

# Objekt- und kartenbasierte Beschreibung der Umgebung ausgehend von einer stereokamerabasierten Umfelderkennung

Dissertation

zur Erlangung des akademischen Grades

Doktoringenieur

(Dr.-Ing.)

von Dipl.-Ing. Thien Nghia Nguyen

geb. am 26.11.1981 in Hatay, Vietnam

genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik  
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr.-Ing. habil. Bernd Michaelis

Prof. Dr.-Ing. Klaus Tönnies

Dr.-Ing. Marc-Michael Meinecke

Promotionskolloquium am 24.01.2012



# Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit am Institut für Elektronik, Signalverarbeitung und Kommunikationstechnik der Otto-von-Guericke-Universität Magdeburg sowie in der Konzernforschung der Volkswagen AG, Bereich Fahrerassistenz und integrierte Sicherheit.

An erster Stelle möchte ich Herrn Prof. Dr. Bernd Michaelis und Herrn Dr. Meinecke danken für die Ermöglichung dieser Arbeit, wertvolle Hinweise und Anregungen und vor allem das große entgegengebrachte Vertrauen. Meinem Zimmerkollegen, Herrn Dr. M. Tornow, danke ich für die zahlreichen Gespräche und Diskussionen, die mich auf viele neuen Ideen gebracht haben, und für die stetige Unterstützung in allen Aspekten der Realisierung einer Doktorarbeit. Herrn Prof. Tönnies danke ich ganz herzlich für die Bereitschaft zur Begutachtung dieser Arbeit.

Dank sagen möchte ich ebenfalls meinen Kollegen, sowohl in der Volkswagen Forschung als auch an der Uni Magdeburg für die gute Zusammenarbeit. Besonders zu erwähnen ist die stetige Unterstützung der Kollegen Jiri Jerhot, Jörn Knaup, Martin Roehder, Simon Steinmeyer, Wojcech Derendarz, Markus Köchy, Helgo Dyckmanns, Andreas Weiser, Dr. Dirk Stüker, Dr. Thanh Binh To und Dr. Marian-Andrzej Obojski während der Umsetzung des Projektes AKTIV-SFR. Herrn Dr. Effertz danke ich für die Aufnahme in sein Team. Herrn Dr. Niese, Christian Bendicks und Michael Heuer danke ich für die Unterstützung bei der Umsetzung und Verfassung der Arbeit.

Ebenso danke ich meinen besten Freunden, Hannes, Tim, Quang und Dr. Thai für die schöne Zeit in Magdeburg. Der Familie Tiep danke ich für die stetige Unterstützung seit meinem ersten Tag in Magdeburg.

Ganz besonderer Dank gilt meinen Eltern und meinem Bruder, die mich stets unterstützt und motiviert haben. Diese Arbeit widme ich ihnen.



# Schriftliche Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Insbesondere habe ich nicht die Hilfe einer kommerziellen Promotionsberatung in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Magdeburg, 01.11.2011

Thien Nghia Nguyen



# Kurzfassung

Die vorliegende Arbeit befasst sich mit der Umfeldwahrnehmung im Stadtverkehr mit Hilfe eines Stereokamera-Systems. Die Herausforderung für diese Arbeit besteht darin, Hindernisse und Randbebauungen in vielfältigen, manchmal unvorhersehbaren, Formen wahrzunehmen sowie die zahlreichen sich in verschiedene Richtungen bewegendenden Verkehrsteilnehmer zu ermitteln und zu verfolgen.

Die Grundlage für diese Arbeit stellt die 3D-Vermessung mit Hilfe des Stereokamera-Systems dar. Die rekonstruierten 3D-Punkte der Umgebung werden auf ein zwei dimensionales Histogramm projiziert. Davon ausgehend werden 3D-Punkte, die aus einem Objekt stammen, in Segmente zusammengefügt. Diese Segmente werden als initiierte Objekte an das Trackingframework übergeben. Das Trackingframework basiert auf einem erweiterten Kalmanfilter. Durch das Framework können bewegte Objekte identifiziert und verfolgt werden. Segmente, die nicht mit bewegten Objekten in Verbindung stehen, werden an eine Belegungskarte übergeben. Die Belegungskarte ist eine Methode zur zuverlässigen Detektion und Darstellung von stationären Hindernissen. Als Ergebnis können zum einen bewegte Verkehrsteilnehmer im Umfeld zuverlässig erfasst und verfolgt und zum anderen Hindernisse und befahrbare Bereiche für das eigene Fahrzeug ermittelt werden.

Im Vergleich zu den in der Literatur bekannten Verfahren werden in dieser Arbeit zwei neue Aspekte eröffnet: die Verbesserung des Tracking durch Nutzung der bildbasierten Merkmale, die aus jedem Objekt extrahiert werden, sowie die Kombination von Objekttracking und Belegungskarte zur verbesserten Umfelderkennung und Umfeldbeschreibung.

# Inhaltsverzeichnis

<b>Vorwort</b> .....	iii
Schriftliche Erklärung.....	v
Kurzfassung .....	vii
Inhaltsverzeichnis.....	viii
Abkürzungen.....	xiii
Notationsvereinbarungen .....	xiv
Formelzeichen .....	xiv
<b>1 Einleitung</b> .....	1
<b>2 Umfeldwahrnehmung in Fahrerassistenzsystemen - Stand der Technik</b> .....	7
2.1 RADAR-Sensorik.....	9
2.2 Laser-Sensorik.....	11
2.3 PMD-Sensorik.....	13
2.4 Ultraschallsensoren.....	15
2.5 Optische Kameras.....	15
2.6 Objektbasierte Umfeldmodellierung .....	17
2.7 Kartenbasierte Umfeldmodellierung .....	20
2.8 Zielsetzung und Aufbau der Arbeit .....	22
<b>3 Grundlagen zur Umfeldvermessung mittels Stereophotogrammetrie</b> .....	25
3.1 Der Stereonormalfall.....	25



3.2	Korrespondenzanalyse zur Disparitätsberechnung .....	28
3.2.1	Blockmatching-Verfahren.....	30
3.2.2	Optimierung der Korrespondenzsuche .....	33
3.3	3D Rekonstruktion .....	39
<b>4</b>	<b>Detektion und Tracking von bewegten Objekten auf der Basis von rekonstruierten 3D-Informationen .....</b>	<b>41</b>
4.1	Objektdetektion durch Segmentierung .....	41
4.2	Objekttracking.....	46
4.2.1	Kalmanfilter .....	49
4.2.2	Erweitertes Kalman Filter .....	53
4.2.3	Parameterdiskussion .....	54
4.2.4	Assoziation neuer Messungen zu bekannten Objekten.....	57
4.2.4.1	Mahalanobis-Abstand.....	58
4.2.4.2	Global-Nearest-Neighbor-Methode (GNN).....	59
4.2.4.3	All-Neighbors-Methode .....	60
4.2.4.4	Verbesserung der Objektassoziation durch zusätzliche Bildmerkmale..	60
4.3	Zusammenfassung.....	69
<b>5</b>	<b>Belegungskarte zur Darstellung stationärer Hindernisse.....</b>	<b>71</b>
5.1	Grundidee der Belegungskarte .....	71
5.2	Stereo-spezifisches Inverses Sensormodell.....	74
5.2.1	Die lokale Karte .....	75

5.2.2	Die globale Karte .....	76
5.2.3	Abbildung lokaler Karte auf globaler Karte.....	78
5.3	Kartenaktualisierung .....	83
5.3.1	Verwendung des Bayes-Theorems .....	83
5.3.2	Verwendung der Dempster-Shafer-Theorie .....	85
5.3.3	Vergleich der beiden Verfahren.....	86
5.3.4	Behandlung von bewegten Objekten.....	89
5.4	Zusammenfassung.....	92
<b>6</b>	<b>Experimentelle Ergebnisse und Analyse .....</b>	<b>95</b>
6.1	Untersuchte Szenarien .....	98
6.1.1	Stationäre Umgebung.....	98
6.1.2	Folgefahrt .....	99
6.1.3	Gegenverkehr .....	102
6.1.4	Folgefahrt und Gegenverkehr.....	104
6.1.5	Abbiegen .....	105
6.1.6	Querverkehr.....	107
6.2	Analyse .....	109
6.2.1	Verfolgung von bewegten Objekten .....	111
6.2.2	Verbesserung der Belegungskarte durch Objekttracking .....	113
<b>7</b>	<b>Schlussbetrachtungen.....</b>	<b>119</b>
7.1	Zusammenfassung.....	119

7.2 Ausblick.....	123
8 Veröffentlichungen.....	CXXV
Literatur .....	CXXVII



# Abkürzungen

Abkürzung	Bedeutung
FAS	Fahrerassistenzsysteme
ACC	Automatic Cruise Control
ABS	Antiblockiersystem
ESP	Elektronisches Stabilitätsprogramm
PMD	Photonic Mixture Devices
3D	Dreidimensional
SAD	Summe der absoluten Differenzen
SADMF	Mittelwertfreie Summe der absoluten Differenzen
SSD	Summe der Differenzenquadrate
SSDMF	Mittelwertfreie Summe der Differenzenquadrate
KKF	Kreuzkorrelationsfunktion
KKFMF	Kreuzkorrelationsfunktion
KF	Kalmanfilter
EKF	Erweiterter Kalmanfilter
HMM	Hidden Markov Model
SURF	Speed Up Robust Features
POI	Point-of-Interest
ROI	Region-of-Interest
GNN	Global nearest neighbor
SLAM	Simultaneous localization and mapping
BS	Bayes Theorem
Dempster-Shafer	Dempster Shafer Theorem
ADTF	Automotive Data and Time triggered Framework
FPGA	Field Programmable Gate Array

# Notationsvereinbarungen

Skalare	$x, y, z, \dots, X, Y, Z$
Vektoren	Klein, unterstrichen $\underline{x}, \underline{y}, \underline{z}$
Matrizen	Groß, unterstrichen $\underline{A}, \underline{B}, \underline{C}$
Mengen	Kalligraphisch, groß $\mathbb{R}, \mathbb{Z}, \mathbb{C}$

# Formelzeichen

Symbol	Bedeutung
$X, Y, Z \in \mathbb{R}$	3D-Koordinaten eines Raumpunktes
$u, v \in \mathbb{N}$	Bildkoordinaten
$I(u, v)$	Grauwert eines Bildes als Funktion
$d$	Disparitätswert
$D(u, v)$	Disparitätsfunktion
$\underline{x}_k$	Zustandsvektor eines Objektes zum Zeitpunkt $k$
$\underline{P}_k$	Kovarianzmatrix
$\underline{z}_k$	Gewonnener Messwert zum Zeitpunkt $k$
$p(\ )$	Wahrscheinlichkeit
$\mu$	Mittelwert
$\underline{F}_k$	Zustandsübergangsmatrix
$\hat{\underline{x}}_k$	Prädizierter Zustand
$\underline{B}_k$	Übersetzungsmatrix
$\underline{u}_k$	Steuerungsvektor
$\underline{w}_k$	Modellrauschen
$\underline{v}_k$	Messrauschen
$\underline{Q}_k$	Varianzmatrix des Modellrauschens

$\underline{\mathbf{R}}_k$	Varianzmatrix des Messrauschens
$\underline{\mathbf{H}}_k$	Messmatrix
$\underline{\mathbf{K}}_k$	Gainmatrix
$x$	Relative Entfernung eines Objektes in longitudinaler Richtung
$y$	Lateraler Versatz eines Objektes
$v_x$	Relative Geschwindigkeit in longitudinaler Richtung
$v_y$	Relative Geschwindigkeit in lateraler Richtung
$B$	Breite eines Objektes
$L$	Länge eines Objektes
$H$	Höhe eines Objektes
$\tilde{\mathbf{I}}$	Durchschnittlicher Grauwert
$\delta I$	Varianz des Grauwertes
$\underline{\mathbf{H}}(\underline{\mathbf{x}}, \sigma)$	Hesse-Matrix
$L_{xx}(\underline{\mathbf{x}}, \sigma)$	Faltung der zweiten partiellen Ableitung der Gaußfunktion mit dem Eingangsbild
$D_{xx}$	Approximation von $L_{xx}(\underline{\mathbf{x}}, \sigma)$
$\sigma$	Skalierungsfaktor
$L$	Lokale Karte als eine Menge von Zellen
$l_k$	Lokale Zelle
$C$	Globale Karte
$c_n$	Globale Zelle
$m_n(\ )$	Glaubensmaße einer Zelle
$\underline{\mathbf{T}}$	Translationsmatrix
$\underline{\mathbf{R}}$	Rotationsmatrix





# Kapitel 1

## Einleitung

„Die weltweite Nachfrage nach Kraftfahrzeugen wird eine Million nicht überschreiten, allein schon aus Mangel an Chauffeuren“ –(Gottlieb Daimler 1834-1901). Mit diesem Satz hat Gottlieb Daimler, ein bekannter deutscher Ingenieur, Konstrukteur und Industrieller, der die heutige Automobileindustrie mit dem ersten schnelllaufenden Benzinmotor und das vierrädrige Kraftfahrzeug mit Verbrennungsmotor geprägt hat, die Zukunft prophezeit. Heutzutage ergibt sich jedoch ein anderes Bild. Bis zum 01.01.2008 wurden allein für Deutschland 56,956,183 Fahrzeuge zugelassen [AKT10]. Die hohe und stetig steigende Anzahl an Fahrzeugen in Deutschland bzw. weltweit stellt große Herausforderungen an Infrastruktur, Umwelt und Verkehrssicherheit.

In Abbildung 1-1 wird eine statistische Zusammenfassung der Unfallgetöteten dargestellt. Es ist zu erkennen, dass zwar die gesamte Anzahl stetig sinkt, 1980 bis 2006, jedoch hat es in Deutschland im Jahr 2006 noch über 8.000 tödliche Unfälle gegeben. Der Rückgang der Unfälle ist zum Teil auf die verbesserte Infrastruktur, Aufklärung der Verkehrsregeln bzw. Fortschritte in der Fahrzeugsicherheitstechnik zurückzuführen. Die Fortschritte beinhalten unter anderem die Einführung des Airbags (ab 1974), des ABS - Antiblockiersystem (ab 1985) sowie des ESP - Elektronisches Stabilitätsprogramm (ab 1995). Dadurch können Fahrer bzw. Insassen in kritischen Situationen geschützt und unterstützt werden.

Des Weiteren wurden in fortgeschrittenen (hochdotierten) Forschungsprojekten wie [PRO95], [PRE05], [AKT10], an denen Partner aus der Automobileindustrie und führenden Forschungsinstituten beteiligt sind, verschiedene Maßnahmen zur Erhöhung der Sicherheit im Verkehr erforscht. Die erzielten Ergebnisse aus solchen Projekten werden in der Regel in

die Fahrzeugentwicklung transferiert, was in neuen Produkten mit erweiterten Sicherheitsfunktionen resultiert.

Erwähnenswert sei zum Beispiel das erste ACC-System (Abstandsregeltempomat) [AUT07], das ab 1999 in den Markt eingeführt wurde. Dieses System detektiert das vorausfahrende Fahrzeug und passt automatisch die Fahrgeschwindigkeit und den Sicherheitsabstand an.

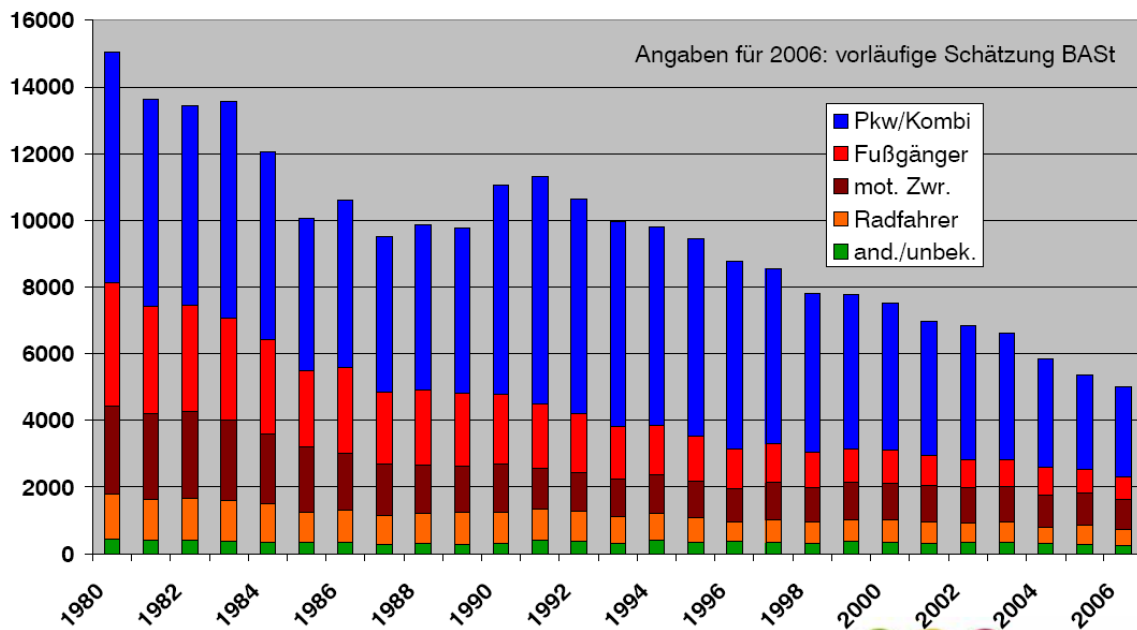


Abbildung 1-1: Getötete nach Art der Verkehrsteilnahme absolut, Deutschland 1980-2006 [AKT11].

Eingesetzt wird das System überwiegend auf der Autobahn oder der Landstraße bei einem Geschwindigkeitsbereich von 30km/h bis 180 km/h. Weiterhin ist seit 2011 das System „Side Assist“ im Markt erhältlich, welches die seitlichen Bereiche hinter dem Fahrzeug überwacht und den Fahrer bei einem Spurwechselvorhaben vor einer gefährlichen Situation warnt [SID08].

Solche Systeme werden als Fahrerassistenzsysteme (FAS) bezeichnet. FAS haben im Allgemeinen die Aufgabe, den Fahrer bei Fahraufgaben zu unterstützen, ohne ihn zu belasten und sind eine hervorragende Ergänzung zu Sicherheitssystemen wie ABS und ESP. Anders als bei Sicherheitssystemen kann der Fahrer die Funktionalität des FAS gegebenenfalls immer übersteuern und hat nach wie vor die Verantwortung für das Geschehen.

Abbildung 1-2 zeigt wo FAS am besten wirken können, nämlich dort wo der Fahrer unter- oder überfordert wird.

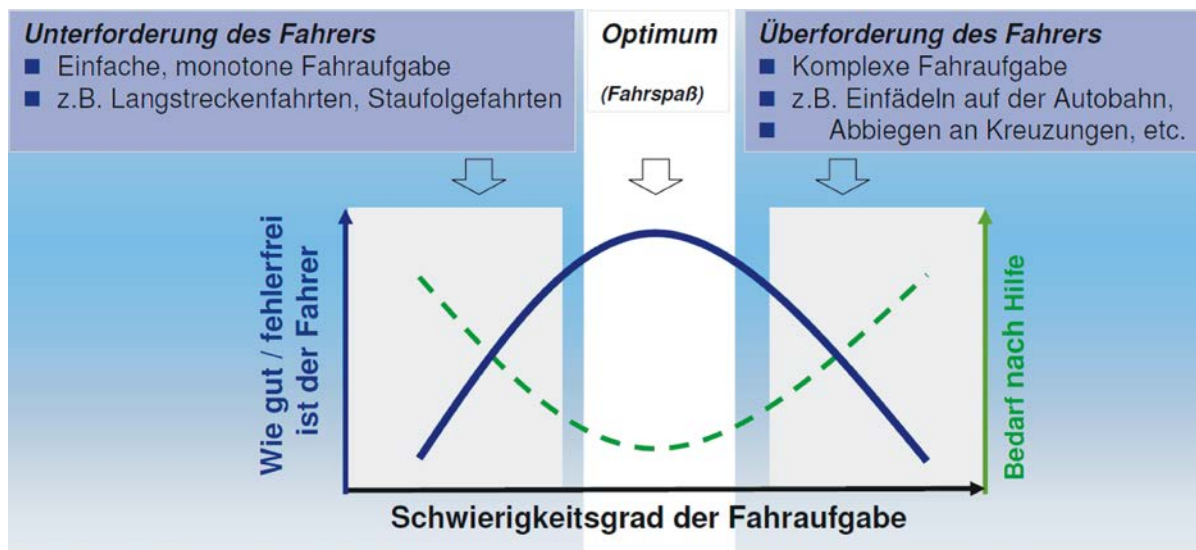


Abbildung 1-2: Wirkungsbereich von Fahrerassistenzsystemen [WEI09].

Unterforderung des Fahrers durch einfache und monotone Fahraufgaben wie Langzeitfahren in Fahrzeugkolonnen auf Autobahnen oder Landstraßen kann unter Umständen Müdigkeit und Ablenkung beim Fahrer verursachen. Da kommen Systeme wie ACC sehr gut zum Einsatz und können den Fahrer wesentlich entlasten. Auf der anderen Seite kann der Fahrer, im Besonderen bei unerfahrenen Fahrern, bei komplexen Fahraufgaben wie Links-Abbiegen an einer Kreuzung mit dem Vorhandensein von Radfahrern überfordert sein. Die Unterstützung für den Fahrer in solchen Situationen ist Gegenstand der heutigen Forschung. Bei der Urban-Challenge 2007 [DAR07] haben führende Forschungsinstitute eine autonome Fahrzeugsführung im städtischen Verkehr demonstriert.

Ein solches FAS besteht aus drei Hauptkomponenten: die Umfeldwahrnehmung, die Situationserkennung und die Fahrzeugsregelung/Mensch-Maschine-Schnittstelle. Aufgabe der Umfeldwahrnehmung ist es, mit Hilfe von sensorischen Vermessungen relevante Informationen über das Umfeld bereitzustellen. Auf der Basis dieser Informationen wird die Fahrsituation analysiert [BER09] bzw. werden mögliche Kollisionen abgeschätzt [MON08], [MEI10]. Des Weiteren kommen von diesem Modul Vorgaben bzw. Entscheidungsmaßnahmen zur Kollisionsvermeidung/-folgenminderung, die das Modul „Fahrzeugsregelung“ um-

setzt [EIG09]. Je nach Anwendungen werden notwendige Informationen dem Fahrer über die Mensch-Maschine-Schnittstelle bereitgestellt [HEE10].

Ziel ist es, dem Fahrer in seinen Fahraufgaben zu assistieren, ohne ihn abzulenken oder gar zu belasten. Die Motivation für die vorgelegte Arbeit stammt aus dem Förderprojekt AKTIV-SFR „Adaptive und kooperative Technologien für den intelligenten Verkehr – Sicherheit für Fußgänger und Radfahrer“ [AKT10].

Es werden unfallrelevante Szenarien zwischen Fußgänger und Fahrzeugen ausgewertet, analysiert, Gegenmaßnahmen entwickelt und erprobt. Eine der größten Herausforderung in diesem Projekt ist die Umfeldwahrnehmung im Stadtverkehr, darunter auch die Erkennung von Fußgängern in Umgebungen mit dichtem Aufkommen von Objekten. Es sollen nicht nur Fußgänger sondern auch Radfahrer und andere Verkehrsteilnehmer wahrgenommen werden. Ausgestattet wird der Versuchsträger mit folgenden Sensoren: nach vorne ein Stereokamera-System, ein PMD-Sensor, ein 77 GHz-Radar-Sensor, seitlich nach hinten zwei 24 GHz-Radar-Sensoren. Abbildung 1-3 veranschaulicht den sensorischen Aufbau des Versuchsträgers [AKT11].

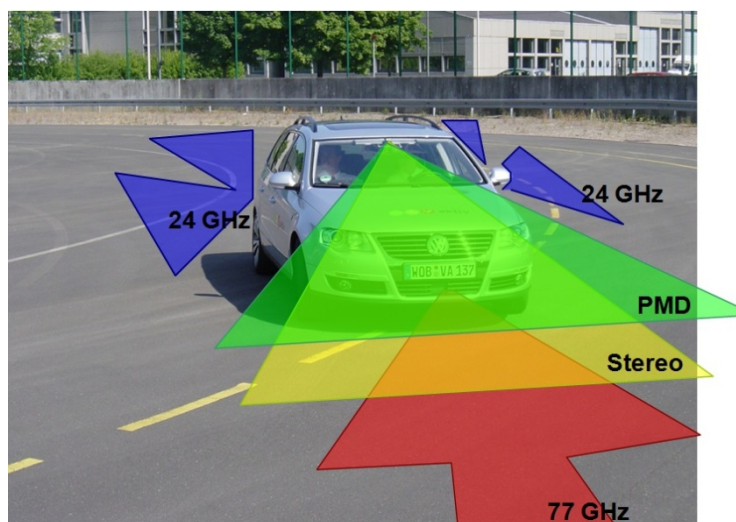


Abbildung 1-3: VW-Versuchsträger des Projektes AKTIV-SFR. Die unterschiedlichen Sensoren und deren Messbereiche werden dargestellt.

Aufgrund der vielfältigen Anzahl der Objekte und Hindernisse im Stadtverkehr ist es mit den herkömmlichen Verfahren nicht gelungen, die bewegten Objekte und Randbebauungen zuverlässig zu erkennen sowie darzustellen. Der objektbasierte Ansatz, der im nächsten

Kapitel ausführlicher beschrieben wird, kann einerseits bewegte Fahrzeuge und Fußgänger gut verfolgen, ist aber zur Darstellung von stationären Hindernissen nicht geeignet.

Auf der anderen Seite kommt der kartenbasierte Ansatz, der auch im nächsten Kapitel näher erläutert wird, nicht ohne Weiteres mit den bewegten Objekten zurecht, während die Randbebauungen mit diesem Ansatz gut beschrieben werden können. Daher ist Ziel dieser Arbeit, eine Methode zu entwickeln, die den Herausforderungen im Stadtverkehr optimal entgegen kommt.

Im nächsten Kapitel wird auf die Umfeldwahrnehmung eingegangen. Weitere Informationen zu FAS können z.B. in [WIN09] nachgelesen werden.



# Kapitel 2

## Umfeldwahrnehmung in Fahrerassistenzsystemen - Stand der Technik

Außerhalb des Kontextes von Fahrerassistenzsystemen ist die Umfeldwahrnehmung, zum Beispiel in der Robotik, seit Jahrzehnten ein wichtiger Bestandteil. Ein Wahrnehmungssystem besteht in der Regel aus Sensoren, Einheiten zur Verarbeitung der Sensormesssignale und einem Modellierungsframework. In der Flugsicherung werden zum Beispiel Radar-Sensoren eingesetzt, um Ziele im Umkreis von mehreren 100 km zu detektieren. Im Rahmen des Projektes PROMETHEUS [PRO95] wurde beispielhaft die Fahrzeugerkennung mit Hilfe einer Kamera umgesetzt. Zurzeit werden in verschiedenen Projekten wie AKTIV [AKT10], PREVENT [PRE05] weitere Assistenz- und Sicherheitsfunktionen auf Basis der Erkennung von Objekten durch Kamera-Systeme realisiert. In Abbildung 2-1 wird ein genereller Aufbau eines Wahrnehmungssystems veranschaulicht.

Eine Reihe von Sensoren, die in Fahrerassistenzsystemen eingesetzt werden, wird zum Beispiel in [AUT07] dargestellt. Die Verarbeitung der Messdaten dient zur Extraktion der relevanten Informationen. Dazu gehören beispielsweise die Befahrbarkeit des geplanten Fahrweges, die Position anderer Objekte im Umfeld, vorhandene Fahrspuren, etc. All diese Informationen werden verwendet, um das Umfeld zu beschreiben bzw. zu interpretieren. Dieser Vorgang wird als Umfeldmodellierung bezeichnet.

Im Folgenden werden die zur Umfeldvermessung häufig eingesetzten Sensoren beschrieben. Jede Art von Sensoren besitzt ein unterschiedliches Messprinzip und wird daher für unterschiedliche Aufgaben sowie Erfassungsbereiche verwendet. In Abbildung 2-2 werden die aktuell häufig eingesetzten Sensoren zur Umfeldwahrnehmung veranschaulicht. Naturgemäß

ist der Umfeldbereich vor dem Fahrzeug intensiver zu überwachen. Der Frontbereich wird in Fern- und Nahbereich eingeteilt. Der Fernbereich ist für Hochgeschwindigkeitsfahrten, zum Beispiel Autobahnfahrten, relevant.

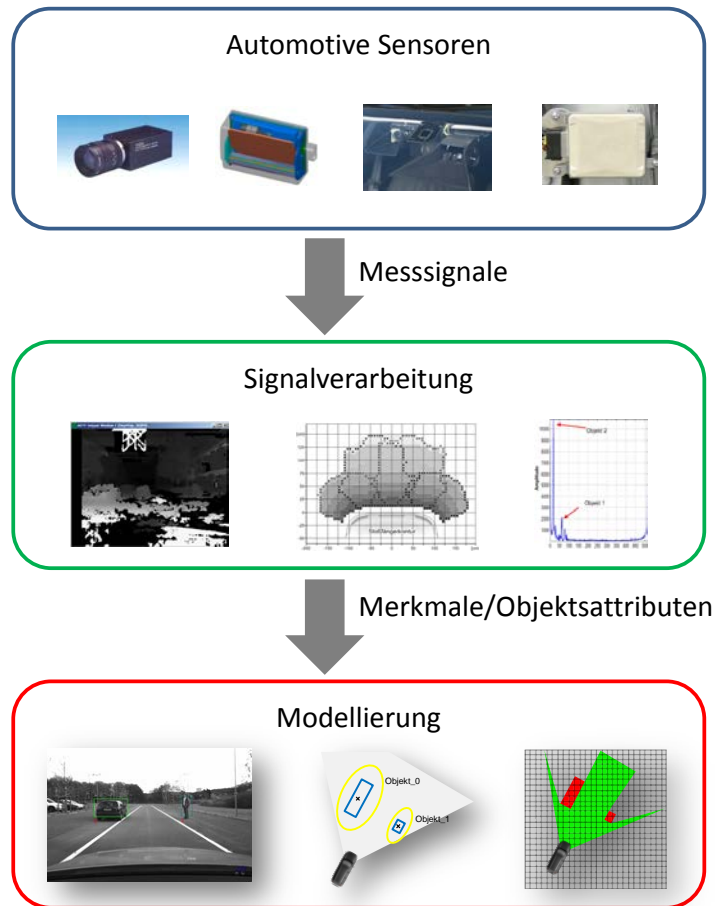


Abbildung 2-1: Genereller Aufbau eines Umfeldwahrnehmungssystems

Es werden häufig Radar-Sensoren für die ACC-Funktion eingesetzt. Im Stadtverkehr spielt dagegen die Erkennung im Nahbereich eine wesentlichere Rolle. Für den Nahbereich wird häufig eine Mono- oder Stereokamera verwendet [AKT10], [INS07]. Die Kamera wird zum Beispiel auch für die Überwachung des Rückraums im Rahmen der Parkassistentenfunktion eingesetzt. Des Weiteren werden Ultraschall-Sensoren im Bereich 0 bis 2 m verwendet, um die für den Fahrer schlecht sichtbaren Hindernisse rund um das Fahrzeug zu detektieren [REI10]. Ferner werden für die Überwachung der hinteren Seitenbereiche, beispielsweise für die Side-Assist-Anwendung [SID08], Radar-Sensoren eingesetzt.

Im Folgenden werden die häufig eingesetzten Sensoren kurz beschrieben.



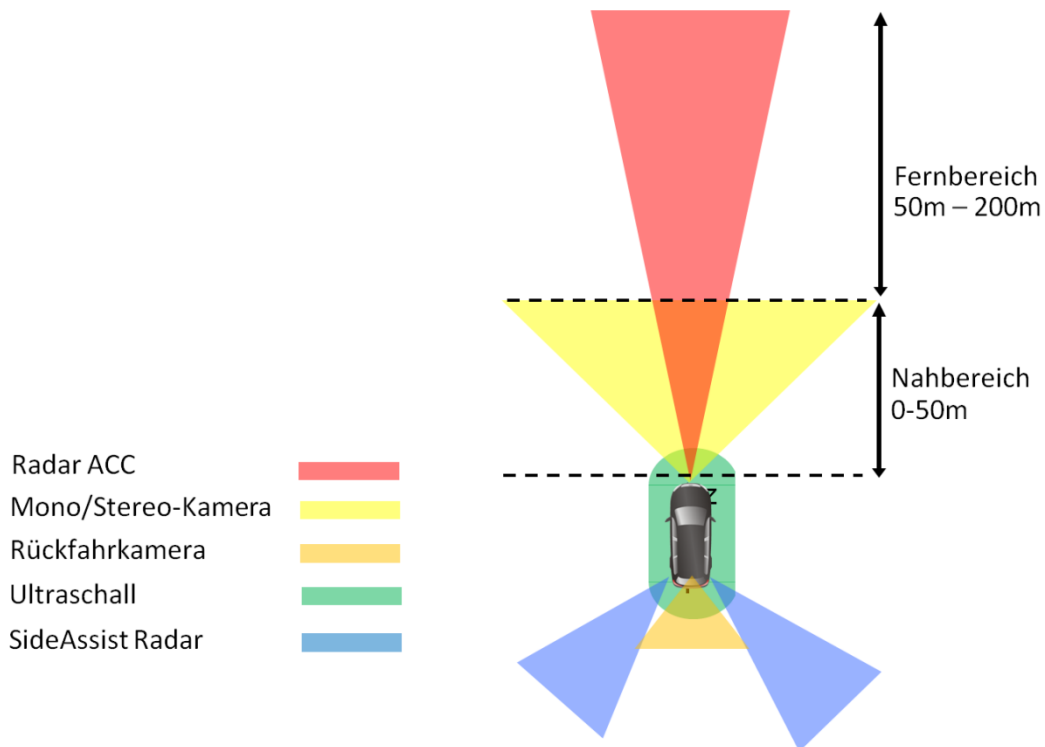


Abbildung 2-2: Die zur Umfeldüberwachung in Fahrzeugtechnik eingesetzte Sensoren mit ihren Erfassungsbereichen.

## 2.1 RADAR-Sensorik

Der Begriff RADAR steht für **R**adio **D**etection **a**nd **R**anging. Das bedeutet Funkortung und Abstandsmessung. Erstmals wurde 1886 von Heinrich Hertz experimentell nachgewiesen, dass Radiowellen von metallischen Gegenständen reflektiert werden. Die Radar-Technik wurde ab Anfang des 19. Jahrhundert stets weiter entwickelt. Die Einsatzgebiete von Radar sind unter anderem Flugzeug-, Schiff-, Wetter- und Bodenradar. Die Grundfunktionsweise eines Radarsystems kann in [BEK64] nachgelesen werden.

Im Bereich der Fahrzeugtechnik wurde Radar in den 70er Jahren in der Forschung zur Detektion anderer Fahrzeuge und Hindernisse eingesetzt. Das erste mit einem Radarsystem zur automatischen Abstandsregelung ausgestattete PKW-Serienfahrzeug erschien 1995 aus dem Haus Mitsubishi. In Deutschland waren ab 1998 solche Systeme erhältlich [MIT95]. Unter den gängigsten Radarmessverfahren können Pulsradar, Dauerstrichradar (CW), linear frequenzmoduliertes Dauerstrichradar (LFMCW), frequenzumgestastetes Radar aufgelistet werden [MEI01].



Abbildung 2-3: Funktionsprinzip des für ACC-Funktion eingesetzten Radarsensors

Der Vorteil beim Einsatz eines Radar-Sensors liegt in der hohen Reichweite von bis 200m, der Witterungsunabhängigkeit sowie geringeren Produktionskosten gegenüber anderen Sensoren. Die derzeit für Radar verwendeten Frequenzen in Fahrerassistenzsystemen liegen bei 24 GHz und 77 GHz. Eine besondere Herausforderung ist neben der präzisen Entfernung- und Relativgeschwindigkeitsmessung die Bestimmung des Abstandswinkels.

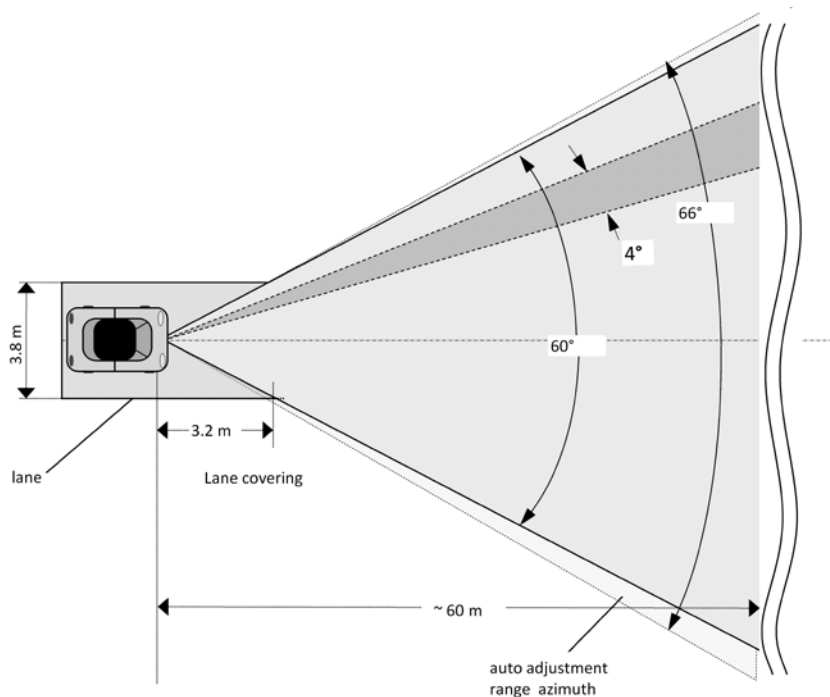


Abbildung 2-4: Funktionsprinzip eines weiter entwickelten Radarsensor mit Verbesserungen zur in Objekttrennung und -lokalisierung [ARS300]

In dem Projekt AKTIV-SFR wurde ein scannender Radarsensor verwendet, der einen im Azimut schwenkbaren Antennenbeam aufweist. Die Winkelmessung erfolgt, wie in Abbildung 2-4 veranschaulicht wird. Durch die Steuerung des Beams kann die Reflektionsstelle für jeden Strahl bestimmt werden. Damit kann zum einen die Position eines Objektes genauer bestimmt und zum anderen mehrere Objekte detektiert werden. Abbildung 2-5 veranschaulicht die Objekterkennung anhand dieses Radar-Sensors.

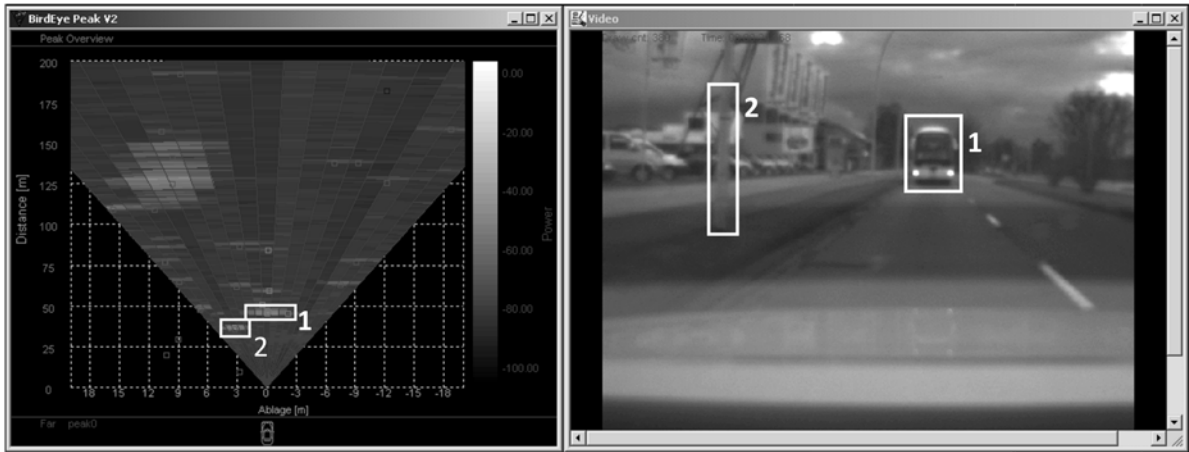


Abbildung 2-5: Umfeldwahrnehmung mit Radarsensor [AKT11]. Das linke Bild stellt die Messwerte dar. Das rechte Bild die daraus resultierende Objektdetektion.

## 2.2 Laser-Sensorik

Laser-Scanner wurden erstmals in der Fahrzeugtechnik zur Umfelderkennung Mitte der 90er Jahre eingesetzt [MIT95]. Es wird ein Laserimpuls engstrahlend zum Objekt gesendet und die Laufzeit zwischen Senden und Empfangen gemessen. Dadurch kann die Entfernung zum Objekt bestimmt werden. Typische Werte für die Pulsdauern liegen im Bereich von Nanosekunden, die Pulslängen liegen somit im 1-m-Bereich.

Durch geeignete Signalverarbeitungsverfahren können Messgenauigkeiten im cm-Bereich realisiert werden. Für ACC-Anwendungen wurde der Laser-Sensor erstmals eingesetzt [LEX03]. Mit einer Reichweite von bis zu 150 m und 30° Sichtfeld sowie Detektionsmöglichkeit von mehreren Objekten bietet der Sensor eine gute Grundlage für die ACC-Funktionalitäten.

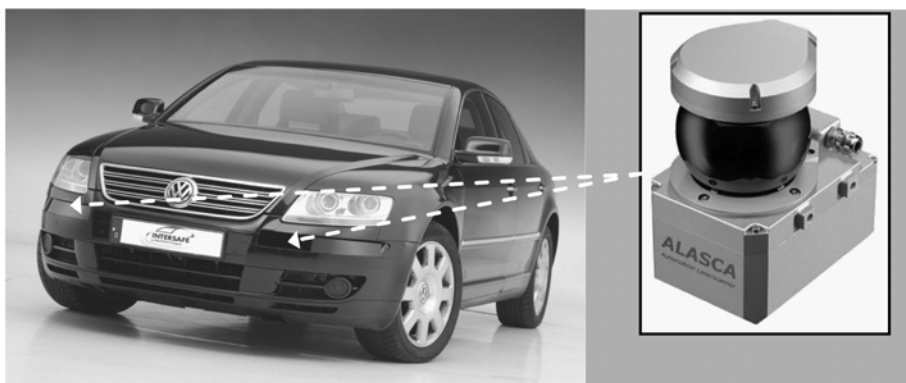


Abbildung 2-6: Beispiel eines Laserscanners, der im Projekt INTERSAFE eingesetzt wurde [INS07].

Für weitere Funktionalitäten wie Assistenz im Stadtverkehr, Fußgängererkennung, etc. wird der Sensor weiter entwickelt. Beispielsweise wird anhand des ALASCA Laser Systems [FUR02] ein Laser-Scanner mit 220° Sichtfeld, 150m Reichweite und 0.5° Winkelauflösung realisiert.



Abbildung 2-7: Umfeldwahrnehmung mit einem Laserscanner. Das Bild wurde aus [FUR0202] zur Veranschaulichung entnommen. Links: das Kamerabild von der Szene. Rechts: das parkende Fahrzeug wird als ein L-förmiges Objekt erkannt. Der Fußgänger dahinter wird auch erkannt und als *Ped* (Fußgänger) klassifiziert.

Durch einen drehenden Spiegel wird der Laserstrahl sozusagen in unterschiedlichen Winkel gelenkt. Der Laser-Scanner kann in der Regel das Umfeld 16-mal pro Sekunde abscannen. Dieser Sensor wurde im Projekt INTERSAFE [INS07] eingesetzt. Die Anforderungen an die Umfeldwahrnehmung im Rahmen dieses Projektes wie Erkennung von Fahrspurmarkierungen, Objekte im Kreuzungsbereich können mit Hilfe dieses Laser-Sensors in Kombination mit anderer Sensorik realisiert werden.

Im Gegensatz zum Radar wird die Detektionsleistung eines Laser-Sensors aufgrund seines optischen Messprinzips von Regen, Schnee sowie Verschmutzung des Sichtbereichs des Sensors negativ beeinflusst. Im Laufe der Entwicklung wird die Wetterunabhängigkeit verbessert. Im Vergleich zum Radar-Sensor weist der Laser-Sensor in der Regel höhere Produktionskosten auf.

In Abbildung 2-7 wird die Umfeldvermessung durch einen Laser-Sensor veranschaulicht. Objekte wie Fahrzeuge, Fußgänger, Randbebauungen können mit ihrer Kontur erfasst werden.

## 2.3 PMD-Sensorik

Einen weiteren auf aktivem Licht basierenden Sensor stellt der PMD-Sensor dar. PMD steht für Photonic Mixer Device, eine 3D-Kamera Technologie, deren Funktionsprinzip auf dem Lichtlaufzeitverfahren (engl.: Time of Flight) beruht und oft als Synonym für alle Time-of-Flight-Sensoren verwendet wird [PMD08].

Es wird an die Umgebung ein Infrarot-Impuls gesendet. Dies wird in der Regel mit Hilfe einer Infrarot-Leiste durchgeführt, siehe Abbildung 2-8. Die von der Szene reflektierte Strahlung trifft auf den PMD Sensor. Dieser ist ebenfalls an die Modulationsquelle gekoppelt. So werden die in Elektronen gewandelten Photonen in Abhängigkeit vom Referenzsignal noch im Halbleiterbereich pixelweise mit Hilfe der so genannten Ladungsträgerschaukel entfernungsselektiv getrennt [PMD08].

Durch diesen einfachen Vergleichsprozess zwischen dem optischen Mess- und dem elektronischen Referenzsignal stellt das resultierende Ausgangssignal des Sensors bereits einen direkten Bezug zur 3D-Information dar. Das Ergebnis ist ein 3D-Tiefenbild von der Umgebung.

Abbildung 2-9 stellt beispielsweise die Aufnahme einer PMD-Kamera dar. Das untere rechte Bild ist das Entfernungsbild. Das obere rechte Bild ist das Amplitudenbild. Die auf Basis des Entfernungsbildes detektierten Objekte bzw. Hindernisse im Umfeld wurden im linken Kamerabild mit weißen Boxen markiert.

Der Vorteil einer PMD-Kamera besteht darin, direkt aus der Messung die Entfernung für jeden Pixel zu bestimmen. Aufgrund der geringen Pixelauflösung der PMD-Kameras, die für den Außenbereich eingesetzt werden [ELI08], reicht es nicht immer für eine Objektklassifikation bzw. Objektunterscheidung aus.

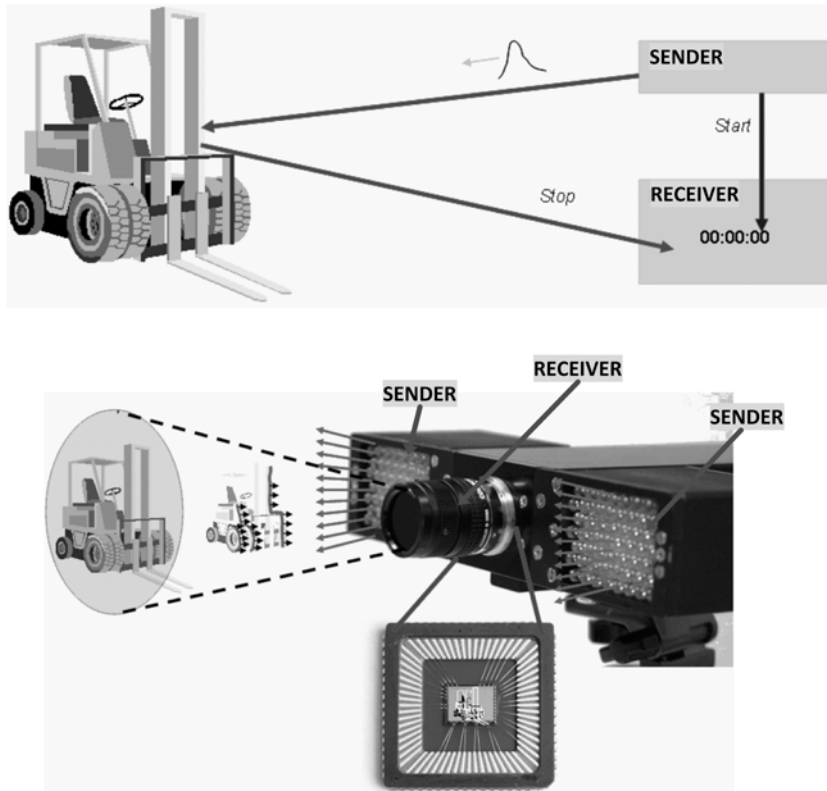


Abbildung 2-8: Das Funktionsprinzip eines PMD-Sensors wird im oberen Bild dargestellt [RIN07]. Das untere Bild gibt ein Beispiel wie der Sender und der Empfänger angebracht werden.



Abbildung 2-9: Umfeldwahrnehmung mit PMD-Sensor [ELI08]. Das linke Bild zeigt die mit PMD erkannten Objekte. Das obere rechte Bild stellt die Intensität der Reflexion dar. Das untere rechte Bild entspricht dem Entfernungsbild.

Beispielsweise kann ein Fußgänger nicht zuverlässig von einem Baum unterschieden werden. Es wird in der Regel eine Kamera zur Verbesserung der Objekterkennung zusätzlich verwendet. Des Weiteren besitzt das Messprinzip „Aktives Licht“ bekannte Probleme wie Übersteuerung, Überblendung etc. und damit auch die Witterungsabhängigkeit. Das Funktionsprinzip des Sensors wird in [PMD08], [PMD09] ausführlich beschrieben.

## 2.4 Ultraschallsensoren

Heutzutage werden Ultraschallsensoren häufig zur Einparkhilfe sowie Rückfahrhilfe eingesetzt. Ultraschallsensoren sind in den Stoßfängern von Kraftfahrzeugen integriert und dienen zur Ermittlung von Abständen zu Hindernissen und zur Überwachung eines Raumes z. B. beim Ein- und Ausparken bzw. Rangieren [REI10].

In Abbildung 2-10, linkes Bild, wird veranschaulicht, dass Ultraschallsensoren bei der automatischen Suche nach freien Parklücken eingesetzt werden. Die neue Generation von Ultraschallsensoren besitzt eine Reichweite von bis 4 m und ist für diesen Zweck gut geeignet.

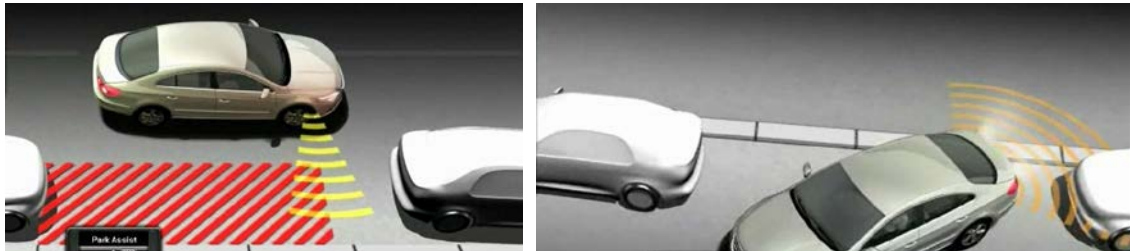


Abbildung 2-10: Ultraschallsensoren eingesetzt zur Einparkhilfe. Links: Parklückenvermessung mit Ultraschallsensor. Rechts: Abstandsmessung beim Einparken

Zur Abstandsmessung senden die Sensoren Ultraschallimpulse mit einer Frequenz von ca. 43,5 kHz aus und detektieren das Zeitintervall zwischen Aussenden der Impulse und Eintreffen der von Hindernissen reflektierten Echoimpulse. Auf Basis der detektierten Laufzeit kann die Entfernung berechnet werden.

