF:	Automotive Data and Time-Triggered Framework
AWV:	Anhaltewegverkürzung
BRIEF:	Binary Robust Independent Elementary Features
CAN:	Controller Area Network
CCD:	Charge-coupled Device
CPU:	Central Processing Unit
CMVS:	Clustering View for Multi-view Stereo
DoG:	Difference of Gaussian
DSO:	Direct Sparse Odometry
ESC:	Electronic Stability Control
FAST:	Features from Accelerated Segment Test
FREAK:	Fast retina keypoint
GNSS:	Globales Navigationssatellitensystem
GPS:	Global Positioning System
GPU:	Graphics Processing Unit
GSM:	Global System for Mobile Communications
ICP:	Iterative Closest Point
IMU:	Inertia Measurement Unit
KKS:	Kamerakoordinatensystem
KLT:	Kanade-Lucas-Tomasi (Tracker)
LiDAR:	Light Detection and Ranging
LVDS:	Low-Voltage Differential Signaling
LSD:	Line Segment Detector
LSD-SLAM:	Large-Scale Direct Monocular SLAM
MLS:	Mobiler Laserscanner
MVS:	Multiple View Stereo
RADAR:	Radio Detection and Ranging
RANSAC:	Random Sample Consensus
PKS:	Pixelkoordinatensystem
PMVS:	Patch-based Multi-view Stereo
RGB:	Rot, Grün, Blau
SfM:	Structure from Motion
SG:	Steuergerät
SIFT:	Scale-Invariant Feature Transform
SLAM:	Simultaneous Localization and Mapping
SURF:	Speeded Up Robust Features
SVO:	Semi-direct Visual Odometry
USK:	Universelles Sensorkoordinatensystem
TLS:	Terrestrischer Laserscanner
TTC:	Time To Collision
WKS:	Weltkoordinatensystem

Mathematische Notation

Im Sinne der 2D- und 3D-Repräsentation werden alle zweidimensionalen Koordinaten kleingeschrieben (x, y) und alle dreidimensionalen Koordinaten großgeschrieben (X, Y, Z). Darüber hinaus gilt:

x, X	Skalar
\mathbf{x}, \mathbf{X}	Vektoren, Matrizen
\mathbf{x}, \mathbf{X}	Vektoren, Matrizen in homogenen Koordinaten
x'	Definition der Bildnummer '=1, ''=2, ...
x_1	Aufzählung oder Präzisierung des Parameters x
\hat{x}	Projektion von x

Abbildungsverzeichnis

Abbildung 1-1 Beispiele für Kollisionen durch unzureichende Durchfahrtshöhe [3, 4, 5]	1
Abbildung 1-2 Beispiele passiver Warnsysteme. Links: Parkhauseinfahrt [7]; Mitte und Rechts: Brückenunterführung [8, 9].....	2
Abbildung 1-3 Beispiel Warnportale. Links: Reflektierende Stangen [10]; Rechts: Ausführung mit Ketten [11]	3
Abbildung 1-4 Laserbasiertes System von [11]	3
Abbildung 1-5 Schema: LaRa-OHVD [13]	4
Abbildung 1-6 Höhenbestimmung nach [15] v.l.n.r: Linienerkennung innerhalb einer Bounding Box, Auswahl der in Fahrtrichtung liegenden Linie mit kürzestem Abstand zur linken oberen Ecke der Bounding Box, Blob Detektion, Canny Edge Filter und Auswahl der untersten Linie, Abstandsberechnung zwischen oberer und unterer Linie	4
Abbildung 1-7 Systeme zur Berechnung der Durchfahrtshöhe. Links oben: Ultraschallbasiertes System [17]; Links unten: Zwei-Sensor System [18]; Rechts: LiDAR-basiertes System [19].....	5
Abbildung 1-8 Schema: Kamerabasierte Durchfahrtshöhenerkennung	7
Abbildung 1-9 Durchfahrtshöhenerkennung – Systemüberblick.....	10
Abbildung 2-1 OXTS RT3002 [31]	16
Abbildung 2-2 Referenzsensorik Laserscanner. Links: Faro Focus3D X 130 HDR [33]; Rechts: Velodyne HDL64e S3 [34]	18
Abbildung 2-3 Beispielhafter Vergleich von TLS (oben) und MLS (unten) anhand eines typischen Hindernisses mit Messpunkten Links, Mitte und Rechts	19
Abbildung 2-4 Vernetzungskonzept	21
Abbildung 2-5 Schema: Notbremsassistent	22
Abbildung 2-6 Definition von WKS und USK.....	24
Abbildung 2-7 Definition von KKS und PKS.....	25
Abbildung 2-8 Prinzip der Zentralperspektiven Kamera	27
Abbildung 2-9 Korrespondenzen durch Epipolargeometrie	30
Abbildung 2-10 Räumlicher Vorwärtsschnitt	33
Abbildung 3-1 Bestimmung der äußeren Orientierung im USK mithilfe eines AICON 3D Systems	38
Abbildung 3-2 Gemittelter Rückprojektionsfehler bei unterschiedlichen Kalibrierfällen	40
Abbildung 3-3 3D-Testfeld bestehend aus acht definierten Markerpositionen	41
Abbildung 3-4 Temperatureinfluss auf die Punktdetektion. Links: Abkühlkurve; Rechts: Änderung der Punktcoordinate 146 (Rot - Starttemperatur, Blau - Endtemperatur)	44
Abbildung 3-5 Temperatureinfluss auf die Punktdetektion. Zur besseren Visualisierung wird die Änderung der Bildkoordinaten mit dem Faktor 50 multipliziert und die Helligkeit des Schachbrettmusters erhöht.	45

