



Vorausschauende Fahrzeug- sensorik mit Photonic Mixer Device und Videokamera für den aktiven Fußgängerschutz

Björn Elias



Vorausschauende Fahrzeugsensorik mit Photonic Mixer Device und Videokamera für den aktiven Fußgängerschutz

Von der Fakultät für Elektrotechnik und Informationstechnik der
Rheinisch-Westfälischen Technischen Hochschule Aachen zur Erlangung
des akademischen Grades eines Doktors der Ingenieurwissenschaften
genehmigte Dissertation

vorgelegt von
Dipl.-Ing. Björn Elias
aus Leverkusen

Berichter: Univ.-Prof. Dr. Petri Mähönen
Univ.-Prof. Dr.-Ing. Til Aach
Tag der Prüfung: 21. April 2009

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Aufl. - Göttingen : Cuvillier, 2009
Zugl.: (RWTH) Aachen, Univ., Diss., 2009
978-3-86727-975-8

Audi Dissertationsreihe, Band 18

D 82 (Dissertation RWTH Aachen, 2009)

© CUVILLIER VERLAG, Göttingen 2009
Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2009
Gedruckt auf säurefreiem Papier

978-3-86727-975-8

DANKSAGUNG

Die vorliegende Arbeit entstand im Rahmen meiner Tätigkeit als Doktorand in der Elektronik Vorentwicklung der AUDI AG in Ingolstadt in den Jahren 2004 bis 2007.

Ich bedanke mich bei meinem Betreuer Dr. Helmut Riedel für sein Engagement und die konstruktiven Gespräche. Ganz besonderer Dank gilt Prof. Dr. Petri Mähönen, der mich seit vielen Jahren betreut. Seine außergewöhnliche Motivationskraft und die vielen Tipps haben mir den Weg zur Erstellung einer Dissertation für die RWTH Aachen gezeigt. Beiden Herren danke ich für das in mich gesetzte Vertrauen.

Ebenso bedanke ich mich bei Martin Roehder, der das Projekt lange begleitet und durch viele Beiträge zur Sensordatenfusion und Bildverarbeitung bereichert hat, bei Volker Darius für seine engagierte Unterstützung bei der Softwareerstellung und Datenauswertung, Gregor Staudte für die Arbeiten an der Klassifikation, Dennis Mronga für die Validierung, meinen Vorgesetzten Walter Streit, Dr. Martin Meyer und Dr. Björn Giesler, die mir die Arbeit bei der AUDI AG und der Audi Electronics Venture GmbH überhaupt erst ermöglicht und sich jederzeit für mich eingesetzt haben, Tina Binder und Jörgen Erichsen für ihre hilfreichen Anmerkungen und Korrekturen, Kerstin Schöberl für all ihre Unterstützung und meinen Eltern für die Freiräume der vergangenen Jahre.

Diese vielseitigen Unterstützungen haben maßgeblich zum Gelingen der Arbeit beigetragen. Vielen Dank.

INHALTSVERZEICHNIS

DANKSAGUNG	I
INHALTSVERZEICHNIS	III
KURZFASSUNG	VII
ABSTRACT	IX
1 EINLEITUNG	1
1.1 MOTIVATION	3
1.2 FÖRDERPROJEKTE	4
1.3 GLIEDERUNG DER ARBEIT	4
2 VERBREITUNG VORAUSSCHAUENDER ASSISTENZSYSTEME	7
2.1 DAIMLER	7
2.2 LEXUS	8
2.3 BMW	10
2.4 CADILLAC	11
2.5 HONDA	11
3 DER SCHUTZ VON FUSSGÄNGERN	13
3.1 UNFALLFORSCHUNG ZU FUSSGÄNGERUNFÄLLEN	14
3.1.1 BEGRIFFE ZUR METHODISCHEN UNFALLFORSCHUNG	14
3.1.2 DEFINITION DER KATEGORIE »FUSSGÄNGER«	15
3.1.3 FUSSGÄNGERUNFALLSTATISTIK DEUTSCHLAND	15
3.1.4 FUSSGÄNGERUNFALLSTATISTIK EUROPA	18
3.1.5 FUSSGÄNGERUNFALLSTATISTIK USA	19
3.1.6 WELTWEITE FUSSGÄNGERUNFALLSTATISTIK	19
3.2 FAHRERASSISTENZ UND SICHERHEIT	20
3.2.1 CHARAKTERISTIKA VON FAHRERASSISTENZ- UND SICHERHEITSSYSTEMEN	20
3.2.2 WIENER KONVENTION	27

3.2.3	HAFTUNG DES AUTOMOBILHERSTELLERS	28
3.3	MÖGLICHE AUSPRÄGUNG EINES SYSTEMS FÜR DEN AKTI- VEN FUSSGÄNGERSCHUTZ (FSS)	29
3.3.1	FOLGERUNGEN AUS DEN UNFALLANALYSEN	30
3.3.2	AKTIONSKONZEPT	30
3.3.3	BERECHNUNG DER TIME TO CONTACT	31
3.3.4	EXISTIERENDE ALTERNATIVE AKTIONSKONZEPTE	34
3.4	VERKAUFSRELEVANTE ASPEKTE	36
4	VORAUSSCHAUENDE SENSOREN	39
4.1	PHOTONIC MIXER DEVICE (PMD)	40
4.1.1	FUNKTIONSWEISE DER PMD-KAMERA	40
4.2	MONOVIDEO-KAMERA	44
4.2.1	FUNKTIONSWEISE EINES CCD-SENSORS	45
4.2.2	VERGLEICH MIT CMOS-BILDAUFNEHMER	46
4.2.3	BESONDERHEITEN DER BELICHTUNGSREGELUNG	46
4.2.4	ERREICHBARE MESSGENAUIGKEIT	48
4.3	STEREOVIDEOSENSORIK	49
4.4	RADAR	51
4.5	LIDAR	52
4.6	LASERSCANNER	52
4.7	AKTIVE NAHINFRAROT-SENSORIK	53
4.8	THERMISCHE SENSORIK	54
4.9	ULTRASCHALL	56
4.10	SENSORIKEMPFEHLUNG	59
5	FUSSGÄNGERERKENNUNG AUS 3D- UND 2D-DATEN	61
5.1	DIE ALLGEMEINE SCHÄTZTHEORIE	62
5.2	MERKMALSEXTRAKTION AUS 3D-DATEN	63
5.2.1	VORVERARBEITUNG DER REGION	63
5.2.2	GRÖSSE	64
5.2.3	Globale geometrische Merkmale	65
5.2.4	MOMENTENBASIERTE OBJEKTFORM	66
5.2.5	KONTURBASIERTE OBJEKTFORM	68
5.2.6	WEITERE MERKMALE ZUR OBJEKTFORM	70
5.2.7	TEXTUR	71
5.2.8	OBERFLÄCHE	73
5.3	FUSION DER 3D- UND 2D-DATEN	74
5.4	MERKMALSEXTRAKTION AUS 2D-DATEN	75
5.4.1	HISTOGRAMME DER KANTENRICHTUNG	75
5.4.2	MUSTERVERGLEICH	77
5.5	EIGNUNG MÖGLICHER KLASSIFIKATIONSVERFAHREN	78

5.5.1	KLASSIFIKATION DURCH DIREKTEN VERGLEICH . . .	79
5.5.2	SCHWELLENENTSCHEIDUNG	79
5.5.3	k-NÄCHSTE-NACHBARN KLASSIFIKATION	80
5.5.4	MAXIMUM LIKELIHOOD KLASSIFIKATION	80
5.5.5	MAXIMUM-A-POSTERIORI KLASSIFIKATION	81
5.5.6	KÜNSTLICHE NEURONALE NETZE (KNN)	84
5.5.7	SUPPORT VECTOR MACHINES	88
6	VERSUCHSDURCHFÜHRUNG	99
6.1	VERSUCHSAUFBAU	99
6.1.1	EINGESETZTES VERSUCHSFAHRZEUG	100
6.1.2	VERWENDETES PMD-SYSTEM	100
6.1.3	VERWENDETES MONOVIDEO-SYSTEM	101
6.1.4	KOORDINATENSYSTEM UND KALIBRIERUNG	101
6.1.5	WIRKUNGSBEREICH DER SENSORIK	104
6.2	SOFTWAREKOMponentEN VON PRESUME	105
6.2.1	VORKLASSIFIKATOR	107
6.2.2	HAUPTKLASSIFIKATION	108
6.2.3	KLASSIFIKATIONSTRACKING	109
6.2.4	TRENNUNG ZWISCHEN NAH- UND FERNBEREICH . .	109
6.3	TRAINING UND TEST DER KLASSIFIKATOREN	109
6.3.1	AUFTEILUNG DES DATENSATZES	110
6.3.2	BRUTE-FORCE PERMUTATION DER PARAMETER . . .	110
6.3.3	EVOLUTIONÄRER ALGORITHMUS	111
6.4	GÜTEPARAMETER UND BEWERTUNG DER KLASSIFIKATION .	112
6.5	TRAININGS- UND VALIDIERUNGSDATEN	117
7	ERGEBNISSE	121
7.1	ANALYSE DER MERKMALSEXTRAKTION	122
7.2	DETEKTIONSRATE	122
7.3	KLASSIFIKATIONS RATEN	123
7.3.1	EINFLUSS DER SVM-PARAMETER	125
7.3.2	ERGÄNZENDE ERGEBNISVALIDIERUNG	126
7.4	ERKENNUNGSRATEN	126
7.5	ERFORDERLICHE RECHENZEIT	128
7.6	VERGLEICH MIT INTERNATIONALEN FORSCHUNGS- & ENT- WICKLUNGS AKTIVITÄTEN	129
7.6.1	UNIVERSITÄT PARMA	131
7.6.2	TECHNISCHE UNIVERSITÄT CLUJ-NAPOCA	132
7.6.3	TECHNISCHE UNIVERSITÄT HAMBURG-HARBURG . .	132
7.6.4	TECHNISCHE UNIVERSITÄT WUPPERTAL	133
7.6.5	AUSTRALIAN NATIONAL UNIVERSITY	133

7.6.6	DAIMLER AG	134
7.6.7	TOYOTA CENTRAL R & D LABS	135
7.6.8	NISSAN	135
8	ZUSAMMENFASSUNG UND AUSBLICK	141
A	HISTOGRAMME DER MERKMALE	145
B	EINFLUSS DER SVM-PARAMETER	155
C	OPTIMIERUNG MIT GENETISCHEM ALGORITHMUS	163
D	SYSTEMGRENZEN DES VERWENDETEN SENSORSYSTEMS	165
	LITERATURVERZEICHNIS	175

KURZFASSUNG

Die Steigerung der Sicherheit der Verkehrsteilnehmer ist Mittelpunkt vieler Forschungs- und Entwicklungsarbeiten in der Automobilindustrie. Zu den Verkehrsteilnehmern zählen neben den Fahrzeuginsassen auch die ungeschützten Verkehrsteilnehmer, also unter anderem alle Fußgänger. Zur Steigerung der Sicherheit werden verstärkt vorausschauende Sensoren zur Wahrnehmung des Fahrzeugumfeldes eingesetzt. Die große Herausforderung zur Umsetzung effektiver Sicherheitsfunktionen liegt dabei sowohl in der Definition der entsprechenden Funktion, als auch in der Auswahl der Sensorik und der Entwicklung der komplexen Sensordatenverarbeitung. Die vorliegende Arbeit beschreibt ein aktives Fußgängerschutzsystem (FSS), also eine aktive Sicherheitsfunktion für den vorausschauenden Fußgängerschutz in Kraftfahrzeugen. Zu Beginn werden die Anforderungen an die Sensorik hergeleitet, um anschließend eine Analyse der für den Einsatz im Fahrzeug geeigneten Sensoren durchzuführen. Die Aufgabe der Sensorik und der nachgeschalteten Verarbeitungseinheiten ist die leistungsstarke Erkennung von Fußgängern, also die Detektion und die Klassifikation, auf Basis von 3D- und 2D-Bilddaten bewegter Sensoren. Das vorgestellte System mit der Bezeichnung PRESUME (Pedestrian Recognition System Using a Multi-Sensor Environment) erzielt hohe Erkennungsraten bei gleichzeitig geringer Anforderung an die Rechenkapazität. Die ausgewählte Sensorik bietet darüber hinaus eine sehr hohe Detektionsrate aller kollisionsrelevanten Objekte bei sehr kleiner und kontrollierbarer Geisterzielrate. Das Ergebnis der Arbeit stellt somit ein Sensorsystem inklusive der für die Erkennung von Fußgängern erforderlichen echtzeitfähigen Bildverarbeitungssoftware dar, welches sich nach dem aktuellen Stand der Technik als eines der besten Systeme präsentiert. Die Arbeit liefert damit einen wesentlichen Beitrag zur Umsetzung einer vorausschauenden Fahrzeugfunktion zur Steigerung der Sicherheit der ungeschützten Verkehrsteilnehmer.

ABSTRACT

The reduction of accidents and injury severity is an important subject of many research and development projects in the automotive industry. It is especially important to reduce the injury severity of pedestrians in case of collisions with passenger cars. In order to increase the safety of these pedestrians the number of forward looking sensors in passenger cars is rising continuously. The aim of the sensors is to collect information about the car's environment and to trigger safety systems. Choosing the best suited sensors and an appropriate sensor data processing is a very difficult task. This work deals with an active safety system that shall reduce the injury severity of pedestrians in case of an accident. Starting with collecting the requirements for moving sensors it derives a selection of the best suited sensors for the recognition, i. e. detection and classification, of pedestrians. Afterwards, we present the developed sensor system called PReSUME (Pedestrian Recognition System Using a Multi-Sensor Environment) which consists of two optical sensors. We describe a 3D and 2D image processing chain and a sensor data fusion in order to perform the detection of obstacles and the classification between obstacles belonging to class »pedestrian« and to class »non-pedestrian«. The sensor system PReSUME achieves a high detection rate, a high correct classification rate and an adequate frame rate. It ranks among the best sensor systems for pedestrian recognition considering state of the art sensor systems. It is a perfect match for active pedestrian safety system.

1

EINLEITUNG

Gegen Ende des 19-ten Jahrhunderts wurden die ersten Automobile entwickelt. Die enorme Steigerung der Verkehrsdichte und der pro Fahrer durchschnittlich zurückgelegten Strecke führten zu einer großen und immer weiter steigenden Zahl von Unfällen mit vielen Verletzten und Getöteten. Seit ungefähr 20 Jahren sinkt jedoch die Zahl folgenschwerer Unfälle [1]. Die Gründe dafür sind zahlreich. So bieten verbesserte passive und crashaktive Systeme, allen voran Rückhaltegurte und Airbags, für alle Fahrzeuginsassen mehr Schutz als früher. Auch das Anti-Blockier-System (ABS) und das Elektronische Stabilitätsprogramm (ESP) tragen nachweislich zu einer höheren Verkehrssicherheit bei. Dennoch starben alleine in Deutschland im Jahr 2007 fast 5 000 Personen bei Verkehrsunfällen, über 430 000 wurden verletzt.

In den letzten 15 bis 20 Jahren sind auch die Zahlen der verletzten und getöteten »schwachen Verkehrsteilnehmer« im Straßenverkehr – dazu zählen hauptsächlich Fußgänger – rückläufig. Die Gründe dafür sind sowohl in einer verbesserten Infrastruktur, als auch in der konstruktiven Optimierung der Frontpartien von Personenkraftwagen (Pkw) zu finden. Beispiele für diese konstruktiven Optimierungen sind weichere Stoßfänger, rundere Fahrzeugfronten und veränderte Scheibenwischerpositionen. Die intensive Nutzung von Computern zur Crashsimulation hat einen großen Anteil an der Wirksamkeit dieser Maßnahmen. Jedoch sind die heute noch auftretenden Zahlen an Unfallopfern keineswegs vernachlässigbar¹. So sind alleine in Deutschland im Jahr 2007 fast 34 500 Fußgänger bei Verkehrsunfällen zu Schaden gekommen, 695 starben [2].

Sowohl passive als auch aktive Systeme können dazu einen Beitrag leisten. Während passive Systeme die Verletzungsgefahr für den Fußgänger beim Zusammenstoß zwischen ihm und dem Fahrzeug durch strukturelle Maßnahmen am Fahrzeug mindern, sollen aktive Systeme durch Hinzunahme von Sensoren und Aktoren den Zusammenstoß vermeiden oder, falls die Unfallvermeidung aus physikalischen oder anderen Gründen nicht möglich ist, das Verletzungsrisiko für den schwachen Verkehrs-

¹<http://www.Unfallforschung-der-versicherer.de/Unfallforschung/Publikationen/Download/Broschueren/Jahresbericht20072008.pdf>, 2008

teilnehmer so weit wie möglich reduzieren. Eine besonders herausfordernde und komplexe Aufgabe kommt dabei den Sensoren, die das Fahrzeugumfeld beobachten, und der Sensordatenverarbeitung zu, die frühzeitig Hinweise auf eine mögliche Kollision liefern sollen. Mit diesen vorausschauenden Systemen ist es prinzipiell möglich, die unfallvermeidende und die unfallfolgenmindernde Wirkung aktiver und crashaktiver Aktoren sinnvoll umzusetzen und zu optimieren.

Doch bei der Erkennung von Fußgängern auf Basis von in Fahrzeugen integrierten Sensoren treten besondere Herausforderungen auf. Auf der einen Seite sind die Sensoren starken Einschränkungen bezüglich Einbauport, Einbauwinkel und natürlich Kosten unterworfen. Es existieren nur wenige Positionen, an denen ein vorausschauender Sensor im Fahrzeug integriert werden kann. Außerdem handelt es sich um einen (mit dem Fahrzeug zusammen) bewegten Sensor. Diese Eigenbewegung erfordert die Entwicklung komplexerer Ansätze als beispielsweise in der Personenerkennung mit statischen Kameras an Bahnhöfen oder in Einkaufszonen. Auf der anderen Seite ist die Anzahl möglicher Erscheinungsformen von Fußgängern quasi unendlich groß, obgleich alle Fußgänger für den Menschen ähnlich erscheinen. Es ist also prinzipbedingt gar nicht möglich, eine vollständige Menge an Fußgängern für Trainings- oder Testzwecke zu erfassen. Die Schwierigkeit besteht vielmehr darin, die Menge so klein und doch gleichzeitig so allgemein zu halten, dass der Testumfang abprüfbar bleibt und gleichzeitig generalisierbar ist. Diese Besonderheiten der Fußgängererkennung für den Einsatz in Kraftfahrzeugen werden in der vorliegenden Arbeit adressiert.

Werden Aktoren allein auf Basis von Sensorinformationen ausgelöst, spricht man von einem *autonomen* System. Werden die Aktoren jedoch erst durch eine zusätzliche Aktivität des Fahrers angesteuert, handelt es sich um ein *fahrergetriggertes* System. Mischformen, also *teilautonome* Systeme, sind ebenfalls denkbar. Findet durch das autonome beziehungsweise teilautonome Auslösen der Aktoren ein aktiver Eingriff in das Verkehrsgeschehen statt, können neben den positiven Einflüssen auch ungewollte Nebeneffekte auftreten. So können Produkthaftungsfälle sowohl bei Fehlauflösungen als auch bei Einbeziehung Dritter in das Unfallgeschehen entstehen. Die darauf möglicherweise folgenden Produkthaftungsklagen können in den USA mit theoretisch beliebig hohen Schadenersatzforderungen einhergehen. Um dieses immens hohe Risiko für die Automobilhersteller auch in Zukunft kontrollierbar zu halten ist es wichtig, die Entwicklung aktiver Systeme, insbesondere die Entwicklung autonomer und teilautonomer aktiver Sicherheitssysteme, präzise zu definieren und sorgfältig umzusetzen.

1.1 MOTIVATION

Um dem Ziel der Reduktion der Anzahl sowohl schwerverletzter als auch getöteter Unfallopfer näher zu kommen, müssen derzeitige Schwachstellen und Missstände analysiert, neue Systeme entwickelt und getestet und mögliche Nebeneffekte dieser Systeme abgeschätzt werden. Aus Unfallstatistiken ist zu erkennen, dass die Hauptunfallverursacher bei Kollisionen zwischen Fußgängern und Fahrzeugen die Pkw sind. Daher spielt der Pkw-Fahrer eine wesentliche Rolle. Sein Verhalten, sein Verantwortungsbereich und die für ihn möglichen technischen Unterstützungen müssen berücksichtigt werden. Für das Jahr 2010 wird prognostiziert, dass die Hälfte aller Premium-Segment Pkw ein oder mehrere Fahrerinformations- oder Fahrerwarnsysteme besitzen [3]. In 2015 soll sogar ein Viertel dieser Fahrzeuge die Fußgänger (und auch Radfahrer) bei einem drohenden Zusammenstoß aktiv schützen können. Vollautonome Systeme werden jedoch in den nächsten 10 bis 15 Jahren voraussichtlich noch keinen großen Marktanteil erlangen.

Es existieren EU-Vorhaben zur Reduktion der Verkehrstoten bis 2010 [4], deren Ziel die schrittweise Reduktion der Zahl der Toten im Straßenverkehr von ca. 50 000 im Jahr 2000 auf jährlich 25 000 ab dem Jahr 2010 [5] ist. Um diesem Ziel stetig näher zu kommen, war zwischen dem Jahr 2000 und 2004 eine Reduktion um circa 20 % notwendig. Zwar haben dieses Zwischenziel bisher die Länder Deutschland, Belgien, Dänemark, Frankreich, Griechenland, Luxemburg, die Niederlande und Portugal erreicht, der EU-Durchschnitt liegt jedoch schlechter und erfüllt das Kriterium bisher nicht. Die teilweise Erreichung des Zwischenziels bei gleichzeitig steigendem Bestand an motorisierten Fahrzeugen ist sicherlich ein Indiz für die Wirksamkeit von verbesserten Sicherheitssystemen und Sicherheitskonzepten bei Pkw. Es ist zum jetzigen Zeitpunkt aber davon auszugehen, dass der mit Abstand größte Beitrag zur Umsetzung dieses Vorhabens durch die Systeme ABS und ESP geleistet wurde und auch weiterhin geleistet wird. Vorausschauende Sicherheitssysteme werden bis 2010 keinen marktrelevanten Anteil erreichen. Daher werden diese Systeme zur Steigerung der Verkehrssicherheit wohl erst in den *nächsten* Jahrzehnten eine große Rolle spielen [6].

Ein zusätzlicher Motivationsaspekt für aktive Fußgängerschutzsysteme ist die Tatsache, dass die bereits geltenden strengen gesetzlichen Regelungen der Europäischen Union für den *passiven* Fußgängerschutz durch Integration unfallvermeidender oder unfallfolgenmindernder *aktiver* Maßnahmen entschärft werden könnten, siehe dazu [7] und [8]. Die zitierten Regelungen der EU definieren Grenzen für die Verletzungsschwere an Beinen, Hüfte und Kopf bei definierten Crash-Tests. Die Festlegung der Rahmenbedingungen für diese Testunfälle berücksichtigt aber bisher lediglich

die konstruktiven Maßnahmen am Fahrzeug und den Einsatz eines fahrergetriggerten Bremsassistenten. Es werden bisher keine zusätzlichen aktiven Systeme berücksichtigt.

Weiterhin werden die Auswahl sowie die Auswahlmethoden der für die jeweilige vorausschauende Schutzfunktion idealen Sensorik kontrovers diskutiert. Diese Probleme sind bisher nicht vollständig gelöst. Hinzu kommt die Schwierigkeit, dass es sich um ein System mit bewegter Sensorik handelt. Diese Arbeit soll einen entscheidenden Beitrag leisten, diese Herausforderungen für eine Beispielfunktion des aktiven Fußgängerschutzes zu lösen.

1.2 FÖRDERPROJEKTE

Im Rahmen von bundes- und europaweiten Förderprojekten wird seit vielen Jahren versucht, die Sicherheit von Fußgängern im Straßenverkehr zu erhöhen. Zu diesen Projekten gehörte unter anderem das Förderprojekt »Sensors and System Architecture for Vulnerable road Users protection (SAVE-U)«² [9], [10], [11] und [12]. Dieses Projekt zielte auf die Entwicklung und Optimierung von Sensoren ab, um schwache Verkehrsteilnehmer zu detektieren und zu klassifizieren. Einen aktuellen Beitrag zur Sicherheit für Fußgänger und Radfahrer leistet die Initiative »Adaptive und Kooperative Technologien für den Intelligenten Verkehr (AKTIV)«³, gefördert vom Bundesministerium für Wirtschaft und Technologie. Im Teilprojekt »Schutz für Fußgänger und Radfahrer (SFR)« sollen bis 2010 Systeme entwickelt und getestet werden, die bei einem potenziellen oder akuten Unfall Schutzmaßnahmen aktivieren. Diese Maßnahmen können sowohl die Warnung aller am Geschehen beteiligter Personen, als auch den Eingriff in das Fahrverhalten des Fahrzeugs, zum Beispiel durch einen aktiven Bremseneingriff und weitere reversible sowie irreversible Schutzmaßnahmen, umfassen. Ebenso sollen im SFR Teilprojekt Testverfahren entwickelt werden, um kritische Verkehrs- und Unfallsituationen inklusive dem Einsatz aktiver Systeme reproduzierbar nachstellen zu können.

1.3 GLIEDERUNG DER ARBEIT

In Kapitel 2 wird eine Marktübersicht über bereits existierende vorausschauende Assistenzsysteme gegeben. Um anschließend in Kapitel 3 die genauen Anforderungen an ein Fußgängerschutzsystem herleiten zu können, werden Unfallstatistiken analysiert, der Unterschied zwischen Fahrerassistenz- und Sicherheitsfunktionen erklärt, gesetzliche Grenzen auf-

²<http://www.save-u.org/>

³<http://www.AKTIV-online.org>

gezeigt sowie die komplexe Problematik bei Produkthaftungsfällen erörtert. Damit werden die wichtigsten nicht sensorspezifischen Aspekte des vorausschauenden Fußgängerschutzes behandelt. Die darauf aufbauende Formulierung einer zwar beispielhaften aber überaus geeigneten Ausprägung eines aktiven Fußgängerschutzsystems führt zur Herleitung der Anforderungen an die vorausschauende Sensorik und die Sensordatenverarbeitung. Kapitel 4 liefert einen Überblick über alle im automotiven Bereich relevanten Sensoren und fasst die Analysen der Leistungsfähigkeit dieser Systeme für den vorausschauenden Fußgängerschutz zusammen. Im Verlauf der Arbeit konnten die wichtigsten sensorspezifischen Vor- und Nachteile festgestellt werden. Diese werden aufgeführt und die Auswahl des Photonic Mixer Device und der Videokamera als sinnvollste Sensorkombination begründet. Damit alleine ist schon ein überaus wertvoller Beitrag für die Automobilindustrie entstanden. Der nächste Schritt, nämlich die zur Erkennung und insbesondere zur Klassifikation von Fußgängern notwendige Bilddatenverarbeitung, ist in Kapitel 5 eingehend beschrieben. Sie stellt die zentrale Komponente der Sensordatenverarbeitung und damit den sensorspezifischen Teil des vorausschauenden Fußgängerschutzes dar. Zusätzlich wird besonderes Gewicht auf die Beschreibung der Support Vector Machines und auf die im Rahmen dieser Arbeit implementierte klassenspezifische Bestrafung gelegt. Die gesamte technische Umsetzung, also der Aufbau des Versuchsfahrzeugs, die modulare Software und die verwendeten Trainingsdaten und -algorithmen sind in Kapitel 6 erklärt. Die mit Hilfe einer neuen Testspezifikation definierten Szenariendaten stellen ebenfalls eine wichtige Komponente dar. Im Anschluss zeigt Kapitel 7 die Leistungsfähigkeit des Gesamtsystems, das die Bezeichnung PReSUME trägt⁴. Abschließend wird ein Überblick über verwandte internationale Forschungs- und Entwicklungsarbeiten gegeben. Die Arbeit wird in Kapitel 8 durch eine Zusammenfassung und einen Ausblick beendet.

⁴PReSUME steht für »Pedestrian Recognition System Using a Multi-Sensor Environment« und beschreibt die vollständige Ausprägung des im Rahmen dieser Arbeit entwickelten Systems zur Erkennung von Fußgängern

VERBREITUNG VORAUSSCHAUENDER ASSISTENZSYSTEME

Einige Automobilhersteller bieten bereits vorausschauende Systeme und crashaktive Schutzsysteme für den Fußgängerschutz an. Bei den vorausschauenden Systemen handelt es sich überwiegend um präventive und unterstützende Systeme, die dem Fahrer zusätzliche Informationen, entweder in Form von Bilddaten oder in Form von akustischen oder optischen Warnungen, über das Fahrzeugumfeld bereitstellen. In manchen Systemen ist zusätzlich eine Aktorik integriert, die in das Fahrgeschehen eingreifen kann, sofern die Sensorik ein relevantes Objekt detektiert. Bei den crashaktiven Schutzsystemen sind bisher ausschließlich aufstellbare Motorhauben im Einsatz, die durch Kontaktsensoren am vorderen Stoßfänger angesteuert werden und die Verletzungsschwere beim Aufprall auf die Fahrzeugfront verringern sollen. Im Folgenden ist eine Übersicht über die Automobilhersteller gegeben, die vorausschauende Systeme beziehungsweise aktive Aktorik anbieten.

2.1 DAIMLER

Daimler bietet ein mit Bosch entwickeltes Night-View-System an. Dieses wurde in Europa zum ersten Mal im Herbst 2005 in der S-Klasse eingesetzt. Zur Zeit ist es als Zusatzausstattung für die S-Klasse, die CL-Klasse und die E-Klasse bestellbar. Das System soll den Fahrer bei Dunkelheit unterstützen, indem die Erkennung des Straßenverlaufs und die frühe Erkennung von Fußgängern und Radfahrern durch Darstellung des Fahrzeugumfeldes in einem Display erleichtert wird.

Das Infrarot (IR)-System besteht aus zwei im Gehäuse der Frontscheinwerfer integrierten IR-Strahlungsquellen und einer an der Innenseite im Bereich des Innenspiegels montierten Kamera, die im Nahinfrarot (NIR) Spektrum empfindlich ist. Diese empfängt die von den Umfeldobjekten reflektierte Strahlung. Das Kamerabild wird in einem Acht-Zoll-Display in der Kombiinstrumententafel angezeigt (siehe Abbildung 2.1).

Das Night-View-System kann vom Fahrer aktiviert und deaktiviert werden. Voraussetzung für die Aktivierung ist Dunkelheit und eingeschaltetes



ABBILDUNG 2.1: Night-View-System einer Mercedes S-Klasse aus Sicht des Fahrers, in der Kombiinstrumententafel ist eine Verkehrsszene mit drei Fußgängern zu erkennen

Daimler	
Bezeichnung	Night-View
Art	anzeigend
Sensorik	NIR-Scheinwerfer, NIR-Kamera
Hersteller	Bosch
Modelle	S-Klasse, CL-Klasse
Preis	1 900 EUR
Serieneinsatz	seit Herbst 2005

TABELLE 2.1: Fußgängerschutzsystem Night-View bei Daimler

Abblendlicht. Die IR-Strahler schalten sich bei einer Fahrzeuggeschwindigkeit größer 15 km/h automatisch ein und bleiben beim Abbremsen bis 10 km/h aktiv. Während des Rückwärtsfahrens ist das System deaktiviert.

Gut reflektierende Objekte werden bereits in einer Entfernung von über 200 m erkannt. Bei schlecht reflektierenden Objekten verringert sich diese Entfernung.

2.2 LEXUS

Im Lexus LS 460 kommt das Advanced Pre Crash Safety System (APCS) zum Einsatz. In Europa ist dieses System seit Januar 2007 als Mehrausstattung erhältlich.

Der Fahrer soll durch das Assistenzsystem vor möglichen Gefahren ge-



ABBILDUNG 2.2: Lexus LS460

Lexus	
Bezeichnung	Advanced Pre Crash Safety System
Art	aktiv und warnend
Sensorik	76 GHz Radar, IR-Stereokamera vorne, Kamera für Gesichtfeldererkennung
Hersteller	Lexus
Modell	LS 460
Preis	ca. 8 000 EUR
Serieneinsatz	Asien: seit 2006, Europa: seit 2007

TABELLE 2.2: Fußgängerschutzsystem bei Lexus

wart und in potenziellen Gefahrensituationen unterstützt werden. Dazu werden mehrere Komponenten eingesetzt, nämlich die Erkennung von Fußgängern, die Warnung beim Verlassen der Fahrspur und eine Warnung, wenn der Fahrer in Gefahrensituationen nicht auf das Fahrgeschehen konzentriert ist. Hinweise auf die Fahreraufmerksamkeit liefert ein IR-Kamerasystem, das den Fahrer beobachtet. Das System zur Erkennung von Objekten und zur Fußgängerklassifikation basiert auf einem 76 GHz Radar, der in der Front des Fahrzeugs eingebaut ist, und einer IR-Stereokamera mit 35 cm Basisweite, die hinter der Windschutzscheibe montiert ist. Die Auslösevorgänge werden in Abhängigkeit von der detektierten Aufmerksamkeit des Fahrers parametrisiert.

Mit diesem System ist laut Hersteller eine Fußgängererkennung am Tag und in der Nacht möglich. Der Radar scannt den Bereich vor dem Fahrzeug in 40 Teilschritten. Damit ist eine Erkennung von unbelebten Hinder-

nissen im Nah- und Fernbereich möglich. Mit der IR-Stereokamera werden Fußgänger und Tiere erkannt. Die dafür notwendigen IR-Scheinwerfer befinden sich im Gehäuse der Frontscheinwerfer. Wenn das Fahrzeug eine Gefahrensituation erkennt, z.B. wenn sich ein Fußgänger auf der Straße befindet, wird der Fahrer akustisch und optisch gewarnt. Falls der Fahrer abgelenkt ist, soll er optisch, akustisch und durch einen Bremsruck gewarnt werden. Ohne weitere Fahrerinteraktion und bei drohender Kollision leitet das System laut Hersteller autonom eine Bremsreaktion mit ungefähr 6 m/s^2 ein.

2.3 BMW

Seit Ende 2005 kann das BMW Night-Vision, welches zusammen mit Autoliv entwickelt wurde, als Sonderausstattung in der 7-er Reihe bestellt werden. Seit März 2006 ist es auch für die 5-er und 6-er Reihe verfügbar. Durch das System soll der Fahrer Objekte in der Dunkelheit besser erkennen und die Unfallrate bei Nacht gesenkt werden.

Das System besteht aus einer Wärmebildkamera, die im vorderen Stoßfänger links untergebracht ist. Zum Schutz wird vor der Kamera eine beschichtete Galliumarsenid-Platte und ein feines Gitter montiert. Dieses Glas wird durch eine Kamerareinigungsdüse gereinigt. Bei Außentemperaturen unter $+5^\circ \text{ C}$ wird das Glas zusätzlich beheizt. Die Kamera hat eine Auflösung von 320×240 Pixel und kann Abstrahlungswärme im Wellenlängenbereich von $8 \mu\text{m}$ bis $14 \mu\text{m}$ erfassen. Bei Geschwindigkeiten unter 80 km/h hat die Kamera einen Öffnungswinkel von 36° . Die Bildinformationen werden auf einen 8,8 Zoll-Monitor in der Mittelkonsole angezeigt (siehe Abbildung 2.3)¹.

Bei mittleren Geschwindigkeiten wird dem Fahrer ein Bildausschnitt von 24° am Monitor angezeigt. Dieser Ausschnitt schwenkt bis zu 6° nach links oder rechts mit dem Straßenverlauf mit. Dies wird als Panning bezeichnet. Ab höheren Geschwindigkeiten kann ein digitaler Zoom aktiviert werden, der weiter entfernte Objekte im Verhältnis $1,5 : 1$ vergrößert darstellt. Das Night-Vision-System wird über einen Schalter neben dem Lichtschalter ein- und ausgeschaltet. Die Feineinstellung wird dann über das iDrive Menü vorgenommen. Ab Oktober 2008 soll das System selbsttätig kritische Situationen erkennen und ein gelbes Warnsignal im Multifunktions- und Head-Up-Display den Fahrer warnen. Fußgänger werden dann in einer Entfernung bis 100 m erkannt.

¹<http://www.7-forum.com>, zuletzt besucht im Januar 2007



ABBILDUNG 2.3: Night-Vision-System eines 7er-BMW aus Sicht des Fahrers, das Display zeigt eine Verkehrsszene mit zwei Fußgängern und einem Hund

BMW	
Bezeichnung	BMW Night Vision
Art	anzeigend und warnend
Sensorik	Ferninfrarot (FIR)-Kamera
Hersteller	Autoliv Inc.
Modelle	5-er, 6-er, 7-er Reihe
Preis	ca. 1 950 EUR
Serieneinsatz	seit Herbst 2005 bzw. Frühling 2006

TABELLE 2.3: Fußgängerschutzsystem bei BMW

2.4 CADILLAC

Cadillac hat das auf Wärmestrahlung basierende System zur Erkennung von Fußgängern und Tieren im Jahr 2004 wegen zu geringer Nachfrage vom amerikanischen Markt genommen. In Deutschland wurde es nicht angeboten.

2.5 HONDA

Honda zeigt im Modell *Legend* seit Herbst 2004 das weltweit erste Fußgängererkennungssystem, welches den Fahrer bei Annäherung an Fußgänger akustisch und visuell warnt. Das System besteht aus zwei unterhalb des vorderen Stoßfängers integrierten FIR-Kameras. Das Display für die visuelle Warnung ist über der Kombiinstrumententafel montiert. Abbildung 2.4 zeigt ein Foto aus Sicht des Fahrers. Die akustische Warnung besteht



ABBILDUNG 2.4: So sieht der Fahrer eines Honda Legend mit integriertem Intelligent Night Vision System eine Verkehrsszene mit einem Fußgänger und einem entgegenkommenden Fahrzeug.

aus zwei kurz hintereinander ertönenden Pieptönen. Die Gestalt des Fußgängers wird im Display mit einer Box umzeichnet. Der Erfassungsbereich ist mit einem Öffnungswinkel von 12° , einer Mindestentfernung von 30 m und einer Maximalentfernung von 80 m angegeben. Die Entwicklung wurde von Honda selbst in ungefähr 10-jähriger Arbeit durchgeführt.

Honda	
Bezeichnung	Intelligent Night Vision System
Art	anzeigend und warnend
Sensorik	Stereo-FIR-Kamera
Hersteller	Honda
Modell	Legend
Preis	ca. 5 000 EUR
Serieneinsatz	seit 2004

TABELLE 2.4: Fußgängerschutzsystem bei Honda

Vom Objekt werden der Abstand und die Form erfasst. Die Halbbilder der Kameras werden überlagert und es findet ein Musterabgleich statt. Aus dem resultierenden Tiefenbild wird der Abstand zwischen Fahrzeug und Objekt bestimmt. Zusätzlich wird die Größe und Form der Objekte und eine mögliche Kopf-Schulter-Partie untersucht [13]. Die Einbauposition der Kameras liegt im unteren Bereich des vorderen Stoßfänger, rechts und links schräg unterhalb des Nummernschildes. Die Funktion ist erst bei einer Fahrzeuggeschwindigkeit von über 30 km/h aktiv.

3

DER SCHUTZ VON FUSSGÄNGERN

Um die zu bewältigende Aufgabe des Schutzes von Fußgängern im Straßenverkehr zu analysieren und um ein adäquates Fußgängerschutzsystem, kurz FSS, entwickeln zu können, muss zunächst eine Bedarfsanalyse durchgeführt werden. Diese detaillierte Erarbeitung der Anforderungen konnte durch die Enge Abstimmung mit den Entwicklungsabteilungen von Audi erfolgen. Dies ist ein wesentlicher Aspekt für die Anwendbarkeit der Arbeit in der Industrie, der bei vielen stärker akademisch ausgelegten Arbeiten fehlt.

Aus den Analysen ergeben sich eine Reihe substantieller Fragen, auf die eine Antwort gefunden werden muss.

1. Welche Verkehrssituationen führen besonders oft zu Unfällen mit Fußgängern?
2. Für welche kritischen Verkehrssituationen existieren schon Assistenzsysteme?
3. Was ist der Unterschied zwischen einem Komfort- und einem Sicherheitssystem?
4. Wie darf ein autonomer Systemeingriff aus rechtlicher Sicht ausgelegt sein?
5. Sollen Fußgängerschutzsysteme als warnende Präventivsysteme oder als autonome aktive Schutzsysteme ausgelegt werden?
6. Wie können vorausschauende aktive Schutzsysteme validiert werden?

Das vorliegende Kapitel enthält Zusammenfassungen und Analysen, die zur Beantwortung dieser Fragen beitragen. Dazu liefert Abschnitt 3.1 eine Untersuchung der Unfallstatistiken zwischen Pkw und Fußgängern. In Abschnitt 3.2 wird eine Übersicht über existierende Assistenzsysteme gegeben und die besonderen Anforderungen an die Entwicklung eines Sicherheitssystems, auch im Hinblick auf rechtliche Aspekte, beschrieben. Die daraus gewonnenen Erkenntnisse liefern in Abschnitt 3.3 die Basis für

eine mögliche Ausprägung eines aktiven FSS. Abschließend befasst sich Abschnitt 3.4 mit psychologischen Aspekten, die auf die Entscheidung von Käufern für oder gegen ein Sicherheitssystem Einfluss haben können.