In aktuellen Forschungsprojekten [INS09] werden Ultraschallsensoren im LKW zur Fußgänger- sowie Radfahrerdetektion beim Rechtsabbiegen verwendet.

Beim Regen oder starken Wind nimmt die Leistungsfähigkeit von Ultraschallsensoren ab. Im Allgemeinen finden Ultraschallsensoren breite Anwendungen, nicht zuletzt aufgrund der geringen Produktionskosten im Vergleich zu anderen Sensoren.

## 2.5 Optische Kameras

Kameras besitzen im Gegenteil zu den aufgeführten Sensoren ein passives Messprinzip [LUH00]. Das Umfeld wird bei der Vermessung nicht bestrahlt. Die dadurch entstehenden

Vorteile sind z.B. keine elektromagnetische Belastung an die Fahrumgebung, keine neuen gesetzlichen Einschränkungen wegen Strahlung, keine Wechselwirkung zwischen den Sensoren.

Des Weiteren ist die Infrastruktur auf visuelle Wahrnehmung ausgerichtet und teilweise (Spurmarkierungen, Verkehrszeichen) nur bildgebend zu erfassen. Außerdem ermöglichen Kameras eine bildhafte Interpretation der Szene. Dies ist nicht vergleichbar mit der Interpretation einzelner Abstandsmessungen.

Die Grundlagen über Bildaufnahme sowie Bildvorverarbeitung werden in [BUR05] umfangreich beschrieben. Ein Kamera-System in der Fahrzeugtechnik wird beispielsweise im Projekt INTERSAFE [INS07] realisiert. Die eingesetzten industriellen Kameras sind robust gegen Hitze und Schlag. Weiterhin besitzen sie eine relativ hohe Framerate. Die Verarbeitung digitaler Bilddaten kann sowohl auf Basis einer Hardwareimplementierung, z. B. mit FPGA [TOR09], als auch per Software [OPENCV] umgesetzt werden.

In Abbildung 2-11 wird das Kamerasystem im Versuchsträger veranschaulicht. Die Einbauposition sichert die Sichtqualität der Kamera gegen Schmutz oder Regen. Zur Fahrspurerkennung wird die Kamera eingesetzt. In Projekten wie SAFE-U [SAV05] wird die Kamera zur Detektion von Fußgängern verwendet [GAV07] [NED09].

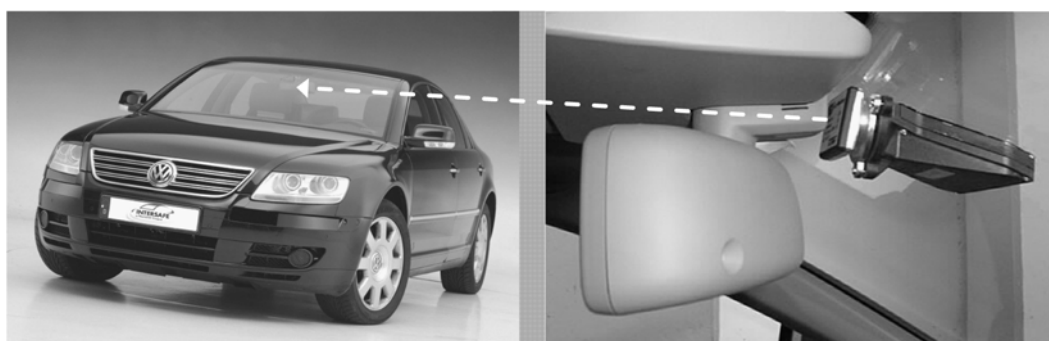


Abbildung 2-11: Mono-Kamera als Sensor für moderne Fahrerassistenzsysteme [INS07]. Links: der Versuchsträger, in dem die Kamera eingebaut wird. Rechts: die Einbauposition der Kamera

Generell gibt es zwei Typen von optischen Kameras: CMOS- und CCD [TOR09]. CCD-Kameras basieren auf passiven Sensorelementen während bei CMOS-Kameras aktive Pixel-Sensoren verwendet werden. Dabei hat jedes Sensorelement einen eigenen Analogsignalver-



stärker. Für den Einsatz im Straßenverkehr haben CMOS-Kameras gegenüber CCD-Kameras Vorteile.

Neben den eingangs genannten Vorteilen hat die Nutzung von Kameras unter Umständen Einschränkungen bei Bestimmung der Entfernung von Objekten sowie der Detektion von nicht trainierten Objektmustern. Diese Nachteile können mit Hilfe von Stereokamera weitgehend ausgeglichen werden. Im Kapitel 3 wird ausführlich das Funktionsprinzip eines Stereokamera-Systems beschrieben.

Zusammenfassend lässt sich sagen, dass jeder Sensor auf Basis seines Messprinzips seine Vor- und Nachteile hat. Um die über mehrere Zeitschritte gewonnenen Sensordaten sinnvoll in Umfeldinformationen zu interpretieren, wird eine passende Umfeldbeschreibung/modellierung benötigt. Im Allgemeinen wird zwischen objektbasierter und kartenbasierter Umfeldmodellierung unterschieden.

## 2.6 Objektbasierte Umfeldmodellierung

Historisch hat dieses Verfahren seinen Ursprung in der Luftabwehrtechnik. Die dort entwickelten Verfahren wurden dementsprechend für die Umfeldwahrnehmung in FAS übernommen und weiter verfolgt. Bei dieser Methode wird das Umfeld durch Objekte beschrieben. Pkws, Lkws, Fußgänger, Radfahrer, stationäre Hindernisse, etc., werden als Objekte ausgegeben. Ein Objekt wird durch einen Zustandsvektor beschrieben. In dem Zustandsvektor sind Informationen zum Objekt, wie zum Beispiel relative Position zum Ego-Fahrzeug, relative Geschwindigkeit, Breite und Länge, Objekttyp (Fahrzeuge, Fußgänger, Radfahrer), etc., enthalten. Die Extraktion der oben genannten Informationen hängt allerdings vom verwendeten Sensor ab. Manche Informationen können mit bestimmten Sensoren einfacher extrahiert werden. In Abbildung 2-12 wird ein Beispiel dieser Methode dargestellt.

In [PET09] wird ein Laserscanner als Hauptsensor eingesetzt und beschrieben, wie ein Objekt aus Messdaten rekonstruiert wird. Parallel dazu wird in [NED09], [BRO05], [LAB05], [GAV07] beschrieben, wie aus Bildsequenzen Objekte extrahiert werden können. Um die

Erkennungssicherheit der Objekte zu gewährleisten, wird ein sogenanntes Objekttracking durchgeführt. Dabei werden zyklisch Messdaten aufgenommen und daraus mit Hilfe stochastischer Ansätze wie z. B. Bayes-Theorem der Zustand des Objektes geschätzt. Stochastische Hypothesen dienen auch zum Eliminieren falscher Objekte. Viele Theorien und Anwendungen im Gebiet Objekttracking wurden in [BLA99] und [SHA00] ausführlich beschrieben.

Ein typischer Anwendungsfall der objektbasierten Umfeldmodellierung ist das ACC-System. Anhand eines Radarsensors wird das Umfeld vermessen und daraus werden Objekte gebildet und verfolgt. Für Fahrten auf Autobahnen oder Landstraßen, wo in der Regel andere Fahrzeuge mit vorhersehbarem Verhalten zu erwarten sind, eignet sich die objektbasierte Methode sehr gut. Auf der anderen Seite kann es z. B. im Stadtverkehr bei der Anwendung der objektbasierten Umfeldmodellierung zu Problemen kommen, wenn die Anzahl und Vielfältigkeit der Objekte sehr hoch sind. Die für Autobahnen und Landstraßen vorbereiteten Objektshypothesen reichen nicht aus, das Umfeld dort ausreichend genau zu beschreiben.

In Abbildung 2-13 ist zu erkennen, dass die Randbebauungen rechts von der Fahrbahn bzw. die dicht beieinander parkenden Autos am linken Fahrbahnrand nicht exakt durch ein Objektmodell beschrieben werden können, da diese in der Objektmodell-Datenbank nicht vorhanden sind. Es ist in der Tat schwierig, die vielfältigen Objekte, die im Stadtverkehr auftreten, lediglich durch einen Quader zu beschreiben (sehr häufig werden Objekte als Quader beschrieben [DAR07], [PET09], [KNO01]). Man kann die Anzahl der Objektmodelle erhöhen, um der Komplexität im Stadtverkehr entgegenzukommen. Dies erhöht aber zugleich den Rechenaufwand dramatisch. Darüber hinaus stellt sich die Frage, wie viele Modelle ausreichend sind, um die Vielfalt der Randbebauungen zu beschreiben, besonders in unstrukturierten Umgebungen wie beim Stadtverkehr.

Auf der anderen Seite können Fahrzeuge mit den bekannten Modellen sehr gut erkannt und verfolgt werden. Ein einfaches Quader-Modell reicht schon aus, Fahrzeuge hinreichend genau zu beschreiben [PET09]. Dieser Vorteil soll weiter genutzt werden. Zur Darstellung der Randbebauungen kann ein Ansatz aus der Robotik herangezogen werden, da die Um-

felderfassung nicht nur ein Forschungsschwerpunkt in FAS ist. In der Robotik wird diese Aufgabe seit mehreren Jahrzehnten bearbeitet. In [ELF89] wird ein alternativer Ansatz zur Umfeldmodellierung dargestellt. Dies wird im nächsten Unterkapitel beschrieben.



Abbildung 2-12: Objektbasierte Umfeldmodellierung auf der Autobahn. PKWs werden als Objekte mit Angaben von relativen Abständen dargestellt.



Abbildung 2-13: Probleme bei der Darstellung von Randbebauungen mit vorhandenen Objektmodellen.

## 2.7 Kartenbasierte Umfeldmodellierung

Die Grundidee des Verfahrens ist es, das Umfeld als eine Art 2D-Gitter darzustellen. Ein Gitter enthält mehrere Zellen. Jede Zelle repräsentiert einen Bereich im Umfeld und enthält dementsprechend Informationen über diesen Bereich. Die wichtigste Information für Fahrerassistenzsysteme ist die Befahrbarkeit des Bereiches.

Liegt ein Hindernis in dem Bereich, wird die korrespondierende Zelle als nicht befahrbar bzw. belegt, bezeichnet. Gehört hingegen die normale Fahrbahn zu dem Bereich, wird die Zelle als befahrbar, oder frei markiert. In verbreiteten Werken [ELF89], [THR05] wird der Zustand der Zelle probabilistisch dargestellt. Dies bedeutet, in jeder Zelle wird eine Belegungswahrscheinlichkeit errechnet. Eine Wahrscheinlichkeit von 1,0 gibt an, die Zelle ist sicher belegt. Eine Wahrscheinlichkeit von 0,0 bedeutet, dass die Zelle frei ist oder dass der Bereich sicher befahrbar ist. Da jede Sensormessung Fehler enthält, erweist sich die probabilistische Darstellung als sehr gut geeignet.

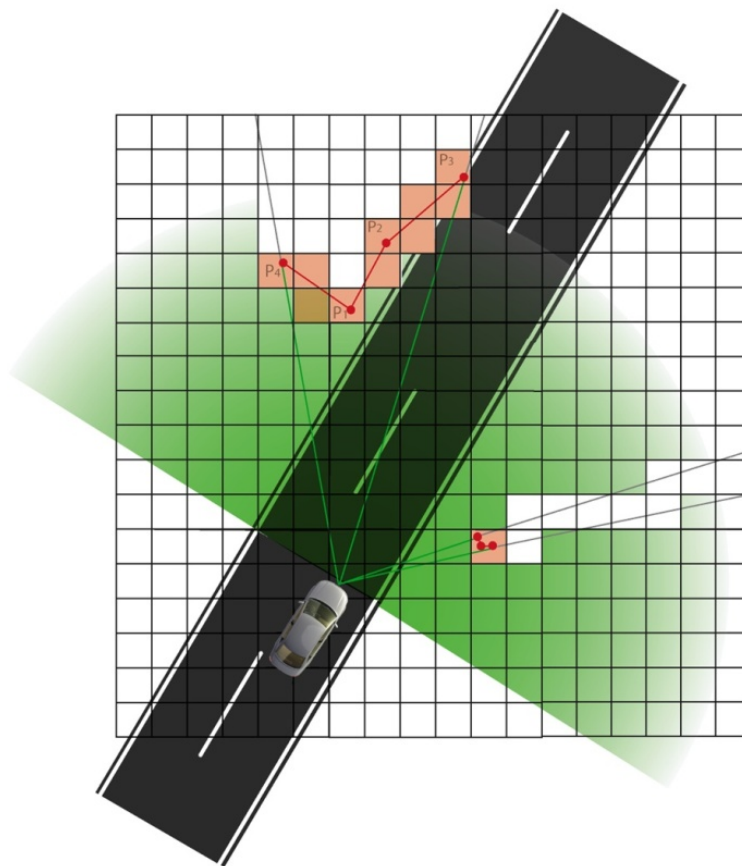


Abbildung 2-14: Kartenbasierte Umfeldmodellierung [WEI06]

Bei diesem Verfahren gibt es zwei Hauptbestandteile: die Abbildung der Messwerte auf der Karte und die Aktualisierung der Zellen bei jeder neuen Messung. Generell kann jeder Sensor für die Karte eingesetzt werden, wenn das Sensormodell entsprechend entwickelt wird. Von Vorteil sind Sensoren, die eine dichte Messung liefern, z.B. Dense-Stereo oder Multi-Ebenen Laserscanner [BAD07], [WAN07].

Das kartenbasierte Verfahren wurde unter anderem im Rahmen der Urban Challenge [DAR07] erfolgreich eingesetzt. Der eingesetzte Sensor war ein 64-Ebene-Laserscanner [HDL64]. In vielen Parkapplikationen [LAM10], [WUT10] wurden mit dem Verfahren gute Ergebnisse erzielt. Da für Parkhäuser und Parkplätze keine große Reichweite des Sensors benötigt wird, wurden hauptsächlich Ultraschallsensoren verwendet.

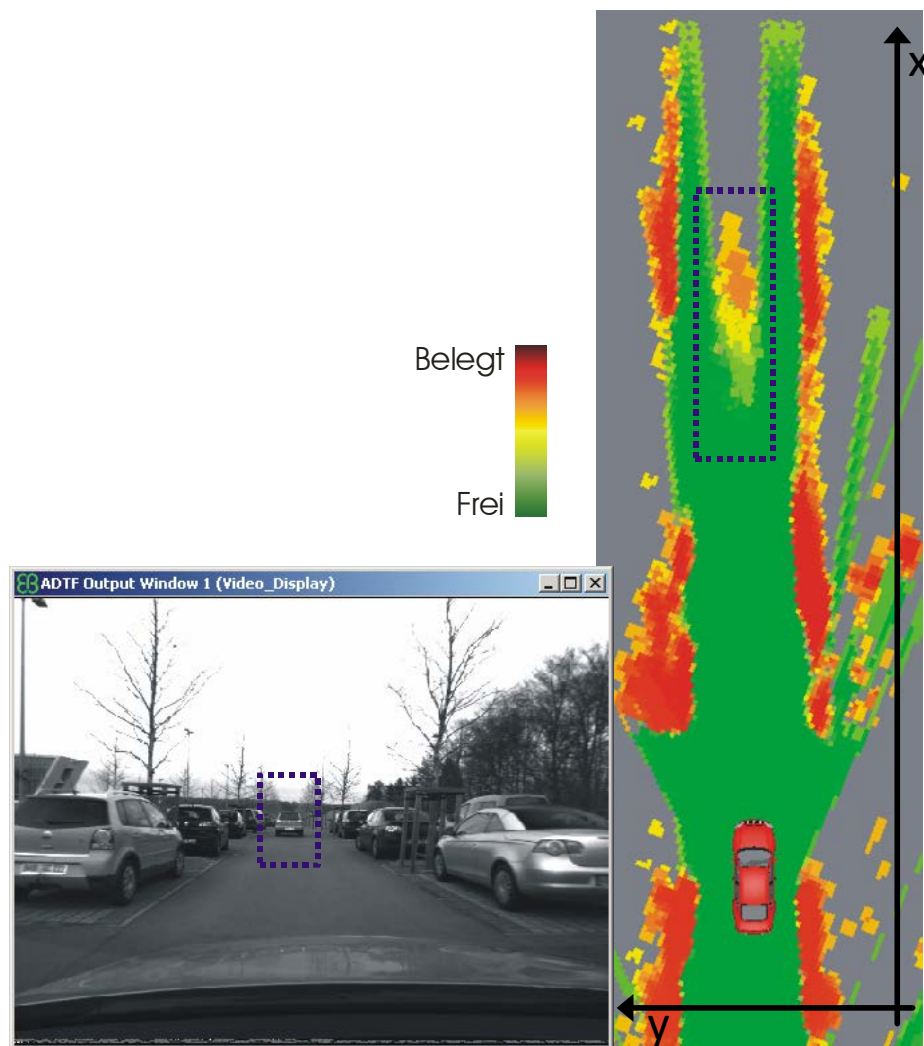


Abbildung 2-15: Dynamisches Objekt in der Belegungskarte. Links: das Kamerabild von der Szene. Rechts: die Belegungskarte mit einem bewegten Objekt, das im blauen Kasten gekennzeichnet wird.

Die Stärke des kartenbasierten Verfahrens liegt darin, dass, im Gegensatz zum objektbasierten Verfahren, nur ein Modell für die Beschreibung des Umfeldes benötigt wird [PLA06]. Dies ist zur Darstellung stationärer Hindernisse in unstrukturierter Umgebung wie dem Stadtverkehr besonders vorteilhaft.

Bei entsprechender Zellenauflösung (jede Zelle deckt in der Regel einen Bereich von 0,2m x 0,2m ab) können Hindernisse sehr gut beschrieben werden. Bei einer überdimensionierten Kartengröße kann es zu großem Rechenaufwand führen. In der Regel reicht im Stadtverkehr eine Kartengröße von 40m x 40m aus. Bei der Erfassung dynamischer Objekte in der Karte hat der Ansatz jedoch Schwächen, die zu Problemen führen. Da die üblichen Aktualisierungsalgorithmen nicht für dynamische Objekte optimiert sind [NGU09], ergeben sich einige Inkonsistenzen in der Darstellung.

In Abbildung 2-15 wird ein Beispiel für eine generierte Belegungskarte in einer Parkplatzsituation dargestellt. Das einzige dynamische Objekt in der Szene ist der vorausfahrende Pkw.

Auf der Karte ist zu erkennen, dass dieses Objekt einen Bereich generiert (in der blauen Box gekennzeichnet), der weder sicher belegt noch frei ist. Diese unsichere Bewertung ist für die auf der Umfeldwahrnehmung basierenden Sicherheitsapplikationen wie Notbremsung oder Fußgängerschutz von großem Nachteil.

An dieser Stelle setzt diese Arbeit an, die Nachteile des einen Verfahrens durch die Stärke des anderen Verfahrens auszugleichen, um insgesamt eine bessere Umfeldbeschreibung zu erreichen.

## 2.8 Zielsetzung und Aufbau der Arbeit

In den vergangenen Abschnitten wurden die üblichen Verfahren zur Umfeldmodellierung dargestellt. Es ist erkennbar, dass jede Methode ihre Vor- und Nachteile hat. Im Folgenden werden kurz die Stärken und Einschränkungen der Verfahren dargestellt.

In Tabelle 2-1 werden die zwei Verfahren gegenüber gestellt. Man sieht, dass die Vor- und Nachteile sich theoretisch gut ergänzen.

	Objektbasierte Modellierung	Kartenbasierte Modellierung
Verfolgung von bewegten Objekten wie Fahrzeuge, Fußgänger, Radfahrer, etc.	Geeignet	Nicht geeignet
Darstellung unstrukturierter stationärer Hindernisse	Nicht geeignet	Geeignet
üblicher Funktionsbereich	Bis 200 m	Bis 50 m

Tabelle 2-1: Vergleich objekt- und kartenbasierter Ansätze

Zwei Verfahren gleichzeitig einzusetzen, löst jedoch nicht das Problem, da beide Verfahren unabhängig funktionieren und es ohne Weiteres nicht möglich ist, die Ergebnisse auf Grund stark unterschiedlicher Darstellungsformen zu fusionieren. Die Zielsetzung dieser Arbeit besteht deshalb darin, ein Verfahren zu entwickeln, dass:

- *sowohl bewegte Objekte als auch stationäre Hindernisse zuverlässig erfasst und darstellt.* Es sollen die Stärken beider oben charakterisierten Verfahren genutzt werden. Konkret heißt das, bewegte Objekte mit Objekttracking zu identifizieren sowie stationäre Hindernisse mit der Karte darzustellen.
- *die Ergebnisse beider Verfahren werden systematisch und konzeptionell kombiniert.* Die Informationen über bewegte Objekte sollen genutzt werden, um die inkonsistenten Bereiche (siehe Abbildung 2-15) zu beseitigen. Es werden Algorithmen entworfen und untersucht, die bewegte Objekte zuverlässig verfolgen.
- *Die Umfeldinformationen sollen nicht doppelt dargestellt werden.* Das heißt, es reicht aus, wenn ein Objekt im Umfeld, seinem Bewegungszustand entsprechend, entweder als stationäres Hindernis oder als bewegtes Ziel dargestellt wird.

Um diese Zielsetzung zu verwirklichen, wird als besonders bedeutsamer Sensor ein Stereokamera-System eingesetzt.

In Kapitel 3 wird die Umfeldvermessung- und rekonstruktion auf Basis einer Stereokamera beschrieben. Es wird erläutert, welche Bedeutung eine dichte Rekonstruktion für eine gute Objektdetektion ist.

Dies stellt die Grundlage für die anschließende Objektdetektion und das Tracking dar, die im Kapitel 4 beschrieben werden. Es wird ein Kalman Filter eingesetzt. Die Parametrisierung des Filters wird ausführlich diskutiert. Was als neu im Vergleich mit den gängigen Trackingverfahren bezeichnet werden kann, ist die Nutzung der Bildattribute eines Objekts, um den Zustandsvektor zu erweitern. Dies führt zur verbesserten Objektverfolgung. Dadurch werden bewegte Objekte präziser identifiziert.

Diese Information wird im Kapitel 5 „Kartenbasierte Umfeldmodellierung“ ausführlich genutzt, um die durch dynamische Objekte verursachte Inkonsistenz zu eliminieren. Dies stellt im Vergleich zu in der Literatur bekannten Verfahren eine neue Alternative dar. Des Weiteren werden das Sensormodell für die Stereokamera sowie das Updateverfahren nach der Dempster-Shafer-Theorie beschrieben.

Kapitel 6 stellt die experimentellen Ergebnisse dar, die auf Basis der Vielzahl an Messdaten erreicht wurden.

Die Arbeit wird mit der Zusammenfassung und einen Ausblick in Kapitel 7 abgeschlossen.



# Kapitel 3

## Grundlagen zur Umfeldvermessung mittels Stereophotogrammetrie

Als Sensor für die Umfeldwahrnehmung in modernen Fahrerassistenzsystemen werden häufig optische Kameras eingesetzt. Die vielseitigen Informationen, die auf Basis der Kamerabilder gewonnen werden können, sind die Vorteile beim Einsatz dieses Sensors [NED07], [AKT10]. Die Bildverarbeitung beschäftigt sich mit einer Vielzahl an Verfahren, die aus dem Bildsignal unterschiedliche Informationen extrahieren [BUR05], [WOE09]. Die Grundlage zum Kameraaufbau sowie die projektive Geometrie wird in [LUH00] beschrieben.

Darüber hinaus werden Mehrkamarasysteme verwendet bzw. entwickelt, um 3D-Informationen über das Umfeld zu gewinnen [LUH00]. Einen Spezialfall stellt der Normalfall der Stereophotogrammetrie dar. Die meisten verwendeten Stereokamarasysteme in Forschungsbereich der Fahrerassistenzsysteme sind im Stereonormalfall. Da in dieser Arbeit die Stereokamera als Hauptsensor eingesetzt wird, werden auf Basis der Kamerabilder die stereophotogrammetrischen Berechnungen sowie die darauf aufbauende Objektdetektion eigenständig durchgeführt.

### 3.1 Der Stereonormalfall

Umfangreiche Grundlagen über die Bildgebung einer Kamera werden in [LUH00] beschrieben. In dieser Arbeit wird von einem Lochmodell der Kameras ausgegangen [SCHR05]. Beim Stereonormalfall werden zwei Kameras gleicher Bauart parallel zu einander gestellt. Die Kameras sollen theoretisch die gleiche Bildkonstante  $c$  besitzen. In der Praxis unterscheiden sich die Bildkonstanten selbst bei baugleichen Kameras geringfügig. Die Bildebe-

nen beider Kameras stehen zum einem auf einer selben Ebene, zum anderen auf gleicher Höhe. In Abbildung 3-1 werden die Lage beider Bildebenen zueinander visualisiert.

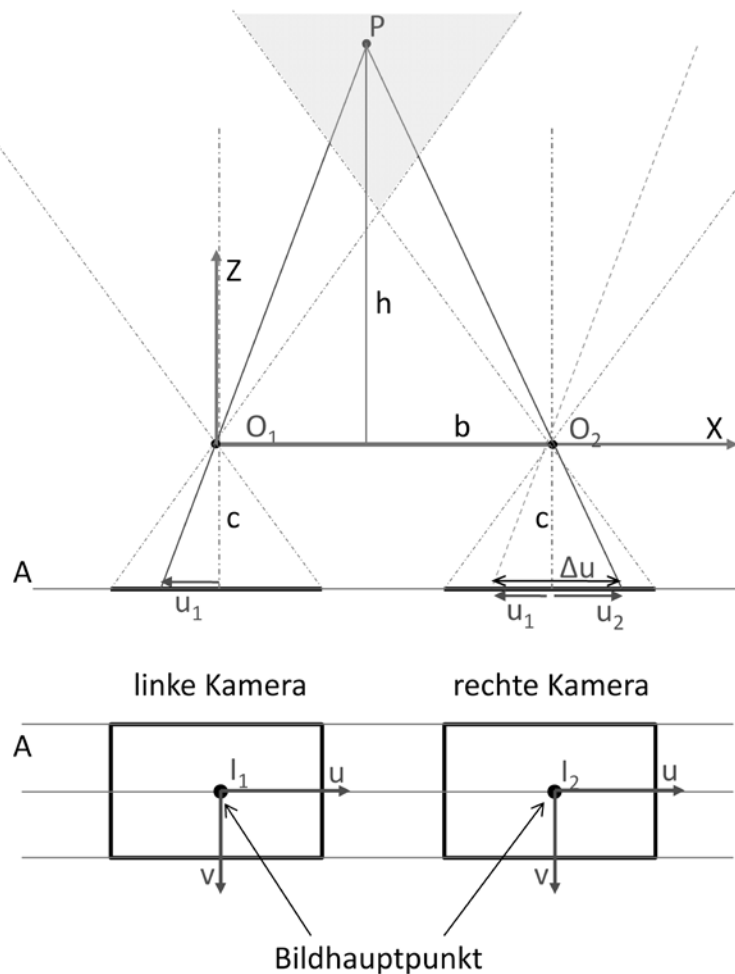


Abbildung 3-1: Prinzip der 3D-Berechnung bei der Stereophotogrammetrie im Normalfall

Die durch die beiden Bildhauptpunkte  $I_1$  und  $I_2$  verlaufende Gerade ist parallel zur oberen und unteren Kante der beiden Bildmatrizen und gehört ebenfalls zur gemeinsamen Ebene  $A$ . Unter dieser Bedingung erhält man eine vereinfachte geometrische Konstellation zur Stereoauswertung in Abbildung 3-1.

Dabei wird der Abstand zwischen zwei Kameras als Basisbreite  $b$ , die Entfernung eines 3D-Punkt  $P$  zum Projektionszentrum als  $h$  bezeichnet. Jeder Pixel auf der Bildebene wird mit den Bildkoordinaten  $(u, v)$  gekennzeichnet. Der graue Bereich stellt den so genannten „Sichtbereich“ des Stereokamera-Systems dar. Alle 3D-Objekte außerhalb dieses Bereiches

können nicht stereoausgewertet werden, da sie nicht gleichzeitig auf den Bildebenen beider Kameras abgebildet werden.

Alle 3D-Punkte werden hier im Kamerakoordinatensystem angegeben. Der Ursprung des Kamerakoordinatensystems wird am Projektionszentrum  $O_1$  der linken Kamera gelegt. Die Entfernung zum Ursprung wird als  $Z$  bezeichnet, der laterale Versatz als  $X$  und die Höhe als  $Y$ . Wird ein 3D-Punkt  $P$  auf beiden Bildebenen abgebildet, kann die Entfernung  $Z$  wie folgt berechnet werden:

$$Z = h = \frac{\mathbf{b} \cdot \mathbf{c}}{\mathbf{u}_1 - \mathbf{u}_2} \quad (3.1)$$

Dabei sind  $\mathbf{u}_1$  die korrespondierende  $u$ -Koordinate des Punktes  $P$  in der linken Kamera,  $\mathbf{u}_2$  die der rechten Kamera. Wir bezeichnen den Pixelversatz  $\Delta \mathbf{u} = \mathbf{u}_1 - \mathbf{u}_2$  zwischen zwei Bildpunkten als *Disparität*.

Des Weiteren lassen sich  $X$  und  $Y$  wie folgt ermitteln:

$$\begin{aligned} X &= \frac{\mathbf{b} \cdot \mathbf{u}_1}{\mathbf{u}_1 - \mathbf{u}_2} \\ Y &= \frac{\mathbf{b} \cdot \mathbf{v}_1}{\mathbf{u}_1 - \mathbf{u}_2} \end{aligned} \quad (3.2)$$

Mit (3.1) und (3.2) können theoretisch für alle 3D-Punkte im grauen Bereich die Koordinaten  $(X, Y, Z)$  bestimmt werden. Auf Basis von synthetischen Bildern wird dies in Abbildung 3-2 veranschaulicht.

Die rot eingekreisten Pixel in beiden Bildern, die aus demselben Objekt im Raum stammen, weisen eine Disparität in  $u$ -Richtung auf. Diese Disparität entspricht der Helligkeit des Pixels im Disparitätsbild, der die gleichen  $(\mathbf{u}_1, \mathbf{v}_1)$ -Koordinaten hat wie der aus dem linken Bild. Der Disparitätswert  $\Delta \mathbf{u}$  wird in Pixel berechnet. Aus (3.1) ist zu erkennen, dass Objekte in kleinen Abständen große Disparitätswerte besitzen. Der Disparitätswert ist umgekehrt proportional zum Abstand. Dies wird in Abbildung 3-2 veranschaulicht.

Es kann ausschließlich für Pixel im linken Bild, die im weiß gestrichelten Rahmen liegen, Disparitätswerte berechnet werden. Pixel außerhalb dieses Bereiches sind im rechten Bild nicht vorhanden.

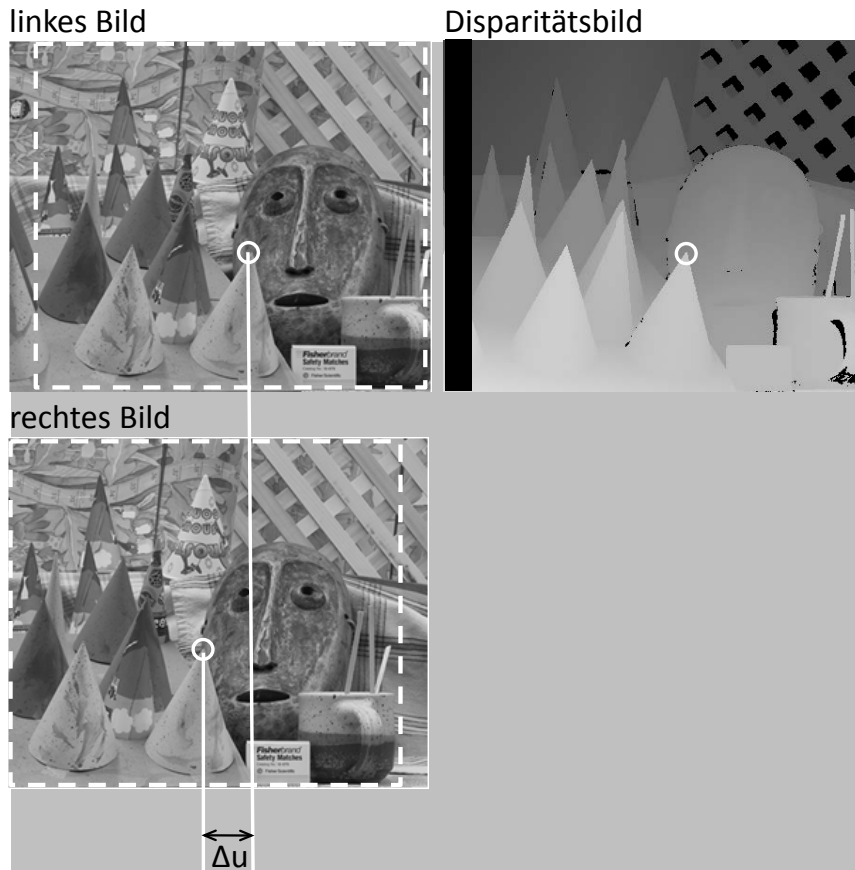


Abbildung 3-2: Disparitätsbild auf der Basis zwei Stereo-Bilder

Die Hauptherausforderung beim Einsatz in der Praxis besteht darin, dass für einen 3D-Punkt nicht immer zwei korrespondierende Punkte in beiden Bildebenen gefunden werden können. Man kann auf Basis seines Vorwissens korrespondierende Pixel leicht erkennen. Eine automatisierte Korrespondenzsuche ist dagegen eine Herausforderung. Im nächsten Unterkapitel wird auf die Korrespondenzanalyse zur Disparitätsberechnung eingegangen.

## 3.2 Korrespondenzanalyse zur Disparitätsberechnung

Das Ziel der Korrespondenzanalyse ist es, für jeden Pixel im linken Bild einen korrespondierenden Pixel im rechten Bild zu finden, der die Abbildung vom selben 3D-Punkt ist. Es existieren in der Literatur [WOE09] verschiedene Suchverfahren nach Korrespondenzen.

Die Grundidee ist dabei, dass korrespondierende Pixel ähnliche Merkmale aufweisen. Ein wichtiges Merkmal stellt die Lage eines Pixels zu seinem benachbarten Pixel dar. Dies bedeutet, es werden innerhalb eines vordefinierten Fensters, dem so genannten *Pixelblock*, die Grauwerte aller Pixel in Betracht gezogen. Der Referenzpixel, für den es im rechten Bild nach Korrespondenz gesucht wird, ist in der Mitte des Blocks. Der Block wird als *Referenzblock* bezeichnet.

Ausgehend von der Annahme des Stereonormalfalls wird auf derselben Zeile im rechten Bild Pixelblöcke mit dem Referenzblock im linken Bild verglichen. Diese Blöcke werden allgemein als *Suchblöcke* bezeichnet. Dieser Prozess wird in Abbildung 3-3 veranschaulicht. Ist ein Suchblock mit dem Referenzblock am ähnlichsten, wird der Suchpixel als Korrespondenzpixel bezeichnet. Die Testbilder wurden aus der Datenbank von [SCHA02] entnommen. Zur Berechnung der „Ähnlichkeit“ der Blöcke werden zahlreiche Verfahren eingesetzt. Im Folgenden werden die relevantesten Verfahren erläutert:

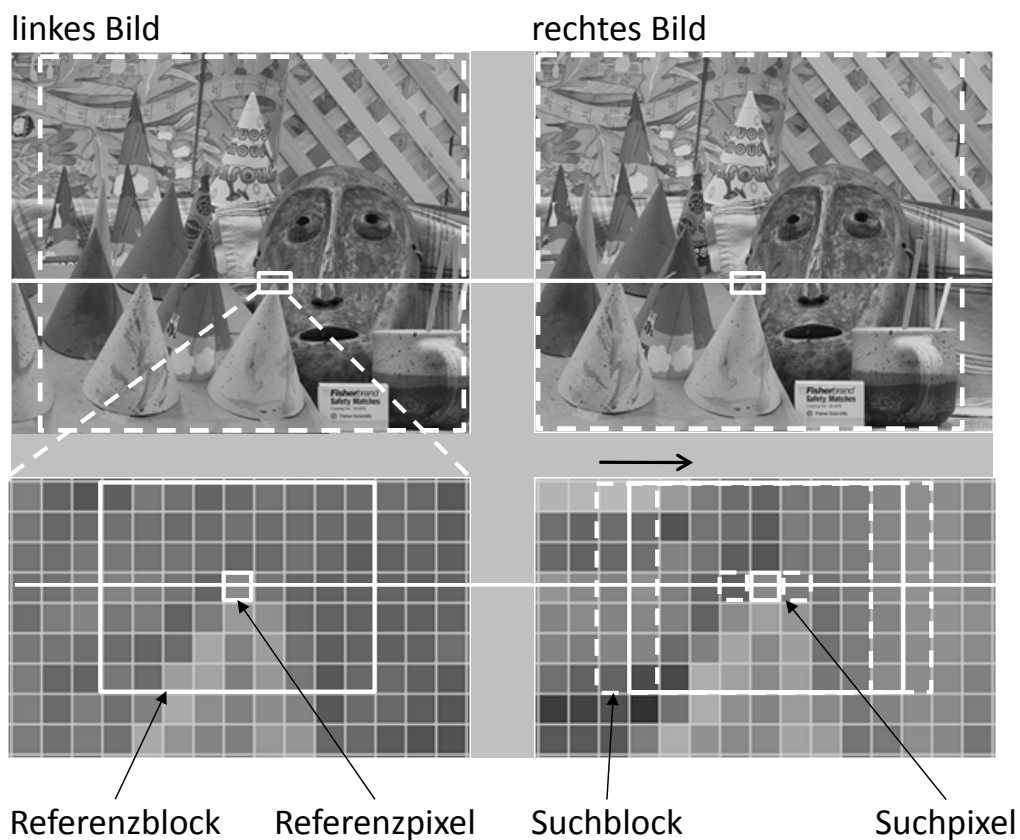


Abbildung 3-3: Korrespondenzsuche anhand des Blockmatching

### 3.2.1 Blockmatching-Verfahren

Am einfachsten ist das SAD-Verfahren (Sum of Absolute Differences). Daher wird in diesem Kapitel dieses Verfahren zur Veranschaulichung verwendet. Bei diesem Verfahren wird im ersten Schritt für jeden Pixel im Suchblock der Grauwertunterschied mit dem jeweiligen Pixel im Referenzblock berechnet. Danach werden alle Unterschiede summiert. Alternative Verfahren werden im Laufe des Kapitels erwähnt und diskutiert. Die SAD-Berechnung wird in die Formel (3.3) umgesetzt:

$$\text{SAD}_{(u,v)}(\xi) = \sum_{\frac{n-1}{2}}^{\frac{n-1}{2}} \sum_{\frac{m-1}{2}}^{\frac{m-1}{2}} \left| I_L(u+i, v+j) - I_R(u+\xi+i, v+j) \right| \quad (3.3)$$

Der Referenzpixel  $I_L$  liegt auf der Position  $(u, v)$  während  $I_R(u+\zeta, v)$  den Suchpixel auf der gleichen Zeile  $v$  darstellt.  $n$  und  $m$  sind dabei die Blockgrößen in lateraler und longitudinaler Richtung und immer ungerade.  $(i, j)$  gibt die Position des Pixels im Block an. Die Verschiebung  $\zeta$  unterliegt einem maximalen Disparitätswert, der auf Basis (3.1) und Pixelgröße berechnet werden kann. In Abbildung 3-3 wird beispielsweise für den Pixel an der Spitze des Kegels im linken Bild der korrespondierende Pixel im rechten Bild gesucht. Die Suchblockgröße wurde mit 7 Pixeln gewählt.

Im rechten Bild wird entlang der gleichen Linie innerhalb des Bereichs  $\xi = 25$  der SAD-Wert für alle Pixel berechnet. Der Verlauf der Funktion wird in Abbildung 3-4 dargestellt.

Bei der Korrespondenzsuche für den Pixel im linken Bild in Abbildung 3-3 wird derjenige Pixel im rechten Bild als „Match“ gewählt, der den kleinsten SAD-Wert aus Formel (3.3) aufweist. In Abbildung 3-4 ist ein Minimum an der Stelle 49 zu erkennen. Nach Groundtruth-Daten beträgt der wahre Disparitätswert 50 Pixel. Damit stimmt der detektierte Wert mit dem tatsächlichen gut überein.

Die Funktion  $\text{SAD}_{(u,v)}(\xi)$  wird auch in verschiedenen Literaturstellen [MUL05] als Kostenfunktion bezeichnet. Neben der SAD existieren noch verschiedene Kostenverfahren wie SSD (Sum of Squared Differences), KKF (Kreuzkorrelationsfunktion).

Für den Fall, dass das linke und rechte Bild Helligkeitsdifferenzen aufweisen, werden mittelwertfreie Verfahren eingesetzt wie SADMF, SSDMF; KKFMF [TOR09]. Bei solchen Verfahren werden die Grauwerte im Bezug auf den Referenzpixel normiert. Dadurch spielen nur die Verhältnisse der Grauwerte zu einander eine Rolle bei der Korrespondenzsuche und nicht mehr die Grauwerte selbst. Ein auch sehr verbreitetes Verfahren ist die CENSUS-Transformation [STE04].

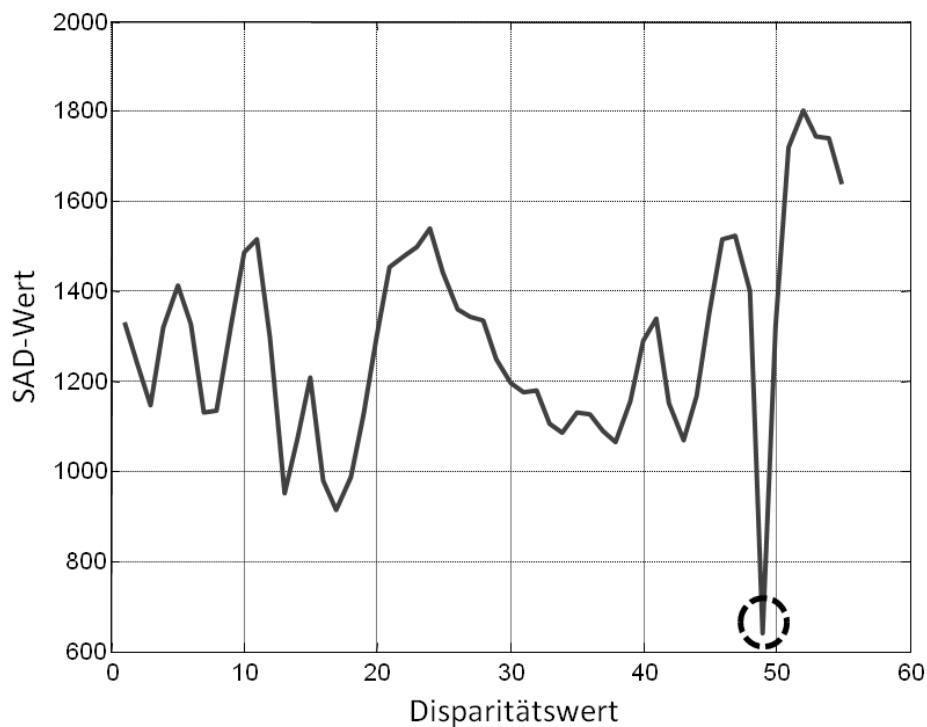


Abbildung 3-4: SAD-Kostenverlauf

Im Gegensatz zum SAD-Algorithmus, der auf den absoluten Helligkeitswerten der Pixel in jedem Block beruht, basiert die Census-Transformation auf einer relativen Aussage der Helligkeitsverteilung in den betrachteten Fenstern. Die Census-Transformation ist somit robuster als die SAD gegenüber schwankenden Belichtungssteuerungen oder unterschiedlichen Helligkeiten beider Bilder.

In dem Verfahren werden die Nachbarschaftspixel entsprechend ihrer relativen Helligkeit bezüglich des Pixels im Zentrum bewertet. Weist ein Pixel eine höhere Helligkeit als das Referenzpixel auf, wird es mit einer Eins beaufschlagt. Bei einem geringeren oder gleichen Helligkeitswert wird eine Null zugewiesen. Daraufhin erhält jeder Pixel einen sogenannten

„Census Vector“ mit der entsprechenden Bit-Folge. Die Länge des „Census-Vectors“ entspricht der Anzahl der verglichenen Bildpunkte im Messfenster.

Die Ähnlichkeit von zwei Bildpunkten wird dann über die Summe von Hamming-Distanzen berechnet (SAD-Prinzip). Die Hamming-Distanz gibt dabei die Anzahl der unterschiedlichen Bits in zwei Bitfolgen an. Des Weiteren zeichnet sich dieses Verfahren durch die Zuverlässigkeit und Kompatibilität zur Echtzeitsimplementierung aus [GRA11]. In diesem Kapitel wird zur Veranschaulichung nur die SAD-Funktion verwendet.

Ein typisches Problem bei der Korrespondenzsuche wird am Beispiel des Pixels in Abbildung 3-5 dargestellt. Es ist zu erkennen, dass die Pixel in ihrer Nachbarschaft ähnliche Grauwerte besitzen. Dies führt zu einem SAD-Verlauf wie in Abbildung 3-6. In dem Versatzbereich von 33 bis 41 Pixel sind ähnlich geringe SAD-Werte entstanden.

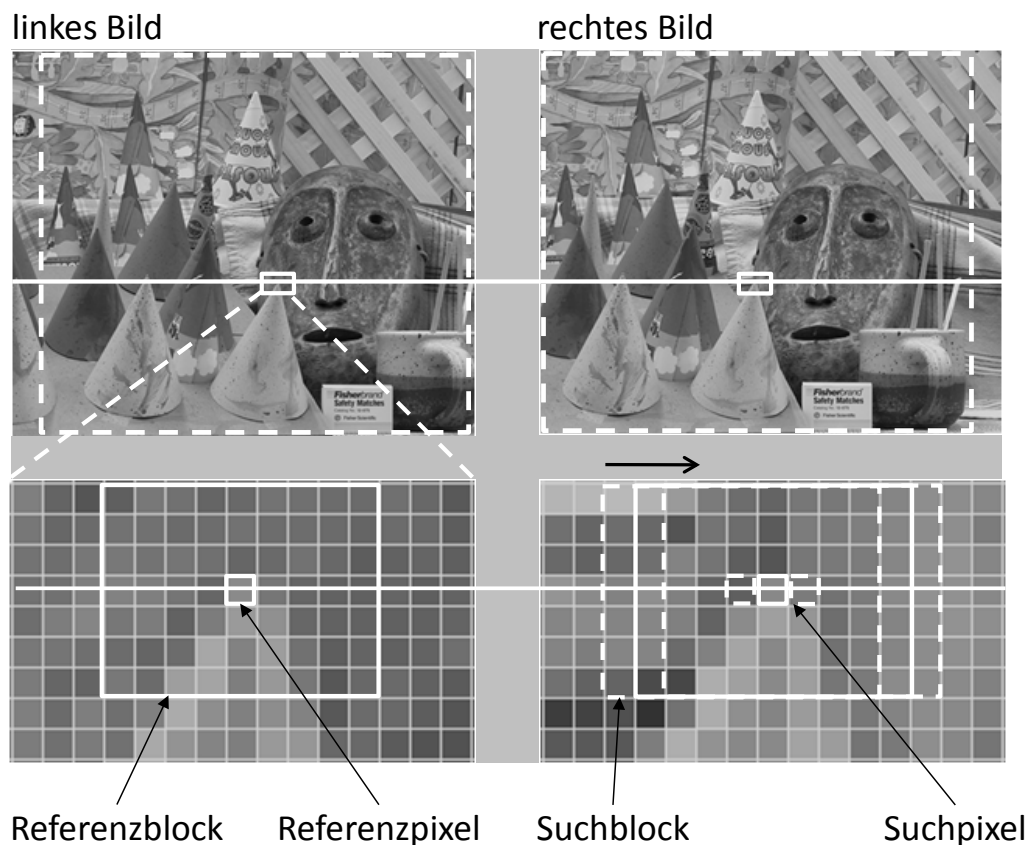


Abbildung 3-5: Problemfall bei der Korrespondenzsuche – texturarme Regionen veranlassen ähnliche Blöcke

Dadurch ist es eine Herausforderung, den richtigen Korrespondenzpixel zu finden. Auf die Behandlung dieses Problems wird im nächsten Unterkapitel eingegangen. Im Allgemeinen



kann man sagen, dass je unterschiedlicher die Pixel innerhalb eines Suchfensters sind, desto eindeutiger sich das Minimum der SAD-Werte ausprägt.

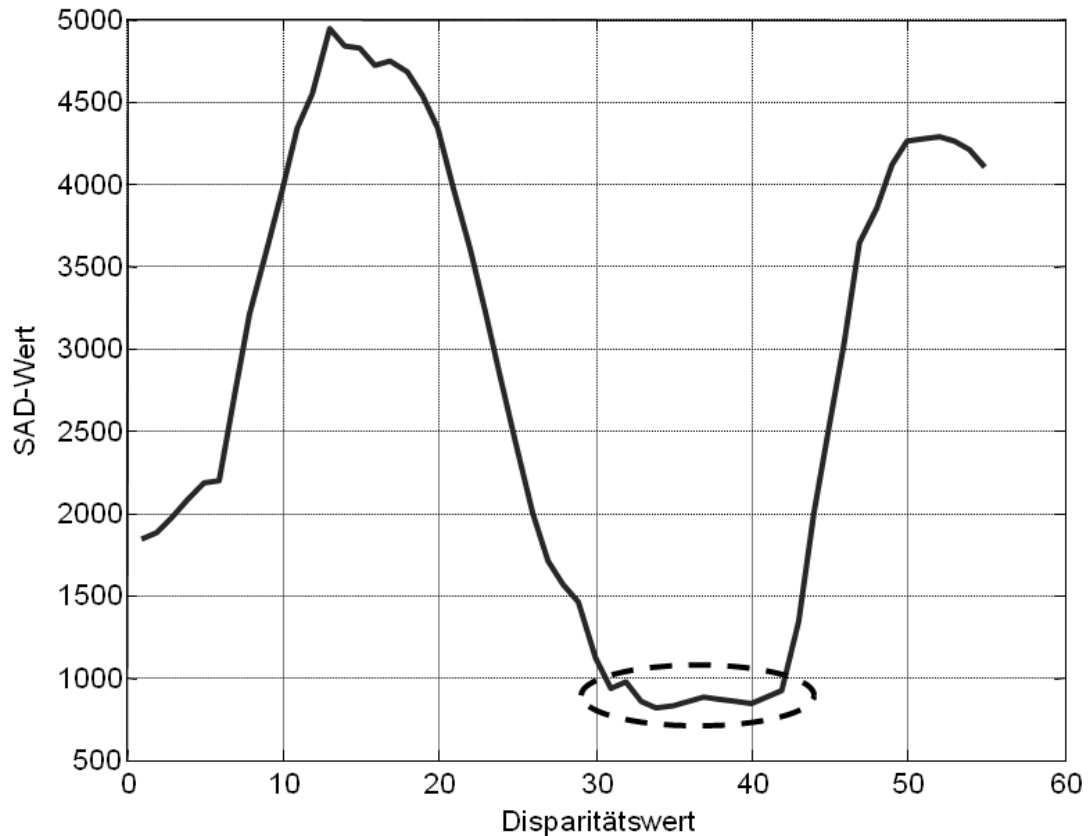


Abbildung 3-6: SAD-Kostenverlauf bei texturarmen Regionen

### 3.2.2 Optimierung der Korrespondenzsuche

Um das Problem von räumlich gestreckten Minima wie in Abbildung 3-6 zu beheben, ist es zuerst wichtig zu erkennen, dass kein eindeutiges Minimum existiert. In [GRA11] werden beispielsweise die drei kleinsten Minimum-Werte mit einander verglichen. Ist der kleinste Wert nicht eindeutig geringer als die zwei nächsten kleinsten, wird das Minimum verworfen.