Abbildung 3-6 Simulationsmodell zur Ermittlung der Genauigkeit der bestimmten Hindernishöhe unter Einfluss der Kamerakalibrierung.....	46
Abbildung 3-7 Einfluss der Brennweite (links) und des Bildhauptpunkts (rechts) auf die Höhenbestimmung für Hindernisse mit einer Höhe von 2 m.	47
Abbildung 4-1 Schema. Kinematisches Einspurmodell.....	52
Abbildung 4-2 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S.....	52
Abbildung 4-3 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S und kombinierter Ansatz 1S+2S	53
Abbildung 4-4 Extremwerte als Kandidaten für die Optimierung	55
Abbildung 4-5 Vergleich Rad- und Referenzodometrie. Einspurmodell 1S, kombinierter 1S+2S Ansatz und optimierter 1S+2S Ansatz.....	58
Abbildung 4-6 Relativer Fehler der Radodometrie. Einspurmodell 1S, kombinierter 1S+2S und optimierter 1S+2S Ansatz.....	59
Abbildung 4-7 Bewegungsvektoren der Radodometrie und visuellen Odometrie	60
Abbildung 4-8 Auswirkung des Skalierungsfehlers auf die Objektdistanz. Oben: Fehler für ausgewählte Objektdistanzen von 5-40 m; Unten: Vergrößerter Ausschnitt bis zu einem maximalen Fehler von 0.1 m.....	61
Abbildung 4-9 Fahrschlauchprädiktion. Links: Prädiktion bei geradlinigen Verlauf durch Extrapolation der momentanen Eigenbewegung möglich; Rechts: Falsche Prädiktion durch Extrapolation (gelb) - Verwendung von Bildinformationen nötig (grün)	63
Abbildung 4-10 Horizont und Fluchtpunkte nach [68]	63
Abbildung 4-11 Fluchtpunkt- und Horizontdetektion.....	64
Abbildung 4-12 Detektion der Fahrbahnmarkierungen. v.l.n.r: Ausgangsbild; Transformation in die Vogelperspektive, Binärbild, Approximation der Linienverläufe; Rücktransformation der Linien in das Ausgangsbild.....	65
Abbildung 4-13 Detektion der Fahrbahnmarkierungen. Links: Fehlgeschlagene Modellierung; Rechts: Erfolgreiche Erkennung und Modellierung der Fahrbahnränder.....	66
Abbildung 4-14 Eigenbewegungsbasierte Fahrschlauchprädiktion in der X/Y -Ebene. Die durchgezogene Linie beschreibt die Fahrzeugmitte. Die gestrichelten Linien stellen die Fahrzeugbreite dar. v.l.n.r: Lenkradwinkel 10° , 30° , 90° , 180°	67
Abbildung 5-1 Ungenauigkeit aufgrund der Schnittgeometrie. Links: Kleiner markierter Bereich durch großen Schnittwinkel; Rechts: Aufnahmeconfiguration ähnlich einer Bildfolge im Fahrzeug. Große Ungenauigkeit vor allem in Kamerablickrichtung	70
Abbildung 5-2 Aus den Ungenauigkeiten der Bildkoordinaten abgeleiteter Fehler (Hindernisentfernung 40 m) - Darstellung im Pixelkoordinatensystem. Links: Fehler in Fahrtrichtung mit einem Maximum von 4.56 m nahe des Bildhauptpunkts; Rechts: Fehler in vertikaler Richtung mit einem Maximum von 0.001 m am Bildrand.....	72
Abbildung 5-3 Optischer Fluss. Oben: Ausgangsequenz (links) und Methode nach Lucas und Kanade (rechts); Unten: Methode nach Horn und Schunck (links) sowie nach Farneäck (rechts)	82
Abbildung 5-4 Gegenüberstellung der durch die Detektoren extrahierten Merkmale	84

Abbildungsverzeichnis

Abbildung 5-5 Detektierte Merkmale in schwach beleuchteten Sequenzen. Links: Initiale Parametrierung; Rechts: Niedriger Schwellwert und Mindestdistanz	85
Abbildung 5-6 Verfolgte Merkmale während der Beleuchtungsänderung in zwei aufeinanderfolgenden Bildern.....	85
Abbildung 5-7 Zufahrt auf Brückenunterführung. Links: Bild der Sequenz; Rechts: Punktwolke der Frontansicht der Brückenunterführung nach der Initialisierungsphase (9 Bilder – Abstand zur Brücke ~32 m).	86
Abbildung 5-8 3D-Punktwolken der Zufahrt auf die Brückenunterführung. Links: Punktwolke nach 29 Bildern; Rechts: Punktwolke nach 49 Bildern	87
Abbildung 5-9 Merkmalsverfolgung bei geringer Umgebungsbeleuchtung; Links: Keine Detektion in dunklen homogenen Bereichen; Rechts: Unterkante der Brücke wird nicht identifiziert.....	87
Abbildung 5-10 Kantenbilder. Oben links: Sobel; Oben rechts: Canny; Unten links: Probabilistische Hough Linien; Unten rechts: Line Segment Detector	89
Abbildung 5-11 Erstellen einer Linienkorrespondenz mit Berücksichtigung der Epipolargeometrie nach [112]. Links: Detektierte Linie lk' im Bild I' mit den Endpunkten pk' und qk' ; Rechts: Detektierte und ins Unendliche verlängerte Linie lk'' im Bild I'' . Schnittpunkte xp und xq der Epipolarlinien (grün) mit der unendlichen Linie, sowie die Endpunkte pk'' und qk'' .	90
Abbildung 5-12 Linienrekonstruktion.....	91
Abbildung 5-13 3D-Hypothesen $H' \rightarrow ''$ (gelb) und $H'' \rightarrow '$ (rot) und die gemeinsame 3D-Linie L . Entnommen aus [124].	91
Abbildung 5-14 Rekonstruktion der 3D-Linien (Blau) mit Fahrzeugtrajektorie (Grün).	92
Abbildung 6-1 Initialisierung eines 3D-Punktes mit $\rho' = 1$ und hoher Varianz im Bild I' . Anschließende Optimierung der inversen Tiefe und Varianz durch die nachfolgenden Bilder I'' und I'''	97
Abbildung 6-2 Direktes Verfahren. 3D-Rekonstruktion mit Fahrzeugtrajektorie (Grün)	98
Abbildung 7-1 Unterteilung und Klassifizierung des 3D-Punktraums. Grün: Fahruntergrund; Blau: Struktur/ Objekte; Orange: Potentielle Hindernisse; Rot: Hindernis im Fahrschlauch.....	99
Abbildung 7-2 Fehler in der 3D-Punktwolke. Grün: Ausreißer bzw. Detektionsfehler; Blau: isolierte Cluster	100
Abbildung 7-3 Isoliertes Cluster aufgrund des bewegten Objekts. Links: Originalbild mit Fokus auf das Fahrzeug auf der gegenüberliegenden Fahrspur; Rechts: Rekonstruktion von Fahrzeugpunkten auf der eigenen Fahrspur.....	101
Abbildung 7-4 Zufahrt auf freihängende Warnschilder zu verschiedenen Zeitpunkten. Links: $I=30$; Mitte: $I=60$; Rechts: $I=90$	103
Abbildung 7-5 Unterteilung und Filterung von 3D-Punkten. Oben: Originale 3D-Punktwolke (SfM); Unten: Raumunterteilende Zellstruktur mit Würfeln der Kantenlänge von 0.25 m.....	104
Abbildung 7-6 Berechnung der Fahruntergrundhöhe und Klassifizierung der Bodenzellen (grün).	105
Abbildung 7-7 Eliminierung falscher Fahruntergrundkandidaten.	106