3.1 UNFALLFORSCHUNG ZU FUSSGÄNGERUNFÄLLEN

Um herauszufinden, welche Verkehrssituationen und Szenarien besonders großen Einfluss auf Unfälle zwischen Fußgänger und Pkw haben, wird die aktuelle Unfallstatistik in diesem Abschnitt dargestellt und analysiert.

3.1.1 *Begriffe zur methodischen Unfallforschung*

Nachfolgend werden einige in der Unfallanalyse gebräuchliche Begriffe erklärt, die häufig als Auswahlkriterien in Statistiken auftauchen und die Verletzungsschwere eines Menschen beschreiben.

Abbreviated Injury Scale (AIS)

Die Abbreviated Injury Scale (AIS) ist eine Klassifikation für Einzelverletzungen. Diese weltweit gültige Einstufung nutzt einen 7-stelligen Code, der für jede Körperregion getrennt definiert ist. Die jeweils letzte Zahl dieses Codes, also die 7. Stelle, beschreibt die Verletzungsschwere der betroffenen Körperregion. Es steht dafür eine Skala von 1 bis 6 zur Verfügung, welche die Verletzungsgrade wie folgt definiert: AIS-0 gilt als unverletzt, AIS-1 als leicht verletzt, AIS-2 bis AIS-4 als schwer verletzt und AIS-5 als schwerst verletzt. AIS-6 entspricht einer nach dem aktuellen wissenschaftlichen Stand nicht behandelbaren Verletzung, die mittelbar zum Tod führt.

Maximum Abbreviated Injury Scale (MAIS)

Der Maximum Abbreviated Injury Scale (MAIS) entspricht dem höchsten AIS der verletzten Person. Dieser Wert wird recht häufig verwendet, da die schwerste Verletzung im Normalfall auch die relevanteste ist. In den im Folgenden analysierten Statistiken wird häufig eine Mindestverletzungsschwere MAIS-2+ benutzt. Im Gegensatz zur Betrachtung aller Verletzungsschweren dient die MAIS-2+ Einstufung zur Kennzeichnung der für uns relevanten Unfälle.

Injury Severity Score (ISS)

Die Gesamtverletzungsschwere wird anhand des Injury Severity Score (ISS) klassifiziert. Der ISS wird aus der Summe der Quadrate der drei höchsten Einzel-AIS-Werte eines Verletzten gebildet. Verletzte mit einem ISS über 25 werden als polytraumatisiert definiert.

Aufprallarten

Sind bei einem Unfall ein Fahrzeug und ein Fußgänger beteiligt, so differenziert man zwischen drei verschiedenen Aufprallarten. Der Primäraufprall bezeichnet den direkten Kontakt zwischen Fußgänger und Fahrzeug. Als Sekundäraufprall bezeichnet man den Sturz des Fußgängers auf die Straße, eventuell kommt es auch noch zu einem Tertiäraufprall, bei dem der Verunglückte vom Unfall- oder einem anderen Fahrzeug an- oder überfahren wird.

3.1.2 Definition der Kategorie »Fußgänger«

In die Kategorie »Fußgänger« fallen alle Fußgänger sowie Fußgänger mit Hunden, Fußgänger mit Kinderwagen, Skiläufer, Inline Skater, Kinder auf Rollern, Schlitten oder Rollschuhen (o.Ä.) und Kinder in Kinderwagen. In [14] ist eine Definition der Kategorie »Fußgänger« gegeben, nach der folgende Personen in Fußgängerunfallstatistiken *nicht* zur Kategorie Fußgänger gehören:

»Personen, die sich arbeitsbedingt auf der Fahrbahn aufhalten oder noch in enger Verbindung zu einem Fahrzeug stehen, wie Straßenarbeiter, Polizeibeamte bei der Verkehrsregelung oder ausgestiegene Fahrzeuginsassen bei Pannen, zählen nicht als Fußgänger.«

Demnach ist die Zahl tatsächlich verunglückter Personen, die durch ein aktives Fußgängerschutzsystem geschützt werden könnten, höher.

3.1.3 Fußgängerunfallstatistik Deutschland

Tabelle 3.1 zeigt eine Übersicht über die bei Verkehrsunfällen beteiligten Personen. Dabei wird zwischen beteiligt (verunglückt oder nicht verletzt) und verunglückt (verletzt oder getötet) differenziert. Die Tabelle zeigt weiterhin auf, wie stark Fußgänger an Unfällen mit Pkw beteiligt sind. Laut dieser Statistik haben Fußgänger einen höheren Anteil bei den getöteten Personen als bei den Verletzten. Dies zeigt, dass Fußgänger wenig geschützt sind.

Die Statistiken aus Tabellen 3.2 und 3.3 zeigen, dass der Unfallschwerpunkt innerhalb von Ortschaften liegt. Das Verhältnis der Unfälle mit Fußgängerbeteiligung zwischen inner- und außerorts liegt bei über 16 : 1. Dabei sind die Pkw als Hauptverursacher anzusehen, was im Übrigen als einer der wichtigsten Motivationsgründe für die Entwicklung einer vorausschauenden Sicherheitsfunktion für Fußgänger anzusehen ist. Zwischen inner- und außerorts liegt das Verhältnis der Verletzten bei 18 : 1 und bei

Auswirkung auf Person	Gesamtanzahl	davon Fußgänger insgesamt	anteilig
beteiligt	ca. 652 487	ca. 37 024	ca. 5,7 %
verunglückt	436 368	34 499	7,8 %
verletzt	431 419	33 804	7,8 %
getötet	4 949	695	14,0 %

TABELLE 3.1: Statistik für Unfälle mit Fußgängern in Deutschland für das Jahr 2007 [2]

Unfallort	Verursacher	Häufigkeit
Innerorts (gesamt: 22 466)	Fußgänger	6 947 31 %
	PKW	15 519 69 %
Außerorts (gesamt: 1 239)	Fußgänger	461 37 %
	PKW	778 63 %

TABELLE 3.2: Unfälle mit Personenschaden für 2005

Unfallort	Verursacher	Häufigkeit
Innerorts (gesamt: 281)	Fußgänger	114 41 %
	PKW	167 59 %
Außerorts (gesamt: 119)	Fußgänger	64 54 %
	PKW	55 46 %

TABELLE 3.3: Statistik über Unfälle mit Todesfolge

den Getöteten bei 7 : 3. In den innerörtlichen Szenarien stirbt nur ein geringer Teil der betroffenen Personen, nämlich ungefähr eine Person von 80. Außerhalb von Ortschaften ist dieser Anteil jedoch deutlich höher, hier stirbt jeder Zehnte nach einem Zusammenstoß mit einem Pkw. Die Zahl der getöteten Fußgänger lag 2005 bei über 650, das sind 17,7 % weniger als in 2004. Die Zahl der getöteten Kinder ist in 2005 zum ersten Mal seit 5 Jahren angestiegen, und zwar um 3,9 % auf 159.

In [1] wird eine Datenbank des Volkswagenkonzerns beschrieben, mit deren Hilfe seit 1980 Stichproben von Unfällen inklusive der genauen Erfassung der vollständigen Unfallsituation erfasst wird. Dort ist zu erkennen, dass sich der Anteil der aufgelisteten Fußgänger- und Radfahrerunfälle von 32,2% auf 50,7% Prozent erhöht, wenn das Auswahlkriterium der Verletzungsschwere auf MAIS-2+ angehoben wird. Dies beschreibt deutlich, dass die Gruppe der schwachen Verkehrsteilnehmer aus der Gruppe der Fußgänger und der Radfahrer besteht.

Typische Unfallszenen

Eine Hochrechnung auf Basis dieser Stichprobenanalyse von Daten der letzten 20 Jahre ergibt, dass der Zusammenstoß zwischen Pkw und Fußgänger mit einer Wahrscheinlichkeit von über 72 % an der Fahrzeugfront stattfindet. Dabei wurden nur solche Unfälle berücksichtigt, bei denen eine Verletzungsschwere der Stufe MAIS-2+ auftrat. Von diesen Frontunfällen ausgehend sind folgende Wahrscheinlichkeiten für die Unfallszene gegeben: Mit einer Wahrscheinlichkeit von über 86 % fährt das Fahrzeug auf einem geraden Streckenabschnitt und der Fußgänger überquert die Fahrbahn von links oder rechts. In circa 33 % der Fälle tritt der Fußgänger aus einer Teil- oder Vollverdeckung hervor. Die Menge der Unfälle in Abbiegesituationen ist dagegen recht klein und betrifft nur ungefähr 7 %. Zu nur 2,3 % bewegt sich der Fußgänger parallel zum Fahrzeug.

Auf Basis dieser Analyse erscheint es sinnvoll, die Aufmerksamkeit des Fahrers in den besonders häufig auftretenden potentiellen Unfallsituationen zu erhöhen. Diese Situationen beinhalten hauptsächlich die Geradeausfahrt mit sich seitlich näherndem Fußgänger.

Tageszeit und Helligkeit

Insgesamt liegt das Verhältnis der Unfallhäufigkeiten zwischen Tag und Nacht bei ungefähr 2 : 1. Dabei weisen die Unfälle bei Nacht eine durchschnittlich höhere Verletzungsschwere im Bereich um MAIS-4 auf.

Geschwindigkeiten

Auf der Basis von [1] wurde Abbildung 3.1 erzeugt. Diese zeigt die kumulierte Anzahl der MAIS-2+ Verletzungen über der Kollisionsgeschwindigkeit. Es ist zu erkennen, dass die typischen Aufprallgeschwindigkeiten der Pkw zwischen 10 km/h und 60 km/h liegen. Nur wenige Unfälle treten mit Aufprallgeschwindigkeiten größer als 70 km/h auf.

Bei einem Zusammenstoß mit 65 km/h liegt die Überlebenswahrscheinlichkeit des Fußgängers bei ungefähr 30 %. Im Zusammenhang mit der Wahrscheinlichkeit für einen Zusammenstoß bei dieser und noch höherer Geschwindigkeiten führt dies zu dem Schluss, dass eine Reduktion der Aufprallgeschwindigkeit außerhalb des schraffierten Bereiches sehr wenig Einfluss auf die Zahl der Verletzten und Getöteten hat. Ebenso hat ein Aktionskonzept für den Geschwindigkeitsbereich unterhalb des schraffierten Bereiches geringe Auswirkung, da dort nur circa 7 % MAIS-2+ Verletzungen auftreten [15]. Der Einfluss der Reduktion der Kollisionsgeschwindig-

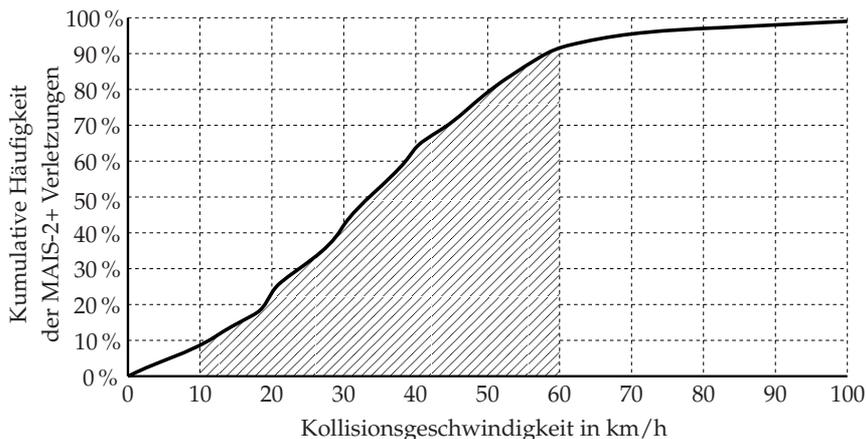


ABBILDUNG 3.1: Unfallschwere bei PKW-Fußgänger-Unfällen in Abhängigkeit der Kollisionsgeschwindigkeit [16]

keiten bei PKW-Fußgänger-Kollisionen wurde im Rahmen des EU-Projektes SAVE-U analysiert [16]. Das Ergebnis ist in Abbildung 3.1 gezeigt.

Kinder

Trotz der häufigen Hinweise auf die Zahlen über verletzte und getötete Kinder im Straßenverkehr [17] macht diese Gruppe nur einen Anteil von knapp 8 % aus. Dies zeigt auch die Angabe über Körpergrößen der verletzten Personen. Ungefähr 12 % der Personen sind kleiner als 1,40 m.

Bremsreaktionen

In ungefähr 50 % der Fußgängerunfälle wurde das Fahrzeug nicht abgebremst [1]. Grund für diese Inaktivität ist häufig, dass der Fahrer die Verkehrssituation nicht als kritisch wahrnimmt. Laut [18] sind bei fast der Hälfte aller Auffahrunfälle die Fahrer unaufmerksam und bremsen daher nicht. Dies ist ein signifikanter Hinweis auf die Notwendigkeit einer Weiterentwicklung bisheriger Sicherheitssysteme.

3.1.4 Fußgängerunfallstatistik Europa

Kollisionen zwischen Fahrzeugen und Fußgängern sind für bis zu 10 000 Todesfälle pro Jahr in den Ländern der Europäischen Union verantwortlich¹ [5]. Bei insgesamt rund 40 000 bis 50 000 Verkehrstoten zählt die Grup-

¹Statistiken von Verkehrsunfällen von 2000 bis 2002 in Deutschland, <http://www.fussgaengerschutz.net>, 2003

pe der Fußgänger zu der am zweitstärksten betroffenen Gruppe. Die Gruppe der ungeschützten Verkehrsteilnehmer insgesamt, also Fußgänger, Radfahrer und Motorradfahrer, sind mit ca. 30 % vertreten (Stand 2003). Zwar sinkt die Zahl der getöteten Verkehrsteilnehmer kontinuierlich, jedoch stagniert die Zahl der Verletzten. Tabelle 3.4 gibt einen quantitativen Überblick.

Auswirkung auf Person	Anzahl insgesamt	Fußgänger insgesamt	Anteil	Zweiräder insgesamt	Anteil
verletzt (1997)	1 677 328	155 151	9,3 %		
verletzt (2001)	ca. 1,9 Mio	ca. 185 000	9,7 %	ca. 465 000	24,5 %
getötet (1998)	42 699	6 618	15,5 %	2 306	5,4 %
getötet (2001)	ca. 48 000	ca. 9 000	18,8 %	ca. 10 000	20,8 %

TABELLE 3.4: Statistik für die Europäische Union (EU-25) [5]

3.1.5 Fußgängerunfallstatistik USA

Sowohl die Anzahl getöteter, als auch die Anzahl verletzter Fußgänger sind in den USA seit 15 Jahren rückläufig. Tabelle 3.5 zeigt eine Zusammenfassung für das Jahr 2003.

Auswirkung auf Person	Anzahl insgesamt	Fußgänger insgesamt	Anteil
verletzt (2003)		69 949	
getötet (2003)	ca. 36 500	4 749	13 %

TABELLE 3.5: Statistik USA

3.1.6 Weltweite Fußgängerunfallstatistik

Die World Health Organization (WHO) berichtet, dass im Jahr 2002 fast 1,2 Millionen Menschen weltweit bei Verkehrsunfällen starben. Die Zahl der Verletzten liegt nach Schätzungen zwischen 20 und 50 Millionen. Eine weitere Abschätzung² führt zu 760 000 getöteten Fußgängern pro Jahr. Der Anteil der getöteten Fußgänger an allen getöteten Unfallbeteiligten beträgt damit über 63 %, er unterliegt aber starken länderspezifischen Schwankungen. Während in Deutschland rund 13% der Getöteten im Straßenverkehr

²aus <http://www.worldbank.org/transport/roads/safety.htm>, zuletzt besucht im Dezember 2006

Fußgänger sind, liegt die Zahl in Thailand bei rund 47%. In Ländern, in denen weniger Personen zu Fuß am Straßenverkehr teilnehmen, liegt auch die Zahl der Unfalltoten in der Gruppe Fußgänger deutlich niedriger³. Tabelle 3.6 zeigt eine Übersicht.

Land	Fußgänger	Pkw-Insassen
Deutschland	13 %	61 %
Frankreich	12 %	63 %
Griechenland	22 %	42 %
USA	13 %	52 %

TABELLE 3.6: Verkehrstote nach Art der Verkehrsteilnahme [19]

3.2 FAHRERASSISTENZ UND SICHERHEIT

Das Ziel von Fahrerassistenzsystemen ist die Unterstützung des Fahrers bei der Erfüllung seiner Fahraufgabe. Die Verkehrssituationen, in denen eine Unterstützung geleistet werden soll, können dabei recht verschieden sein. Abhängig von diesen Verkehrssituationen werden sehr unterschiedliche Informationen des Fahrzeugumfeldes benötigt. In den folgenden Unterkapiteln werden sowohl die Fahrsituationen, also auch die existierenden und zukünftigen Assistenz- und Sicherheitssysteme kurz vorgestellt und kategorisiert. Ebenso wird die besondere Herausforderung bei der Entwicklung eines Sicherheitssystems im Hinblick auf rechtliche Aspekte beschrieben.

3.2.1 Charakteristika von Fahrerassistenz- und Sicherheitssystemen

Beginnen wir mit den fünf möglichen Fahrsituationen [20]. Abbildung 3.2 zeigt sowohl die zu entscheidenden Situationen wie auch die möglichen Übergänge. Der erste Zustand ist die »normale Fahrt«. Die in der horizontalen Ellipse liegenden Phasen »normale Fahrt«, »Warnung« und »Unfallvermeidung möglich« können abwechselnd auftreten, ohne dass ein Unfall eintritt. Erst beim Übergang in die vertikale Ellipse tritt tatsächlich ein Unfall ein. Die Phasen »Unfallvermeidung möglich«, »Unfall unausweichlich« und die »Rettungsphase« treten sequentiell und nur in der hier gegebenen Reihenfolge auf. Ein Unfall ist dort unvermeidbar. Zwischen dem normalen Fahrbetrieb (waagerechte Ellipse) und dem eigentlichen

³aus International Road Traffic and Accident Database (IRTAD), <http://www.irtad.net>, zuletzt besucht im April 2002

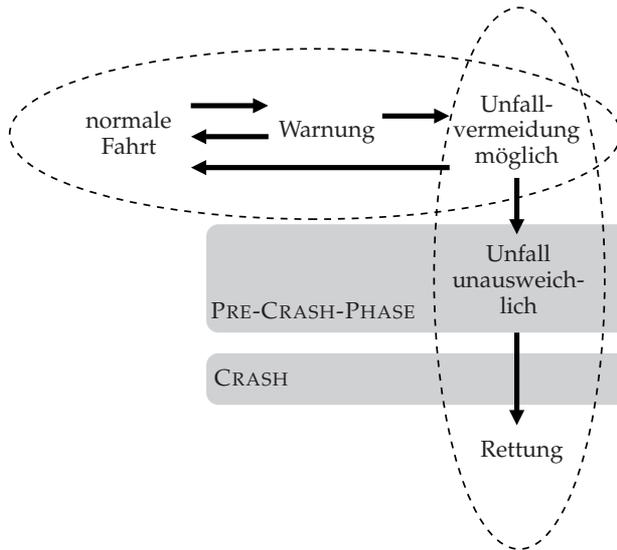


ABBILDUNG 3.2: Fahrsituationen und mögliche Übergänge

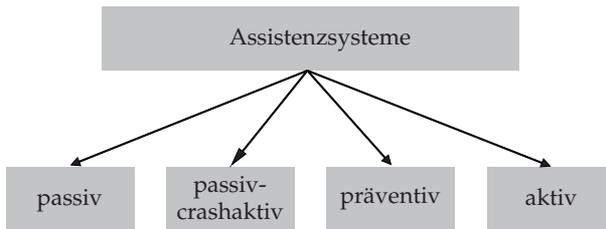


ABBILDUNG 3.3: mögliche Kategorien von Sicherheitssystemen – präventive und aktive Methoden können durch vorausschauende Sensorik gesteuert werden

Unfallereignis liegt eine Phase, die als »Pre-Crash-Phase« bezeichnet werden kann.

Es existieren viele Ansätze, eine Klassifizierung der Sicherheitssysteme in verschiedenen Systemmethoden vorzunehmen. Jedoch sind die Einteilungen oft recht unterschiedlich, da eine präzise Beschreibung der Sicherheitssysteme notwendig ist. Aus diesem Grund wird hier noch einmal eine Einteilung vorgenommen. Die in dieser Arbeit verwendeten Kategorien sind in Abbildung 3.3 gezeigt. Sie teilen die Systeme in passive, passiv-crashaktive, präventive und aktive Methoden ein. Aktive Systeme unterscheiden sich insofern von passiven Systemen, als dass sie schon im Vorfeld auf gefahrenbringende Situationen reagieren können. Präventive Sy-

steme sind selbständig arbeitende Systeme, die frühzeitig Aktionen auslösen können, jedoch keinen aktiven Einfluss auf das Fahrverhalten haben.

Schauen wir nun auf die Beschreibung relevanter Fahrzeugsysteme aus dem Bereich Komfort und Sicherheit. Einige Systeme existieren bereits⁴, andere befinden sich derzeit im Entwicklungsstadium.

GRA: Die Geschwindigkeitsregelanlage (GRA) regelt das Fahrzeug auf eine gesetzte Zielgeschwindigkeit. Weitere Bezeichnungen des Systems sind Speed Control oder Cruise Control.

ACC: Das Adaptive Cruise Control (ACC) System ist eine Erweiterung der GRA und reagiert auf vorausfahrende Fahrzeuge. Das Ziel ist es, eine vom Fahrer vorgegebene Zeitlücke zwischen dem vorausfahrenden und dem eigenen Fahrzeug zu erreichen. Dazu finden Eingriffe in die Motorsteuerung und die Bremse statt. Die Regelgröße ist die Beschleunigung. Synonyme dafür sind die Automatische Distanzregelung (ADR) oder die Abstandsregelanlage (ARA).

ESP: Das Elektronische Stabilitätsprogramm (ESP) greift bei Überschreitung der erreichbaren Seitenführungskräfte automatisch in das Bremssystem und das Motormanagement ein. Ein verwandter Begriff ist Dynamic Stability Control (DSC).

ABS: Das Anti-Blockier-System (ABS) verhindert das Blockieren der Räder, so dass das Fahrzeug auch bei einer Vollbremsung lenkbar bleibt. Der englische Begriff ist Antilock Braking System.

ASR: Wird ein Schlupf an einem Rad festgestellt, so greift die Anti-Schlupf-Regelung (ASR) in die Motorsteuerung ein und reduziert die Motorleistung.

EDS: Die Elektronische Differentialsperre (EDS) ist ein Teil des ESP und der ASR. Sie sorgt dafür, dass bis zu einer Fahrzeuggeschwindigkeit von 40 km/h beziehungsweise 80 km/h ein Rad bei Auftreten von Schlupf automatisch abgebremst wird. Der englische Begriff ist Electronic Differential Lock (EDL).

ASA: Das System Audi Side Assist (ASA) warnt den Fahrer, wenn er mit seinem Fahrzeug auf eine Fahrspur wechseln will, auf der sich ein von hinten schnell näherndes anderes Fahrzeug befindet. Eine international benutzte Bezeichnung ist unter Anderem Lateral Control Assistant.

⁴aus dem Internet Lexikon der AUDI AG, Übersicht der Fahrzeugsysteme, <http://www.audi.de/audi/de/de2/tools/glossary.html>, 2007

ALA: Der Audi Lane Assist (ALA) erkennt mit Hilfe eines optischen Sensors die Fahrspurmarkierungen und warnt den Fahrer beim unbeabsichtigten Verlassen dieser Spur. Das englische Synonym ist Lane Keeping Assistant.

Nachtsichtsystem: Ein Nachtsichtsystem erhöht die Chance, dass der Fahrer bei Dunkelheit relevante Objekte vor seinem Fahrzeug erkennt und so bei kritischen Verkehrssituationen früher reagieren kann. Die genaue Systemausprägung kann dabei stark variieren und reicht von rein darstellenden Systemen über warnende Systeme bis zu eingreifenden Systemen. Die Funktion ist im Allgemeinen unter dem Begriff Night Vision (NV) bekannt.

APS: Das Acoustic Parking System (APS) vereinfacht das Einparken durch abstandsabhängige Warntöne.

PLA: Ein Parklenkassistent (PLA) misst während der Vorbeifahrt die Abmessungen von Parklücken aus. Möchte der Fahrer in eine Parklücke einparken, übernimmt Fahrzeug die Steuerung der Lenkung, der Fahrer kontrolliert nur die Geschwindigkeit.

Stadt-ACC: Der Funktionsumfang orientiert sich am ACC. Durch den innerstädtischen Einsatz liegt jedoch beim Stadt-ACC der relevante Geschwindigkeitsbereich bei 0 km/h bis 60 km/h. Das Aktionskonzept konzentriert sich auf langsamer fahrende und stehende Fahrzeuge, die sich im prädierten Fahrschlauch⁵ befinden.

ANB: Die Automatische Notbremse (ANB) bremst das Fahrzeug, wenn ein Auffahrunfall aus physikalischen Gründen nicht mehr zu vermeiden ist.

FSS: Unter dem Begriff Fußgängerschutzsystem (FSS) sind alle aktiven Maßnahmen zu sehen, die in dieser Arbeit anhand der Beispielapplikation aus Abschnitt 3.3 vorgestellt werden. Dadurch soll die Sicherheit eines Fußgängers beim Zusammenstoß mit einem Kraftfahrzeug erhöht werden. Im Fokus steht dabei der Frontalzusammenstoß im innerörtlichen Verkehr. Durch die Verwendung von vorausschauenden Sensoren kann eine mehrstufige Auslösestrategie eingesetzt werden, die zu verschiedenen Zeitpunkten vor dem prädierten Zusammenstoß Fahrerwarnungen oder Bremsaktionen vorsieht.

Man kann alle diese Systeme in eine Matrix eintragen, die durch die zuvor beschriebenen Kategorien »Fahrsituationen« und »Assistenzmethode«

⁵Der prädierte Fahrschlauch ist der Bereich, der als wahrscheinlichster Pfad vom Fahrzeug in den folgenden Sekunden überfahren wird.

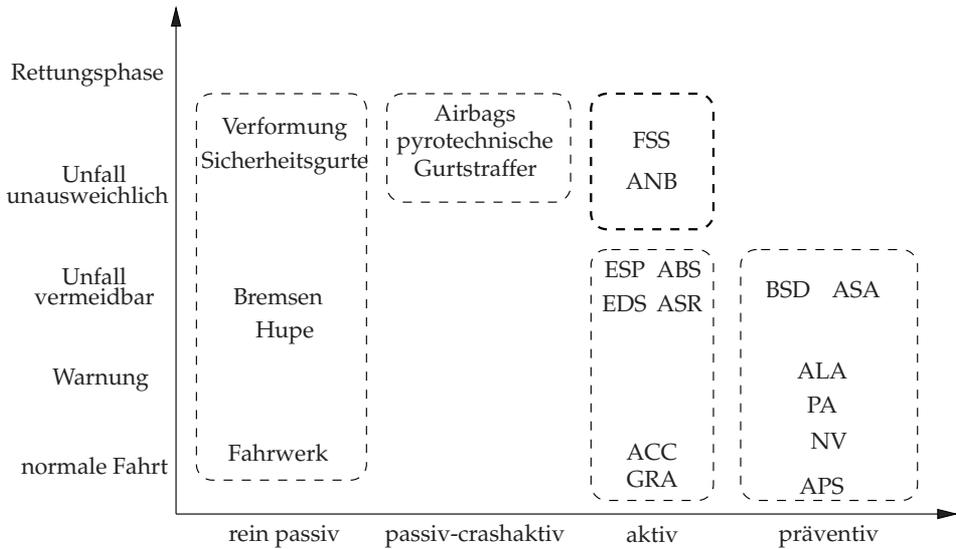


ABBILDUNG 3.4: Übersicht einiger präventiver, passiver und aktiver Systeme und Vorrichtungen mit Einstufung in die entsprechenden Fahrsituationen

aufgespannt ist (siehe Abbildung 3.4). Dabei ist zu erkennen, in welchen Bereichen die bisher eingesetzten Systeme ihre Wirkung entfalten. Die passiven Systeme decken alle Fahrsituationen ab, die passiv-crashaktiven Systeme konzentrieren sich per Definition auf das Unfallereignis. Aktive und präventive Systeme bieten bisher jedoch hauptsächlich für die Situationen Schutz, die laut Abbildung 3.2 in den Bereich der möglichen Unfallvermeidung fallen. Erst die beiden Applikationen Fußgängerschutzsystem (FSS) und Automatische Notbremse (ANB) erstrecken ihren Wirkungsbereich vollständig in das Unfallereignis selbst hinein. Das bedeutet, dass die heutigen passiven Systeme bereits den Fahrer bei einem Unfall schützen, die aktiven Systeme bisher aber nur dabei helfen, den Unfall zu vermeiden. Gründe dafür liegen in der Tatsache, dass sowohl den passiven als auch den aktiven Systeme die Informationen über den bis zum Zusammenstoß verbleibenden Zeitraum, die prädizierte Unfallschwere, die Art des Unfallgegners und die Kritikalität der Verkehrssituation nicht zur Verfügung stehen.

Trägt man nun die Systemautonomie, also die Stärke des Einflusses des Fahrers auf den Systemeingriff, in Abhängigkeit der Fahrsituationen auf, so erkennt man in Abbildung 3.5 folgendes: Je kritischer die Fahrsituation ist, desto weniger Einfluss hat der Fahrer auf die weitere Fahrzeugdynamik und auf die Auslösung einzelner Aktoren. Dies hat durchaus seine

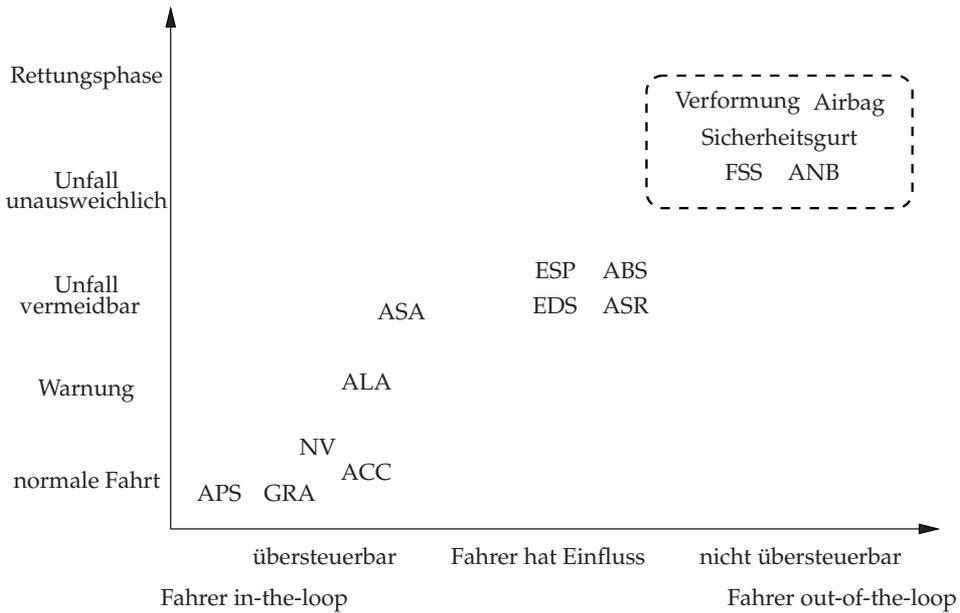


ABBILDUNG 3.5: Kategorisierung der Systeme in Abhängigkeit der Fahr-situation und der Systemautonomie

Berechtigung, denn die Systeme *müssen* dem Fahrer einen Teil seiner Aufgaben abnehmen, für die er bis zu diesem Zeitpunkt verantwortlich war, denn der Fahrer kann die Fahrsituation – offensichtlich – mit zunehmender Kritikalität beziehungsweise mit abnehmender Zeitspanne bis zum prädi-zierten Unfallzeitpunkt immer weniger beherrschen.

Dem Kunden werden beim Kauf eines Fahrzeugs Komfort- und Sicher-heitssysteme angeboten. Dies geschieht entweder über Serienausstattun-gen (besonders bei den passiven Sicherheitssystemen) als auch über Mehr-ausstattungen (häufig als Komfortsysteme bezeichnet). Dabei bezieht sich der Begriff Komfort nicht auf den *Fahr*komfort. Alle hier diskutierten Assi-stenzsysteme sind nämlich darauf ausgerichtet, entweder Unfälle zu ver-meiden oder Unfallfolgen zu minimieren. Das heißt nichts anderes, als dass die Sicherheit erhöht wird. Der Fahrer soll sich bei einem Teil sei-ner Aufgaben unterstützen lassen. Die Eingriffe ins Fahrverhalten sind dabei im Allgemeinen recht moderat. Der Fahrer kann jederzeit das Sys-tem übersteuern. Im Normalfall können diese Systeme auch deaktiviert werden. Beispiele für diese Komfortsysteme sind die Geschwindigkeitsre-gelanlage (GRA), das Adaptive Cruise Control (ACC), der Audi Side Assi-st (ASA), der Audi Lane Assist (ALA) und bekannte Ausprägungen von Nachtsicht-Systemen.

Demgegenüber steht hier der Begriff des Sicherheitssystems. Gerät das Fahrzeug in einen kritischen Fahrzustand, so sind deutlich drastischere Maßnahmen möglich und sinnvoll. Es können sowohl reversible als auch irreversible Maßnahmen eingeleitet werden. Sicherheitssysteme wirken in den Bereichen, in denen Fahrer entweder nie oder zumindestens nicht in der kurzen Zeit der Pre-Crash-Phase Einfluss auf die Fahrdynamik hat. Ein Beispiel dafür ist das ABS. Obwohl der Fahrer den Bremsvorgang kontrolliert, kann er die ABS-Funktion nicht beeinflussen. Er kann daher nicht jederzeit das System übersteuern. Er ist zwar noch in den Auslösevorgang integriert, aber die Art des Eingriffs entzieht sich seinem Einfluss. Weitere Beispiele sind auch das ESP und jegliche Ausprägung einer ANB und eines FSS. Bei Sicherheitssystemen gilt es also, dem Fahrer einen Teil seines Verantwortungsbereiches temporär abzunehmen oder eine teilautonome Funktion hinzuzufügen. Dadurch trägt das Sicherheitssystem einen Teil der Gesamtverantwortung.

Die Entwicklung von Assistenzsystemen und Sensorik lief bisher parallel und häufig in enger gegenseitiger Abstimmung. Dabei wurden die Anforderungen der Applikation an den jeweils zugehörigen Sensor iterativ an dessen Fähigkeiten angepasst. Dies ist immer dann möglich, wenn das Assistenzsystem als Komfortsystem ausgelegt ist, und in seinem Aufgabenbereich quasi frei definierbar ist. Man definiert das Erreichbare als Ziel. Mit dem Einsatz von Sicherheitssystemen ist dieser Ansatz kritischer zu bewerten. Die hier durch die Sensorik, die Applikation und die Aktorik zu erfüllenden Aufgaben unterliegen zum Teil vordefinierten Rahmenbedingungen. Diese Rahmenbedingungen stammen aus zwei Bereichen. Einerseits sind es gesetzliche Vorgaben über die Einhaltung von Grenzwerten, wie beispielsweise die maximal zulässige Beschleunigung des Kopfes einer Person beim Aufprall auf die Frontpartie des Fahrzeugs, andererseits leiten sich Anforderungen aus den definierten Verkehrssituationen ab, in denen die Systeme eine Verbesserung erzielen müssen. Weiterhin erschweren der Trend zur Sensordatenfusion (SDF) und die gemeinsame Nutzung von Sensoren für mehrere Fahrzeugfunktionen die enge Abstimmung zwischen Applikation und Sensorik.

Die Unterscheidung der Fahrerassistenzsysteme (FAS) in Komfort- und Sicherheitssysteme ist im Übrigen aus einem Herstellerinteresse entsprungen [21]. Sie soll das Risiko einer grundsätzlichen Haftung des Herstellers mindern, wenn ein System ausdrücklich als ein System zur Steigerung des Fahrkomforts bezeichnet ist. Jedoch muss auch diese Systemkategorie dem Fahrer bei der Erfüllung seiner Fahraufgabe dienlich sein. Als Fahraufgabe wird dabei im Allgemeinen verstanden⁶:

»[...] sich so zu verhalten, dass kein anderer geschädigt, gefähr-

⁶Paragraph 1 Absatz 2 der Straßenverkehrsordnung (StVO)

det oder mehr als nach den Umständen unvermeidbar behindert oder belästigt wird. [...]«

Somit wird grundsätzlich an alle FAS per Gesetz der Anspruch erhoben, die Sicherheit jedes anderen beliebigen Dritten »nach den Umständen« zu gewährleisten. Dies gilt also auch für die Systeme, die den Fahrer bei der Fahraufgabe lediglich unterstützen sollen. Durch den Schutz Dritter ergibt sich auch eine entsprechende Sicherheit für den Fahrer selbst.

Es besteht daher nicht unbedingt ein großer technischer Unterschied zwischen Komfort- und Sicherheitssystemen. Die Zielsetzungen dieser Systeme sind prinzipiell ähnlich. Jedoch besteht bei Sicherheitssystemen eine deutlich höhere Anforderung an die Freigabe. Wie genau diese Freigabe erreicht werden kann, muss bei jedem System von Neuem betrachtet werden. Die Freigabeanforderungen sind deshalb höher, weil die Eingriffe in das Fahrverhalten stärker und die Möglichkeiten zur Übersteuerung durch den Fahrer geringer sind.

3.2.2 Wiener Konvention

Neben den technischen Entwicklungen müssen auch rechtliche Aspekte bei der Umsetzung von Fahrerassistenzsysteme betrachtet werden. Zu diesen rechtlichen Grundpfeilern gehört unter Anderem die Wiener Konvention über den Straßenverkehr von 1968. Sie wurde in 5 Sprachen verfasst, und zwar in Englisch, Französisch, Russisch, Chinesisch und Spanisch. Durch die Ratifikation, die in Deutschland 1973 erfolgte, hat die Konvention in Deutschland den Rang eines einfachen Bundesgesetzes erlangt. Jedoch ist die deutsche Übersetzung, also der deutsche Vertragstext, nicht maßgeblich, da nur die Versionen in den 5 authentischen Sprachen relevant sind. Mit der Konvention, deren Präambel die Steigerung der Verkehrssicherheit als Ziel erklärt, verpflichten sich die Vertragsstaaten auf die Verabschiedung einheitlicher Verkehrsregeln und einheitlicher Zulassungsregeln.

Zwei Absätze aus der Wiener Konvention stellen zentrale und für die Entwicklung von Assistenzsystemen relevante Punkte dar. Die Zitate aus Artikel 8, Absatz 5

»[...] Every driver shall at all times be able to control his vehicle or to guide his animals. [...]«

und Artikel 13, Absatz 1

» [...] Every driver of a vehicle shall in all circumstances have his vehicle under control [...] to perform all manoeuvres required of him. [...]«

lassen eine erste Interpretationsmöglichkeit zu, die systeminitiierte und nicht-übersteuerbare Assistenzsysteme verbietet. Die beiden Artikel, die im 2. Kapitel der Konvention stehen, beziehen sich jedoch ausschließlich auf verhaltensrechtliche Vorgaben, also auf Aspekte, die den Fahrer betreffen⁷. Die zulassungsrechtlichen Aspekte sind in Kapitel 3 und in Anhang 5 der Konvention definiert. Dort tauchen Formulierungen wie in Kapitel 2 jedoch nicht auf. Daher ist die Auslegung von Assistenzsystemen – in Anbetracht des Stands der Technik von 1968 nicht verwunderlich – nicht eingeschränkt, sondern es muss lediglich sichergestellt sein, dass eine vertragliche Regelung bezogen auf den Fahrer mit seinen motorischen und kognitiven Fähigkeiten festgelegt ist. Daher ist eine weitere Interpretationsmöglichkeit gegeben, nach der Bauartzulassungen jeglicher Art *nicht* gegen die Wiener Konvention verstoßen und systeminitiierte, nicht-übersteuerbare Assistenzsysteme erlaubt wären⁸. Als Grenzfall dieser Interpretation wären teilautomatisierte Start-Ziel-Systeme nicht vertragswidrig, da diese Systeme im Übereinkommen nicht geregelt sind. Unter »Kontrolle des Fahrers« kann verstanden werden, dass der Fahrer vom Start bis zum Ende der Fahrt anwesend ist und durch seine Zieleingabe, zum Beispiel in das Navigationssystem, das Ziel der Fahrt selbst bestimmt, also kontrolliert. Erst fahrerunabhängige Verkehrssysteme wären nicht mehr mit der Konvention vereinbar, Assistenzsysteme mit der Möglichkeit zur Übersteuerung sind es aber auf jeden Fall. Als Beispiel sei angeführt, dass der Unterschied zwischen ABS und ANB nach dieser zweiten Interpretation nur quantitativ und daher vertragsrechtlich geklärt ist. Solange es also um den Individualverkehr geht, ist demnach eine Kontrolle, wie in Kapitel 2 gefordert, gewährleistet.