Des Weiteren wird die Korrespondenzsuche parallel sowohl für das linke als auch für das rechte Bild durchgeführt. Nur Pixelpaare, die sich in beiden Bildern als Korrespondierende zeigen, werden zur Disparitätsberechnung verwendet. Durch diese Verifikationsmechanis-

men kann die Zuverlässigkeit der Disparitätsberechnung erhöht werden [SCHR05], [GRA11].

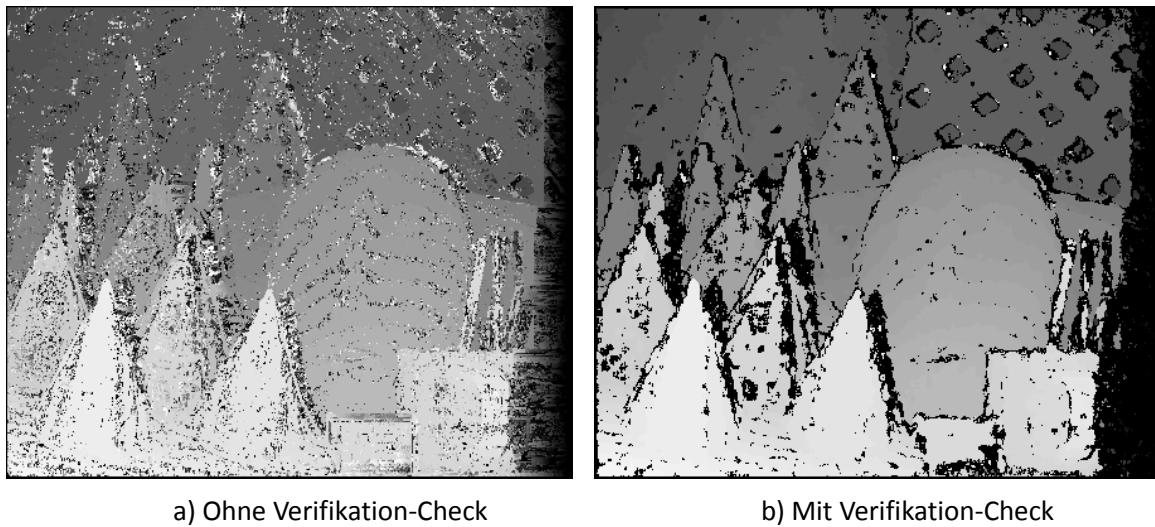


Abbildung 3-7: Gegenüberstellung der Disparitätsbilder. Bild a) wird mit der SAD-Funktion berechnet, Suchfenstergröße: 5x5 Pixeln. Bild b) wird auch mit der SAD-Funktion gleicher Suchfenstergröße berechnet, danach aber mit obigen Verifikationsmechanismen geprüft.

In Abbildung 3-7 sind zwei Disparitätsbilder dargestellt. In Teil a) ist das Disparitätsbild ohne die oben beschriebene Verifikation der korrespondierenden Pixel. Teil b) ist das Disparitätsbild mit Pixelverifikation. Es ist zu erkennen, dass das Disparitätsbild in Teil b) deutlich weniger verrauscht ist, im Vergleich zum Teil a). Es sind insgesamt zwar weniger Disparitätspunkte in Teil b) vorhanden, die Punkte sind aber überwiegend korrekt gematcht.

Um dennoch in texturarmen Regionen Korrespondenzen zu finden, werden verschiedene Ansätze praktiziert.

Einen guten Ansatz liefert zum Beispiel das hierarchische Verfahren [TOR09]. Es werden mehrere Stereo-Bildpaare mit unterschiedlichen Auflösungen aus dem originalen Paar erzeugt. Dieses entspricht einer so genannten Auflösungs-pyramide. Es wird als erstes für die niedrigste Auflösung die Disparitätskarte berechnet. Diese stellt die Grundlage für die Berechnung mit der nächsten Auflösung dar. Die Disparitätswerte werden mit steigender Auflösung verfeinert [TOR09].

Neben pixelbasierten Disparitätsberechnungsverfahren, auch als *lokale Methoden* bezeichnet, sind noch *globale Methoden* entwickelt worden [COX96]. Diese betrachten die Korrespondenzsuche als eine Energie-Funktion und adressieren die Lösung der Funktion als ein Optimierungsproblem [MAT89], [BOY01]. Die Kernidee dieser Verfahren ist es, nicht für jeden einzelnen Pixel ein richtiges Match zu finden, sondern die gesamte Disparitätskarte muss eine optimale Lösung sein. Dafür wird folgende Energie-Funktion definiert:

$$E_{\text{total}}(\mathbf{D}) = E_{\text{data}}(\mathbf{D}) + \lambda \cdot E_{\text{smooth}}(\mathbf{D}) \quad (3.4)$$

Wobei  $\mathbf{D}$  die Disparitätskarte,  $\lambda$  der Gewichtungsfaktor zwischen zwei Größen  $E_{\text{data}}(\mathbf{D})$  und  $E_{\text{smooth}}(\mathbf{D})$  der gesamten Energiefunktion  $E_{\text{total}}(\mathbf{D})$  sind.

Die Disparität-Energiefunktion  $E_{\text{data}}(\mathbf{D})$  ist definiert als eine Funktion der Bildintensitätsdifferenzen:

$$E_{\text{data}}(\mathbf{D}) = \sum_{\underline{\mathbf{p}}} \left| I_{\text{L}}(\underline{\mathbf{p}}) - I_{\text{R}}(\underline{\mathbf{p}} + \mathbf{D}(\underline{\mathbf{p}})) \right| \quad (3.5)$$

Wobei  $\underline{\mathbf{p}} = [u, v]$  die Pixelposition,  $I_{\text{L}}(\underline{\mathbf{p}})$  die Intensität des Pixel an der Position  $\underline{\mathbf{p}}$  im linken Bild,  $I_{\text{R}}(\underline{\mathbf{p}} + \mathbf{D}(\underline{\mathbf{p}}))$  die Pixelintensität des korrespondierenden Pixel im rechten Bild sind.

Der zweite Term in (3.4)  $E_{\text{smooth}}(\mathbf{D})$  sorgt dafür, dass benachbarte Disparitätspixel zusammenhängende Werte haben und wird wie folgt definiert:

$$E_{\text{smooth}}(\mathbf{D}) = \sum_{\underline{\mathbf{p}}} \sum_{\underline{\mathbf{q}}} \mathbf{V}[\mathbf{D}(\underline{\mathbf{p}}), \mathbf{D}(\underline{\mathbf{q}})] \quad (3.6)$$

Die Funktion  $\mathbf{V}[\mathbf{D}(\underline{\mathbf{p}}), \mathbf{D}(\underline{\mathbf{q}})]$  wird als *Interagierende Funktion* (Engl.: *interacting function*) eingeführt. Die Form von  $\mathbf{V}$  prägt maßgeblich die gesamte Energie-Funktion aus (3.4). Die Definition von Gleichmäßigkeit zwischen benachbarten Disparitätspixeln wird durch  $\mathbf{V}$  ausgedrückt. Diese Funktion wird in vielen Literaturstellen unterschiedlich definiert [POG85], [GAM87]. Zum Beispiel als monoton steigende Funktion [POG85]:

$$V[\mathbf{D}(\underline{\mathbf{p}}), \mathbf{D}(\underline{\mathbf{q}})] = |\mathbf{D}(\underline{\mathbf{p}}) - \mathbf{D}(\underline{\mathbf{q}})|^\alpha, \quad \alpha = 1 \text{ or } 2 \quad (3.7)$$

Oder alternativerweise als eine Sprung-Funktion [POG85]:

$$V[\mathbf{D}(\underline{\mathbf{p}}), \mathbf{D}(\underline{\mathbf{q}})] = \begin{cases} 0 & \dots \text{if } |\mathbf{P}(\underline{\mathbf{p}}) - \mathbf{D}(\underline{\mathbf{q}})| < T \\ c & \dots \text{if } |\mathbf{P}(\underline{\mathbf{p}}) - \mathbf{D}(\underline{\mathbf{q}})| > T \end{cases} \quad (3.8)$$

Mit (3.5), (3.7) oder (3.8) werden die gesamte Energie-Funktion für alle Pixel der Disparitätskarte bestimmt. Eine Disparitätskarte wird als die „richtige“ angenommen, wenn die gesamte Energie-Funktion minimal wird.

$$\mathbf{D}^* = \arg \min \mathbf{E}_{\text{total}}(\mathbf{D}) \quad (3.9)$$

Aus allen möglichen Disparitätskarten wird mit  $\mathbf{D}^*$  die minimale gesamte Energie-Funktion erreicht. Damit soll theoretisch die korrekteste Disparitätskarte erzielt werden.

In der Praxis, besonders bei echtzeitfähigen Systemen, sind globale Ansätze wegen hohem Rechenaufwand nicht immer geeignet. Des Weiteren wurde in [MAR06] gezeigt, dass Algorithmen, die globale Ansätze verwenden, mehr von der Schwankung der getroffenen Szenebedingungen beeinflusst werden als lokale Ansätze, die auf einfacheren Korrespondenzkriterien basieren. Als Alternative dazu wird der Ansatz „Semiglobal-Matching“ in [MUL05] vorgestellt. Dieser Ansatz approximiert das Minimum der 2D-Energie-Funktion aus (3.4) durch Minimierung mehrerer 1D-Energie-Funktionen [HER09]. Diese 1D-Funktionen werden entlang der Kostenpfade berechnet, die am Rand des Bildes starten und beim gesuchten Pixel enden.

Dies wird in Abbildung 3-8 veranschaulicht. Die Energie-Funktion für jeden Pfad setzt sich aus einer Data-Funktion und zwei Glätte-Funktionen zusammen. Die Data-Funktion wird wie die in 3.2.1 aufgeführten Kostenfunktionen berechnet. Die erste Glätte-Funktion vergibt Kosten  $c_1$  für kleine Disparitätsänderungen zwischen benachbarten Pixeln. Die Kosten  $c_1$  werden gering gehalten, um schräge Ebenen zuzulassen.

Die zweite Glätte-Funktion vergibt größere Kosten für größere Disparitätsänderungen. Die Kosten für einen Pfad werden rekursiv wie folgt berechnet:

$$L_r(\underline{p}_i, d) = C(\underline{p}_i, d) + \min \left[ L_r(\underline{p}_{i-1}, d), L_r(\underline{p}_{i-1}, d-1) + c_1, L_r(\underline{p}_{i-1}, d+1) + c_1, \right. \\ \left. \min_{\Delta} L_r(\underline{p}_{i-1}, \Delta) + c_2 \right] - \min_{\Delta} L_r(\underline{p}_{i-1}, \Delta) \quad (3.10)$$

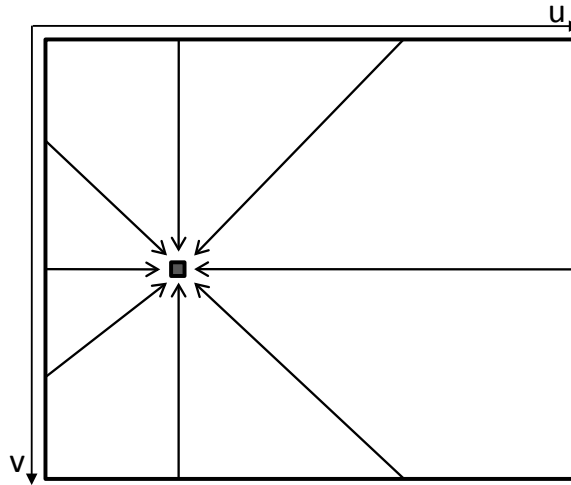


Abbildung 3-8: Kostenpfade für die Berechnung der gesamten 1D-Energie. Für den rot gekennzeichneten Pixel wird ein Disparitätswert gesucht.

Wobei  $\mathbf{r}$  der aktuelle Pfad,  $C(\underline{p}_i, d)$  die Data-Funktion,  $\underline{p}_i$  der aktuelle Pixel,  $\underline{p}_{i-1}$  der vorige Pixel,  $\Delta$  die größere Disparitätsabweichung,  $L_r(\underline{p}_i, d)$  die bis Stelle  $i$  rekursiv berechneten Kosten sind. In der Regel reichen acht Pfade aus, um eine gute Disparitätskarte zu erhalten. Es werden die Kosten aller Pfade addiert. Der Disparitätswert, durch den die minimalen Kosten ergeben, wird angenommen.

$$d^* = \arg \min_{\mathbf{r}} \sum_{\mathbf{r}} L_r(\underline{p}, d) \\ \text{mit } d \in [d_{\min}, d_{\max}] \quad (3.11)$$

Neben pixelbasierter Disparitätsberechnung sind noch segmentbasierte Verfahren zu erwähnen. Bei dem Verfahren in [BLE04], [TAO00] wird angenommen, dass homogene Regionen eine Ebene im Disparitätsbild darstellen. Dadurch kann für alle Pixel auf der Ebene dementsprechend ein Disparitätswert zugeordnet werden. Pixel im Randbereich solcher

Segmente haben durch Blockmatching einen guten Disparitätswert. Dieser Wert wird für alle dem Segment zugehörigen Pixel angenommen.

Neben nicht eindeutigen Minima des Gütekriteriums sind Verdeckungen eine Fehlerquelle in der Disparitätsberechnung [ZIT99]. Es ist nicht möglich, Korrespondenzen für verdeckte Pixel zu finden. Jedoch ist es sinnvoll, verdeckte Pixel zu identifizieren, um falsches Matching zu vermeiden. Lösungsansätze dazu werden in [ZIT99] und [BEL92] dargestellt. Eine ausführliche Evaluierung aktueller Algorithmen zur Disparitätsberechnung kann unter [SCHA02] gefunden werden.

In dieser Arbeit wurde der Semiglobal-Matching-Algorithmus eingesetzt. Das Ergebnis wird beispielsweise in Abbildung 3-9 gezeigt. Es ist zu erkennen, dass das Disparitätsbild anhand des klassischen Blockmatchingverfahrens im Vergleich zu dem Ergebnis mit Semiglobal-Matching weniger Matches hat. Für die Objektdetektion auf Basis von Stereokamera ist eine große Anzahl an 3D-Punkten von entscheidendem Vorteil.

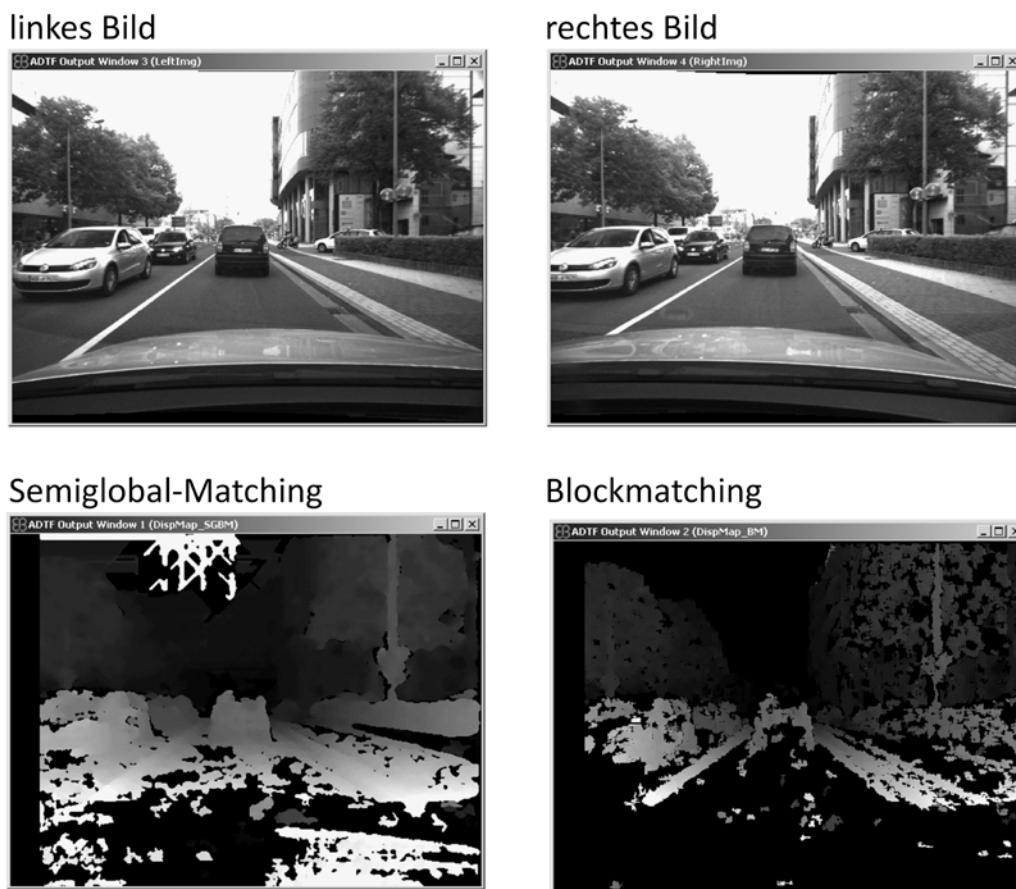


Abbildung 3-9: Disparitätsberechnung mit Semiglobal-Matching und Blockmatching

### 3.3 3D Rekonstruktion

Auf Basis des berechneten Disparitätsbild können unter Annahme des Stereonormalfalls die 3D-Koordinaten jedes einzelnen Pixels mit den Formeln (3.1) und (3.2) berechnet werden. Die Genauigkeit der 3D-Rekonstruktion hängt direkt von der Qualität der Disparitätsberechnung ab. In [LUH00] wird die Genauigkeit einer 3D-Rekonstruktion genau beschrieben.

Im Allgemeinen lässt sich sagen, dass die Genauigkeit mit steigender Entfernung antiproportional abnimmt. In Abbildung 3-10 werden die rekonstruierten 3D-Punkte des Verkehrsumfeldes veranschaulicht.

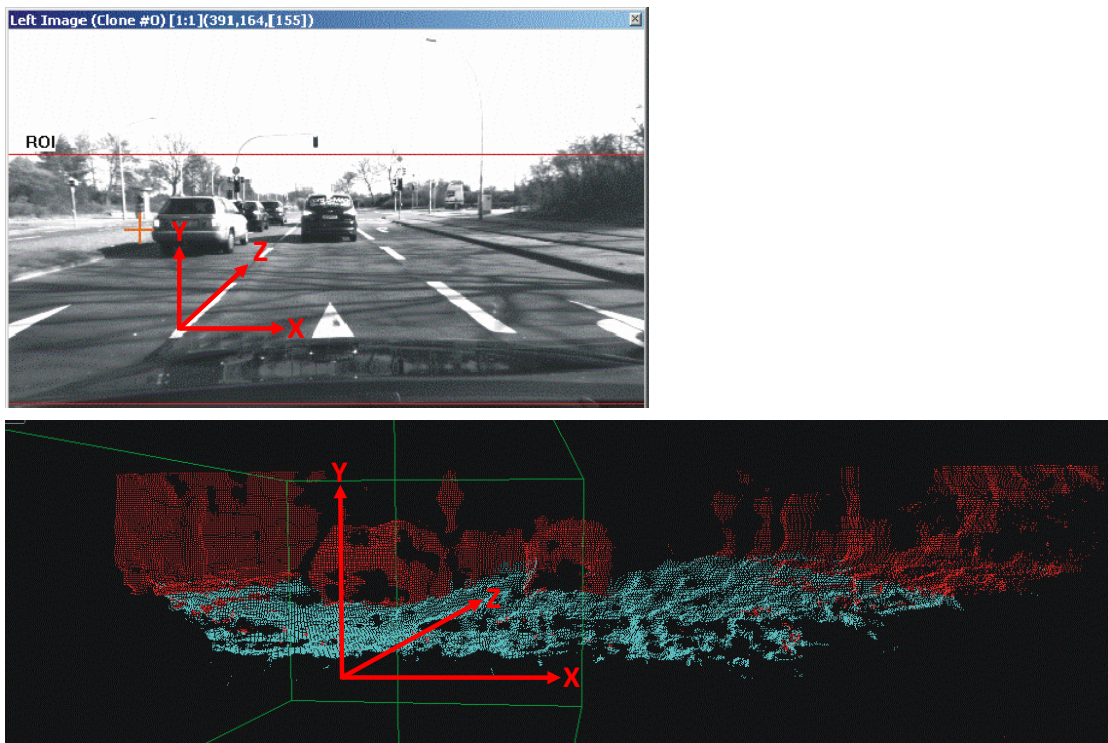


Abbildung 3-10: rekonstruierte 3D-Punkte einer Verkehrsszene. Blaue Punkte repräsentieren quasi den Boden. Rote Punkte repräsentieren Objekte und Hindernisse. [NED07]

Alternativ kann zur 3D-Rekonstruktion auch der optische Fluss verwendet werden [WAR85]. Als optischer Fluss wird ein Vektorfeld bezeichnet, das die Bewegungsrichtung und -Geschwindigkeit für jeden Bildpunkt einer Bildsequenz angibt. In der Literatur sind verschiedene Verfahren zur Berechnung des optischen Flusses bekannt [HOR81], [LUC81], [FLE05].

Die Erkenntnis über die Bewegung eines Pixels in einer Bildsequenz kann in Kombination mit der Information über die Bewegung der Kamera/des eigenen Fahrzeugs verwendet werden, um die 3D-Information desjenigen Pixels zu berechnen [FAU93].

Das Ergebnis der Anwendung dieser Methode zeigt im Vergleich zur 3D-Rekonstruktion mit Stereokamera (Normalfall) jedoch keine relevante Verbesserung, besonders beim Auftreten mehrerer bewegter Objekte. Zum gegenwärtigen Zeitpunkt bedarf die 3D-Rekonstruktion auf Basis des optischen Flusses mehr Rechenleistung. Deshalb wird in dieser Arbeit lediglich die 3D-Rekonstruktion auf Basis der Stereokamera durchgeführt.



# Kapitel 4

## Detektion und Tracking von bewegten Objekten auf der Basis von rekonstruierten 3D-Informationen

Die rekonstruierten 3D-Informationen bieten zusammen mit den Bildinformationen ein großes Potenzial zur Objektdetektion. Aktuelle Arbeiten wie [BRO05],[NED07] zeigen gute Beispiele wie aus solchen Informationen Objekte im Umfeld wahrgenommen werden können. Am Anfang dieses Kapitels wird eine effiziente Methode zur Objektdetektion beschrieben. Davon ausgehend wird eine Tracking-Strategie erläutert und angewendet, um bewegte Objekte zuverlässig über mehrere Messungen zu verfolgen. Das Kapitel schließt mit einer Zusammenfassung und Diskussion ab.

### 4.1 Objektdetektion durch Segmentierung

Das Ziel der stereobasierten Objektdetektion ist es, aus dem Tiefenbild 3D-Punkte, die zu einem Objekt im Fahrzeugsumfeld gehören, zu einem Segment zusammenzufügen. Abbildung 4-1 veranschaulicht dies. Man kann sagen, der Begriff der Objektdetektion wird in dieser Hinsicht auf die Segmentierung zusammengehöriger 3D-Punkte reduziert. Da wesentliche Objekte sowie Hindernisse im Verkehrsgeschehen zwischen 0m - 2m Höhe liegen, werden nur 3D-Punkte in diesem Bereich berücksichtigt. Weiterhin erfolgt noch eine Bodenextraktion, um alle Punkte, die zum Boden gehören, herauszufiltern [LAB05].

In den Arbeiten von [REH97], [AIH06b] wurden verschiedene Techniken zur Bildsegmentierung bei Farbbildern dargestellt. Bei Anwendung dieser Techniken auf das Tiefenbild wur-

den jedoch nicht die gewünschten Ergebnisse erzielt. Ein Disparitätsbild besitzt in der Tat andere Merkmale als Farbbilder.

Als weiterer Grund dafür können die kleineren Lücken im Disparitäts-/Tiefenbild genannt werden. Die in [REH97] beschriebenen Algorithmen sind nicht für diesen Fall optimiert. Deshalb wurde ein eigenes Verfahren entwickelt, um die Segmentierung auf Basis der Disparitätskarte durchzuführen.



Abbildung 4-1: Ziel der Objektdetektion ist es, Pixel, die zu einem Objekt in realer Welt gehören, zu einem Segment zusammenzuführen.

Auf Basis der rekonstruierten 3D-Punkte wird ein zweidimensionales Histogramm aufgebaut. (Siehe Abbildung 4-2). In x-Richtung (Fahrzeugkoordinaten) hat das Histogramm 500 Einheiten oder Zellen, in y-Richtung 200. Jede Zelle repräsentiert einen Bereich von 0,1m x 0,1m.

Damit deckt das Histogramm einen Bereich von -10m bis 10m von links nach rechts und bis 50m vor dem Ego-Fahrzeug ab. Für jeden 3D-Punkt, der in diesem Bereich liegt und dessen Höhe geringer als 1.5m ist, wird der Wert der Zelle an der Position des 3D-Punktes *inkrementiert*. 3D-Punkte, die höher sind als 1.5m, sind für das Ego-Fahrzeug irrelevant.

Das Histogramm drückt sozusagen die Häufigkeit der 3D-Punkte über der Fläche aus. Hat die Zelle mehr als eine bestimmte Anzahl von Punkten, wird die Zelle als *Objektzelle* eingestuft. Auf Basis dieses Histogramms wird der Segmentierungsalgorithmus durchgeführt.

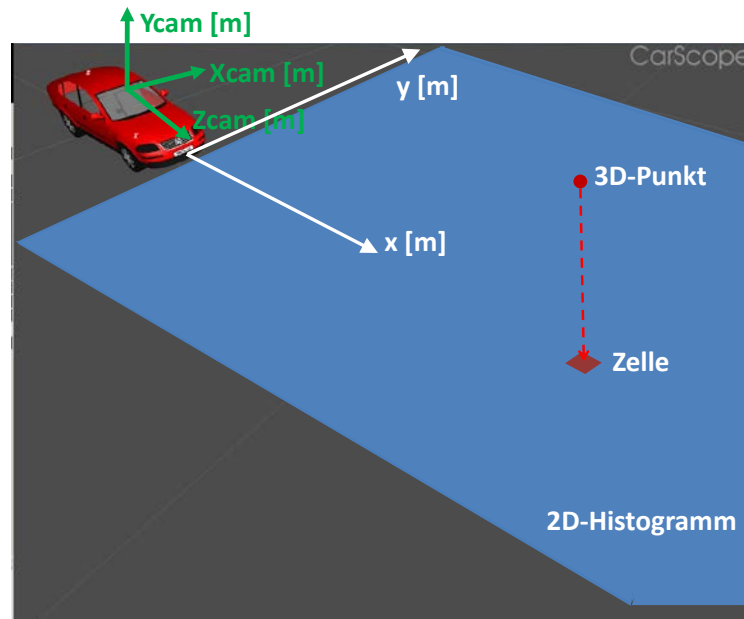


Abbildung 4-2: Histogramm zur Beschreibung der Häufigkeit der 3D-Punkte über der Fläche (2D-Histogramm)

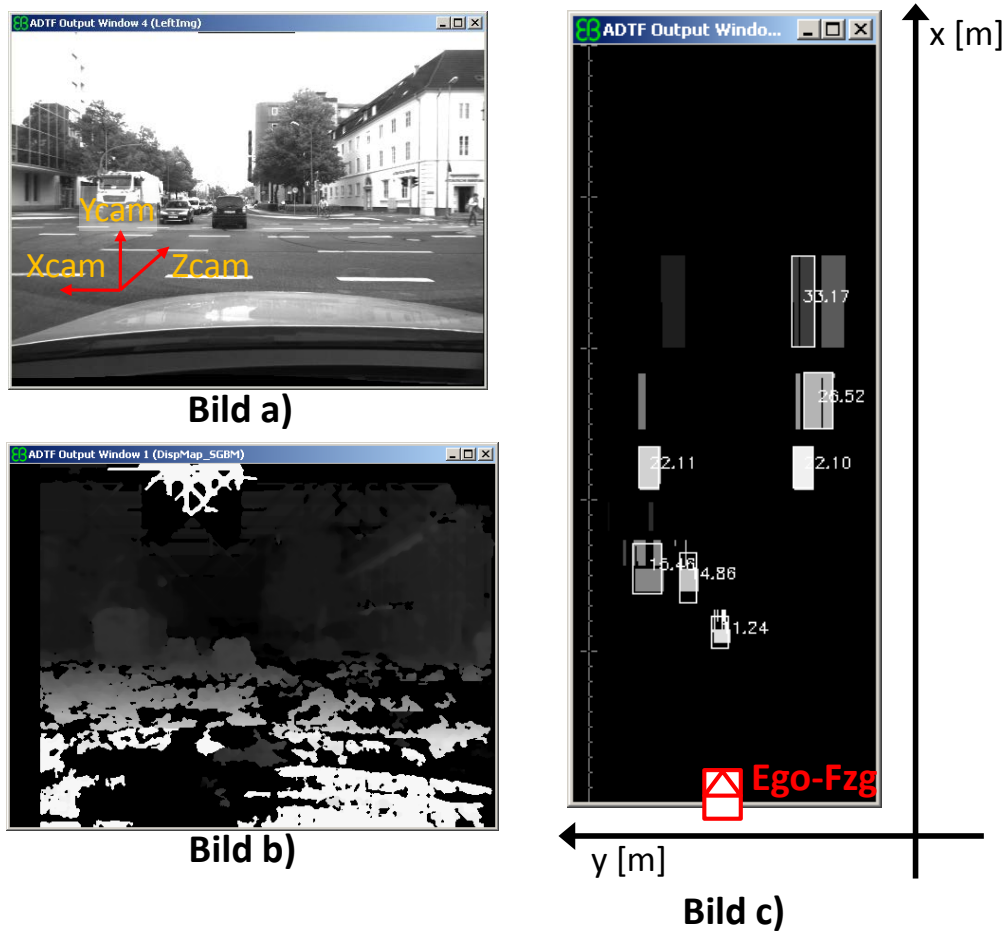


Abbildung 4-3: Teil a) das linke Kamerabild. Teil b) Das Disparitätsbild. Teil c) Das 2D-Histogramm, das auf Basis der aus dem Disparitätsbild rekonstruierten 3D-Daten erstellt wird.

Der Segmentierungsalgorithmus beinhaltet zwei Schritte, eine *Linienoperation* und eine *Blockoperation*. In Abbildung 4-4 wird der erste Schritt veranschaulicht. Es wird von links nach rechts entlang der Linie nach Zellen gesucht, die als Objektzellen klassifiziert wurden. Wird eine Objektzelle gefunden, wird überprüft, ob die Zelle eine ID-Nummer besitzt (am Anfang hat keine Zelle eine ID-Nummer). Besitzt die Objektzelle keine ID-Nummer, wird der Zelle eine noch nicht vergebene ID-Nummer zugewiesen, falls in einem vorher festgelegten Radius um diese Zelle (in Abbildung 4-4 grün eingezeichnet) keine Objektzelle mit ID-Nummer vorhanden ist.

Ist dagegen eine ID-Nummer vorhanden, wird diese Nummer derjenigen Zelle zugewiesen. Nach diesem Schritt hat die gefundene Objektzelle entweder eine neue ID oder die ID einer benachbarten Objektzelle. Sobald das Ende der Zeile erreicht wird, bearbeitet man die nächste Zeile bearbeitet.

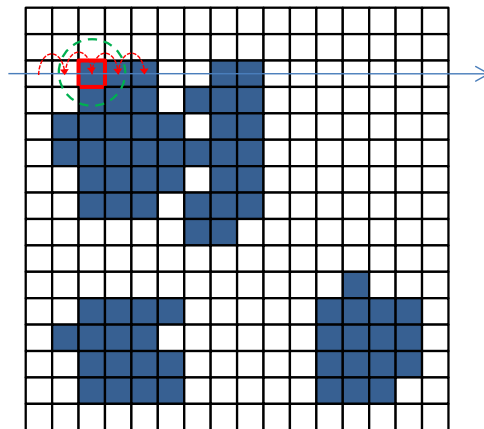


Abbildung 4-4: Linienoperation

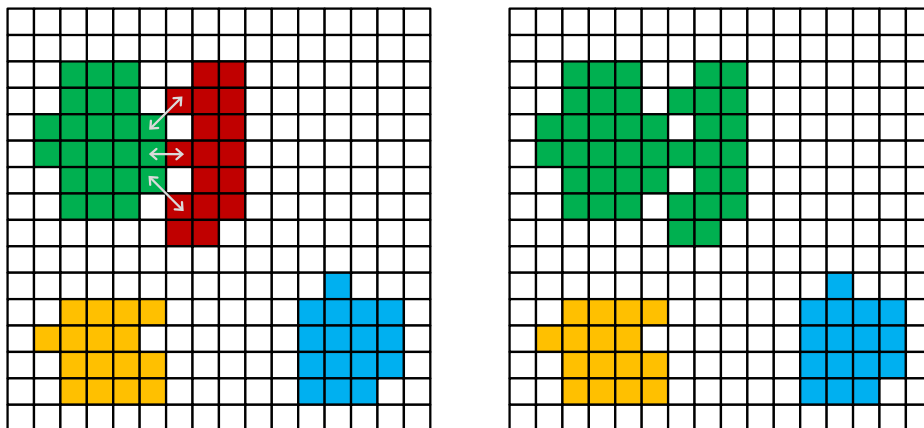


Abbildung 4-5: Blockoperation

Die Linienoperation beginnt mit der obersten Zeile des Histogramms. Das Ergebnis des ersten Schrittes der Segmentierung wird in Abbildung 4-5 links beispielhaft dargestellt. Jede Farbe repräsentiert ein unterschiedliches Objekt. In Wirklichkeit gehören das grüne und rote Segment zu einem Objekt und sollten daher als ein Segment dargestellt werden. Im zweiten Schritt werden die Segmente auf Abstände zueinander überprüft. Grenzen zwei Segmente durch mehrere Zellen aneinander, wie das rote und grüne Segment in Abbildung 4-5, werden sie zu einem Segment verschmolzen.

Die Position und Größe jedes Segments werden am Ende des Vorgangs berechnet. Die mittlere Zellenposition eines Segments stellt die Y- und X-Position eines Objektes dar.

$$x_{\text{Obj}} = \frac{1}{N} \sum_{i=1}^N x_{\text{cell}}(i), \quad y_{\text{Obj}} = \frac{1}{N} \sum_{i=1}^N y_{\text{cell}}(i) \quad (4.1)$$

Wobei N die Anzahl der Zellen im Segment ist.

Die Objektlänge und -breite werden wie folgt berechnet:

$$L_{\text{Obj}} = z_{\text{cell}}^{\max} - z_{\text{cell}}^{\min}, \quad B_{\text{Obj}} = x_{\text{cell}}^{\max} - x_{\text{cell}}^{\min} \quad (4.2)$$

Alternativerweise kann die Kovarianz der Objektposition in Längs- und Querrichtung als Maß für die Länge und Breite des Objektes verwendet werden. Allerdings ist dabei die Kenntnis über die Orientierung des Objektes notwendig. Dies ist nicht immer möglich, da in vielen Fällen nur eine Seite des Objektes beobachtet werden kann (siehe Abbildung 4-6). Im Allgemeinen reicht die Gleichung (4.2) für unsere Anwendung aus. Die Höhe eines Objekts wird durch die Höhe des zum Objekt gehörigen höchsten 3D-Punkt beschrieben.

Das Ergebnis der stereobasierten Segmentierung/Objektdetektion wird in Abbildung 4-6 veranschaulicht. Ein segmentiertes Objekt beinhaltet damit folgende Informationen:

$$\text{Obj: } x, y, B, L, H \quad (4.3)$$

Dabei sind  $x$  und  $y$  die Position des Objektes im Bezug auf das Ego-Fahrzeug (siehe Abbildung 4-2),  $B$  die Breite,  $L$  die Länge und  $H$  die Höhe.



Abbildung 4-6: Objekte, die über das 2D-Histogramm segmentiert wurden, werden im Kamerabild zur Kontrolle veranschaulicht.

Es ist an dieser Stelle wichtig zu erwähnen, dass alle Bildpixel, die zu einem segmentierten Objekt gehören, als eine *zusätzliche Matrix* an das Objekt angehängt werden. Damit wird die Bildinformation eines Objektes gespeichert. Dies eröffnet weitere Möglichkeit für das Objekttracking. Darauf wird im Laufe dieses Kapitel noch ausführlich eingegangen (Abschnitt 4.2.4).

## 4.2 Objekttracking

Die Objektdetektion bildet eine wichtige Grundlage für die Verfolgung von bewegten Objekten. Dieses Thema wird seit über dreißig Jahren erforscht. Das Ziel eines Trackingsystems ist es, den *Zustand* eines Objektes auf Basis gewonnener Messdaten möglichst korrekt zu schätzen.

Der Zustand eines Objektes beinhaltet hauptsächlich Informationen über die Objektdynamik wie Position, Geschwindigkeit, Beschleunigung, Gierrate, etc. In verschiedenen Literaturstellen [SHA00] und [BLA99] werden Verfahren zum Objekttracking ausführlich beschrieben. Im Allgemeinen basieren die meisten der Trackingverfahren auf der Annahme einer versteckten Markov-Kette [RAB89] und dem Bayes Theorem [BAY63]. Es wird angenommen, dass der Zustand eines Objektes nicht direkt beobachtbar ist und ausschließlich

durch Messungen geschätzt werden kann. Beobachtbar sind hingegen Messdaten  $\underline{z}_k, \underline{z}_{k+1}, \dots$ , die dem Objekt zugeordnet werden können. In Abbildung 4-7 stellt  $\underline{x}_k$  den Zustand eines Objektes zum Zeitpunkt  $k$  als Vektor dar. Die bedingte Wahrscheinlichkeit  $p(\underline{z}_k | \underline{x}_k)$  drückt aus, dass die Messung nur vom Objektzustand abhängig ist, und nicht von den vorhergehenden Messungen. Die dynamische Veränderung des Objektzustandes,  $\underline{x}_k \rightarrow \underline{x}_{k+1}$ , z. B. das Objekt bewegt sich von A nach B, wird durch die bedingte Wahrscheinlichkeit  $p(\underline{x}_{k+1} | \underline{x}_k)$  modelliert. Dabei wird eine versteckte Markov-Kette (engl. Hidden Markov Model, HMM) angenommen, bei der der Zustand  $\underline{x}_k$  lediglich vom unmittelbar vorigen Zustand  $\underline{x}_{k-1}$  abhängt. Es handelt sich hier um ein zeitdiskretes System, wobei die Messrate das Zeitintervall bestimmt.

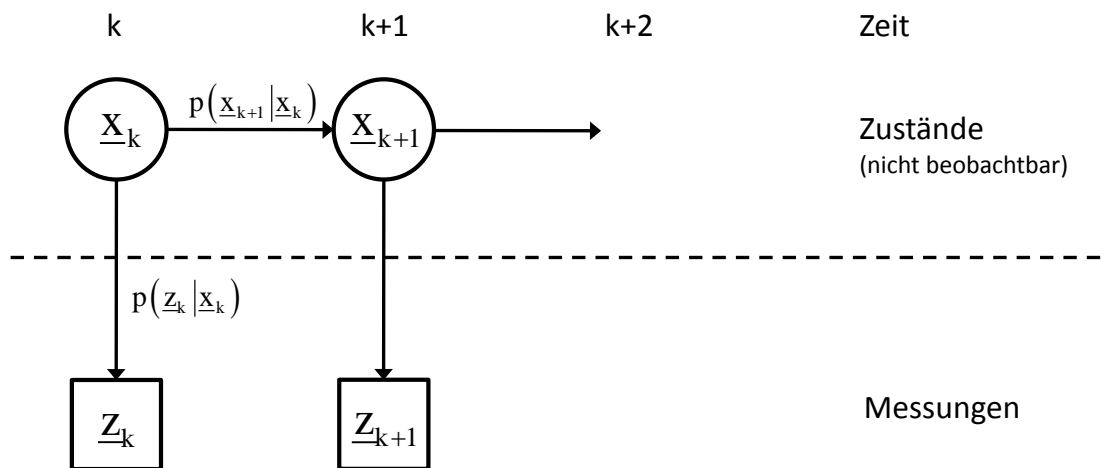


Abbildung 4-7: Hidden-Markov-Modell eines Objektes

Da die Zustände wie schon angenommen nicht immer direkt beobachtbar sind, werden sie durch folgende bedingte Wahrscheinlichkeit geschätzt:  $p(\underline{x}_k | \underline{z}_k)$ . Diese Wahrscheinlichkeit wird nach Bayes-Theorem wie folgt berechnet:

$$p(\underline{x}_k | \underline{z}_k) = \frac{p(\underline{z}_k | \underline{x}_k) \cdot p(\underline{x}_k)}{p(\underline{z}_k)} \quad (4.4)$$

Die Wahrscheinlichkeit  $p(\underline{z}_k)$  erhält man nach dem Gesetz der totalen Wahrscheinlichkeit aus [THR05]:

$$p(\underline{z}_k) = \sum_{\underline{x}_k} p(\underline{z}_k | \underline{x}_k) \cdot p(\underline{x}_k) \quad (4.5)$$

Liegt eine Wahrscheinlichkeitsverteilung für  $\underline{x}_k$  vor, sprich ein kontinuierlicher Systemzustand, kann (4.5) wie folgt ausgedrückt werden [THR05]:

$$p(\underline{z}_k) = \int p(\underline{z}_k | \underline{x}_k) \cdot p(\underline{x}_k) d\underline{x}_k \quad (4.6)$$

Das Ziel ist es, den Zustand auf Basis mehrerer Messungen  $p(\underline{x}_k | \underline{z}_k, \underline{z}_{k-1}, \dots, \underline{z}_1)$  optimal zu schätzen. Nach (4.4) kann  $p(\underline{x}_k | \underline{z}_k, \underline{z}_{k-1}, \dots, \underline{z}_1)$  wie folgt berechnet werden:

$$p(\underline{x}_k | \underline{z}_{k:1}) = \frac{p(\underline{z}_k | \underline{x}_k, \underline{z}_{k-1:l}) \cdot p(\underline{x}_k | \underline{z}_{k-1:l})}{p(\underline{z}_k | \underline{z}_{k-1:l})} \quad (4.7)$$

Nach der Annahme des HMM bedingt nur  $\underline{x}_k$  die Messung  $\underline{z}_k$ . Deshalb wird  $p(\underline{z}_k | \underline{x}_k, \underline{z}_{k-1:l})$  vereinfacht zu  $p(\underline{z}_k | \underline{x}_k)$ . Die Wahrscheinlichkeiten  $p(\underline{x}_k | \underline{z}_{k-1:l})$  und  $p(\underline{z}_k | \underline{z}_{k-1:l})$  lassen sich mit Hilfe von (4.5) so berechnen:

$$p(\underline{x}_k | \underline{z}_{k-1:l}) = \int p(\underline{x}_k | \underline{x}_{k-1}) \cdot p(\underline{x}_{k-1} | \underline{z}_{k-1:l}) d\underline{x}_{k-1} \quad (4.8)$$

$$p(\underline{z}_k | \underline{z}_{k-1:l}) = \int p(\underline{z}_k | \underline{x}_k) \cdot p(\underline{x}_k | \underline{z}_{k-1:l}) d\underline{x}_k \quad (4.9)$$

Damit lässt sich (4.7) wie folgt umschreiben:

$$p(\underline{x}_k | \underline{z}_{k:1}) = \frac{p(\underline{z}_k | \underline{x}_k) \cdot \int p(\underline{x}_k | \underline{x}_{k-1}) \cdot p(\underline{x}_{k-1} | \underline{z}_{k-1:l}) d\underline{x}_{k-1}}{\int p(\underline{z}_k | \underline{x}_k) \cdot \int p(\underline{x}_k | \underline{x}_{k-1}) \cdot p(\underline{x}_{k-1} | \underline{z}_{k-1:l}) d\underline{x}_{k-1} d\underline{x}_k} \quad (4.10)$$

Da die Wahrscheinlichkeiten  $p(\underline{z}_k | \underline{x}_k)$  und  $p(\underline{x}_k | \underline{x}_{k-1})$ , wie in Abbildung 4-7 veranschaulicht, als bekannt angenommen werden, lässt sich  $p(\underline{x}_k | \underline{z}_{k:1})$  in (4.10) auf Basis der vorigen Schätzung  $p(\underline{x}_{k-1} | \underline{z}_{k-1:l})$  berechnen.  $p(\underline{x}_{k-1} | \underline{z}_{k-1:l})$  kann wiederum auf Basis  $p(\underline{x}_{k-2} | \underline{z}_{k-2:l})$  ermittelt werden. Damit kann  $p(\underline{x}_k | \underline{z}_{k:1})$  rekursiv ausgehend von gewonnenen Messdaten  $\underline{z}_k, \underline{z}_{k-1}, \dots, \underline{z}_1$  bestimmt werden. Die Lösung der Integrale in (4.10) stellt einen messbaren



Aufwand dar. Des Weiteren muss ein Mittelwert für  $p(\underline{x}_{k-1} | \underline{z}_{k-1:l})$  aus allen  $\underline{x}_k$ -Zustände berechnet werden.

$$\underline{\mu}_{x_k} = \int p(\underline{x}_k | \underline{z}_{k:l}) d\underline{x}_k \quad (4.11)$$

Um diesen Aufwand zu reduzieren, werden auf Basis einiger Annahmen verschiedene Filtertypen konstruiert, die die Berechnung deutlich vereinfachen.

Bei einem kontinuierlichen Systemzustand kommen entweder das Kalmanfilter oder das Partikelfilter zum Einsatz [EFF09]. Das Partikelfilter [RIS04] ist dazu konstruiert, das Objekttracking bei beliebigem Rauschprozess zu realisieren. Das Kalmanfilter basiert dagegen darauf, dass das Rauschen normalverteilt ist. Für den angedachten Einsatz zeigt das Partikelfilter dem Kalmanfilter gegenüber keinen eindeutigen Vorteil, benötigt jedoch wesentlich mehr Rechenaufwand. Daher wird in dieser Arbeit das Kalmanfilter eingesetzt. Im nachfolgenden Abschnitt wird das Prinzip des Kalmanfilters beschrieben [THR05].

### 4.2.1 Kalmanfilter

Das Kalmanfilter wurde von Rudolf E. Kalman 1960 vorgestellt. Es ist ein Bayes Minimum Varianz Schätzer [KAL60], der aus (4.10) und (4.11) hergeleitet wurde, mit der Annahme, dass der Zustandsübergang von  $\underline{x}_{k-1}$  zu  $\underline{x}_k$  linear ist. Die Grundlage eines Kalmanfilter stellt die zeitdiskrete lineare stochastische Differenzgleichung dar:

$$\underline{x}_k = \underline{F}_k \cdot \underline{x}_{k-1} + \underline{B}_k \cdot \underline{u}_k + \underline{w}_k \quad (4.12)$$

Dabei ist  $\underline{x}_k$  der Objektzustand zum Zeitpunkt  $k$  und  $\underline{F}_k$  die Zustandsübergangsmatrix von  $k-1$  zu  $k$ . Die Matrix  $\underline{F}_k$  wird auch in vielen Literaturstellen als Systemmatrix bezeichnet.  $\underline{u}_k$  beschreibt den Steuerungsvektor des eigenen Fahrzeugs, z. B. die Fahrzeugssteuerungsparameter des Fahrers.

Der Zustandsvektor eines Objektes besteht in der Regel aus folgenden Komponenten:

$$\underline{x}_k = \left[ x \quad y \quad v_x \quad v_y \right]^T.$$

Die Übersetzungsmatrix  $\underline{\mathbf{B}}_k$  übersetzt die Steuerungsparameter in die Ego-Fahrzeugposition.  $\underline{\mathbf{w}}_k$  repräsentiert das Modellrauschen, für den Fall, dass eine Abweichung zwischen angenommener Systemübergangsmatrix und tatsächlicher Bewegung des Objektes auftritt. Die zweite wichtige Annahme ist, dass das Modellrauschen normalverteilt ist:

$$\underline{\mathbf{w}}_k \sim \mathbf{N}(\mathbf{0}, \underline{\mathbf{Q}}_k); \quad p(\underline{\mathbf{x}}_k | \underline{\mathbf{x}}_{k-1}, \underline{\mathbf{u}}_k) \sim \mathbf{N}(\underline{\mathbf{F}}_k \cdot \underline{\mathbf{x}}_{k-1} + \underline{\mathbf{B}}_k \cdot \underline{\mathbf{u}}_k, \underline{\mathbf{Q}}_k)$$

Des Weiteren gilt zwischen Messung und Schätzung folgender Zusammenhang:

$$\underline{\mathbf{z}}_k = \underline{\mathbf{H}}_k \cdot \underline{\mathbf{x}}_k + \underline{\mathbf{v}}_k \tag{4.13}$$

Dabei ist  $\underline{\mathbf{z}}_k$  der durch Messung gewonnene Messwert zum Zeitpunkt  $k$ .  $\underline{\mathbf{H}}_k$  ist die Messmatrix, die den Systemzustand  $\underline{\mathbf{x}}_k$  auf den Messwert  $\underline{\mathbf{z}}_k$  abbildet.  $\underline{\mathbf{v}}_k$  stellt das Messrauschen dar. Sowohl das Messrauschen  $\underline{\mathbf{v}}_k$  als auch das Modellrauschen  $\underline{\mathbf{w}}_k$  werden als normalverteilt angenommen und sind statistisch unabhängig voneinander.

$$\underline{\mathbf{v}}_k \sim \mathbf{N}(\mathbf{0}, \underline{\mathbf{R}}_k); \quad p(\underline{\mathbf{z}}_k | \underline{\mathbf{x}}_k) \sim \mathbf{N}(\underline{\mathbf{H}}_k \cdot \underline{\mathbf{x}}_k, \underline{\mathbf{R}}_k)$$

Die Gleichungen (4.12) und (4.13) sowie die Annahmen bezüglich  $\underline{\mathbf{w}}_k$  und  $\underline{\mathbf{v}}_k$  stellen die Grundlage eines Filtersystems dar. Die Kernidee des Kalmanfilter-Algorithmus besteht darin, die Fehler (oder Abweichungen) zwischen Prädiktion und Messung zu nutzen, um den Objektszustand zu aktualisieren.