Abbildung 7-8 Delaunay Triangulation des Fahruntergrunds. Oben links: Aufnahme der Umgebung; Oben rechts: Identifizierte Kandidaten des Bodens und gefahrene Trajektorie; Unten links: Delaunay Triangulation; Unten rechts: Detailaufnahme des Bordsteins	107
Abbildung 7-9 Delaunay Triangulation des Fahruntergrunds. Oben links: Aufnahme der Umgebung; Oben rechts: Identifizierte Kandidaten des Bodens und gefahrene Trajektorie; Unten links: Delaunay Triangulation; Unten rechts: Detailaufnahme des Ausreißers.....	108
Abbildung 7-10 Extrahierter Fahruntergrund. Schwarz: Gefilterte Punktwolke; Grün: Geschätzter Fahruntergrund an jeder Stützstelle	109
Abbildung 7-11 Klassifizierte Zellstruktur. Weiß: Unklassifiziert; Grün: Zellen des Fahruntergrunds.....	110
Abbildung 7-12 Ausgewählte Sequenzen zur Fahrbahnmodellierung. Links: Kurvenfahrt; Rechts: Auffahrt zu einem Parkplatz.....	110
Abbildung 7-13 Fahrtuntergrundmodellierung und Klassifizierung ausgewählter Sequenzen. Oben: Kurvenfahrt; Unten: Auffahrt zu einem Parkplatz	111
Abbildung 7-14 Klassifizierte Zellstruktur. Grün: Fahruntergrund; Blau: Struktur; Orange: Potentielles Hindernis; Rot: Identifizierte Hinderniszellen innerhalb des prädizierten Fahrschlauchs; Blauer Rahmen: Rekonstruiertes Verkehrsschild ohne Rohrpfeiler	112
Abbildung 7-15 Klassifizierte 3D-Gitterstruktur ausgewählter Sequenzen.	113
Abbildung 8-1 Vergleich der absoluten Position und Rotation von Rad- und visueller Odometrie (SfM, DSO) sowie der Referenz (IMU)	114
Abbildung 8-2 Relativer Fehler der einzelnen Dimensionen von Radodometrie und visueller Odometrie (SfM, DSO)	115
Abbildung 8-3 Vergleich der verfolgten Bildpunkte. Links: Indirektes SfM-Verfahren mit Shi-Tomasi Detektor; Rechts: Direktes DSO-Verfahren	117
Abbildung 8-4 Fehlgeschlagene Initialisierung/Rekonstruktion mit dem direktem Ansatz. Links: Helle Umgebungsbeleuchtung und starke Beleuchtungsänderung; Mitte: Regen, Scheibenwischer, Reflektionen auf der nassen Fahrbahn; Rechts: Keine Objekte im Bildvordergrund.....	118
Abbildung 8-5 Testsequenzen zur Fahrbahnrekonstruktion. Links: Asphaltierte Straße mit Markierungen, Bordsteinkanten und Schattenwurf; Mitte: Gepflasterte Fahrbahn; Rechts: Geringe Umgebungsbeleuchtung	119
Abbildung 8-6 Fahrbahnrekonstruktion (Draufsicht) und maximaler Abstand D zu Bodenpunkten bei $I = 100$. Links: SfM; Rechts: DSO; Oben: Asphaltierte Straße mit Markierungen, Bordsteinkanten und Schattenwurf; Mitte: Gepflasterte Fahrbahn; Unten: Geringe Umgebungsbeleuchtung	120
Abbildung 8-7 Ausgewählte Hindernisse. Links: Einfahrt zu einem Parkdeck; Mitte: Tunnelleinfahrt; Rechts: Freihängende Warnschilder und -baken	121
Abbildung 8-8 3D-Punktwolke der Hindernisse. Links: SfM; Rechts: DSO; Oben: Einfahrt zu einem Parkdeck; Mitte: Tunnelleinfahrt; Unten: Freihängende Warnschilder und -baken.....	122
Abbildung 8-9 Linienbasierte Rekonstruktion des Fahruntergrunds. Oben: Vornehmlich quer zur Fahrbahn rekonstruierte Linien; Unten: Rekonstruktion von Schlagschatten.....	123

Abbildungsverzeichnis

Abbildung 8-10 Linienbasierte Rekonstruktion von Hindernissen. Oben: Rekonstruktion der wahren Brückenunterkante, Fehlschlagen der Detektion mittels SfM-Verfahren; Mitte: Identifizierung von zwei räumlich getrennten Unterkanten; Unten: Gesamtheitliche Rekonstruktion des Portals durch Linien.....	125
Abbildung 8-11 Fehler der Fahrbahnmodellierung aller Sequenzen. Links: Mittelwert und Standardabweichung des Höhenfehlers in der SfM- bzw. DSO-Punktwolke; Rechts: Exemplarische Verteilung bei 5 m (SfM).....	126
Abbildung 8-12 Gefilterter 3D-Punkt- und Linienraum mit Trajektorie (hellgrün), Fahruntergrund (dunkelgrün) und identifizierten Hindernispunkten innerhalb des Fahrschlauchs. Oben: SfM und Linien; Unten: DSO	128
Abbildung 8-13 Manuelle Auswahl des Hindernisses. Links: Draufsicht einer SfM-Punktwolke, Trajektorie (grün), Fahrschlauch (grün - gepunktet) und detektierte Linien (blau); Rechts: Vergrößerter Ausschnitt des Hindernisses mit manuell gesetzten Marken (cyan-violett) und berechnetem Schnittpunkt mit der Trajektorie (rot).....	129
Abbildung 8-14 Ausgewählte Sequenzen mit Durchfahrtshindernissen. Oben: große gut sichtbare Hindernisse; Unten: kleine freihängende Hindernisse	131
Abbildung 8-15 Punkthäufigkeiten in Abhängigkeit von der Distanz zum Hindernis. Links: große Hindernisse; Rechts: kleine Hindernisse.....	131
Abbildung 8-16 Analyse der Hindernisdetektion in Abhängigkeit von der Distanz zum Hindernis. Mittelwert und Standardabweichung der Distanzberechnung. Links: große Hindernisse; Rechts: kleine Hindernisse.....	132
Abbildung 8-17 Analyse der Hindernisdetektion in Abhängigkeit von der Distanz zum Hindernis. Mittelwert und Standardabweichung der Höhenberechnung. Links: große Hindernisse; Rechts: kleine Hindernisse.....	133
Abbildung 8-18 Zweifache Zufahrt auf eine freihängende Warnbake mit einer Höhe von 2.216 m. Zunächst erfolgt die Erkennung der größeren Schildunterkante. Bei Zufahrt auf das Hindernis schwankt die Höhe zwischen der Ober- und Unterkante der Warnbake. Die Hindernisdetektion mittels Linien schlägt fehl.....	134
Abbildung 9-1 Warnhinweise zur Anzeige im Kombiinstrument. Ohne Balken: Kein Hindernis detektiert; Grauer Balken: Hindernis wurde detektiert und Höhe noch nicht zuverlässig bestimmt; Grüner Balken: Hindernis ist unterfahrbar; Roter Balken: Hindernis ist nicht unterfahrbar	136
Abbildung 9-2 Hysterese zur Stabilisierung der Warnausgabe. Die Hindernishöhe wird mit 2 m und die TTC mit 2.5 s angenommen. In Grau wird die Unsicherheit dargestellt, Grün beschreibt das Limit für unterfahrbare Hindernisse - Rot für nicht unterfahrbare Hindernisse. Die Fahrzeughöhe ist in schwarz aufgetragen. Die Fahrzeuggeschwindigkeit beträgt links 10 km/h und rechts 20 km/h.	137
Abbildung 11-1 Pixelgenau klassifizierte Bilder aus einem neuronalen Netz.	143