Jedoch ist diese zweite Interpretation keineswegs unumstritten. In der russischen Version der Wiener Konvention ist nämlich das Verb *to control* derart übersetzt, dass sich die Anforderungen an den Fahrer etwas präziser auf den Verlauf der Fahrt, also unter anderem auf die Längs- und Querführung, beziehen. Zur Zeit wird versucht, eine allgemein anerkannte Regelung herbeizuführen.

3.2.3 Haftung des Automobilherstellers

Unabhängig von der Wiener Konvention muss der Hersteller bei der Entwicklung seiner Systeme sicherstellen, dass er seiner Sorgfaltspflicht nachgekommen ist. Das Produkt muss also als sicher gelten. Da es in unserem

⁷Diese Unterscheidung ist deshalb zulässig, da *Verkehrsregeln* die Aufgaben des Fahrers festlegen, *Zulassungsregeln* jedoch nur das Fahrzeug betreffen.

⁸Diese Interpretation wurden von Professor Dr. Bernhard Kempen der Universität Köln in Bergisch Gladbach, Deutschland, am 15. März 2007 im Rahmen des 1. Workshops zum Thema „Rechtliche Rahmenbedingungen“ des Förderprojektes AKTIV vorgestellt.

Fall aber um *neue* Systeme geht, bei denen Aspekte wie die Funktionsfähigkeit nur in eingeschränktem Umfang getestet werden können, wird derzeit ein »Code of Practice« entwickelt, der eine Orientierungshilfe für die Entwicklung gibt. Im Allgemeinen muss dabei der Stand der Wissenschaft und Technik berücksichtigt werden. Dies ist vor allem bei Produkthaftungsfällen wichtig. In jedem Fall muss die Entwicklung nach bestem Wissen und Gewissen erfolgen, Präzedenzfälle mit Vorfällen von einzuführenden Systemen gibt es durch die Neuheit des Systeme nicht. Die vor Beginn der Entwicklung erstellten Lastenhefte – Bauteillastenhefte für die Hardware, Funktionslastenhefte für die Fahrzeugfunktion – sind die wichtigste Orientierungshilfe für abschließende Tests und die Freigabe. Können einzelne Punkte des Lastenhefts nicht eingehalten werden, müssen diese mit Begründung abgeändert werden. Der vollständige Funktionsumfang sollte immer in enger Absprache mit den Herstellern und den betroffenen Organisationen, also zum Beispiel Behörden, erfolgen. Er sollte so robust ausgelegt sein, dass die Fahrerbelastung gering bleibt. Umfassende Studien mit Probanden müssen helfen, eine Systemauslegung und die Mensch-Maschine Schnittstellen derart umzusetzen, dass insgesamt ein Beitrag zur Verkehrssicherheit garantiert ist.

Damit sich der Fahrer jederzeit seiner Verantwortung bewusst ist, muss er wissen und verstehen, dass das System nicht in jedem Fall einen Unfall verhindert oder die Unfallfolgen verringert. Ein Produkthaftungsfall wäre somit tendenziell eher ausgeschlossen. Jedoch kann auch der umgekehrte Fall eines automatischen Fahrzeugeingriffs ohne akute Gefahrensituation auftreten. Dieser Falsch-Positiv-Alarm kann einen Unfall mit theoretisch beliebig hoher Unfallschwere hervorrufen, eine Produkthaftungsklage ist dann nicht mehr mit Sicherheit abwendbar. Als Schutz vor dem Produkthaftungsfall bleibt nur noch ein Nachweis, dass alle Entwicklungs- und Konstruktionsschritte nach den im »Code of Practice« beschriebenen Richtlinien und nach dem aktuellem Stand der Technik erfolgt sind. Unzuverlässigkeiten, Schwächen oder andere Probleme des Systems in der Serienentwicklung oder in der Serienfertigung müssen nachweisbar abgestellt werden. Je stärker der Eingriff in das Fahrverhalten bei einem FSS ist, desto geringer muss die Wahrscheinlichkeit einer Fehlauflösung, also eines Falsch-Positiv-Alarms, sein.

3.3 MÖGLICHE AUSPRÄGUNG EINES SYSTEMS FÜR DEN AKTIVEN FUSSGÄNGERSCHUTZ (FSS)

Aus den Ergebnissen der vorangegangenen Abschnitten wird nun eine mögliche Ausprägung eines aktiven Fußgängerschutzsystem (FSS) hergeleitet.

3.3.1 *Folgerungen aus den Unfallanalysen*

Die Analyse der Unfallstatistiken zeigt, dass eine bestimmte für den aktiven Fußgängerschutz relevante Unfallsituation besonders häufig auftritt. Diese Situation lässt sich wie folgt zusammenfassen:

1. Unfallort im innerstädtischen Bereich
2. initiale Fahrzeuggeschwindigkeit kleiner 60 km/h
3. Fußgänger tritt (mit oder ohne vorherige Sichtverdeckung) auf die Straße
4. Fahrer bremst nicht oder zu spät
5. Zusammenstoß zwischen Fußgänger und Pkw an der Fahrzeugfront
6. Aufprallgeschwindigkeit beim Primäraufprall 20... 50 km/h
7. Fußgänger erleidet Verletzungen der Kategorie MAIS-2+

Der Einsatz von vorausschauender Sensorik scheint in dieser Beispielsituation sinnvoll. Da typischerweise ein sehr kurzer Zeitraum zwischen der Erkennung und dem Primäraufprall liegt, muss ein adäquates Aktionskonzept für das Zusammenspiel zwischen Sensorik und Aktorik entwickelt werden.

3.3.2 *Aktionskonzept*

In diesem Abschnitt wird beispielhaft eine Realisierung eines vorausschauenden aktiven Fußgängerschutzsystem (FSS) dargestellt, welches die Verbindung von Sensorik und Aktorik zeigt. Das FSS kann Fußgänger schützen. Dazu wird online und in Echtzeit die Wahrscheinlichkeit eines Zusammenstoßes zwischen Fahrzeug und Fußgänger bestimmt. Die Funktion muss reversible und nicht-übersteuerbare Schutzmaßnahmen am Fahrzeug ansteuern, wenn definierte Schwellwerte erreicht werden. Zur Integration dieser Fahrzeugfunktion ist eine vorausschauende Sensorik erforderlich. Diese erkennt Verkehrsobjekte vor dem Fahrzeug und klassifiziert sie in die Klassen Fußgänger und Nicht-Fußgänger. Ebenso werden zu jedem Objekt Eigenschaften ermittelt. Diese Eigenschaften sind zum Beispiel die Position, Größe und Form. Mit Hilfe der Fahrzeugeigenbewegung und der Sensorinformation können mehrere Auslöseindikatoren bestimmt werden. Zu möglichen Auslöseindikatoren zählt eine Kritikalität K der Verkehrssituation, die Objektklasse S_i , die Klassifikationsgüte G_j , die prädizierte Aufprallgeschwindigkeit v_c und Time To Contact (TTC).

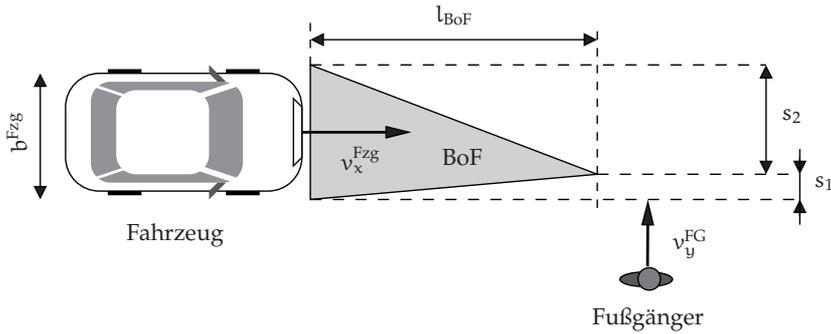


ABBILDUNG 3.6: Gelangt der Fußgänger in den BoF, so ist ein Unfall unvermeidbar; die Asymmetrie entsteht durch das Bewegungsmodell des Fußgängers

Prinzipiell gilt natürlich, dass viele verschiedene Typen von Umfeldobjekten, also Pkw, Lkw, Bebauung o. Ä., zu Unfallgegnern werden können. Natürlich ist es auch dann wünschenswert, den Unfall zu verhindern oder die Unfallfolgen zu minimieren. Genau das ist aber die Aufgabe einer Pre-Crash-Funktion. Um zu zeigen, dass eine Objektklassifikation – also unter anderem eine Unterscheidung zwischen Fußgängern und Nicht-Fußgängern – sinnvoll ist, wird die Abgrenzung zwischen FSS und einer Pre-Crash Funktion erläutert. Eine Klassifikation der Umfeldobjekte ist aber auch für diejenigen Objekte sinnvoll, für die kein Aktionskonzept aktiviert werden soll. Dies gilt zum Beispiel für alle Objekte, deren Höhe und Breite unterhalb definierter Schwellwerte liegen.

3.3.3 Berechnung der Time to Contact

Berechnen wir, ausgehend von der in Abschnitt 3.3.1 beschriebenen typischen Unfallsituation, die Größe eines Bereiches vor dem Fahrzeug, den wir als Bereich ohne Fluchtmöglichkeit (BoF) bezeichnen. Befindet sich ein Fußgänger innerhalb dieses BoF, so ist ein Zusammenstoß aus physikalischen Gründen nicht mehr vermeidbar. Die BoF Abmessungen hängen von statischen und dynamischen Größen des Fahrzeugs und von Grundeigenschaften des Fußgängers ab. Wir definieren nun beispielhaft die Bewegungsmöglichkeiten eines Fahrzeugs und nehmen eine maximale Agilität eines Fußgängers an. Somit beschreiben wir die Mindestgröße des BoF. Die Szene in Abbildung 3.6 beschreibt die verwendeten Begriffe. Die Zahlenwerte für dieses Beispiel sind in Tabelle 3.7 aufgelistet. Das verwendete Koordinatensystem wird in Abschnitt 6.1.4 beschrieben.

Fahrzeug: Ein Fahrzeug mit der Breite b^{Fzg} bewegt sich mit einer Geschwindigkeit $\vec{v}^{\text{Fzg}} = v_x^{\text{Fzg}}$ in x -Richtung.

Fußgänger: Ein Fußgänger bewegt sich mit $\vec{v}^{\text{FG}} = v_y^{\text{FG}}$ in den Fahrschlauch des Fahrzeugs hinein. Zu einem bestimmten Zeitpunkt beschleunigt er entweder nach vorne mit $\vec{a}^{\text{FG}} = a_y^{\text{FG}}$ oder nach hinten mit $\vec{a}^{\text{FG}} = -a_y^{\text{FG}}$. Somit versucht er, aus dem Gefahrenbereich zu flüchten.

Parameter	Variable	Wert
Fahrzeugbreite	b^{Fzg}	2,0 m
Geschwindigkeit des Fahrzeugs	v_x^{Fzg}	50,4 km/h
Beschleunigung des Fahrzeugs	a_x^{Fzg}	-4 m/s^2
Geschwindigkeit des Fußgängers	v_y^{FG}	2 m/s
Beschleunigung des Fußgängers	a_y^{FG}	10 m/s^2
Totzeit	t_1	0,1 s

TABELLE 3.7: Beispielwerte zur Berechnung des BoF

Der gesuchte Fall tritt genau dann ein, wenn t_c für beide Fluchtmöglichkeiten identisch ist. Einsetzen für beide Varianten in $s(t_c) = \frac{1}{2}at_c^2 + vt_c + s_0$, Auflösen nach t_c , Gleichsetzen und $s_1 + s_2 = b^{\text{Fzg}}$ ergibt

$$s_1 = \frac{1}{2}a_y^{\text{FG}}t_c^2 - v_y^{\text{FG}}t_c \quad (3.1)$$

$$\wedge s_2 = \frac{1}{2}a_y^{\text{FG}}t_c^2 + v_y^{\text{FG}}t_c \quad (3.2)$$

$$\Rightarrow t_c = +\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}} + \sqrt{\left(\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}}\right)^2 + \frac{2}{a_y^{\text{FG}}}s_1} \quad (3.3)$$

$$\wedge t_c = -\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}} + \sqrt{\left(\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}}\right)^2 + \frac{2}{a_y^{\text{FG}}}s_2} \quad (3.4)$$

$$\Rightarrow s_1 \approx 0,106 \text{ m} \quad (3.5)$$

$$\wedge s_2 \approx 1,894 \text{ m} \quad (3.6)$$

Die Zeit für die Flucht beträgt somit

$$t_c = -\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}} + \sqrt{\left(\frac{v_y^{\text{FG}}}{a_y^{\text{FG}}}\right)^2 + \frac{2s_2}{a_y^{\text{FG}}}} \quad (3.7)$$

$$\Rightarrow t_c \approx 0,447 \text{ s.} \quad (3.8)$$

Es bleiben also maximal t_c Sekunden bis zum Zusammenstoß. Wenn nun das Fahrzeug für die Zeit t_1 mit einer konstanten Geschwindigkeit v_x^{Fzg} weiterfährt und danach für die Dauer $t_2 = t_c - t_1$ mit einer konstanten Verzögerung a_x^{Fzg} abbremst, ergibt sich die Länge des BoF zu

$$l_{BoF} = v_x^{Fzg}(t_1 + t_2) + \frac{1}{2} a_x^{Fzg} t_2^2 \quad (3.9)$$

$$\Rightarrow l_{BoF} \approx 6,017 \text{ m}, \quad (3.10)$$

falls $t_1 = 0,1 \text{ s}$ angenommen wird.

Die Form des BoF entsprechend Abbildung 3.6 gilt genau dann, wenn das Fahrzeug nicht ausweicht. Wenn das Fahrzeug nicht durch Bremsen, sondern durch Ausweichen in die Richtung entgegen der Fußgängerbewegungsrichtung steuert, so wird $t_2 = 0$ gesetzt und die Beschleunigung des Fußgängers a_y^{FG} um die Querbeschleunigung des Fahrzeugs a_y^{Fzg} additiv erweitert. Dies resultiert in einem kleineren t_c und somit kleinerer BoF.

Es ist damit ersichtlich, dass bei einer prädierten Time To Contact (TTC) \hat{t} mit $\hat{t} > t_c$ nicht sicher ist, ob es überhaupt zu einem Zusammenstoß kommt. Die Aktivierung einer Schutzmaßnahme mehr als t_c vor einem möglichen Zusammenstoß, sei es die Warnung des Fahrers oder ein autonomer Fahrzeugeingriff, würde also nur auf einer Kollisionsschätzung basieren. Aus [22] ist zu entnehmen, dass die Reaktionszeit eines sehr aufmerksamen Fahrers bei $0,5 \dots 0,7 \text{ s}$, die Reaktionszeit eines übermüdeten oder abgelenkten Fahrers deutlich über $0,7 \text{ s}$ liegt. Es ist also wirkungslos, eine Fahrerwarnung bei $\hat{t} \leq t_c$ zu erzeugen. Daher wird in dieser Arbeit keine Fahrerwarnung, sondern ausschließlich der autonome Fahrzeugeingriff bei $\hat{t} \leq t_c$ propagiert. Dieser Eingriff kann durch Ausweichen oder Bremsen umgesetzt werden. Zusätzlich zum aktiven Eingriff kann eine crash-aktive Haube den Primäraufprall des Fußgängers auf das Fahrzeug dämpfen (siehe dazu Abschnitt 3.3.4).

Autonomer Bremsingriff

Ein Bremsingriff muss von mehreren Einflussgrößen gesteuert werden. Dazu zählt die Art des Umfeldobjektes entsprechend der Objektklasse S_i , die Kritikalität K , die TTC und die Kollisionsgeschwindigkeit v_c ab. So kann der Nutzen und das Risiko abgeschätzt werden, welche durch die Reduktion der Fahrzeuggeschwindigkeit zu erwarten sind. Der Nutzen der Geschwindigkeitsreduktion ist bei Unfällen mit Fußgängern innerhalb eines gewissen Geschwindigkeitsfensters sehr hoch. Der besonders wirksame Geschwindigkeitsbereich leitet sich unmittelbar aus Abbildung 3.1 ab. Eine Reduktion der Aufprallgeschwindigkeit, sofern diese im Bereich 15 km/h bis 65 km/h liegt, ist demnach wirkungsvoll. Es ist zu beachten, dass durch einen autonomen Bremsingriff – besonders bei *hohen* Geschwin-

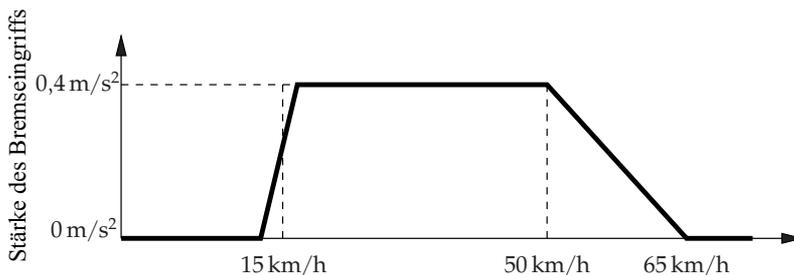


ABBILDUNG 3.7: Stärke des autonomen Bremsseingriffs in Abhängigkeit der prognostizierten Aufprallgeschwindigkeit

digkeiten – ein zusätzliches Unfall- und Verletzungsrisiko entsteht. Übersteigt dieses Risiko den zu erwartenden Nutzen, so sollte der Bremsseingriff nicht erfolgen. Abbildung 3.7 zeigt den für das FSS vorgeschlagenen Verlauf des Bremsseingriffs in Abhängigkeit der prognostizierten Aufprallgeschwindigkeit v_c . Er ist abgeleitet aus Abbildung 3.1, da diese den Zusammenhang zwischen Risiko, Nutzen und Aufprallgeschwindigkeit widerspiegelt.

Somit unterscheidet sich der Bremsseingriff beim FSS vom Eingriff beispielsweise bei der ANB. Dort ist der abzudeckende Geschwindigkeitsbereich deutlich höher. Unabhängig von v_c hat eine Reduktion der kinetischen Energie einen positiven Einfluss auf die Unfallschwere bei der ANB. Lediglich bei sehr niedrigen Geschwindigkeiten, zum Beispiel beim Einparken, ist eine Bremsung nicht erforderlich. Natürlich gibt es auch bei der ANB einen Zusammenhang zwischen Ausgangsgeschwindigkeit und Risiko für eine automatische Bremsung, jedoch wird die Parametrierung anders gestaltet sein.

3.3.4 Existierende alternative Aktionskonzepte

Im Folgenden werden bereits existierende alternative Aktionskonzepte beschrieben.

Jaguar Seit Anfang 2006 bietet Jaguar im aktuellen Modell XK serienmäßig eine crash-aktive Fußgängerschutzfunktion mit der Bezeichnung Pedestrian Deployable Bonnet System (PDBS) an. Diese besteht aus einem Kontaktsensor und einer aktiven Haube. Der Kontaktsensor kann verschiedene Typen von Unfallgegnern unterscheiden. Im Fall eines Unfalls mit einem Fußgänger oder Radfahrer wird die Haube mit Hilfe von Pyrotechnik um bis zu 15 cm angehoben. Dieser Vorgang dauert ungefähr 50 ms, es treten Beschleunigungen der Haube bis zu 50 g auf. Da die Messungen der

Sensorik erst bei Kontakt mit dem Unfallgegner starten kann, handelt es sich hierbei nicht um ein vorausschauendes, sondern um ein crash-aktives System.



ABBILDUNG 3.8: Crash-aktives Fußgängerschutzsystem: aktive Haube beim Jaguar XK

Citroën Auch Citroën bietet im Modell C6 eine crash-aktive Fußgängerschutzfunktion auf Basis von Kontaktsensorik und sogenannter pyrotechnischer Pop-Up Haube an.



ABBILDUNG 3.9: Crash-aktives Fußgängerschutzsystem: aktive Haube beim Citroën C6

Honda Honda bietet im Modell Legend eine crash-aktive Haube an. Abbildung 3.10 zeigt den Auslösevorgang bei einer simulierten Unfallszene.

Infinity Infinity bietet in der G37-Reihe eine sensorgesteuerte Motorhaube für den Fußgängerschutz an. Das Datum für die Markteinführung dieser Modelle war 2007 (USA) und Ende 2008 (Europa).



ABBILDUNG 3.10: Crash-aktives Fußgängerschutzsystem: aktive Haube beim Honda Legend

3.4 VERKAUFSRELEVANTE ASPEKTE

Die Entwicklung von Sicherheitssystemen wird aber nicht nur durch die Analyse der Unfallstatistiken beeinflusst, es werden auch die Kaufkriterien der Kunden berücksichtigt. Es muss daher der Frage nachgegangen werden, was *Sicherheit* für den Kunden bedeutet und welche Erwartung der Fahrer an solche Systeme hat.

Bei Befragungen von Kunden über deren Kaufgewohnheiten hängt das Ergebnis im Allgemeinen stark von der Art der Befragung ab. Wurde eine repräsentative Menge der Zielgruppe befragt? Konnten sie sich frei äußern? Wussten sie, welche Mechanismen sich im Unterbewußtsein tatsächlich hinter ihren Entscheidungen verbergen?

Es ist daher fraglich, ob eine forsa-Umfrage⁹, wie sie zum Beispiel in [22] wiedergegeben ist, im besonderen Maße aussagekräftig ist. Eine dieser Umfragen ergibt, dass das Kriterium »Sicherheit« insgesamt eine Wichtigkeit von 89 % beim Kauf eines Pkw besitzt. Bei Frauen habe dieser Faktor eine höhere Priorität (94 %) als bei Männern (86 %). Personen im Alter von 45 bis 59 Jahren legen nach dem Ergebnis der Umfrage größeren Wert auf eine gute Sicherheitsausstattung (92 %) als die Altersgruppe 18 bis 29 Jahre (86 %). Das Design (58 %), der Fahrspaß (51 %) und die Motorleistung

⁹forsa ist die Gesellschaft für Sozialforschung und statistische Analyse mbH, die 1984 in Deutschland gegründet wurde [23]. Es ist eines der größten Markt- und Meinungsforschungsinstitute Deutschlands und führt sogenannte Computer-Assisted Telephone Interviewing (CATI) in Bereichen der Sozialforschung, Wahl- und Politikforschung, Medienforschung, Marketingforschung und Onlineforschung durch.

(43 %) spielen der Umfrage zu Folge eine sehr geringe Rolle bei der Kaufentscheidung [23].

Das Umfrage-Institut Emnid ist in der Media- und Meinungsforschung tätig. Bei der Frage nach der Relevanz verschiedener Kaufkriterien wird laut Emnid-Statistik aus dem Jahr 2003 [24] die Sicherheit nicht einmal auf die ersten vier Plätze gewählt. Platz 1 belegt die Umweltverträglichkeit (89 %), Platz 2 sind niedrige Reparatur- und Ersatzteilkosten (84 %), auf Platz drei wird der Kfz-Versicherungsbeitrag (82 %) geführt. Platz 4 sind Automarke und Image (40,9 %). Doch schon im Jahr 2004 änderten sich die Zahlen komplett; nun steht plötzlich das Thema Sicherheit mit 89 % auf Platz eins. In [25] wurde veröffentlicht, dass 75 % aller 1 000 befragten Autofahrer beim nächsten Autokauf mehr Wert auf eine umfangreiche Ausstattung mit elektronisch gesteuerten Sicherheitssystemen legen.

Insgesamt erwecken diese Zahlen den Eindruck, dass die Ergebnisse der einzelnen Umfragen individuell dem Institut beziehungsweise dem Unternehmen angepasst werden, das diese Befragung beauftragt. Typischerweise findet man daher in Umfragen des Allgemeiner Deutscher Automobil Club (ADAC), dass die Sicherheit beim Kauf die höchste Priorität besitzt, denn Sicherheitsstandards stehen seit der Gründung des ADAC im Vordergrund.

Trotz widersprüchlicher Aussagen und der oft klar zu erkennenden Doppelmoral kann man aber in jedem Fall festhalten, dass Autokäufer in ihrem Fahrzeug hochwertige Sicherheitssysteme als Serienausstattung ohne Aufpreis integriert haben möchten. Es wird erwartet, dass die Assistenzsysteme ihren Zweck in dem Moment erfüllen, in dem sie benötigt werden, zumal der Fahrer während normaler Fahrsituationen die Systeme nicht oder nur bedingt erleben kann [26]. Das Ziel muss es also sein, den Fahrer nicht zu bevormunden, sondern mit wenig ablenkender Wirkung »im Hintergrund« zu arbeiten. Denn auch in der Zukunft sollte der Fahrspaß auf keinen Fall verloren gehen.

Das Durchschnittsalter der Käufer von Fahrzeugen des Premiumsegments steigt. Kleine körperliche Defizite können mit Assistenzsystemen kompensiert werden und damit auch körperlich behinderten oder älteren Personen noch individuelle Mobilität ermöglichen. Daher wird die Nachfrage nach Sicherheits- und Komfortsystemen in den nächsten Jahren weiter steigen.

Laut [27] könnten 90 % aller Auffahrunfälle – mit 25 % eine der häufigsten Unfallursachen – vermieden werden, würde hinter dem Lenkrad kein Mensch, sondern ausschließlich ein »elektronischer Copilot« sitzen. Daher sollten Systeme wie das FSS entweder als Serienausstattung angeboten oder gekoppelt an ein bestehendes Komfortsystem angeboten werden, welches als Mehrausstattung verfügbar ist. In diesem Fall bietet sich zum Beispiel das Stadt-ACC an, dass eine vorausschauende Sensorik und

eine Bremsaktorik verwendet. Das Gesamtsystem würde dann nicht mehr nur auf Fahrzeuge, sondern auch auf Fußgänger reagieren, die sich in den Fahrschlauch hineinbewegen. Im Idealfall kann der Hersteller sogar ein komplettes Sicherheitspaket anbieten, welches die Zusammenarbeit aller Assistenz- und Sicherheitssysteme koordiniert. Dadurch kann eine Optimierung des Fahrer-Workloads und der verschiedenen Auslösestrategien erfolgen.

In [28] kommt man zu der Überzeugung, dass Assistenzsysteme folgende Aufgabe zu erfüllen haben, um die Anzahl an Verkehrsunfällen und die damit verknüpfte Zahl Schwerverletzter und Toter zu verringern: Die Systeme sollen dann und nur dann eingreifen, wenn der Fahrer überfordert ist. Sie sollen ihn weder ablenken noch ihm das Gefühl von scheinbar vorhandener Sicherheit geben. Sie müssen außerdem in allen Situationen robust funktionieren.

4

VORAUSSCHAUENDE SENSOREN

Nachdem in den vorangegangenen Kapiteln die Notwendigkeit von Maßnahmen zum Schutz von Fußgängern festgestellt und ein Vorschlag für ein Fußgängerschutzsystem (FSS) gegeben wurde, konzentriert sich dieses Kapitel auf die Herausarbeitung der jeweiligen Vor- und Nachteile der existierenden vorausschauenden Sensoren. Für die Entwicklung jeder vorausschauenden Funktion spielt die Frage nach der zu verwendenden Sensorik eine elementare Rolle. Die Auswahl des idealen Sensors und der entsprechenden Datenverarbeitung gestaltet sich dabei also so schwierig, dass ein beträchtlicher finanzieller und zeitlicher Aufwand betrieben werden muss. Gesucht ist dabei nicht nur der beste Sensor »auf dem Papier«, sondern ein robustes System, welches auch in der anstehenden Serienproduktion kostengünstig ist, alle automotiven Anforderungen an die Robustheit erfüllt und sowohl die Anforderungen der zu bedienenden Funktion wie auch deren Weiterentwicklungen erfüllen muss. Der hohe Kostendruck und der geringe Bauraum fordern darüber hinaus, dass neben dem vorausschauenden aktiven Fußgängerschutz auch noch andere Fahrzeugfunktionen mit ausreichend guten Sensordaten versorgt werden können. Die gleichzeitige Erfüllung dieser Art von Konglomerat und die Aufgabe, die Ergebnisse hinreichend gut mit Tests bestätigen zu können, gehört zu den großen Aufgaben der Automobilhersteller und Zulieferer im Bereich der Fahrerassistenzsysteme. Es wird damit deutlich, dass bereits eine erfolgreiche Sensorauswahl ein elementares und wertvolles Teilergebnis darstellt.

Vor diesem herausfordernden Hintergrund wurde daher eine Analyse der verfügbaren Sensoren durchgeführt. Die ersten beiden Unterkapitel 4.1 und 4.2 sind dem Photonic Mixer Device (PMD)-Sensor und der Monovideo (MV)-Kamera mit ihren Funktionsweisen und Charakteristika gewidmet, da diese beiden Sensortypen ein besonders großes Potential für die Bereitstellung der notwendigen Umfeldinformationen für ein FSS bieten. Danach werden in den Abschnitten 4.3 bis 4.9 Stereokameras, Radare, Laserscanner mit festen oder beweglichen Laserstrahlen, Infrarotkameras mit spektraler Empfindlichkeit im Nah- und Ferninfrarotbereich und Ultraschallsensoren auf die spezifischen Vor- und Nachteile betrachtet. Abschnitt 4.10 liefert abschließend eine Empfehlung.



ABBILDUNG 4.1: PMD-Kamera (ohne Strahlungsquellen), Stand 2007

4.1 PHOTONIC MIXER DEVICE (PMD)

Das Photonic Mixer Device (PMD)-System ist ein aktives optoelektronisches Sensorsystem, bestehend aus Infrarotstrahlern und Kamera. Die in der Regel aus Light Emitting Diodes (LED) aufgebauten Strahlungsquellen emittieren Infrarotstrahlung, die durch ein überlagertes Rechtecksignal der Frequenz f_{mod} moduliert wird. Die PMD-Kamera empfängt die von den Objekten in der Szene reflektierte Strahlung und erzeugt 3D-Bildinformationen, wobei jedes Pixel des Bildaufnehmers einen Entfernungs- und einen Amplitudenwert liefert. Prinzipiell handelt es sich dabei um ein Time Of Flight (TOF) Messverfahren. Abbildung 4.1 zeigt ein Muster einer PMD-Kamera aus dem Jahr 2007. Die Entwicklung des Systems wurde von der Firma PMDTechnologies¹ begonnen, deren Gesellschafter die Audi Electronics Venture GmbH (AEV) und die Firma i.f.m. Electronic GmbH sind². Schon früh wurde das Potenzial für den Einsatz im automotiven Umfeld erkannt. Mit Beginn dieser Arbeit wurde auch der Einsatz dieses Sensorsystems für den vorausschauenden aktiven Fußgängerschutz untersucht [29], [30]. Diese Ansätze wurden als erfolgsversprechend betrachtet und in der folgenden Zeit stetig weiterentwickelt.

4.1.1 Funktionsweise der PMD-Kamera

Exemplarisch ist die Funktionsweise eines Pixels in Abbildung 4.2 gezeigt. Das vom Oszillator erzeugte Rechtecksignal wird gleichzeitig für die Modulation der Strahlungsquellen und für die Ansteuerung des Pixels verwendet. Die Lichtquellen werden mit dem Rechtecksignal ein- und ausge-

¹aus <http://www.pmdtec.com>, zuletzt besucht im Januar 2008

²aus <http://www.ifm-electronic.com>, zuletzt besucht im Januar 2008

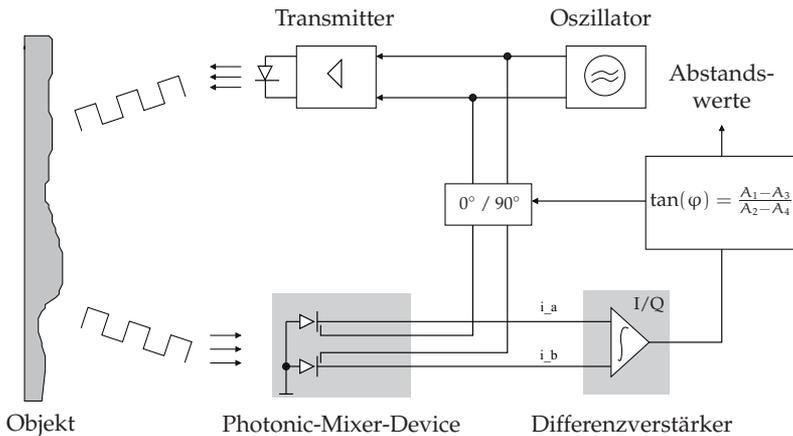


ABBILDUNG 4.2: PMD-Messverfahren eines einzelnen Pixels [31]

schaltet, an den PMD-Pixeln liegt das Oszillatorsignal in Form der beiden Spannungen $(U_0 + u_m)$ und $(U_0 - u_m)$ (siehe Abbildung 4.3) an und erzeugt ein Spannungsgefälle. Durch die Reflexion der IR-Strahlung an einem Objekt werden ein oder mehrere PMD-Pixel mit dem um φ_L phasenverschobenen Signal bestrahlt. Der innere photoelektrische Effekt erzeugt freie Ladungsträgerpaare, die getrennt in die beiden Potentialtöpfe, das sind die n^+ -dotierten Bereiche, fließen. Die Richtung des Spannungsgefälles entscheidet dabei, in welche Richtung die Elektronen wandern. Die Ladungsansammlungen in der Verarmungszone des Pixels werden also mit dem Oszillatorsignal »gemischt«. Die Ladungsmenge, also das Potenzial, wird nach jedem Zyklus bestimmt. Dies geschieht über die Integration der Ströme \bar{i}_a und \bar{i}_b im Differenzverstärker. Die Ströme \bar{i}_a und \bar{i}_b sind dabei schon das Ergebnis eines Korrelationsprozesses, bei dem das Reflektions- und das Oszillatorsignal überlagert werden. Wir bezeichnen diese Abtastwerte mit A_i [32].

Um die durch die Laufzeit des Signals resultierende Phasenverschiebung φ_L zu rekonstruieren, sind mehrere Messvorgänge notwendig. Dazu wird das vom Oszillator an jedem Pixel zur Verfügung stehende Signal in n aufeinanderfolgenden Messungen um jeweils $\varphi_O = 2\pi/n$ phasenverschoben. Jedes A_i mit $i = 1 \dots n$ liefert nun eine von n Stützstellen der Autokorrelationsfunktion des Modulationssignales. Mindestens $n = 3$ Stützstellen werden für eine eindeutige Rekonstruktion der Autokorrelationsfunktion – und somit für die Rekonstruktion des durch die Laufzeit des Signals resultierenden Phasenunterschieds φ_L – benötigt. Bei $n = 4$ erhält man einen zusätzlichen, redundanten Messwert. Die nacheinander vom Differenzverstärker bestimmten Werte A_1, A_2, A_3 und A_4 entsprechen also

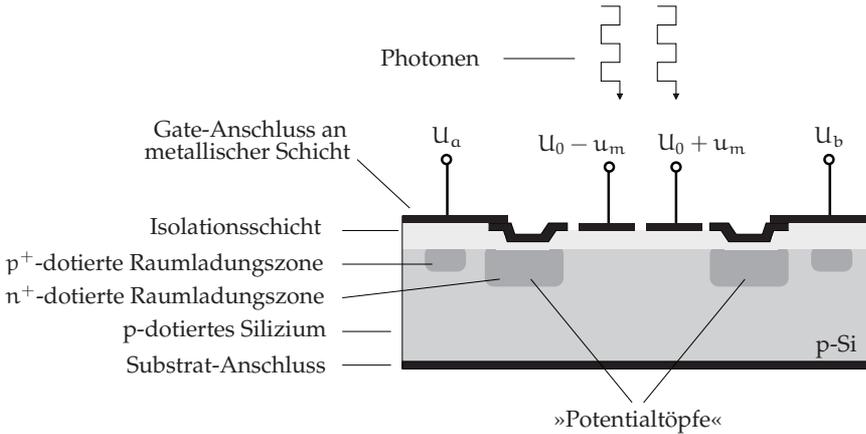


ABBILDUNG 4.3: Aufbau des Halbleitermaterials eines PMD-Pixels [32]

den Phasenverschiebungen $\varphi_O = \{0^\circ; 90^\circ; 180^\circ; 270^\circ\}$. Daraus lässt sich, wie in [33] beschrieben, die Phasenverschiebung φ_L über

$$\varphi_L = \arctan \left(\frac{A_1 - A_3}{A_2 - A_4} \right) \quad (4.1)$$

bestimmen. Über die Lichtgeschwindigkeit c kann nun die insgesamt zurückgelegte Strecke

$$R' = \frac{c}{f_{\text{mod}}} \frac{\varphi_L}{360^\circ} \quad (4.2)$$

berechnet werden. Die Distanz R zwischen einem Objekt und der Fahrzeugfront (bei identischer Einbaulage von IR-Strahlungsquellen und PMD-Kamera) ergibt sich demnach zu

$$R = \frac{1}{2} R'. \quad (4.3)$$

Der Amplitudenwert Γ eines Pixels ergibt sich aus der Summe

$$\Gamma = \sum_{i=1}^n A_i \quad (4.4)$$

aller einzelnen Werte des Differenzverstärkers. Aufgrund der Periodizität des Modulationssignals werden Phasenverschiebungen größer als 360° nicht erkannt. Deshalb liefert

$$\lambda_d = \frac{c}{2f_{\text{mod}}} \quad (4.5)$$

den maximal möglichen Distanzwert einer Messung, im Folgenden Eindeutigkeitsbereich genannt. Um zu garantieren, dass Objekte, die sich außerhalb des Eindeutigkeitsbereichs befinden, nicht fälschlicherweise in diesen hinein abgebildet werden, müssen spezielle Techniken angewendet werden, bei denen zwischen verschiedenen Modulationsfrequenzen f_{mod} umgeschaltet wird. Da kleine Frequenzen einen großen Eindeutigkeitsbereich, große Frequenzen hingegen eine erhöhte Messgenauigkeit bewirken, gibt es folgende Möglichkeiten zur gleichzeitigen Realisierung eines großen Eindeutigkeitsbereichs und einer hohen Messgenauigkeit auf Basis eines Mehrfrequenzverfahrens mit 2 Frequenzen. Entweder verwendet man alternierend eine niedrige Modulationsfrequenz $f_{\text{mod},1}$ zur Definition des Eindeutigkeitsbereiches und eine hohe Modulationsfrequenz $f_{\text{mod},2} = n \cdot f_{\text{mod},1}$ mit $n \in \mathbb{N}$ zur Erreichung einer hohen Messgenauigkeit, oder man verwendet zwei sich geringfügig unterscheidende hohe Modulationsfrequenzen $f_{\text{mod},1}$ und $f_{\text{mod},2} = (m/n) \cdot f_{\text{mod},1}$ mit $m, n \in \mathbb{N}$. In diesem Fall erreicht man eine für die Größe des Eindeutigkeitsbereiches verantwortliche Frequenz

$$f_{\text{mod},3} = \frac{f_{\text{mod},1} - f_{\text{mod},2}}{2} \quad (4.6)$$

$$\Rightarrow f_{\text{mod},3} = \left(\frac{1}{2} - \frac{m}{2n}\right) \cdot f_{\text{mod},1}. \quad (4.7)$$

Bei der richtigen Wahl von m und n ist $f_{\text{mod},3}$ deutlich kleiner als $f_{\text{mod},1}$ und $f_{\text{mod},2}$. Man erhält somit einen großen Eindeutigkeitsbereich bei gleichzeitig sehr hoher Messgenauigkeit.

Eignung

Ein großer Vorteil dieses Systems sind seine aktiven Lichtquellen. Dies erzeugt eine gewisse Robustheit gegenüber Witterungseinflüssen. Darüber hinaus bietet das Sensorsystem eine Lösung zur inhärenten Unterdrückung von Störsignalen. Diese unkorrelierten Signale resultieren hauptsächlich aus Fremdlicht und Rauschen. Die Besonderheit des PMD-Empfängers im Vergleich zu anderen Empfängersystemen besteht aus der Integration von Bandpassfilter, elektronischem Mischer und Tiefpassfilter auf Hardwareebene in einem Pixelelement. Dieses Prinzip erhöht die Robustheit insbesondere durch Einflüsse der Sonne und anderen starken Lichtquellen. Als weitere Vorteile des PMD-Systems sind die parallele Abtastung der Pixel und die hohe Bildwiederholrate zu nennen. Die im Vergleich zu einem Monovideosystem relativ geringe Gesamtzahl an Pixeln ermöglicht eine kostengünstige Sensordatenverarbeitung.

Eine wesentliche Einschränkung des Sensors ist die begrenzte Sichtweite, sofern es sich um die Detektion von kleinen Objekten handelt. Die im

Verhältnis zu bekannten Videosystemen recht geringe horizontale Auflösung bei gleichzeitig relativ großem Öffnungswinkel (zum Beispiel 64 Pixel auf 55° Öffnungswinkel) führt zu einer relativ großen Raumwinkelabdeckung, die ein Pixel in einer bestimmten Entfernung belegt. Fehler in der Entfernungsmessung entstehen genau dann, wenn das Objekt diese Fläche, genauer gesagt diesen Kugelausschnitt, nicht vollständig abdeckt. Die daraus resultierende Distanzmessung kann unter Umständen größere als dem tatsächlichen Objektabstand entsprechende Werte liefern. Ebenso führt eine zu geringe Reflektivität der Objekte zu begrenzter Sichtweite.

Ergänzende Details zum Trade-Off zwischen Auflösung, Reichweite und Empfindlichkeit sind in Anhang D näher beschrieben. Weiterführende Literatur ist in [31] und [34] zu finden.