Unter Prädiktion wird verstanden, dass auf Basis des letzten Objektzustandes  $\underline{\mathbf{x}}_{k-1}$  und angenommener Zustandsübergangsmatrix  $\underline{\mathbf{F}}_k$  der aktuelle Objektzustand  $\underline{\mathbf{x}}_k$  geschätzt wird, ohne die Messung einzubeziehen. Der Filteralgorithmus wird in Abbildung 4-8 veranschaulicht. Man sieht, dass der Algorithmus der Struktur eines Regelkreises ähnelt.

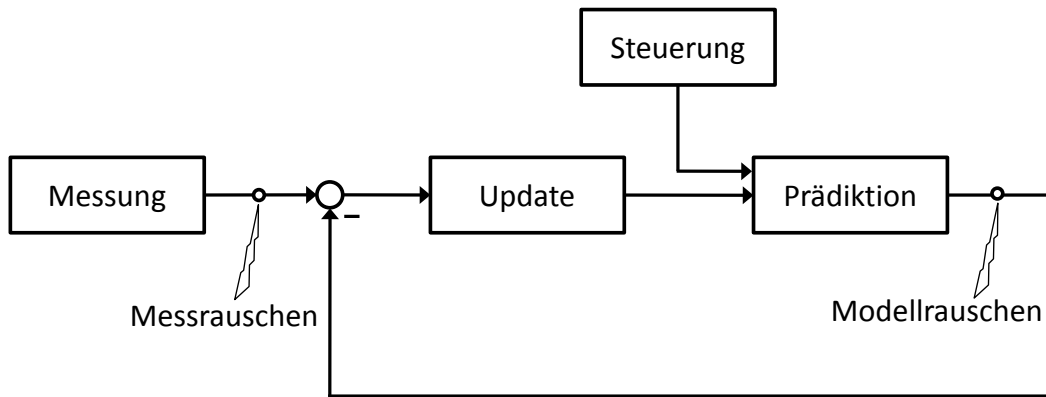


Abbildung 4-8: Kalmanfilter-Algorithmus

Die Filterung erfolgt rekursiv in drei Schritten:

1. *Prädiktion*
2. *Differenz Messwert/Prädiktion*
3. *Update*

Prädiktion:

Bei diesem Schritt werden sowohl der Objektzustand als auch die Kovarianzmatrix prädi-  
ziert:

$$\begin{aligned}\hat{\underline{x}}_k &= \underline{F}_k \cdot \underline{x}_{k-1} + \underline{B}_k \cdot \underline{u}_k \\ \hat{\underline{P}}_k &= \underline{F}_k \cdot \underline{P}_{k-1} \cdot \underline{F}_k^T + \underline{Q}_k\end{aligned}\quad (4.14)$$

$\hat{\underline{x}}_k$  repräsentiert damit den prädierten Zustand und  $\hat{\underline{P}}_k$  die prädierte Kovarianzmatrix. Mit der Addition der Modellrauschenkovarianzmatrix  $\underline{Q}_k$  wird in die Prädiktion die Abweichung des Bewegungsmodells berücksichtigt.

Abweichung Messwert/Prädiktion:

Der prädierte Zustand  $\hat{\underline{x}}_k$  wird mit der Messmatrix  $\underline{H}_k$  multipliziert, um eine so genannte *prädierte Messung* darzustellen. Damit kann die Abweichung  $\Delta \underline{z}_k$  berechnet werden.

$$\Delta \underline{z}_k = \underline{z}_k - \underline{H}_k \cdot \hat{\underline{x}}_k \quad (4.15)$$

Die Abweichung entsteht durch zwei Faktoren, Messrauschen und Modellrauschen. Das Kalmanfilter hat einen Mechanismus, der diese Abweichung verwendet, um den Objektzustand optimal zu schätzen, trotz Rauschens. Dies wird im Update-Schritt beschrieben.

Update:

Der prädizierte Zustand wird anhand der Abweichung mit folgenden Gleichungen korrigiert:

$$\begin{aligned}\underline{x}_k &= \hat{\underline{x}}_k + \underline{K}_k \cdot \Delta \underline{z}_k \\ \underline{P}_k &= (\underline{I} - \underline{K}_k \cdot \underline{H}_k) \cdot \hat{\underline{P}}_k\end{aligned}\tag{4.16}$$

Die Matrix  $\underline{K}_k$ , auch Matrix des Kalman-Verstärkungsfaktors oder Gainfaktors genannt, bestimmt, in welchem Maße die Abweichung verwendet wird, um die Schätzung zu korrigieren.

Der Matrix wird wie folgt ermittelt:

$$\begin{aligned}\underline{K}_k &= \hat{\underline{P}}_k \cdot \underline{H}_k^T \cdot (\underline{R}_k + \underline{H}_k \cdot \hat{\underline{P}}_k \cdot \underline{H}_k^T)^{-1} \\ \underline{S}_k &= \underline{R}_k + \underline{H}_k \cdot \hat{\underline{P}}_k \cdot \underline{H}_k^T\end{aligned}\tag{4.17}$$

Aus (4.17) ist zu erwähnen, dass jedes Element von  $\underline{K}_k$  zwischen 0 und 1 liegt [THR05]. Ist ein Element von  $\underline{K}_k$  gleich 0, wird die Messung für die Schätzung des aktuellen Zustandes an dieser Stelle nicht berücksichtigt. Die Prädiktion wird sozusagen als aktueller Schätzwert übernommen. Ist ein Element von  $\underline{K}_k$  gleich 1, wird die Abweichung vollständig für die Korrektur der Schätzung verwendet. Da das Modellrauschen  $\underline{Q}_k$  in  $\hat{\underline{P}}_k$  berücksichtigt wurde, kann  $\underline{K}_k$  als Mittelung zwischen Messrauschen und Modellrauschen betrachtet werden. Ist das Modellrauschen  $\underline{Q}_k$  im Vergleich zum Messrauschen  $\underline{R}_k$  überwiegend, folgt das Filter dem Messwert schneller [WEL06]. Im Gegenzug ist das Filter gegenüber Ausreißern empfindlicher. Ist das Messrauschen  $\underline{R}_k$  viel größer als das Modellrauschen  $\underline{Q}_k$ , reagiert das Filter weniger empfindlich auf Ausreißer, ist aber gegenüber von unerwarteten Objektbewegungen träger. Im Abschnitt 4.2.3 wird ausführlicher auf die Einflüsse

von  $\underline{\mathbf{R}}_k$  und  $\underline{\mathbf{Q}}_k$  auf die Performance von Kalmanfiltern anhand eines Beispiels eingegangen. Die  $\underline{\mathbf{S}}_k$ -Matrix wird als Residual-Kovarianzmatrix bezeichnet. Auf Basis von (4.16) kann der Objektzustand geschätzt werden. Das gleiche gilt ebenfalls für die Kovarianz.

In unserer konkreten Anwendung sieht der Zustands- und Messvektor folgendermaßen aus:

$$\begin{aligned}\underline{\mathbf{x}}_k &= [\mathbf{x}, \mathbf{y}, \mathbf{v}_x, \mathbf{v}_y, \mathbf{B}, \mathbf{L}, \mathbf{H}]^T \\ \underline{\mathbf{z}}_k &= [\mathbf{x}, \mathbf{y}, \mathbf{B}, \mathbf{L}, \mathbf{H}]^T\end{aligned}\tag{4.18}$$

Größen wie Objektbreite  $\mathbf{B}$ , Objektlänge  $\mathbf{L}$ , Objekthöhe  $\mathbf{H}$  werden nicht geschätzt wie Objektposition und Geschwindigkeit sondern lediglich aus neuen Messungen „übernommen“. Praktische Untersuchungen haben gezeigt, dass das Tracking von Objektbreite, -länge und -höhe keine relevante Verbesserung bringt. Daher wird aufgrund der Rechenzeiteffizienz dieser Schritt erspart.

#### 4.2.2 Erweitertes Kalman Filter

Die Einschränkung des linearen oder einfachen Kalman Filters liegt darin, dass nur lineare Systemübergänge modelliert werden können. In Gleichung (4.14) ist zu erkennen, dass die Objektbewegung mit der Matrix  $\underline{\mathbf{F}}$  in Form eines linearen Zusammenhangs beschrieben wird. Dies beschreibt das reale Verhalten eines Objekts in vielen Fällen ausreichend genau. Dagegen kann das Verhältnis zwischen Steuerungsvektor und resultierender Zustandsänderung nicht linear sein. In Gleichung (4.14) geht es um die  $\underline{\mathbf{B}}$ -Matrix. In einem automobilen Fahrzeug beinhaltet der Steuerungsvektor  $\underline{\mathbf{u}}_k$  Größen wie Radgeschwindigkeit, Lenkwinkel, Fahrzeugsbeschleunigung, etc. Anhand eines Fahrzeugsmodell [MIT04] werden diese Größen in der Fahrzeugsbewegung abgebildet. Diese Abbildung ist nicht linear und kann somit nicht mit einer Matrixmultiplikation mit Matrix  $\underline{\mathbf{B}}$  beschrieben werden. An dieser Stelle wird statt  $\underline{\mathbf{B}} \cdot \underline{\mathbf{u}}$  eine Funktion  $\mathbf{b}(\underline{\mathbf{u}})$  eingeführt.

Um trotzdem das Filterupdate nach (4.16) und (4.17) weiter nutzen zu können, wird die Funktion  $\mathbf{b}(\underline{\mathbf{u}})$  linearisiert, d.h. nach allen Komponenten von  $\underline{\mathbf{u}}$  abgeleitet. Die partiellen

Ableitungen bilden quasi die Matrix  $\underline{\mathbf{B}}$  in dem gewählten Arbeitspunkt und ein quasi linearer Zusammenhang kann wiederhergestellt werden mit  $\mathbf{b}(\underline{\mathbf{u}}) \cdot \underline{\mathbf{u}}$ .

### 4.2.3 Parameterdiskussion

Die Parameter des Kalmanfilters werden abhängig vom Anwendungsfall gewählt. In diesem Abschnitt werden die Parameter anhand eines vereinfachten Beispiels diskutiert. Angenommen es wird ein Objekt mit einem Kalmanfilter verfolgt, das sich nur in eine Richtung bewegt. Damit ergibt sich folgender Zustandsvektor eines Objektes:

$$\underline{\mathbf{x}}_k = \begin{bmatrix} \mathbf{x} \\ \mathbf{v} \end{bmatrix} \quad (4.19)$$

Messbar ist in diesem Beispiel die Entfernung zum Objekt, also der  $x$ -Komponente im Zustandsvektor  $\underline{\mathbf{x}}_k$ .

$$\begin{aligned} \underline{\mathbf{z}}_k &= \underline{\mathbf{H}} \cdot \underline{\mathbf{x}} = \begin{bmatrix} 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{x} \\ \mathbf{v} \end{bmatrix} \\ \underline{\mathbf{z}}_k &= [\mathbf{x}]; \underline{\mathbf{v}}_k \sim \mathbf{N}(0, \sigma_R) \end{aligned} \quad (4.20)$$

Nach [BLA99] ist es sinnvoll, ein Bewegungsmodell ohne Beschleunigung zu verwenden, wenn, wie in unserem Fall, nur die Position mit der Kamera gemessen werden kann. Die Objektbewegung lässt sich mit folgender Übergangsmatrix beschreiben:

$$\underline{\mathbf{F}}_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (4.21)$$

Damit kann die Prädiktion  $\hat{\mathbf{x}}_k$  wie folgt berechnet werden:

$$\hat{\underline{\mathbf{x}}}_k = \begin{bmatrix} \hat{\mathbf{x}}_k \\ \hat{\mathbf{v}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{k-1} + \mathbf{v}_{k-1} \cdot \Delta t \\ \mathbf{v}_{k-1} \end{bmatrix} \quad (4.22)$$

Die Beschleunigung des Objektes wird als Messrauschen angenommen [BLA99]:

$$\underline{w}_k \sim N(0, \underline{Q}_k); \quad \underline{Q}_k = 2\sigma_Q^2 \tau \begin{bmatrix} \frac{\Delta t^3}{3} & \frac{\Delta t^3}{2} \\ \frac{\Delta t^3}{2} & \Delta t \end{bmatrix} \quad (4.23)$$

$\sigma$  ist dabei die Standardabweichung der Beschleunigung,  $\tau$  die empirische Beschleunigungszeit eines Objektes. Im Folgenden werden anhand eines Simulationsbeispiels die Einflüsse von  $\underline{R}_k$  und  $\underline{Q}_k$  bei der Anwendung des Kalmanfilter veranschaulicht.

Es wird jetzt angenommen, dass das Objekt für die ersten 5 Sekunden mit einer Beschleunigung von  $1 \text{ m/s}^2$  anfährt. Danach wird mit einer Beschleunigung von  $-1 \text{ m/s}^2$  für 5 Sekunden abgebremst. Dann steht das Objekt für 5 Sekunden still. Es wird von einer Messfehlervarianz  $\sigma_R^2$  von  $1 \text{ m}^2$  und einer Modellfehlervarianz  $\sigma_Q^2$  von  $1 \text{ m/s}^2$  ausgegangen.

Der Verlauf des geschätzten Objektzustandsvektors wird in Abbildung 4-9 und Abbildung 4-10 grafisch dargestellt. In Abbildung 4-9 stellt die blaue Kurve die tatsächliche Position des Objekts über die Zeit dar, die roten Punkte die Messwerte und die schwarze Kurve die geschätzte Position. In Abbildung 4-10 wird die tatsächliche Geschwindigkeit mit der blauen Kurve visualisiert, die geschätzte Geschwindigkeit mit der grünen Kurve. Es ist erkennbar, dass die geschätzte Position von der tatsächlichen ein wenig abweicht.

Das gleiche gilt für die geschätzte Geschwindigkeit. Darüber hinaus hat die geschätzte Geschwindigkeit eine kleine zeitliche Zögerung gegenüber der tatsächlichen. Dies liegt daran, dass die Geschwindigkeit nicht direkt gemessen wird, sondern aus mehreren Positionsmessungen geschätzt werden muss. Abhängig von den Parametern des Kalman-Filters verändert sich der durchschnittliche Fehler.

Nach der Kalmanfilter-Theorie ist es entscheidend, die Messfehlervarianz  $\sigma_R^2$  und Modellfehlervarianz  $\sigma_Q^2$  möglichst genau abschätzen zu können. Das Messrauschen  $R$  gibt die Genauigkeit der messbaren „Größen“ an, hier in unserem Fall die Position.  $Q$  auf der anderen Seite steht für die Genauigkeit von Größen, die geschätzt werden müssen. In unserem Fall ist es die Geschwindigkeit.

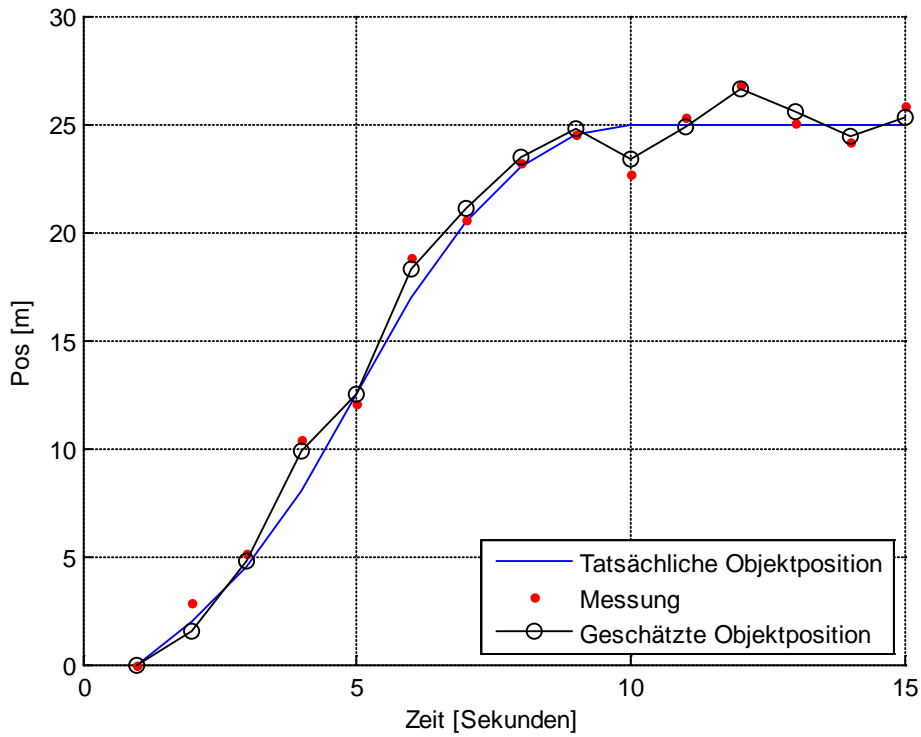


Abbildung 4-9: Gegenüberstellung der tatsächlichen Objektposition und der auf Basis der Messung geschätzter Objektposition

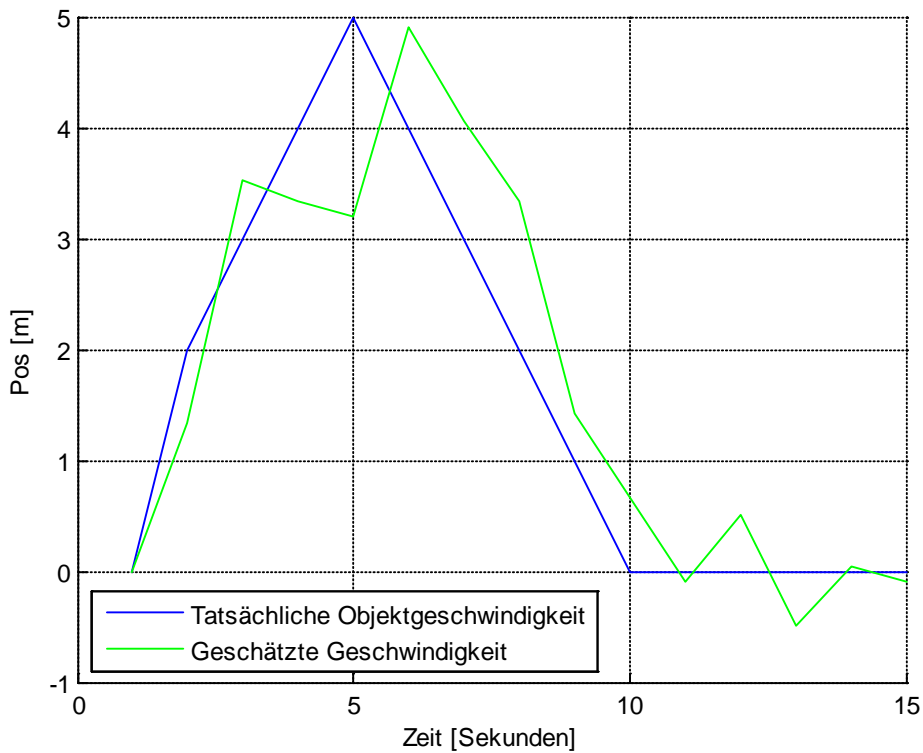


Abbildung 4-10: Gegenüberstellung der tatsächlichen Objektgeschwindigkeit und der geschätzten Geschwindigkeit.



Insgesamt lassen sich mit der Parameteruntersuchung folgende Aussagen treffen:

- In Situationen, in denen die Abweichungen vom Modell häufig auftreten (häufige Beschleunigung oder Abbremsung beim Fahrverhalten), ist es besser, die Q-Matrix möglichst ähnlich wie vom tatsächlichen Rauschen zu dimensionieren, damit das Filter relativ schnell auf die Systemzustandsänderung reagieren kann. Das System ist aber gleichzeitig bei Messausreißern empfindlicher.
- Bei Fahrten mit großem Anteil an konstanter Geschwindigkeit, empfiehlt es sich, die Q-Matrix kleiner als die tatsächliche anzunehmen. Die Geschwindigkeitsschätzung kann mit steigender Anzahl an Messungen stabiler geschätzt werden. Nach [WEL06] sollen die beiden Matrizen Q- und R situationsabhängig angepasst werden.

Das Beispiel in Abbildung 4-9 beinhaltet ein einziges Objekt. In der Praxis spielt beim Auftreten mehrerer Objekte die Assoziation neuer Messungen zu bekannten Objekten (in Englisch: Measurement-To-Track-Assoziation) eine entscheidende Rolle für die Qualität des Objekttracking. Dies wird im nächsten Abschnitt erläutert.

#### 4.2.4 Assoziation neuer Messungen zu bekannten Objekten

Die Measurement-To-Track-Assoziation beinhaltet den Schritt, eine Prädiktion mit einer Messung zu verknüpfen und diese für das Update des Objektzustandes zu verwenden. Dies wird in Abbildung 4-11 veranschaulicht.

Die Qualität des Tracking hängt nicht nur von der Messgenauigkeit, der Modellgenauigkeit sowie der Parametrisierung des Kalmanfilters ab, sondern auch noch von der richtigen Assoziation zwischen Messung und Objektprädiktion. Eine falsche Assoziation führt zu einem falschen Objektupdate und somit zum Objektverlust. In [SHA00], [BLA99] werden mehrere Verfahren zur Measurement-To-Track-Assoziation beschrieben.

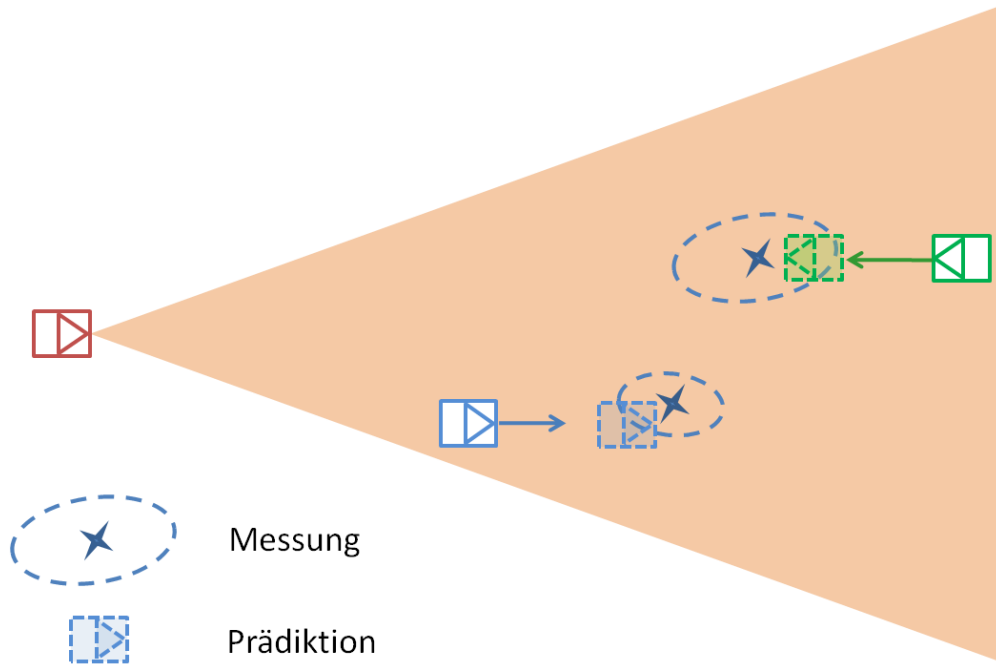


Abbildung 4-11: Measurement-to-track-Association

All diese Verfahren basieren darauf, die Ähnlichkeit zwischen Messung und Prädiktion auszuwerten. Der Mahalanobis-Abstand [MAH36] eignet sich gut für diesen Zweck und wird deshalb verbreitet verwendet. Im Folgenden wird der Mahalanobis-Abstand beschrieben.

#### 4.2.4.1 Mahalanobis-Abstand

Der Mahalanobis-Abstand wird nach Prasanta Chandra Mahalanobis benannt und ist ein Distanzmaß zwischen zwei Vektoren in einem mehrdimensionalen Vektorraum [MAH36]. Anders als der euklidische Abstand wird bei der Berechnung von Mahalanobis-Abstand die Korrelation zwischen den Vektorzuständen mit berücksichtigt. Der Mahalanobis-Abstand wird verwendet, um ein Ähnlichkeitsmaß zwischen zwei Vektoren zu bestimmen und kann größenunabhängig wie folgt berechnet werden:

$$d(\hat{\underline{z}}_j, \underline{z}_k) = \sqrt{(\hat{\underline{z}}_j - \underline{z}_k)^T \cdot \underline{\underline{S}}^{-1} \cdot (\hat{\underline{z}}_j - \underline{z}_k)} \quad (4.24)$$

Dabei ist  $\underline{\underline{S}}$  die Kovarianzmatrix,  $\hat{\underline{z}}_j$  ist der prädizierte Messvektor des Track-Objekts  $j$ ,  $\underline{z}_k$  die gewonnene  $k$ -te Messung. Theoretisch wird für alle Messungen und Prädiktionen paarweise der Mahalanobis-Abstand berechnet.

In der Praxis wird, um den Rechenaufwand effizient zu halten, ein so genanntes „Gate“ eingeführt. Wenn der euklidische Abstand zwischen Prädiktion und Messung größer ist als Gate, wird dieses Paar bei der Berechnung des Mahalanobis-Abstandes nicht mehr berücksichtigt. Nach [BLA99] gibt es verschiedene Arten von Gates (Kreis, Rechteck, Ellipsoid). In unserer Anwendung wird ein Kreis-Gate verwendet. Nachdem der Mahalanobis-Abstand für alle Paare berechnet wurde, können die Paare mit einer der nach folgenden Methoden einander zugeordnet werden [BLA99]:

#### 4.2.4.2 Global-Nearest-Neighbor-Methode (GNN)

Bei dieser Methode wird einer Prädiktion eine Messung zugeordnet, die innerhalb des Gates liegt und den minimalen Mahalanobis-Abstand zu der Prädiktion besitzt [BLA99].

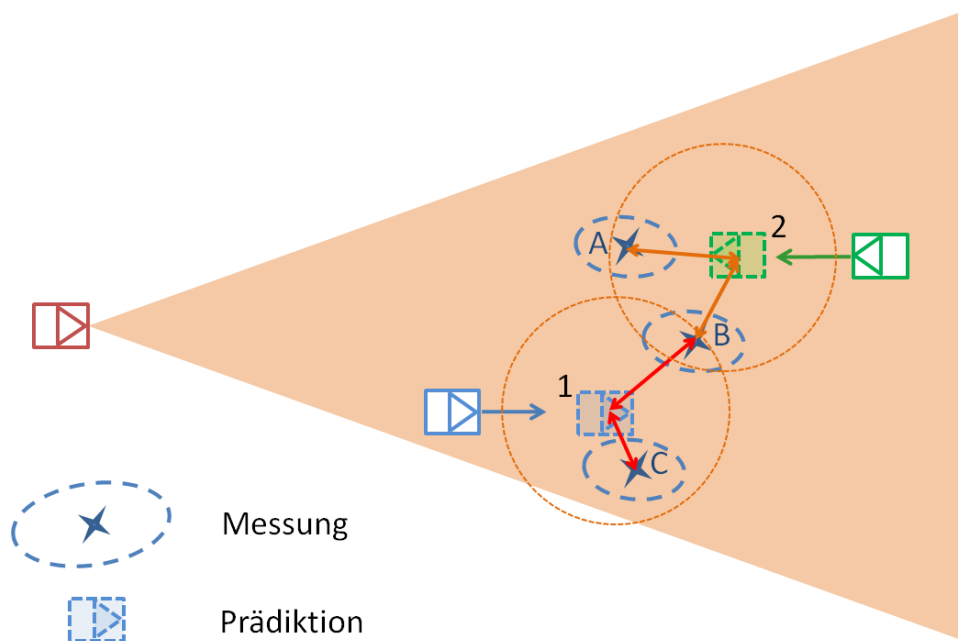


Abbildung 4-12: Assoziation nach der Methode Global-Nearest-Neighbor

In Abbildung 4-12 wird die Idee des Verfahrens veranschaulicht. Innerhalb des Gates um Objekt 1 sind zwei Messungen B und C. Messung C wird mit dem Objekt assoziiert, da der Abstand zum Objekt am geringsten ist.

Der Nachteil dieses Verfahrens wird sichtbar, wenn Messungen und Prädiktion dicht nebeneinander auftreten, was zu keinem eindeutigen Minimum führt. Damit steigt die Chance einer falschen Assoziation.

#### 4.2.4.3 All-Neighbors-Methode

Bei dieser Methode werden einer Prädiktion mehrere Messungen zugeordnet [BLA99]. Bei einem Mehrensensoren-System kommt diese Methode häufig zur Anwendung. Dies kann den Nachteil der GNN-Methode ausgleichen, indem keine harte Entscheidung für die Assoziation getroffen wird. Die Mehrfach-Assoziation kann beim Update-Schritt, im Vergleich zu (4.15) wie folgt berechnet werden [STU04]:

$$\underline{\mathbf{x}}_k = \hat{\underline{\mathbf{x}}}_k + \underline{\mathbf{K}}_k \cdot \Delta \tilde{\underline{\mathbf{z}}}_k \quad (4.25)$$

Wobei  $\Delta \tilde{\underline{\mathbf{z}}}_k$  die gesamte Innovation für alle assoziierte Messungen darstellt.  $\Delta \tilde{\underline{\mathbf{z}}}_k$  wird wie folgt bestimmt:

$$\Delta \tilde{\underline{\mathbf{z}}}_k = \sum_i^N \mathbf{p}_i \cdot \underline{\mathbf{z}}_i - \underline{\mathbf{H}} \cdot \hat{\underline{\mathbf{x}}}_k \quad (4.26)$$

Dabei ist  $N$  die Anzahl der assoziierten Messungen,  $\underline{\mathbf{z}}_i$  die jeweilige Messung und  $\mathbf{p}_i$  die Wahrscheinlichkeit, dass die Messung  $i$  zur Prädiktion gehört. Diese Wahrscheinlichkeit  $\mathbf{p}_i$  wird in [BLA99] bestimmt.

In unserer Anwendung wird diese Methode nicht eingesetzt, da nur die Stereo-Kamera zur Fußgängerdetektion verwendet wird. Damit ergibt sich für ein Objekt, z.B. einen Fußgänger, in der Regel nur eine Messung. Die Multiassoziation ergibt damit keine Vorteile.

Es wird daher in unserer Anwendung die GNN-Assoziation verwendet. Um das Problem der nicht eindeutigen Assoziation zu lösen (siehe Abbildung 4-12), wird im Folgenden ein neues leistungsfähiges Verfahren vorgestellt.

#### 4.2.4.4 Verbesserung der Objektassoziation durch zusätzliche Bildmerkmale

Je mehr Informationen aus einer Messung extrahiert werden, desto einfacher können die Objekte voneinander unterschieden werden. In der bekannten Literatur [BLA99] werden für die Assoziation lediglich Informationen wie Objekt-Position, Objekt-Geschwindigkeit verwendet. In unserer Anwendung bietet das Kamerabild großes Potenzial, mehr Informatio-

nen über das Objekt zu gewinnen. Dies wird in veröffentlichten Arbeiten zum Objekttracking mit Kamera nicht erwähnt [BRO05], [NED07].

Deshalb wird in dieser Arbeit Verfahren zur Merkmalsextraktion eingesetzt [NIX08], [SURF06]. Es werden für jedes nach Abschnitt 4.1 segmentierte Objekt bildbasierte Merkmale extrahiert. Dadurch wird der Merkmalvektor einer Messung angesichts der Beschreibung aus (4.18) mit verschiedenen Bildmerkmalen erweitert. Diese spielen jedoch für die Zustandsübergang-Gleichungen keine Rolle (4.14) - (4.17).

Es ist wichtig zu erwähnen, dass nur Pixel, die zu einem Objekt gehören, zur Merkmalsextraktion verwendet werden. Diese Pixel werden bei der Segmentierung mitgespeichert. Im Folgenden werden die untersuchten Merkmale beschrieben.

#### Durchschnittlicher Grauwert

Diese Größe wird häufig in der Bildverarbeitung verwendet und repräsentiert den durchschnittlichen Grauwert aller Pixel, die zu einem Objekt gehören, und wird wie folgt berechnet:

$$\tilde{I} = \frac{1}{N} \sum_j^N I(j) \quad (4.27)$$

Wobei  $I(j)$  der Grauwert eines zum Objekt gehörigen Pixels ist und  $N$  die Anzahl der Pixel, die das Objekt besitzt. Der durchschnittliche Grauwert  $\tilde{I}$  wird in Kombination mit der Varianz des Grauwertes eingesetzt.

#### Varianz des Grauwertes

Diese Größe  $\delta I$  wird ausgehend vom durchschnittlichen Wert wie folgt berechnet:

$$\delta I = \frac{1}{N} \sum_j^N [I(j) - \tilde{I}]^2 \quad (4.28)$$

Mit den beiden Größen „durchschnittlicher Grauwert“ und „Varianz des Grauwertes“ werden zusätzliche Informationen zum Objekt gewonnen:

In Abbildung 4-13 werden diese Merkmale veranschaulicht.

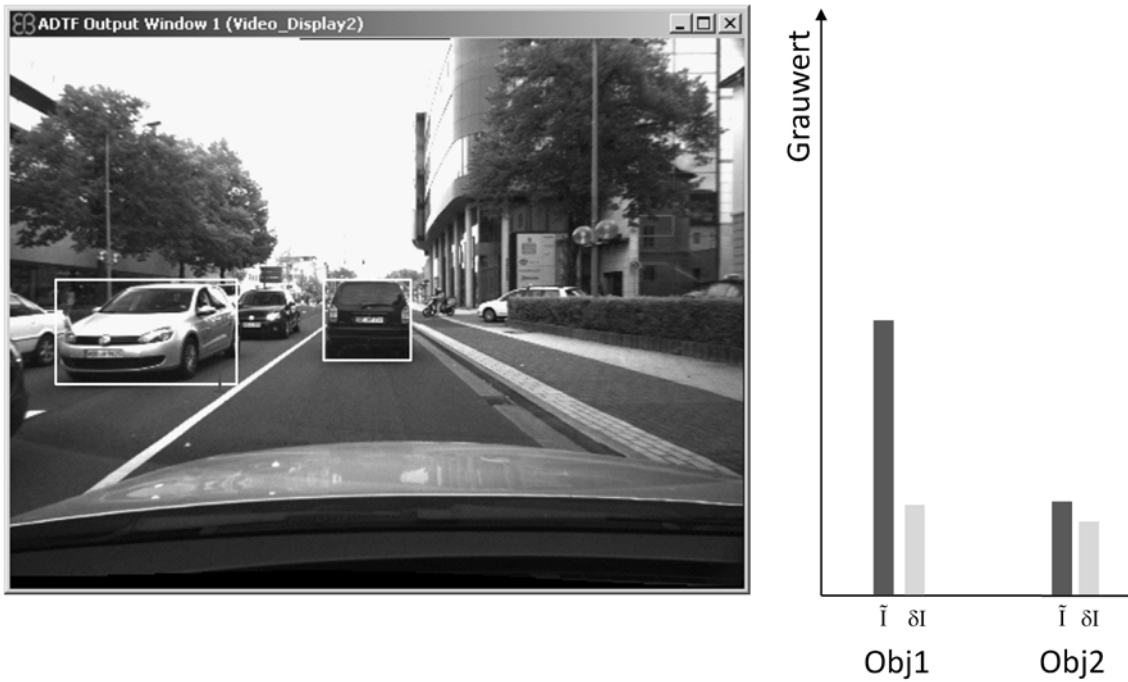


Abbildung 4-13: Durchschnittlicher Grauwert und Varianz des Grauwertes der Objekte

### Grauwerthistogramm

Das Histogramm ist bei der Beurteilung der Eigenschaften eines Bildes oder Objektes ein nützliches Hilfsmittel. Es zeigt die Häufigkeitsverteilung der einzelnen Grauwerte.

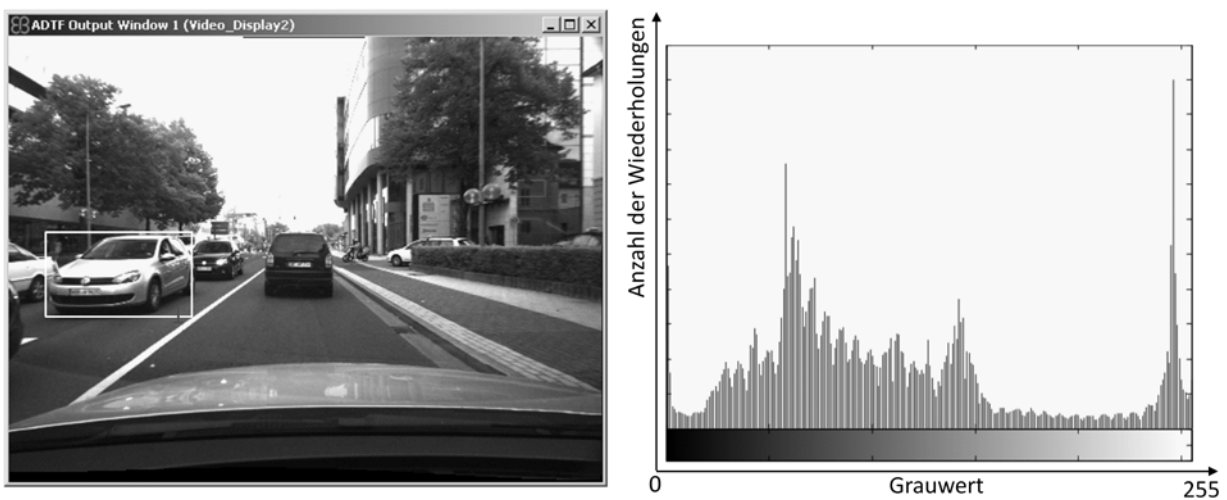


Abbildung 4-14: Grauwerthistogramm über die Pixel eines detektierten Objektes

Für ein Grauwertbild  $I$  mit möglichen Intensitätswerten im Bereich  $I(u, v) \in [0, K - 1]$  enthält das zugehörige Histogramm  $H$  genau  $K$ -Einträge, wobei für ein typisches 8-Bit-Grauwertbild  $K = 2^8 = 256$  ist. Jeder Histogramm-eintrag ist definiert als [BUR05]:

$$\begin{aligned} h(i) &= \text{die Anzahl der Pixel von } I \text{ mit dem Intensitätswert } i \\ i &\in [0, K] \end{aligned} \quad (4.29)$$

In Abbildung 4-14 wird ein Beispiel eines Histogramm  $H(i)$  dargestellt. Zur Berechnung des Histogramms werden ausschließlich die zum Objekt gehörigen Pixel verwendet.

### SURF-Deskriptoren

Speeded Up Robust Features (SURF) ist ein robuster Algorithmus zur Erkennung von Bildmerkmalen und ist skalierungs- und rotierungsunabhängig [SURF06]. Der Algorithmus wurde von Herbert Bay et al. [SURF06] vorgestellt und beinhaltet zwei Hauptschritte:

1. Detektion von Point-of-Interest (Punkte, die Merkmale enthalten), kurz als POIs bezeichnet.
2. Extraktion der Merkmale

Beim Schritt 1 wird als erstes das Integralbild berechnet [VIO01]. Auf dieser Basis werden die Hesse-Matrix bzw. ihre Determinante für jeden Pixel im Integralbild berechnet. Die Hesse-Matrix wird als eine Funktion  $\underline{H}$  der Pixelposition  $\underline{x}=(x,y)$  und der Skalierung  $\sigma$  dargestellt:

$$\underline{H}(\underline{x}, \sigma) = \begin{bmatrix} L_{xx}(\underline{x}, \sigma) & L_{xy}(\underline{x}, \sigma) \\ L_{xy}(\underline{x}, \sigma) & L_{yy}(\underline{x}, \sigma) \end{bmatrix} \quad (4.30)$$

Wobei  $L_{xx}(\underline{x}, \sigma)$  die Faltung der zweiten partiellen Ableitung nach  $x$  der Gauß-Funktion  $\partial^2 g(\sigma)/\partial x^2$  mit dem Eingangsbild am Pixel  $\underline{x}$  ist. Bei  $L_{yy}(\underline{x}, \sigma)$  geht es um die zweite Ableitung nach  $y$  und bei  $L_{xy}(\underline{x}, \sigma)$  wird es nach  $x$  und  $y$  abgeleitet. In Abbildung 4-15 wird  $L_{yy}$  und  $L_{xy}$  bei einer Skalierung von  $\sigma=1.2$  veranschaulicht. Um den Rechenaufwand gering zu halten, wird an dieser Stelle stattdessen die Approximation  $D_{yy}$  und  $D_{xy}$  verwendet wie auf der rechten Seite der Abbildung 4-15. Die damit verbundene Beeinflussung in der Genauigkeit wurde als nicht relevant gezeigt.

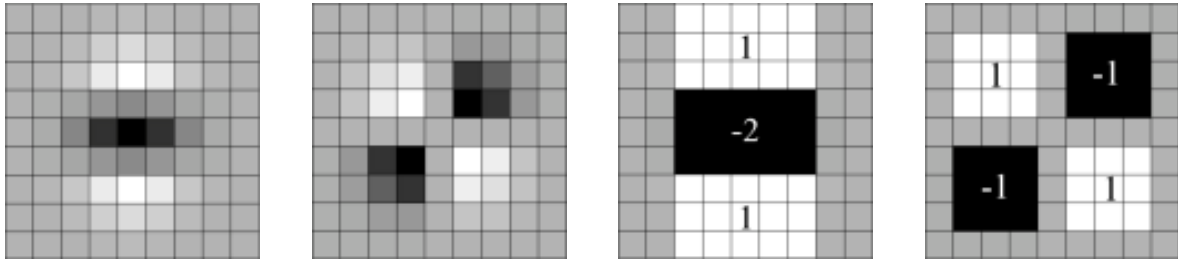


Abbildung 4-15: von links nach rechts, die zweite partielle Ableitung der Gaußfunktion nach y, nach xy sowie die entsprechenden SURF-Approximationen [SURF].

Es wird die Determinante der Hesse-Matrix berechnet:

$$\det(\underline{\mathbf{x}}) = D_{xx} D_{yy} - (D_{xy})^2 \quad (4.31)$$

Der Gewichtungsfaktor  $w$  dient als Ausgleich für den Unterschied zwischen Approximationen  $D_{xx}, D_{xy}$  und  $D_{yy}$  und ursprünglichen  $L_{xx}, L_{xy}$  und  $L_{yy}$ .  $w$  wird wie folgt berechnet:

$$w = \frac{|L_{xy}(\sigma)| \cdot |D_{yy}(b)|}{|L_{yy}(\sigma)| \cdot |D_{xy}(b)|} \quad (4.32)$$

Wobei  $b$  die Breite der Filtermaske ist. Das Gewicht  $w$  variiert geringfügig abhängig von der Filterbreite  $b$  und der Skalierung  $\sigma$ , wird aber für den Einsatz in SURF als 0.9 angenommen. Es hat sich gezeigt, dass dies keinen maßgeblichen Einfluss auf das Ergebnis hat [SURF06]. Der Wert der Determinante wird als *Blob* an Position  $\underline{\mathbf{x}} = (x, y, \sigma)$  bezeichnet.

Um die Skalierungsunabhängigkeit zu gewährleisten, müssen die POIs (Point-of-Interest) bei unterschiedlichen Skalierungen gesucht werden. Statt die Bildauflösung zu verändern wie bei SIFT [LOW99], wird bei SURF die Filtermaskengröße variiert. Die Filtergröße in Abbildung 4-15 ist 9x9 und entspricht einer Skalierung mit  $\sigma = 1.2$ . Bei einer Vergrößerung der Filtergröße vergrößert sich auch die Skalierung. Die Filterbreite  $b$  und die Skalierung  $\sigma$  sind wie folgt voneinander abhängig:

$$\sigma = b \cdot \frac{\sigma_R}{b_R} \quad (4.33)$$

Wobei  $b_R$  und  $\sigma_R$  die Bezugsfiltergröße und -skalierung ist ( $b_R = 9, \sigma_R = 1.2$ ).

Abbildung 4-16 veranschaulicht unterschiedliche Filtergrößen.



Ein Pixel hat daher unterschiedliche Blob-Werte bei unterschiedlichen Skalierungen. Die Suche nach dem Point-of-Interest definiert sich in diesem Falle als die Suche nach lokalen Extrema im  $x - y - \sigma$ -Raum. Die  $3 \times 3 \times 3$ -Nachbarschaft jedes Blobs wird untersucht. Ist ein Blob-Wert betragsmäßig größer als alle seine 26 Nachbar-Blobs bzw. größer als ein Schwellwert, wird diese Skalenraum-Position  $\underline{x} = (x, y, \sigma)$  als ein Interessenpunkt gekennzeichnet.

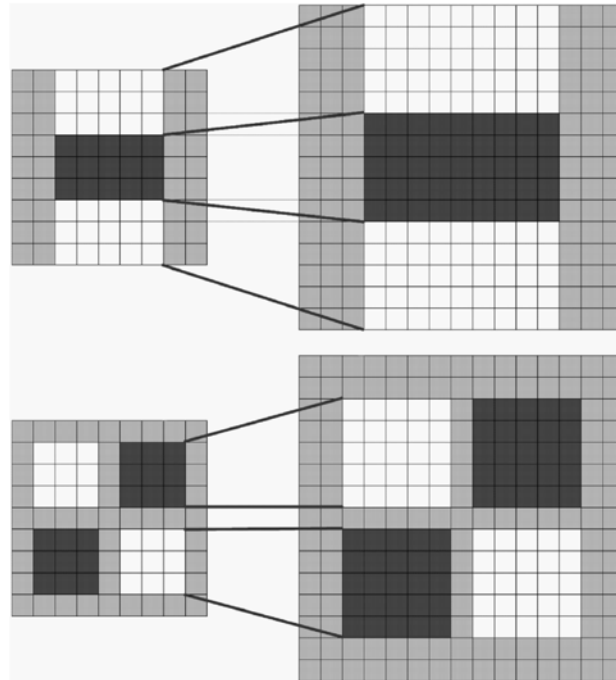


Abbildung 4-16: Unterschiedliche Skalierungen durch unterschiedliche Filtergröße [SURF06]

Nachdem die Points-of-Interest detektiert wurden, ist es wichtig, die Punkte pixelgenau zu lokalisieren. Dies ist bei größeren Skalierungen nicht selbstverständlich, da gröber abgetastet wurde als bei kleineren Skalierungen. In diesem Fall werden die gefundenen Punkte auf Subpixelgenauigkeit interpoliert.

In Abbildung 4-17 werden die detektierten POIs veranschaulicht. Die eingezeichneten Boxen stellen die segmentierten Objekte dar. Es ist zu erkennen, dass für jedes Objekt mehrere POIs detektiert wurden. Dies ist für das Tracking eine wichtige Zufuhr an Information.



Abbildung 4-17: SURF-Points of Interest

Als nächster Schritt werden die Merkmale der detektierten POIs extrahiert, damit die Punkte miteinander von Bild zu Bild verglichen werden können. Zunächst wird für jeden Interessenspunkt die Orientierung ermittelt. Dafür wird die Antwort der Haar-Wavelet-Transformation in x- und y-Richtung für alle Nachbarn des POI innerhalb eines Radius von  $6 \times \sigma$  berechnet.

Dabei ist  $\sigma$  die Skalierung, an der der Interessenspunkt nach (4.31) detektiert wurde. Die Haar-Wavelet-Filter sind in Abbildung 4-18 visuell dargestellt. Die Größe der Wavelet-Masken ist skalierungsabhängig und beträgt  $4 \times \sigma$ . Nachdem alle Wavelet-Antworten berechnet wurden, werden sie mit einer Gauß-Verteilung gewichtet, die am Interessenspunkt zentriert ist. Die Antwort wird dann in einem 2D-Histogramm aufgetragen (siehe Abbildung 4-19). Anhand eines Verschiebungsfensters von  $60^\circ$  wird die gesamte Richtung für die aktuelle Fensterposition berechnet. Das Fenster mit dem stärksten Richtungsvektor wird als Orientierung des POI angenommen.

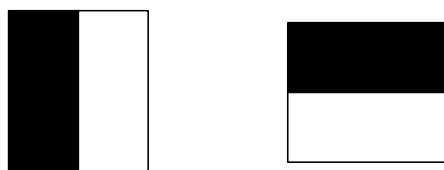


Abbildung 4-18: Grundformen der Haar-Wavelet-Filter [SURF06]

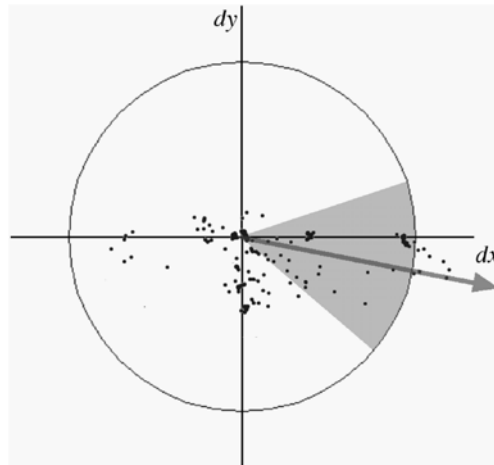


Abbildung 4-19: Schätzung der Orientierung eines SURF-POI [SURF06]

Des Weiteren wird ein Quadratfenster um den Punkt herum mit der Größe  $20 \times \sigma$  auf Basis der vorher bestimmten Orientierung konstruiert. In diesem Fenster werden die Haar-Wavelet-Antworten berechnet. Es werden am Ende alle Wavelet-Antworten in x- und y-Richtung ( $\sum dx$  und  $\sum dy$ ) sowie die absolute Werte der Antworten ( $\sum |dx|$  und  $\sum |dy|$ ) aufsummiert.

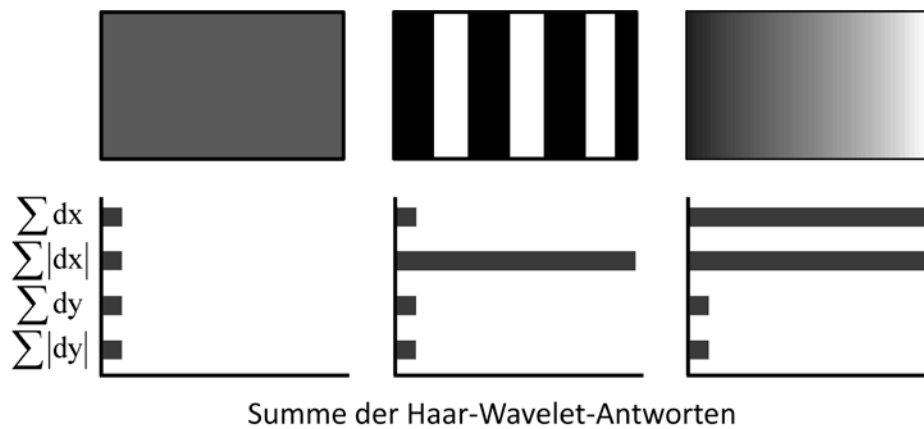


Abbildung 4-20: SURF-Deskriptor eines POI [SURF06]

In Abbildung 4-20 wird ein typischer Deskriptorvektor veranschaulicht. Ein solcher Vektor enthält vier Komponente:  $\sum dx$ ,  $\sum dy$ ,  $\sum |dx|$  und  $\sum |dy|$ .