Tabellenverzeichnis

Tabelle 1-1 Anforderungen an das Gesamtsystem	8
Tabelle 2-1 Vergleich der Umfeldsensorik	15
Tabelle 2-2 Kamerakonfiguration	16
Tabelle 2-3 Messgenauigkeit OXTS RT3002 [31].....	17
Tabelle 2-4 Vergleich des TLS und MLS	18
Tabelle 2-5 Vergleich der berechneten Hindernishöhen zwischen TLS und MLS	20
Tabelle 3-1 Äußere Orientierung der Kamera im Fahrzeug-USK.....	39
Tabelle 3-2 Kalibrierfälle mit 2D-Testfeld.....	39
Tabelle 3-3 Parameter der Kalibrierung (2D- Testfeld)	40
Tabelle 3-4 Parameter der Kalibrierung (3D- Testfeld)	42
Tabelle 3-5 Korrelationen der inneren Orientierung. Oben: Nach Kalibrierung mit 2D-Testfeld; Unten: Nach Kalibrierung mit 3D-Testfeld ohne bekannte Markerpositionen	43
Tabelle 3-6 Zeitliche Veränderung der Kalibrierparameter	43
Tabelle 4-1 Maßstabsabhängiger Fehler [cm] bei unterschiedlichen Basislängen und festen Objektdistanzen	62
Tabelle 5-1 Merkmalsdetektoren und ihre Eigenschaften (Bewertung nach [97]).....	75
Tabelle 5-2 Vergleich der Detektoren	75
Tabelle 5-3 Berechnungszeiten der Deskriptoren [ms]	78
Tabelle 5-4 Gemeinsame (oben) und geometrische (unten) Korrespondenzen zwischen einem Bildpaar [%]	79
Tabelle 5-5 Mittelwerte der Rückprojektionsfehler [Pixel].....	79
Tabelle 5-6 Mittelwerte der Tracklänge [Bilder].....	80
Tabelle 5-7 Vergleich der Merkmalsdetektoren unter Nutzung des KLT Trackers	83
Tabelle 8-1 Gemittelte Anzahl an 3D-Punkten beim indirekten und direkten Verfahren in den untersuchten Sequenzen	119
Tabelle 8-2 Übersicht der Fehler der Fahrbahnmodellierung aller Sequenzen [cm].....	126
Tabelle 8-3 Resultierender Höhenfehler des Hindernisses in ausgewählten Distanzen zur Bestimmung der Ausdehnung der Evaluierungszone.....	130

Appendix A Kamerakalibrierung

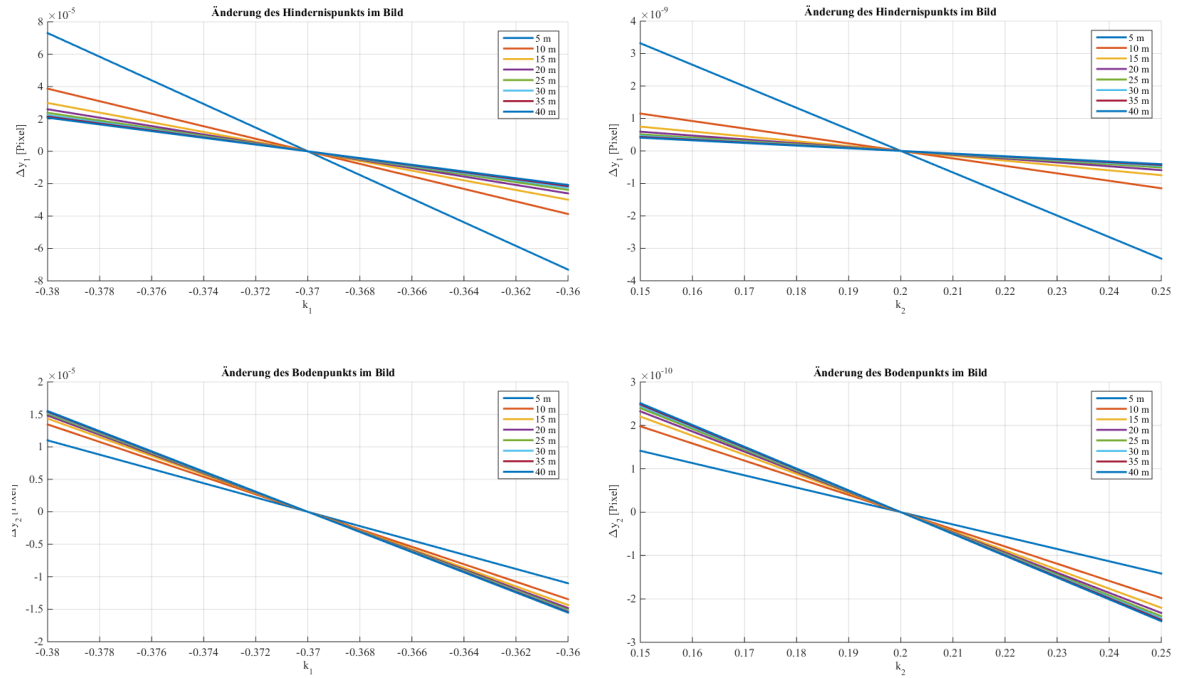


Abbildung A-1 Einfluss der Verzeichnungsparameter k_1 und k_2 auf die Bildpositionen für Hindernisse mit einer Höhe von 2 m und einer Entfernung zwischen 5 m und 40 m