4.2 MONOVIDEO-KAMERA

Videokameras finden schon heute in Fahrzeugfunktionen wie dem Audi Lane Assist (ALA) Anwendung. Dabei werden, wie auch in dieser Arbeit, Graustufenkameras mit einem Bildaufnehmer eingesetzt. Diese Kameraart wird im Folgenden als Monovideo (MV) bezeichnet. Es handelt sich um passive Sensorik, die keine zusätzliche Beleuchtung benötigt. Die geringe Baugröße von oft kaum mehr als 3 cm Kantenlänge des Gehäuses ermöglicht eine technisch einfache und optisch günstige Fahrzeugintegration im Bereich des Rückspiegels. Im einfachsten Fall wird dem Fahrer ein unverarbeitetes Videobild in Echtzeit dargestellt. Komplexere Funktionen, wie Spur- oder Objekterkennung, erfordern jedoch eine nachfolgende digitale Bildverarbeitung (BV).



ABBILDUNG 4.4: BlueFOX Monovideo-Kamera, Modell mvBlueFOX-120G mit Objektiv

Für die 2D-Erfassung des Fahrzeugumfeldes kommt in dieser Arbeit eine Industriekamera der Firma Matrix-Vision³ zum Einsatz, wie sie in Abbildung 4.4 gezeigt ist. Die USB-Kamera mit der Bezeichnung mvBlueFOX-120G besitzt einen Full-Frame-Charge-Coupled Device (CCD) Sensor und liefert maximal 60 Graustufenbilder pro Sekunde bei einer Video Graphics

³aus <http://www.matrix-vision.de>, zuletzt besucht im Januar 2008

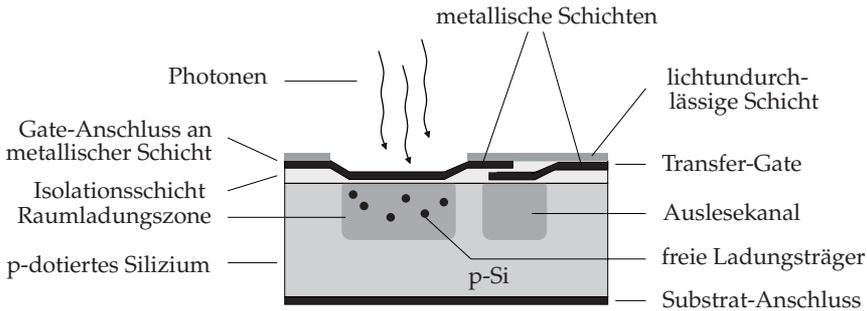


ABBILDUNG 4.5: schematischer Aufbau eines CCD-Bildaufnehmers

Array (VGA) Auflösung mit 640×480 Bildpunkten. Die Auflösung eines Bildpunktes beträgt 10 Bit. Weitere Details der Kamera sind in Abschnitt 6.1.3 und in Tabelle 6.2 beschrieben.

4.2.1 Funktionsweise eines CCD-Sensors

Bei einem CCD-Sensor bewirkt der innere photoelektrische Effekt, dass die auf einen Festkörper treffenden Photonen freie Ladungsträgerpaare, also Elektronen – Loch – Paare, erzeugen [35]. Das Halbleiterbasismaterial ist ein p-dotiertes Siliziumsubstrat, auf das eine nichtleitende SiO_2 -Schicht aufgebracht ist. Die Chipoberfläche ist mit einer aufgedampften Metallschicht versehen, die in eine Matrixstruktur unterteilt ist⁴. Jede Zelle dieser Matrix bildet ein Pixel. Da die Metallschicht eine positive Vorspannung besitzt, verdrängt sie einige Defektelektronen im oberen Bereich des Substrates. Dadurch entsteht eine Raumladungszone mit negativem Ladungsüberschuss. Fallen Photonen auf die Chip-Oberfläche, entstehen durch den photoelektrischen Effekt freie Elektronen, die von der metallischen Schicht angezogen werden. Die Elektronen sammeln sich an der Unterseite der Isolationsschicht in der Verarmungszone. Die Ladungsgröße ist dabei proportional zur gemessenen Helligkeit.

Nach der Belichtung müssen die Ladungen ausgelesen werden. Bei einem Progressive-Scan Transferverfahren sind dazu neben den Pixelelementen zusätzliche Bereiche auf dem Chip integriert, die von einer lichtundurchlässigen Schicht abgedeckt sind (siehe Abbildung 4.5). Somit können nach der Belichtung die freien Ladungsträger in den Auslesekanal verschoben werden. Da dieser Auslesebereich vor weiterer Belichtung geschützt ist, spricht man von einem elektronischen Shutter.

⁴Dieser Aufbau von Halbleitermaterial, Isolator und Metalloberfläche entspricht einer Reihenschaltung von Photodiode und Kondensator zur Ladungsspeicherung.

CCD-Chips bestechen durch ihre hohe Lichtempfindlichkeit aufgrund der geringen Anzahl lichtunempfindlicher Bauteile auf der Oberfläche und durch die hohe Bildqualität, die durch die geringe Anzahl an Verstärkerelementen erreicht wird. Das Überbelichten eines Pixel mit der Konsequenz, dass benachbarte Pixel ebenfalls überblenden, muss durch zusätzliche Elektronenabflüsse, sogenannte Overflow Drains, auf Chipebene kompensiert werden. Dieser Effekt wird als Blooming bezeichnet. Die zusätzlichen Bereiche verringern die Apertur, also den Anteil der für die Lichtaufnahme nutzbaren Chipfläche, kurz Füllfaktor. Die Nachbearbeitungsmodule benötigen im Allgemeinen relativ viel Platz, und Bildbereiche können nicht einzeln adressiert werden.

4.2.2 Vergleich mit CMOS-Bildaufnehmer

Im Vergleich zu CCD-Chips sind bei Complementary Metal-Oxide Semiconductor (CMOS)-Bildaufnehmern systembedingte Vor- und Nachteile zu finden. Bei CMOS-Bildaufnehmern ist jedes Pixel separat verdrahtet und kann daher einzeln angesteuert werden. Dadurch können viele einfache Bildoperationen parallel ausgeführt werden, vor allem die Analog-zu-Digital (A/D)-Wandlung, Glättung, Schärfung, Gammakorrektur, Weißabgleich und erste Bildverarbeitungsschritte. Jedoch verringern die zusätzlichen Bauteile das Verhältnis aus lichtempfindlicher zu lichtunempfindlicher Chipoberfläche und verringern dadurch die Empfindlichkeit. Es treten keine Bloomingeffekte auf, da die Ladung direkt für jedes Pixel einzeln in einen Spannungswert umgewandelt wird. Die herstellungsbedingten Schwankungen der Bauteileigenschaften resultieren jedoch in einer geringeren Uniformität, also einer Ungleichmäßigkeit, und somit in einer geringeren Bildqualität. Die als *Windowing* bezeichnete Möglichkeit des Auslesens einzelner Bildpunkte ermöglicht es, sehr hohe Bildwiederholraten beim Auslesen kleiner Bildbereiche zu erzielen. Grundsätzlich sind sowohl CCD- als auch CMOS-Chips für die Erfüllung der Fußgängererkennung geeignet.

4.2.3 Besonderheiten der Belichtungsregelung

Jedes Pixel eines Bildaufnehmers einer Digitalkamera liefert analoge Spannungswerte. Ein A/D-Wandler bildet diese Analogwerte über eine Quantisierungsfunktion digital ab. Da mit einer Bittiefe von 10 Bit das Quantisierungsrauschen bei konstanter linearer Quantisierung sehr groß wird, wird häufig eine Belichtungsregelung angewendet. Diese Electronic Light Control (ELC), auch als Automatic Exposure Control (AEC) bezeichnet, versucht, bei wechselnden Lichtverhältnissen den analogen Spannungspegel



ABBILDUNG 4.6: Beispielsequenz einer ungleichmäßigen, aber über der Zeit gleichbleibenden Ausleuchtung einer Szene mit Personen. Die Person am rechten Bildrand »verschwindet« durch die Belichtungssteuerung der Videokamera.

konstant zu halten und damit stets eine optimale Belichtung zu garantieren. Bei zu wenig Licht regelt die Automatic Gain Control (AGC) nach.

Bei einer Belichtungsregelung auf analoger Seite kann zum Beispiel der Istwert in Form der maximalen oder mittleren Bildhelligkeit bestimmt werden. Bei Verwendung der maximalen Bildhelligkeit wird ein einziger heller Bildpunkt den Kontrast des übrigen Gesamtbildes verringern. Verwendet man den mittleren Helligkeitswert als Regelgröße, so wird eventuell ein großer Teil des Bildes überbelichtet sein. In beiden Fällen können die Ergebnisse ungenügend sein.

In Abbildung 4.6 ist solch eine Bildfolge gezeigt, bei der zwei Fußgänger in einer Halle vor einem Fahrzeug stehen. Die Halle ist mit normaler Deckenbeleuchtung ausgeleuchtet, zusätzlich beleuchtet eine helle Lichtquelle hinter der Kamera den Aufnahmebereich⁵. Über die Bildfolge hinweg blieb die Beleuchtung konstant, lediglich die Fußgängerpositionen wurden verändert. Die Belichtungssteuerung ist auf den zentralen Bildbereich fokussiert, das heißt die Pixel in der Bildmitte werden nie übersteuert oder zu dunkel sein. Anhand dieser Sequenz kann das Problem der Überbelichtung, Unterbelichtung und Belichtungssteuerung gezeigt werden. Die Person im Vordergrund ist zuerst leicht überbelichtet (Bild 2), im weiteren Verlauf (Bild 3 und 4) durch die Belichtungssteuerung korrekt belichtet. Die Person im Hintergrund ist zu Beginn (Bild 1 und 2) kontraststark abgebildet, »verschwindet« jedoch nahezu in der Folge (Bild 3 und 4). Diese kontrastschwache Abbildung ist ausschließlich auf die Belichtungsregelung der Kamera zurückzuführen. Dieser Effekt tritt bei eingeschaltetem Abblend- oder Fernlicht oder bei tief stehender Sonne auf und verhindert, dass Personen in dunklen Bildbereichen erkannt werden können, obwohl sie eigentlich ausreichend beleuchtet sind.

Ein Optimierungsansatz ist die Regelung auf Basis der digitalen Pi-

⁵Die Lichtquelle kann in realen Situationen zum Beispiel einem Fahrzeugscheinwerfer entsprechen.

xelwerte. Dazu ist ein A/D Wandler mit einer deutlich höheren Bittiefe notwendig. Die maximale Bittiefe muss sich an der maximalen Übertragungskapazität der Architektur orientieren. Die lineare und nicht geregelte Abbildung sollte erst in einer nachfolgenden Bearbeitung mit dem Kriterium des maximalen Kontrasts und Auswahl einer Bittiefe entsprechend der Rechnerarchitektur erfolgen. Dabei können die einzelnen separat zu untersuchenden Bildbereiche einer jeweils unterschiedlichen Bitauswahl unterliegen.

4.2.4 Erreichbare Messgenauigkeit

Die hohe Auflösung von MV-Kameras, typischerweise VGA, ermöglicht eine sehr genaue laterale Messgenauigkeit \hat{y} . Diese ist als Funktion des Objektabstands x_{Objekt} , der Brennweite f und der Pixelgröße p mit

$$\hat{y} = \frac{p}{f} \cdot x_{\text{Objekt}} \quad (4.8)$$

darstellbar. Unter der Annahme einer Kamera mit $f = 3,38 \text{ mm}$ und $p = 5,3 \mu\text{m}$ (wie in 6.1.3 beschrieben) und einem Objektabstand von beispielsweise 20 m ergibt sich somit eine maximale Messgenauigkeit der lateralen Position von 3,14 cm. Die Messung von x_{Objekt} selbst ist jedoch nicht möglich, es kann immer nur eine Schätzung unter Annahme eines konkreten Fahrbahnverlaufs und eines gegebenen Fußpunktes des Objektes gemacht werden. Nimmt man an, dass der Fußpunkt exakt bestimmt wurde und die Schätzung über den Fahrbahnverlauf fehlerfrei ist, so führt eine Ungenauigkeit in der Winkelschätzung, zum Beispiel hervorgerufen durch eine Nickwinkelveränderung des Fahrzeugs um 1° nach oben beziehungsweise nach unten, und einem Objektabstand von $x_{\text{Objekt}} = 20 \text{ m}$ zu einem Schätzfehler von

$$\hat{x} \approx \begin{cases} x_{\text{Objekt}} - 4,32 \text{ m} \\ x_{\text{Objekt}} + 7,62 \text{ m} \end{cases} \quad (4.9)$$

Eignung

Die Monovideokamera liefert grundsätzlich ausreichend viele Informationen, die für eine Klassifikation notwendig sind. Sie ist genauer gesagt der einzige Sensor, der genug Informationen zur Verfügung stellt, um eine ausreichend hohe Klassifikationsgüte zu erzielen. Doch genau diese Tatsache bedeutet auch, dass die Datenverarbeitung beliebig aufwendig ist. Die ausschließliche Verwendung einer Monovideokamera zur Erkennung von Fußgängern ist daher ohne eine Reduktion der Datenmenge nicht zielführend. Da es sich weiterhin um im Raum bewegte Sensorik handelt, sind einfache Bildoperationen, die Veränderungen in den Daten der Bildfolgen

extrahieren, nicht zielführend. Die Detektion von Hindernissen ist somit um ein Vielfaches aufwendiger als mit stationärer Sensorik und größtenteils statischem Hintergrund. Weiterhin ist zu bezweifeln, dass die oben beschriebene Art der Entfernungsmessung ausreichend präzise ist.

4.3 STEREOVIDEOSENSORIK

Jede Videokamera bildet bei Einzelaufnahmen 3D-Raumkoordinaten auf 2D-Bildkoordinaten ab. Der Verlust der Entfernungsinformation kann nur durch Annahmen (etwa über die zu erwartende Größe des Objektes), nicht jedoch durch Rücktransformation ausgeglichen werden. Um mit Hilfe von 2D-Bilddaten dennoch Entfernungsmessungen vornehmen zu können, kann das Prinzip des stereoskopischen Sehens verwendet werden⁶. Dies ist am Beispiel von Stereovideosensorik erklärt. Abbildung 4.7 skizziert die Anordnung von zwei Bildaufnehmern einer solchen binokularen Abbildung. Beide Kameras beobachten die gleiche Szene, jedoch aus unterschiedlichen Perspektiven und daher mit unterschiedlichem Informationsgehalt⁷. Die korrespondierenden Bildpunkte, beziehungsweise Bildbereiche, müssen etwa mit Hilfe von Block-Matching-Verfahren gefunden und anschließend in ein einheitliches Koordinatensystem transformiert werden. Das durch die Bildaufnehmer und deren Anordnung gegebene Gleichungssystem mit zwei Gleichungen und zwei Unbekannten führt zu einer eindeutigen Lösung des Abstandes Z aus $P = [X, Y, Z]^T$. Dieser Punkt P spannt zusammen mit den Kameraursprungspunkten O_l und O_r eine Epipolarebene auf. Die Epipolarlinie der rechten Kamera ist eine Abbildung der Verbindungslinie zwischen dem Ursprung O_l der linken Kamera und dem Punkt P und umgekehrt. Das oben beschriebene Block-Matching-Verfahren muss nur auf den Epipolarlinien erfolgen, dies verringert den Rechenaufwand erheblich. Die Lage der beiden Bildaufnehmer zueinander ist frei wählbar, muss aber bekannt sein, um die Transformation durchführen zu können. Durch dieses binokulare Verfahren entsteht ein Disparitätenbild, in dem Abstandswerte zu einzelnen Punkten oder Bereichen zugeordnet werden. Das Disparitätenbild entspricht daher einem 3D-Tiefenbild.

⁶Es existieren ebenso Verfahren zur Bestimmung von 3D-Raupunkten auf Basis einer einzigen Kamera. Die Tiefeninformation kann über zeitlich aufeinander folgende Aufnahmen einer bewegten Kamera erfolgen. Dieses Verfahren wird als Structure from Motion (SfM) bezeichnet.

⁷Beide Bildaufnehmer liefern bei biokularem Sehen die gleiche, bei binokularem Sehen unterschiedliche Informationen.

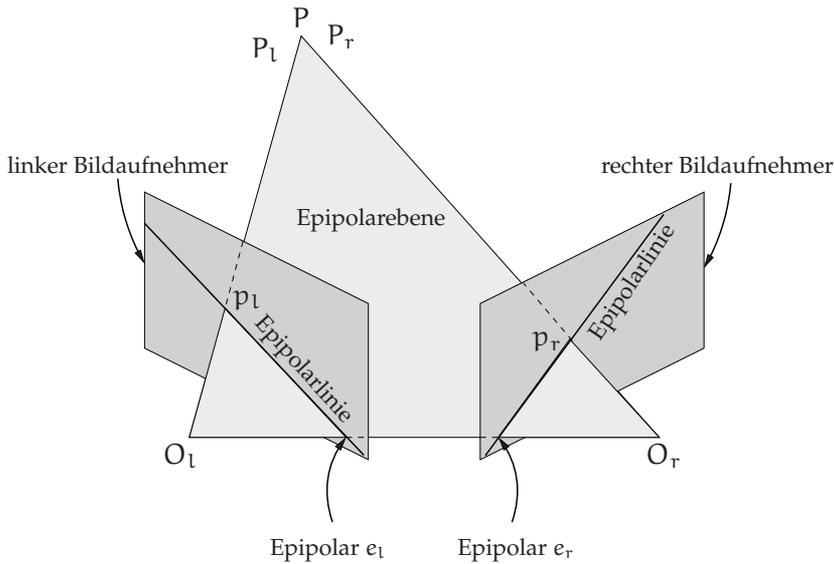


ABBILDUNG 4.7: Schematische Abbildung einer Stereokameraanordnung

Eignung

Um aus den einzelnen Raumpunkten Objekte zu bilden – und zwar ohne die Annahme von Modellen – werden Clusterverfahren angewandt. Häufig findet dazu eine Gruppierung auf Basis der räumlichen Abstände statt. Die Merkmale für die Klassifikation der Objekte, etwa die vertikale und horizontale Abmessung, die Geschwindigkeit und diverse Bewegungsmerkmale wie Varianz und Periodizität, werden in nachfolgenden Bildverarbeitungsschritten berechnet [36]. Ebenso ist auch der Objektabstand messbar, und zwar im Gegensatz zur Monovideo-Sensorik ohne Annahme eines bekannten Fahrbahnverlaufs und Kenntnis der Position des Fußpunktes. Die Merkmale der Bewegungsart, also schwingende Arme oder Beine, werden häufig als wichtige Merkmale zur Detektion und Klassifikation herangezogen. Wie auch bei Monovideo-Kameras ist eine angemessene Detektions- und Klassifikationsleistung stark abhängig von Umweltbedingungen. Die Detektion von Objekten erfordert ein erfolgreiches Block-Matching, das die BV bei schlechten Sichtverhältnissen vor eine große Herausforderung stellt.

Zwar existieren bereits verschiedene Algorithmen zur Objektbildung aus Punktwolken von Stereosystemen, jedoch sind diese Algorithmen nach Meinung des Autors bisher für den Einsatz bei Sicherheitssystemen noch nicht robust genug. In [37] werden mehrere neu entwickelte Werkzeuge und Methoden vorgestellt, um Grundeigenschaften von Stereosystemen

bewerten zu können. Zusätzlich dazu wurden prinzipielle Überlegungen angestellt, die zu der Schlussfolgerung führten, dass passive Stereosysteme für Sicherheitssysteme in den nächsten 5 bis 10 Jahren nicht geeignet sein werden.

4.4 RADAR

Seit einigen Jahren gibt es verstärkt Forschungs- und Entwicklungsarbeiten im Bereich der Fußgängerdetektion und -klassifikation durch Radar. Vorteile dieser Sensoren liegen in der Robustheit gegenüber Witterungsbedingungen und Helligkeiten. Sie können hinter Kunststoffabdeckungen, häufig im Bereich des Frontstoßfängers, montiert werden, jedoch nicht hinter metallischen Karosserieteilen. Die unterschiedlichen Verfahren wie Pulse Doppler (PD) und Continuous Wave (CW) beziehungsweise Frequency-Modulated Continuous Wave (FMCW) ermöglichen die Messung von Objektposition und -geschwindigkeit, beides vornehmlich in longitudinaler Richtung. Die hohe Messgenauigkeit liefert Positionswerte für Objektentfernungen bis zu 200 m, je nach Sensortyp, verwendeter Frequenz, Antennencharakteristik, Bandbreite und gewähltem Modulationsprinzip. Anwendung finden Radarsensoren bereits in den Funktionen ACC und ASA. Die Messfehler der Geschwindigkeitsmessungen liegen typischerweise unter 1 km/h.

Eignung

Nachteile der Sensoren in Bezug auf die Klassifikation von Fußgängern sind die geringe laterale Auflösung und die schlechte Messbarkeit lateraler Bewegungen. Im Allgemeinen besitzen sie einen relativ geringen Öffnungswinkel. Öffnungswinkel $\pm 45^\circ$ sind zwar möglich, aber es ist dann mit Einbußen im Entfernungsbereich zu rechnen.

Eine Möglichkeit zur Klassifikation von Fußgängern mit Radar funktioniert über die Analyse des Spektrums und der Periodizitäten, also der Änderungen der Messdaten über der Zeit. Diese Änderungen werden durch die Gehbewegungen der Personen, hauptsächlich messbar an den Beinen und dem Rumpf, hervorgerufen [38]. Der entscheidende Nachteil dieser Methode ist die Tatsache, dass der Fußgänger über mehrere Sekunden hinweg eine Gehbewegung in radialer Richtung zum Sensor ausführen muss. Diese Voraussetzung kann nach der im vorherigen Abschnitt erfolgten Unfallanalyse *nicht* als gegeben angenommen werden. Sie ist sogar als sehr unwahrscheinlich zu betrachten.

4.5 LIDAR

Ein Light Detection and Ranging (LIDAR)⁸ System arbeitet mit elektromagnetischen Wellen im Bereich um 800 nm und ist in der Lage, Entfernungsbestimmungen in einzelnen Strahlen durchzuführen. Typischerweise werden dazu Standardlaserdioden verwendet. Das in Abbildung 4.8 gezeigte System der Firma Hella ist vorrangig für den Einsatz als ACC Sensor konzipiert. Es arbeitet mit 16 festen Strahlen. Durch Reflexionen an Objekten und ein laufzeitbasiertes Verfahren werden 16 Entfernungswerte bestimmt. Im Steuergerät kann daraus die Geschwindigkeit und die Position eines Umfeldobjektes bestimmt werden.



ABBILDUNG 4.8: Ein LIDAR System der Firma Hella [39]

Eignung

Durch die geringe Anzahl an Strahlen kann zwar ein großes Objekt vor dem Fahrzeug erfasst, jedoch auf Grund des geringen Informationsgehaltes kaum klassifiziert werden.

4.6 LASERSCANNER

Ein Laserscanner nutzt einen rotierenden Infrarotstrahl, der das Fahrzeugumfeld abtastet. Wie beim Lidar System kann die jeweilige Objektentfernung über die Laufzeit des Signals bestimmt werden. Die Genauigkeit liegt bei wenigen Zentimetern. Abbildung 4.9 zeigt einen ALASCA Laserscanner der Firma IBEO, einer Tochter der SICK AG⁹.

⁸Der Begriff LIDAR hat mit Absicht eine gewisse Ähnlichkeit zum Wort Radar.

⁹aus <http://www.sick.com/home/de.html>, zuletzt besucht im Oktober 2006

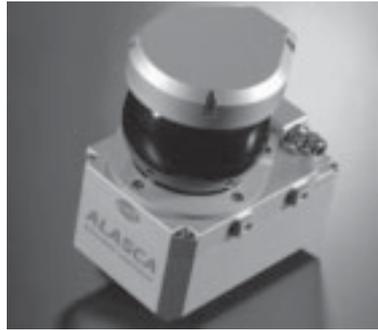


ABBILDUNG 4.9: Beispiel für einen Laserscanner der Firma IBEO

Der Erfassungsbereich erstreckt sich bei diesem Sensor in $0,25^\circ$ -Schritten über 240° horizontal und in $0,8^\circ$ -Schritten über $3,2^\circ$ vertikal. Ein Scandurchlauf dauert 100 ms.

Eignung

Durch die relativ lange Scandauer ist die Geschwindigkeit, mit der eine Szene sequentiell abgetastet wird, beschränkt. Zu diesem System existiert eine Objekterkennung, die nach [40] die Erkennung und Klassifikation von Fußgängern und anderen Umfeldobjekten ermöglicht. Fußgänger werden dort durch die horizontalen Abmessungen und durch die periodischen Beinbewegungen klassifiziert. Auch hier besteht jedoch wieder die Herausforderung, zusätzliche nicht-periodische Merkmale aus den Sensordaten zu generieren.

4.7 AKTIVE NAHINFRAROT-SENSORIK

Als aktive Nahinfrarot (NIR) Sensorik bezeichnet man ein aus IR-Strahler und Kamera bestehendes System, bei dem die Empfindlichkeit der Kamera in einem Wellenlängenbereich um 800 nm bis 900 nm und damit knapp über dem Wellenlängenbereich des sichtbaren Lichtes (380 nm bis 780 nm) liegt. Allgemein bestehen NIR-Kameras aus einem CCD- oder CMOS-Chip. Die IR-Bilder entstehen sowohl durch die Reflexionsprojektion der eigenen Strahlungsquellen als auch durch weitere, externe IR-Quellen. Beispiele für diese oft störenden IR-Quellen sind die Scheinwerfer des Gegenverkehrs, Ampelanlagen und Straßenbeleuchtungen. Alle Objekte ohne eigene Strahlungsquelle werden nur durch die Reflexion sichtbar. Dabei ist zu beachten, dass sich die Reflektivität im IR-Spektrum von der Reflektivität im sichtbaren Spektrum unterscheidet. Den höchsten Reflektionsgrad lie-



ABBILDUNG 4.10: Typische Ausgabe einer Nahinfrarotkamera mit CCD-Detektor (links) und CMOS-Detektor (rechts), am linken Bildrand ist jeweils ein Fußgänger zu erkennen

fern Retroreflektoren, zum Beispiel Verkehrsschilder und Leitpfosten. Bäume und Sträucher weisen einen hohen, die menschliche Haut, Kleidung und Fell hingegen einen schwachen Reflexionsgrad auf [41].

Eignung

Die Reichweite solcher Systeme ist von der Bestrahlungsstärke der Strahler abhängig. Deren Abstrahlung ist durch gesetzliche Vorgaben bezüglich der Augensicherheit beschränkt. Regen, Schnee und Nebel reduzieren die Reichweite des NIR-Systems durch die Absorption und Streuung der NIR-Strahlung durch die Wassermoleküle. Typische Reichweiten, je nach Art der Reichweitenbestimmung, liegen bei 80 bis 120 m.

4.8 THERMISCHE SENSORIK

Jeder Gegenstand, dessen Temperatur über dem absoluten Nullpunkt liegt, absorbiert und emittiert Strahlung entsprechend des Planck'schen Strahlungsgesetzes. Diese Strahlung hängt von der Temperatur und der Oberflächenbeschaffenheit ab. Thermische Sensoren messen die emittierte Strahlung und können somit Temperaturänderungen beziehungsweise Temperaturverläufe messen. Existiert zwischen einem Objekt und dessen Hintergrund ein Kontrast im Sinne einer unterscheidbaren spektralen Strahlungsdichte – hervorgerufen durch unterschiedliche Temperaturen – kann das Objekt detektiert werden. Der spektrale Abschnitt der Wärmestrahlung im Bereich von 8 000 nm bis 12 000 nm wird als FIR-Strahlung bezeichnet. Thermische Sensorsysteme sind passiv, sie benötigen keine eigene IR-Bestrahlung.



ABBILDUNG 4.11: Typische Ausgabe einer FIR-Kamera mit Wechselstrom-Detektor (links) und Gleichstrom-Detektor (rechts), auf der linken Straßenseite ist jeweils ein Fußgänger zu erkennen

Ursprünglich für den militärischen Einsatz konzipiert existieren heute zwei verschiedene Arten passiver IR-Systeme für den automotiven Einsatz. Die beiden Sensorprinzipien unterscheiden sich in der Art der Detektoren: Gleichstrom- oder Wechselstromdetektoren.

Bei Gleichstromdetektoren erzeugt die Wärmestrahlung eine Widerstandsänderung des betroffenen Detektorelementes. Höhere Wärmestrahlung führt zu einem Anstieg des Widerstands und damit zu einem Spannungsabfall, der durch die Kamera als hellerer Bildpunkt umgesetzt wird. Eine zyklische Kalibrierung ist jedoch notwendig, um die Widerstandskoeffizienten der einzelnen Bildpunkte zueinander abzugleichen. Diese Kalibrierungen führen zu einer kurzzeitigen Unterbrechung des Videosignals im Bereich von 0,3 s [41].

Wechselstromdetektoren nutzen hingegen kapazitive Sensorelemente und ein Flügelrad, welches abwechselnd die direkte Wärmestrahlung für ein Pixel und einen durch Mikrolinsen generierten Mittelwert auf ein Sensorelement abbildet. Die für die Darstellung des Helligkeitswertes benötigte Spannung wird nun durch die Differenz zwischen direkter und gemittelter Wärmestrahlung erzeugt. Diese Technik führt zu einer lokalen Kontrastanpassung. Gleiche Helligkeitswerte an verschiedenen Bildpunkten entsprechen nun nicht mehr, wie das bei Gleichstromdetektoren nach einer Kalibrierung der Fall ist, gleichen Temperaturen. Dafür kann der Wechselstromdetektor auch in Bildbereichen mit geringen Temperaturunterschieden kontraststarke Bilder erzeugen.

In [42] wird ein Sensorarray aus Thermopiles für den Einsatz zur Fußgängererkennung in Kraftfahrzeugen beschrieben. Die Sensorelemente nutzen den Seebeck-Effekt, der eine Temperaturdifferenz in eine Spannungsdifferenz wandelt. Der Erfassungsbereich erstreckt sich auf eine Messreich-

weite von ungefähr 8 m bei einem Öffnungswinkel von 10° . Um eine Entfernungsmessung zu realisieren, kann eine Triangulation eingesetzt werden. Dafür ist eine Sensorfusion von Einzelsensoren mit überlappendem Erfassungsbereich notwendig. Die Daten der Messungen, also die Information über den Kontrast warmer Objekte zum kälteren Hintergrund, werden in einem Occupancy Grid [43] eingetragen. Diese Darstellung ermöglicht eine kartenähnliche Sicht des Fahrzeugumfeldes. Jedem Kartenelement wird, sobald eine neue Messung vorliegt, über das Bayes-Theorem eine neue Besetzungswahrscheinlichkeit zugeordnet.

Eignung

Die Öffnungswinkel von FIR-Kameras liegen im Bereich von 22° bis 36° . Da Glas eine wärmedämmende Eigenschaft besitzt, müssen die Kameras ohne jegliche Abdeckung im Fahrzeug integriert sein. Um die Linse möglichst gut schützen und bei Verschmutzung reinigen zu können, ist bei Pkw bisher lediglich eine Einbauposition in Höhe der Scheinwerfer oder im Bereich des Kühlergrills möglich. Jedoch ist eine Erfassung der gesamten Fahrzeugfront notwendig, wie in der Analyse der Fußgängerunfallszenen in Abschnitt 3.3.1 bereits festgestellt wurde. Um diese Anforderung dennoch umsetzen zu können, muss entweder ein horizontaler Öffnungswinkel des Sensors im Bereich von bis zu 180° oder eine nach außen gedrehte Sensoreinbaulage von bis zu $90^\circ - (\phi/2)$ vorliegen, wobei ϕ dem Öffnungswinkel des Sensors entspricht. Obwohl die FIR-Kameras relevante Objekte, nämlich Fußgänger und auch Radfahrer, besonders hervorheben und damit die Bildverarbeitung in vielen Situationen vereinfachen, verhindert die Verbaulage der Kameras in der Fahrzeugfront die Erfüllung der Anforderungen des in dieser Arbeit beschriebenen FSS.

Zwei Gründe sprechen gegen den Einsatz von Thermopiles. Zum Einen liefern Arrays – ähnlich wie Lidar-Systeme – relativ wenig Information über die Objekte. Zum Anderen sind die in [42] vorgestellten Occupancy Grids in dynamische Szenen mit Objekteigenbewegungen umso schlechter geeignet, je länger der Zeitraum ist, für den Belegungsinformationen eines Grid-Elements gesammelt werden. Falls sich also ein Objekt in ein Grid-Element hineinbewegt, welches zuvor als unbesetzt erkannt wurde, erhöht sich nur die Unsicherheit der Belegung. Es wird nicht sofort ein Objekt erkannt. Verzichtet man daher auf jegliche Historie, so wird auch kein Occupancy Grid erstellt.

4.9 ULTRASCHALL

Die Messung von Entfernungen mittels Ultraschall (US)-Sensoren basiert auf der Laufzeitmessung von Ultraschallwellen bei einer Frequenz von



ABBILDUNG 4.12: Beispiel eines Ultraschallsensors der Firma Valeo

über 16 kHz. Da sich Schallwellen kugelförmig ausbreiten, kann nur eine radiale Objektentfernung pro Objekt und Sensor gegeben werden. Mit Hilfe mehrerer Sensoren, die überlappende Erfassungsbereiche besitzen, können Schnittpunkte bestimmt werden, die eine ungefähre Objektposition auf Basis der Trilateration berechnen. Die zur Zeit in Fahrzeugen verbauten Sensoren werden für Einparkhilfen eingesetzt, da die Entfernungsmessungen nur in einem Bereich bis 3 m genau und über 5 m kaum sinnvoll sind.

Abbildung 4.12 zeigt einen US Sensor der Firma Valeo¹⁰. Die verwendeten Frequenzen liegen zwischen 38,4 kHz und 51,2 kHz. Die Zykluszeit dieses Sensors liegt bei 100 ms.

Eignung

Die Anfälligkeit des akustischen Signals gegenüber Wind verschlechtert die Messgenauigkeit, so dass der Niedergeschwindigkeitsbereich bei Messungen favorisiert wird. Aus diesen Gründen werden US-Sensoren bisher bei keinen weiteren Fahrzeugfunktionen eingesetzt. Eine zusätzliche Abhängigkeit ergibt sich durch die unterschiedliche Reflexion an unterschiedlichen Materialien. Manche Materialien, zum Beispiel Pelz, absorbieren US fast vollständig. Gefrorener Schnee am Boden ist ebenfalls sehr problematisch, da er zu Geisterobjekten führt. Regen ruft im Allgemeinen keine Geisterziele hervor. Unterschiedliche Temperaturen führen zwar zu einer unterschiedlichen Schallausbreitungsgeschwindigkeit, jedoch ist dadurch das Messprinzip nicht beeinträchtigt.

¹⁰aus <http://www.valeo.de>, zuletzt besucht im August 2006

Je kürzer die Zykluszeit, desto geringer ist die maximale Reichweite, die gemessen werden kann. Bei einer Objektdistanz von 5 m benötigt der Schall mindestens 29 ms für die Hin- und Rückstrecke. Da man aber aus Gründen der Kreuzpeilung mehrere Sensoren zeitlich nacheinander betreiben möchte, ist eine weitere Reduzierung der Zykluszeit schwierig. Fehldetektionen sind selten länger als zwei Zyklen und treten tendenziell eher auf, wenn sich ein Objekt in den Erfassungsbereich hinein bewegt.

Einschränkungen bezüglich der Reichweite sind durch Faktoren wie Sendeleistung, Richtcharakteristik des Trichters, Witterungsbedingungen, Oberflächenmaterial des Objektes und Sensorverschmutzung gegeben. Vor allem Materialien mit groben, rauen Oberflächen wie Fell, Cord, Baumwolle und Leder dämpfen das Signal stark ab. Ebenso können Luftverwirbelungen am Objekt die Signalqualität beeinträchtigen. Der Einfluss von Regen ist eher klein. Störungen des Signals können aber auch durch externe US-Quellen auftreten. Im Automobilbereich sind hier etwa Abrollgeräusche bei nasser Fahrbahn, Kehrmaschinen und Lkw (Ablassen von Druckluft) zu erwähnen.

Abhängig von der gewünschte Erkennung sind folgende Tatsachen festzuhalten:

Detektion mit einem Sensor Es kann lediglich die radiale Entfernung ausgegeben werden. Die Zykluszeit ergibt sich aus der maximal geforderten Reichweite, die einen Bereich von vermutlich $d = 10$ m nicht überschreiten kann. Es ergibt sich eine maximale Systemtakttrate von 17 Hz. Da jedes Hindernis im Erfassungsbereich zum gleichen Sensorsignal führt, ist diese Art der Umfelderkennung für das FSS nicht geeignet.

Detektion mit mehreren Sensoren Die für die Trilateration notwendigen Mehrfachmessungen erfordern mindestens 2, eher 4 Sensoren. Da die Signale zeitlich entkoppelt ausgesandt und verarbeitet werden müssen, ergibt sich eine entsprechend der Anzahl verwendeter Sensoren reduzierte Systemtakttrate. Die Genauigkeit der Positionsbestimmung ist stark windabhängig und zeitlich sowohl durch die Laufzeit des Signals als auch durch die erforderlichen Mehrfachmessungen entsprechend spät verfügbar. Prinzipiell ist somit eine Hindernisdetektion für das FSS möglich, jedoch mit zu geringer Systemtakttrate und sehr ungenauer Positionsbestimmung.

Klassifikation Eine Klassifikation ist mit der hier vorgestellten Sensorik nicht realistisch, da zu wenig Informationen ausgewertet werden können.

Zusammenfassend ist zu sagen, dass der Einsatz von US-Sensoren für ein FSS kritisch zu bewerten ist. Bisher sind keine Sensoren im Einsatz, die

die erforderliche Reichweite aufweisen. Alle anderen Nachteile bezüglich Positionsungenauigkeit und nicht realistischer Klassifikation bleiben prinzipbedingt bestehen.

4.10 SENSORIKEMPFEHLUNG

Aus den Beschreibungen und den zusätzlich durchgeführten eigenen Analysen konnten eine Reihe von Erkenntnissen über die Eignung der unterschiedlichen vorausschauenden Sensoren gewonnen werden. Dabei zeigt sich, dass eine Kombination von PMD- und MV-Sensorik besonders reizvoll erscheint. Die Gründe dafür sind zahlreich. So liefert die PMD-Kamera Objektlisten, die unter anderem die relevanten Verkehrshindernisse beinhalten. Trotz der Nichtstationarität der Sensorik ist diese Objektdetektion zuverlässig und effizient. Das Sichtfeld der PMD-Kamera entspricht relativ gut dem erforderlichen Erfassungsbereich der Schutzfunktion FSS. Die aktive Beleuchtung unterstützt die witterungsrobuste Bereitstellung der gewünschten Informationen. Zusätzlich liefert sie klassifikationsrelevante Informationen über die Objektform, die etwa ein Laserscanner oder ein Radarsensor nicht liefern können. Die notwendige Kapazität für die Sensordatenverarbeitung ist, vor allem in Gegensatz zu Stereovideosystemen, deutlich kleiner und eignet sich daher mit dem heutigen Stand der Technik besser für den automotiven Einsatz. Die für eine zuverlässige Klassifikation noch fehlende Information kann nur aus einer Monovideokamera kommen, da sie der einzige Sensor ist, der eine ausreichende Menge an Informationen über die Objekte zur Verfügung stellt. Die Vorauswahl der Objektdetektion durch einen entfernungsmessenden Sensor ermöglicht weiterhin die Reduktion der zu verarbeitenden Videodaten. Aus diesen Gründen wird im Folgenden nur noch die Kombination aus PMD und MV zur Erkennung von Fußgängern betrachtet. Genau dieses System trägt die Bezeichnung Pedestrian Recognition System Using a Multi-Sensor Environment (PRESUME).

FUSSGÄNGERERKENNUNG AUS 3D- UND 2D-DATEN

Mit Hilfe der im vorherigen Kapitel beschriebenen vorausschauenden Sensoren können Fernmessungen der Umgebung durchgeführt werden. Die für den vorausschauenden Fußgängerschutz zu lösende Aufgabe ist es, Objekte der Umwelt und deren Merkmale aus den Messdaten zu extrahieren und ihren Bedeutungsklassen zuzuordnen. Dieser Vorgang der Klassifikation als wesentlicher Bestandteil der Mustererkennung ist in vielen Bereichen vorzufinden. In der Produktion gibt es Klassifikatoren zur Entscheidung von Gütekriterien und Fehlerfreiheit eines Produktes, in der Medizin werden Blutparameter von Patienten verwendet um Krankheiten zu diagnostizieren, und natürlich finden Klassifikatoren in der Bildverarbeitung (BV) breite Anwendung. Es bleibt anzumerken, dass die Anzahl möglicher Erscheinungsformen aller Fußgänger beliebig groß ist, und dass die Erkennung dieser Fußgänger auf Basis räumlich bewegter Sensoren geschehen muss.