Unterschiedliche Regionen ergeben unterschiedliche Werte. Alle Merkmalsvektoren, die zu einem Objekt gehören, werden in ein Vektor-Array eingefügt. Dieses Array wird dann an den Messvektor eines Objektes angehängt.

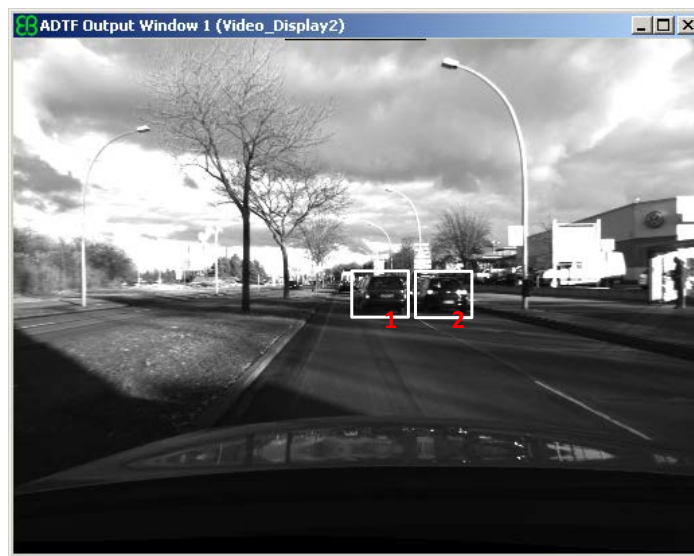
#### 4. Detektion und Tracking von bewegten Objekten

Damit wird der Mess- und Zustandsvektor aus (4.18) mit den zusätzlichen Bildmerkmalen wie folgt erweitert:

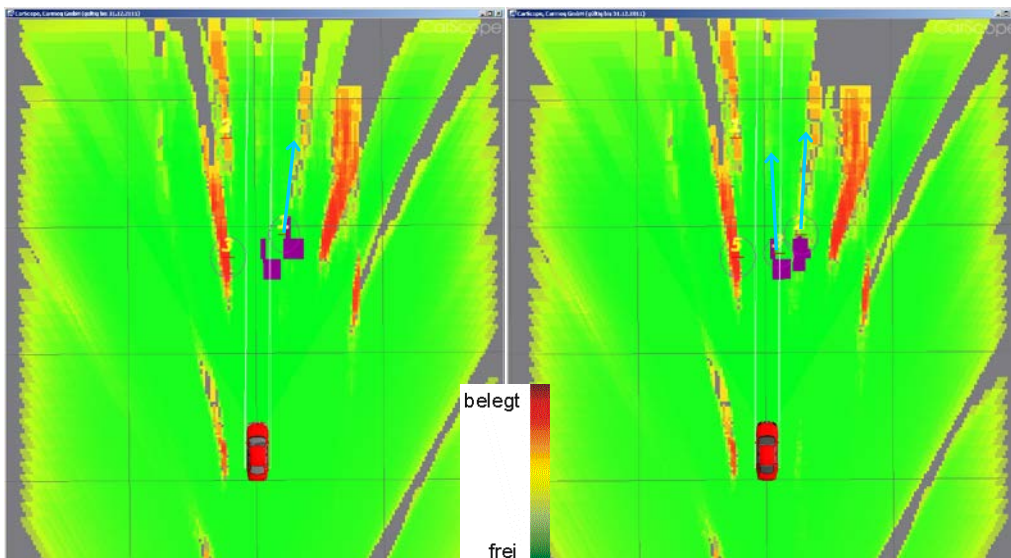
$$\underline{z}_k = [x, y, B, L, H, \tilde{I}, \delta I, \underline{p}, \underline{s}]^T \quad (4.34)$$

Wobei  $\underline{p}$  das Histogramm und  $\underline{s}$  das Array der SURF-Punkte darstellen.

Wie schon vorher erwähnt, hat die Verwendung zusätzlicher Bildattribute das Ziel, prädi-zierte Objekte mit gemessenen Objekten besser zu assoziieren.



Teil a)



Teil b)

Teil c)

Abbildung 4-21: Verbesserung der Assoziation mit SURF-Merkmalen. Bild a) Das Kamerabild mit den bewegten Objekten Bild b): Assoziation ohne SURF-Merkmale. Bild c): Assoziation mit SURF-Merkmalen.

Der Mahalanobis-Abstand kann damit mit den Erweiterungen berechnet werden. Das Gate wird demzufolge auch angepasst. Bei der Durchführung dieser Methode auf vorhandenen Testdaten hat sich gezeigt, dass die SURF-Deskriptoren  $\underline{s}$  maßgeblich zur Verbesserung der Assoziation beitragen. Auf Merkmale wie durchschnittlicher Grauwert, Varianz oder Grauerthistogramm kann verzichtet werden. Die Verbesserung in einer Measurement-To-Track-Assoziation wird in Abbildung 4-21 veranschaulicht.

Es ist in der Abbildung 4-21 zu erkennen, dass die zwei gekennzeichneten vorausfahrenden PKWs im Ergebnis bezüglich Form, Grauwert, etc. recht ähnlich sind. Dadurch kommt es zur falschen Assoziation wie in Abbildung 4-21, Bild b). Das Objekt 1 im Kamerabild hat einen Spurwechsel von der rechten Spur auf die linke Spur durchgeführt.

Die Assoziation im mittleren Bild erfolgt ohne SURF-Merkmale zwischen Objekt 1 und Messung von Objekt 2. Dadurch wird die Position des Objekts 1 falsch aktualisiert. Dagegen erfolgt die Assoziation im rechten Bild auf der Basis von SURF-Merkmalen. Dadurch können zwei bewegte Objekte korrekt verfolgt werden.

### 4.3 Zusammenfassung

In diesem Kapitel wurde eine auf Daten einer Stereokamera basierende Objektdetektion und das Objekttracking beschrieben. Die Objektdetektion erfolgt mit Hilfe eines 2D-Histogramms bzw. eines Segmentierungsalgorithmus. Im Anschluss wird ein erweitertes Kalman-Filter zum Objekttracking verwendet. Eine Analyse über die Auswahl der Kalmanfilter-Parameter wurde präsentiert. Auf dieser Basis kann eine optimale Festlegung der Parameter erreicht werden.

Des Weiteren wird die Problematik der Assoziation bei hoher Dichte von Objekten geschildert. Je dichter die Objekte nebeneinander auftreten, desto größer ist die Wahrscheinlichkeit einer falschen Assoziation.

Um die Anzahl der falschen Assoziationen in solchen Situationen gering zu halten, wird der Messvektor mit den Bildmerkmalen erweitert. Dadurch kann jedes Objekt von anderen

eindeutiger unterschieden werden bzw. leichter im nächsten Frame gefunden werden. Dieses neue Verfahren trägt dazu bei, das Objekttracking erheblich zu verbessern.

Die auf Testdaten erzielten Ergebnisse des Objekttracking werden in Kapitel 6 dargestellt. Der Fokus des Objekttracking besteht darin, wie schon in Kapitel 2 erwähnt wurde, bewegte Objekte zu verfolgen. Die Repräsentation stationärer Hindernisse wird mit dem kartenbasierten Ansatz im nächsten Kapitel umgesetzt.

# Kapitel 5

## Belegungskarte zur Darstellung stationärer Hindernisse

Eine alternative Möglichkeit zur objektbasierten Umfeldmodellierung stellt die Belegungskarte dar. Zum ersten Mal wurde dieses Verfahren von Alberto Elfes 1988 [ELF89] vorgestellt. Seitdem findet diese Methode breite Anwendung, vor allem in der Robotik und in den letzten Jahren auch in der Automobilindustrie. Die Stärke dieser Methode liegt darin, dass unstrukturierte Umgebungen sehr gut und genau beschrieben werden können, ohne dass ein Objektmodell dafür benötigt wird.

### 5.1 Grundidee der Belegungskarte

Die Grundidee der Belegungskarte ist es, das Umfeld in diskreter Form darzustellen. Die Umfeldinformation wird in so genannten *Zellen* gespeichert. Jede Zelle repräsentiert einen bestimmten Bereich des Umfeldes. Die Zellen bilden zusammen entweder ein zweidimensionales oder ein drei dimensionales Gitter um das Ego-Fahrzeug [EFF09]. In der Praxis wird meist ein 2D-Gitter verwendet, da es deutlich recheneffizienter ist und in der Regel bereits die relevanten Umfeldinformationen im Vergleich zum 3D-Gitter enthält und dessen Mehrwert somit gering ist. In Abbildung 5-1 wird an einem Beispiel eine Belegungskarte in Form eines 2D-Gitters veranschaulicht. Informationen, die eine Zelle beinhaltet, sind in der Regel die **Befahrbarkeit** des, der Zelle zugeordneten, Umfeldbereichs. In Abbildung 5-1 befinden sich in rot markierten Zellen Hindernisse wie Häuser oder Bäume. Dagegen ist die Straße mit grünen Zellen gekennzeichnet. Dabei spielt es keine Rolle, um welche Art von

Hindernissen es sich handelt. Jede Zelle wird voneinander unabhängig modelliert und gibt dementsprechend die Befahrbarkeit ihres Bereichs an.

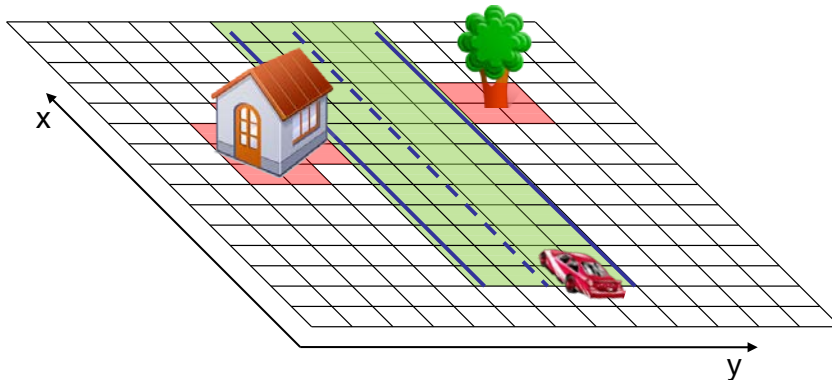


Abbildung 5-1: Beispiel einer Belegungskarte

In verschiedenen Literaturstellen sind mehrere Varianten [HEU10] der Belegungskarte beschrieben. Grundsätzlich unterscheiden sich die Varianten im Bezugssystem zum Ego-Fahrzeug. **Ortsfeste Belegungskarten** haben ihren Ursprung an einem vorher festgelegten Punkt, z.B. nach GPS-Koordinaten. Das Ego-Fahrzeug wird anhand seiner Position dementsprechend in der Karte „weiter geschoben“. **Fahrzeugsfeste Belegungskarten** dagegen haben ihren Ursprung am eigenen Fahrzeug. Jede Variante hat ihre Vor- und Nachteile.

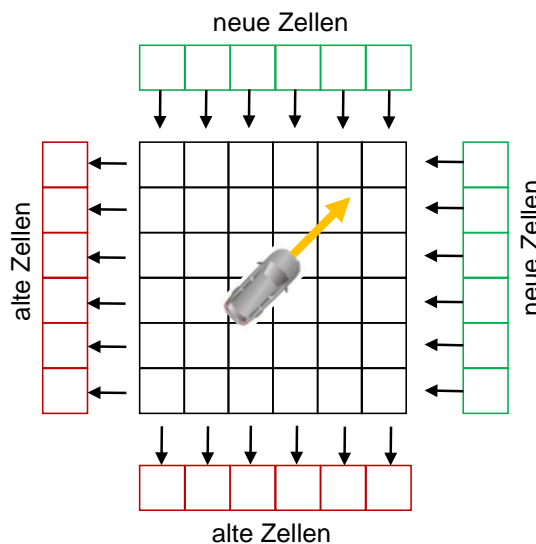


Abbildung 5-2: Verschiebung der Speicher nach Bewegung des Ego-Fahrzeugs

Bei der fahrzeugsfesten Karte muss der Inhalt jeder Zelle nach Bewegung des Ego-Fahrzeugs adaptiert werden. Bewegt sich z. B. das Ego-Fahrzeug 1m nach vorn, werden die



Inhalte der Zellen 1m nach hinten durchgereicht. Dafür werden neue Sensordaten einfacher auf die Karte eingetragen. Dies wird im Verlauf des Kapitels noch ausführlich beschrieben. Bei der ortsfesten Karte muss dagegen nur die Position des Ego-Fahrzeugs in der Karte aktualisiert werden. Dafür hat man die Herausforderung in Verwaltung des Speichers bzw. beim Mappen von Sensordaten zu bewältigen.

In unserer Anwendung wurde eine ortsfeste Belegungskarte ausgewählt. Aufgrund der großen Anzahl an Zellen ist es aufwändig, alle Zellen entsprechend der Bewegung des Ego-Fahrzeugs zu verschieben. Stattdessen ist es effizienter, lediglich die Ego-Position in der Karte zu aktualisieren. In [THR05] wird das Problem von *Simultaneous Localization and Mapping*, kurz geschrieben SLAM, ausführlich dargestellt. Für unsere Anwendung trifft jedoch diese Darstellung nicht zu, da im normalen Verkehrsgeschehen *close-looping* [WAN07] nicht vorkommt. Close-looping adressiert die Problematik, bei der das Ego-Fahrzeug eventuell mehrfach an einer selben Stelle im Kreis fährt und die Umgebung sich kaum verändert [THR05]. Stattdessen wird Wert darauf gelegt, dass der für die Karte benötigte Speicher konstant gehalten wird. Dazu werden Umfeldbereiche, die weit hinter dem Ego-Fahrzeug liegen, für die neuen Bereiche freigegeben. Dieses Prinzip wird in Abbildung 5-2 veranschaulicht. In unserer Anwendung bedeckt eine Zelle einen Bereich von 0.2m x 0.2m. Es wird ein Bereich von 60m x 60m um das Ego-Fahrzeug überwacht bzw. mitgeführt. Das Ego-Fahrzeug wird quasi in der Mitte dieser Teilkarte gehalten. Überschreitet das Fahrzeug den vorher definierten „mittleren Kreis“, in unserem Fall 2m, werden die gegen die Fahrtrichtung am weitesten zurückliegenden Speicherzellen freigegeben, oder anders gesagt, der im Speicher befindliche Inhalt wird gelöscht. Diese Speicher sind in Abbildung 5-2 rot markiert.

Gleichzeitig werden diese Speicher mit neuer Umfeldinformation gefüllt, die den Umfeldbereich repräsentieren, der am weitesten in Fahrtrichtung liegt. Die Umfeldinformation wird durch Sensormessung gewonnen. Darauf wird später ausführlich eingegangen. Durch diesen Mechanismus kann die Menge an Speichern, die für den definierten Umfeldbereich um das Ego-Fahrzeug verwendet wird, konstant gehalten werden.

In der Regel hängt es von der Anwendung ab, wie viele Zellen für einen Umfeldbereich ausreichend sind, oder anders gesagt, wie groß die Abmessung einer Zelle sein sollte. Bei unserer Anwendung [AKT10] reicht eine Zellgröße von 0,2m x 0,2m sowie eine Reichweite von 30m (60m insgesamt) aus. Der gesamte Ablauf bei der Generierung einer Belegungskarte wird in Abbildung 5-3 veranschaulicht. Generell wird die Belegungskarte von zwei Eingangsgrößen verändert bzw. aktualisiert: Sensordaten und Ego-Daten. Auf der Basis erfasster Ego-Daten wie Radgeschwindigkeit, Lenkradwinkel, Egobeschleunigung, etc., wird die Ego-Position anhand eines Einspurmodells geschätzt [SCHI07]. Die Kartensteuerung wertet die Ego-Position aus, damit neue und alte Bereiche verschoben werden können. Auf der anderen Seite werden Sensordaten, in unserem Fall 3D-Punkte des Stereokamera-Systems, verwendet, um den Inhalt der Zellen zu aktualisieren. Dazu gehören ein entsprechendes *inverses Sensormodell* und *Zellaktualisierungsmechanismen*. Dies wird im folgenden Unterkapitel beschrieben.

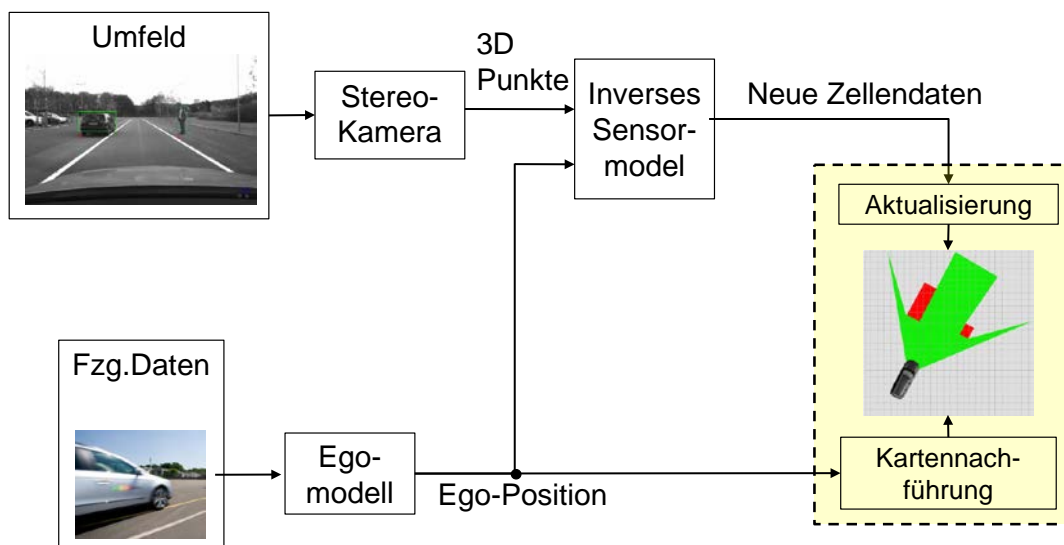


Abbildung 5-3: Abläufe in einem kartenbasierten Umfeldmodellierungsframework

## 5.2 Stereo-spezifisches Inverses Sensormodell

In Kapitel 2 wurde die 3D-Rekonstruktion beschrieben, womit die Umfeldinformation auf Basis von Stereobildern gewonnen wurde. Des Weiteren wurde in Kapitel 4 ein 2D-Histogramm zur Objektdetektion dargestellt. Da unsere Belegungskarte ebenfalls zweidimensional ausgerichtet ist, kann dieses 2D-Histogramm naturgemäß in die Karte als Sen-

sordaten aufgenommen werden. Einheitlich können dieses 2D-Histogramm als „lokale Karte“ und die Belegungskarte als „globale Karte“ bezeichnet werden. Die Information von der lokalen Karte muss auf die globale Karte korrekt abgebildet werden. Dazu werden die zwei Karten wie folgt definiert:

### 5.2.1 Die lokale Karte

Die lokale Karte  $L$  entspricht einer Menge von Zellen  $l_k$ :

$$L = \{l_1, l_2, \dots, l_K\} \quad (5.1)$$

$K$  repräsentiert die gesamte Anzahl von lokalen Zellen während  $k$  der Index jeder Zelle ist. Jede Zelle ist ein Vektor mit folgenden Komponenten:

$$l_k = [b, i, j, \Delta i, \Delta j]^T \quad (5.2)$$

Wobei  $b$  der Belegungsgrad der Zelle,  $i$  und  $j$  die Koordinaten der Zelle im Bezug auf Ego-Fahrzeug,  $\Delta i$  und  $\Delta j$  die Breite und Länge der Zelle sind. In unserer Anwendung ist  $\Delta i$  gleich  $\Delta j$ . Je mehr 3D-Punkte eine Zelle  $l_k$  hat, desto sicherer ist die Zelle als belegt eingestuft [BRA06]. Für den Fall, dass eine Zelle nur wenige 3D-Punkte hat, ist der Belegungsgrad  $b$  dementsprechend niedrig und die Ungewissheit über die Zelle hoch, da weniger Informationen über die Zelle bekannt sind oder gewonnen wurden. Wenn einer Zelle kein 3D-Punkt zugeordnet wird, ist der Belegungsgrad Null. Dies bedeutet aber nicht dass die Zelle frei ist, sondern dass keine Information über die Zelle vorhanden ist. Das Verhältnis zwischen Belegungsgrad  $b$  und Anzahl der 3D Punkte kann über folgende Funktion dargestellt werden:

$$b = f(s_{3D})$$

$$b = \frac{s_{3D}}{\kappa \cdot d^2} \quad (5.3)$$

Dabei ist  $s_{3D}$  die Anzahl der zugehörigen 3D-Punkte,  $d$  der Abstand der Zelle zum Ego-Fahrzeug und  $\kappa$  die Abstand-Belegungsgrad-Konstante. Die theoretische Anzahl der 3D

Punkte, die eine Zelle haben kann, sinkt quadratisch mit der Entfernung [LUH00]. Daher wird die Entfernung  $d$  entsprechend im Nenner aufgeführt.

### 5.2.2 Die globale Karte

Die globale Karte  $C$  ist eine Menge von Zellen  $c_n$ :

$$C = \{c_1, c_2, \dots, c_N\} \quad (5.4)$$

$N$  ist dabei die Anzahl der gesamten Zellen. Jede globale Zelle besitzt folgende Komponente:

$$c_n = [\underline{p}, x, y, \Delta x, \Delta y]^T \quad (5.5)$$

Dabei stellen  $\underline{p}$  der Belegungsgrad einer globalen Zelle dar.  $x$  und  $y$  sind die globalen Koordinaten und  $\Delta x$  und  $\Delta y$  sind die Zellengröße.

Zur Modellierung des Belegungsgrads einer globalen Zelle werden in der Literatur häufig entweder das Bayes-Theorem oder die Dempster-Schafer-Theorie verwendet [MUR00].

#### Nach dem Bayes-Theorem

Um das Bayes-Theorem zu verwenden, wird vorausgesetzt, dass eine Zelle einen von zwei Zuständen hat, *belegt* oder *frei*. Da der Zustand einer Zelle in der Praxis nicht immer 100% sicher festgestellt werden kann, wird jedem Zustand eine Wahrscheinlichkeit zugewiesen. Damit ergibt sich folgende Beziehung zwischen den Wahrscheinlichkeiten [ELF89]:

$$p(\text{belegt}) = 1 - p(\text{frei}) \quad (5.6)$$

Je höher  $p(\text{belegt})$  ist, desto wahrscheinlicher ist die Zelle belegt bzw. desto unwahrscheinlicher ist die Zelle frei. Liegen  $p(\text{belegt})$  und  $p(\text{frei})$  dagegen in der Nähe von 0.5, bedeutet dies eine Ungewissheit in Zustandsbestimmung der Zelle. Hat eine Zelle nach (5.3) wenig 3D-Punkte, ist der lokale Belegungsgrad  $b$  dementsprechend niedrig.

Als Beispiel wird der lokale Belegungsgrad  $b = 0.2$  angenommen. Wenn für die Wahrscheinlichkeit  $p(\text{belegt})$  gleich der Belegungsgrad  $b = 0.2$  übernommen wird, heißt es, dass die Zelle mit einer Wahrscheinlichkeit von  $p(\text{frei}) = 1 - p(\text{belegt}) = 0.8$  nicht belegt ist. Bei genauer Betrachtung bedeutet  $b = 0.2$  aber, dass es mit einer Wahrscheinlichkeit von  $0.8$  nicht sicher gestellt werden kann, wie der Zustand der Zelle ist. Auf diesem Grund werden  $p(\text{belegt})$  und  $p(\text{frei})$  wie folgt berechnet:

$$\begin{aligned} K_{\text{Norm}} &= 2 \\ p(\text{belegt}) &= 0.5 + \frac{b}{K_{\text{Norm}}} = 0.5 + \frac{0.2}{K_{\text{Norm}}} = 0.6 \\ p(\text{frei}) &= 1 - p(\text{belegt}) = 0.4 \end{aligned} \tag{5.7}$$

Der Belegungsgrad fließt in die Berechnung der Wahrscheinlichkeit ein und wird durch die Normierungskonstante  $K_{\text{Norm}}$  gewichtet. Der geringe Unterschied zwischen  $p(\text{belegt})$  und  $p(\text{frei})$  deutet eine unsichere Aussage über die Belegung der Zelle an.

#### Nach der Dempster-Shafer-Theorie

Während bei dem Bayes-Theorem eine Wahrscheinlichkeitsfunktion auf eine Menge von sich ausschließenden Hypothesen, in unserem Fall  $H = \{\text{belegt}, \text{frei}\}$  gesetzt wird, wird nach Dempster-Shafer Theorie eine *Mass Function* definiert, die auf die Menge  $\Theta = 2^H = \{\emptyset, \{\text{belegt}\}, \{\text{frei}\}, \{H\}\}$  setzt [SHA76]. Die Teilmenge  $\{H\} = \{\text{belegt}, \text{frei}\}$  repräsentiert die Ungewissheit. Die Zelle kann entweder *belegt* oder *frei* sein. Die *Mass Function* repräsentiert das Gewicht entsprechend der Zuverlässigkeit für ein Ereignis aus der Menge  $\Theta$ . Angelehnt am Beispiel im vorigen Abschnitt, werden nach der Dempster-Shafer Theorie folgende Glaubensmaße definiert:

$$\begin{aligned} m(\text{belegt}) &= b = 0.2 \\ m(\text{frei}) &= 0 \\ m(\text{unbekannt}) &= 0.8 \\ m(\emptyset) &= 0 \end{aligned} \tag{5.8}$$

Wie schon beschrieben, existiert keine Information oder „Nachweis“ darüber, dass die Zelle frei ist, beträgt das Glaubensmaß  $m(\text{frei})$  Null. Die Leermenge  $\emptyset$  repräsentiert widersprüchliche Aussagen z. B. „belegt **und** frei“, welche in diesem Fall noch nicht auftritt und hat deshalb auch ein Glaubensmaß von Null. Beim Schritt Kartenupdate wird auf  $m(\emptyset)$  ausführlich eingegangen.

### 5.2.3 Abbildung lokaler Karte auf globaler Karte

Eine Aufgabe des Sensormodells ist es, die Menge  $L$  in die Menge  $C$  abzubilden:  $f : L \rightarrow C$ . Es ist auch hervorzuheben, dass nur **belegte Zellen** aus  $L$  auf  $C$  gemappt werden. Die Lage der Karten zueinander wird in Abbildung 5-4 veranschaulicht.

Die Abbildung  $f$  erfolgt in zwei Schritten:

1. Lokalisierung einer lokalen Zelle auf die globale Karte:  $[u, v] \rightarrow [i, j]$
2. Übertragung des Belegungsgrades einer lokalen Zelle auf die Wahrscheinlichkeiten einer globalen Zelle:  $b \rightarrow p$

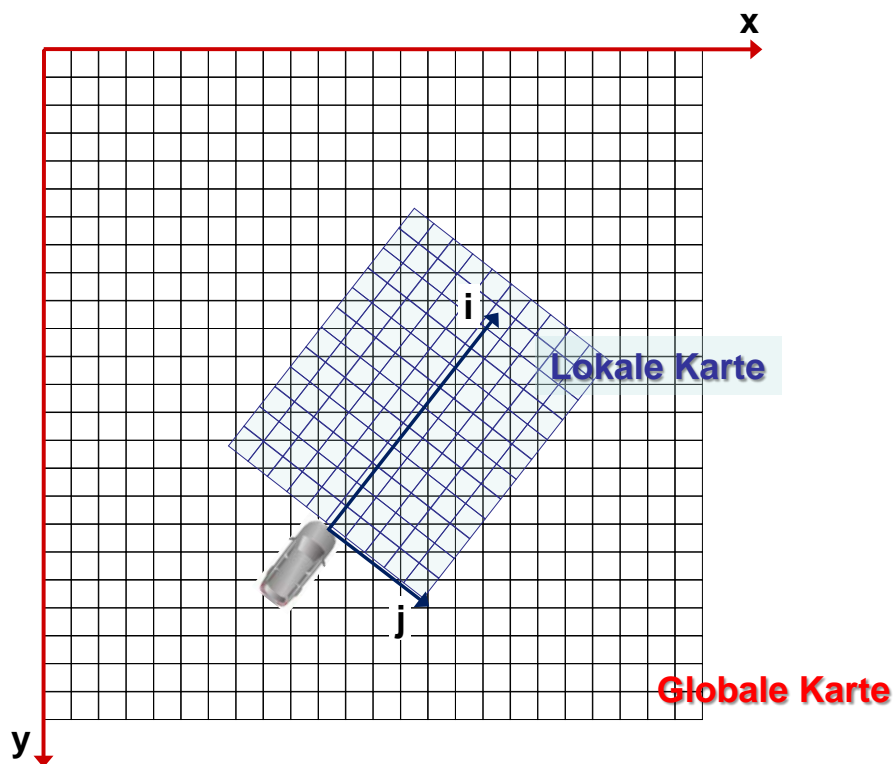


Abbildung 5-4: Gegenüberstellung von lokaler und globaler Karte

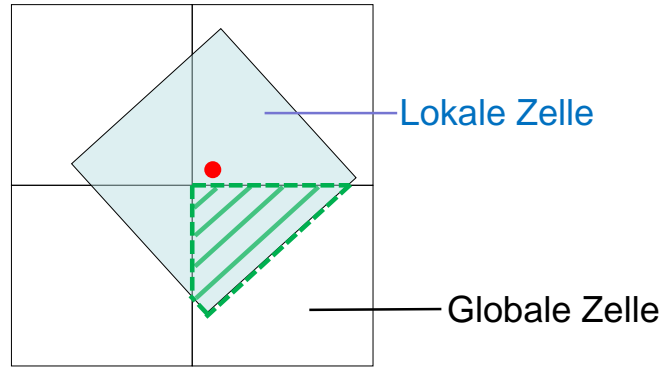


Abbildung 5-5: Übereinanderlegung der Zellen

Beim Schritt 1 wird folgende Koordinatentransformation durchgeführt:

$$\begin{bmatrix} i \\ j \end{bmatrix} = \underline{\mathbf{T}} + \underline{\mathbf{R}} \cdot \begin{bmatrix} u \\ v \end{bmatrix} \quad (5.9)$$

Dabei ist  $\underline{\mathbf{T}}$  die Positionsverschiebung des Ego-Fahrzeug im Bezug auf den Ursprung bei den globalen Koordinaten,  $\underline{\mathbf{R}}$  die Rotation nach dem Winkel zur ursprünglichen Richtung. Damit kann die Mitte einer lokalen Zelle auf die globale Karte positioniert werden. Dies stellt der rote Punkt in Abbildung 5-5 dar.

Schritt 2: Es ist zu erkennen, dass die Fläche einer lokalen Zelle über mehreren globalen Zellen liegen kann. In diesem Fall wird die gemeinsame Fläche, die aus der lokalen Zelle stammt und innerhalb einer globalen Zelle liegt, verwendet um den Belegungsgrad  $p$  zu bestimmen. In Abbildung 5-5 ist diese gemeinsame Fläche grün schraffiert.

Wie schon vorhin beschrieben, kann der Belegungsgrad  $p$  entweder mit dem Bayes-Theorem oder mit der Dempster-Shafer-Theorie modelliert werden. Nach dem BS (Bayes-Theorem) werden die Wahrscheinlichkeiten wie folgt berechnet:

$$\begin{aligned} p(\text{belegt}) &= b \cdot \frac{A_{\text{Cut}}}{A_{\text{Cell}}} \\ p(\text{frei}) &= 1 - p(\text{belegt}) \end{aligned} \quad (5.10)$$

Die *Mass Function* wird nach der Dempster-Shafer Theorie wie folgt berechnet:

$$\begin{aligned}
 m(\text{belegt}) &= b \cdot \frac{A_{\text{Cut}}}{A_{\text{Cell}}} \\
 m(\text{frei}) &= 0 \\
 m(\text{unbekannt}) &= 1 - m(\text{belegt}) \\
 m(\emptyset) &= 0
 \end{aligned}
 \tag{5.11}$$

Dabei ist  $b$  der vorher bestimmte Belegungsgrad dieser lokalen Zelle.  $A_{\text{Cut}}$  entspricht der schraffierten gemeinsamen Fläche von lokaler und globaler Zelle,  $A_{\text{Cell}}$  die gesamte Fläche derjenigen globalen Zelle.

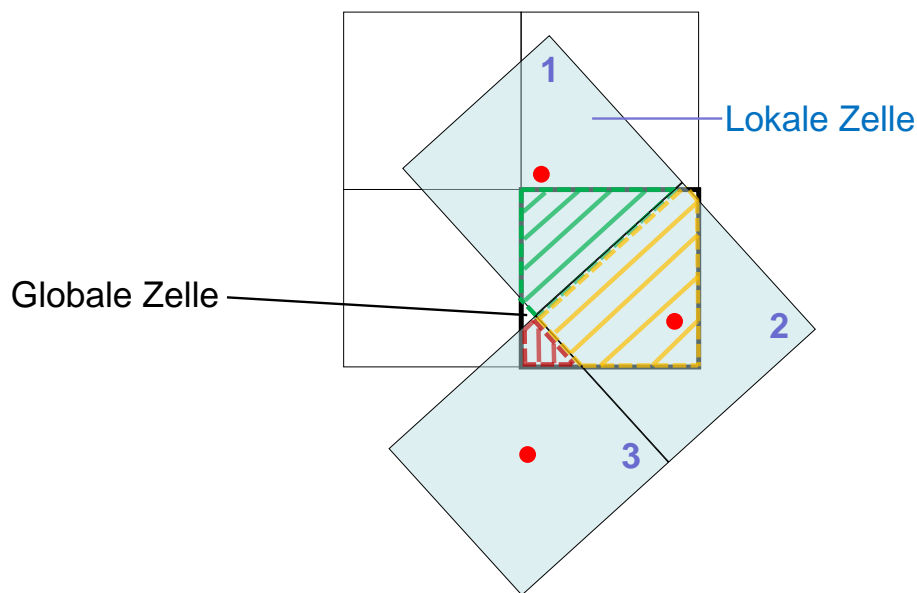


Abbildung 5-6: Zusammensetzung der Zellen

Für den Fall, dass mehrere belegte lokale Zellen nebeneinander liegen wie in Abbildung 5-6, wird die Belegungswahrscheinlichkeit nach dem Bayes-Theorem wie folgt berechnet:

$$\begin{aligned}
 p(\text{belegt}) &= \sum_{i=1, j=1}^{I, J} b_{i,j} \cdot \frac{A_{\text{Cut}}^{i,j}}{A_{\text{Cell}}} \\
 p(\text{frei}) &= 1 - p(\text{belegt})
 \end{aligned}
 \tag{5.12}$$

Nach der Dempster-Shafer-Theorie setzen sich die *Mass Functions* einer globalen Zelle aus:



$$\begin{aligned}
 m(\text{belegt}) &= \sum_{i=1, j=1}^{I, J} b_{i,j} \cdot \frac{A_{\text{Cut}}^{i,j}}{A_{\text{Cell}}} \\
 m(\text{frei}) &= 0 \\
 m(\text{unbekannt}) &= 1 - m(\text{belegt}) \\
 m(\emptyset) &= 0
 \end{aligned}
 \tag{5.13}$$

Dabei gibt  $i$  und  $j$  den Index der lokalen Zelle an, die eine gemeinsame Fläche mit der globalen Zelle hat. Anhand (5.9), (5.10), (5.11), (5.12) und (5.13) werden alle belegten Zellen von der lokalen Karte  $L$  auf die globale Karte  $M$  umgerechnet.

Es werden nun freie Zellen ermittelt. Dabei wird von den Projektionsstrahlen (Lichtstrahlen), die aus den detektierten Objekten stammen, ausgegangen. Wenn ein Hindernis von der Kamera erfasst wird, welches in belegten Zellen resultiert, wird davon ausgegangen, dass zwischen Kamera und Hindernis ein befahrbarer Bereich existiert. Sonst wäre der Lichtstrahl nicht in die Kamera gekommen. In Abbildung 5-7 wird dieses Prinzip veranschaulicht. Dieser befahrbare Bereich entspricht den freien Zellen in Abbildung 5-7.

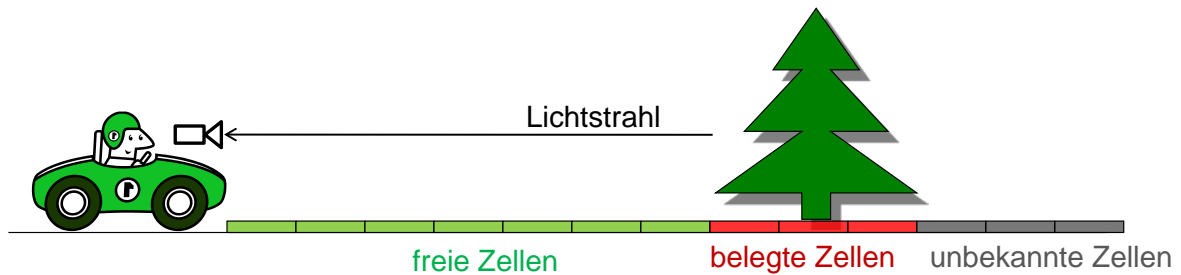


Abbildung 5-7: Initiierung freier Zellen nach dem Lichtstrahl

Mit diesem Prinzip werden freie Zellen auf der globalen Karte ausgehend von vorher umgemappten belegten Zellen initiiert. Auf der Basis des Kameraöffnungswinkels wird, wie in Abbildung 5-8 dargestellt, von links nach rechts ein virtueller Lichtstrahl initiiert. Trifft der Lichtstrahl auf eine belegte Zelle, werden alle Zellen entlang des Strahls bis zur belegten Zelle als *frei* klassifiziert. Dies bedeutet, diesen Zellen wird durch die *Mass Function* ein Gewicht  $m(\text{frei})$  zugeordnet, welches genau so groß wie das Gewicht  $m(\text{belegt})$  der vom Lichtstrahl getroffenen belegten Zelle. Um Zellen entlang des Strahls effizient anzusprechen, wird der Bresenham-Algorithmus verwendet [BRE65].

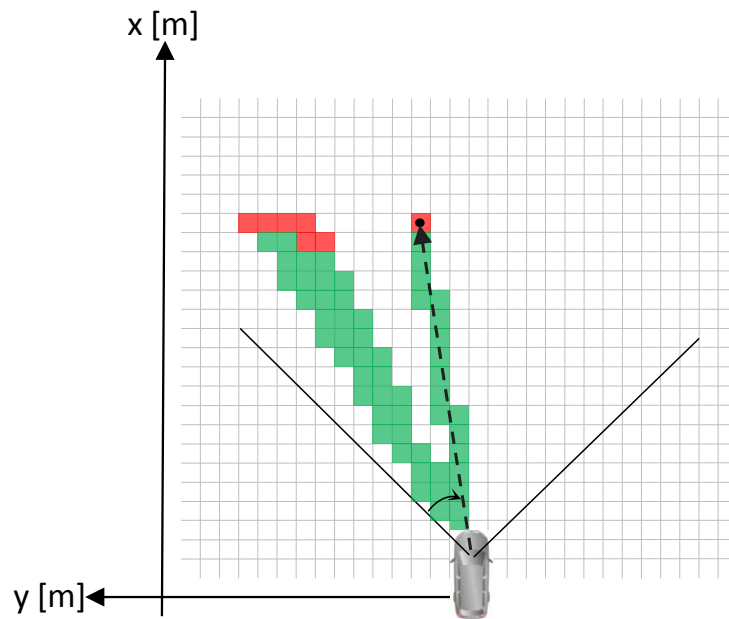


Abbildung 5-8: Generierung der freien Bereiche mit dem Bresenham-Algorithmus

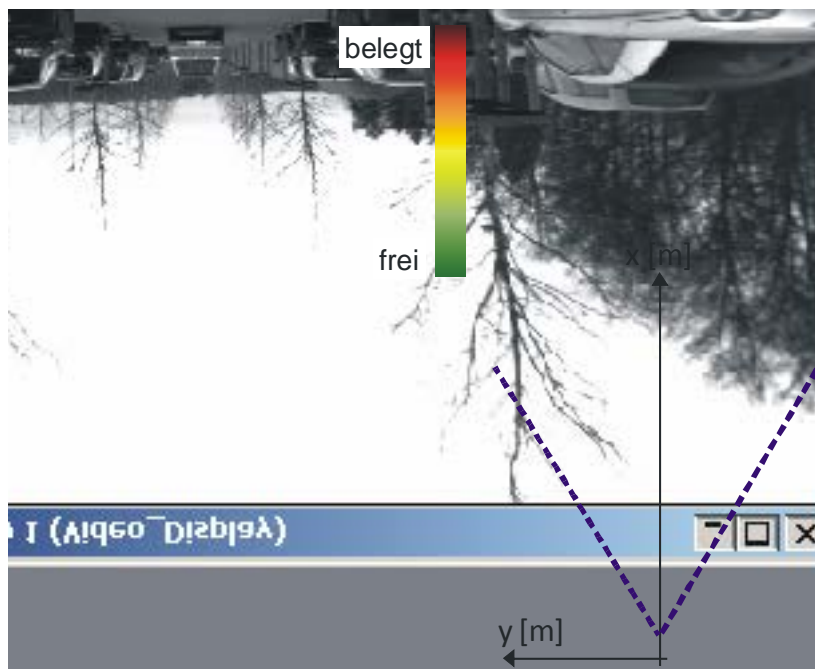


Abbildung 5-9: Die globale Belegungskarte, die auf Basis einer Messung erstellt wird. Unten links: das Kamerabild der Szene. Rechts: die Belegungskarte.

Der Bresenham-Algorithmus ist ein Algorithmus in der Computergrafik zum Zeichnen von Geraden auf Rasteranzeigen. Das Besondere an dem Algorithmus ist, dass er Rundungsfehler, die durch die Diskretisierung von kontinuierlichen Koordinaten entstehen, minimiert, und gleichzeitig einfach implementierbar ist, mit der Addition von ganzen Zahlen als komplexeste Operation, und somit ohne Multiplikation, Division und Gleitkommazahlen aus-

kommt. In Abbildung 5-8 werden freie Bereiche, die auf dieser Weise entstehen, veranschaulicht. Globale Zellen, die weder durch Ummappen nach (5.9) und (5.13) als *belegt* noch durch das Lichtstrahlprinzip als *frei* dargestellt werden, werden als *unbekannt* klassifiziert. Nach der Dempster-Shafer-Theorie bedeutet dies  $m(\text{unbekannt}) = 1$ ,  $m(\text{belegt}) = m(\text{frei}) = m(\emptyset) = 0$ . Nach dem Bayes-Theorem bedeutet dies  $p(\text{belegt}) = p(\text{frei}) = 0.5$ .

Als Ergebnis wird die globale Karte in Abbildung 5-9 dargestellt. Grüne Zellen sind mit hoher Wahrscheinlichkeit *frei*, rote Zellen stellen Bereiche mit Hindernissen dar und graue Zellen sind Bereiche worüber durch Messung keine Information gewonnen werden konnten. Es ist zu erwähnen, dass die Umsetzung von der lokalen Karte **L** zur globalen Karte **C** für jede neue Messung, also für jedes Stereobildpaar erfolgt.

### 5.3 Kartenaktualisierung

Beschrieben durch den obigen Begriff werden zum einen Zellen, auf die neue Messungen abgebildet werden, aktualisiert und zum anderen Zellen, die nicht mehr im Beobachtungsbereich sind, entsprechend *gealtert*. Das Ziel der Aktualisierung ist es, die über mehrere Messungen gewonnenen Informationen sinnvoll zu kombinieren, um eine zuverlässige Aussage über den Belegungszustand der Zelle treffen zu können. Abweichungen, die durch Messrauschen entstehen, z. B. eine belegte Zelle wird als frei gemessen, sollen durch eine Vielzahl an Messungen ausgeglichen werden. Dabei wird angenommen, dass das Messrauschen normalverteilt ist. Da im vorigen Unterkapitel nach dem Bayes-Theorem und der Dempster-Shafer-Theorie unterschiedliche Ansätze zur Beschreibung einer Zelle vorliegen, werden zur Zellenaktualisierung weiterhin diese zwei Ansätze verwendet.

#### 5.3.1 Verwendung des Bayes-Theorems

Die Wahrscheinlichkeit, dass eine globale Zelle zum Zeitpunkt 1 belegt ist, vorausgesetzt dass die Messung  $z_1$  erfolgt ist, wird als  $p(\mathbf{B}|z_1)$  bezeichnet. Dagegen wird  $p(\bar{\mathbf{B}}|z_1)$  als Wahrscheinlichkeit, dass die Zelle frei ist, bezeichnet. Die a-priori Wahrscheinlichkeit für

eine Zelle, die bis dato unbekannt ist, werden als  $p_0(\mathbf{B}) = p_0(\overline{\mathbf{B}}) = 0.5$  angenommen. Gemäß [ELF89] wird  $p(\mathbf{B}|z_1)$  und  $p(\overline{\mathbf{B}}|z_1)$  auf Basis von a-priori Wahrscheinlichkeiten wie folgt berechnet:

$$p(\mathbf{B}|z_1) = \frac{p(z_1|\mathbf{B}) \cdot p_0(\mathbf{B})}{p(z_1|\mathbf{B}) \cdot p_0(\mathbf{B}) + p(z_1|\overline{\mathbf{B}}) \cdot p_0(\overline{\mathbf{B}})} \quad (5.14)$$

Wobei  $p(z_1|\mathbf{B})$  die Wahrscheinlichkeit darstellt, dass die Messung  $z_1$  entsteht wenn die Zelle belegt ist. Diese Wahrscheinlichkeit wird zur Messwertaktualisierung (engl. Measurement Update) verwendet. Diese Wahrscheinlichkeit kann auch als die True-Positive-Rate von einem Sensor gesehen werden. Dagegen stellt  $p(z_1|\overline{\mathbf{B}})$  die Wahrscheinlichkeit dar, dass die Messung  $z_1$  gewonnen wird wenn die Zelle frei ist. Diese entspricht der True-Negative-Rate.

Es ist hervorzuheben, dass (5.14) nur unter der Annahme, dass *der Zustand einer Zelle über die Zeit konstant bleibt*, gültig ist. Dies gilt auch für alle darauf aufbauenden Formeln.

Wird bei der nächsten Messung der Messwert  $z_2$  gewonnen, kann eine Belegungswahrscheinlichkeit auf Basis von  $z_1$  und  $z_2$  nach der Bayes-Formel [BAY63] wie folgt berechnet werden:

$$\begin{aligned} p(\mathbf{B}|z_1, z_2) &= \frac{p(z_2|\mathbf{B}, z_1) \cdot p(\mathbf{B}|z_1)}{p(z_2|z_1)} \\ &= \frac{p(z_2|\mathbf{B}) \cdot p(\mathbf{B}|z_1)}{p(z_2|\mathbf{B}) \cdot p(\mathbf{B}|z_1) + p(z_2|\overline{\mathbf{B}}) \cdot p(\overline{\mathbf{B}}|z_1)} \end{aligned} \quad (5.15)$$

Es wird angenommen, dass die Messungen voneinander unabhängig sind bzw. sich der Zustand der Zelle nicht verändert. Anhand von (5.15) können alle durchgeführten Messungen  $\{z_1, z_2, \dots, z_n\}$  zur Schätzung der Belegung der Zelle verwendet werden:

$$p(\mathbf{B}|z_k, Z_{k-1}) = \frac{p(z_k|\mathbf{B}) \cdot p(\mathbf{B}|Z_{k-1})}{p(z_k|\mathbf{B}) \cdot p(\mathbf{B}|Z_{k-1}) + p(z_k|\bar{\mathbf{B}}) \cdot p(\bar{\mathbf{B}}|Z_{k-1})} \quad (5.16)$$

Wobei  $Z_{k-1} = \{z_1, z_2, \dots, z_{k-1}\}$  die gesamte Messungen bis  $k-1$  repräsentiert. Die Wahrscheinlichkeit, dass die Zelle nicht belegt ergibt sich zu:

$$p(\bar{\mathbf{B}}|z_k, Z_{k-1}) = 1 - p(\mathbf{B}|z_k, Z_{k-1}) \quad (5.17)$$

### 5.3.2 Verwendung der Dempster-Shafer-Theorie

Wie oben auf der Seite 77 bereits dargelegt, kann der Zustand einer Zelle nach der Dempster-Shafer Theorie durch die Menge  $\Theta = \{\emptyset, \{\mathbf{B}\}, \{\mathbf{F}\}, \{\mathbf{U}\}\}$  beschrieben werden.  $\mathbf{B}$  steht für belegt,  $\mathbf{F}$  steht für frei,  $\mathbf{U}$  steht für den unbekanntem Zustand, also entweder belegt oder frei. Für die Messung 1 werden durch die *Mass Function* folgende Gewichte definiert:  $m_{z_k}(\mathbf{B})$ ,  $m_{z_k}(\mathbf{F})$  und  $m_{z_k}(\mathbf{U})$ .

Bei der zweiten Messung können die Gewichte mit der Dempster-Kombinationsregel [MUR00] fusioniert werden.

$$\begin{aligned} m_2(\mathbf{B}) &= \frac{m_{z_1}(\mathbf{B}) \cdot m_{z_2}(\mathbf{B}) + m_{z_1}(\mathbf{B}) \cdot m_{z_2}(\mathbf{U}) + m_{z_1}(\mathbf{U}) \cdot m_{z_2}(\mathbf{B})}{1 - m_{z_1}(\mathbf{B}) \cdot m_{z_2}(\mathbf{F}) - m_{z_1}(\mathbf{F}) \cdot m_{z_2}(\mathbf{B})} \\ m_2(\mathbf{F}) &= \frac{m_{z_1}(\mathbf{F}) \cdot m_{z_2}(\mathbf{F}) + m_{z_1}(\mathbf{F}) \cdot m_{z_2}(\mathbf{U}) + m_{z_1}(\mathbf{U}) \cdot m_{z_2}(\mathbf{F})}{1 - m_{z_1}(\mathbf{B}) \cdot m_{z_2}(\mathbf{F}) - m_{z_1}(\mathbf{F}) \cdot m_{z_2}(\mathbf{B})} \end{aligned} \quad (5.18)$$

$$m_2(\mathbf{U}) = 1 - m_2(\mathbf{B}) - m_2(\mathbf{F})$$

Das Gewicht für widersprüchliche Informationen  $m(\emptyset)$  wird nach der Dempster-Kombinationsregel in der neuen Verteilung der Gewichte nicht aufgeführt. Es wurde verwendet, um die restlichen Gewichte zu „normieren“. Die Normierung entsteht durch die Teilung der Maße durch  $1 - m(\emptyset)$  wobei  $m_2(\emptyset) = m_{z_1}(\mathbf{B}) \cdot m_{z_2}(\mathbf{F}) + m_{z_1}(\mathbf{F}) \cdot m_{z_2}(\mathbf{B})$ . Damit summieren sich  $m_2(\mathbf{B})$ ,  $m_2(\mathbf{F})$  und  $m_2(\mathbf{U})$  auf 1.