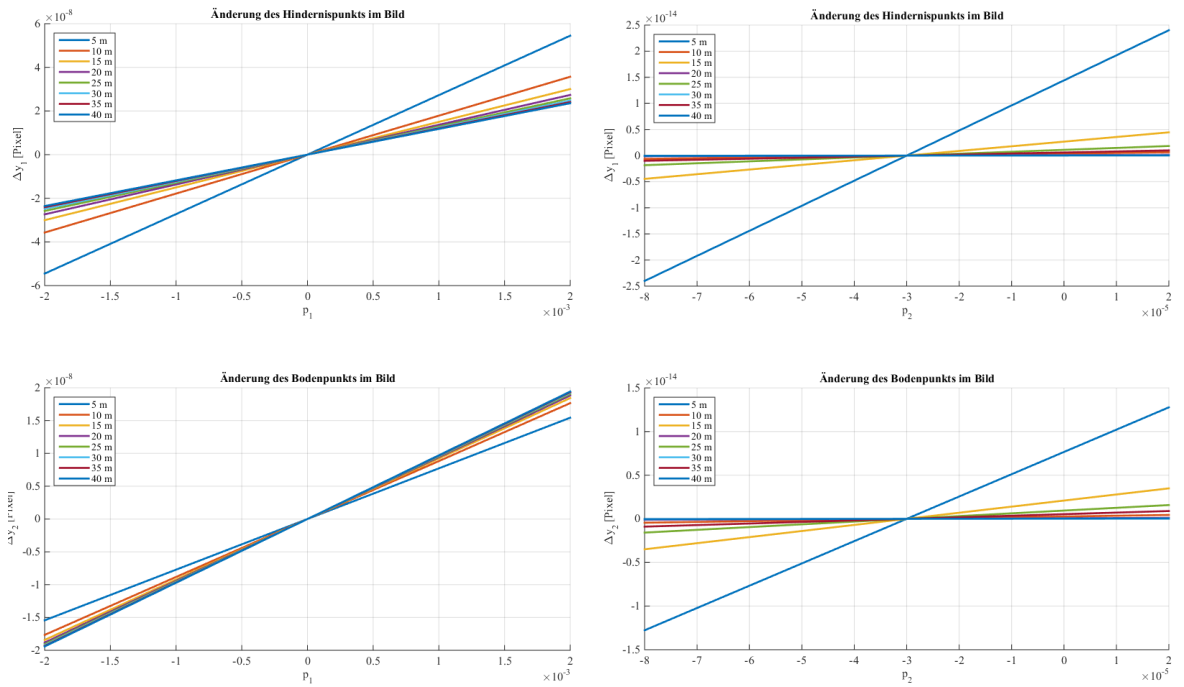


Abbildung A-2 Einfluss der Verzeichnungsparameter p_1 und p_2 auf die Bildpositionen für Hindernisse mit einer Höhe von 2 m und einer Entfernung zwischen 5 m und 40 m

Appendix B Fahruntergrundmodellierung

Die Fahruntergrundmodellierung wird anhand von ausgewählten Sequenzen entwickelt, die eine umfassende Repräsentation von Straßenoberflächen, Steigungen, Gefällen und Kurven beinhaltet. Die Ergebnisse von neun dieser Sequenzen werden hier dargestellt.