Kapitel 5 beschreibt nun das zu lösende Problem der Mustererkennung und liefert einige grundlegenden Kenntnisse über die Theorie der Klassifizierung. Dazu wird im Abschnitt 5.1 kurz auf die mathematische Beschreibung der Schätztheorie eingegangen. Die aus den 3D-PMD- und 2D-MV-Daten extrahierbaren Merkmale zur Unterscheidung der Klassen Fußgänger und Nicht-Fußgänger werden in den Abschnitten 5.2 und 5.4 ausführlich beschrieben. Die Fusion aus PMD- und MV-Daten wird in Abschnitt 5.3 erklärt. Anschließend stellt Abschnitt 5.5 mögliche Ansätze der eigentlichen Klassifikation vor.

Beispielabbildung 5.1 zeigt an einer Szene mit einem die Straße kreuzenden Fußgänger, welche Informationen für die nachfolgenden Bildverarbeitungsschritte zur Verfügung stehen.

Einige Ansätze der Klassifikation wurden bereits in [29], [30], [44] und [45] erwähnt und in begrenztem Umfang analysiert. Im Rahmen dieser Arbeit wurde jedoch zum ersten Mal der aufwendige und vollständige Umfang der Mustererkennung für die Fußgängerklassifikation mit PMD- und MV-Daten hergeleitet, implementiert, getestet, optimiert und dokumentiert.

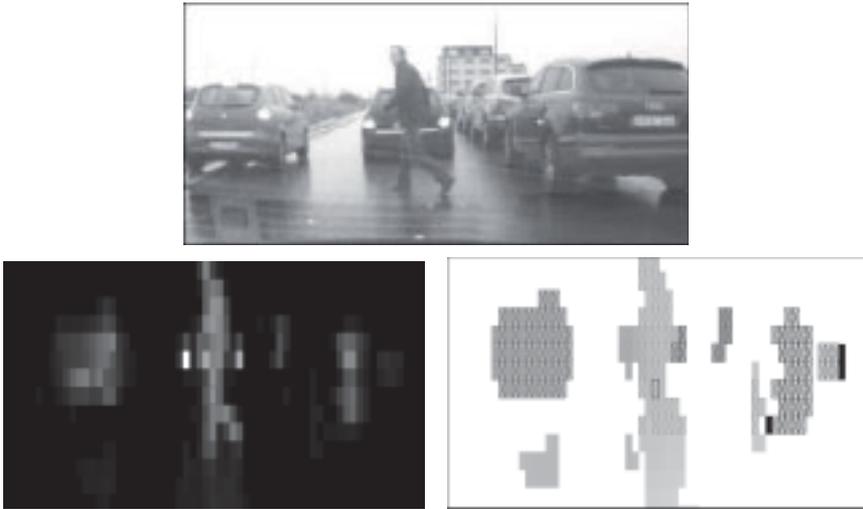


ABBILDUNG 5.1: alle für die Merkmalsextraktion zu einem Zeitpunkt zur Verfügung stehenden Bildinformationen. Hochauflösendes Videokamerasignal (oben), PMD-Amplitudenbild (links), PMD-Distanzbild (rechts)

5.1 DIE ALLGEMEINE SCHÄTZTHEORIE

Um die mathematischen Werkzeuge einsetzen zu können, betrachten wir zunächst die mathematische Beschreibung eines Sensors. Es sei \mathcal{X} der Objektraum und \mathcal{M} der Muster- oder Merkmalsraum. Die Eingangsdaten eines Sensors, der den Objektraum erfasst, werden mit der Funktion

$$f : \mathcal{X} \rightarrow \mathcal{M}, \mathbf{x} \mapsto \mathbf{m} \text{ mit } \mathbf{m} = A(\mathbf{x}) + \mathbf{n} \quad (5.1)$$

auf den Musterraum abgebildet. Das Ergebnis ist die Beobachtung \mathbf{m} , auch Beobachtungsvektor oder Merkmalsvektor genannt. \mathbf{n} ist ein Rauschterm.

Die Mustererkennung ist Teil der allgemeinen Schätztheorie. Mit ihrer Hilfe kann aus gegebenem \mathbf{m} auf ein \mathbf{x} geschlossen werden. Die Schätztheorie unterscheidet dabei zwischen einer Signalschätzung, einer Signaldetektion und einer Mustererkennung. Bei der Schätzung ist jeder beliebige Wert \mathbf{x} möglich. Man spricht von einem vollbesetzten Objektsignalraum $\mathbf{x} \in \mathcal{X}$. Die Signaldetektion unterscheidet sich insofern, dass die gesuchten Werte nur eine Untermenge aller möglichen Werte $\mathbf{x} \in \{\mathbf{x}_i\} \subset \mathcal{X}$ darstellen, \mathbf{x}_i entspricht einem Hyperkubus. Im Fall der Mustererkennung existieren nun endlich viele Äquivalenzklassen S_i so dass $\mathbf{x} \in S_i \{\mathbf{x}_i\} \subset \mathcal{X}$ gilt. Es ist die Abbildung \tilde{f} gesucht mit

$$\tilde{f} : \mathcal{M} \rightarrow \{S_1, S_2, \dots, S_K\}, \mathbf{m} \mapsto \mathbf{x}, \quad (5.2)$$

wobei K der gegebenen Anzahl an Klassen entspricht. Die Begriffe Äquivalenzklasse und Bedeutungsklasse, im Folgenden einfach als Klasse bezeichnet, werden in der Literatur synonym benutzt. Dabei ist eine Extraktion von Merkmalen mit hoher Äquivalenzklassenrelevanz und im Allgemeinen eine Reduktion der Merkmalsraumdimension notwendig. Die Mustererkennung teilt sich weiter auf in die zwei Bereiche Merkmalsextraktion und Klassifikation beziehungsweise Interpretation. Mit Hilfe einer Stichprobe kann eine Klasse S_i repräsentiert werden. Um eine gute Klassifikationsleistung zu erreichen, sollte jede Klasse eine gewisse Kompaktheit im Sinne einer Metrik besitzen, die man sich als Cluster im Merkmalsraum vorstellen kann.

5.2 MERKMALSEXTRAKTION AUS 3D-DATEN

Die Merkmalsextraktion ist für die Reduktion der Information in den Sensordaten verantwortlich. Aus den ursprünglich von der PMD-Kamera gelieferten Amplituden- und Distanzbildern werden skalare Merkmale berechnet, die anschließend in die Klassifikation einfließen. Jedes Merkmal $m_i \in \mathbf{m}$ ist Teil des Merkmalsvektors $\mathbf{m} = [m_1, m_2, \dots, m_M]^T$ der Dimension M des Merkmalsraums \mathcal{M} . Im Folgenden werden nun die Berechnungen der Merkmale im Detail erläutert.

5.2.1 Vorverarbeitung der Region

Eine Kostenfunktion bewertet die Amplituden- und Distanzdaten mit dem Ziel, ein Objekt zu erkennen. Dabei werden benachbarte Pixel genau dann zusammengefasst, wenn ihre Amplituden- und Distanzdaten korrelieren. Die dadurch entstehenden Pixelgruppen werden als Region bezeichnet. Als Beispiel ist eine schematische Repräsentation einer Region in Abbildung 5.2 gezeigt. Diese dient in allen nun folgenden Merkmalsberechnungen als Berechnungsgrundlage. Eine Bounding-Box der Größe $(M \times N)$ umschließt das Objekt vollständig.

Für einige Merkmalsberechnungen ist ein reines 2D-Binärbild erforderlich. Dieses besteht nur noch aus Pixelwerten 1 und 0. Zur Erzeugung wird die Region mit der diskreten Bildfunktion $f(x, y)$ multipliziert. Diese liefert einen Wert 1, wenn der Bildpunkt zum Objekt gehört, sonst 0. Die Anzahl der mit 1 gekennzeichneten Pixel ist mit

$$F = \sum_{i=1}^M \sum_{j=1}^N f(x, y) \quad (5.3)$$

beschrieben. Zur Formbeschreibung der Region ist es außerdem notwendig, das Binärbild auf seine Randpixel auszudünnen. Unter Berücksichti-

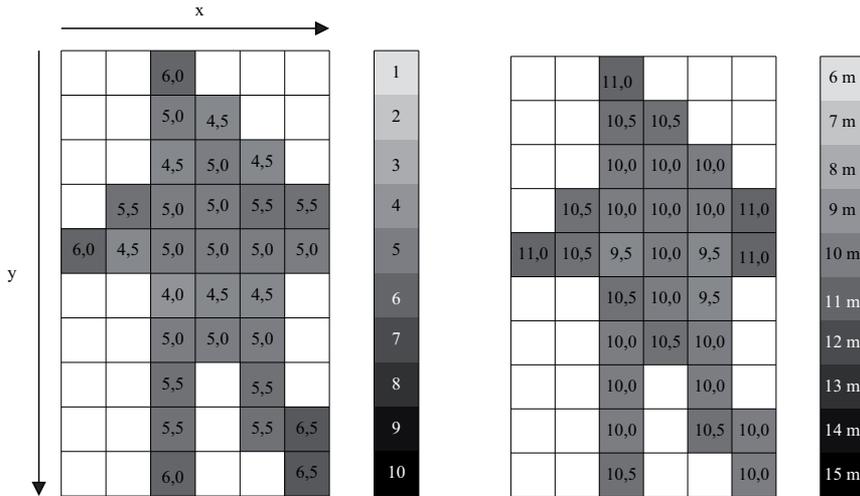


ABBILDUNG 5.2: Schematische Abbildung einer Region aus dem Amplitudenbild (links) und dem Distanzbild (rechts)

gung einer 8-er Nachbarschaft wird eine Kontur $k(x, y)$ erzeugt. Die Länge der Kontur beträgt

$$K = \sum_{i=1}^M \sum_{j=1}^N k(x, y) \quad (5.4)$$

Bei der anschließenden Sortierung nach [46] werden die Konturelemente $k(x, y)$ in

$$\mathbf{r}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix}, \quad i = 1 \dots K \quad (5.5)$$

abgelegt. Siehe dazu Abbildung 5.3.

5.2.2 Größe

Die Objektgröße wird mit einer homogenen Koordinatentransformation unter Berücksichtigung der Entfernung in für alle Objekte eindeutige Fahrzeugkoordinaten abgebildet. Daraus resultieren die Merkmale für die Höhe HEIGHT und die Breite WIDTH, die die umschließende Bounding-Box beschreiben. Zusätzlich wird noch der höchste Punkt des Objektes über das Merkmal MAXTOPPOINT beschrieben. Die Histogramme über die verwendeten Merkmale sind auf Basis des vollständigen Trainingsdatensatzes im Anhang A jeweils einzeln abgebildet.

Es gibt einen als qualitativ zu betrachtenden sogenannten Nahbereich vor dem Fahrzeug, in dem der vertikale Erfassungsbereich des Sensors zur

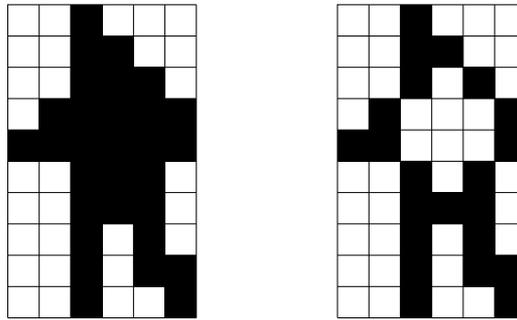


ABBILDUNG 5.3: Reduktion der Region auf das Binärbild durch $f(x, y)$ (links) und die Kontur $k(x, y)$ beziehungsweise r_i (rechts)

vollständigen Erfassung eines durchschnittlich großen Fußgängers nicht mehr ausreicht. Die Trennung des gesamten Erfassungsbereiches in einen Nah- und Fernbereich kann verwendet werden, um dem Merkmal HEIGHT in beiden Bereichen eine unterschiedliche Bedeutung zukommen zu lassen.

Bei der Berechnung der Objektiefe DEPTH ist zu beachten, dass eine Bestimmung zwar erfolgt, diese Angaben jedoch unter Umständen nicht den tatsächlichen Werten entsprechen. Das ist zum Beispiel der Fall, wenn der seitliche Versatz des Objektes nahe Null ist und damit die Objektiefe aus perspektivischen Gründen unsichtbar ist (siehe Koordinatensystem in Abbildung 6.4). Zwar spannen in diesem Fall auch die größten und kleinsten PMD-Abstandswerte das Objekt in seiner Tiefe auf, jedoch sind diese Abstandswerte auf das Sensorrauschen und nicht auf die tatsächliche Objektiefe zurückzuführen.

5.2.3 Globale geometrische Merkmale

Diese Art der Merkmale basieren auf den Binär- und Konturbildern und stellen sehr einfach zu berechnende Beschreibungen der Objekte dar.

Füllgrad

Da die Fläche F der Region skalierungsvariant ist, wird sie mit der Fläche der umschließenden Bounding-Box $M \cdot N$ normiert.

$$\text{COVERAGERATIO} = \frac{F}{M \cdot N} \quad (5.6)$$

Normierter Objektumfang

Der Objektumfang K ist ebenfalls skalierungsvariant und wird daher mit dem Umfang der Bounding Box normiert.

$$\text{LENGTH} = \frac{K}{2M + 2N} \quad (5.7)$$

Formfaktor

Der Formfaktor beschreibt das Verhältnis aus Umfang zu Fläche. Er berechnet sich zu

$$\text{FORMFACTOR} = \frac{K^2}{4\pi \cdot F} \quad (5.8)$$

und ist umso größer, je mehr die Kontur im Sinne einer Zerklüftung von der Form eines idealen Kreises abweicht.

Radienverhältnis

Zu jedem Pixel kann ein Radius bestimmt werden, der sich durch den geometrischen Abstand zum Objektschwerpunkt berechnet. Das Radienverhältnis ist der Quotient aus maximalem und minimalem Radius. Wie beim Merkmal FORMFACTOR wird damit die Zerklüftung in Bezug auf einen Kreis beschrieben. Für stark zerklüftete Objekte, bei denen der Objektschwerpunkt außerhalb der Binärregion liegt, liefert das Radienverhältnis jedoch zu kleine Werte. Es wird daher nicht verwendet.

5.2.4 Momentenbasierte Objektform

In Bezug auf einen Referenzpunkt oder eine Referenzachse liefern Momente numerische Eigenschaften eines Objektes. Sie werden in der Statistik zur Beschreibung von Verteilungen oder in der Mechanik zur Beschreibung der räumlichen Gestaltung von Körpern eingesetzt. In unserem Fall beschreiben sie die Eigenschaften der 2D-Binärbilder. Für eine Region der Größe $M \times N$ lassen sich Momente berechnen, die sowohl eine Aussage über den Schwerpunkt, also auch über dessen Fläche, Varianz, Schiefe und Wölbung liefern. Um die Region eindeutig und vollständig zu beschreiben, sind prinzipiell unendlich viele Momente erforderlich. Die Merkmalsextraktion hat an dieser Stelle aber nicht das Ziel einer sehr präzisen Beschreibung einer einzelnen Form, sondern der Charakterisierung einer typischen Form, die eine Klasse repräsentiert. Daher besteht die Aufgabe darin, diejenigen Momente zu finden, die die Bildinformation ausreichend genau beschreiben. Als Referenzbild dient wiederum Abbildung 5.3.

Allgemeine Momente

Es können die allgemeine Momente der Ordnung p und q über

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y) \tag{5.9}$$

berechnet werden. Tabelle 5.1 zeigt die allgemeinen Momente nullter und erster Ordnung und deren Bedeutung [47].

Momente	Ordnung	Bedeutung bzgl. der Region
m_{00}	0	Fläche
$\bar{x} = \frac{m_{10}}{m_{00}}$	1	Schwerpunkt in x
$\bar{y} = \frac{m_{01}}{m_{00}}$	1	Schwerpunkt in y

TABELLE 5.1: Momente nullter und erster Ordnung

Zentrale Momente

Die zentralen Momente entsprechen den um den Objektschwerpunkt verschobenen allgemeinen Momenten und werden über die Funktion

$$\mu_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{5.10}$$

berechnet. Sie sind damit translationsinvariant. Die wesentliche Bedeutung dieser Momente liegt in der Möglichkeit zur Bestimmung der Regionorientierung und ihrer Exzentrizität. Tabelle 5.2 zeigt eine Auflistung der zentralen Momente bis zur vierten Ordnung und deren Bedeutungen.

Momente	Ordnung	Bedeutung
$\mu_{00} = m_{00}$	0	Fläche der Region
μ_{01}, μ_{10}	1	per Definition Null
$\mu_{11} = m_{11} - \frac{m_{10}m_{01}}{m_{00}}$	1	Trägheit, σ_{xy} von x und y
$\mu_{20} = m_{20} - \bar{x}m_{10}$	2	Trägheit σ_x^2
$\mu_{02} = m_{02} - \bar{y}m_{01}$	2	Trägheit σ_y^2
μ_{03}, μ_{30}	3	Schiefe
μ_{04}, μ_{40}	4	Wölbung

TABELLE 5.2: Zentrale Momente nullter bis vierter Ordnung

Die Momente zweiter Ordnung, auch Trägheitsmomente genannt, dienen zur Bestimmung der Hauptachse. Das Moment der vierten Ordnung μ_{04} , die Wölbung, beschreibt dabei die Abweichung des Verlaufs der Region zum Verlauf einer Normalverteilung. Die Normalverteilung hat die Wölbung $\mu_4 = 0$, ein positiver Wert entspricht einer spitzeren Verteilung und wird als leptokurtisch bezeichnet. Negative Wölbungen sind platykurtisch und weisen einen abgeflachteren Verlauf auf.

In PReSUME wurden verschiedene Momente untersucht und implementiert. Als trennstärkstes Moment hat sich die Regionenorientierung Θ herausgestellt. Ebenso ist die Exzentrizität implementiert, welche die Dehnung des Objektes beschreibt.

$$\text{ORIENTATION} = \frac{1}{2} \left| \arctan \left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \right| \quad (5.11)$$

$$\text{EXCENTRICITY} = \frac{(|\mu_{20}| - |\mu_{02}|)^2 + 4\mu_{11}^2}{(|\mu_{20}| + |\mu_{02}|)^2} \quad (5.12)$$

Normierte zentrale Momente

Durch Normierung der Zentralmomente mit der Fläche erhält man translations- und rotationsinvariante Momente

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{(p+q+2)/2}} \quad (5.13)$$

Hu-Momente

Aus den normierten Momenten η_{pq} lassen sich wiederum neue Momente ableiten, die sogenannten Hu-Momente ϕ_i [48]. Diese sind invariant bezüglich Translation, Skalierung und Rotation. Es existieren 10 Momente bis zur 3. Ordnung, von denen die ersten 4 in Tabelle 5.3 aufgelistet sind.

Momente

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

TABELLE 5.3: Hu-Momente aus [49]

5.2.5 Konturbasierte Objektform

Typischerweise analysieren Bildverarbeitungssysteme die Form und die Anordnung von Kanten oder ähnlichen Grauwertübergängen der zu un-

tersuchenden Objekte, siehe dazu zum Beispiel [50], [51] und [52]. Diese Ansätze nutzen jedoch, im Vergleich zu dem hier verwendeten PMD-System mit 1024 Pixeln, höher auflösende Bilddaten. Dennoch soll der Objektumriss zur Analyse in die Klassifikation einfließen.

Fourieranalyse

In dieser Arbeit wird eine 2D-Fourieranalyse eingesetzt. Die Kontur der Objekte kann mit Hilfe von K Messpunkten in einer komplexen Darstellung mit

$$\mathbf{a}_k = \begin{pmatrix} 1 \\ i \end{pmatrix} \mathbf{r}_k^T \quad (5.14)$$

in Pixelkoordinaten beschrieben werden, wobei \mathbf{r}_k die sortierten Konturelemente $(x_k, y_k)^T$ aus Gleichung (5.5) enthält. Die Fouriertransformation

$$c_m = \frac{1}{K} \sum_{k=0}^{K-1} \mathbf{a}_k e^{-2\pi i \frac{k m}{K}} \quad (5.15)$$

liefert aus K Messpunkten genau K komplexe Fourierkoeffizienten c_m , die das Frequenzverhalten der Kontur beschreiben. Mit genügend großem K kann die ursprüngliche Kontur beliebig genau rekonstruiert werden, in unserer Anwendung ist K jedoch durch die Konturlänge entsprechend Gleichung (5.4) beschränkt. Es entspricht

$$e^{-2\pi i \frac{k}{K}} = \cos\left(\frac{2\pi}{K}k\right) + i \sin\left(\frac{2\pi}{K}k\right) \quad (5.16)$$

der K -ten Einheitswurzel, also K gleichmäßig verteilte Punkte auf dem Einheitskreis der komplexen Zahlenebene. Da die Größenmerkmale separat analysiert und die Objektposition explizit unterdrückt werden, muss eine Unabhängigkeit gegenüber Konturposition und -größe in der Bildebene erreicht werden. Dazu wird der erste Koeffizient c_0 ignoriert und der zweite Koeffizient c_1 als Normierungsfaktor für alle übrigen c_m verwendet. Die Zahl der in die Klassifikation einfließenden Fourierkoeffizienten D ist frei wählbar. Für $D > K - 2$ liefert die Transformation jedoch nicht genug Koeffizienten, da in diesem Fall der Objektumriss zu wenig Pixel beinhaltet. Um diesen Fall dennoch abzudecken, erfolgt eine Skalierung der Kontur.

Als Merkmale werden die i -ten Koeffizienten $c_i = |c_i| \exp(j\phi)$ der Fouriertransformation

$$\text{FOURIERAMPLITUDE_I} \quad |c_i| = \frac{1}{|c_1|} \left(\{\Re\{c_i\}\}^2 + \{\Im\{c_i\}\}^2 \right) \quad (5.17)$$

$$\text{FOURIERPHASE_I} \quad \phi_i = \arctan\left(\frac{\Im\{c_i\}}{\Re\{c_i\}}\right) \quad (5.18)$$

mit $i \in \{1; D\}$ getrennt nach Amplitude und Phase betrachtet. Durch die Normierung mit $|c_1|$ wird die Skalierungsinvarianz erreicht. Das Histogramm über Betrag und Phase des ersten Fourierkoeffizienten ist im Anhang A grafisch dargestellt.

Eine Schwäche der Fourieranalyse auf niedrig aufgelösten Bildern besteht in der Tatsache, dass eine geforderte äquidistante Aufteilung der Abtastwerte a_k nicht präzise erfolgen kann. Damit sinkt die Aussagekraft der einzelnen Koeffizienten c_m .

Histogramme der Kantenrichtung

Zusätzlich werden die normierten Häufigkeiten der Richtung jedes Pixelübergangs bestimmt und aufsummiert. Da für die Richtungsbestimmung nur jeweils der erste nächste Nachbar in der Kontur betrachtet wird, ergeben sich vier Richtungen, sofern die Orientierung vernachlässigt wird.

$$\text{NUMBEREDGES0DEGREE} = \frac{1}{K} \sum_{i=1}^K g_0(\mathbf{r}_i, \mathbf{r}_{i+1}) \quad (5.19)$$

$$\text{NUMBEREDGES45DEGREE} = \frac{1}{K} \sum_{i=1}^K g_{45}(\mathbf{r}_i, \mathbf{r}_{i+1}) \quad (5.20)$$

$$\text{NUMBEREDGES90DEGREE} = \frac{1}{K} \sum_{i=1}^K g_{90}(\mathbf{r}_i, \mathbf{r}_{i+1}) \quad (5.21)$$

$$\text{NUMBEREDGES135DEGREE} = \frac{1}{K} \sum_{i=1}^K g_{135}(\mathbf{r}_i, \mathbf{r}_{i+1}). \quad (5.22)$$

Dabei ist $g_0(\mathbf{r}_i, \mathbf{r}_{i+1}) = 1$ genau dann, wenn die beiden Nachbarpixel \mathbf{r}_i und \mathbf{r}_{i+1} horizontal nebeneinander liegen. Entsprechendes gilt für g_{45} , g_{90} und g_{135} mit einer Nachbarschaftslage in 45° -Schritten im mathematisch positiven Sinn. Einige Histogramme sind in Anhang A abgebildet.

5.2.6 Weitere Merkmale zur Objektform

Um Objekte mit linearer horizontaler oder vertikaler Randstruktur von Objekten mit unregelmäßigen Grenzen unterscheiden zu können, werden ergänzend die folgenden statistischen Merkmale erzeugt:

Verhältnis der Randpixel

Aus dem Verhältnis der Anzahl der Randpixel der ersten h Spalten in Bezug zur Höhe, beziehungsweise b Reihen in Bezug zur Breite, lassen sich

weitere Merkmale ableiten. Histogramme dazu sind in Anhang A dargestellt.

$$\text{UPPERBORDERTOWIDTH} = \frac{1}{bM} \sum_{i=1}^b \sum_{j=1}^M f(j, i) \quad (5.23)$$

$$\text{LOWERBORDERTOWIDTH} = \frac{1}{bM} \sum_{i=M-b+1}^M \sum_{j=1}^M f(j, i) \quad (5.24)$$

$$\text{LEFTBORDERTOHEIGHT} = \frac{1}{hN} \sum_{i=1}^h \sum_{j=1}^N f(j, i) \quad (5.25)$$

$$\text{RIGHTBORDERTOHEIGHT} = \frac{1}{hN} \sum_{i=N-h+1}^N \sum_{j=1}^N f(j, i) \quad (5.26)$$

Abweichung von der durchschnittlichen Spalten- und Reihenlänge

Die Summe der quadratischen Abweichung der Anzahl Pixel m_i pro Reihe i von der durchschnittlichen Anzahl Pixel pro Reihe \bar{m} zeigt die Abweichung der Regionform von einem achsenparallel ausgerichteten Rechteck. Gleiches gilt für die Anzahl Pixel n_j in den Bildspalten j . Siehe dazu auch Anhang A.

$$\text{SQUAREDROWDIFFFROMAVG} = \frac{1}{M^2} \sum_{i=1}^M (\bar{m} - m_i)^2 \quad (5.27)$$

$$\text{SQUAREDCOLDIFFFROMAVG} = \frac{1}{N^2} \sum_{j=1}^N (\bar{n} - n_j)^2 \quad (5.28)$$

5.2.7 *Textur*

Zwar gibt es keine formale Definition des Texturbegriffes, jedoch kann man sie informal als sich wiederholende, mehr oder weniger regelmäßige Muster beschreiben, die als zusammenhängende Region interpretiert werden. Grundlegende Texturmerkmale sind die Grobheit, die Periodizität, der Kontrast und die Gerichtetheit. Eine Texturanalyse muss die Regelmäßigkeiten erkennen und die Abweichungen ignorieren [53]. Im Gegensatz zur Fouriertransformation eine Reversibilität, das heißt die Wiederherstellung der Originaltextur auf Basis der extrahierten Statistiken, nicht gegeben.

Auf Basis der Bildwerte g_i lassen sich der Mittelwert \bar{g}_i , der Maximalwert $g_{i,\max}$ und die Standardabweichungen s_r und s_c mit

$$\text{AVERAGE} = \frac{1}{MN} \sum_{i=1}^{MN} g_i \quad (5.29)$$

$$\text{MAXIMUM} = \frac{1}{d_i} \max_{i \in [1; MN]} g_i \quad (5.30)$$

$$\text{STDDEVATION} = \sqrt{\frac{MN \sum_{i=1}^{MN} g_i^2 - (\sum_{i=1}^{MN} g_i)^2}{MN(MN - 1)}} \quad (5.31)$$

$$\text{STDDEVATIONROW} = \frac{1}{N} \sum_{i=1}^N \sqrt{\frac{M \sum_{j=1}^M g_j^2 - (\sum_{j=1}^M g_j)^2}{M(M - 1)}} \quad (5.32)$$

$$\text{STDDEVATIONCOL} = \frac{1}{M} \sum_{i=1}^M \sqrt{\frac{N \sum_{j=1}^N g_j^2 - (\sum_{j=1}^N g_j)^2}{N(N - 1)}} \quad (5.33)$$

bestimmen. Der Wert d_i bezeichnet den Distanzwert des Pixels (siehe Abschnitt 5.2.8).

Eine weitere Analysemöglichkeit der Grauwerttextur geschieht mit Hilfe von Haralick'schen Texturmaßen über die Spatial Greylevel Dependence (SGLD) Methode [54]. Mit Hilfe einer Statistikbeschreibung der 2. Ordnung werden die Verhältnisse der Graustufen $g_i \in G$ zweier benachbarter Pixel beschrieben. G ist gleich der Anzahl möglicher Grauwerte, also zum Beispiel $G = 256$ bei einer Bittiefe von 8 bit. Für einen gewählten Abstandsvektor \mathbf{d} zwischen zwei Pixeln wird eine SGLD Matrix \mathbf{h}_d erstellt. Dabei repräsentiert ein Element $h_d(m, n)$ dieser Matrix die Häufigkeit des Grauwertübergangs von $m \in G$ auf $n \in G$ zwischen allen Nachbarn des Abstandes \mathbf{d} . Dadurch ist $\dim(\mathbf{h}_d) = G \times G$.

Für die SGLD-Analyse stehen in dieser Verarbeitungsstufe meist nur wenige Pixel zur Verfügung, nämlich die Pixel eines Objektes. Daher betrachten wir ausschließlich Nachbarschaften $|\mathbf{d}| = 1$. Zusätzlich wird die Anzahl verwendeter Graustufen G auf $G' = 32$ reduziert, was einer Bittiefenreduktion der Eingangsbilder von 16 bit auf 5 bit entspricht. Mit diesem Schritt werden zu schwach besetzte SGLD-Matrizen vermieden und der Rechenaufwand drastisch verringert.

Die Matrizen können noch nicht direkt als Merkmal benutzt werden. Erst durch die Bestimmung von Merkmalen erster oder zweiter Ordnung, siehe Gleichungen (5.34) bis (5.43), können Skalare gewonnen werden, die jeweils ein Merkmal m_j repräsentieren. Zu diesen *Amplitudentexturmerkmalen* zählt zum Beispiel das Differenzenmoment (DM) oder die Energie der Matrix. Das Histogramm der Homogenität ist in Anhang A visualisiert.

$$\text{UNIFORMITY} = \sum_{i_1=1}^G \sum_{i_2=1}^G h_d(i_1, i_2) \quad (5.34)$$

$$\text{AUTOKORRELATION} = \sum_{i_1=1}^G \sum_{i_2=1}^G i_1 \cdot i_2 \cdot h_d(i_1, i_2) \quad (5.35)$$

$$\text{CONTRAST} = \sum_{i_1=1}^G \sum_{i_2=1}^G (i_1 - i_2)^2 h_d(i_1, i_2) \quad (5.36)$$

$$\text{DM OF ORDER } q = \sum_{i_1=1}^G \sum_{i_2=1}^G (i_1 - i_2)^q h_d(i_1, i_2) \quad (5.37)$$

$$\text{HOMOGENEITY} = \sum_{i_1=1}^G \sum_{i_2=1}^G \frac{h_d(i_1, i_2)}{1 + |i_1 - i_2|} \quad (5.38)$$

$$\text{INVERSE DM} = \sum_{i_1=1}^G \sum_{i_2=1}^G \frac{h_d(i_1, i_2)}{|i_1 - i_2|^q}; \quad i_1 \neq i_2 \quad (5.39)$$

$$\text{ENERGY} = \sum_{i_1=1}^G \sum_{i_2=1}^G h_d(i_1, i_2)^2 \quad (5.40)$$

$$\text{ENTROPY} = - \sum_{i_1=1}^G \sum_{i_2=1}^G h_d(i_1, i_2) \text{ld}[h_d(i_1, i_2)] \quad (5.41)$$

$$\text{MAX. FREQUENCY} = \max_{i,j} h_d(i_1, i_2); \quad \arg \max_{i,j} h_d(i_1, i_2) \quad (5.42)$$

$$\text{MAIN DIAGONAL} = \sum_{i=1}^G h_d(i, i)^2 \quad (5.43)$$

Homogene Bilder liefern Punkte, die auf der Hauptdiagonalen von \mathbf{h}_d liegen. Bereiche mit regelmäßiger und stark ausgeprägter Struktur liefern Punktwolken auf der Nebendiagonalen von \mathbf{h}_d .

5.2.8 Oberfläche

Ebenso wie die Amplitudenwerte g_i in Abschnitt 5.2.7 können auch die Distanzwerte d_i dieser Analyse unterzogen werden. Die resultierenden Werte werden als Distanztexturmerkmale bezeichnet. Die Berechnung erfolgt ebenfalls nach Formel (5.29) bis (5.43).

5.3 FUSION DER 3D- UND 2D-DATEN

Eine Fusion aus 3D- und 2D-Sensordaten kann die Leistungsfähigkeit der Erkennung erhöhen. Dies wurde in [45] auf Basis des PMD-Systems und einer Monovideokamera bereits gezeigt. Dort wird die Fusion aus 3D- und 2D-Sensordaten durch die Bildung einer Region-Of-Interest (ROI) in der MV-Bildverarbeitung auf Basis der PMD-Daten realisiert. Dabei übergibt der PMD-Sensor die Objektschwerpunkte der detektierten Objekte an die 2D-Bildverarbeitung, die jeweils eine ROI um die Objektschwerpunkte herum aufspannt. Die Größe einer ROI orientiert sich am Maximalwert der erwarteten Größe eines Fußgängers und ist auf eine Höhe von 2,2 m und eine Breite von 1,1 m festgelegt. Der Fußpunkt der Region liegt dabei 0,2 m unter der prädierten Fahrbahnoberfläche, um eventuell auftretende Fehler bei der Detektion und Ungenauigkeiten bei der Transformation abzudecken. Anschließend werden die Eckpunkte der ROI in die 2D-Bildebene transformiert. Die damit definierte Menge M_1 der Pixel bildet die ROI_1.

Zusätzlich zur Verwendung des Objektschwerpunktes können auch die zur PMD-Region der Größe F gehörenden Pixel benutzt werden, die mit 3D-Koordinaten im Fahrzeugkoordinatensystem beschrieben sind. Die Abbildung dieser Pixel in die Bildebene ergibt F einzelne Bereiche. Die Transformation unter Nutzung des Entfernungswertes und des Kameraöffnungswinkels in die 2D-Bildebene liefert eine Menge M_2 , die als ROI_2 bezeichnet wird. Im Idealfall beinhaltet M_2 die Region vollständig, die Form der Region unterliegt natürlich der Pixelquantisierung des PMD-Sensors. Abbildung 5.4 zeigt die beiden Regionen an einem Beispielfeld.



ABBILDUNG 5.4: die 2D-Bildverarbeitung nutzt lediglich einen Ausschnitt des Videobildes (links), und zwar entweder die ROI_1 aus der Bounding-Box (Mitte) oder ROI_2 aus der PMD-Region (rechts)

Bei der in dieser Arbeit gewählten Fusion der 3D- und 2D-Sensordaten handelt es sich um eine Kombination aus komplementärem und kooperativem Fusionsansatz. Der komplementäre Anteil ergibt sich aus der Unabhängigkeit der PMD-Sensordatenverarbeitung von der MV-Sensordaten-

verarbeitung. Selbst wenn die 2D-Kamera, zum Beispiel auf Grund beleuchtungs- oder funktionstechnischer Probleme, keinerlei Daten liefert, kann das Gesamtsystem eine vollständige Objektdetektion und -klassifikation durchführen. Lediglich die Leistungsfähigkeit der Klassifikation ist in diesem Fall beeinträchtigt. Der kooperative Anteil wird dann deutlich, wenn der PMD-Sensor keine Ausgangsdaten liefert. In diesem Fall erfolgt keine Objektdetektion, somit kann auch die MV-Kamera keinen Beitrag zur Merkmalsextraktion leisten und der Klassifikation stehen keine Sensordaten zur Verfügung. In dieser Situation sind die beiden Sensoren voneinander abhängig. Diese Art der Sensordatenfusion ermöglicht eine Minimierung des Rechenaufwands in der Bildverarbeitung, da alle erforderlichen Bildverarbeitungsalgorithmen lediglich auf den Bildausschnitten erfolgen, die durchschnittlich nur circa 10 % bis 20 % der Bildfläche belegen.

5.4 MERKMALSEXTRAKTION AUS 2D-DATEN

Ein wesentliches Kriterium bei der Herangehensweise der Merkmalsberechnung im 2D-Bild ist der erforderliche Rechenaufwand. Da die Echtzeitfähigkeit und die Realisierbarkeit im Fahrzeug auf einem Steuergerät ein wichtiges Ziel sind, basiert nicht nur der Fusionsansatz, sondern auch die nachfolgenden Bildverarbeitungsschritte auf einfachen und effizient umsetzbaren Funktionen.

5.4.1 *Histogramme der Kantenrichtung*

Mit Hilfe des Fusionsansatzes aus Abschnitt 5.3 steht der 2D-BV ein Bildausschnitt M_1 zur Verfügung, der das Objekt, sofern es sich um einen Fußgänger handelt, vollständig umschließt. Die Anwendung der Sobel-Operatoren S_x und S_y auf ROI_1 liefert zwei Gradientenbilder G_x und G_y . Die Überlagerung $G = G_x + G_y$ der Gradientenbilder dient schließlich als Eingangsgröße für die Merkmalsextraktion (siehe Abbildung 5.5).

Es kann nun optional eine Verschmelzung der Konturen ROI_3 als Schnittmenge der Pixelmengen $M_3 = M_1 \cap M_2$ auf den Gradientenbildern erfolgen. Dies führt zu einer Reduktion der betrachteten Bildfläche mit dem Ziel, möglichst viele Gradienten beziehungsweise Kanten des Hintergrunds zu eliminieren, da diese nicht zur Beschreibung der Objektregion beitragen. Als weitere Option wird die Schnittmenge M_3 halbiert und nur noch die untere Hälfte verwendet, um die durch die Beine erzeugten Gradientenwerte stärker zu gewichten.

Um nun aus dem erzeugten und vorverarbeiteten Bildausschnitt ein skalares Merkmal zu generieren, werden die Häufigkeiten über Pixel mit Kanten in bestimmten Winkelbereiche gebildet. Abbildung 5.6 zeigt die Art der Aufteilung in 9 Winkelbereiche Φ_w mit $i = 1 \dots 9$. Im Allgemeinen



ABBILDUNG 5.5: Gradientenbild G auf der Pixelmenge M_1

wird die Richtung eines Gradienten zu

$$\alpha(x, y) = \arctan \left(\frac{G_y(x, y)}{G_x(x, y)} \right) \quad (5.44)$$

berechnet. Da aber nur die Zuweisung eines Winkels in einen Winkelbereich notwendig ist, und um eine Effizienzsteigerung zu erzielen, werden die Werte $\tan(\alpha_w)$, $w = 0 \dots 9$, in 20° -Schritten schon vorab berechnet. Für jedes Pixel, dessen Kantenrichtung untersucht werden soll, müssen dann

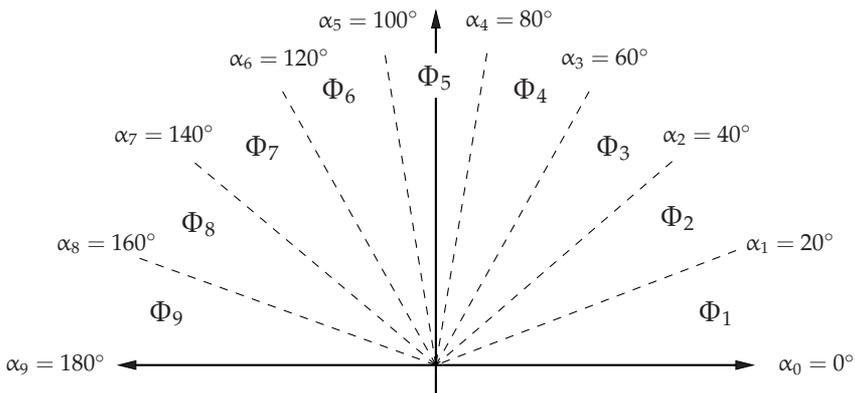


ABBILDUNG 5.6: Einteilung der Winkel α_w in Bereiche Φ_w

lediglich die Bedingungen

$$\tan(\alpha_w) < \frac{G_y(x, y)}{G_x(x, y)} < \tan(\alpha_{w+1}) \quad \text{mit } w = 0 \dots 8 \quad (5.45)$$

überprüft werden, und zwar ohne eine Neuberechnung der Tangens-Werte. Ist Gleichung (5.45) für ein w erfüllt, so gehört der untersuchte Gradientenwert zum Winkelbereich Φ_{w+1} . Die daraus ableitbaren Merkmalstypen berechnen sich mit $i, j \in \mathbb{N}^+$ über

$$\text{EDGEORIENTHIST} = \frac{\Phi_i + \delta}{\Phi_j + \delta} \quad (5.46)$$

$$\text{mit } i = 1 \dots 8 \wedge j = (i + 1) \dots 9$$

$$\text{SYMEDGEORIENTHIST} = \frac{\Phi_{5+i} + \Phi_{5-i} + \delta}{\Phi_{5+j} + \Phi_{5-j} + \delta} \quad (5.47)$$

$$\text{mit } i = 1 \dots 3 \wedge j = (i + 1) \dots 4$$

Da immer $\Phi_{i,j} \geq 0$ gilt, verhindert der additive Term $\delta \in \mathbb{R}^+$ mögliche Divisionen durch Null. Aus Gleichung (5.46) ergeben sich 36 mögliche Merkmale, aus Gleichung (5.47) ergeben sich 6 mögliche Merkmale.