Die Berechnungen aus (5.18) können rekursiv für alle nachfolgenden Messungen durchgeführt werden:

$$\begin{aligned}
 m_k(\emptyset) &= m_{k-1}(\mathbf{B}) \cdot m_{z_k}(\mathbf{F}) + m_{k-1}(\mathbf{F}) \cdot m_{z_k}(\mathbf{B}) \\
 m_k(\mathbf{B}) &= \frac{m_{k-1}(\mathbf{B}) \cdot m_{z_k}(\mathbf{B}) + m_{k-1}(\mathbf{B}) \cdot m_{z_k}(\mathbf{U}) + m_{k-1}(\mathbf{U}) \cdot m_{z_k}(\mathbf{B})}{1 - m_k(\emptyset)} \\
 m_k(\mathbf{F}) &= \frac{m_{k-1}(\mathbf{F}) \cdot m_{z_k}(\mathbf{F}) + m_{k-1}(\mathbf{F}) \cdot m_{z_k}(\mathbf{U}) + m_{k-1}(\mathbf{U}) \cdot m_{z_k}(\mathbf{F})}{1 - m_k(\emptyset)} \\
 m_k(\mathbf{U}) &= 1 - m_k(\mathbf{B}) - m_k(\mathbf{F})
 \end{aligned} \tag{5.19}$$

Wobei  $m_{k-1}(\mathbf{B}), m_{k-1}(\mathbf{F})$  und  $m_{k-1}(\mathbf{U})$  die bis dato auf Basis aller vorhandenen Messungen berechneten Gewichte sowie  $m_{z_k}(\mathbf{B}), m_{z_k}(\mathbf{F})$  und  $m_{z_k}(\mathbf{U})$  die zum Zeitpunkt  $k$  per Messung gewonnenen Gewichte sind. Dies stellt eine Alternative im Vergleich zum Verfahren nach Bayes in (5.16) dar.

### 5.3.3 Vergleich der beiden Verfahren

Der grundsätzliche Unterschied zwischen den beiden Verfahren ist die Möglichkeit, Ungewissheit bzw. Unwissen auszudrücken. Um dies zu veranschaulichen, wird das Beispiel aus Abschnitt 5.2 verwendet und erweitert.

Nach der Dempster Shafer Theorie	Nach dem Bayes Theorem
Bevor die Zellen gemessen wurden, sind die a-priori-Werte der Zellen wie folgt angenommen:	
$m_0(\mathbf{B}) = 0$ $m_0(\mathbf{F}) = 0$ $m_0(\mathbf{U}) = 1$	$p_0(\mathbf{B}) = 0.5$ $p_0(\mathbf{F}) = 0.5$
<u>Beispiel:</u> Bei der ersten Messung kann aufgrund der Entfernung z.B. nur mit einer Wahrscheinlichkeit von 0.6 festgestellt werden, dass die Zelle belegt ist. Es gibt sonst keinen Hinweis ob die Zelle frei ist. Damit ergeben sich:	

$m_1(\mathbf{B}) = 0.2$ $m_1(\mathbf{F}) = 0$ $m_1(\mathbf{U}) = 0.8$	$p_1(\mathbf{B}) = 0.6$ $p_1(\mathbf{F}) = 0.4$
<p>Nach der Dempster-Shafer Theorie beträgt die Ungewissheit <math>m_1(\mathbf{U}) = 0.8</math>. Nach dem Bayes Theorem kann die Ungewissheit nicht ausgedrückt werden. Stattdessen werden <math>p(\mathbf{B})</math> und <math>p(\mathbf{F})</math> nach Gl. (5.7) berechnet.</p> <p><u>Beispiel:</u> Bei einer zweiten Messung wird durch das Annähern des Ego-Fahrzeugs an das Objekt ein höheres Belegungsgrad, nämlich 0.6 gemessen. Damit ergeben sich:</p>	
Aktualisierung nach (5.19): $m_2(\mathbf{B}) = 0.68$ $m_2(\mathbf{F}) = 0$ $m_2(\mathbf{U}) = 0.32$	Aktualisierung nach (5.16): $p_2(\mathbf{B}) = 0.85$ $p_2(\mathbf{F}) = 0.15$
<p><u>Beispiel:</u> Bei der dritten Messung wird die Zelle mit einer Wahrscheinlichkeit von 0.7 als frei gemessen, da das Objekt, das über der Zelle stand, sich fort bewegt hat. Damit ergeben sich:</p>	
Nach (5.19): $m_3(\mathbf{B}) = 0.389$ $m_3(\mathbf{F}) = 0.4$ $m_3(\mathbf{U}) = 0.21$  Das Gewicht für Konflikt beträgt $m_3(\emptyset) = 0.476$	Nach (5.16): $p_3(\mathbf{B}) = 0.5$ $p_3(\mathbf{F}) = 0.5$
<p>Nach dem Bayes Theorem entsprechen die obigen Wahrscheinlichkeiten denen zum Anfang, als die Zellen noch nicht vermessen wurden. Es kann also nicht unterschieden werden, ob eine Zelle unbekannt ist oder ob konfliktbehaftete Messungen vorliegen, die auf ein bewegtes Objekt hindeuten könnte. Auf Basis der Gewichte gemäß der Dempster-Shafer-Theorie lässt sich erkennen, dass die Zelle länger beobachtet wurde, <math>m_3(\mathbf{U}) = 0.21</math>, bzw. dass eine konfliktbehaftete Situation vorliegt, <math>m_3(\emptyset) = 0.476</math></p>	

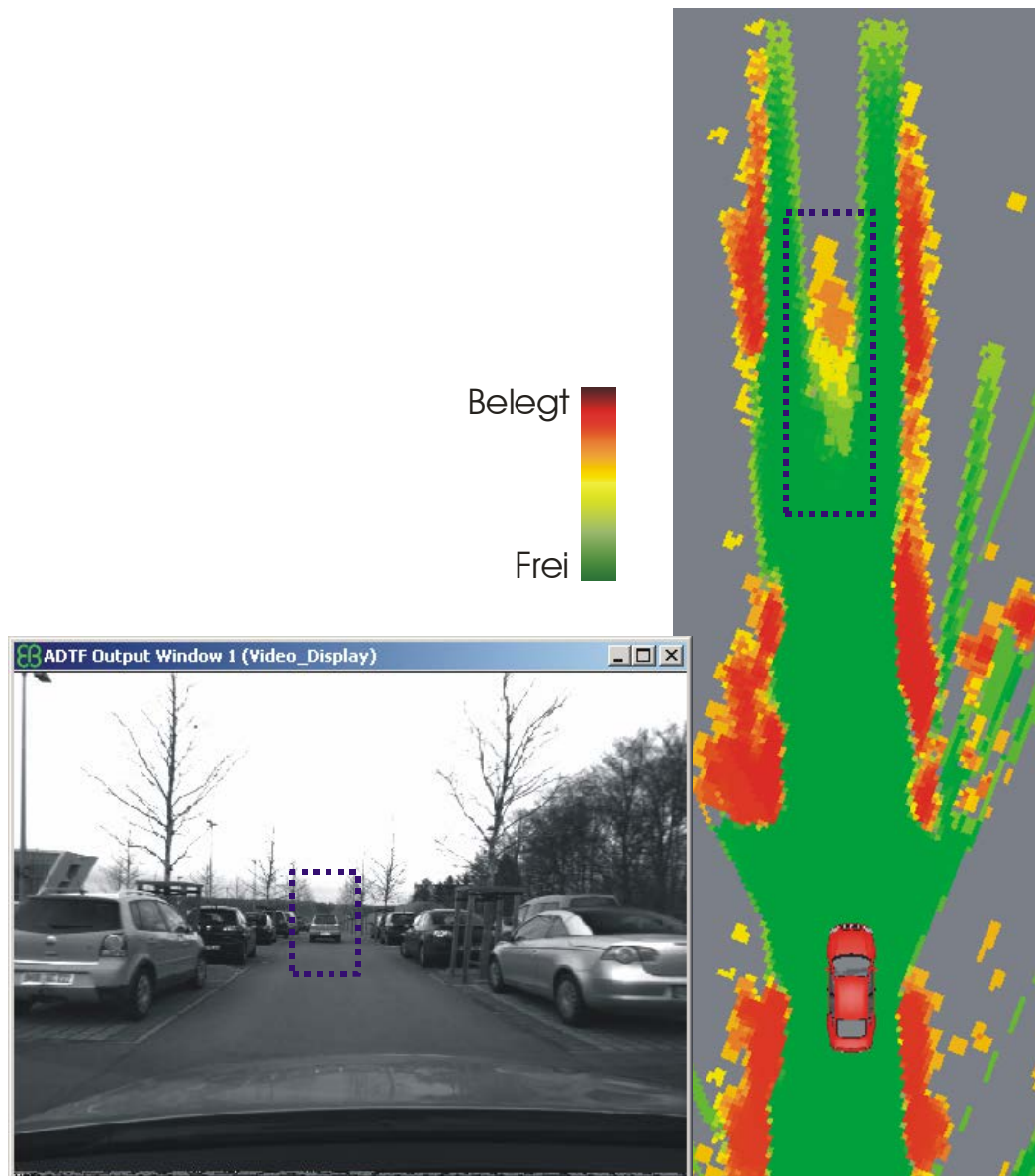


Abbildung 5-10: Links: das Kamerabild. Rechts: die globale Belegungskarte nach der Fusion mehrerer Messungen

Aus diesem Grund wird in dieser Arbeit zur Modellierung und Aktualisierung der Zustände der Zellen das Verfahren nach der Dempster-Shafer Theorie eingesetzt. Eine detailliertere Analyse und Vergleiche findet man in [HEU10]. Das Ergebnis der Anwendung mit der Dempster-Shafer-Theorie ist beispielsweise in Abbildung 5-10 dargestellt. Im Vergleich zur Abbildung 5-9 erkennt man, dass die Farben gesättigter sind. Dies deutet auf eine „eindeutigere“ Belegungswahrscheinlichkeit, entweder ist die Zelle mit hoher Wahrscheinlichkeit belegt oder mit hoher Wahrscheinlichkeit frei.



Dies folgt aus der Fusion aller gewonnenen Messwerte, welches zu einer zuverlässigeren Schätzung der Zelle führt. Die belegten und freien Bereiche auf der Karte decken die im Videobild visualisierten Hindernissen und Fahrbahnen gut ab (Abbildung 5-10). Der einzige Bereich, der nicht eindeutig belegt oder frei ist, wird durch den blauen Kasten in Abbildung 5-10 gekennzeichnet.

Es wird ein voraus fahrendes Fahrzeug betrachtet. Durch seine Bewegung wird der Zustand von Zellen, die entlang des Fahrwegs liegen, kontinuierlich verändert. Dadurch entstehen für dieselbe Zelle unterschiedliche Messwerte (einmal *belegt*, einmal *frei*). Dies verursacht nach (5.19) erhebliche Konfliktmaße  $m(\emptyset)$ . Dies führt dazu, dass die Belegt- und Freimaße  $m(\mathbf{B})$  und  $m(\mathbf{F})$  ähnlich groß sind und damit der Zustand der Zelle nicht eindeutig angegeben werden kann. Die Genauigkeit der Belegungskarte wird dadurch beeinflusst.

#### 5.3.4 Behandlung von bewegten Objekten

Es ist deshalb sinnvoll, Messungen von Zellen, die in diesem Moment von einem bewegten Objekt belegt sind, zu identifizieren bzw. bei der Zellenaktualisierung nicht zu berücksichtigen.

Davon ausgehend ist es problematisch, die richtigen Konfliktzellen zu identifizieren, um sie von der Aktualisierung in (5.19) herauszunehmen. Des Weiteren stellen Konfliktzellen häufig lediglich einen Teil des bewegten Objekts dar. Der Rest des Objekts wird durch belegte Zellen dargestellt, die durch die Überlappung bei der Bewegung ihren Zustand nicht ändern. Dies wird in Abbildung 5-11 veranschaulicht. Dies bedeutet, dass nicht nur die Konfliktzellen sondern auch die zum bewegten Objekt gehörigen Zellen zu identifizieren sind. Nach [VU09] wurde eine modellbasierte Segmentierung durchgeführt, um dieses Problem zu lösen.

Für unsere Anwendung bringt dieser Ansatz aufgrund von Messrauschen (siehe Abbildung 5-12) kein sicheres Ergebnis. An vielen Stellen entstehen Konfliktzellen aufgrund von Messrauschen, obwohl dort kein bewegtes Objekt existiert.

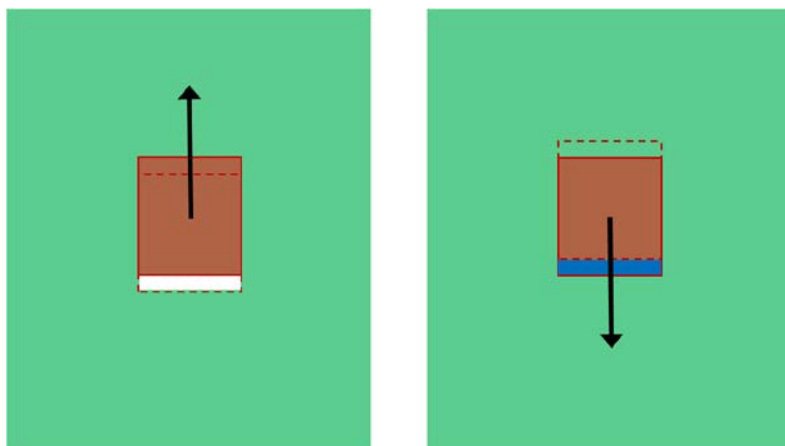


Abbildung 5-11: Konfliktbehaftete Zellen entstehen durch Objektbewegung. Links: Objekt bewegt sich vom Ego-Fahrzeug weg. Rechts: Objekt bewegt sich zum Ego-Fahrzeug.

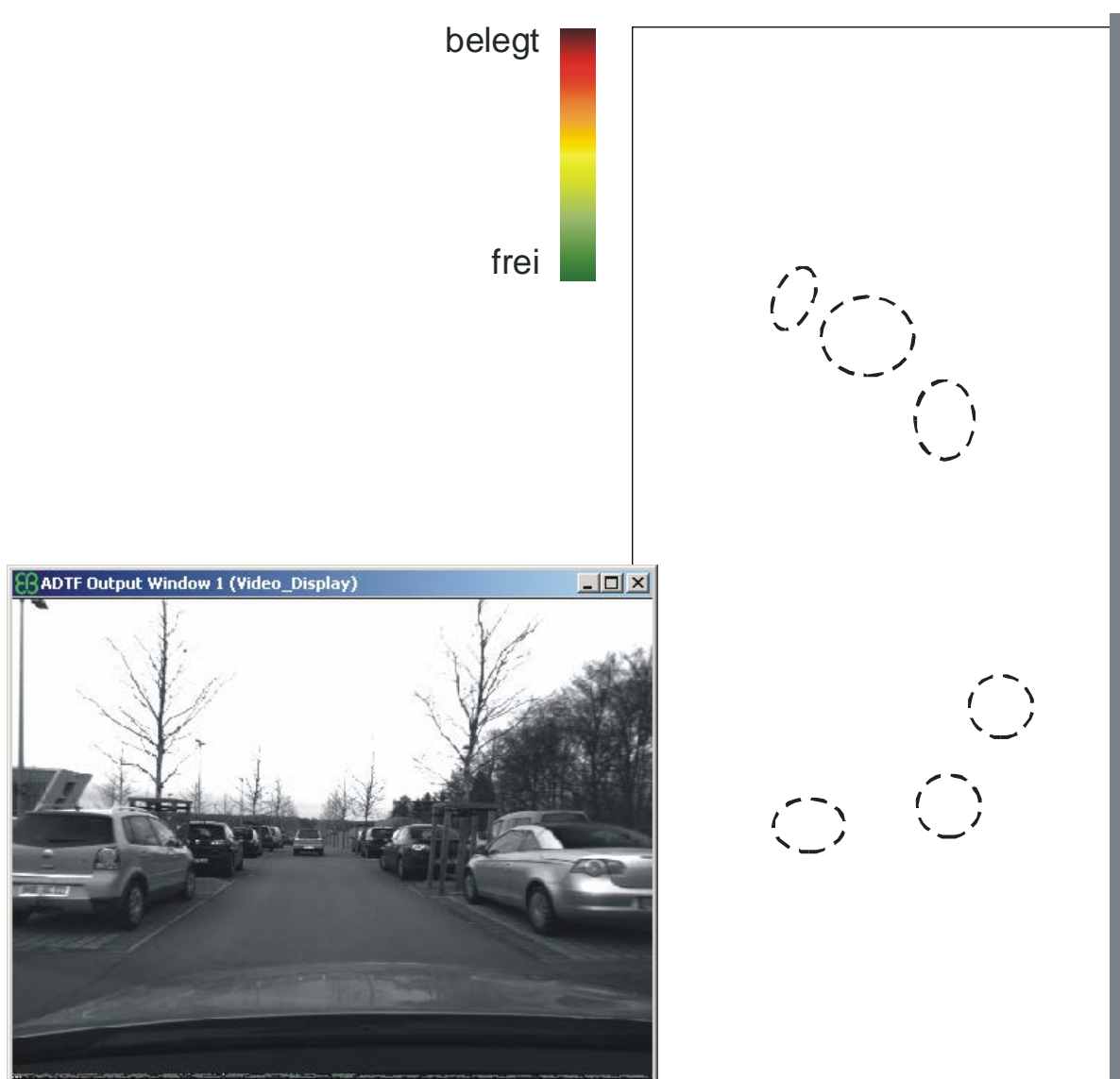


Abbildung 5-12: Links: das Kamerabild. Rechts: die Belegungskarte, mit Zellen, deren Zustände bei zwei aufeinanderfolgenden Messungen sich ändern. Diese Zellen werden blau und weiß gekennzeichnet.

Auf der anderen Seite wurde in Kapitel 4 bereits die Segmentierung bzw. das Objekttracking durchgeführt. Daher ist es an dieser Stelle sinnvoll, die Information aus dem Objekttracking zu benutzen, um Messungen, die zu bewegten Objekten gehören, zu identifizieren. Dies wird wie folgt schrittweise umgesetzt:

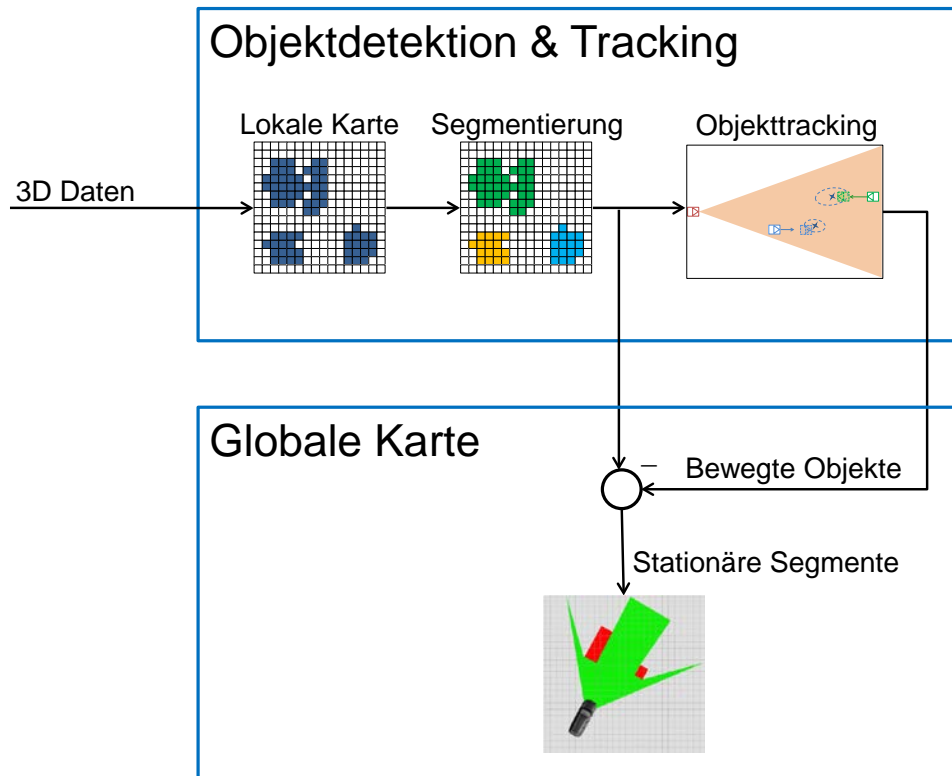


Abbildung 5-13: Verwendung des Tracking zur Selektion von stationären Segmenten

Gemäß Abbildung 5-13 werden Objekte, die aus dem Tracking als „bewegt“ geschätzt wurden ( $|\mathbf{v}_x| + |\mathbf{v}_y| > v_{\min}$ ), verwendet, um Segmente, die bereits mit diesen Objekten assoziiert wurden, zu identifizieren. Die minimale Geschwindigkeit  $v_{\min}$  wird in der Regel in dieser Anwendung als ca. 2 km/h definiert. Eine genauere Geschwindigkeitsschätzung ist mit dem vorhandenen Objekttracking noch nicht möglich.

Dementsprechend werden Segmente, die mit den bewegten Objekten assoziiert wurden, nicht auf die globale Karte gemappt. Es werden ausschließlich Segmente, die aus stationären Hindernissen stammen, in der globalen Karte behandelt. Dadurch wird die Annahme in (5.14) eingehalten. Das Ergebnis wird beispielsweise in Abbildung 5-14 veranschaulicht. Dabei ist das zum bewegten Objekt gehörige Segment mit Lila-Farbe gekennzeichnet, wurde jedoch nicht bei der Kartenaktualisierung berücksichtigt. Im Vergleich zur Abbildung

5-10 ist zu erkennen, dass die Region, in der sich das bewegte Objekt befindet, eindeutiger Information über den Belegungszustand aufweist. Eine quantitative Auswertung der Verbesserung in diesem Zusammenhang wird in Kapitel 6 gegeben.

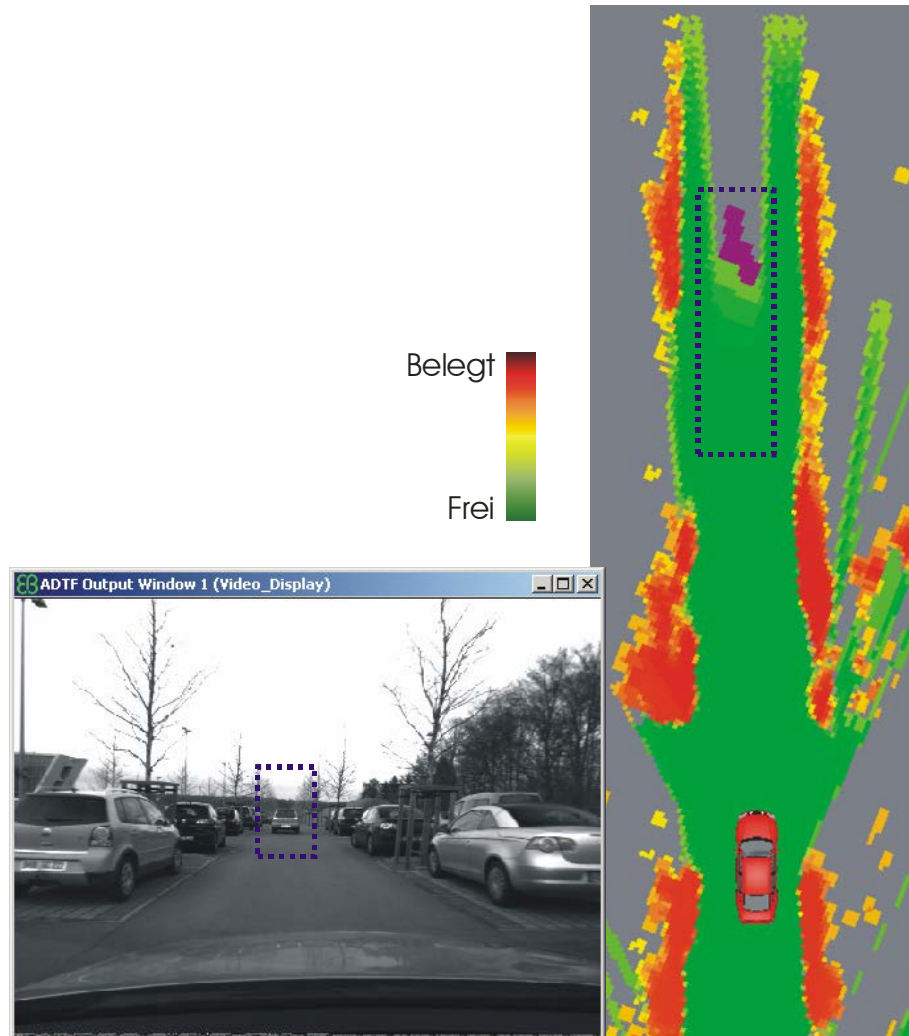


Abbildung 5-14: Verwendung des Tracking zur Verbesserung der Qualität der Belegungskarte. Links: das Kamerabild. Rechts: die Belegungskarte. Die in Lila markierte Region visualisiert, wo das bewegte Objekt sich befindet. Die Messung vom bewegten Objekt wird nicht in der Karte fusioniert. Dadurch ist die Belegungskarte im Vergleich zur Abbildung 5-10 genauer und zuverlässiger.

## 5.4 Zusammenfassung

In diesem Kapitel wurden die Grundlagen zur Anwendung der Belegungskarte beschrieben, als eine Alternative im Vergleich zum objektbasierten Ansatz. Für die Beschreibung des Umfeldes wurde die globale Karte eingesetzt, während für die Interpretation der Kameradaten die lokale Karte in Anwendung gebracht wurde. Des Weiteren wurden zur Modellie-

rung des Belegungszustandes Methoden basierend auf dem Bayes-Theorem und der Dempster-Shafer-Theorie beschrieben bzw. gegenübergestellt. Das inverse Sensormodell setzt die gewonnenen Sensordaten in neue Zellinformationen um, die mit der vorhandenen Information der Zelle entweder nach dem Bayes-Theorem oder der Dempster-Shafer-Theorie fusioniert werden können. Des Weiteren wurde gezeigt, dass bewegte Objekte die Genauigkeit und Zuverlässigkeit der Karte negativ beeinflussen. In [VU09] wird ein Ansatz präsentiert, der aus Differenzen zwischen zwei nacheinander folgenden Messungen bewegte Objekte identifiziert.

In dieser Arbeit wird ein anderer Ansatz verfolgt. Nämlich das Ergebnis aus dem Objekttracking wird verwendet, um Messungen, die von bewegten Objekten stammen, zu identifizieren und somit nicht in der globalen Karte zu berücksichtigen. Damit können stationäre Hindernisse im Verkehrsumfeld sowie befahrbare Bereiche mit Hilfe der Belegungskarte modelliert und zuverlässig beschrieben werden. Die Testergebnisse werden im nachfolgenden Kapitel dargestellt.



# Kapitel 6

## Experimentelle Ergebnisse und Analyse

Die in den vorigen Kapiteln dargestellten Methoden wurden in der Programmiersprache C++ umgesetzt. Die implementierte Software wurde in einem Versuchsträger im Rahmen des Projektes AKTIV-SFR [AKT10] integriert und getestet. Dabei wird ein C++ Bibliothek und Framework namens ADTF (Automotive Data and Time triggered Framework) verwendet. Dieses Framework wurde für Echtzeitfähigkeit und Ansteuerung von automotiven Einheiten wie CAN, Flexray, Ethernet entwickelt. Mit dem in Abbildung 6-1 gezeigten Versuchsträger wurden auch umfangreiche Messdaten aufgenommen. In Abbildung 6-1 werden der Versuchsträger und das eingesetzte Stereokamera-System veranschaulicht. Die Bildaufnahme wurde auf Basis einer Bildauflösung von 512 x 383 Pixeln durchgeführt.



Abbildung 6-1: das Testfahrzeug mit dem eingesetzten Stereokamera-System

Der Gesamtprozess von der Softwareentwicklung bis zur Erprobung im Testgelände sowie die Messdatenaufnahme im realen Verkehr wird in Abbildung 6-2 dargestellt. Es wurde insgesamt eine Datenmenge von 5 Stunden Video mit 20 Frames pro Sekunde aufgenommen.

Da das Ziel des Projektes die Verbesserung der Sicherheit für Fußgänger und Radfahrer ist, wurden überwiegend Daten im innerstädtischen Verkehr aufgezeichnet. Abgeleitet von der Unfallanalyse in Kapitel 1 wurden gezielt Daten in unfallrelevanten Verkehrssituationen aufgenommen und ausgewertet. Bevor auf ein einzelnes relevantes Szenario eingegangen wird, wird eine Fehleranalyse für die gesamte Algorithmen-Kette durchgeführt.

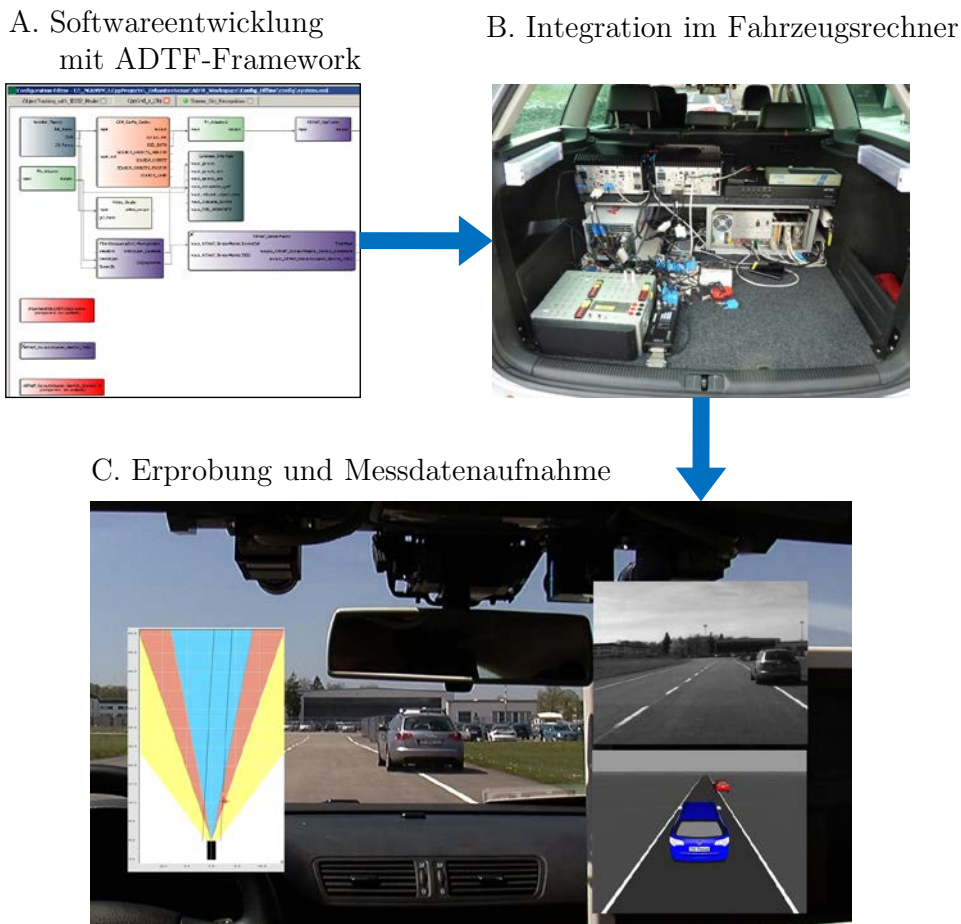


Abbildung 6-2: Prozess zur Umsetzung entwickelter Algorithmen

In Abbildung 6-3 werden die Hauptbestandteile der verwendeten Algorithmen, die in den vorangehenden Kapiteln in einzelnen Verarbeitungsschritten beschrieben wurde, dargestellt. Anhand der Abbildung 6-3 ist zu erkennen, dass die bei einem Schritt entstandenen Fehler sich auch auf die nachfolgenden Schritte auswirken.

Dabei beinhalten der Schritt „ObjektTracking“ sowie der Schritt „Belegungskarte“ die größte Unsicherheit, weil die in den vorigen Schritten entstandenen Fehler übernommen werden. Auf die Auswertung einzelner Teile wird in diesem Kapitel nicht ausführlich eingegangen.



gangen. Es ist wichtiger auszuwerten, wie sich die Ergebnisse der gesamten Kette in relevanten Szenarien zeigen.

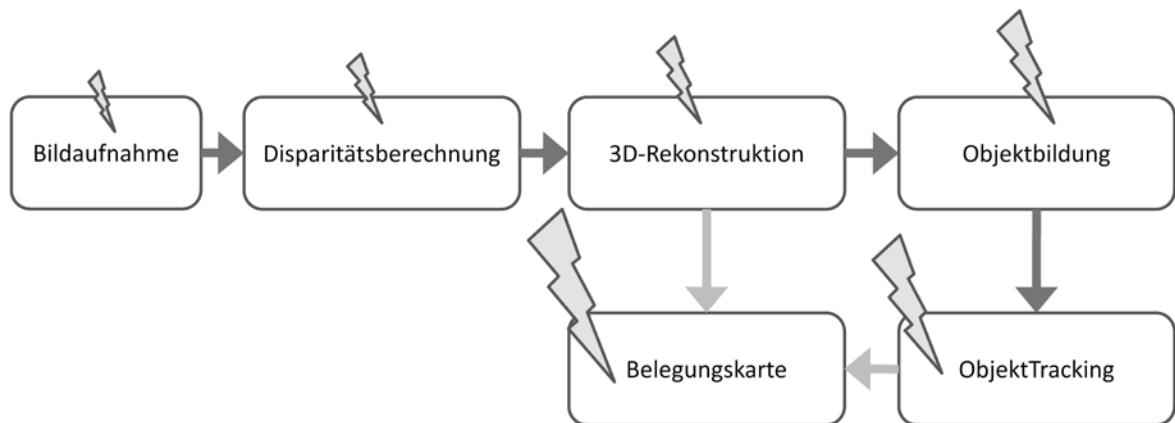


Abbildung 6-3: Hauptkomponenten der Algorithmen, bei denen Fehler auf treten können

Deshalb wird die Performance von ObjektTracking und Belegungskarte in folgenden relevanten Szenarien/Situationen untersucht:

- Stadtfahrt nur mit stationären Hindernissen
- Folgefahrt
- Stadtfahrt mit Gegenverkehr
- Rechts-/Links-Abbiegen
- Querverkehr
- Fußgängerszenarien

Zum Vergleich wird die Belegungskarte, die ohne Objekttracking auf Basis des Bayes-Theorems konstruiert wird, gezeigt. Zur Auswertung bestand nicht die Möglichkeit, alle im Umfeld aufgetretene Objekte bezüglich deren Positionen, Bewegungszustand sowie Abmessungen weltgetreu zu referenzieren. Deshalb können lediglich anhand der Bildinformation die Ergebnisse des Objekttracking und der Belegungskarte ausgewertet werden. Für das Objekttracking sind folgende Parameter vorhanden:

- Detektionsreichweite: bis 50m
- Detektionsrate: 20Hz
- Minimale Objektgröße: 0.2m x 0.2m x 0.2m

Für die Belegungskarte werden folgende Parameter festgelegt:

- Kartengröße: 80m x 80m
- Zellengröße: 0.2m x 0.2m
- Updaterate: 9Hz

Im Folgenden werden die Ergebnisse in den relevanten Szenarien dargestellt.

### 6.1 Untersuchte Szenarien

#### 6.1.1 Stationäre Umgebung

In Abbildung 6-4 wird ein städtisches Szenario dargestellt. Die obere linke Grafik repräsentiert das Kamerabild, die untere linke Grafik das 2D-Histogramm und die rechte Grafik die Belegungskarte.

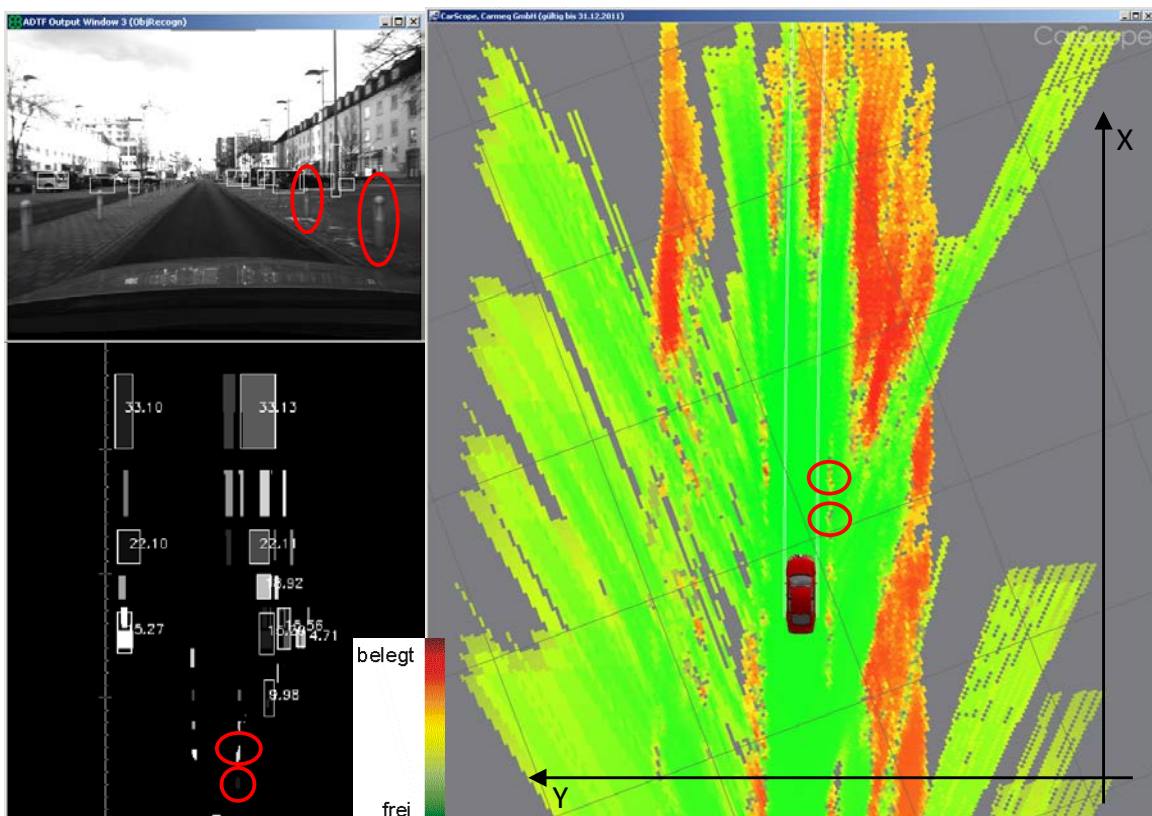


Abbildung 6-4: Innenstadt ohne bewegte Autos. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte von der Szene

Auf dem 2D-Histogramm stellen die weiß eingerahmten Kästen die segmentierten Objekte dar. Die longitudinale Entfernung jedes Objekt wird ebenfalls verzeichnet. Die rechte Gra-

fik stellt die Belegungskarte dar. Wie im vorherigen Kapitel beschrieben, repräsentieren grüne Zellen freie Bereiche, rote Zellen nicht befahrbare Bereiche und graue Zellen unbekannte Bereiche.

In Abbildung 6-4 ist zu erkennen, dass die Infrastruktur wie Häuser, Straßenlaternen, etc. bzw. parkende Autos gut in der Belegungskarte abgebildet wurden. Da kein bewegtes Objekt in der Szene vorhanden ist, wird vom Objekttracking kein Objekt ausgegeben. Ein kleines Problem stellen jedoch die Straßenpfosten dar, die im Bild rot eingekreist sind. Im 2D-Histogramm werden sie gut abgebildet.

Auf der Belegungskarte sind sie jedoch nicht eindeutig erkennbar. Dies ist darauf zurückzuführen, dass das Histogramm immer für ein Frame aufgebaut wird während die Belegungskarte mit der zeitlichen Integration der Messwerte funktioniert. Dies wurde im Kapitel 5 ausführlich erläutert. Da die Lokalisierung des Ego-Fahrzeugs auf der Karte nicht immer genau ist, sowie die Sensormessungen gewisse Abweichungen haben, werden auf der Karte die Pfosten nicht immer in den gleichen Zellen abgebildet. Dadurch sind die Zellen der Pfosten nicht eindeutig auf der Karte als belegt markiert. Zur Verbesserung der Ego-Lokalisierung können in Zukunft GPS-Daten verwendet werden. Im Allgemeinen lässt sich ausgehend von einer Fahrgeschwindigkeit bis 60 km/h aussagen, dass die Belegungskarte, die mit 9 Hz aktualisiert wird, die stationären Bereiche im Stadtverkehr gut darstellt.

### 6.1.2 Folgefahrt

Bei dieser Art von Verkehrssituationen sind sowohl bewegte Objekte als auch stationäre Hindernisse vorhanden. In Abbildung 6-5 befindet sich das Ego-Fahrzeug auf der linken Spur und überholt die auf der rechten Spur anhaltenden Fahrzeuge.

Das vorausfahrende Fahrzeug wurde als bewegtes Objekt vom Tracking detektiert. Der Bereich des bewegten Objektes wird mit der Farbe Lila markiert. Der blaue Pfeil zeigt die Bewegungsrichtung des Objektes. Die Pfeillänge deutet den Betrag der Geschwindigkeit an. Da die auf der rechten Spur befindlichen Fahrzeuge wegen einer roten Ampel anhalten, werden sie als stationäre Hindernisse ausgewiesen.

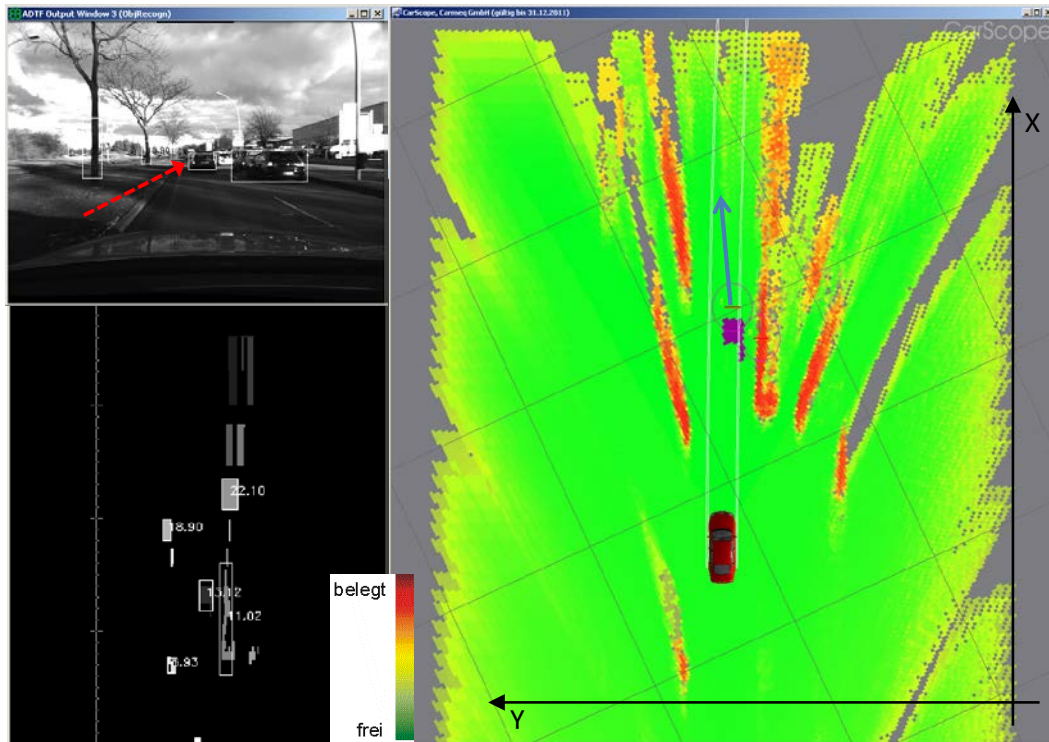


Abbildung 6-5: Folgefahrt ohne Gegenverkehr. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Unten rechts: die Belegungskarte wird mit dem Objekttracking unterstützt.

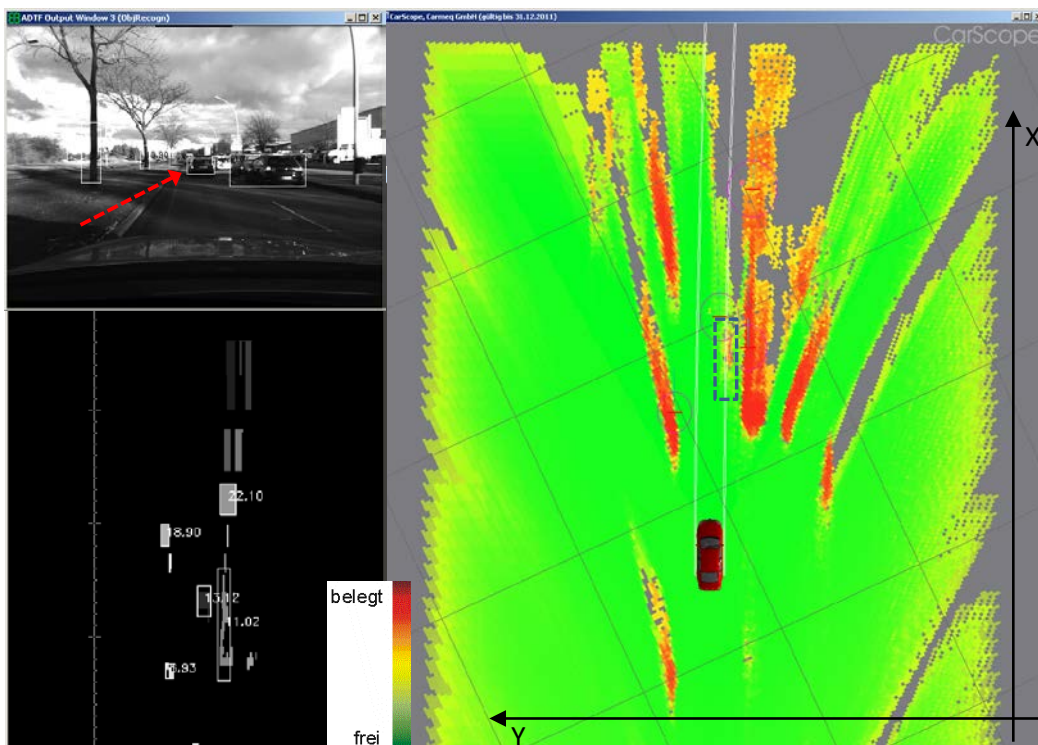


Abbildung 6-6: Folgefahrt ohne Tracking Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Unten rechts: die Belegungskarte in Betrieb ohne Objekttracking. Der Unterschied zu Abbildung 6-5 ist zu erkennen

Dies resultiert in einen roten Bereich auf der rechten Seite der eigenen Fahrspur. Anders sieht es aus, wenn die Belegungskarte ohne Objekttracking aufgebaut wird. Dies wird in Abbildung 6-6 veranschaulicht.

Es ist zu erkennen, dass an einer Stelle vom vorausfahrenden Fahrzeug ein Bereich mit einer gewissen Belegungswahrscheinlichkeit ist, obwohl in der Realität dieser Bereich frei ist. Dieser Bereich wird in einem blauen Kasten markiert. Da der Bereich nicht 100% belegt und auch nicht 100% frei ist, ist die Vorausberechnung einer möglichen Kollision sehr schwierig.

Damit lässt sich klar aussagen, dass mit Hilfe des Objekttrackings die Belegungskarte zuverlässiger und genauer aufgebaut wird. Das Objekttracking erfasst nicht nur ein Objekt, sondern auch mehrere bewegte Objekte in gleicher Fahrtrichtung wie im Folgenden gezeigt wird.

In Abbildung 6-7 befindet sich das Ego-Fahrzeug auf der rechten Fahrspur. Vorausfahrend sind mehrere Fahrzeuge. Zwei auf der linken Fahrspur, eines auf der rechten. Durch das Objekttracking konnten alle drei Fahrzeuge verfolgt werden. Dadurch werden Segmente, die zu bewegten Objekten gehören, nicht in die Karte eingetragen. Es ist auch zu erkennen dass die Bordsteinbereiche von links und rechts nicht als belegt dargestellt werden, da die Mindesthöhe einer Zelle, ab der sie als belegt klassifiziert wird, 20 cm beträgt.

In Abbildung 6-7 ist ersichtlich, dass die Belegung der Zellen im Fernbereich, z.B. ab 25 m, nicht immer eindeutig bestimmt werden kann. Unter 20 m decken die Zellen mit guter Genauigkeit reale Bereiche ab. Dies ist für sicherheitskritische Funktionen, z.B. automatische Notbremsung oder Lenkempfehlung, unter Berücksichtigung einer maximalen Fahrgeschwindigkeit von 50 km/h ausreichend [AKT10].

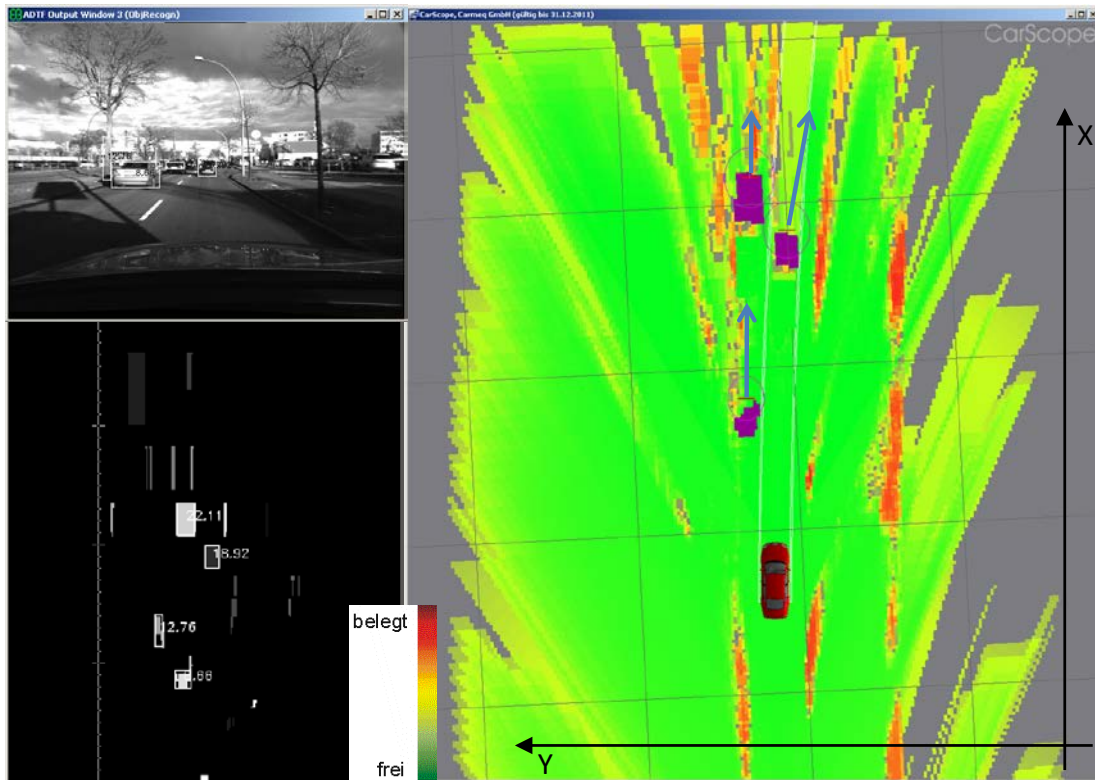


Abbildung 6-7: Folgefahrt mit Objekttracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte funktioniert beim Auftreten mehrerer bewegter Autos gut mit Hilfe des Objekttrackings.