Asphalt – Markierungen 1

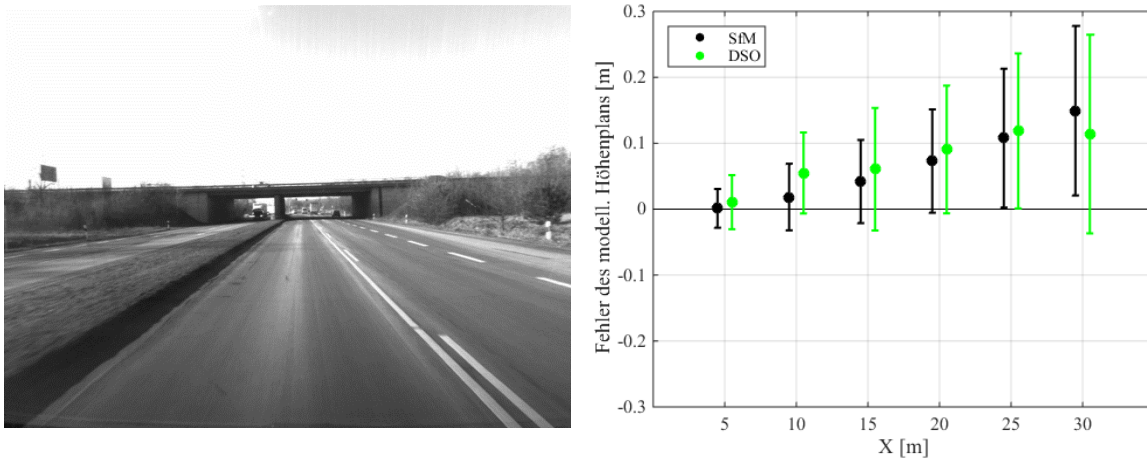


Abbildung B-1 Fehler der Fahrbahnmodellierung. Asphalt – Markierungen 1

Asphalt – Markierungen 2

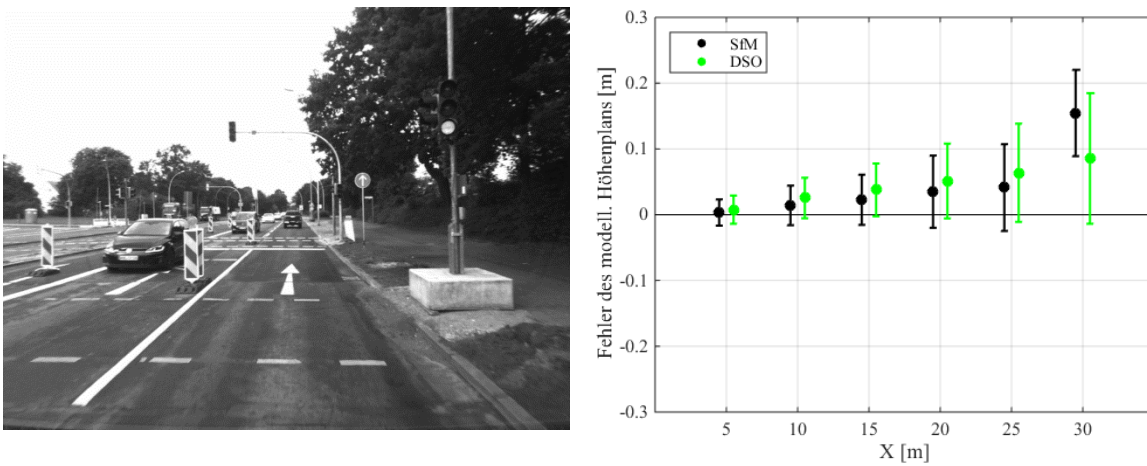


Abbildung B-2 Fehler der Fahrbahnmodellierung. Asphalt – Markierungen 2

Asphalt – Markierungen 3

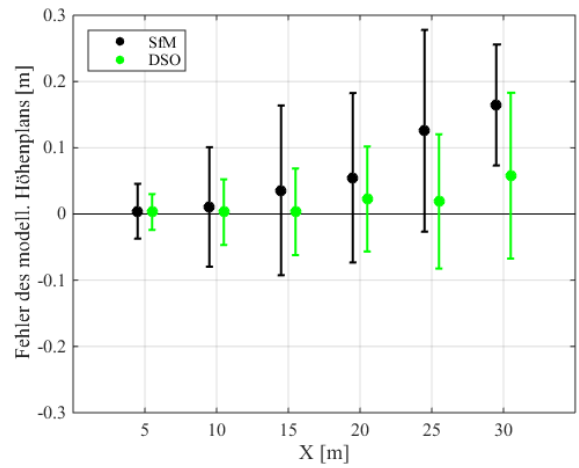


Abbildung B-3 Fehler der Fahrbahnmodellierung. Asphalt – Markierungen 3

Asphalt – Markierungen 4

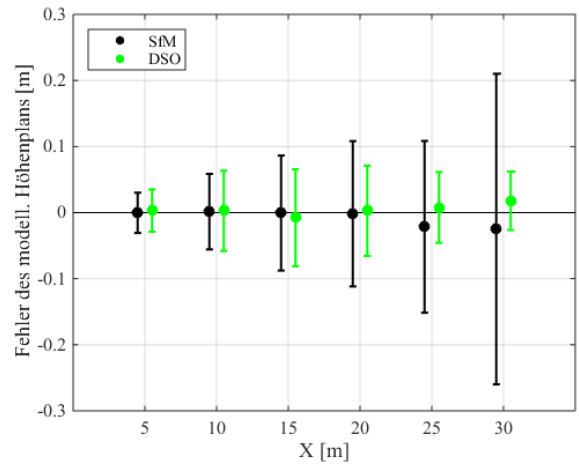


Abbildung B-4 Fehler der Fahrbahnmodellierung. Asphalt – Markierungen 4

Asphalt – Markierungen – Schatten 1

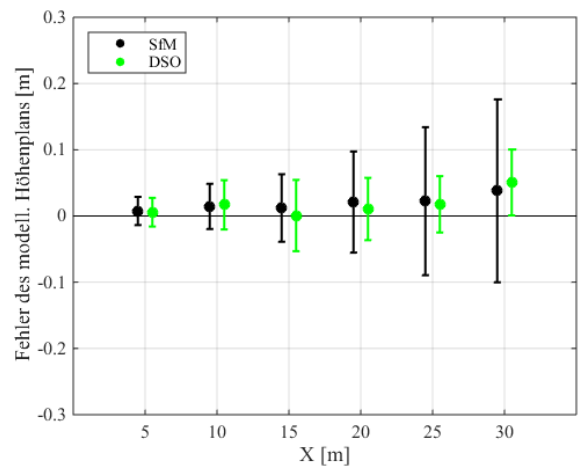


Abbildung B-5 Fehler der Fahrbahnmodellierung. Asphalt – Markierungen – Schatten 1

Asphalt – Markierungen – Schatten 2

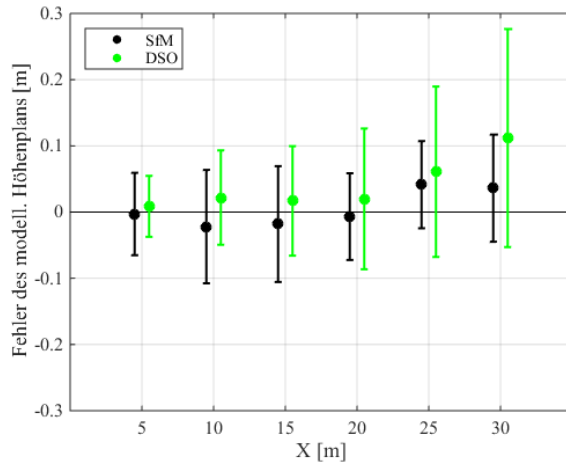


Abbildung B-6 Fehler der Fahrbahnmodellierung, Asphalt – Markierungen – Schatten 2