5.4.2 Mustervergleich

Ein Vergleichsmuster $p_k(m', n')$ wird auf Basis einer Kreuzkorrelation mit den Bilddaten $g(x, y)$ verglichen. Die Skalierung des Musters $p_k(m', n')$, und damit die Wahl der Größe, werden von einem Parameter k bestimmt. Als Merkmal kann der größte Korrelationswert aus k unterschiedlich skalierten Vergleichsmustern erzeugt werden. Normiert man die Signale, so dass ihre Energien den Wert 1 haben, so ergibt sich für ein einzelnes bestimmtes k

$$\text{CORRELATION} = \max(p_k(m', n') \star g(x, y)) \quad (5.48)$$

$$= \max_{m, n} \left(\sum_{m', n'} p(m', n') g(m+m', n+n') \right). \quad (5.49)$$

Somit entspricht das Gesamtmerkmal dem größten Wert aller k Korrelationswerte

$$\text{MAXCORRELATION} = \max_k(\text{CORRELATION}). \quad (5.50)$$

Es sei an dieser Stelle angemerkt, dass Pose und Orientierung des Fußgängers starken Einfluss auf den maximal erreichbaren Korrelationswert haben.

5.5 EIGNUNG MÖGLICHER KLASSIFIKATIONSVERFAHREN

Zu Beginn dieses Abschnitt müssen einige Begriffe erklärt werden, die im Umgang mit der Klassifikation häufig uneinheitlich verwendet werden. Die in dieser Arbeit auftretenden Bezeichnungen sind in der folgenden Aufzählung kurz beschrieben [47].

Klasse Eine Zusammenfassung von »Dingen« mit gleichen oder ähnlichen Merkmalen, in unserem Fall also Fußgänger und Nicht-Fußgänger

Klassenbildung Zerlegung der Gesamtheit der Objekte in Klassen, in unserem Fall ist damit die Bildung von 2 Klassen gemeint

Klassifikation, Klassierung Einteilung der Objekte in Klassen, also das Treffen der Entscheidung über die Klassenzugehörigkeit eines Objektes

Klassifizierung Erstellung von Klassengrenzen, in unserem Fall das Herausarbeiten von Merkmalsgrenzen

Klassifikator, Klassifizierer Instanz, die eine Klassifizierung vornimmt (typische Klassifikatoren sind in den folgenden Abschnitten erklärt)

Klassifikationsverfahren Vorgehensweise des Klassifikators, also die Methoden des Klassifikators

Klassifikator-Framework (Software-)Umgebung, die mit Hilfe von Klassifizierungsmethoden das Erstellen von Klassifikatoren ermöglicht

Das nun zu lösende Problem der Klassifikation ist die Entscheidung darüber, mit welchen Algorithmen die gewonnenen Merkmale im Merkmalsraum \mathcal{M} zu einer Klassenentscheidung führen. Ziel ist es, einen unbekanntem Eingangsdatensatz

$$\mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_M \end{pmatrix} \quad (5.51)$$

auf die richtige Zielklasse S_i abzubilden. Da es sich bei der Klassifizierung um eine Methode des überwachten Lernens handelt¹, müssen Stichproben, also Trainings- und Testdaten, zur Verfügung stehen. Die Stichproben

¹Im Gegensatz dazu kommen Clusterbildungsverfahren mit der Methode des unüberwachten Lernens aus.

müssen in einem interaktiven Prozess von einem Benutzer mit Ground-Truth-Daten versehen werden, das heißt jedem \mathbf{m} muss a-priori ein S_i zugeordnet sein.

Mit Hilfe von Testreihen gewinnt man Wissen darüber, welche Messungen $\mathbf{m}_j \in \mathcal{M}$, $j \in \{1, 2, \dots, J\}$, ein Sensor liefert, der eine genügend große Anzahl Objekte aller Klassen S_i mit $i \in \{1, 2, \dots, K\}$ beobachtet. J repräsentiert die diskrete Anzahl möglicher Messungen, K ist gleich der Anzahl unterschiedlicher Klassen. Die Aufteilung der Eingangsdaten erfolgt dabei entweder in zwei Klassen (binäre Klassifikation) mit $\tilde{f} : \mathcal{M} \rightarrow \{S_1, S_2\}$, in endlich viele Klassen (Multiklassen-Klassifikation) mit $\tilde{f} : \mathcal{M} \rightarrow \{S_1, \dots, S_K\}$ oder in eine Abbildung auf reelle Zahlen (Regression) mit $\tilde{f} : \mathcal{M} \rightarrow \mathbb{R}$. In unserem Fall ist ein Zweiklassenproblem zu lösen. In den Abschnitten 5.5.1 bis 5.5.7 werden die zur Verfügung stehenden Klassifikationsalgorithmen vorgestellt und deren Eignung analysiert.

5.5.1 Klassifikation durch direkten Vergleich

Als trivialste Lösung einer Klassifikation ist ein direkter Vergleich denkbar. Dabei wird die Messung mit allen möglichen Erscheinungsformen der Zielklasse verglichen. Da die Variabilität der repräsentativen Objekte – und damit der benötigte Trainingsmessdatensatz – extrem groß ist, ist dieser Ansatz auf Grund des hohen Rechenaufwands nicht zielführend.

5.5.2 Schwellenentscheidung

Diese Klassifikationsmethode verwendet T Merkmale. Für jedes Merkmal m_t mit $t \in \mathcal{T} = \{1 \dots T\}$ eines Merkmalsvektors \mathbf{m} wird eine obere und untere Schranke $m_{t,\min}$ beziehungsweise $m_{t,\max}$ auf Basis von Expertenwissen angegeben. Die Schranken müssen derart festgelegt sein, dass sie alle typischen Objekte von einer der beiden Klassen möglichst vollständig umschließen. Da die Merkmale und die zugehörigen Schranken explizit repräsentiert sind, handelt es sich um ein symbolisches System². Die Umsetzung der Schwellenentscheidung ist wie folgt realisiert. Ist die Bedingung

$$m_{t,\min} \leq m_t \leq m_{t,\max} \quad \forall t \in \mathcal{T} \quad (5.52)$$

für einen Merkmalsvektor eines Objektes erfüllt, so wird keine Aussage über eine Klassenzugehörigkeit getroffen. Ist die Bedingung nicht erfüllt, wird es unmittelbar der Klasse Nicht-Fußgänger zugeordnet. Da nur einzelne Merkmale und keine Abhängigkeiten zwischen den Merkmalen betrachtet werden, führt jede einzelne Verletzung eines einzigen Grenzwert-

²Im Vergleich zu *symbolischen* Systemen ist in *subsymbolischen* Systemen ein Einblick in die erlernten Lösungswege nicht möglich

tes zu einer Klassifikation. Daher müssen die Grenzen großzügig gewählt werden. Dieser Algorithmus führt zur Reduktion der Anzahl der Objekte. Er wird daher als Vorklassifikation eingesetzt, siehe auch Abbildung 6.6.

Im Prinzip entspricht dies einem Boosting Algorithmus, da die einzelnen Entscheidungsschwellen jeweils einzelne Merkmale betreffen, die isoliert betrachtet nur einen unbedeutenden Beitrag leisten. Bei einem Boosting werden solche schwachen Klassifikatoren, auch *base classifiers* genannt, zu einem wirkungsvollen Klassifikator kombiniert.

5.5.3 *k*-Nächste-Nachbarn Klassifikation

Bei der Klassifikation mit Hilfe der nächsten Nachbarn handelt es sich um einen nicht-parametrischen Klassifikator. Es werden die zu jedem Merkmalsvektor \mathbf{m}_j nächstgelegenen $k \in \{1, 3, 5, \dots\}$ Stichprobenvektoren, die jeweils a-priori einer Klasse zugeordnet sind, bestimmt. Die Zuordnung von \mathbf{m}_j erfolgt dann genau in die Klasse, zu der die meisten dieser *k*-Nächsten-Nachbarn gehören. Ein Merkmalsvektor \mathbf{m}_j enthält dabei alle zur Verfügung stehenden Merkmale.

Die Leistungsfähigkeit und der Rechenaufwand des *k*-Nächste-Nachbarn Klassifikators ist stark abhängig von *k* und von der Gesamtanzahl verwendeter Merkmalsvektoren. In Bereichen des Merkmalsraumes mit hoher Stichprobendichte kann die Berechnung jedoch vereinfacht werden, wenn man die Stichprobenvektoren zu Häufungen zusammenfügt. Genau diese Clustermethode führt über die Hinzunahme von statistischen Kriterien, zum Beispiel Wahrscheinlichkeiten der einzelnen Klassen, zur Maximum-a-Posteriori Methode. Daher wird der Klassifikationsansatz mit Hilfe des *k*-Nächste-Nachbarn Klassifikators nicht weiter verfolgt.

5.5.4 *Maximum Likelihood Klassifikation*

Der in den 1920-er und 1930-er Jahren entwickelte Begriff der Likelihood Inferenz ist eine elementare Methode der statistischen Inferenz [55]. Als Basis dient die Likelihoodfunktion $L(\theta) = p(\mathbf{m}; \theta)$. Dabei sind \mathbf{m} die Beobachtungen einer Zufallsvariablen, wobei die dem Zufallsprozess X zu Grunde liegende Verteilung einer – bis auf den Parameter $\theta \in \Theta$ – bekannten Wahrscheinlichkeitsdichtefunktion $p(\mathbf{m}; \theta)$ entspricht. Die Wahrscheinlichkeit, dass nun die Beobachtungen \mathbf{m} des Prozesses X bei Wahl eines festen Parameters θ auftreten, wird als Likelihoodfunktion (kurz: Likelihood) bezeichnet. Insbesondere sind nun die Werte $\theta \in \Theta$ gesucht, die zu einer hohen beziehungsweise der höchsten Likelihood führen. Es entspricht also

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Theta} L(\theta) \quad (5.53)$$

dem Maximum-Likelihood-Schätzer (ML-Schätzer). Die Bestimmung des Maximums geschieht über die Suche nach Nullstellen der ersten Ableitung der Likelihood oder der ersten Ableitung der Log-Likelihoodfunktion $S(\theta) = dl(\theta)/d\theta$, die als Score-Funktion bezeichnet wird. Die Log-Likelihoodfunktion $l(\theta)$ ist häufig leichter zu berechnen und darf auf Grund der Monotonie der Logarithmusfunktion an Stelle der Likelihoodfunktion verwendet werden.

Der ML-Schätzer berechnet keine Wahrscheinlichkeiten – diese würden sich auf zukünftige Ereignisse auswirken – sondern er berechnet die Parameter einer aufgenommenen Verteilung, die von der Stichprobe am ehesten erzeugt worden sein könnte. Voraussetzung dafür ist natürlich die vom Anwender beziehungsweise Experten aufzustellenden Hypothese über die Wahrscheinlichkeitsdichteverteilung des mustererzeugenden (Zufalls-)Prozesses. Es handelt sich somit um einen parametrischen verteilungsbasierten Klassifikator. Der feste aber unbekannte Parameter θ – also in unserem Fall die gesuchte Klasse $S_i = \{S_1, S_2\}$ – entspricht beim ML-Schätzer jedoch nicht einer Zufallsvariablen, wohl aber bei dem im nächsten Abschnitt beschriebenen Maximum-a-Posteriori Klassifikator.

5.5.5 Maximum-a-Posteriori Klassifikation

Ist also die Beschaffenheit eines Systems bekannt, so können einige Ergebnisse und Effekte mit Hilfe der Wahrscheinlichkeitsrechnung in ihrer Auftretenswahrscheinlichkeit bestimmt werden. Die Vorgehensweise der Maximum-A-Posteriori (MAP) Klassifikation wird auch als Bayes-Methode bezeichnet. Das Wissen der Testreihen wird durch die bedingten Wahrscheinlichkeiten $p(\mathbf{m}_j|S_i)$ notiert. Trägt man diese als Matrix auf, erhält man ein Sensormodell

$$p(\mathbf{m}|S) = \begin{pmatrix} p(\mathbf{m}_1|S_1) & \cdots & p(\mathbf{m}_J|S_1) \\ \vdots & \ddots & \vdots \\ p(\mathbf{m}_1|S_K) & \cdots & p(\mathbf{m}_J|S_K) \end{pmatrix}. \quad (5.54)$$

Die Einträge $p(\mathbf{m}_j|S_i)$ entsprechen der a-priori Wahrscheinlichkeit (WSK), dass ein Merkmalsvektor \mathbf{m}_j auftritt, wenn es sich um ein Objekt der Klasse S_i handelt. Dies wird auch als *likelihood of observed measurements* bezeichnet. Die Elemente einer Zeile müssen aufsummiert Eins ergeben, also

$$\sum_{k=1}^J p(\mathbf{m}_k|S_i) = 1. \quad (5.55)$$

Für $K = J$ beschreiben die Beobachtungen direkt die existierenden Klassen, das heißt jeder Beobachtung ist eine Klasse zugeordnet. In diesem Fall

entsprechen die Einträge auf der Hauptdiagonalen den Wahrscheinlichkeiten für eine korrekte Klassifikation.

Im Allgemeinen kann allerdings nicht $J = K$ vorausgesetzt werden, denn es kann eine größere Anzahl an Beobachtungstypen als Klassentypen auftreten. In unserem Fall betrachten wir ausschließlich $K = 2$ Klassen, J kann jedoch sehr groß werden. Man erhält für jede Zeile in Gleichung (5.54), also für jede Klasse, eine Wahrscheinlichkeitsverteilungsfunktion bestehend aus J diskreten Werten. Geht J gegen unendlich, so bildet sich eine Wahrscheinlichkeitsdichtefunktion. Dies ist der Fall, wenn die Beobachtungen noch nicht fest zu den K Klassen hinzugeordnet sind bzw. nicht zugeordnet werden können.

Um aus dem Sensormodell (5.54) das inverse Sensormodell $P(S|\mathbf{m})$ entsprechend (5.56) herzuleiten, ist zusätzlich a-priori-Wissen $P(S_i)$ und eine Normierung $p(\mathbf{m})$ notwendig. Das inverse Sensormodell gibt die a-posteriori-Wahrscheinlichkeit von S gegeben \mathbf{m} an, also welche Klasse S_i *tatsächlich* vorliegt, wenn der Sensor die Beobachtung \mathbf{m}_j macht. $P(S_i)$ gibt die Auftretenswahrscheinlichkeit der einzelnen Klassen an. Es ist nach dem Satz von Bayes

$$P(S_i|\mathbf{m}_j) = \frac{P(S_i) \cdot p(\mathbf{m}_j|S_i)}{p(\mathbf{m}_j)} \quad (5.56)$$

$$\Leftrightarrow P(S_i|\mathbf{m}_j) = \frac{P(S_i) \cdot p(\mathbf{m}_j|S_i)}{\sum_{k=1}^K p(\mathbf{m}_j|S_k)P(S_k)} \quad (5.57)$$

und das inverse Sensormodell ergibt sich zu

$$P(S|\mathbf{m}) = \begin{pmatrix} P(S_1|\mathbf{m}_1) & \cdots & P(S_K|\mathbf{m}_1) \\ \vdots & \ddots & \vdots \\ P(S_1|\mathbf{m}_J) & \cdots & P(S_K|\mathbf{m}_J) \end{pmatrix}. \quad (5.58)$$

Auch in (5.58) ergibt die Summe über jede Zeile Eins, also ist

$$\sum_{k=1}^K P(S_k|\mathbf{m}_j) = 1. \quad (5.59)$$

Der Bayes-Klassifikator ist also eine Erweiterung des ML-Schätzers um das a-priori-Wissen der Klassenhäufigkeiten. Sinnvollerweise wählt man als Optimalitätskriterium die Minimierung von Fehlentscheidungen. Es ergibt sich ein Klassifikator, der die a-posteriori-Wahrscheinlichkeit entsprechend Gleichung (5.61) maximiert, also eben ein Maximum-A-Posteriori Klassifikator (MAP-Klassifikator).

$$\hat{\theta}_{\text{MAP}} = \arg \max_{\theta \in \Theta} P(S_i|\mathbf{m}) \quad (5.60)$$

$$= \arg \max_{\theta \in \Theta} p(\mathbf{m}|S_i)P(S_i) \quad (5.61)$$

Daraus leitet sich die Entscheidungsbedingung ab, mit der beim Zweiklassenproblem eine Klasse $S_i \in \{S_1, S_2\}$ ausgewählt wird. Aus (5.56) und (5.61) ergibt sich der Vergleich

$$P(S_1)p(\mathbf{m}|S_1) \stackrel{?}{\geq} P(S_2)p(\mathbf{m}|S_2). \quad (5.62)$$

Falls der größere der beiden Terme aus Gleichung (5.62) im Vergleich zum normierenden Nenner aus Gleichung (5.56) zu klein ist, ist die Entscheidung nicht sicher genug. Aus diesem Grund kann man eine Zurückweisung unter der Bedingung

$$p(\mathbf{m}|S_i)P(S_i) < \beta \cdot p(\mathbf{m}) \quad (5.63)$$

$$\Rightarrow P(S_i|\mathbf{m}) < \beta \quad (5.64)$$

integrieren. β gibt die Schwelle für die Wahrscheinlichkeit der Klassenzugehörigkeit $P(S_i|\mathbf{m})$ an, die für eine Klassenzuweisung mindestens erreicht sein muss. Ist diese erfüllt, so sollte keine Klassenzuweisung erfolgen. Für den Fall identischer Klassenhäufigkeiten $P(S_1) = P(S_2)$ vereinfacht sich Gleichung (5.62) zu

$$p(\mathbf{m}|S_1) \stackrel{?}{\geq} p(\mathbf{m}|S_2) \quad (5.65)$$

und ergibt wieder den ML-Klassifikator aus Gleichung 5.53.

Es muss wiederum eine Annahme über die klassenspezifische Verteilung getroffen werden. In unserem Fall bietet sich die Wahl einer Gauß-Verteilung an. Diese ist im N -dimensionalen Raum mit

$$p(\mathbf{m}) = \frac{1}{\sqrt{(2\pi)^N \det(\mathbf{C}_{xx})}} e^{-\frac{1}{2}(\mathbf{m}-\mu_m)^T \mathbf{C}_{xx}^{-1}(\mathbf{m}-\mu_m)} \quad (5.66)$$

beschrieben. Dabei ist $\mathbf{C}_{xx} = E\{(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T\}$ die Autokovarianzmatrix und $\mu_m = E(\mathbf{m})$ der Vektor des Erwartungswertes von \mathbf{m} . Aus der Positiv-Semidefinitheit und der Symmetrie von \mathbf{C}_{xx} folgt, dass eine Ortskurve mit konstanter Wahrscheinlichkeitsdichte jeweils ein Ellipsoid ergibt. Für den Fall $N = 2$ degeneriert die Autokovarianzmatrix zur Varianz σ^2 und das Ellipsoid zur Ellipse. Eine Visualisierung ist in Abbildung 5.7 gegeben.

Bei dem hier vorgestellten Ansatz mit Hilfe der Bayes-Methode wurde also die klassenspezifische Verteilungsdichte $p(\mathbf{m}|S)$ durch statistische Modelle angenähert, also durch die Schätzung der Parameter einer Gauß-Verteilung. Es existiert jedoch eine weitere Herangehensweise zum Entwurf eines Klassifikators, indem man die a-posteriori Wahrscheinlichkeitsdichte $p(S|\mathbf{m})$ auswertet³. Diese kann durch eine Funktionsapproximation

³Anhand des Bayes-Theorem aus Gleichung (5.56) ist die Analogie im Sinne einer Gleichwertigkeit zwischen beiden Ansätzen erkennbar, sofern die Klassenhäufigkeit $P(S_i)$ als gleichverteilt zu $1/K$ angenommen werden, und der Nenner unabhängig von jeglicher Klassenverteilung ist.

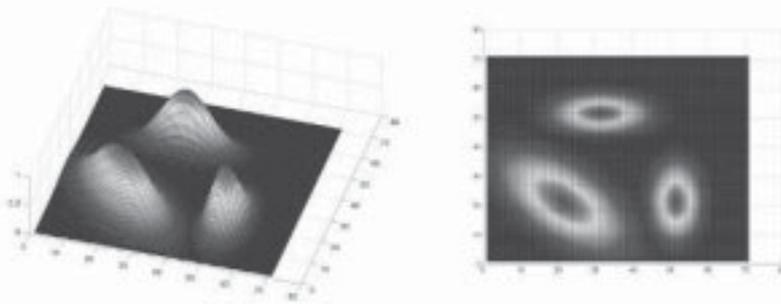


ABBILDUNG 5.7: Zwei perspektivische Darstellungen von drei Normalverteilungen im Fall eines zweidimensionalen Merkmalraumes

mit Hilfe von linear verknüpften linearen oder nichtlinearen Basisfunktionen beschrieben werden. Sie kann zum Beispiel durch nichtlineare Regression mit Polynomen oder mit Hilfe Künstlicher Neuronaler Netze durchgeführt werden.

Zusammenfassend lässt sich festhalten, dass der populäre Bayes Klassifikator durch Einfachheit, sowohl im Rechenaufwand als auch in der Verständlichkeit, überzeugt. Er findet in vielen Alltagsproblemen Anwendung. Bei sehr komplexen Problemen stößt man jedoch schnell an die Grenzen der Leistungsfähigkeit. Aus diesem Grund wurde der Klassifikator nur in der Entwicklungsphase untersucht und fließt in die Ergebnisauswertung nicht mit ein.

5.5.6 Künstliche Neuronale Netze (KNN)

KNN wurden in Anlehnung an die Funktionsweise des menschlichen Gehirns entwickelt. Dabei werden mehrere Eingangssignale auf Knoten oder Neuronen weitergeleitet, die Informationsverarbeitung leisten. Die Verarbeitung komplexer Aufgaben werden jedoch erst durch die Vernetzung der Eingangssignale ermöglicht. Somit ist prinzipiell die Nachbildung des zentralen Nervensystems von Lebewesen möglich, bei denen mit Hilfe von Neuronen, bestehend aus jeweils einem Soma (Zellkörper), einem oder mehreren Dendriten (Eingängen) und einem Axon (Ausgang) stärkende und schwächende Wirkungen auf Nervenimpulse möglich sind. Die Vernetzung der Neuronen, in Analogie zur linearen Verknüpfung der Basisfunktionen, erfolgt über Synapsen. Eine Synapse verbindet ein Axon mit einem oder mehreren Dendriten nachfolgender Neuronen. Zusätzlich spielen für die Aktivierung eines Neurons die Stärken der Verbindungen zwischen Axone und Dendriten eine Rolle.

Als einfachstes Beispiel dient ein einschichtiges Perzeptron aus Abbildung 5.8. Dort existieren n Eingänge, die über n Kanten mit den Gewichten w_i mit genau einem Neuron verbunden sind. Im Perzeptron selbst findet eine Schwellwertentscheidung statt. Zusätzlich fließt ein Offset b , ein sogenannter Bias, in die Schwellwertfunktion ein, der als Gewicht w_0 einer konstanten Eingabe m_0 interpretiert werden kann. Die Aktivierungsfunktion $f(s)$ kann einer linearen Funktion

$$f(s) = k(b + \sum_{i=1}^n m_i w_i), \quad (5.67)$$

einer Sprungfunktion $f(s) = 1$ falls $\sum_i m_i w_i \geq b$ oder einer Sigmoid-Funktion, zum Beispiel der Fermifunktion

$$f(s) = \frac{1}{1 + \exp(-s)} \quad \text{mit} \quad s = b + \sum_{i=1}^n m_i w_i, \quad (5.68)$$

entsprechen. Im Rahmen des Lernprozesses sind nun die Gewichtsvektoren w_i und ein Schwellwert für das Perzeptron gesucht. Diese werden in einem $(n + 1)$ -dimensionalen Gewichteraum abgebildet.

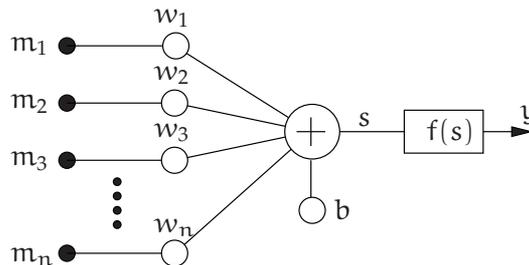


ABBILDUNG 5.8: Perzeptron mit n gewichteten Eingängen für den Eingangsvektor \mathbf{m} , einem Bias b , einer Schwellwertfunktion $f(s)$ und einem Ausgang $y = f(\langle \mathbf{w}, \mathbf{m} \rangle + b)$.

Daneben existieren die Vektoren des Eingaberaumes der Dimension n . Jeder Punkt im Gewichteraum entspricht einer Realisierung des Perzeptrons und teilt den Eingaberaum durch den ihm per Schwellwert zugewiesenen Ausgangswert y in zwei Teile. Diese Raumhälften sind durch eine Hyperebene der Dimension $(n - 1)$ getrennt. Damit ist eine Analogie zu der Support Vector Machine (SVM) aus dem weiter unten folgenden Abschnitt 5.5.7 erkennbar. Umgekehrt entspricht ein Punkt im Eingaberaum einer Hyperebene im Gewichteraum. Da für komplexe Klassifikationsaufgaben die einfache Trennung des Merkmalsraumes, also des Eingaberau-

mes, nicht ausreicht, verwendet man im Allgemeinen mehrschichtige Perzeptrons, so wie sie in Abbildung 5.9 beispielhaft skizziert sind. Dort ist

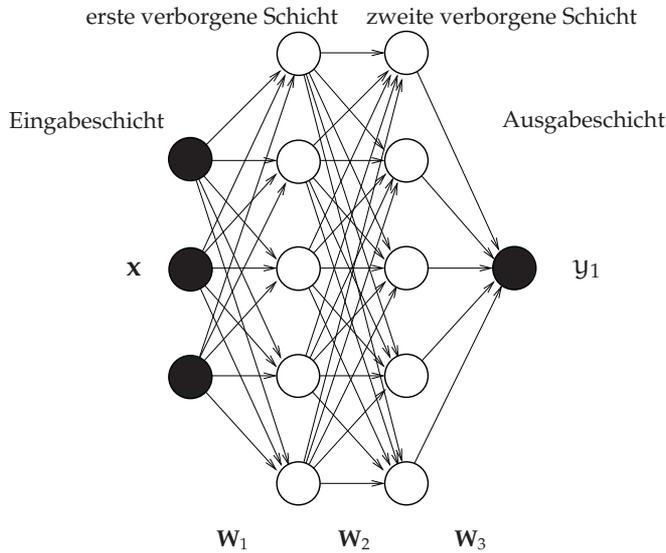


ABBILDUNG 5.9: Topologie eines dreischichtigen Perzeptrons zur Realisierung eines Klassifikators mit beliebig linear begrenzten Gebieten

jedes Neuron einer Schicht, mit Ausnahme der Eingabeschicht, mit jedem Neuron der vorherigen Schicht verbunden und reduziert dabei die eingehenden Signale auf jeweils ein Ausgangssignal.

$$\mathbf{W}_1 = \begin{pmatrix} w_{1,1} & w_{1,2} & w_{1,3} \\ w_{2,1} & w_{2,2} & w_{2,3} \\ w_{3,1} & w_{3,2} & w_{3,3} \\ w_{4,1} & w_{4,2} & w_{4,3} \\ w_{5,1} & w_{5,2} & w_{5,3} \end{pmatrix} \quad (5.69)$$

Die erste verborgene Schicht mit n_1 Neuronen besitzt eine Matrix \mathbf{W}_1 der Gewichtungskoeffizienten der Dimension $n_1 \times n_0$ und erzeugt n_1 Trennebenen im Merkmalsraum. n_0 entspricht dabei der Anzahl der Neuronen in der unmittelbar davorliegenden Eingabeschicht. Die Trennebenen dieser ersten Schicht liefern Eckpunkte eines Hyperwürfels. In der zweiten Schicht werden die Eckpunkte im Sinne einer logischen UND-Verknüpfung derart miteinander in Beziehung gesetzt, dass sich linear begrenzte konvexe Gebiete, sogenannte Polyeder, ausbilden. Sie wird mit \mathbf{W}_2 der Dimension $n_2 \times n_1$ beschrieben. In einer dritten Schicht können nun mit Hilfe der Gewichtematrix \mathbf{W}_3 diese Gebiete so gruppiert werden, dass auch nicht zusammenhängende Polyeder verknüpft werden. Damit sind logi-

sche ODER-Operationen realisierbar. Dies ist in Abbildung 5.10 dargestellt, die Berechnung der Ausgabe y_3 erfolgen gemäß Gleichung 5.70 bis 5.72. y_i entspricht der Ausgabe der i -ten verborgenen Schicht.

$$y_3 = W_3 y_2 \quad (5.70)$$

$$\Leftrightarrow y_3 = W_3 W_2 y_1 \quad (5.71)$$

$$\Leftrightarrow y_3 = W_3 W_2 W_1 m \quad (5.72)$$

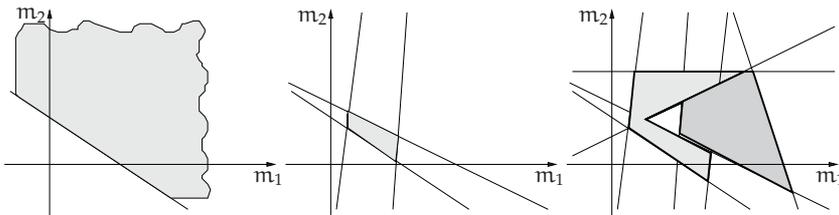


ABBILDUNG 5.10: Wirkungen der ersten, zweiten und dritten Schicht eines mehrschichtigen Perzeptrons zur Bildung beliebig geformter zusammenhängender Polyeder

Man unterscheidet bei KNN, in Abhängigkeit der Signalübertragungsrichtung, zwischen Feed-Forward Netzen ohne Rückkopplung und Feed-Back Netzen mit Rückkopplung. Da durch die rekursiven Zweige die Ausgaben auch von zeitlich zurückliegenden Eingaben abhängen, spricht man von Speicherelementen.

Für das Training der KNN werden zu Beginn die Gewichtematrizen W und die Schwellenwerte b mit Zufallswerten initialisiert. Diese müssen nun so angepasst werden, dass die Klassifikationsaufgabe möglichst gut erfüllt werden kann⁴. Die Trainingsmethoden werden dabei in die Kategorien überwachtes Lernen, unüberwachtes Lernen und Lernen mit Bestätigung unterteilt. Diese werden im Folgenden kurz erläutert.

Im Fall des überwachten Lernens liefert der Vergleich aus Soll- und Ist-Ausgabe eine Fehlerrate, die eine Neuberechnung der Netzparameter steuert. Nach diesem Prinzip funktioniert der Error-Back-Propagation Algorithmus. Man unterscheidet dort noch zwischen verstärkendem Lernen, bei dem die Neuberechnung der Gewichte durch die Anzahl an Fehlklassifikationen gesteuert wird, und dem korrigierendem Lernen, bei dem die Neuberechnung der Gewichte noch zusätzlich von der Fehlergröße abhängt. Der Error-Back-Propagation Algorithmus ist die bekannteste Form

⁴In sehr forschungslastigen Anwendungen kann auch die Art der Topologie durch das Lernen verändert werden.

des überwachten Lernens von KNN. Im Fall des unüberwachten Lernens versucht das Netz, eine Clusterung ähnlicher Eingabevektoren selbst zu organisieren. Ein Werkzeug dieser Netze sind Self-Organizing Maps (SOM), auch Kohonenkarten oder -netze genannt. Dort wird eine Gesetzmäßigkeit der Menge von Eingangsvektoren ohne den Soll-Ist-Vergleich festgestellt [56]. Als Mittelweg zwischen überwachtem und unüberwachtem Lernen im Sinne der Kooperation zwischen dem KNN und dem Umfeld dient das Lernen mit Bestätigung. Dort ist ein Signal der Bestätigung beziehungsweise des Ablehnens benötigt, welches durch einen *Trainer* gesetzt werden muss.

Für den meist iterativen Lernprozess nimmt die Lernregel die zentrale Stellung ein. Zwar gibt es in der Theorie eine Vielzahl unterschiedlicher Modelle, doch lassen sich die meisten auf eine Grundhypothese zurückführen. Diese wird mit der Hebb'schen Lernregel [57] beschrieben und am Beispiel des Perzeptrons aus Abbildung 5.8 mathematisch durch

$$\Delta \mathbf{w} = \varepsilon \cdot \mathbf{y} \cdot \mathbf{m} \quad (5.73)$$

$$\Leftrightarrow \Delta \mathbf{w} = \varepsilon \cdot f(\langle \mathbf{w}, \mathbf{m} \rangle + b) \cdot \mathbf{m} \quad (5.74)$$

beschrieben. Sie besagt, dass genau dann die Synapsenstärke erhöht wird, wenn der Stimulus *und* die Aktivität hohe Werte gleichen Vorzeichens annehmen. $\varepsilon > 0$ ist ein Parameter, der die Größe des Lernschrittes angibt.

Die weiteren Realisierungs- und Trainingsmöglichkeiten von KNN sind sehr vielseitig, weitere Informationen sind in [58] gegeben.

Als Vorteil gilt die gute Parallelisierbarkeit, die besonders im automotiven Bereich die Chance auf kostengünstige und schnelle Field Programmable Gate-Array (FPGA) Umsetzung erlaubt. Nachteile sind die Herausforderung bei der Auswahl des Netzwerkmodells, der Wahl der Anzahl an Schichten und den jeweiligen Knoten. Da ausgehend von einem oder mehreren Startparametern jeweils entsprechend einer Kosten- beziehungsweise Zielfunktion optimiert wird, ist nicht sichergestellt, dass das gefundene Optimum dem globalen Optimum entspricht. Der Trainingsaufwand kann sehr hoch sein, ein fertig trainiertes KNN benötigt im Allgemeinen jedoch relativ wenig Rechenaufwand.

5.5.7 Support Vector Machines

Support Vector Machine (SVM) sind eine relativ neue Methode in der Klassifikation und werden auch im Bereich der Fußgängererkennung eingesetzt, ein Beispiel dazu liefert [59]. Sie gelten als sehr mächtig, insbesondere im Hinblick auf die Trennbarkeit von stark nichtlinear verzahnten Verteilungen. Im Prinzip nutzen sie den Ansatz, die Dimensionalität des Merkmalsraums \mathcal{M} zu erhöhen, um die Trennfähigkeit im neuen Merkmalsraum \mathcal{H} zu vereinfachen. Das Besondere bei SVM ist nun, dass die Komplexität

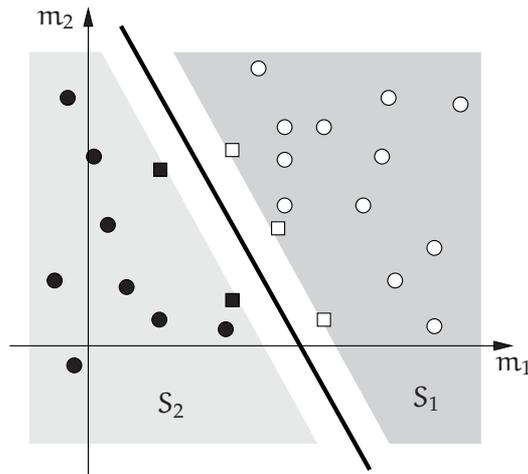


ABBILDUNG 5.11: 2D-Merkmalraum mit einer Trennebene, die alle Vektoren der beiden Klassen (Weiss und Schwarz) optimal trennt. Alle SV sind als Quadrate gekennzeichnet

nicht mehr von der Dimensionalität des Merkmalsraumes, sondern von der Anzahl Trainingssamples abhängt. Dadurch kann $\dim[H]$ prinzipiell einen beliebig hohen Wert annehmen. Im Folgenden ist das Grundprinzip der SVM erklärt. Dies beinhaltet eine kurze Abhandlung der Vapnik-Chervonenkis-Theorie (kurz VC-Theorie), den Fall der linearen Trennbarkeit der Trainingssamples, den Fall der nicht-linearen Trennbarkeit und den Fall der nicht-linearen Erweiterung mit Hinzunahme von Kernelfunktionen. Zusätzlich wurde eine Erweiterung mit klassenspezifischer Bestrafung entwickelt, die im Anschluss beschrieben wird.

Eine SVM [60] nutzt einen statistischen Ansatz der Lerntheorie. Das Training arbeitet dabei mit der Methode des überwachten Lernens. Bei Unterscheidung zwischen insgesamt 2 Klassen wird, ähnlich wie bei einschichtigen KNN, eine Entscheidungsgrenze

$$\langle \mathbf{w}, \mathbf{m} \rangle + b = 0 \quad (5.75)$$

festgelegt. Gleichung 5.75 ist für alle Punkte \mathbf{m} erfüllt, die auf dieser Entscheidungsgrenze liegen. \mathbf{w} ist der Normalenvektor im \mathbb{R}^M und b eine Konstante. Abbildung 5.11 dient zur Veranschaulichungen für den 2-dimensionalen Fall.

Es sei zunächst angemerkt, dass die SVM zwar kein »Allheilmittel« darstellt, da die Merkmalsgewinnung nach wie vor sehr entscheidend ist, jedoch können stark nichtlinear verzahnte Verteilungen der Klasseneigenschaften deutlich besser als mit Bayes und KNN behandelt werden. Der

Ansatz geht davon aus, dass alle Daten mit einem stochastischen Prozess unbekannter Verbundverteilung $P(\mathbf{m}, S)$ beschrieben sind. Weiterhin ist kein a-priori Wissen notwendig.

VC-Theorie

Gehen wir an dieser Stelle kurz auf die Themen der Generalisierungsfähigkeit und der Fehlerabschätzung ein. Gesucht ist eine Zuordnungsfunktion

$$f(\mathbf{m}, \vec{\alpha}) : \mathbf{m} \rightarrow S, \quad (5.76)$$

welche durch die unbekannt Parameter des Vektors $\vec{\alpha}$ charakterisiert wird und den Erwartungswert des Zuordnungsfehlers $R(\vec{\alpha}) = E\{R_{\text{test}}(\vec{\alpha})\}$ minimiert. Man braucht natürlich einen großen Testdatensatz, um nahe an die asymptotische Aussage zu kommen.

Im Allgemeinen wird mit einem endlich großen Lerndatensatz trainiert und mit einem endlich großen Testdatensatz getestet. Man erhält damit den empirischen Fehler $R_{\text{emp}}(\vec{\alpha})$, $\vec{\alpha}$ beschreibt dabei den Parametersatz der jeweiligen SVM. Auf Basis dieses Fehlers muss nun die Generalisierungsfähigkeit abgeschätzt werden, um eine Aussage über den wahren Fehler $R_{\text{wahr}}(\vec{\alpha})$, also die wahre Klassifikationsleistung der SVM, zu bekommen. Der wahre Fehler ist asymptotisch definiert als der Grenzwert des empirischen Fehlers für eine immer weiter wachsende Zahl von Beobachtungen [47].

Wird ein Klassifikator auf einer sehr kleinen Datenmenge optimiert, so ist er überangepasst und hat eine geringe Generalisierungsfähigkeit. Da immer nur ein endlich großer Datensatz zur Verfügung steht, muss die Abschätzung der Generalisierungsfähigkeit beziehungsweise die Abschätzung von R_{wahr} algorithmisch gestützt werden. Dieses Vorgehen ist im Folgenden kurz erläutert.

Ist ein Set von Trainingsvektoren mit L Datensätzen und der Verbundverteilung $P(\mathbf{m}, S)$ gegeben, so gilt für dieses Trainingsset mit den dazugehörigen Labels

$$\{(\mathbf{m}_i, S_i) \in P(\mathbf{m}, S)\} \quad (5.77)$$

mit

$$i = 1 \dots L, \quad \mathbf{m}_i \in \mathbb{R}^M, \quad S_i \in \{+1, -1\}. \quad (5.78)$$

Weiterhin ist keinerlei Vorwissen über die Verteilungsdichten vorhanden und auch nicht notwendig. Der Zuordnungsfehler

$$R(\vec{\alpha}) = E\{R_{\text{test}}(\vec{\alpha})\} \quad (5.79)$$

$$\Leftrightarrow R(\vec{\alpha}) = \int \frac{1}{2} |S - f(\mathbf{m}, \vec{\alpha})| dP(\mathbf{m}, S) \quad (5.80)$$

$$\Leftrightarrow R(\vec{\alpha}) = \int \frac{1}{2} |S - f(\mathbf{m}, \vec{\alpha})| p(\mathbf{m}, S) d\mathbf{m} dS \quad (5.81)$$

ist nicht lösbar, da $p(\mathbf{m}, S)$ unbekannt ist. Es muss ein anderer Ansatz für das zentrale Problem der statistischen Lerntheorie gefunden werden, also eine Antwort auf die Frage, wann ein niedriger Trainingsfehler zu einem niedrigen wahren Fehler führt. Da sich der empirische Fehler, also die Fehlerrate für den gegebenen Trainingsdatensatz, über

$$R_{\text{emp}}(\vec{\alpha}) = \frac{1}{2L} \sum_{i=1}^L |S - f(\mathbf{m}, \vec{\alpha})| \quad (5.82)$$

berechnen lässt, brauchen wir jetzt noch eine Abschätzung, um vom empirischen auf den wahren Fehler zu schließen. Die geforderte Abschätzung, also der Extrapolation, ist über die Vapnik-Chervonenkis Abschätzung [61] möglich. Die VC-Theorie besagt, dass mit einer WSK von $(1 - \eta)$ die Abschätzung

$$R_{\text{wahr}}(\vec{\alpha}) \leq R_{\text{emp}}(\vec{\alpha}) + \Phi(h, L, \eta) \quad (5.83)$$

mit

$$\Phi(h, L, \eta) = \sqrt{\frac{h(\log(2L/h) + 1) - \log(\eta/4)}{L}} \quad (5.84)$$

gilt. Dabei entspricht L der Anzahl an Trainingsbeispielen und h der VC-Dimension des verwendeten Hypothesenraums. Man kann sich h als Maß für die »Anschmiegsamkeit« einer Funktion an gegebene Punkte, in unserem Fall sind dies die Trainingsamples aus Gleichung (5.77) im \mathbb{R}^M , vorstellen. Der Wert η kann zum Beispiel zu $\eta = 0,1$ gewählt werden, somit erreicht man eine WSK für die Gültigkeit der Abschätzung von 90 %. Die Funktion $\Phi(h, L, \eta)$ aus (5.84) wird als VC-Konfidenz bezeichnet. Sie ist bemerkenswerter Weise *unabhängig* von der Verbundverteilung $P(\mathbf{m}, S)$, vorausgesetzt die Trainings- und Testdaten werden statistisch unabhängig ausgewählt. Mit steigendem h wächst auch Φ . Ein größerer Lerndatensatz L führt zu einem kleineren Funktionswert, also zu einer besseren Abschätzung des echten Fehlers. Als Schwachpunkt sei an dieser Stelle angemerkt, dass die VC-Abschätzung nur sehr ungenau und praktisch wenig verwertbar ist.