### 6.1.3 Gegenverkehr

Die Situation mit Gegenverkehr stellt auch einen wichtigen Punkt bei der Auswertung dar. Aufgrund des großen Geschwindigkeitsunterschiedes (im urbanen Umfeld typisch  $2 \times 50 \text{ km/h} = 100 \text{ km/h}$ ) ist es wichtig, den Gegenverkehr zuverlässig zu erkennen bzw. zu verfolgen und anhand des Objekttrackings sowie auf der Belegungskarte korrekt abzubilden.

In Abbildung 6-8 wird gezeigt, dass der PKW in der Gegenrichtung erfasst und erfolgreich verfolgt wird. Die Bewegungsrichtung des PKWs wird durch den blauen Pfeil angezeigt. Die restlichen stationären Hindernisse sowie Infrastrukturen können durch die Karte gut abgebildet werden. Als Vergleich wird für das gleiche Szenario die Belegungskarte ohne Objekttracking aufgebaut (siehe Abbildung 6-9).

An der Stelle, wo sich das bewegte Objekt befindet, ist ein Bereich, bei dem die Befahrbarkeit unklar ist. Für die Kollisionsschätzung ist dies nachteilig. Mit dem Ergebnis aus Abbildung 6-8 können mögliche Kollisionen sicher berechnet werden.

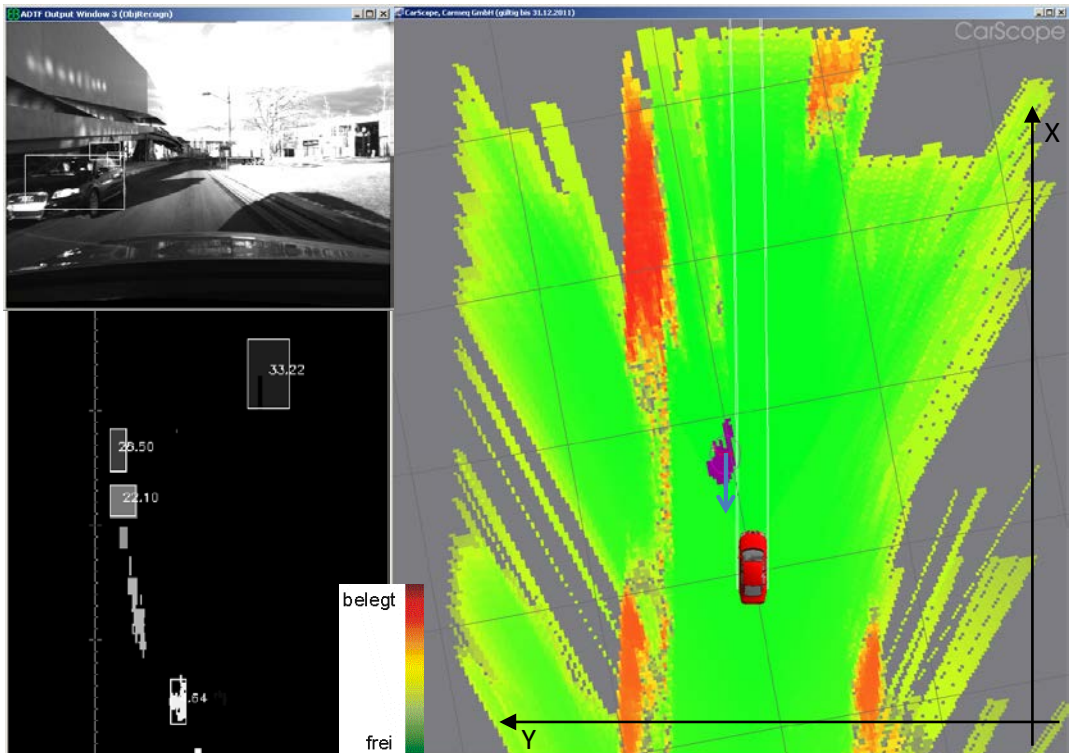


Abbildung 6-8: Gegenverkehr mit ObjektTracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte mit dem erkannten bewegten Objekt.

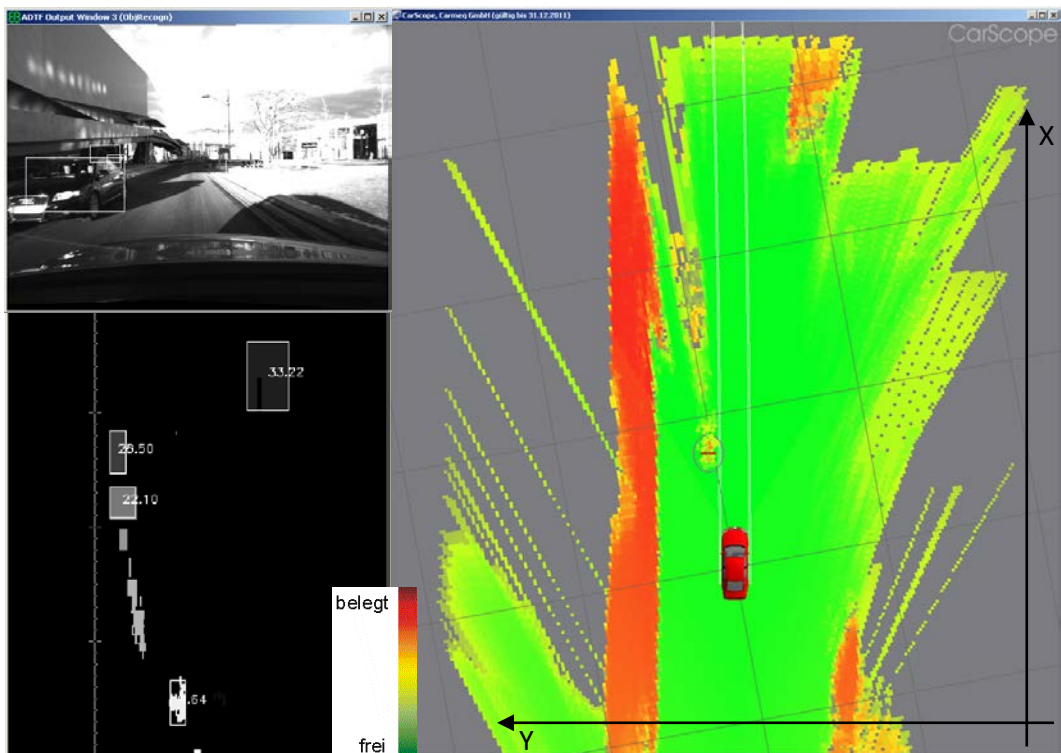


Abbildung 6-9: Gegenverkehr ohne ObjektTracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte zeigt, dass der Bereich, in dem das bewegte Objekt sich befindet, weder eindeutig frei noch belegt ist.

### 6.1.4 Folgefahrt und Gegenverkehr

Im Stadtverkehr kommt es sehr häufig vor, dass auf einer Straße ohne Begrenzung in der Mitte in beide Richtungen gefahren wird. Da ist es für die Umfeldwahrnehmung notwendig, sowohl Gegenverkehr als auch vorausfahrenden Verkehr gut zu erkennen. In Abbildung 6-10 wird gezeigt, dass dies anhand des Objekttracking gut realisiert wurde. Das Tracking ist in der Lage, für zwei sich in Gegenrichtung bewegende Objekte zwei unterschiedliche Bewegungsmodelle zu zuordnen. Die Assoziation der Messung zum Tracking funktioniert ebenfalls gut, so dass die Objekte konstant über die gesamte Beobachtungszeit verfolgt werden können. Des Weiteren kann durch die Erkennung von bewegten Objekten die Belegungskarte sauber aufgebaut werden.

Die befahrbaren und nicht befahrbaren Bereiche können damit eindeutig beschrieben werden.

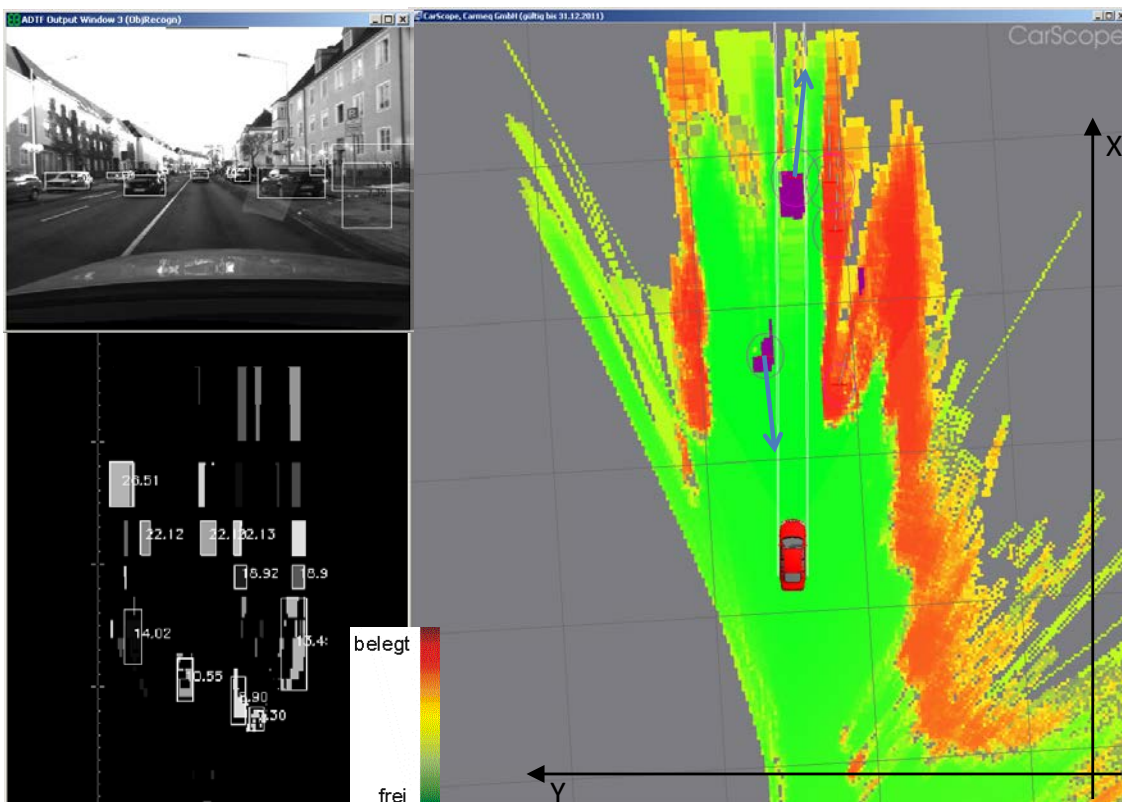


Abbildung 6-10: Folgefahrt mit Gegenverkehr. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Man sieht, dass Objekttracking funktioniert gut sowohl für Gegenverkehr als auch für den vorausfahrenden Verkehr.



### 6.1.5 Abbiegen

Ein weiteres wichtiges Szenario im innenstädtischen Verkehr ist das Abbiegen. In diesem Fall spielt der Öffnungswinkel des Kamerasensors eine wichtigere Rolle als die Reichweite, da seitliche Objekte oder Hindernisse von der Seite schnell detektiert werden sollen. In Abbildung 6-11 wird dies veranschaulicht. Es handelt sich dabei um ein Abbiegszenario nach rechts. Auf der rechten Seite der Belegungskarte ist ein grauer, also unbekannter, Bereich vorhanden, da das Ego-Fahrzeug quasi um eine Achse nach rechts dreht. Mit einer Abtastrate von 9 Hz lässt sich die Belegungskarte gut abbilden. Auf der Belegungskarte ist ein voraus berechneter Eigenfahrtschlauch dargestellt.

Der geschätzte Fahrtschlauch wird in der Regel für die Kollisionsprädiktion verwendet. Man sieht, dass alle Zellen im Fahrtschlauch bis zu 17 m klar detektiert wurden. Damit können unter Berücksichtigung der Fahrgeschwindigkeit beim Abbiegen mögliche Kollisionen mit Echtzeit vorausberechnet werden.

Des Weiteren werden auf der Karte neben dem abbiegenden Fahrzeug andere Fahrzeuge als stationäre Hindernisse angezeigt, da sie wegen der roten Ampel anhalten. Auch in dieser Situation ist es wichtig, das vorausfahrende abbiegende Fahrzeug als bewegtes Objekt zu identifizieren.

In Abbildung 6-12 wird die Belegungskarte ohne Tracking aufgebaut. Dadurch ist der Bereich, in dem eigentlich das bewegte Objekt ist, mit unsicheren Informationen gefüllt. Die Kollisionsberechnung wird dadurch eindeutig erschwert.

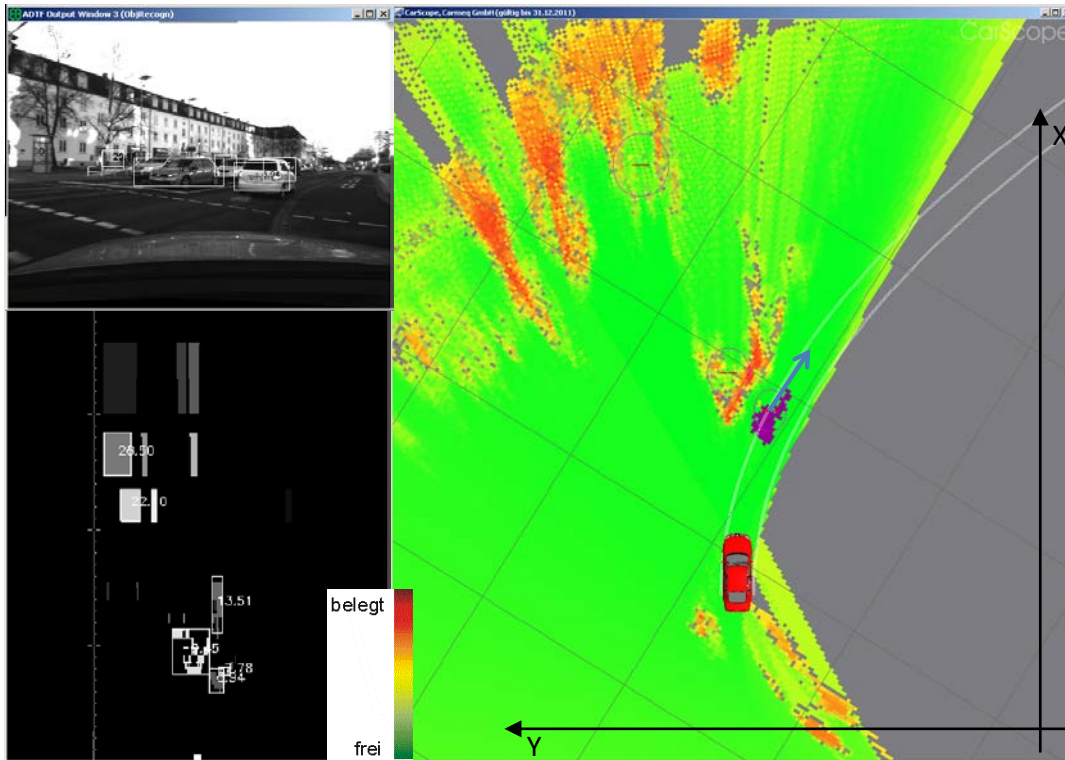


Abbildung 6-11: Rechts abbiegen mit Objekttracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte von der Szene. Man sieht, dass das Objekttracking auch für abbiegenden Verkehr funktioniert.

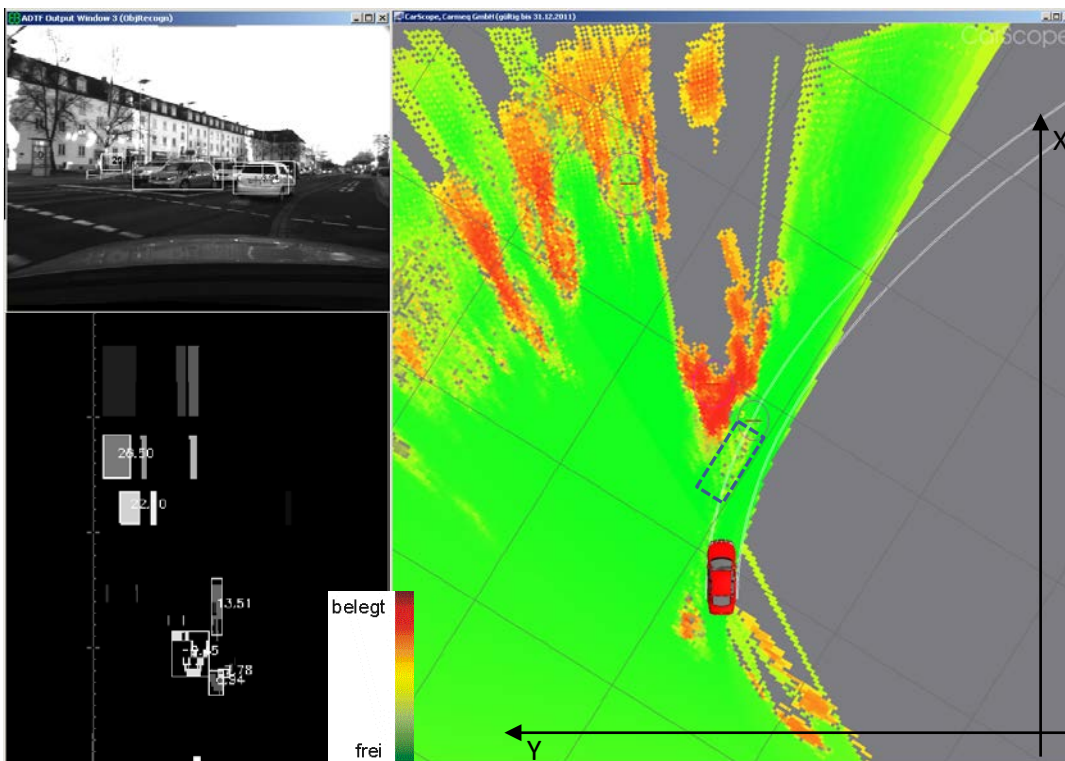


Abbildung 6-12: Rechts abbiegen (ohne ObjektTracking). Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte von der Szene

### 6.1.6 Querverkehr

Ein weiteres relevantes Szenario stellt die Erkennung des Querverkehrs dar. Dabei ist es wichtig, den Querverkehr möglichst früh und genau zu erfassen. Mit einem Öffnungswinkel von  $63^\circ$  ist die Kamera in Stadtverkehr im der Lage, die zur Fahrtrichtung querenden Fahrzeuge genügend früh zu detektieren.

Ein Beispiel wird in Abbildung 6-13 gezeigt. Die detektierte Bewegungsrichtung des PKWs zeigt, dass es sich um Querverkehr handelt. Diese Information aus dem Objekttracking wird in der Karte benutzt um den Bereich des querenden PKWs zu lokalisieren.

Die Querstraße wird zum besseren Verständnis der Szene auf der Karte gekennzeichnet. Zum Vergleich wird in Abbildung 6-14 das gleiche Szenario ohne Tracking dargestellt. Es ist zu erkennen, dass die Stelle, an der das bewegte Objekt ist, als frei angezeigt wird. Dies lässt sich darauf zurückführen, dass der querende PKW sich wegen seiner relativ hohen Fahrgeschwindigkeit nur für eine Messung an dieser Stelle befindet. Dadurch überwiegen die bisherigen Messungen, bei denen der Bereich als frei detektiert wurde.

Für die Kollisionsprädiktion wäre das Ergebnis aus Abbildung 6-14 nicht ausreichend. Da zwar die Belegungskarte an sich im Vergleich zur Abbildung 6-13 nahezu identisch ist, beeinträchtigt das Fehlen der Informationen von bewegten Objekten die genaue Prädiktion.

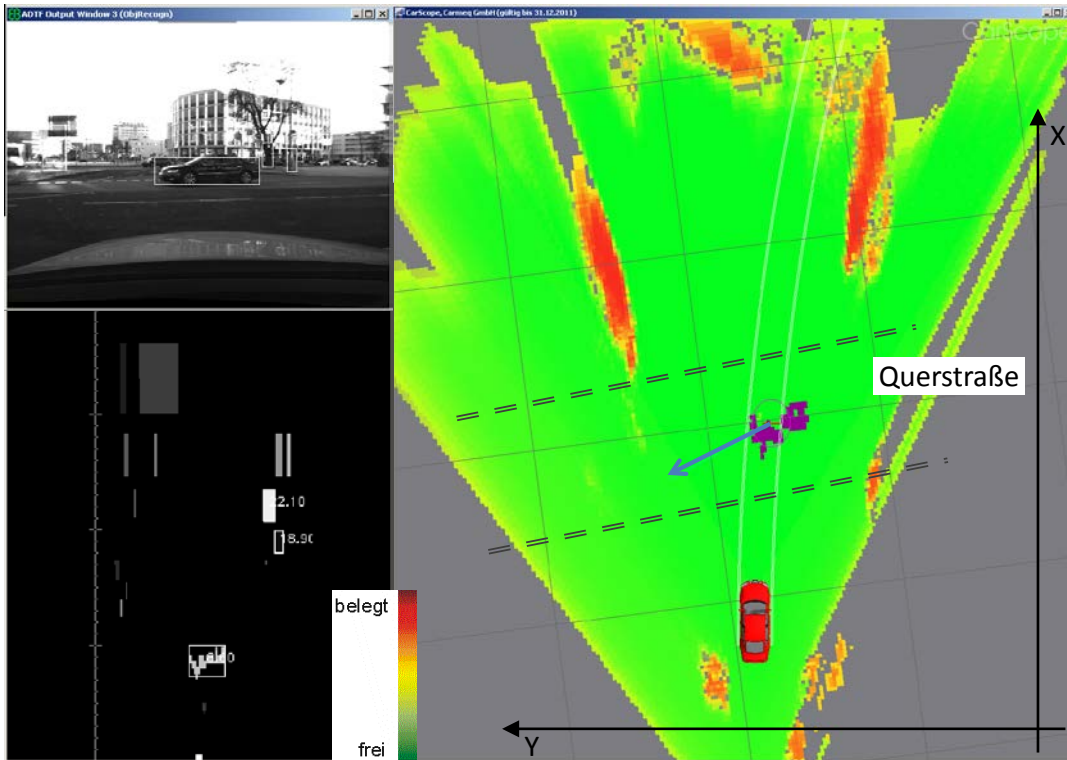


Abbildung 6-13: Querverkehr mit ObjektTracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte von der Szene.

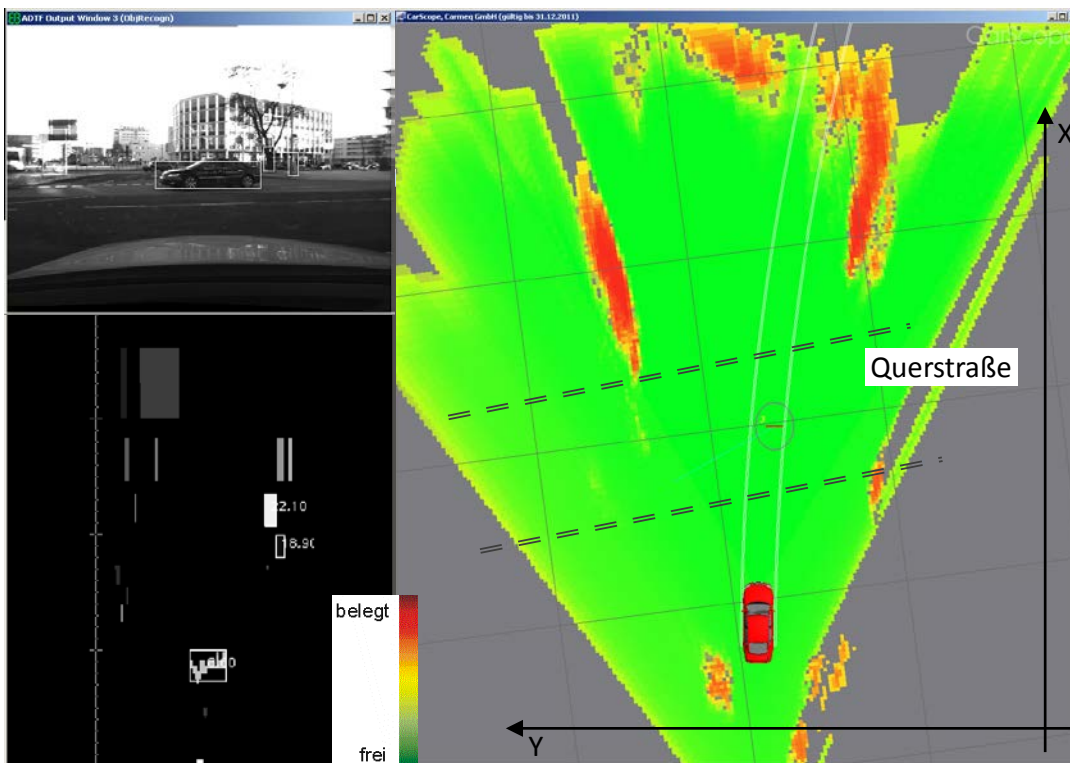


Abbildung 6-14: Querverkehr ohne Tracking. Oben links: das Kamerabild. Unten links: das 2D-Histogramm, auf dem die Objektdetektion durchgeführt wird. Rechts: die Belegungskarte von der Szene

## 6.2 Analyse

In diesem Abschnitt werden die erzielten Ergebnisse der Objektdetektion, des Objekttracking und der Belegungskarte analysiert. Wie schon erwähnt, stehen für die Fahrten in realen Umgebungen keine 100% Groundtruth-Daten zur Verfügung. Die Ergebnisse werden ausschließlich anhand der Bilddaten ausgewertet. In der Statistik wird die Fehlerrate zur Auswertung in vier Kategorien eingeteilt: False Negative, False Positive, True Negative, True Positive.

True Positive Rate entspricht dem Anteil der korrekt als „bewegt“ detektierten Objekte an der Gesamtheit der tatsächlich bewegten Objekte.

False Negative Rate ist Anteil der fälschlicherweise als „nicht-bewegt“ detektierten Objekten an der Gesamtheit der tatsächlich bewegten Objekte

True Negative Rate bedeutet den Anteil der korrekt als „nicht-bewegt“ detektierten Objekte an der Gesamtheit der in Wirklichkeit stationärer Objekte.

False Positive Rate ist letztendlich der Anteil der fälschlicherweise als „bewegt“ detektierten Objekte an der Gesamtheit stationärer Hindernisse.

Wie in Kapitel 4 beschrieben, beinhaltet die Objektbildung die Segmentierung zusammengehöriger 3D-Punkte, die Schätzung der Höhe, Länge und Breite bzw. Position des Segments, das als Objekt bezeichnet wird. Ein Faktor, der die Genauigkeit der Segmentierung beeinflusst, ist unter anderem die Unsicherheit in der 3D-Rekonstruktion. Diese Unsicherheit führt dazu, dass nahe beieinander befindliche Objekte zu einem großen Objekt zusammengefügt werden, oder dass ein großes Objekt als zwei kleine Objekte dargestellt wird. In Abbildung 6-15 wird dieser Effekt veranschaulicht. Auf der linken Seite ist zu erkennen, dass der lange Bus als zwei Objekte dargestellt wird, die hintereinander in einer Reihe stehen.

Dies ist darauf zurückzuführen, dass in der Mitte des Busses keine 3D-Punkte gemessen werden konnten, wodurch die kleine „Lücke“ in dem 2D-Histogramm verursacht wird. Dadurch haben die Segmentierungsalgorithmen die zwei Blöcke als separate Objekte klassifi-

ziert. Auf der anderen Seite werden Autos, die nahe hintereinander fahren als ein Objekt dargestellt. Die 3D-Punkte der Autos sind so nah zueinander, dass alles in einen Block zusammengefasst wird. Eine quantitative Aussage, wie häufig solche Fehler auftreten, ist im Diagramm in Abbildung 6-16 dargestellt. In der Regel kommt es häufiger vor, dass Objekte zusammengefügt werden.

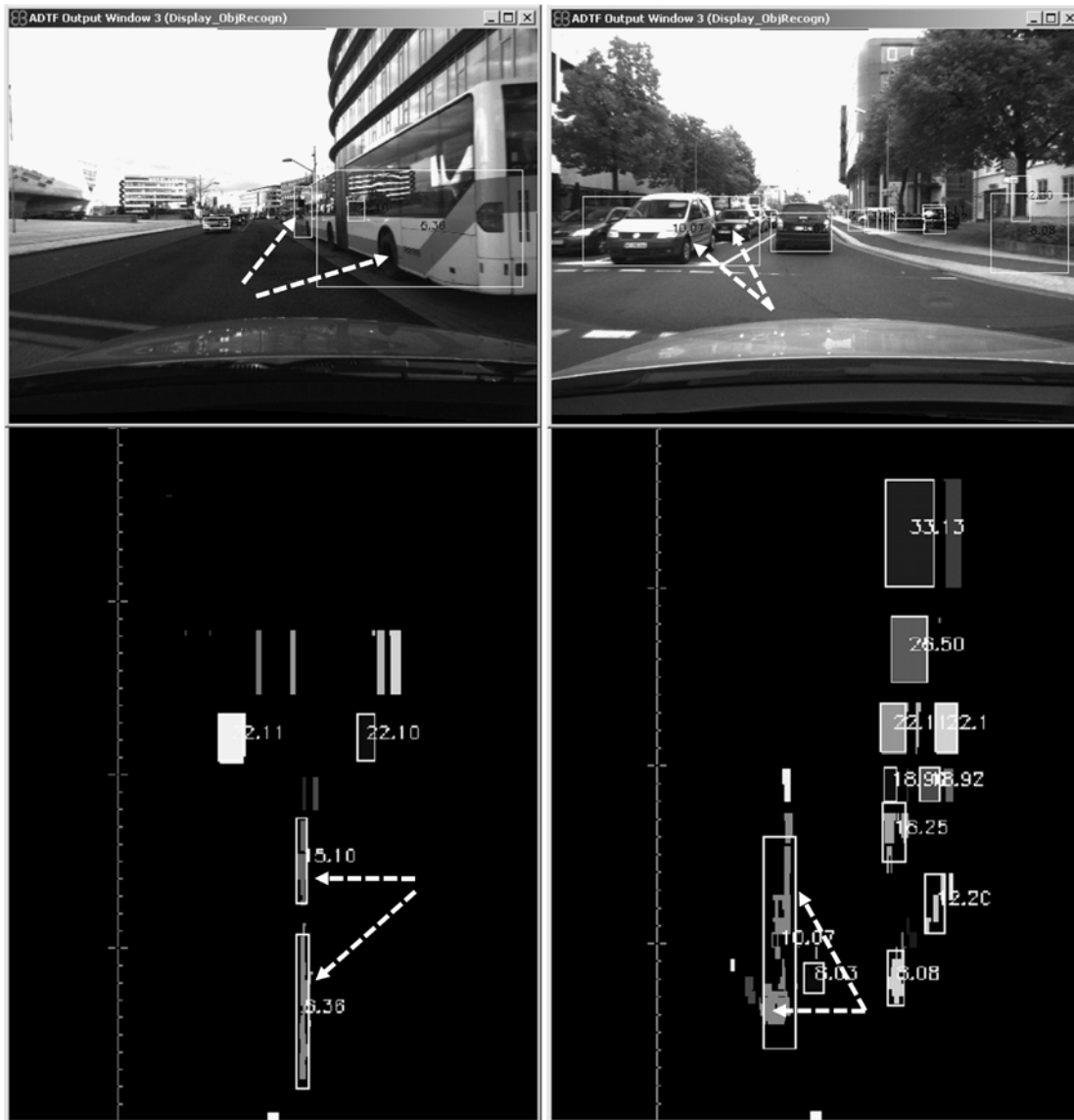


Abbildung 6-15: typische Fehler der Segmentierung. Linke Seite: ein großes Objekt wird als zwei Objekte detektiert. Rechte Seite: zwei Objekte werden als ein Objekt zusammengefügt.

In Abbildung 6-16 repräsentiert die Abszissenachse den Abstand zwischen den Objekten. Die Rate der korrekt erkannten Objekte ist abhängig vom Abstand zwischen den Objekten und wird durch die Ordinatenachse repräsentiert. Die rote Kurve zeigt, dass bei einer Entfernung bis 5m die Detektionsrate mit steigendem Abstand zwischen Objekten besser wird.

Bei 3m-Abstand liegt die Detektionsrate bei 80%. Je weiter die Objekte auseinander sind, desto höher ist die Detektionsrate. Die grüne und blaue Kurve zeigen ein ähnliches Verhalten bei 10m- und 15m-Entfernung.

Es ist festzuhalten, dass bei zunehmender Entfernung die Fähigkeit, Objekte voneinander zu trennen, abnimmt. Dies ist auf die zunehmende Ungenauigkeit in der 3D-Schätzung bei größeren Entfernungen zurückzuführen. Fehler in der Objektdetektion verursachen dementsprechend Unsicherheiten im Objekttracking. Wichtig ist die Wirkung in Verfolgung von bewegten Objekten. Dies wird im folgenden Abschnitt ausführlicher analysiert.

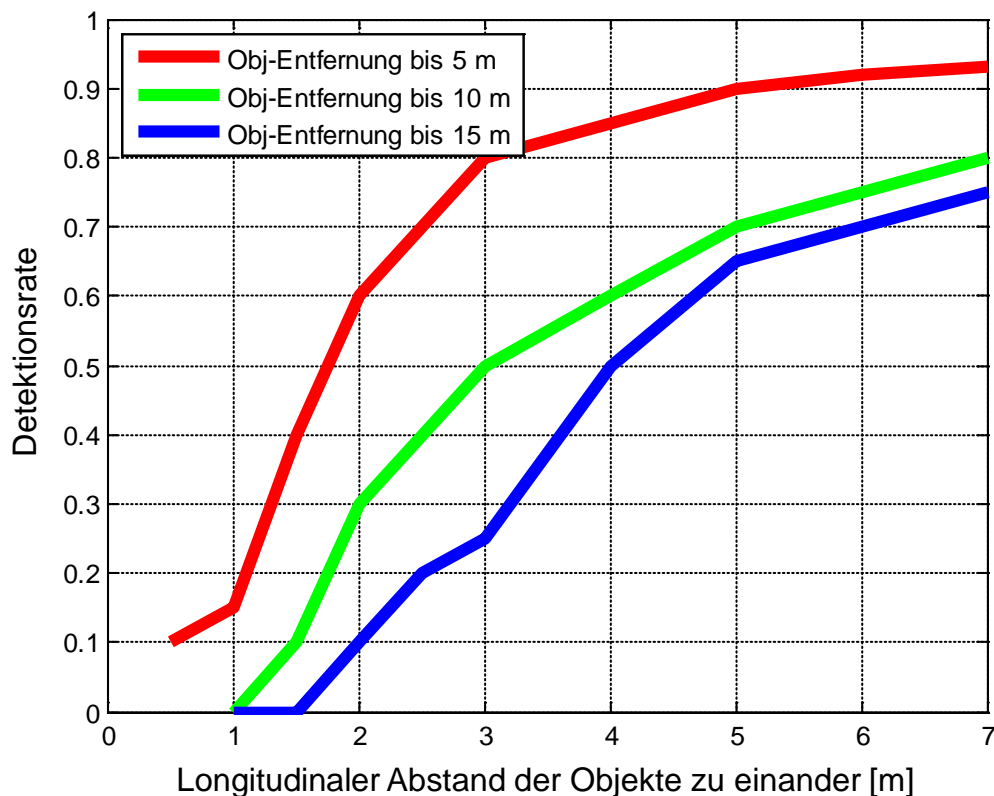


Abbildung 6-16: Fehlerrate der Objektdetektion bei unterschiedlichen Entfernungen

### 6.2.1 Verfolgung von bewegten Objekten

Für die Erkennung von bewegten Objekten hat die Ungenauigkeit in der Objektdetektion die Konsequenz, dass für zwei bewegte Objekte nur ein Track (Objekthypothese) initialisiert wird, was zu Inkonsistenzen im Laufe der Track-Verwaltung führt. Die Grundidee der Erkennung und Verfolgung von bewegten Objekten besteht darin, dass ein stabiler Track,

welcher ein Objekt in der Realität repräsentiert, eine Mindestgeschwindigkeit haben muss, um als bewegtes Objekt klassifiziert zu werden. Aufgrund von Unsicherheiten in der Geschwindigkeitsschätzung, ist folgendes Verhalten bei der Erkennung von bewegten Objekten beobachtet worden:

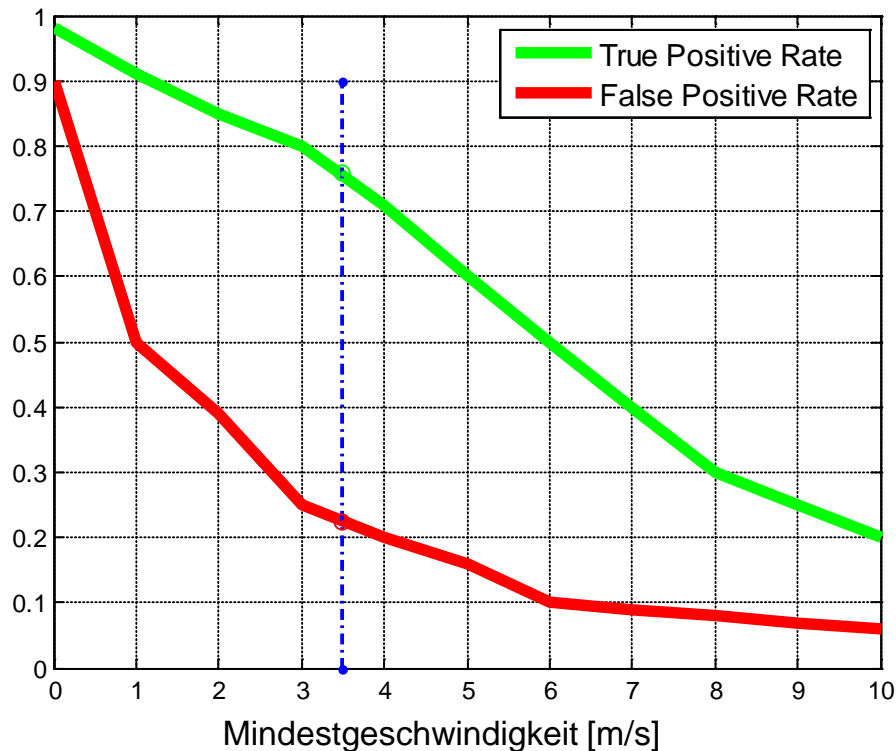


Abbildung 6-17: Abhängigkeit der True-Positive und False-Positive-Rate von einer vorgegebenen Mindestgeschwindigkeit

In Abbildung 6-17 stellt die grüne Kurve die Detektionsrate von wirklich bewegten Objekten auf Basis des Vergleichs der getrackten Geschwindigkeit mit der vorgegebenen Mindestgeschwindigkeit dar. Man sieht, je niedriger die Mindestgeschwindigkeit ist, desto höher ist die Detektionsrate von bewegten Objekten. Gleichzeitig ist die False-Positive-Rate, also die Rate der falschen Erkennung, bei der stationäre Objekte als bewegt klassifiziert werden, auch höher. Dies ist auf die Ungenauigkeit in der Geschwindigkeitsschätzung zurückzuführen. Die Geschwindigkeitsschätzung basiert sowohl auf der 3D-Position-Messung des Objektes als auch auf der Schätzung der Eigenbewegung. Die bekannte Ungenauigkeit in diesen zwei Prozessen trägt dazu bei, dass ein stationäres Objekt eine kleine Geschwindigkeit hat. Wenn die Mindestgeschwindigkeit höher gesetzt wird, dann sinken die False-



Positive-Rate, zugleich aber auch die True-Positive-Rate. Eine Mindestgeschwindigkeit von 3m/s stellt einen guten Kompromiss dar. Damit liegt die True Positive Rate auf fast 80% während die False Positive Rate bei ca. 20% liegen. Bei Betrachtungen über mehrere Frames werden die als falsch erkannten bewegten Objekte jedoch schnell herausgefiltert. Dies bedeutet, dass ein stationäres Objekt in der Regel nur für ein Frame als bewegtes Objekt klassifiziert wird.

### 6.2.2 Verbesserung der Belegungskarte durch Objekttracking

Es ist in Einzelfällen möglich, dass ein stationäres Hindernis für ein, zwei Frames als ein bewegtes Objekt klassifiziert wird. In diesem Fall wird die dem Objekt zugehörige Messung nicht in der Belegungskarte berücksichtigt. Da die Karte mit 9Hz aktualisiert wird, hat die Fehlerkennung in ein, zwei Frames keinen erheblichen Einfluss auf die gesamte Qualität der Karte.

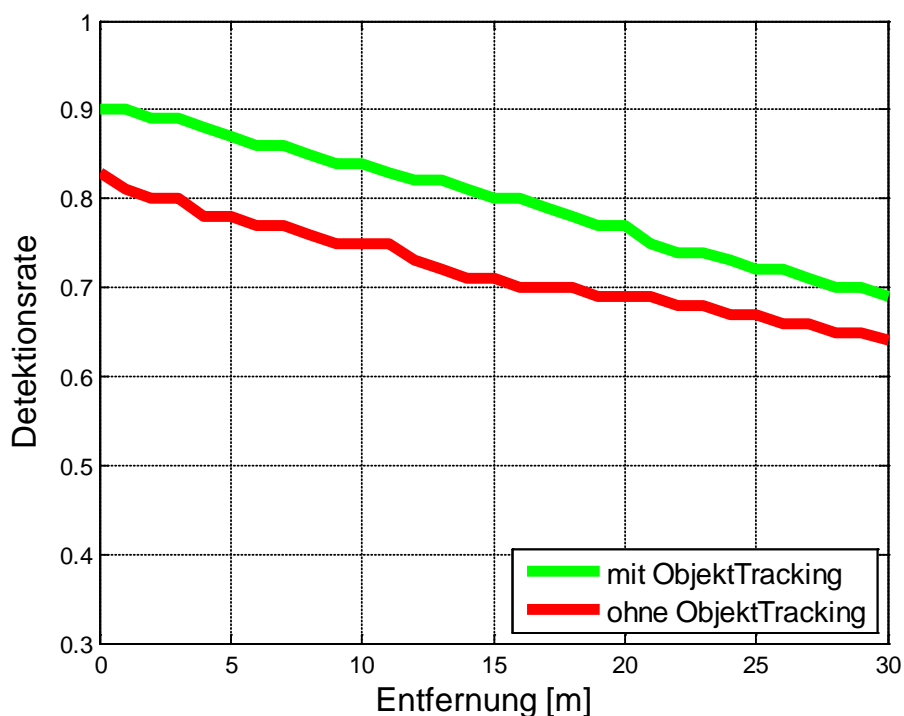


Abbildung 6-18: Vergleich der Belegungskarte mit und ohne ObjektTracking –(nur ein bewegtes Objekt im Umfeld)

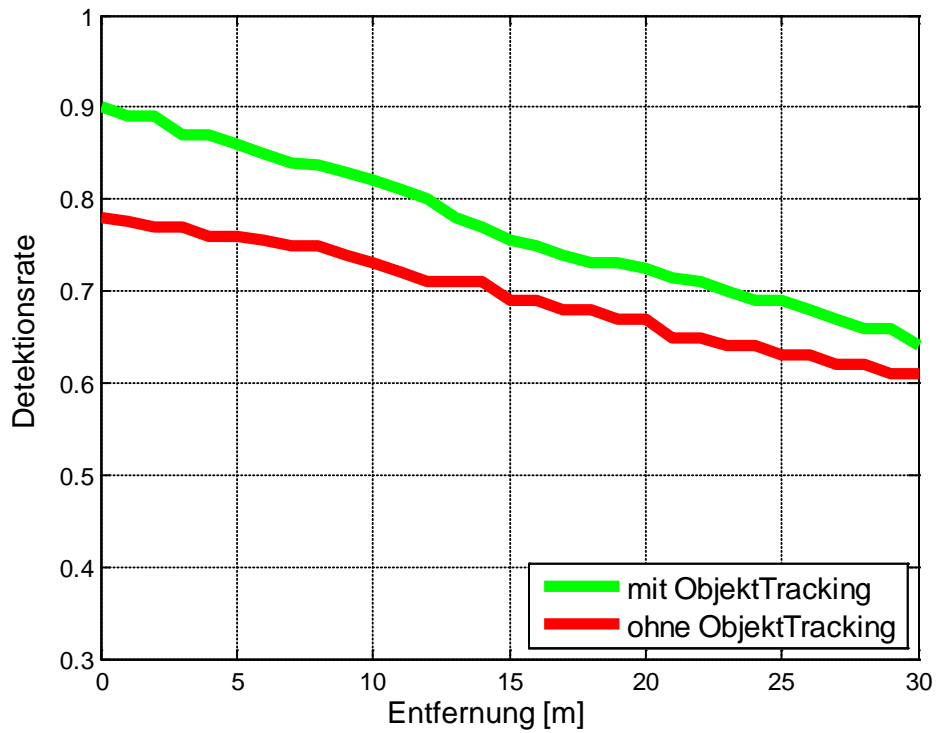


Abbildung 6-19: Vergleich Belegungskarte mit und ohne ObjektTracking-(zwei bewegte Objekte im Umfeld)

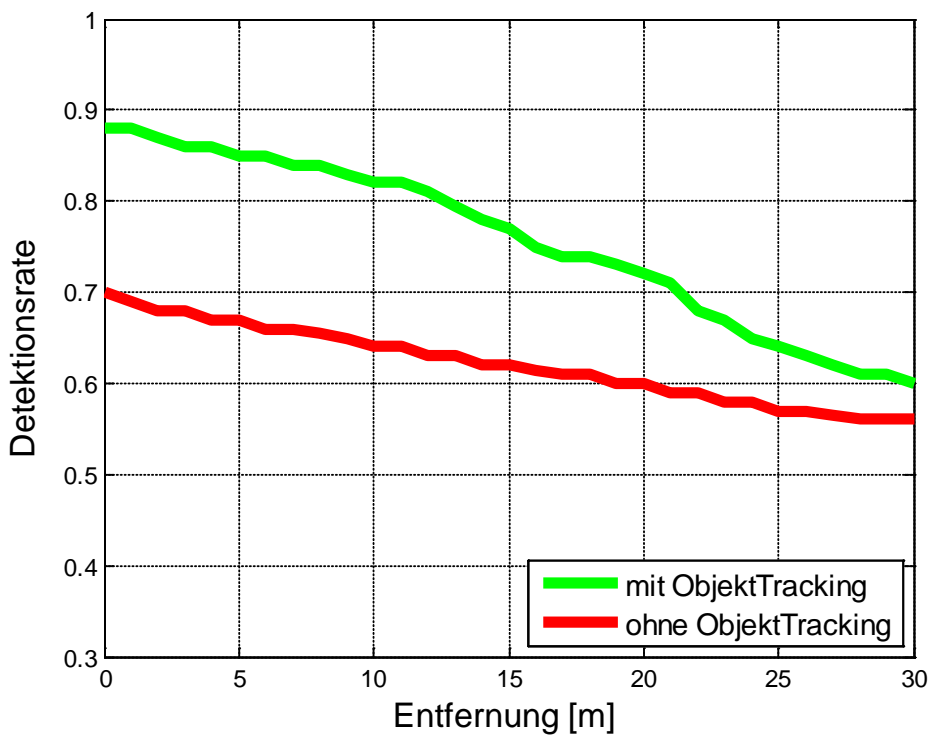


Abbildung 6-20: Vergleich Belegungskarte mit und ohne ObjektTracking (drei Objekte)

In Fällen wo das Objekttracking das bewegte Objekt nicht erkennt, besonders bei Verfolgungsfahrten, hat die Karte in den Regionen des bewegten Objekts eine merkbare Ungenauigkeit.

In Abbildung 6-18, Abbildung 6-19 und Abbildung 6-20 werden zusammengefasste Evaluierungen der Belegungskarte dargestellt. Die durchgezogenen Kurven stellen die Detektionsrate mit Hilfe des Objekttracking dar. Die roten Kurven stellen im Gegensatz dazu die Detektionsrate ohne Objekttracking dar. In Abbildung 6-18 handelt sich um Situationen, in denen ein bewegtes Objekt im Umfeld vorhanden ist. Es ist zu erkennen, dass mit Hilfe des Objekttrackings die Richtigkeit der Karte ca. um 8% verbessert wird. Mit der Entfernung nimmt die Richtigkeit der Karte ab. Im Fernbereich (ab 25 m) ist der Unterschied zwischen mit und ohne Tracking nicht mehr so groß wie im Nahbereich (unter 10m). Dies ist durch die Ungenauigkeit beim Objekttracking im Fernbereich zu erklären.

In Abbildung 6-19 handelt es sich um Szenarien, bei denen zwei bewegte Objekte vorhanden sind. Auch hier kann festgestellt werden, dass mit Hilfe des Objekttrackings die Richtigkeit der Karte verbessert wird. Im Nahbereich (unter 10m) ist eine Verbesserung von min. 12% zu erkennen. Der Unterschied im Vergleich zum Ein-Objekt-Fall (8%) ist damit zu erklären, dass zwei bewegte Objekte mehr Inkonsistenz auf der Karte verursachen als ein bewegtes Objekt. Deshalb kann durch Einsatz des Objekttrackings eine größere Verbesserung auf der Karte erreicht werden. Dieser Trend setzt sich im Fall von drei bewegten Objekten fort. Da liegt der Verbesserungsgrad bei ca. 20% im Nahbereich. Im Fernbereich ist der Unterschied im Allgemeinen zwischen den Karten mit und ohne Tracking geringer. Im Endeffekt kann festgestellt werden, dass die Qualität der Belegungskarte mit Hilfe des Objekttrackings erheblich verbessert wird, besonders wenn mehrere bewegte Objekte auftreten.

Durch den Einsatz des in dieser Arbeit beschriebenen Verfahrens werden bei der Kollisionsschätzung verbesserte Ergebnisse erreicht. Dies wird in Abbildung 6-21 veranschaulicht.