Pflaster

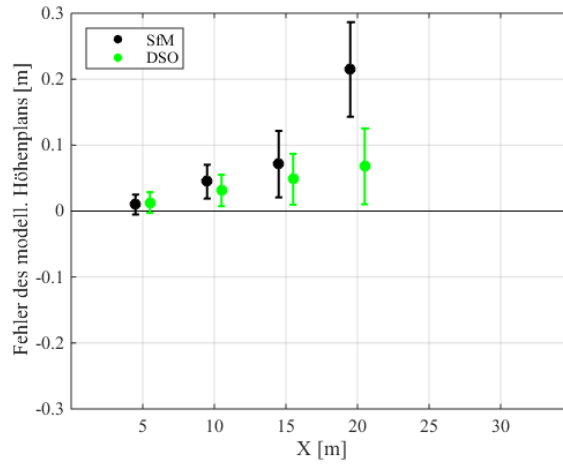


Abbildung B-7 Fehler der Fahrbahnmodellierung, Pflaster

Schatten

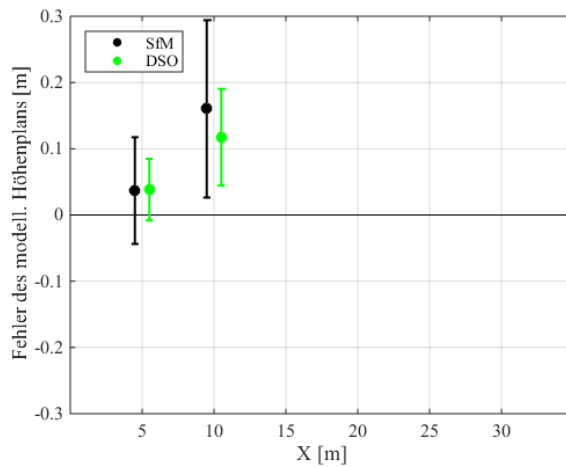


Abbildung B- 8 Fehler der Fahrbahnmodellierung, Schatten

Steigung

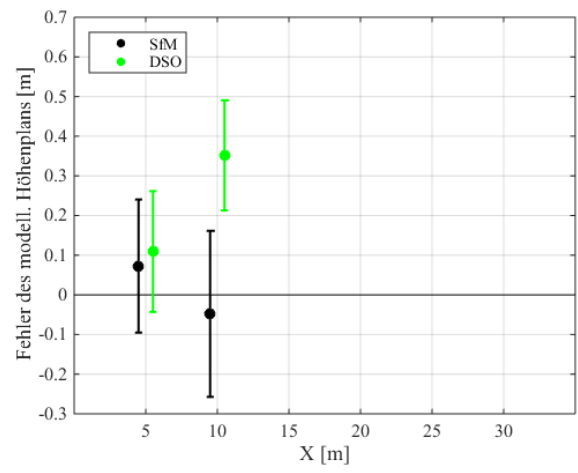


Abbildung B-9 Fehler der Fahrbahnmodellierung. Steigung

Appendix C Hindernisdetektion

Die Hindernisdetektion wird hier, ebenso wie in Abschnitt 8.4, in große und kleine Hindernisse getrennt. Dabei werden für ein Hindernis stets mehrere Sequenzen aufgenommen, um die Detektion bei unterschiedlichen Umgebungssituationen zu analysieren.