Nun kann natürlich der empirische Fehler beliebig klein gemacht werden, indem eine entsprechende Funktionenklasse, also ein hohes h , gewählt wird, etwa Polynome mit sehr hohem Polynomgrad. Man kommt dann jedoch zum »Overfitting«, also zu einer geringeren Generalisierungsfähigkeit. Ein zu kleines Polynom bzw. eine zu kleine Funktionenklasse hat jedoch einen hohen empirischen Fehler zur Folge. Der Gesamtfehler ist in Abbildung 5.12 dargestellt. Zur Lösung einer Klassifikationsaufgabe mit einer SVM ist daher der Trade-Off zwischen dem empirischen Fehler und der Generalisierungsfähigkeit zu optimieren.

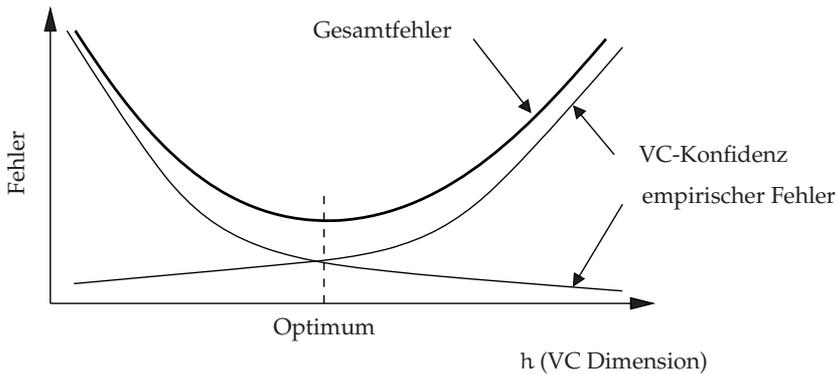


ABBILDUNG 5.12: der Gesamtfehler der Klassifikation ergibt sich empirischem Fehler und VC-Konfidenz

Allgemein gilt, dass Hyperebenen in \mathbb{R}^M eine VC-Dimension von $M+1$ ergeben. Bei Polynomen wächst die VC-Dimension stark mit dem Grad des Polynoms. Da wir eine möglichst kleine VC-Konfidenz suchen, sollten wir ein Polynom minimaler VC-Dimension wählen, dies entspricht idealerweise einer Hyperebene. Eine Hyperebene mit maximalem Rand ist also gleich der trennenden Hyperebene mit minimaler VC-Dimension. Die Support-Vektoren sind dann genau die Vektoren, die den maximalen Rand der Hyperebene mit kleinster VC-Dimension stützen. Da Aussagen über die VC-Dimension bei SVM möglich sind, ist auch die Abschätzung über den Gesamtfehler möglich. Wir können also die VC-Abschätzung mit einer gewissen Schranke η anwenden.

Fall der linearen Trennbarkeit

Im einfachsten Fall lassen sich alle Trainingsamples mit einer Hyperebene

$$\langle \mathbf{w}, \mathbf{m} \rangle + b = 0 \quad (5.85)$$

trennen, so wie es zu Beginn dieses Abschnitts erläutert wurde. Jedoch reicht es nicht aus, die Trennebene an eine beliebige Position zu setzen. Es wird nämlich zusätzlich gefordert, dass der Abstand der Trennebene zu den Trainingsamples optimal ist. Optimal bedeutet, dass die Trennebene orthogonal zur kürzesten Verbindungslinie zwischen den konvexen Hüllen mit gleich großem Abstand zu den Support-Vektoren beider Klassen liegt. Neben die Hyperebene können zusätzliche Hyperebenen H_1 und H_2 gelegt werden, so dass gilt: Für jeden Vektor \mathbf{m}_1 der Klasse $S_1 = +1$ gilt

$$\langle \mathbf{w}, \mathbf{m}_1 \rangle + b \geq +1, \quad (5.86)$$

für jeden Vektor \mathbf{m}_2 der Klasse $S_2 = -1$ entsprechend

$$\langle \mathbf{w}, \mathbf{m}_2 \rangle + b \leq -1. \quad (5.87)$$

Zusammengefasst ergibt sich

$$S_i (\langle \mathbf{w}, \mathbf{m}_i \rangle + b) \geq 1. \quad (5.88)$$

Alle Vektoren, für die die Gleichheitsbedingung gilt, liegen dann genau auf den Hyperebenen $H_1 : \langle \mathbf{w}, \mathbf{m}_1 \rangle + b = +1$ und $H_2 : \langle \mathbf{w}, \mathbf{m}_2 \rangle + b = -1$. Bei richtiger Wahl von \mathbf{w} ist nun der Abstand d_+ beziehungsweise d_- zwischen der trennenden Hyperebene und H_1 beziehungsweise H_2 durch $1/\|\mathbf{w}\|$ gegeben. Der Abstand zwischen H_1 und H_2 wird als »Margin« bezeichnet. Da die Ebenen den gleichen Normalenvektor haben, liegen sie parallel zueinander. Unser Ziel ist eine trennende Hyperebene mit maximaler Margin, daher muss der Abstand $2/\|\mathbf{w}\|$ maximiert werden. Das primale Optimierungsproblem lautet somit

$$\text{Minimiere} \quad J(\mathbf{w}, b) = \frac{1}{2} \|\mathbf{w}\|^2 = \frac{1}{2} \langle \mathbf{w}, \mathbf{w} \rangle \quad (5.89)$$

$$\text{Nebenbedingung} \quad [S_i (\langle \mathbf{w}, \mathbf{m}_i \rangle + b) \geq 1] \quad \forall i. \quad (5.90)$$

Um das Optimierungsproblem mit seiner Nebenbedingung handhabbarer zu gestalten, kann es durch die Einführung einer Lagrange Funktion

$$L_P(\mathbf{w}, b, \vec{\alpha}) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^L \alpha_i (S_i (\langle \mathbf{w}, \mathbf{m}_i \rangle + b) - 1) \quad \text{mit } \alpha_i \geq 0 \quad (5.91)$$

umformuliert werden. α_i bezeichnen darin die Lagrange-Multiplikatoren. Diese Lagrangefunktion beschreibt das primale Problem, für das Dank seiner Konvexität ein äquivalentes duales Problem formuliert werden kann. Dieses wird als Wolf-duales Optimierungsproblem bezeichnet und kann numerisch iterativ gelöst werden.

$$\text{Maximiere} \quad L_D = \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i=1}^L \sum_{j=1}^L S_i S_j \alpha_i \alpha_j \mathbf{m}_i \mathbf{m}_j \quad (5.92)$$

$$\text{Nebenbedingung} \quad \alpha_i \geq 0 \quad \text{und} \quad \sum_{i=1}^L \alpha_i S_i = 0 \quad (5.93)$$

Die Lösung ist eindeutig, da es sich um ein positiv semidefinites Problem – also konvexes Optimierungsproblem mit linearer Nebenbedingung – handelt, bei dem immer ein globales Maximum existiert. Das ist einer der Stärken der SVM im Vergleich zu Neuronalen Netzen, die häufig nur suboptimale Lösungen liefern. Der die Ebene beschreibende Normalenvektor \mathbf{w}

wird in dieser Funktion nicht mehr explizit aufgeführt, sondern er wird durch die Summe über den Skalarprodukten der Merkmalsvektoren beschrieben. Daher wird die Trennebene gar nicht explizit berechnet, sondern durch die Linearkombination der Trainingsvektoren. Die Komplexität hängt somit auch nicht mehr von der Dimensionalität M des Merkmalsraumes ab, sondern von der Anzahl Support Vektoren. Die Lösung von Gleichung (5.92) und (5.93) liefert die gewünschten Lagrange Parameter α_i

$$\mathbf{w}^* = \sum_{i=1}^L \alpha_i S_i \mathbf{m}_i = \sum_{\mathbf{m}_i \in SV} \alpha_i S_i \mathbf{m}_i \quad (5.94)$$

$$\Delta b^* = -\frac{1}{2} \max_{S_i=-1} (\langle \mathbf{w}, \mathbf{m}_i \rangle) + \min_{S_i=+1} (\langle \mathbf{w}, \mathbf{m}_i \rangle). \quad (5.95)$$

Für alle Support-Vektoren gilt $0 < \alpha_i < \infty$, für alle anderen ist $\alpha_i = 0$. Aus diesem Grund ist die Lösung nur von den Support-Vektoren abhängig.

Fall der nicht-linearen Trennbarkeit

Sind die Trainingssamples jedoch nicht linear trennbar, so existiert für das duale Problem aus Gleichung (5.92) kein lokales Maximum, da die Funktion nach oben nicht beschränkt ist. Man kann dann die VC-Dimension beliebig hoch setzen, aber das verschlechtert die Generalisierungsfähigkeit, wie wir bereits in Abbildung 5.12 gesehen haben. Um das zu verhindern, kann man für den nicht-separablen Fall zusätzliche Support-Vektoren mit berücksichtigen, die sich nicht nur auf dem »harten Rand« sondern auch etwas weiter weg befinden (siehe dazu Abbildung 5.13). Dies wird als Soft-Margin bezeichnet. Mit der Einführung einer neuen Variable ξ_i für jedes Merkmal i mit $i = 1 \dots L$ und Einsetzen in Gleichungen (5.89) und (5.90) erhalten wir

$$\text{Minimiere} \quad J(\mathbf{w}, b) = \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^L \xi_i \right) \quad (5.96)$$

$$\text{Nebenbedingung} \quad [S_i (\langle \mathbf{w}, \mathbf{m}_i \rangle) + b] \geq 1 - \xi_i \quad \forall i. \quad (5.97)$$

Die Slack-Variable ξ_i ermöglicht es jetzt dem Optimierungsalgorithmus, die sonst nicht lösbare Nebenbedingung zu beherrschen. Der Vorteil eines linearen Bestrafungstermes aus Gleichung (5.96) liegt darin, dass im dualen Problem der Term ξ nicht mehr auftaucht. Es bleibt lediglich C als zusätzliche Nebenbedingung. Die Formulierung des dualen Problems, in

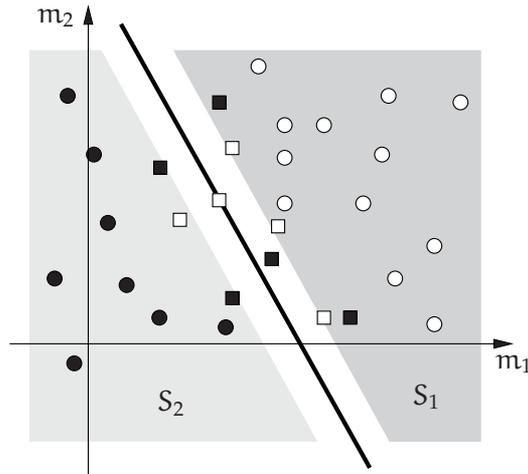


ABBILDUNG 5.13: 2D-Merkmalraum mit einer Trennebene, die Klassentrennung (weiss und schwarz) ist nicht vollständig. Die SV sind als Quadrate gekennzeichnet

Analogie zu Gleichung (5.92) und (5.93), ergibt

$$\text{Maximiere} \quad L'_D = \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i=1}^L \sum_{j=1}^L S_i S_j \alpha_i \alpha_j \mathbf{m}_i \mathbf{m}_j \quad (5.98)$$

$$\text{Nebenbedingung} \quad 0 \leq \alpha_i \leq C \text{ und } \sum_{i=1}^L \alpha_i S_i = 0 \quad (5.99)$$

Wir fassen noch einmal zusammen, wo Trainings- und Testvektoren im Merkmalsraum liegen können. Abbildung 5.14 zeigt dies an 4 Vektoren der Klasse S_1 . Die folgenden Gleichungen zeigen, wie dann die zugehörige Nebenbedingung aus Gleichung (5.97) gestaltet ist:

$$\langle \mathbf{w}, \mathbf{m}_1 \rangle + b > 1, \xi_i = 0, \alpha_i = 0 \Rightarrow \text{kein SV} \quad (5.100)$$

$$\langle \mathbf{w}, \mathbf{m}_2 \rangle + b = 1, \xi_i = 0, \alpha_i > 0 \Rightarrow \text{SV ohne Bestrafung} \quad (5.101)$$

$$\langle \mathbf{w}, \mathbf{m}_3 \rangle + b = 0, \xi_i = 1, \alpha_i = C \Rightarrow \text{SV mit Bestrafung} \quad (5.102)$$

$$\langle \mathbf{w}, \mathbf{m}_4 \rangle + b < 0, \xi_i > 1, \alpha_i = C \Rightarrow \text{SV mit Bestrafung} \quad (5.103)$$

Für jedes nicht richtig klassifizierbare Sample i gilt Gleichung (5.103). Die Summe $\sum_i \xi_i$ gibt eine obere Grenze an Trainingsfehlern an. Diese wird auch als »Upperbound« bezeichnet.

$C \in \mathbb{R}^+$ ist ein frei wählbarer Faktor, der den Trade-Off zwischen Trainingsfehlerminimierung und Marginmaximierung kontrolliert. Er wird natürlich nur für die Vektoren i angewendet, für die $\xi_i > 0$ gilt. Es werden

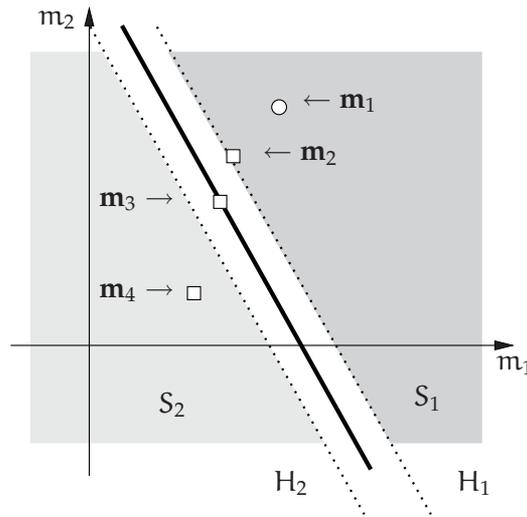


ABBILDUNG 5.14: Beispielvektoren aus Klasse S_1 (weiss), die SV sind als Quadrate gekennzeichnet

damit in der Klassifikation zu den SV der richtigen Seite der Zone auch SV der falschen Seite der Zone berücksichtigt. Für SV der richtigen Seite gilt $\xi_i = 0$ und α_i zwischen 0 und C (kleine Gewichtung der Bestrafung), für SV in der falschen Zone $\xi_i > 0$ und $\alpha_i = C$ (volle Gewichtung der Bestrafung). Für alle übrigen Vektoren ist nach wie vor $\alpha_i = 0$, sie werden nicht berücksichtigt. Je höher der Trade-Off C gewählt wird, desto stärker fließt die Bestrafung ein. Für $\lim C \rightarrow \infty$ bekommt man wieder die »harte« Lösung. Das Training läuft dann so lange, bis alle Vektoren richtig getrennt sind. Mit kleinerem C wird die Bestrafung geringer, also können mehr Vektoren als SV eingesetzt werden. Der wählbare Parameter $\varepsilon \in [0; C[$ entscheidet darüber, ob ein Vektor als SV ausgewählt wird. Gilt für einen Vektor $\alpha_i > \varepsilon$, so wird dieser als SV ausgewählt. Sinnvollerweise wird $\varepsilon > 0$ gewählt, damit eine Reduktion der Anzahl verwendeter Vektoren erzielt wird und damit $\dim[\mathbf{w}^*] < L$ einhergeht. Je größer ε gewählt wird, desto weniger Vektoren werden zu SV. Idealerweise wird sogar $\dim[\mathbf{w}^*] \ll L$ erreicht.

Fall der nicht-linearen Erweiterung

Reicht dieser Ansatz nicht aus, weil die Trainingsamples stark nichtlinear verzahnt sind, so kann unabhängig von einer Soft-Margin eine höhere Funktionsklassen benutzt werden. Damit erhöht man die VC-Dimension

der trennenden Kurve. Bei dieser Erweiterung des Hypothesenraumes

$$\Psi : \mathbb{R}^M \rightarrow \mathcal{H}, \mathbf{m}_i \mapsto z_i \quad (5.104)$$

benutzt man den Kernel-Trick $K(\mathbf{m}_i, \mathbf{m}_j) = \langle \Psi(\mathbf{m}_i), \Psi(\mathbf{m}_j) \rangle$. Die Kernelfunktion $K(\mathbf{m}_i, \mathbf{m}_j)$, so fordert es die Mercer-Bedingung, muss ein Skalarprodukt einer Funktion Ψ abbilden. \mathcal{H} ist der neue Merkmalsraum, der uns erlaubt, nun doch eine lineare Trennfunktion einzufügen. Diese Trennfläche entspricht einer nichtlinearen Trennfunktion im Originalraum \mathbb{R}^M . Die Optimalität der Hyperebene bleibt erhalten. Ein weiterer Effekt der Transformation ist, dass die Margin im neuen Merkmalsraum parallel, nach der Rücktransformation in den Ursprungsraum aber nicht mehr parallel liegt.

Häufig benutzte Kernelfunktionen sind in Gleichung (5.105) bis (5.108) beschrieben.

$$\text{Linear – Kernel} \quad K(\mathbf{m}_i, \mathbf{m}_j) = \langle \mathbf{m}_i, \mathbf{m}_j \rangle \quad (5.105)$$

$$\text{Polynom – Kernel} \quad K(\mathbf{m}_i, \mathbf{m}_j) = (\langle \mathbf{m}_i, \mathbf{m}_j \rangle + 1)^\rho \quad (5.106)$$

$$\text{Gauß – Kernel} \quad K(\mathbf{m}_i, \mathbf{m}_j) = \exp\left(-\frac{\|\mathbf{m}_i - \mathbf{m}_j\|^2}{2\sigma^2}\right) \quad (5.107)$$

$$\text{Sigmoid – Kernel} \quad K(\mathbf{m}_i, \mathbf{m}_j) = \tanh(\kappa \langle \mathbf{m}_i, \mathbf{m}_j \rangle + \theta) \quad (5.108)$$

Linearkombinationen von gültigen Kernel liefern wiederum gültige Kernel. Zum Finden einer guten Kernelfunktion ist eine empirische Vorgehensweise zu wählen. Beim Design ist zu beachten, dass bei Polynomfunktionen ein höherer Grad ρ des Polynoms zu einer höherer VC-Dimension führt. Bei Gauß-Funktionen, auch radiale Basisfunktionen (RBF) genannt, entspricht σ , also die Breite der Gauß-Funktion, dem frei wählbaren Parameter. Ein kleines σ führt zu »Overfitting«, also zu weniger Generalisierung. Zu den Sigmoid-Funktionen, auch als Schwanenhals-Funktionen bezeichnet, zählt zum Beispiel die Tangens Hyperbolicus Funktion. Der S-förmige Verlauf wird durch die Parameter κ und θ beeinflusst. Ergänzend zur Wahl der Kernelfunktion und deren Parametern hat man über den Bestrafungsterm C eine Designmöglichkeit.

Erweiterung mit klassenspezifischer Bestrafung

Die Optimierungsalgorithmen der SVM liefern das optimale Ergebnis, sofern es sich um ein linear trennbares Problem handelt. Ist eine Trennung mit den gegebenen Parametern, insbesondere mit dem angegebenen Kernel, nicht möglich, so bricht der Algorithmus den Optimierungsvorgang bei Erreichen eines Abbruchkriteriums ab. Es ist also ein optimales Ergebnis unter Berücksichtigung der Abbruchbedingungen gefunden.

In beiden Fällen werden die Klassen S_1 und S_2 mit gleicher Priorität behandelt. Das bedeutet, dass keine der beiden Klassen während des Op-

timierungsvorgangs bevorzugt wird. In Abhängigkeit der Fahrzeugfunktion kann es aber wünschenswert sein, die Falschklassifikationsraten der einzelnen Klassen unterschiedlich gewichten zu wollen. Es ist vorstellbar, dass etwa auf das Erkennen von vielen Fußgängern zugunsten einer niedrigeren Falschalarmrate verzichtet werden kann. Um dies mit Hilfe der SVM umsetzen zu können, wurde die SVM um eine klassenspezifische Bestrafung erweitert. Dies ist durch eine klassenspezifische Auswahl der SV realisiert.

Galt bisher

$$\mathbf{m}_i \in SV \Leftrightarrow \alpha_i \geq \varepsilon \quad (5.109)$$

für alle Vektoren \mathbf{m}_i mit $i = 1 \dots L$ beider Klassen S_1 und S_2 , so ist nun durch die klassenspezifische Gewichtung g_ε das neue Auswahlverfahren

$$\mathbf{m}_i \in SV \Leftrightarrow \begin{cases} \alpha_i \geq \varepsilon \cdot g_\varepsilon & \text{falls } f(\mathbf{m}_i, \vec{\alpha}) = S_1 \\ \alpha_i \geq \varepsilon/g_\varepsilon & \text{falls } f(\mathbf{m}_i, \vec{\alpha}) = S_2 \end{cases} \quad (5.110)$$

gegeben. Für unterschiedliche Wahl von g_ε erreicht man unterschiedliche Sensitivitäten und Fehlalarmraten, die als Grundlage für die Erstellung von Receiver Operating Characteristic (ROC)-Kurven dienen (siehe dazu Abschnitt 6.4).

Zusammenfassung SVM

Zusammenfassend lässt sich feststellen, dass SVM gute Klassifikatoren für nichtlinear verzahnte Klassenverteilungen darstellen. Es wird kein a-priori Wissen über die Klassenhäufigkeiten oder über die klassenspezifischen Verteilungsdichten $p(\mathbf{m}_j|S_i)$ benötigt. Das Risiko des Übertrainierens ist nicht gegeben, da nur die SV auf die Bestimmung der Entscheidungsgrenze Einfluss haben. Jedoch gibt es wiederum viele Parameter, die schließlich ein empirisches Training und Testen erfordern. Die Parameter sind hauptsächlich die Art der Kernelfunktion und deren spezifische Parameter sowie die Höhe der Bestrafungsgewichtung C im nichtseparablen Fall.

Ausschlaggebend für die Wahl einer SVM als Klassifikationsverfahren ist die Tatsache, dass eine möglichst starke Klassentrennung sichergestellt werden sollte. Die benötigte Rechenzeit für das Training und für die Klassifikation sind dabei weniger wichtig als eine Antwort auf die Frage zu bekommen, ob die aus den zur Verfügung stehenden Sensordaten berechneten Merkmale klassenspezifische Eigenschaften aufweisen. Dafür bietet die SVM das größte Potenzial.

6

VERSUCHSDURCHFÜHRUNG

Nachdem in Kapitel 3 das Fußgängerschutzsystem, in Kapitel 4 die Sensorik analysiert und ausgewählt und in Kapitel 5 die Bildverarbeitung und die Klassifikation ausführlich beschrieben wurden, befasst sich Kapitel 6 mit der technischen Umsetzung. Diese gliedert sich wie folgt. Abschnitt 6.1 erläutert den Aufbau des Versuchsfahrzeugs inklusive der genauen Parametrierung der verwendeten Sensorik. Abschnitt 6.2 liefert eine Übersicht über die Softwaremodule des Erkennungssystems mit der Bezeichnung Pedestrian Recognition System Using a Multi-Sensor Environment (PRE-SUME). Abschnitt 6.3 erklärt der Vorgang des Trainings, der sowohl mit einer Brute-Force-Methodik als auch mit einem Evolutionären Algorithmus ausgeführt werden kann. Die zur Bewertung der Klassifikation verwendete Fitnessfunktion wird in Abschnitt 6.4 definiert und begründet. Zusätzlich sind alle notwendigen Klassifikationsraten aufgelistet. Abschließend beschreibt Abschnitt 6.5 den Umfang der verwendeten Trainings-, Test- und Validierungsdaten.

Es sei angemerkt, dass dieser praktische Teil der Arbeit viele unterschiedliche und umfangreiche Arbeitspakete enthält. Das Versuchsfahrzeug wird durch den Umbau zu einem Einzelstück von hohem Wert, welches mit besonderer Vorsicht behandelt werden muss. Der Aufwand von der Lieferung des Fahrzeugs bis zum Abschluss der Umbauarbeiten umfasst dabei mehrere Monate. Ebenso beinhalten die Arbeiten an der Sensorik nicht nur die im vorherigen Kapitel beschriebene Algorithmenentwicklung, sondern auch die Parametrierung und die Kalibrierung. Da es sich bei den PMD-Systemen jeweils um Prototypen handelt, bedarf diese Arbeit viel Abstimm- und Versuchsaufwand. Steht das Fahrzeug nach Abschluss der Umbauarbeiten bereit, so müssen umfangreiche und vor allem relevante Verkehrsszenen definiert und aufgenommen werden.

6.1 VERSUCHSAUFBAU

Der vorliegende Abschnitt beinhaltet die Beschreibung des Versuchsfahrzeugs, der verwendeten Sensorik, deren Wirkungsbereich und das zu Grunde liegende Koordinatensystem.



ABBILDUNG 6.1: Audi A8 Versuchsfahrzeug

6.1.1 *Eingesetztes Versuchsfahrzeug*

Für die Aufzeichnung der Messdaten und für die Evaluierung der Erkennungsleistung wurde ein Versuchsfahrzeug aufgebaut, das mit einem PMD-System, einer Monovideo-Kamera, einem PC und der erforderlichen Elektronik zur Anbindung der Sensorik und des Fahrzeugbussystems ausgestattet wurde. Abbildung 6.1 zeigt ein baugleiches Fahrzeug des Versuchsfahrzeugs. Die Sensordaten wurden mit Hilfe des PC visualisiert, aufgezeichnet und ausgewertet. Abbildung 6.2 zeigt die Integration der PMD-Kamera und der MV-Kamera.

Im Lauf der Untersuchungen sind noch weitere Versuchsfahrzeuge eingesetzt worden, die teilweise mit identischen, teilweise mit unterschiedlichen Sensorgenerationen ausgerüstet waren und zur zusätzlichen Ergebnisvalidierung eingesetzt wurden.

6.1.2 *Verwendetes PMD-System*

Die PMD-spezifischen Sensoreigenschaften sind in Tabelle 6.1 aufgelistet. Das System hat eine maximale Bildwiederholrate von 100 Hz. Es existieren jedoch verschiedene Verfahren, bei denen mehrere Belichtungs- und Verarbeitungsschritte benötigt werden, um ein Bild zu liefern. Dadurch wird die Wiederholrate verringert. In der hier verwendeten Konfiguration erreicht das PMD-System eine Bildwiederholrate von 60 Hz. Die Integration der IR-Strahlungsquellen erfolgte wie in Abbildung 6.3 gezeigt.

Häufig kommt der Wunsch eines größeren Erfassungsbereiches auf, sowohl im Sinne des Öffnungswinkels als auch bezogen auf die Reichweite. Ebenso ist die Auflösung des Bildaufnehmers Gegenstand vieler Diskussionen, da die Bildverarbeitung häufig eine hohe Auflösung fordert. Eine

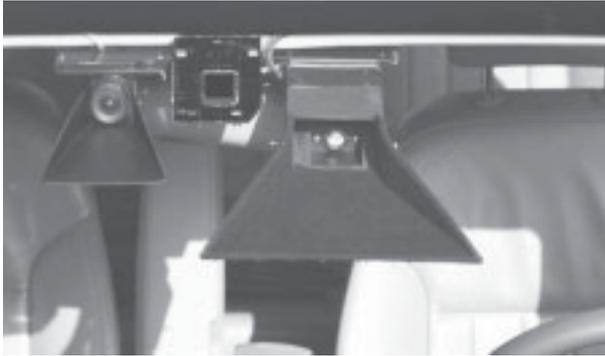


ABBILDUNG 6.2: Versuchsaufbau mit hinter der Windschutzscheibe montiertem PMD-Sensor (im Bild rechts) und MV-Sensor (im Bild links) inklusive je einer Blende zum Schutz gegen Reflexionen

Erhöhung der Auflösung fordert zwar mehr Rechenkapazität bei der Bildverarbeitung, liefert jedoch im Allgemeinen auch bessere Ergebnisse. Um eine Antwort auf diese Punkte zu liefern, detailliert Anhang D sowohl den Trade-off zwischen Öffnungswinkel und Strahldichte als auch den Trade-off zwischen Auflösung und Reichweite.

Um bei einem gegebenen vertikalen Öffnungswinkel eine durchschnittlich große Person in einem möglichst großen Entfernungsbereich vor dem Fahrzeug vollständig zu erfassen, wurde eine Optimierung des Nickwinkels durchgeführt. Die Randbedingung der Optimierungsberechnung ist ein möglichst kleiner Nahbereich, das heißt die vollständige Objekthöhe muss in einem möglichst kleinen Abstand erfassbar sein. Das Ergebnis ist eine Neigung des Sensors um 3° bis 6° nach unten zur idealen Erfassung eines unmittelbar vor dem vorderen Stoßfänger stehenden Mannes mit einer Körpergröße von 1,93 m. Für diese Körpergröße gilt, dass nur 5 % aller männlichen Personen größer sind. Er wird daher als 95 %-Mann bezeichnet.

6.1.3 *Verwendetes Monovideo-System*

Die MV-spezifischen Sensoreigenschaften sind in Tabelle 6.2 aufgelistet.

6.1.4 *Koordinatensystem und Kalibrierung*

Für die Darstellung der Umfeldobjekte gilt ein AutoSAR-ähnliches Koordinatensystem entsprechend Abbildung 6.4. Der Ursprung des kartesischen Koordinatensystems liegt in einem Punkt, der von der Mitte des vorde-

PMD-Systeminformationen	
Hersteller	i.f.m. Electronic GmbH und PMD Technologies GmbH
Sensorbezeichnung	PhotonICs(R) PMD 1k-S
Schnittstelle	Ethernet
Reichweite (Fahrzeuge)	> 45 m
Reichweite (Fußgänger)	> 15 m
Nickwinkel	6,5°
Auflösung	64 × 16
Abmessungen (L × H × B) (horiz./vert.)	103 × 52 × 74 mm
Wiederholfrequenz (pro Bild)	60,0 Hz
Modulationsfrequenz 1	$f_{\text{mod},1} \approx 7,481 \text{ MHz}$
Modulationsfrequenz 2	$f_{\text{mod},2} \approx 6,234 \text{ MHz}$
Spannungsversorgung	12 V
Intrinsische Parameter	
Brennweite	9,85 mm
Pixelgröße	155,31 × 210,63 μm
Öffnungswinkel ($\theta_h \times \theta_v$)	53,55° × 19,41°

TABELLE 6.1: Parameter des verwendeten PMD-Systems



ABBILDUNG 6.3: IR-Strahlungsquellen im Kühlergrill

ren Stoßfängers des Fahrzeugs auf die Fahrbahnebene projiziert ist. Dieser ist in den Abbildungen dargestellt. Bei dem unten beschriebenen Koordinatensystem ist die x -Achse vom Fahrzeugaufbau waagrecht nach vorne gerichtet, die y -Achse zeigt parallel zur Hinterachse nach links bei Sicht in Vorwärtsrichtung des Fahrzeugs. Die z -Achse steht orthogonal auf den

MV-Systeminformationen	
Sensorhersteller	Sony
Sensorname	ICX098AL/BL
Schnittstelle	USB 1.1 / USB 2.0
Wiederholfrequenz	30,0 Hz genutzt
Spannungsversorgung	5 V über USB
Leistungsaufnahme	< 2,5 W
Abmessungen (L × H × B) (ohne Objektiv)	58,8 × 38,8 × 38,8 mm
Gewicht (ohne Objektiv)	ca. 120 g
Bildaufnehmer / intrinsische Parameter	
Technologie	CCD
Auflösung	640 × 480
Art des Auslesens	Progressiv Scan
Transfertyp (horiz./vert.)	Full-Frame-Shutter
zulässige Umgebungstemperatur	0° C bis 45° C
Autom. Belichtungsregelung (AEC)	ja
Autom. Verstärkungsregelung (AGC)	ja
ADC Auflösung	10 Bit
Übertragung	8 Bit
Objektivgewinde	C-Mount
Brennweite	3,38 mm
Pixelgröße	5,6 × 5,6 μm
Öffnungswinkel	55,86° × 43,37°

TABELLE 6.2: Parameter des verwendeten Monovideo-Systems

beiden anderen Achsen und ist nach oben gerichtet. Der Bezugspunkt auf dem Objekt liegt im Objektzentrum. Alle Bewegungsmerkmale sind Relativgrößen, die auf die Fahrzeugbewegung bezogene sind. Länge, Breite und Höhe eines Objektes beziehen sich auf eine Bounding-Box, die das Objekt umschließt. Die Länge ist immer in x -Richtung, die Breite immer in y -Richtung definiert. Im Gegensatz zu Gauß-Krüger-Koordinaten, bei denen die Lage eines Objektes in 2D und einem zusätzlichen Höhenwert vorliegen, stehen in dem hier benutzten Koordinatensystem alle 3 Raumkoordinaten für die Lage inklusive Höhe und Breite zur Verfügung. Die mathematische Beschreibung der Transformation der Koordinatensysteme ist in [62] beschrieben.

Alle Sensoren müssen in der Art und Weise kalibriert werden, dass ei-

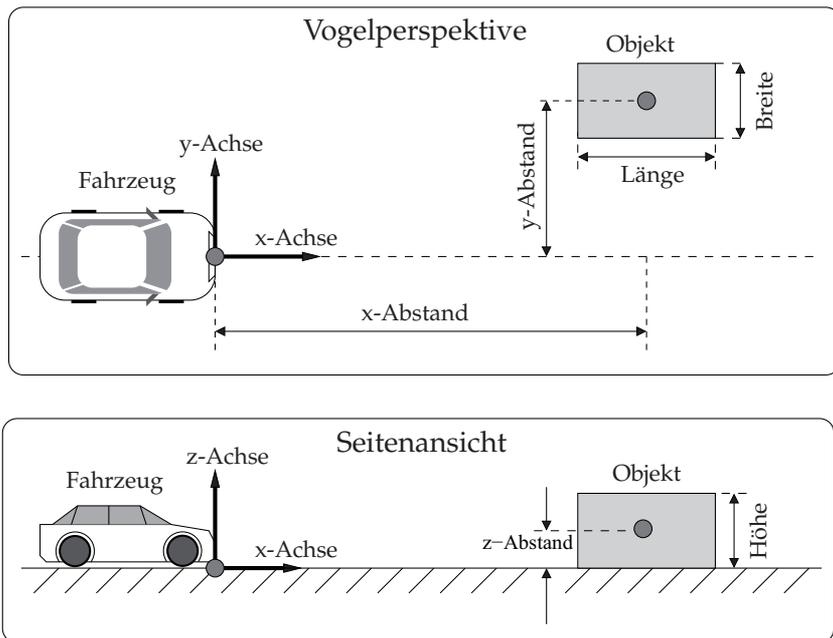


ABBILDUNG 6.4: Verwendetes Fahrzeugkoordinatensystem

ne präzise Transformation der Daten in die jeweiligen Sensorkoordinaten möglich ist. Die Kalibrierung muss regelmäßig überprüft werden. Hierbei ist eine erhöhte Sorgfaltspflicht angebracht, um eine hohe Qualität der Sensordaten zu gewährleisten. Die Kalibrierung wird nach der Methode nach Zhang [63] durchgeführt und ist in eine intrinsische und eine extrinsische Kalibrierung aufgeteilt. Die intrinsische Kalibrierung ermittelt die Parameter, die für die Projektion der Punkte im Kamerakoordinatensystem auf die Bildpunkte in Bildkoordinaten zuständig sind. Diese Parameter sind die Brennweite, der Kamerahauptpunkt, der Scherungsparameter, die radiale Verzerrung und die tangentielle Verzerrung. Die extrinsische Kalibrierung ermittelt die Lage der Kamera, also damit die Lage des Kamerakoordinatensystems, in Bezug auf das in Abbildung 6.4 gezeigte Referenzkoordinatensystem. Die Lage wird über die Parameter der Position und der Rotation beschrieben. Weitere Informationen zur Kalibrierung und zu Kalibrierverfahren sind in [64] und [65] beschrieben.

6.1.5 Wirkungsbereich der Sensorik

Durch die Wahl der Sensorik mit ihrer Optik, der Einbaulage und den Einstellungen der Erkennungssoftware ergibt sich ein Bereich vor dem Fahrzeug, in dem Objekte erkannt werden können. Abbildung 6.5 zeigt diesen

Wirkungsbereich.

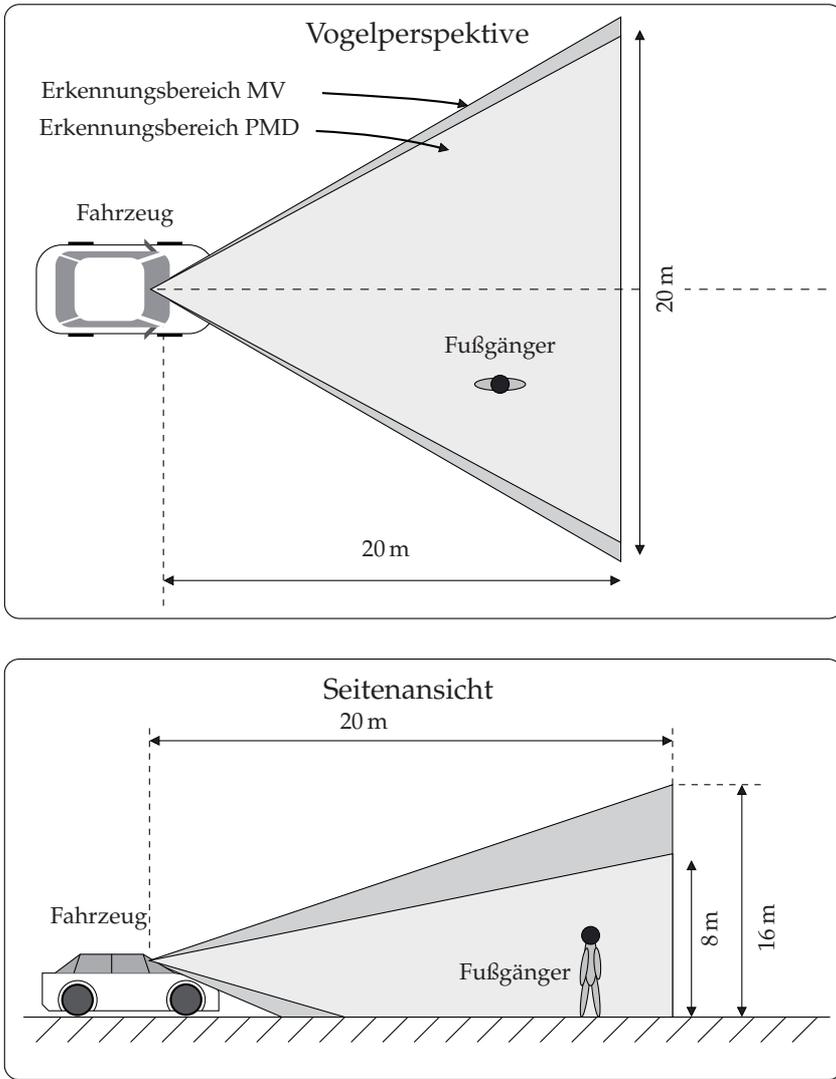


ABBILDUNG 6.5: Der Erkennungsbereich ist Teil des Sichtfeldes, in ihm werden Objekte detektiert und klassifiziert

6.2 SOFTWAREKOMPONENTEN VON PRESUME

Die gesamte Übersicht der Softwarekomponenten ist in Abbildung 6.6 illustriert. Sie beschreibt den Aufbau des mit PRESUME bezeichneten Sy-

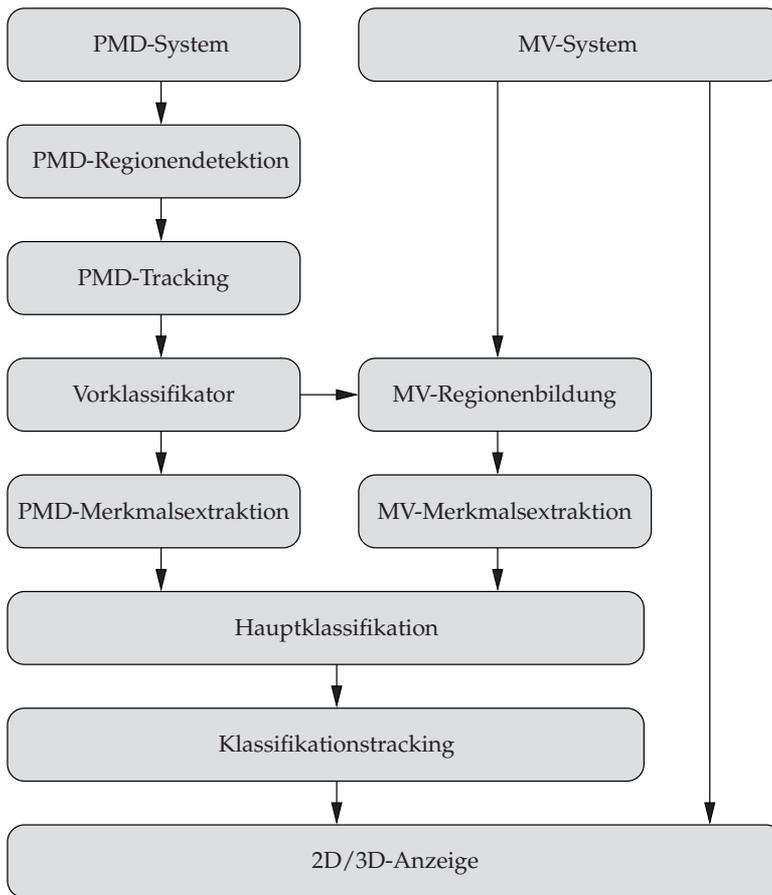


ABBILDUNG 6.6: Schematische Darstellung von PreSUME zur Klassifikation von Fußgänger auf Basis von 2D- und 3D-Daten und zur Visualisierung der Ergebnisse

stems. Alle Softwarekomponenten, auch Filter genannt, sind als Teil des Automotive Data and Time Triggered Framework (ADTF) entwickelt worden. Das ADTF ist ein in der Audi Electronics Venture GmbH (AEV) entwickeltes Softwareframework zur Unterstützung der Funktionsentwicklung, vor allem im Bereich der Fahrerassistenzsysteme. Es handelt sich dabei um eine PC-basierte C++ - Plattform, die modular aufgebaut ist und die Aufnahme und Wiedergabe von Sensordaten in Echtzeit ermöglicht. In dieser Arbeit wurde die Version 1.1.3 Build 15 des Frameworks verwendet.