Die blauen Kästen in Abbildung 6-21 unten links sind Objekte, die mit Tracking detektiert wurden. Es ist zu erkennen, dass die geparkten Autos und Infrastruktur, im Bild im Kreis

hervorgehoben, mit den Objektkasten sich schwer darstellen lassen. Die Objektmodelle können die Vielfalt der Umgebung nicht ausreichend darstellen. Des Weiteren werden bei Messung zur Messung unterschiedliche Objekte nach dem Segmentierungsalgorithmus im Abschnitt 4.1 erzeugt. Dies hat für die Kollisionsberechnung nach [AET10] erhebliche Nachteile.

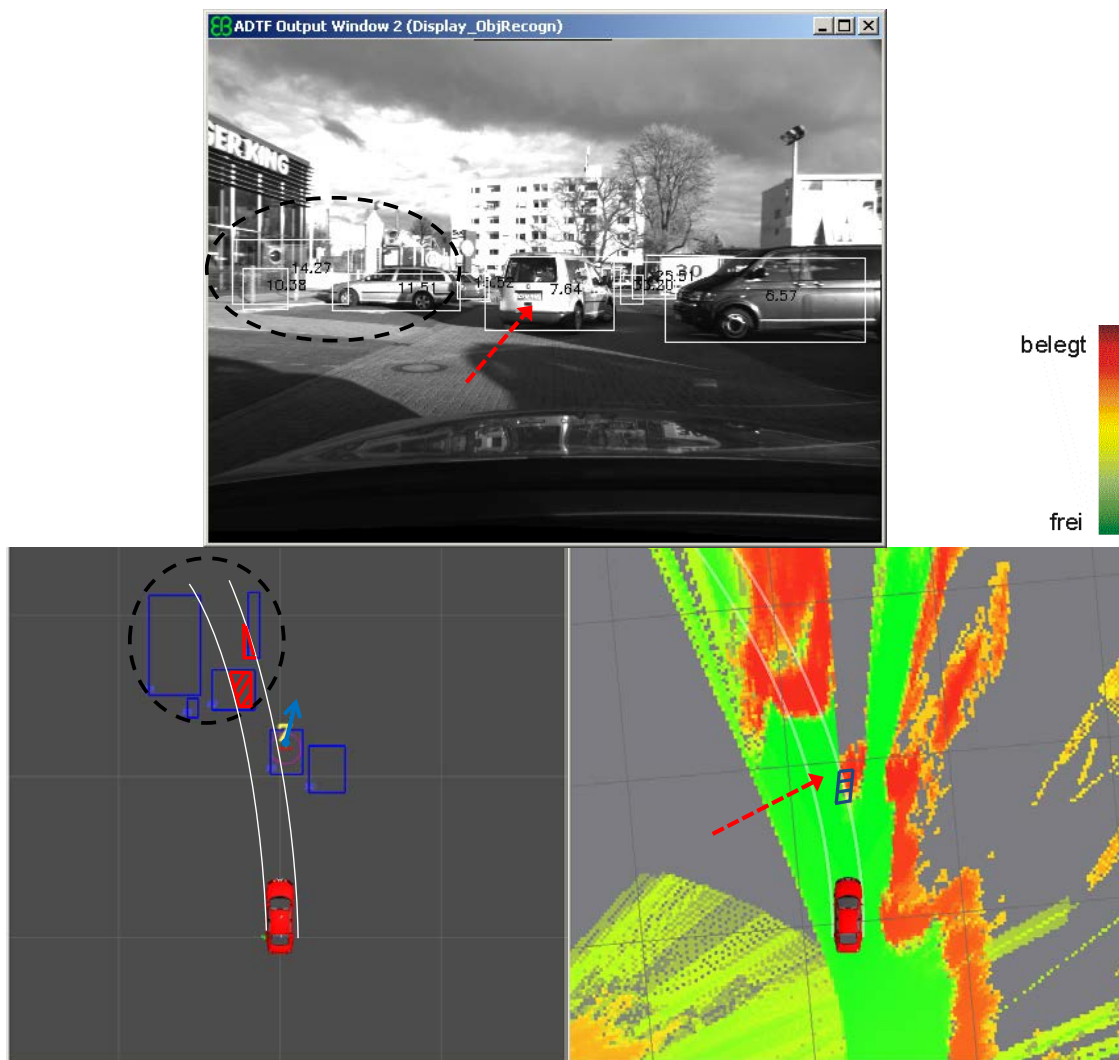


Abbildung 6-21: Kollisionsprädiktion. Oben: das Kamerabild von der Szene. Das bewegte Objekt wird mit einem roten Pfeil gekennzeichnet. Unten links: prädizierte Kollisionsregionen auf Basis von Objekten. Unten rechts: prädizierte Kollisionsregion auf Basis von Belegungskarte.

Die Kollision wird auf Basis der Überschneidungsfläche zwischen prädiziertem Fahr-schlauch und Objekt-abmessungen prädiziert. Wenn die Objekte nicht kontant sind, ist die Kollisionsprädiktion dementsprechend unzuverlässig. Weiterhin kann über Bereiche, in denen sich kein Objekt befindet, keine eindeutige Aussage getroffen werden, ob diese Bereiche

frei und befahrbar sind. Die objektbasierte Umfeldmodellierung hat diese Beschreibungsmöglichkeit nicht.

Auf der anderen Seite erkennt man in Abbildung 6-21 unten rechts, dass auf der Belegungskarte das bewegte Objekt als Kollisionspotential betrachtet wird, da die Geschwindigkeitsinformation in den Zellen nicht modelliert wird. Es handelt sich um die Belegungskarte, die noch nicht mit Objekttracking verbessert wurde.

Durch den neuen Ansatz werden in Abbildung 6-22 folgende Ergebnisse erreicht:

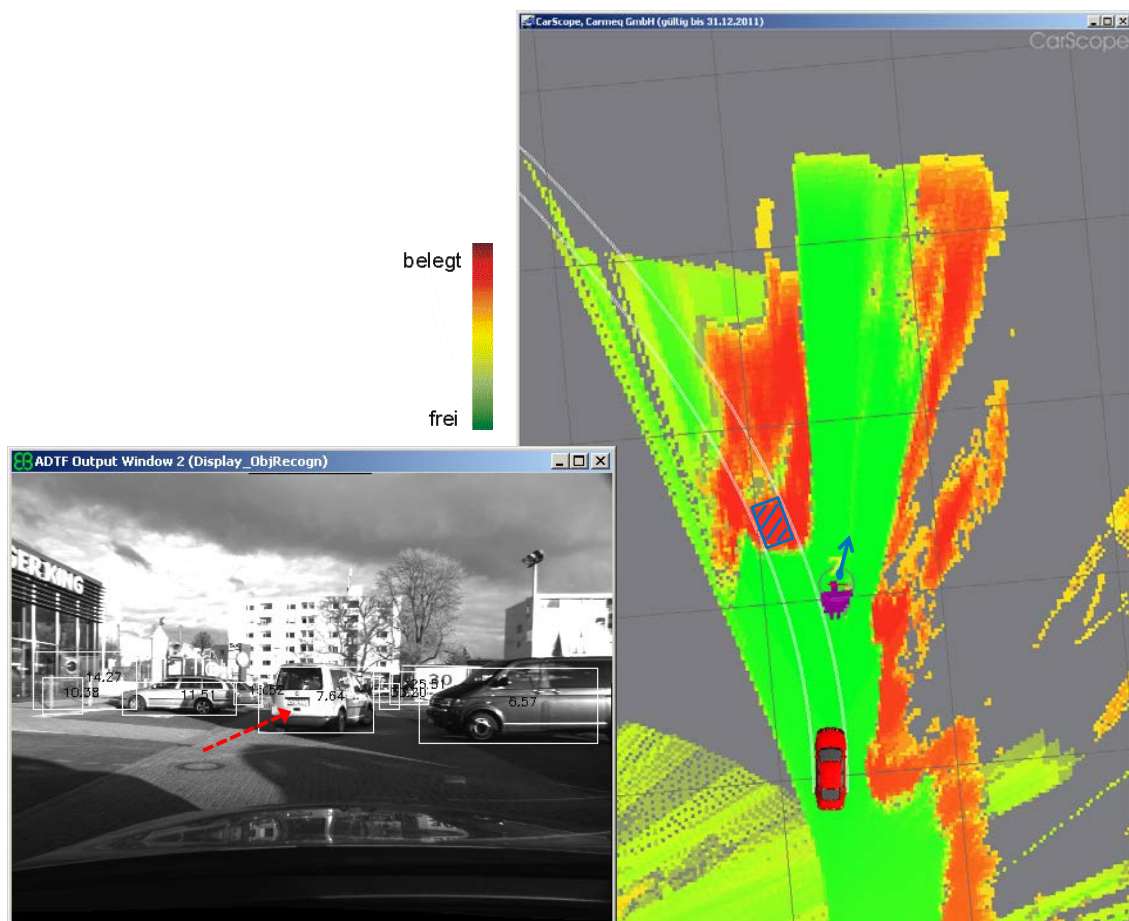


Abbildung 6-22: Links: das Kamerabild von der Szene. Das bewegte Objekt wird mit einem roten Pfeil gekennzeichnet. Rechts: die Kollisionsprädiktion ist im Vergleich zu beiden Fällen in Abbildung 6-21 zuverlässiger.

In Abbildung 6-22 wird gezeigt, dass die Kollisionsprädiktion auf Basis von bewegten Objekten und der mit Objekttracking unterstützten Belegungskarte am besten funktioniert. Mit den Erkenntnissen der Objektbewegung kann der künftige Aufenthaltsraum des Objektes berechnet werden. Des Weiteren lassen sich mögliche Kollisionen mit stationären Hin-

dernissen zuverlässig auf der Belegungskarte ermitteln. Die freien Bereiche können mit der Karte ebenfalls sicher gestellt werden.

Der neue Ansatz wurde im Projekt AKTIV-SFR eingesetzt und erprobt [AKT10]. Die Maßnahmen zur Erhöhung der Sicherheit für Fußgänger wurden erfolgreich erprobt.

# Kapitel 7

## Schlussbetrachtungen

### 7.1 Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Umfeldwahrnehmung auf Basis eines Stereokamera-Systems in Fahrerassistenzsystemen. Die Motivation für die Arbeit liegt darin, die Sicherheit für Fußgänger und Radfahrer im Stadtverkehr durch neue Entwicklungen in Fahrerassistenzsystemen zu erhöhen. Gefördert wurde die Arbeit im Rahmen des Projektes AKTIV-SFR [AKT10].

In diesem Projekt wurde auf aktive und präventive Maßnahmen gesetzt. Dies bedeutet, dass Fußgänger und Radfahrer frühzeitig erkannt werden bzw. der Fahrer über eine mögliche Kollision durch einen Lenkruck informiert wird und ggf. eine automatische Notbremsung zur Kollisionsfolgenminderung eingeleitet wird.

Um dieses Konzept zu realisieren, muss als erstes das Umfeld des Ego-Fahrzeugs ausreichend erfasst werden. Fußgänger, Radfahrer und andere am Verkehr teilnehmende Fahrzeuge sowie Fahrbahn und Randbebauungen, etc. müssen über eine lange Zeit erfasst und verfolgt werden. Für diesen Zweck werden in der Automobilindustrie verschiedene Sensoren zur Umfeldvermessung und -abtastung eingesetzt. Jede Sensorart besitzt ein eigenes Messprinzip und hat dadurch unterschiedliche Vor- und Nachteile beim Einsatz. Im Allgemeinen können die Sensoren in zwei Gruppen grob aufgeteilt werden, diese sind die aktiven Sensoren und passiven Sensoren.

Bei aktiven Sensoren werden elektromagnetische Wellen einer bestimmten Frequenz oder andere Informationsträger in die Umgebung gesendet. Die Reflektion der Welle wird aus-

gewertet, um Information der Umgebung zu extrahieren. Ein sehr verbreitet eingesetzter Sensor auf dieser Basis ist der Radarsensor. Aufgrund seiner hohen Reichweite und Witterungsrobustheit wird der Sensor für ACC-Anwendungen eingesetzt. Die Zieltrennung sowie größere Winkelauflösung wird derzeit weiterentwickelt. Ein anderer Sensor, der ebenfalls auf dem Time-Of-Flight-Prinzip basiert, ist der Lasersensor. Dieser Sensor zeichnet sich durch eine relative hoher Winkelauflösung aus und stellt eine gute Alternative dar. Ein weiterer Sensor auch in dieser Richtung ist der PMD-Sensor. Dieser Sensor gibt ein grobes Entfernungsbild der Umgebung wieder. Damit besteht die Möglichkeit, nicht nur die Entfernung der Objekte sondern auch ihre Höhe und Breite zu bestimmen.

Auf der anderen Seite werden häufig Kameras als passiver Sensor eingesetzt. Ein Vorteil von passiven Sensoren liegt eindeutig in der Umweltverträglichkeit. Durch Auswertung der aufgenommenen Kamerabilder werden Informationen zu Objekten im Fahrzeugumfeld gewonnen. Der Nachteil der Kameras liegt in der genauen Bestimmung der Entfernung zum Objekt. Anders als aktive Sensoren kann die Entfernung zum Objekt nicht gemessen sondern nur geschätzt werden. Dies setzt eine Mustererkennung voraus, die in unbekanntem Situationen zu Problemen führen könnte. Eine Erweiterung vom Mono-Kamera-System auf ein Stereokamera-System bringt erhebliche Verbesserungen in der Entfernungsbestimmung. In dieser Arbeit wurde ein Stereokamera-System zur Umfeldwahrnehmung eingesetzt. Im Vergleich zu anderen Sensoren kann bei einem Stereokamera-System zusätzlich zur Entfernungsschätzung noch die Bildinformation ausgenutzt werden. Dies erweist sich als ein großer Vorteil.

Der Normalfall eines Stereokamera-Systems besteht aus zwei Kameras gleicher Art, die parallel zu einander und auf gleicher Höhe aufgestellt werden. Auf Basis von zwei Bildern, die durch die linke und die rechte Kamera aufgenommen werden, kann theoretisch die Entfernungsinformation für jedes Pixel bestimmt werden, wenn in zwei Bildern zwei Pixel identifiziert werden können, die zum selben 3D-Punkt eines Objektes im Raum gehören. Diese Pixel werden als korrespondierende Pixel bezeichnet. Durch den Versatz beider Pixel zueinander kann die Entfernung ausgerechnet werden.



Die wesentliche Herausforderung in diesem Prozess ist die computerbasierte Korrespondenzsuche. Es wurden die klassischen Verfahren wie die Anwendung von SAD oder KKFMF vorgestellt. Die Zuverlässigkeit der berechneten Disparitätskarte kann durch Forderung nach einem eindeutigen Minimum der Korrelationskosten erhöht werden. Falsch zusammengehörige Pixel werden dadurch weitgehend heraus gefiltert. In texturarmen Regionen kann mit dem Verfahren jedoch keine Disparität gefunden werden. Deshalb wurden neben den genannten pixelbasierten Ansätzen auch globale Ansätze vorgestellt. Dabei wird die Disparitätskarte ausgewählt, die die minimalen globalen Kosten verursacht. Da diese Verfahren in der Regel großen Rechenaufwand benötigen, was zu Problemen bei echtzeitfähigen Anwendungen führt, wird der Ansatz Semiglobal-Matching zur Disparitätsberechnung verwendet. Dieser zeichnet sich durch die Rechenzeiteffizienz und eine dichte Disparitätskarte aus. Auf Basis der berechneten Disparitätskarte erfolgt eine 3D-Rekonstruktion der Umgebung. Die 3D-Information ist sozusagen die resultierende Messung eines Stereokamera-Systems.

In einem Wahrnehmungssystem werden gewonnene Messungen zur Umfeldbeschreibung, oder auch Umfeldmodellierung, verwendet. Die Aufgabe der Umfeldmodellierung ist es, das reale Umfeld zuverlässig und präzise durch eine geeignete Darstellungsform wiederzugeben, damit die darauf aufbauenden Sicherheitsmaßnahmen zur Kollisionsvermeidung eingesetzt werden können. Generell gibt es zwei Verfahren zur Umfeldmodellierung: Das objektbasierte und das kartenbasierte Verfahren. Das objektbasierte Verfahren beschreibt das Umfeld als eine Menge von Objekten. Andere Verkehrsteilnehmer sowie Hindernisse werden sozusagen als Objekte beschrieben. Jedes Objekt wird durch einen Zustandsvektor dargestellt und repräsentiert. Die Bewegung eines Objekts kann mit Hilfe eines Kalman-Filters unter Anwendung eines geeigneten Bewegungsmodells verfolgt werden. Der objektbasierte Ansatz ist sehr gut geeignet, bewegte Objekte zu verfolgen. Bei der Darstellung von vielfältigen stationären Hindernissen zeigt dieser Ansatz Einschränkungen. Die endliche Anzahl der Zustände im Zustandsvektor eines Objektes reichen nicht immer aus, um die große Vielfalt der Objekten und Hindernissen, besonders im Stadtverkehr, zu beschreiben.

Dagegen kann mit dem kartenbasierten Ansatz das Umfeld, trotz seiner Vielfalt, sehr gut beschrieben werden, solange es um stationäre Hindernisse geht. Bei diesem Ansatz wird das Umfeld mit einem 2D-Gitter modelliert. Jede Zelle deckt einen Bereich im Umfeld ab und enthält die Information über die Befahrbarkeit dieses Bereiches. Dadurch können im Prinzip Hindernisse jeder Form dargestellt werden. Der Nachteil des kartenbasierten Verfahrens besteht aber darin, dass bewegte Objekte nicht ohne weiteres verfolgt werden können. Sie verursachen dabei unzuverlässige Bereiche auf der Karte. An dieser Stelle ist es nötig, beide Verfahren miteinander zu kombinieren, um dadurch insgesamt bessere Ergebnisse zu erzielen. Dies gelingt jedoch nicht automatisch, da beide Verfahren viele Vorkommnisse gleichzeitig beschreiben, außerdem möchte man die dadurch entstehende doppelte Beschreibung vermeiden.

Im Bezug darauf wird in dieser Arbeit ein neues Verfahren vorgestellt. Dieses neue Verfahren verwendet an geeigneten Stellen die Ergebnisse des einen Verfahrens als Indiz für das andere Verfahren:

Als erstes wird die Objektverfolgung durch Bildinformation verbessert. Die Unsicherheit in der Messung-Prädiktion-Assoziation kann mit Hilfe der SURF-Merkmale, die aus der Bildinformation jedes Objekts extrahiert werden, erheblich verbessert werden. Durch die SURF-Merkmale können Objekte, die ähnliche Abmessungen und Positionen haben, gut voneinander unterschieden werden. Durch eine verbesserte Objektverfolgung sind bewegte Objekte einfacher zu klassifizieren. Diese Erkenntnis wird dann verwendet, um die kartenbasierte Darstellung zu verbessern. Es werden nur Messungen in die Belegungskarte fusioniert, die nicht mit bewegten Objekten in Verbindung stehen. Dadurch beschreibt die Belegungskarte sozusagen lediglich stationäre Bereiche. Am Ende stehen folgende Ergebnisse zur Verfügung: Bewegte Objekte werden mit dem objektbasierten Verfahren erfolgreich erfasst und getrackt. Gleichzeitig werden stationäre Hindernisse und Bereiche mit dem kartenbasierten Verfahren dargestellt.

Das Verfahren wurde auf Messdaten, die in verschiedenen Verkehrssituationen in der Praxis aufgenommen wurden, getestet. Die Testergebnisse zeigen zum einen, dass das Objekttracking auf Basis von Stereokamera-Daten in der Lage ist, bewegte Fahrzeuge mit un-

122

terschiedlichen Geschwindigkeiten und Fahrtrichtungen zu verfolgen, zum anderen, dass die Belegungskarte durch Kopplung mit dem Objekttracking zuverlässiger wird. 8 - 20% der falsch detektierten Zellen werden korrigiert. Die darauf basierende Kollisionsprädiktion ist damit wesentlich zuverlässiger. Im Projekt AKTIV-SFR beinhalten die Fußgänger-schutzmaßnahmen neben der Notbremsung auch noch die Lenkempfehlung, damit eine Kollision eventuell noch vermieden werden kann. Dazu sind die Erkenntnisse über freie befahrbare Bereiche entscheidend. Dies kann mit der kartenbasierten Darstellung, wo freie Bereiche eindeutig markiert werden, sehr gut realisiert werden.

Mit den erzielten Ergebnissen leistet diese Arbeit einen wichtigen Fortschritt im Forschungsbereich Umfeldwahrnehmung auf Basis eines Stereokamera-Systems in Fahrerassistenzsystemen.

## 7.2 Ausblick

In zukünftigen Arbeiten kann die Leistungsfähigkeit des Verfahrens weiterhin verbessert werden, wenn die Schätzung der Ego-Position und die Objektbildung durch Segmentierung auf Basis der 3D-Daten verbessert werden. Durch Ungenauigkeit in der Ego-Positionsschätzung werden langsam bewegte Objekte als stationäre Hindernisse, aber auch stationäre Hindernisse als bewegte Objekte, klassifiziert. Die Schätzung der Ego-Position erfolgt in dieser Arbeit lediglich durch Anwendung des Einspurmodells auf gewonnene Fahrzeugsdaten wie Radgeschwindigkeit, Gierrate, etc. Jedoch kann unabhängig davon auf Basis der Kamerabilder der optische Fluss berechnet werden. Dadurch kann die Bewegung der Kamera, die der Bewegung des Ego-Fahrzeugs gleicht, bestimmt werden. Dies dient als eine zusätzliche Messung für die Bestimmung der Ego-Position.

Des Weiteren ist der optische Fluss von Pixeln, die zum bewegten Objekt gehören, anders als der von stationären Hindernissen. Weiterhin kann die Möglichkeit der Umsetzung des Verfahrens auf Basis von hardwarebasierten eingebetteten Systemen, zum Beispiel FPGA + GPU, betrachtet werden. Da sich viele Bildverarbeitungsoperationen auf FPGA realisieren lassen, was zu erheblich kürzeren Rechenzeiten führt, kann die gesamte Rechenzeit ver-

ringert und die Effizienz des Verfahrens gesteigert werden. Dies ist für die darauf aufbauenden Applikationen wie die automatische Notbremsung von entscheidendem Vorteil.

# Veröffentlichungen

## Journal Publikationen und Zeitschriftenbeiträge

**Thien-Nghia Nguyen, Bernd Michaelis, Ayoub Al-Hamadi, Michael Tornow and Marc-Michael Meinecke 2011.** *Stereo-Camera-Based Urban Environment Perception Using Occupancy Grid and Object Tracking*. IEEE Transactions on Intelligent Transportation Systems 2011 (in Druck).

**Michael Tornow, Michael Grasshoff, Thien-Nghia Nguyen and Bernd Michaelis 2011,** *Fast Computation of Dense and Reliable Depth Maps from Stereo Images*. Machine Vision, ISBN 979-953-307-677-4, Editor: Fabio Solari. InTech (in Druck).

## Konferenzartikel

**Thien-Nghia Nguyen, Marc-Michael Meinecke, Bernd Michaelis 2009.** *Optimized Grid-based Fusion Approach for Vehicle Environment Perception*. IEEE Intelligent Vehicles Symposium 2009, Xian, China, ISBN 13: 978-3-937655-19-2, pp. 425-430, 2009.

**Jiri Jerhot, Marc-Michael Meinecke, Thomas Form, Thien-Nghia Nguyen, Ganymed Stanek, Jörn Knaup 2009.** Integrated probabilistic approach to environmental perception with self-diagnosis capability for advanced driver assistance systems. 12th International Conference on Information Fusion, Seattle, WA, USA, July 6-9, 2009, ISBN 13: 978-0-9824-4380-4, pp. 1347-1354.

**Thien-Nghia Nguyen, Marc-Michael Meinecke, Jan Effertz, Ellen Wykowski, Mark Gonter 2010.** *Aktive Maßnahmen zur Erhöhung der Sicherheit für Fußgänger*. AAET 2010- Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel, ISBN 13: 978-3-937655-23-9, pp. 258-273.

**Marc-Michael Meinecke, Martin Roehder, Thien-Nghia Nguyen, Marian Andrzej Obojski, Michael Heuer, Björn Giesler, Bernd Michaelis 2010.** *Motion Model Estimation for Pedestrians in Street-Crossing Scenarios*. 7th Workshop on Intelligent Transportation WIT 2010, Hamburg/Germany, 23rd-24th March 2010.

**Thien-Nghia Nguyen, Marc-Michael Meinecke, Bernd Michaelis, Jiri Jerhot, Jörn Knaup 2009.** *Ein Kombinationsansatz kartenbasierter und objektbasierter Umfeldmodellierung für Fahrerassistenzsysteme*. AAET 2009 - Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel, ISBN 13:978-3-937655-19-2, pp. 264-277.

**Jiri Jerhot, Marc-Michael Meinecke, Thomas Form, Thien-Nghia Nguyen 2009.** *Methodik zur probabilistischen Selbstdiagnose der Umfeldwahrnehmung von Fahrerassistenzsystemen*. AAET 2009 - Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel, ISBN 13:978-3-937655-19-2, pp. 330-340.

## Patentanmeldungen

DE 10 2009 006 113A1. **Thien-Nghia Nguyen, Marc-Michael Meinecke, Thanh-Binh To, Bernd Michaelis**. Vorrichtung und Verfahren zur Sensorfusion mit dynamischen Objekten.

DE 10 2010 006 828A1. **Marc-Michael Meinecke, Jan Effertz, Thorsten Graf, Dirk Stüker, Jörn Christian Knaup, Thien-Nghia Nguyen**. Verfahren zur automatischen Erstellung eines Modells der Umgebung eines Fahrzeugs sowie Fahrerassistenzsystem und Fahrzeug.

# Literatur

- [AET10] Nguyen, T., Meinecke, M., Effertz, J., Wykowski, E. & Gonter, M. 2010. *Aktive Manahmen zur Erhhung der Sicherheit fuer Fussgaenger*. AAET 2010 Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel, pp. 258-273, 2010.
- [AIH06b] Al-Hamadi, A., Niese, R. & Michaelis, B. 2006. *A fast and robust approach for the segmentation of moving objects*. Journal of Computer Vision and Graphics, Vol. 32pp. 13-19, 2006.
- [AKT10] 2010. *AKTIV - Adaptive and Cooperative Technologies for the Intelligent Traffic - a German Research Initiative 2006-2010*. *AKTIV - Adaptive and Cooperative Technologies for the Intelligent Traffic - a German Research Initiative 2006-2010* , 2010.
- [AKT11] Galla, A., Kleen, A., Meinecke, M., Obojski, M. A., Scholz, S. & Nguyen, T. 2011. *Schlussbericht Aktiv-AS der Volkswagen AG zur Forschungsinitiative AKTIV*. Volkswagen Aktiengesellschaft, 2011.
- [ARS300] ADC, C. 2011. *ARS 300 + ARS 308-2 - Long Range Radar*. Firmenpress. Continental, 2011.
- [AUT07] Bosch, R. 2007. *Autoelektrik Autoelektronik. 5. vollstaendig bearbeitete und erweiterte Auflage*. Vieweg \& Sohn Verlag, 2007.
- [AYC08] Burlet, J., Vu, T. & Aycard, O. 2008. *Grid-based Localization and Online Mapping with Moving Object Detection and Tracking*. Intelligent Vehicles Symposium, pp. 684-689, 2008.
- [BAD07] Badino, H., Franke, U. & Mester, R. 2007. *Free Space Computation Using Stochastic Occupancy Grids and Dynamic Programming*. In Dynamic Vision Workshop for ICCV, pp. 2007.

- [BAR94] **Barron, J. L., Fleet, D. J. & Beauchemin, S. 1994.** *Performance of optical flow techniques.* Int. J. Comput. Vis, Vol. 12pp. 43-77, 1994.
- [BAY63] **Bayes, T. 1763.** *An essay towards solving a problem in the doctrine of chances.* Philosophical Transactions of the Royal Society of London, Vol. 53pp. 370-418, 1763.
- [BEK64] **Bekker, C. 1964.** *Augen durch Nacht und Nebel. Die Radarstory. 2. verbesserte Auflage.* Stalling Verlag, Oldenburg, 1964.
- [BEL92] **Belhumeur, P. N. & Mumford, D. A. 1992.** *A bayesian treatment of the stereo correspondence problem using half-occluded regions.* IEEE Conf. On Computer Vision and Pattern Recognition, pp. 1992.
- [BER09] **Berger, C., Krahn, H., Rumpe, B. & Bartels, A. 2009.** *Qualitaetsgesicherte Fahrentscheidungsunterstuetzung fuer automatisches Fahren auf Schnellstrassen und Autobahnen.* AAET 2009 - Automatisierungs-, Assistenzsysteme und eingebettete Systeme fuer Transportmittel, pp. 2009.
- [BLA99] **Blackman, S. S. & Popoli, R. F. August 1999.** *Design and Analysis of Modern Tracking Systems.* Artech House Publishers, August 1999.
- [BLE04] **Bleyer, M. & Gelautz, M. 2004.** *A Layered Stereo Algorithm Using Image Segmentation And Global Visibility Constraints.* IEEE International Conference on Image Processing, pp. 2997-3000, 2004.
- [BLO88] **P., H. A. & Bar-Shalom, Y. 1988.** *The interacting multiple model algorithm for systems with markovian switching coefficients.* IEEE Trans. Automatic Control, Vol. 33(8), pp. 780-783, 1988.
- [BOY01] **Boykov, Y., Veksler, O. & Zabih, R. 2001.** *Fast approximate energy minimization via Graph Cuts.* IEEE Transactions Pattern Analysis and Machine Intelligence, Vol. 23pp. 1222-1239, 2001.



- [BRA06] **Brailion, C., Pradalier, C., Usher, K., Crowley, J. & Laugier, C. 2006.** *Occupancy grids from stereo and optical flow data.* Proc. of the Int. Symp. on Experimental Robotics, pp. 2006.
- [BRE65] **Bresenham, J. E. 1965.** *Algorithm for computer control of a digital plotter.* IBM Systems Journal, Vol. 4pp. 25-30, 1965.
- [BRO05] **Broggi, A., Caraffi, C., Fedriga, R. I. & Grisleri, P. 2005.** *Obstacle Detection with Stereo Vision for Off-Road Vehicle Navigation.* Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1pp. 56-65, 2005.
- [BUR05] **Burger, W. & Burge, M. J. 2005.** *Digitale Bildverarbeitung: Eine Einfuehrung mit Java und ImageJ.* Springer, 2005.
- [COX96] **Cox, I., Hingorani, S. & Rao, S. 1996.** *A Maximum Likelihood Stereo Algorithm.* Computer Vision and Image Understanding, Vol. 63(3), pp. 542-567, 1996.
- [DAN07] **Dang, T. 2007.** *Kontinuierliche Selbstkalibrierung von Stereokameras.* Dissertation. Universitaet Karlsruhe, 2007.
- [DAR07] *Urban Challenge - a prize competition for driverless vehicles, funded by the Defense Advanced Research Projects Agency. 2004 - 2007. .*
- [DEM67] **Dempster, A. P. 1967.** *Upper and lower probabilities induced by a multivalued mapping.* Ann. Math. Statist, Vol. 38(2), pp. 325-339, 1967.
- [EFF09] **Effertz, J. 2009.** *Autonome Fahrzeugfuehrung in urbaner Umgebung durch Kombination objekt- und kartenbasierter Umfeldmodelle.* Dissertation. Technische Universität Braunschweig, 2009.
- [EIG09] **Eigel, T., Giebel, T. & Schnieder, E. 2009.** *AKTIV IQF: Integrierte Laengs- und Querfuehrung durch strukturvariable Regelung.* AAET 2009 - Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel, pp. 2009.

- [ELF89] **Elfes, A. 1989.** *Using Occupancy Grids for Mobile Robot Perception and Navigation.* Computer, IEEE Computer Society, Vol. 22(6), pp. 46-57, 1989.
- [ELI08] **Elias, B. 2008.** *Improvement of a PMD Based Classification of Pedestrians in Road Traffic Scenarios.* 7th Workshop on Intelligent Transportation, March 2010, Hamburg, Germany, pp. 2008.
- [FAU93] **Faugeras, O. 1993.** *Three Dimensional Computer Vision: A Geometric Viewpoint (Artificial Intelligence).* Mit Pr, ISBN: 978-0262061582, 1993.
- [FLE05] **Fleet, D. J. & Weiss, Y. 2005.** *Optical Flow Estimation.* In & N. Paragios, Y. Chen & O. Faugeras (Editor), *Mathematical Models in Computer Vision: The Handbook*, pp. 239-258, Springer, 2005.
- [FUR02] **Streller, D., Furstenberg, K. & Dietmayer, K. 2002.** *Vehicle and object models for robust tracking in traffic scenes using laser range images.* Intelligent Transportation Systems, 2002, pp. 2002.
- [FUR0202] **Fuerstenberg, K. & Lages, U. 2002.** *Pedestrian Detection and Classification by Laserscanners.* In Procs. IEEE Intelligent Vehicles Symposium 2002, pp. 2002.
- [FUR02ITS] **Fuerstenberg, K. C., J., K. C., Eisenlauer, S. & Willhoeft, V. 2002.** *Multilayer Laserscanner for robust Object Tracking and Classification In Urban.* Proceedings of ITS 2002, 9th World Congress on Intelligent Transport Systems, ITS 2002 Chicago, pp. 2002.
- [GAM87] **Gamble, E. & Poggio, T. 1987.** *Visual integration and detection of discontinuities: The key role of intensity edges.* Massachusetts Institute of Technology, 1987.
- [GAV07] **Gavrila, D. M. & Munder, S. 2007.** *Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle.* International Journal of Computer Vision 73, Vol. 1pp. 41-59, 2007.
- [GRA11] **Graßhoff, M. 2011.** *Dichte Tiefenkarten aus stereophotogrammetrischen Bildern auf FPGAs berechnen.* Diplomarbeit. Universität Magdeburg, 2011.

- [HAS09] **Hastie, T., Tibshirani, R. & Friedman, J. 2009.** *The Elements of Statistical Learning (2nd ed.)*. In (Editor), pp. 520-528, Springer, 2009.
- [HDL64] **2010.** *High Definition Lidar - HDL64E S2*. Firmenpress, 2010.
- [HEE10] **Heesen, M., Kelsch, J., Löper, C. & Flemisch, F. 2010.** *Haptisch-multimodale Interaktion fuer hochautomatisierte, kooperative Fahrzeugf?hrung bei Fahrstreifenwechsel-, Brems- und Ausweichmanoevern*. AAET 2010 - Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel, pp. 2010.
- [HER09] **Hermann, S., Klette, R. & Destefanis, E. 2009.** *Inclusion of a Second-Order Prior into Semi-Global Matching*. The 3rd Pacific-Rim Symposium on Image and Video Technology (PSIVT2009), Tokyo, Japan, January 13th—16th, 2009, pp. 2009.
- [HEU10] **Heuer, M. 2010.** *Aufbau einer Belegungskarte anhand eines Stereo-Kamera-Systems im Kontext moderner Fahrerassistenzsysteme*. Diplomarbeit. Universität Magdeburg, 2010.
- [HOR81] **Horn, B. K. & Schunck, B. G. 1981.** *Determining optical flow*. Artificial Intelligence, Vol. 1-3(17), pp. 185-203.
- [INS07] **Furstenberg, K., Hopstock, M., Obojski, M. A. & Deutschle, S. 2007.** *INTERSAFE Final Report*. IBEO, VW, MBW, TRW, IKA, 2007.
- [INS09] **Roessler, B., Fuerstenberg, K., Ahrholdt, M., Obojski, A. & Hopstock, M. 2009.** *INTERSAFE 2 - Project Presentation*. IBEO, VTEC, VW, BMW, 2009.
- [JER09] **Jerhot, J., Meinecke, M., Form, T., Stanek, G., Nguyen, T. & Knaup, J. 2009.** *Integrated probabilistic approach to environment perception with selfdiagnosis capability for advanced driver assistance systems*. The 12th International conference on information fusion FUSION '09, pp. 1347-1354, 2009.
- [KAL60] **Kalman, R. E. 1960.** *A New Approach to Linear Filtering and Prediction Problems*. Transaction of the ASME, Journal of Basic Engineering, Vol. 82pp. 35-45, 1960.

- [KNO01] **Knöppel, C.** 2001. *Stereobasierte und spurgenaue Erkennung von Straßenfahrzeugen im Rückraum eines Straßenfahrzeuges*. Dissertation. Universität Magdeburg, 2001.
- [LAB05] **Labayrade, R.** 2005. *GENERIC , REAL TIME AND ROBUST ROAD OBSTACLES DETECTION USING ONBOARD STEREOVISION*. Young researcher Seminar ECTRI, Vol. 1pp. 1-13, 2005.
- [LAM10] **Lambert, G., Katzwinkel, R. & Schulze, K.** 2010. *Assistenten zur Rangierhilfe am Beispiel des neuen Volkswagen Touareg*. Fahrerassistenz und Integrierte Sicherheit. 26. VDI/VW-Gemeinschaftstagung, Wolfsburg, 6. und 7. Oktober 2010, pp. 467-476, VDI Nachrichten, 2010.
- [LEX03] *Lexus LS 400, LS 430 and LS 460 Generations: Setting the Standard in the Most Comfortable Way Possible. Lexus LS 400, LS 430 and LS 460 Generations: Setting the Standard in the Most Comfortable Way Possible .*
- [LOW99] **Lowe, D. G.** 1999. *Object recognition from local scale-invariant features*. Proceedings of the International Conference on Computer Vision. 2, pp. 1150-1157, 1999.
- [LUC81] **Lucas, B. D. & Kanade, T.** 1981. *An iterative image registration technique with an application to stereo vision*. Proceedings of Imaging Understanding Workshop, pp. 121-130, 1981.
- [LUH00] **Luhmann, T.** 2000. *Nahbereichsphotogrammetrie*. Wichmann Verlag Heidelberg, 2000.
- [MAH36] **Mahalanobis, P. C.** 1936. *On the generalised distance in statistics*. Proc. Nat. Inst. Sci. India, Vol. 2(2), pp. 49-55, 1936.
- [MAR06] **der, W. v. & Gavrila, D. M.** 2006. *Real-Time Dense Stereo for Intelligent Vehicles*. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, Vol. 7(1), pp. 38-50, 2006.

- 
- [MAT89] **Matthies, L., Kanade, T. & Szeliski, R. 1989.** *Kalman filter-based algorithms for estimating depth from image sequences.* International Journal of Computer Vision, Vol. 3pp. 209-238, 1989.
- [MEI01] **Meinecke, M. 2001.** *Zum optimierten Sendesignalentswurf fuer Automobilradare.* Dissertation. TU Hamburg, 2001.
- [MEI10] **Meinecke, M., Roehder, M., Nguyen, T. & Obojski, M. 2010.** *Motion Model Estimation for Pedestrians in Street-Crossing Scenarios.* 7th Workshop on Intelligent Transportation, March 2010, Hamburg, Germany, pp. 2010.
- [MIT04] **Mitschke, M. 2004.** *Dynamik der Kraftfahrzeuge - Auflage: 4.* Springer Berlin Heidelberg, 2004.
- [MIT95] *Preview Distance Control - Press Releases Date: 15 December, 1998. Preview Distance Control - Press Releases Date: 15 December, 1998 .*
- [MON08] **Montemerlo, M. 2008.** *Junior: The Stanford Entry in the Urban Challenge.* Journal of Field Robotics, Vol. 25(9), pp. 569-597, 2008.
- [MUL01] **Hirschmüller, H. 2001.** *Improvements in Real-Time Correlation-Based Stereo Vision.* Proceedings of IEEE Workshop on Stereo and Multi-Baseline Vision, Kauai, Hawaii, 9-10 December 2001, pp. 141-148, 2001.
- [MUL05] **Hirschmuller, H. 2005.** *Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information.* IEEE Conference on Computer Vision and Pattern Recognition, pp. 2005.
- [MUR00] **Murphy, R. 2000.** *An Introduction to AI Robotics.* (ed. 1st), MIT Press, 2000.
- [NED07] **Nedeveschi, S., Graf, T., Meinecke, M. & To, T. 2007.** *A Sensor for Urban Driving Assistance Systems Based on Dense Stereovision.* Proceedings of Intelligent Vehicles, pp. 278-286, 2007.

- [NED09] Nedeveschi, S., Bota, S. & Tomiuc, C. 2009. *Stereo-Based Pedestrian Detection for Collision-Avoidance Applications*. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, Vol. 10(3), pp. 380-391, 2009.
- [NGU09] Nguyen, T., Meinecke, M., Tornow, M. & Michaelis, B. 2009. *Optimized Grid-Based Environment Perception in Advanced Driver Assistance Systems*. IEEE Intelligent Vehicles Symposium, pp. 425-430, 2009.
- [NIX08] Nixon, M. & Aguado, A. 2008. *Feature Extraction & Image Processing*. Elsevier Academic Press, 2008.
- [OPENCV] 2008. *Learning OpenCV: Computer Vision with the OpenCV Library*. (G. Bradski, A. Kaehler, M. Loukides & R. Romano, ) O'Reilly Media, 2008.
- [PAR10] *Assistenzsysteme: Rear Assist.. Assistenzsysteme: Rear Assist. .*
- [PEA90] Pearl, J. 1990. *Reasoning with belief functions: An analysis of compatibility*. Int. J. Approx. Reason, Vol. 4(5/6), pp. 363-389, 1990.
- [PET09] Petrovskaya, A. & Thrun, S. 2009. *Model Based Vehicle Detection and Tracking for Autonomous Urban Driving*. Autonomous Robots, Vol. 26(2-3), pp. 123-139, 2009.
- [PLA06] Plagge, M., Eggert, E., Beyer, S., Smidt, W., Anders, G. & Katzwinkel, R. 2006. *Kartenbasierte Modellierung des Fahrzeugnahbereichs*. VDI-Berichte Nr. 1960, 2006, pp. 499-508, 2006.
- [PMD08] Heckenkamp, C. 2008. *Grundlagen der Bildverarbeitung: Das PMD-Prinzip*. Inspect - Whitepaper, 2008.
- [PMD09] PMDTechnologiesGmbH, 2009. *Whitepaper - A performance review of 3D TOF vision systems in comparison to stereo vision systems*. PMDTechnologies GmbH, 2009.
- [POG85] Poggio, T., Torre, V. & Koch, C. 1985. *Computational vision and regularization theory*. Nature, Vol. 317pp. 314-319, 1985.

- 
- [PRE05] *PREVENT - a European automotive industry activity co-funded by the European Commission to contribute to road safety.* .
- [PRO95] **1987-1995.** *PROMETHEUS - PROgramMme for a European Traffic of Highest Efficiency and Unprecedented Safety. PROMETHEUS - PROgramMme for a European Traffic of Highest Efficiency and Unprecedented Safety* , 1987-1995.
- [RAB89] **Rabiner, L. R. 1989.** *A tutorial on Hidden Markov Models and selected applications in speech recognition.* Proceedings of the IEEE, Vol. 77(2), pp. 257-286, 1989.
- [RAM10] **Narasimha, R. 2010.** *Depth Recovery from Stereo Matching Using Coupled Random Fields. Dissertation.* UNIVERSITÉ DE GRENOBLE, 2010.
- [REH97] **Rehrmann, V. & Priese, L. 1997.** *Fast and Robust Segmentation of Natural Color Scenes.* In Proceedings of the 3 rd Asian Conference on Computer Vision, pp. 598-606, Srpinger Verlag, 1997.
- [REI10] **2010.** *Fahrstabilisierungssysteme und Fahrerassistenzsysteme.* (K. Reif, ) Vieweg + Teubner, 2010.
- [RIN07] **Ringbeck, T. 2007.** *a 3D time of flight camera for object detection.* Optical 3-D Measurement Techniques 09-12.07.2007 ETH Zürich, pp. 2007.
- [RIS04] **Ristic, B., Arulampalam, S. & Gordon, N. 2004.** *Beyond the Kalman Filter: Particle.* Artech House Publishers, 2004.
- [SAV05] *SAVE.U - Sensors and system Architecture for Vulnerable road Users protection. SAVE.U - Sensors and system Architecture for Vulnerable road Users protection .*
- [SCHA02] **Scharstein, D. & Szeliski, R. 2002.** *A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms.* International Journal of Computer Vision, Vol. 47pp. 7-42, 2002.

- [SCH107] **Schindler, E.** 2007. *Fahrdynamik. Grundlagen des Lenkverhaltens und ihre Anwendung für Fahrzeugregelsysteme. Kontakt und Studium, Band 685.* Expert-Verlag GmbH, 2007.
- [SCHR05] **Schreer, O.** 2005. *Stereoanalyse und Bildsynthese.* Springer, 2005.
- [SHA00] **Bar-Shalom, Y.** 2000. *Multitarget/Multisensor tracking: Applications and Advances.* Artech House Publishers, 2000.
- [SHA76] **Shafer, G.** 1976. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.
- [SHA89] **Bar-Shalom, Y., Chang, K. & Blom, H.** 1989. *Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm.* IEEE Transactions on Aerospace and Electronic Systems, Vol. 25(2), pp. 296-300, 1989.
- [SHLE05] **Shlens, J.** 2005. *A tutorial on Principal Components Analysis.* Measurement, Vol. 51(10003), pp. 52, 2005.
- [SID08] *Side Assist - radarbasiertes Spurwechselassistenten - Fachartikel-Archive - Ausg: 2008-02. Side Assist - radarbasiertes Spurwechselassistenten - Fachartikel-Archive - Ausg: 2008-02 .*
- [STE04] **Stein, F.** 2004. *Efficient Computation of Optical Flow Using the Census Transform.* In & C. E. Rasmussen, H. H. Bülthoff, M. A. Giese & B. Schölkopf (Editor), *Pattern Recognition*, pp. 79-86, Springer, 2004.
- [STU04] **Stüker, D.** 2004. *Heterogene Sensordatenfusion zur robusten Objektverfolgung im automobilen Straßenverkehr.* Dissertation. Universität Oldenburg, 2004.
- [SURF06] **Bay, H., Ess, A., Tuytelaars, T. & Gool, L. V.** 2006. *SURF: Speeded-Up Robust Features.* European Conference on Computer Vision ECCV, Vol. 1pp. 404-417, 2006.



- 
- [TAO00] **Tao, H. & Sawhney, H. S. 2000.** *Global Matching Criterion and Color Segmentation Based Stereo*. IEEE Workshop on Applications of Computer Vision, pp. 246, IEEE Computer Society, 2000.
- [THR05] **Thrun, S., Burgard, W. & Fox, D. 2005.** *Probabilistic Robotics*. The MIT Press, 2005.
- [TÖN05] **Tönnies, K. D. 2005.** *Grundlagen der Bildverarbeitung*. Pearson, ISBN: 978-3-8273-7155-3, 2005.
- [TOR09] **Tornow, M. 2009.** *Untersuchung und Entwicklung von Algorithmen zur Stereobildauswertung für die Erfassung von Objekten im Umfeld von Fahrzeugen und Realisierung einer Hindernisdetektion in Echtzeit mittels einer Hardwareimplementierung auf einem FPGA*. Dissertation. University of Magdeburg, 2009.
- [VIO01] **Viola, P. & Jones, M. 2001.** *Rapid Object Detection using a Boosted Cascade of Simple Features*. Conference on Computer Vision and Pattern Recognition, pp. 2001.
- [VU07] **Vu, T., Aycard, O. & Appenrodt, N. 2007.** *Online Localization and Mapping with Moving Object Tracking in Dynamic Outdoor Environments*. In Proc. IEEE Intelligent Vehicles Symposium, pp. 2007.
- [VU09] **Vu, T. & Aycard, O. 2009.** *Laser-based Detection and Tracking Moving Objects using Data-Driven Markov Chain Monte Carlo*. IEEE International Conference on Robotics and Automation, pp. 2009.
- [WAN07] **Wang, C., Thorpe, C., Thrun, S., Hebert, M. & Durrant-Whyte, H. 2007.** *Simultaneous Localization, Mapping and Moving Object Tracking*. The International Journal of Robotics Research, Vol. 26(9), pp. 889-916, 2007.
- [WAR85] **Warren, D. H. & Strelow, E. R. 1985.** *Electronic Spatial Sensing for the Blind: Contributions from Perception*. Springer, ISBN 9024726891.

- [WEH00] **Weisser, H., Schulenberg, P., Bergholz, R. & Lages, U. 2000.** *Laserscanners for Obstacle Detection in Automotive Applications.* IEEE Intelligent Vehicles Symposium 2000, pp. 2000.
- [WEI06] **Weiser, A. 2006.** *Kartenbasierte Umfeldwahrnehmung: Entwurf und Implementierung einer Fahrschlauchschätzung basierend auf den Daten eines Lasersensors.* Diplomarbeit. Technische Universität Ilmenau, 2006.
- [WEI09] **Weiser, A. & Bartels, A. 2009.** *Intelligent Car – Teilautomatisches Fahren auf der Autobahn.* AAET 2009 - Automatisierungs-, Assistenzsysteme und eingebettete Systeme für Transportmittel, pp. 2009.
- [WEL06] **Welch, G. & Bishop, G. An Introduction to the Kalman Filter, Updated: Monday, July 24, 2006.** *An Introduction to the Kalman Filter, Updated: Monday, July 24, 2006.*
- [WIN09] **2009.** *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme fuer aktive Sicherheit und Komfort.* (H. Winner, S. Hakuli & G. Wolf, ) Vieweg+Teubner, 2009.
- [WOE09] **Woehler, C. 2009.** *3D Computer Vision - Efficient Methods and Applications.* Springer, 2009.
- [WUT10] **Wuttke, U. & Barth, H. 2010.** *Park Assist 2.0 - Komplexe Funktionen mit Ultraschalltechnik.* Fahrerassistenz und Integrierte Sicherheit. 26. VDI/VW-Gemeinschaftstagung, Wolfsburg, 6. und 7. Oktober 2010, pp. 459-466, 2010.
- [ZIT99] **Zitnick, C. L. & Kanade, T. 1999.** *A Cooperative Algorithm for Stereo Matching and Occlusion Detection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22pp. 675-684, 1999.

# Lebenslauf

Name: Thien Nghia Nguyen

Geboren am 26.11.1981 in Ha Tay, Vietnam

Geschlecht: männlich

Familienstand: ledig

Nationalität: vietnamesisch

## Schulbildung

1987 – 1992: Grundschule in Ho-Chi-Minh-Stadt, Vietnam

1993 – 1996: Sekundarschule in Ho-Chi-Minh-Stadt, Vietnam

1996 – 1999: Gymnasium in Ho-Chi-Minh-Stadt, Vietnam

## Studium

1999 – 2000: Studium an der Technischen Universität Ho-Chi-Minh-Stadt

Fachrichtung: Maschinenbau

2000 – 2001: Deutschkurs in Magdeburg

2001 – 2007: Studium an der Otto-von-Guericke Universität Magdeburg

Fachrichtung: Informationstechnologie.

Abschluss: Diplomingenieur (Note 1.7)

## Berufstätigkeit

2007 – dato: Wissenschaftlicher Mitarbeiter am IESK/FEIT an der Otto-von-Guericke Universität Magdeburg (in Zusammenarbeit mit der Volkswagen Konzernforschung).