Große Hindernisse – Brücken, Tunnel, Unterführungen

Tunnel – 4.555 m

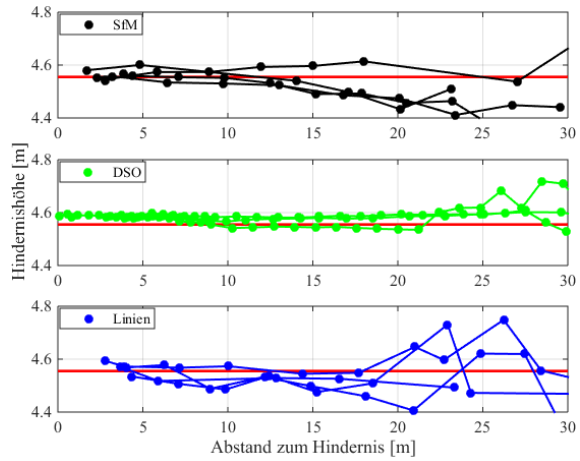


Abbildung C-1 Fehler der Höhenberechnung, Tunnel

Parkdeck – 2.996 m

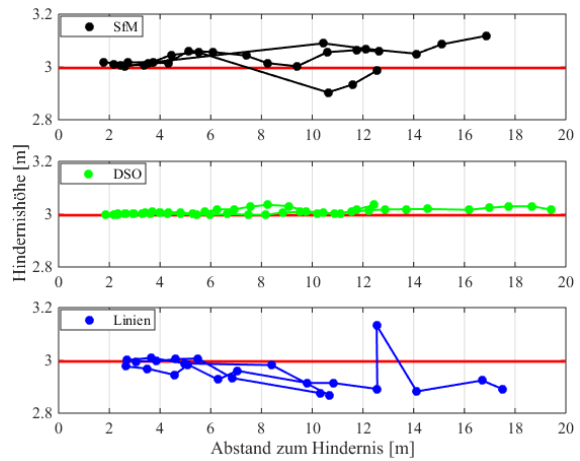


Abbildung C-2 Fehler der Höhenberechnung, Parkdeck

Appendix C

Dachvorsprung – 2.583 m

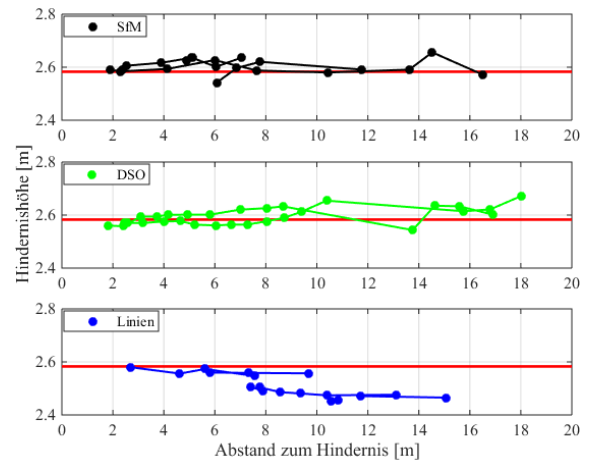


Abbildung C-3 Fehler der Höhenberechnung. Dachvorsprung

Brückenunterführung 1 – 2.957 m

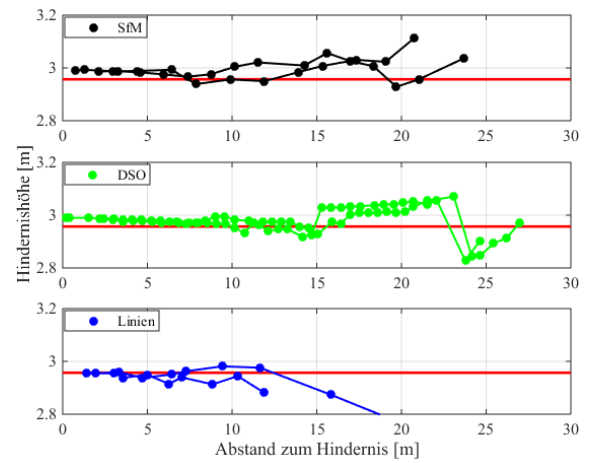


Abbildung C-4 Fehler der Höhenberechnung. Brückenunterführung 1

Brückenunterführung 2 – 2.729 m

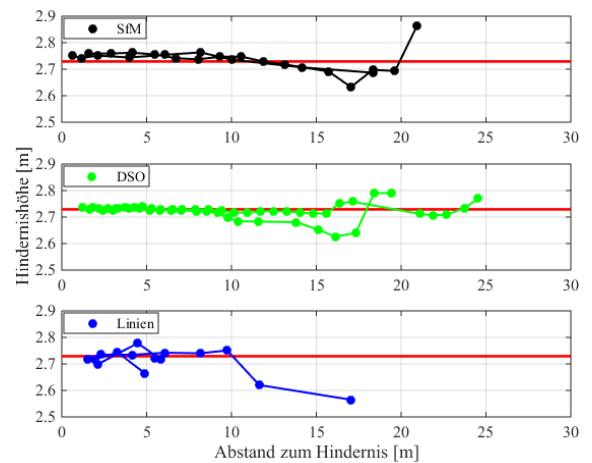


Abbildung C-5 Fehler der Höhenberechnung. Brückenunterführung 2

Kleine Hindernisse – freihängende Schilder und Warnbaken

Parkhaus 1 – 1.951 m

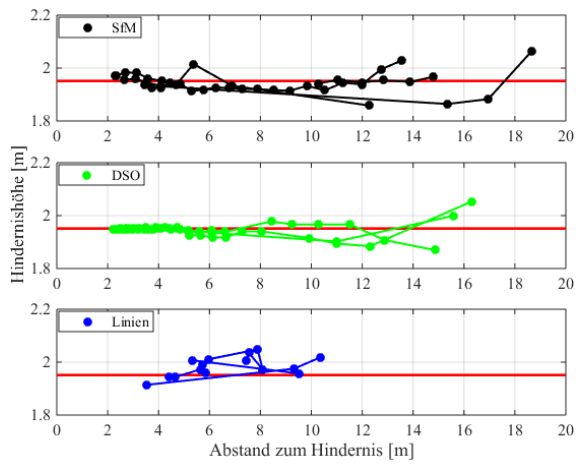


Abbildung C-6 Fehler der Höhenberechnung, Parkhaus 1

Parkhaus 2 – 2.443 m

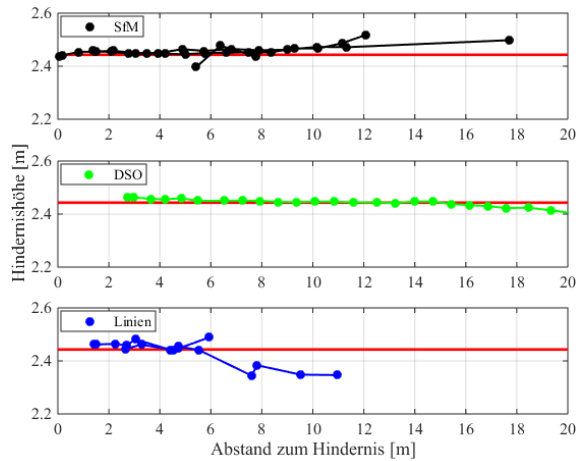


Abbildung C-7 Fehler der Höhenberechnung, Parkhaus 2

Freihängende Warnbake – 2.137 m

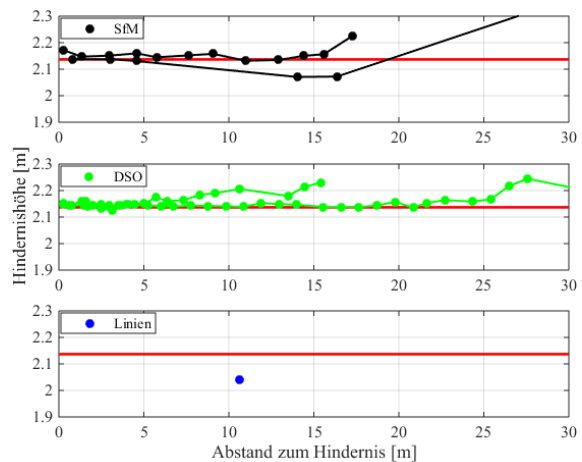


Abbildung C-8 Fehler der Höhenberechnung, Freihängende Warnbake

Freihängendes Schild 1 – 2.369 m

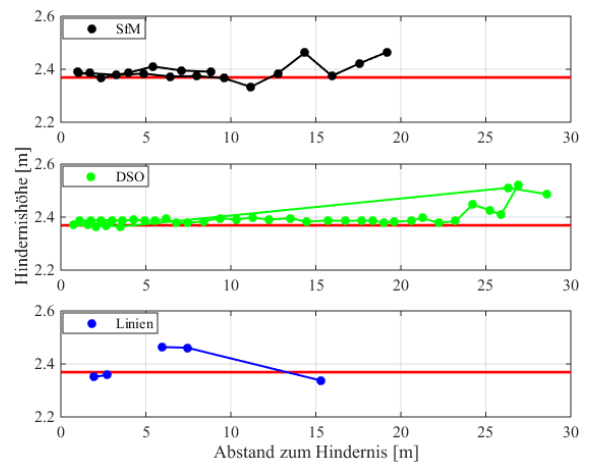


Abbildung C-9 Fehler der Höhenberechnung. Freihängendes Schild 1

Freihängendes Schild 2 – 2.378 m

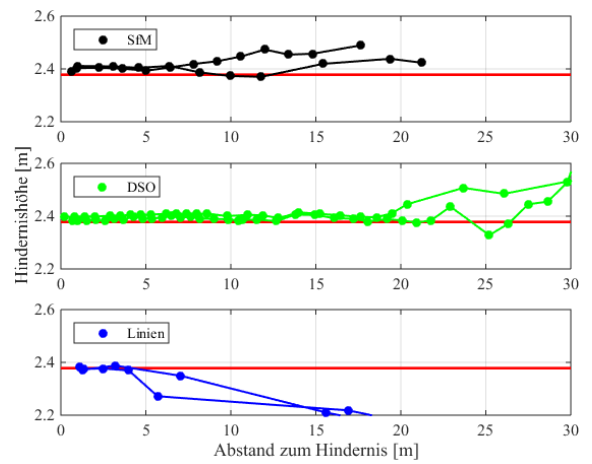


Abbildung C-10 Fehler der Höhenberechnung. Freihängendes Schild 2