Der erste Schritt der Verarbeitungskette besteht aus dem Gruppieren zusammenhängender Regionen auf Basis der PMD-Bilddaten. Das Resultat dieser PMD-REGIONENDETEKTION sind diskrete Objekte. Der nachfol-

gende Kalman-Filter im Filter PMD-TRACKING verfolgt diese Pixelgruppen über der Zeit und kann somit die Objektfindung optimieren und eine Bewegungsschätzung der Objekte liefern. Diese ersten beiden Verarbeitungsschritte wurden nicht im Rahmen der vorliegenden Arbeit erstellt. Ihr Einfluss auf die nachfolgenden Module ist allerdings extrem hoch, daher ist ein Grundverständnis notwendig.

Alle nachfolgenden Filter bekommen als Eingangsgröße die Objektinformationen und bearbeiten diese. Zuerst leistet der VORKLASSIFIKATOR eine Klassifikation mit Hilfe von Schwellenentscheidungen. Dadurch kann die Anzahl zu betrachtender Objekte verringert werden. Anschließend wird eine Transformation der Objektpositionen in die 2D-Bildebene durchgeführt. Dort werden in der MV-REGIONENBILDUNG eigene Regionen aufgespannt, aus denen die MV-Merkmale extrahiert werden. Ebenso wird die PMD-MERKMALSEXTRAKTION angestoßen. Die so gesammelten Merkmale aus der 3D- und 2D-Bildverarbeitung fließen in die HAUPTKLASSIFIKATION ein. Anschließend wird ein Tracking des Klassifikationsergebnisses im Filter KLASSIFIKATIONSTRACKING durchgeführt.

Zur zusätzlichen Unterstützung im Entwicklungsprozess kann das Videobild, das Klassifikationsergebnis und die Leistungsfähigkeit der Klassifikation durch eine 2D/3D-ANZEIGE visualisiert werden.

Der Softwareumfang beläuft sich auf über 54 000 Zeilen ausführbaren C++-Code. Dies zeigt, dass zur Lösung der Aufgabe eine sehr strukturierte Softwarearchitektur notwendig ist. Es fällt dabei auf, dass der Hauptteil der Software auf die Erstellung der Klassifikatoren entfällt. Diese Umfänge beinhalten die Datenaufnahme, das Labeln, Schneiden, Trainieren, Validieren, Visualisieren und Optimieren. Existiert nach erfolgreichem Training ein fertiger Klassifikator, so müssen die Sensordatenanbindung, die Merkmalsextraktion und die Klassifikationskomponenten verwendet werden. Es muss betont werden, dass die Software streng modular aufgebaut ist. Zwar erhöht diese Modularität den initialen Entwicklungsaufwand, jedoch erreicht man nur dadurch eine nachhaltige Erweiter- und Wartbarkeit, die für den Einsatz in zukünftigen Projekten mit neuen Mitarbeitern ermöglicht. Die Software wird aktuell in der Elektronik-Vorentwicklung eingesetzt. Sie wird um zusätzliche Komponenten erweitert und auch in den nächsten Jahren für die Entwicklung diverser Fahrerassistenzsysteme verwendet.

6.2.1 Vorklassifikator

Der Filter für die Vorklassifikation entspricht algorithmisch dem Klassifikationsprinzip aus Abschnitt 5.5.2. Er weist jedem Objekt j die Klasse Nicht-Fußgänger mit einer Klassifikationsgüte $G_j \in [0; 1]$ zu. $G_j = 0$ wird vergeben, wenn für jedes einzelne Merkmal $m_i \in \mathbf{m}$ des Merkmalsvektors

genau $m_i = m_{i,\text{mid}} = 1/2(m_{i,\text{max}} + m_{i,\text{min}})$ gilt. $G_j = 1$ wird vergeben, falls die Bedingung $m_i \leq m_{i,\text{min}}$ oder $m_i \geq m_{i,\text{max}}$ für mindestens ein Merkmal erfüllt ist. Für alle dazwischen liegenden Werte wird eine quadratische Funktion angewendet und der Mittelwert der Klassifikationsgüte bestimmt (siehe Abbildung 6.7). Der Vorteil der quadratischen Funktion ist, dass ein großer Bereich um $m_{i,\text{mid}}$ mit einer kleinen Klassifikationsgüte belegt wird und somit nur die Randwerte, also Werte nahe dem definierten Minimum oder Maximum, die Wahrscheinlichkeit für die Wahl der Klasse Nicht-Fußgänger erhöhen. Ebenso ist eine Sprungfunktion denkbar, die jedoch jeglichen Boosting-Effekt verhindert, der in Abschnitt 5.5.2 erwähnt wurde.

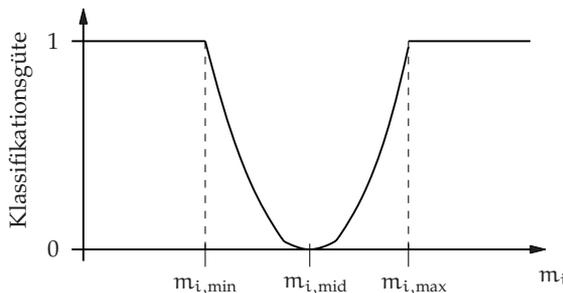


ABBILDUNG 6.7: Zuweisung einer Objektklasse und einer Güte zu jedem Merkmal $m_i \in \mathbf{m}$ im Vorklassifikator

Weiterhin wird $G_j = 1$ gesetzt, wenn mindestens 1 Pixel des Objektes in der ersten oder letzten Spalte des PMD-Bildes liegt. Das Objekt wird nämlich in diesem Fall als Randobjekt betrachtet, für welches die notwendigen Konturanalysen speziell untersucht werden müssten. Außerdem umgeht man damit die Schwierigkeit der nicht bekannten Objektbreite und Objektkontur. Tabelle 6.3 listet die vollständigen Einstellungen des Vorklassifikators auf.

6.2.2 Hauptklassifikation

In der HAUPTKLASSIFIKATION kommt eine SVM zum Einsatz. Die dadurch frei wählbaren Parameter sind die Kernelfunktion K mit dem spezifischen Parameter ρ (Polynom-Kernel), σ (Gauß-Kernel) beziehungsweise κ und θ (sogmoid-Kernel), die Größe des Trade-Off Parameters C , der für die Auswahl der eines Vektors als Support Vektor relevante Parameter ε und die zugehörige klassenspezifische Gewichtung g_ε .

Merkmal	Minimum	Maximum
Höhe	0,4 m	2,0 m
Breite	0,2 m	2,0 m
Tiefe	0,0 m	2,5 m
Abstand	-1,90 m	18,1 m
Höhe des Fußpunktes	-0,80 m	1,5 m
Höhe des Kopfpunktes	0,75 m	2,5 m

TABELLE 6.3: Einstellungen des Vorklassifikator-Filters

6.2.3 Klassifikationstracking

Das PMD-TRACKING stellt den nachfolgenden Filtern die Information bereit, ob ein Objekt auf Basis des Assoziationsschrittes schon existierte oder neu initialisiert wurde. Bei einem bereits existierenden Objekt kann das aktuelle Klassifikationsergebnis $S_{i,1}$ mit dem vorherigen Klassifikationsergebnis $S_{i,2}$ verglichen werden. Im KLASSIFIKATIONSTRACKING findet nun eine Historienbetrachtung über die letzten H Klassifikationsergebnisse statt, der endgültige Ausgabewert der Objektklasse \hat{S}_i ergibt sich aus

$$\hat{S}_i = \frac{\sum_{h=1}^H (H - h + 1) S_{i,h}}{\sum_{h=1}^H (h)}. \tag{6.1}$$

6.2.4 Trennung zwischen Nah- und Fernbereich

In den Abschnitten 5.2.2 und 6.1.2 wurde über die Besonderheit eines Nahbereiches gesprochen. Dieser zeichnet sich dadurch aus, dass ein durchschnittlich großer Fußgänger nicht vollständig erfasst werden kann. Entscheidend für den vertikalen Erfassungsbereich des Sensors ist dessen Öffnungs- und Nickwinkel. Im Nahbereich werden Kopf und Beine nicht vollständig erfasst. Dies hat unmittelbar Auswirkungen auf das Höhenmerkmal HEIGHT, aber auch Konsequenzen für alle Konturmerkmale. Als Grenze zwischen Nah- und Fernbereich kann eine Ebene definiert werden, die in einem bestimmten Abstand x parallel zur y - z -Ebene liegt.

6.3 TRAINING UND TEST DER KLASSIFIKATOREN

Im Folgenden wird die vollständige Evaluierung, das heißt das Trainieren und Testen, beschrieben. Abbildung 6.11 zeigt das Flussdiagramm, das dafür verwendet wird. Der Prozess orientiert sich am Verfahren der k -fachen Kreuzvalidierung, ein Durchgang unter Verwendung eines Messdatensatzes läuft dabei nach folgendem Prinzip ab.

1. **Einlesen:** Einlesen des vollständigen Datensatzes und Berechnung der zu jedem detektierten Objekt gehörenden Merkmale
2. **Datensatz teilen:** Aufspaltung des Datensatzes in einen Trainings- und einen Testdatensatz (wie in Abschnitt 6.3.1 beschrieben). Die Aufteilung wird bei jedem Durchlauf verändert, so dass jeder Datensatzteil pro Klassifikationsschritt genau einmal auch als Testdatensatz dient. Test- und Trainingsdaten sind jedoch zu keinem Zeitpunkt identisch.
3. **Training:** Erstellen eines Klassifikators mit dem Ziel der optimalen Trennung des Trainingsdatensatzes. Die Parameter für den Klassifikator werden aus einer Parameterliste entnommen.
4. **Güte-Auswertung:** Testen des erstellten Klassifikators auf dem Testdatensatz und speichern der Ergebnisse.
5. **Permutation:** Permutation der Parameter des Klassifikators entsprechend der ausgewählten Parameterpermutation. Für diese Parameterpermutation kommen die Verfahren aus Abschnitt 6.3.2 und 6.3.3 zum Einsatz. Ein Rücksprung zu Punkt 2 findet so lange statt, bis alle Parameter getestet wurden.

Das Endergebnis dieser Durchgänge ist eine Liste mit den verwendeten Klassifikatoren und der jeweiligen Anzahl von Richtig- und Falschklassifikationen aller detektierten Objekte.

6.3.1 Aufteilung des Datensatzes

Da in jedem Durchlauf sowohl ein Trainings- als auch ein Testdatensatz benötigt wird, teilen wir den vorhandenen Datensatz nach dem Prinzip der k -fachen Kreuzvalidierung auf. Ein kleiner Prozentsatz, einstellbar zwischen $p_B = 1 \dots 25\%$, wird vom Training ausgenommen und zum Testen verwendet. Das Verschieben des Testdatensatzes, wie in Abbildung 6.8 gezeigt, garantiert, dass er repräsentativ ist. Eine Permutation der Parameter der Hauptklassifikation findet erst nach Abschluss von k Verschiebungen statt. Es ergeben sich $k = \lfloor 1/p_B \rfloor$ Blockpermutationen, also k Verschiebungen des Testdatenbereiches auf dem gesamten Datensatz. In dieser Arbeit wurde $k = 5$ ausgewählt, um eine annähernd gleiche Verteilung der Testdaten zu garantieren. Diese Vorgehensweise wird als k -fache stratifizierte Kreuzvalidierung bezeichnet.

6.3.2 Brute-Force Permutation der Parameter

Alle Parameter der Hauptklassifikation werden in jedem Durchlauf permutiert. Dazu wird ein Brute-Force-Ansatz benutzt, der jede mögliche Kombination aller Parameter vornimmt. Die Anzahl notwendiger Iterationen

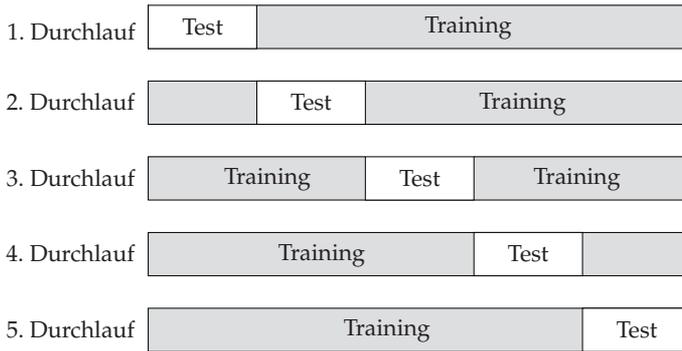


ABBILDUNG 6.8: Methode der k-fachen stratifizierten Kreuzvalidierung, hier im Beispiel mit $k = 5$ auf einem vollständigen Datensatz

I_{all} entspricht dem Produkt aus der Anzahl der in Abschnitt 6.2.2 erläuterten Parameter. Dieser Ansatz ist sinnvoll, wenn nur wenige Permutationen benötigt werden. Dies ist zum Beispiel der Fall, wenn a-priori genug Wissen über die Parameter vorhanden ist, oder wenn der Parameterraum sehr klein ist.

6.3.3 Evolutionärer Algorithmus

Falls eine ungefähre a-priori Einschätzung der Parametersätze nicht gegeben ist, zum Beispiel weil das Wissen über den vorhandenen Trainingsdaten zu gering ist, ist der Brute-Force Ansatz jedoch nicht sinnvoll. In diesen Fällen können evolutionäre Algorithmen (EA) zur Anwendung kommen [56]. Bei Evolutionären Algorithmen, deren Vorbild die biologische Evolution ist, werden Individuen, beschrieben durch Eigenschaften, durch zufällige und deterministische Veränderungen weiterentwickelt. Im Laufe der Entwicklung nähern sich die Individuen dem Optimum an. Wie stark sich ein Individuum weiterentwickelt und dabei dem Optimum genähert hat, kann mit Hilfe einer zu definierenden Fitnessfunktion berechnet werden. Die Berechnung erfolgt auf Basis der das Individuum beschreibenden Zahlenwerte [66].

Im Rahmen dieser Arbeit wurde ein genetischer Algorithmus implementiert, der an den Individuen J_p der Population \mathcal{P}

$$J_p \in \mathcal{P} \text{ mit } p = 1 \dots P \tag{6.2}$$

Mutationen, Rekombinationen und Selektionen vornimmt. Verändert werden in diesem Fall die Parameter der Klassifikation, da diese die Eigenschaften von J_p beschreiben. Die obligatorische Mutation stellt dabei eine zufällige Komponente dar, die den ungerichteten Prozess der Erzeu-

gung von Varianten vorantreibt. Die optionale Rekombination (Kreuzung) kann die Effektivität des Algorithmus steigern. Bei ihr werden Parameter von zwei Individuen miteinander vertauscht. Die Selektion stellt schließlich eine streng zielgerichtete Komponente dar, die das Überleben beziehungsweise das Sterben der Individuen regelt. Diese Regelung kann rein deterministisch sein, oder auch indeterministisch falls eine gewisse Zufälligkeit den eigentlich deterministischen Selektionsprozess stört. Zusätzlich zu der aus P Individuen bestehenden Startpopulation \mathcal{P}_0 ist eine Zielfunktion $f(J_p)$ definiert, die beim Übergang von einer Generation mit der Population \mathcal{P}_i in die nächste Generation \mathcal{P}_{i+1} die Fitness $Y_{\mathcal{P}_i}$ aller Individuen bestimmt. Falls

$$Y_{\mathcal{P}_i} \geq \hat{Y} \quad (6.3)$$

erfüllt ist, bricht der Algorithmus erfolgreich ab. Es ist dann ein Individuum, beschrieben durch einen Parametersatz, gefunden, das eine zuvor definierte Güte der Klassifikation, also einen Mindestfitnesswert \hat{Y} , liefert. Die Anwendung des genetischen Algorithmus ist in Anhang C an einem Beispiel gezeigt. Dabei wurde über mehrere Generationen hinweg die Mutation, Rekombination und Kreuzung durchgeführt.

Durch den Einsatz von EA können mehrere potentielle Lösungen, sofern sie existieren, gefunden werden. Sie können auch dann angewandt werden, wenn das Wissen über die Eigenschaften der Individuen und deren Einfluss auf die Fitness sehr gering ist. Durch die Zufallskomponenten und die Parallelität können lokale Optima überwunden und das globale Optimum erreicht werden. Jedoch bieten die EA keine Garantie, dass das globale Optimum in vernünftiger Zeit gefunden wird. Falls also für ein gegebenes Optimierungsproblem ein spezielles Verfahren existiert bekannt ist, welches sich nicht an EA orientiert, so sollte diesem Verfahren der Vorrang gegeben werden. Dennoch bieten Evolutionäre Algorithmen die enorme Leistungsfähigkeit des natürlichen Evolutionsprozesses bei gleichzeitiger Einfachheit der Teilprozesse [67].

6.4 GÜTEPARAMETER UND BEWERTUNG DER KLASSIFIKATION

Generell lässt sich sagen, dass bei Klassifikationen ein Trade-off zwischen Richtig-Positiv-Rate und Falsch-Positiv-Rate gegeben ist. Ebenso hat auch eine Ungleichverteilung der Klassenhäufigkeiten eine signifikante Auswirkung auf die gemessene Leistungsfähigkeit eines Klassifikationssystems. Tabelle 6.4 erklärt die für die Bestimmung der Güteparameter und die Bewertung relevanten Begriffe. Um eine Aussage über die Genauigkeit des

Parameter	Beschreibung
a	richtig positiv: Einstufung eines Fußgänger (FG) in die Klasse FG
b	falsch positiv: Einstufung eines Nicht-Fußgänger (nFG) in die Klasse FG
c	falsch negativ: Einstufung eines Fußgängers in die Klasse nFG
d	richtig negativ: Einstufung eines nFG in die Klasse nFG
$\frac{a}{a+c}$	Sensitivität r_p (auch Signifikanz, Richtig-Positiv-Rate oder Recall): WSK, dass ein FG als solcher erkannt wird. Beschreibt die Empfindlichkeit des Erkennungssystems
$\frac{d}{b+d}$	Spezifität r_n , Richtig-Negativ-Rate: WSK, dass ein nFG als nFG erkannt wird. Also eine WSK dafür, dass es keinen Fehlalarm gibt
$\frac{b}{b+d}$	Falsch-Positiv-Rate f_p : WSK einer zur Fehlauflösung führenden Falsch- klassifikation
$\frac{a}{a+b}$	Relevanz (auch Precision): WSK, dass das Objekt wirklich ein FG ist, dass also ein vom Klassifikator ausgewählter Fußgänger tatsächlich ein Fußgänger ist
$\frac{d}{c+d}$	Segreganz: WSK, dass ein Objekt wirklich kein FG ist
$\frac{c}{a+c}$	Falsch-Negativ-Rate f_n (auch Missed-Positive-Rate): WSK, dass der Klassifikator einen Fußgänger übersieht
$\frac{c}{c+d}$	Fall-Out: WSK, dass ein vom Klassifikator ausgewählter Nicht-Fußgänger tatsächlich ein Nicht-Fußgänger ist

TABELLE 6.4: Parameter für die Bewertung der Klassifikationsergebnisse

Systems machen zu können, liefert die Korrektklassifikationsrate

$$r_k = \frac{a + d}{a + b + c + d} \quad (6.4)$$

die Wahrscheinlichkeit einer richtigen Systemaussage. Sie ist ein Maß für die Anzahl korrekter Vorhersagen. Umgekehrt dazu liefert die Falschklassifikationsrate

$$r_f = \frac{b + c}{a + b + c + d} \quad (6.5)$$

die Wahrscheinlichkeit einer falschen Systemaussage. Es ist $r_f = 1 - r_k$. Wir benötigen nun eine optimierbare Zielfunktion. Der Nachteil von r_k als Zielfunktion liegt in der Tatsache, dass bei nicht identischer Klassenhäufigkeit $P(S_1) \neq P(S_2)$ diejenige Klasse mit der geringen Anzahl Objekte weniger gewichtet wird. Da wir aber kein a-priori-Wissen über $P(S_i)$ machen können (und auch nicht wollen), kann die Sensitivität und die Spezifität getrennt berechnet und anschließend zusammen betrachtet werden¹. Dies lässt sich mit dem Youden-Index

$$Y = \frac{a}{a+c} + \frac{d}{b+d} - 1 \quad (6.6)$$

$$\Leftrightarrow Y = \frac{ad - bc}{(a+c)(b+d)} \quad (6.7)$$

erreichen [68]. Zwar werden hier die beiden Klassen mit $P(S_1) = P(S_2)$ behandelt, doch genau das ermöglicht eine faire Betrachtung der Gesamtleistung. Zur Erklärung nehmen wir an, die Objekte $(a+c)$ bilden die Klasse 1, $(b+d)$ die Klasse 2. Die Klassenhäufigkeiten $P(S_1) = k \cdot P(S_2)$ mit $k \in \mathbb{R}$ seien ungleich. In (6.6) wird nun diese Ungleichheit durch die getrennten Brüche wieder ausgeglichen. Die Ergebnisse der Klassentrennung fließen unabhängig von der Objektanzahl mit identischem Gewicht in Y ein. Aus diesem Grund wurde Y als optimierbare Zielfunktion ausgewählt. Es gilt $Y \in [-1; 1]$, ein höheres Y bedeutet eine bessere Klassentrennung.

Für die graphische Visualisierung kann eine sogenannte ROC Kurve herangezogen werden. Diese beschreibt die Abhängigkeit der Sensitivität r_p über dem Wert $1 - r_n$ mit

$$1 - r_n = f_p, \quad (6.8)$$

also die Abhängigkeit zwischen der Richtig-Positiv-Rate r_n und der Falsch-Positiv-Rate f_p . Es ergibt sich ein rechtgekrümmter Kurvenverlauf mit den Schnittpunkten $(0/0)$ und $(1/1)$. Der maximale Youden-Index bei Gleichgewichtung von r_p und f_p entspricht daher dem Punkt der ROC-Kurve, der den größten Abstand zur direkten Verbindung zwischen den beiden Punkten aufweist, siehe dazu Abbildung 6.9. Dieser Punkt wird als optimaler Cutpoint bezeichnet.

¹Anwendung auf Diagnoseevaluation, <http://www.imbe.med.uni-erlangen.de/lehre/Querschnittsbereich1/Unterlagen/diagnose.pdf>, Uni Erlangen, zuletzt besucht im Mai 2006

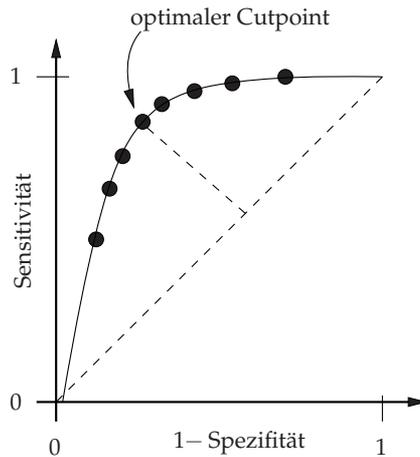


ABBILDUNG 6.9: Schematische Darstellung einer ROC-Kurve mit optimalem Cutpoint

In dieser Arbeit wurde eine Zielfunktion

$$Y_k = \frac{1}{N_B} \sum_{i=1}^{N_B} Y_{k,i} \quad (6.9)$$

$$\Leftrightarrow Y_k = \frac{1}{N_B} \sum_{i=1}^{N_B} \left(\frac{a_i}{a_i + c_i} + \frac{d_i}{b_i + d_i} - 1 \right) \quad (6.10)$$

implementiert. Diese wird bei jeder Parameterpermutation, also nach einer vollständigen Blockpermutation mit N_B Durchgängen, neu bestimmt. Bei Einsatz des genetischen Algorithmus wird bei Erreichen eines Wertes $Y_k \geq Y_{\text{exit}}$ ein (erfolgreicher) Abbruch der Parameterpermutation vorgenommen und der Klassifikator mit den aktuellen Parametern erstellt².

Abschließend sei bemerkt, dass im Fall einer starken Ungleichverteilung entweder die Relevanz – wenn wenige FG vorhanden sind – oder die Segreganz – wenn sehr viele FG vorhanden sind – einen niedrigen Wert annehmen. Daher sind in Kapitel 7 zur Beurteilung der Klassifikations- und Erkennungsleistung zu jedem Datensatz sowohl die Sensitivität r_p als auch die Spezifität r_n und die daraus ableitbare Falsch-Positiv-Rate f_p angegeben.

In Abhängigkeit der genauen Auslegung eines FSS ist eine Entscheidung darüber zu treffen, wie empfindlich oder wie konservativ die Klas-

²Für alle $a_i, b_i, c_i, d_i \neq 0$ liefert die Summation aller Werte und anschließende Berechnung des Youden-Index identische Ergebnisse. Falls aber in mindestens einem der Blockpermutationsdurchgänge Repräsentanten einer Gruppe fehlen, also falls $a_i + c_i = 0$ oder $b_i + d_i = 0$, so wird $Y_{k,i} = 0$ gesetzt und die Gleichheit gilt nicht mehr.

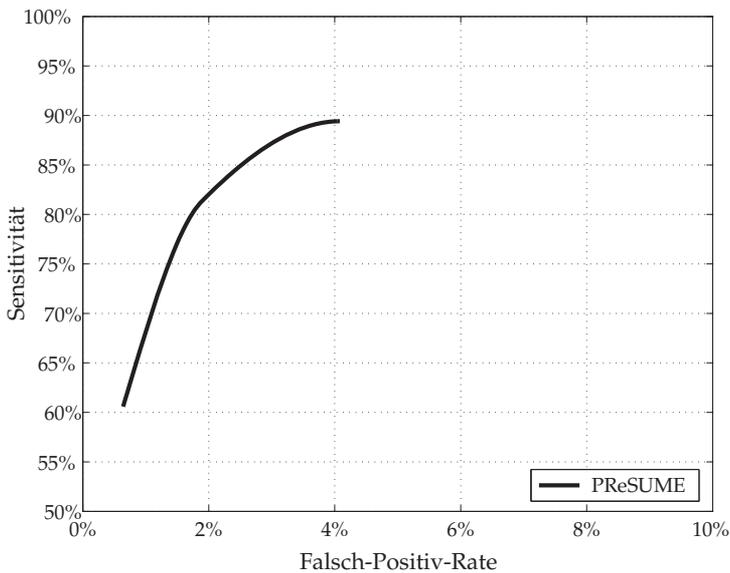


ABBILDUNG 6.10: ROC-Kurve, beispielhaft erstellt auf Basis einiger Validierungsdatensätze

sifikation bezüglich der Richtig-Klassifikation von Fußgängern sein soll. Da die Empfindlichkeit, also die Sensitivität, als Trade-Off zur Fehlalarmwahrscheinlichkeit steht, muss ein Kompromiss gewählt werden. Diese Kompromissmöglichkeiten sind beispielhaft in Abbildung 6.10 dargestellt. Die dort eingetragenen Klassifikationsergebnisse ergeben sich aus unterschiedlicher Wahl des klassenspezifischen Gewichtungsfaktors g_ε .

Bei Gleichgewichtung von Spezifität und Sensitivität, also bei $g_\varepsilon = 1$, stellt der Punkt mit dem größten Abstand zur Winkelhalbierenden das Optimum dar. Dies sei als »Optimalitätskriterium mit Geradensteigung gleich eins« bezeichnet. Bei Ungleichgewichtung und Hinzunahme der in Abschnitt 5.5.7 beschriebenen Erweiterung der SVM durch klassenspezifische Bestrafung mit $g_\varepsilon \neq 1$ erhält man Punkte in der ROC-Kurve, die dem »Optimalitätskriterium mit Geradensteigung ungleich eins« entsprechen.

Für die Umsetzung einer Funktion, insbesondere der FSS, sind Klassifikatoren mit sehr niedriger Falsch-Positiv-Rate, also hoher Spezifität, interessant.

6.5 TRAININGS- UND VALIDIERUNGSDATEN

Die Qualität der Sensordaten erfordert nicht nur eine präzisen Kalibrierung der Sensoren, sondern auch eine umfassende Sammlung von Sensordaten zuvor definierter Verkehrsszenen. Dazu wurde im Rahmen dieser Arbeit eine Testspezifikation entwickelt, die Aufzeichnungen von Szenarien mit applikationsspezifischen Daten ermöglicht und koordiniert. Die Szenarien beinhalten verschiedene Fußgänger in typischen Verkehrssituationen³. Eine vollständige Übersicht über die Trainings- und Validierungsdaten ist in Tabelle 6.5 gegeben. Es folgt eine kurze Zusammenfassung der verwendeten Szenariendaten.

Training Die zum Training und Test verwendeten Messdaten wurden mit dem in Abschnitt 6.1.1 beschriebenen Versuchsfahrzeug aufgezeichnet. Die verwendeten Messdaten aus Datensatz #0₆ beinhalten 8 283 vom PMD-Sensor aufgenommene Objekte. Diese beinhalten 57 Fußgänger und 103 Nicht-Fußgänger. Die Videosequenzen haben eine kumulierte Länge von 11 min 57 s und bestehen aus 73 Einzelszenen. Bei den aufgenommenen Fußgängerszenen traten ausschließlich Einzelfußgänger auf.

Validierung Die Validierungsdatensätze beinhalten insgesamt 16 032 Objektdaten über Fußgänger und 57 412 Objektdaten von Nicht-Fußgängern. Dies ergibt insgesamt 73 444 Objekte. Die Daten sind in Videoszenen mit einer kumulierten Länge von 21 min 4 s enthalten und beinhalten 65 verschiedene Fußgänger und 492 unterschiedliche Nicht-Fußgänger-Objekte. Die Validierungsdateien zeigen 19 unterschiedliche Verkehrsszenen.

³Die dabei aufgebaute Datenbank enthält im Übrigen keine Verkehrsszenen mit Unfällen oder Beinaheunfällen. Bei der Definition der Verkehrsszenen und der Aufnahme wurde streng darauf geachtet, die Sicherheit des Versuchspersonals und aller Verkehrsteilnehmer in vollem Umfang zu gewährleisten.

Daten- satz	Typ	Anzahl FG	Anzahl nFG	Anzahl FG- Objekte L ₂	Anzahl nFG- Objekte L ₁	Anzahl Szenen	Länge (mm:ss)	Dateiname
#0	Training/Test	57	103	12 181	37 506	73	11:57	TrainTest_merge.dat
#0 ₄	Training/Test	57	103	3 039	9 394	73	11:57	TrainTest_merge.dat
#0 ₆	Training/Test	57	103	2 032	6 251	73	11:57	TrainTest_merge.dat
#0 ₁₀	Training/Test	57	103	1 211	3 747	73	11:57	TrainTest_merge.dat
#0 ₂₀	Training/Test	57	103	615	1 843	73	11:57	TrainTest_merge.dat
#1	Validierung	2	2	7 622	5 194	1	03:02	File_01.dat
#2	Validierung	3	3	633	1 749	1	00:10	File_02.dat
#3	Validierung	1	1	82	134	1	00:03	File_03.dat
#4	Validierung	3	60	220	13 904	1	02:28	File_04.dat
#5	Validierung	1	3	330	517	1	00:12	File_05.dat
#6	Validierung	3	13	4 694	3 160	6	02:33	File_06.dat
#7	Validierung	3	13	330	2 558	1	00:40	File_07.dat
#8	Validierung	3	14	169	1 408	1	00:36	File_08.dat
#9	Validierung	5	49	295	4 487	1	01:12	File_09.dat
#10	Validierung	3	23	173	2 015	1	00:28	File_10.dat
#11	Validierung	4	60	231	2 322	1	02:41	File_11.dat
#12	Validierung	20	208	628	17 333	1	05:38	File_12.dat
#13	Validierung	10	28	368	1 466	1	00:48	File_13.dat
#14	Validierung	4	15	257	1 165	1	00:33	File_14.dat
Σ	Validierung	65	492	16 032	57 412	19	21 : 04	

TABELLE 6.5: Tabelle über die Inhalte der verwendeten Datensätze: Spalten 3 und 4 beschreiben die Anzahl unterschiedlicher FG (Fußgänger) und nFG (Nicht-Fußgänger), Spalten 5 und 6 die Anzahl der aufgenommenen Objekte. Es ist $L = L_1 + L_2$ wie in Gleichung (5.78) verwendet.

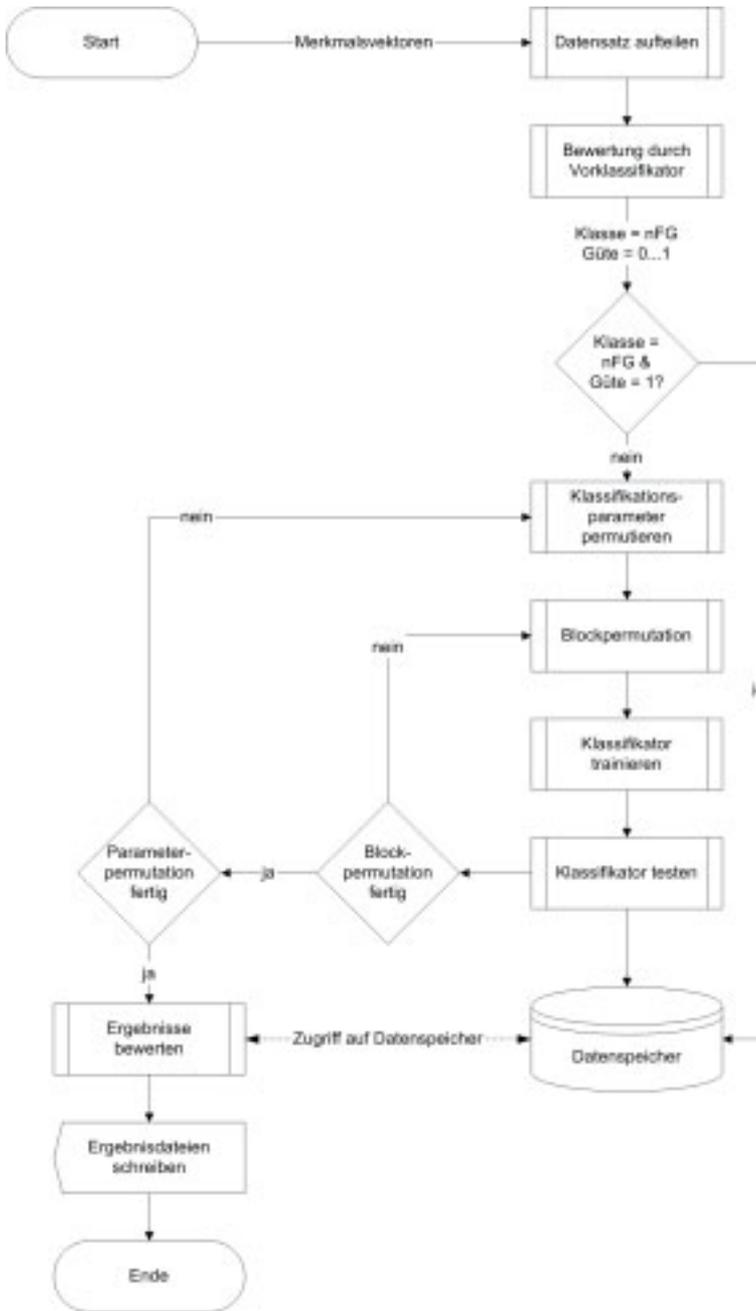


ABBILDUNG 6.11: Flussdiagramm für das automatisierte Trainieren und Testen eines Klassifikators mit verschiedenen Parametern auf Basis eines einzigen Messdatensatzes

7

ERGEBNISSE

Auf Basis der in Kapitel 6 beschriebenen Versuchsdurchführung konnte eine Reihe von Erkenntnissen über die Leistungsfähigkeit von PReSUME gewonnen werden. Details über die Merkmalsextraktion werden in Abschnitt 7.1 vorgestellt, es folgen die Ergebnisse der Detektions- und Klassifikationsraten von Fußgängern in den Abschnitten 7.2, 7.3 und 7.4. Auf das Ziel der Echtzeitfähigkeit der vollständigen Verarbeitungskette wird in Abschnitt 7.5 näher eingegangen. In Abschnitt 7.6 ist eine Übersicht über die internationalen Arbeiten von Universitäten und Forschungsabteilungen zu dieser Thematik und den dabei erzielten Ergebnissen gegeben. Im Anhang D sind zusätzlich einige Systemgrenzen der verwendeten Sensorik beschrieben, die die Grenzen einer weiteren Performancesteigerung erklären.

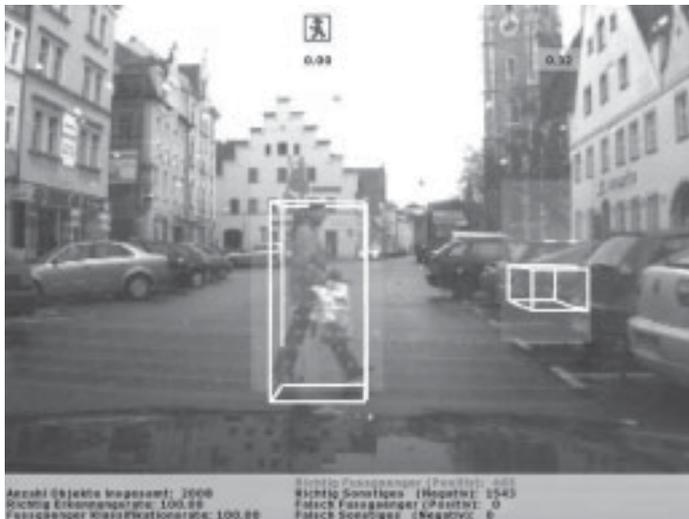


ABBILDUNG 7.1: typische Szene mit zwei vom PMD-Sensor detektierten Objekten – die Person wurde erfolgreich als Fußgänger klassifiziert, das Fahrzeugheck erfolgreich als Nicht-Fußgänger

7.1 ANALYSE DER MERKMALSEXTRAKTION

Der Wertebereich eines Merkmals bei Verwendung der SVM hat einen kritischen Einfluss auf die Gesamtperformance. Eine Vergrößerung des Wertebereiches eines einzelnen Merkmals erhöht dessen Einfluss im Vergleich zu allen anderen Merkmalen. Da in jedem Fall die Trennfähigkeit eines einzelnen Merkmals geringer als die Trennfähigkeit aller Merkmale zusammen ist, führt eine solche Skalierung zu einer Verschlechterung der Gesamtklassifikationsrate. Eine präzise Anpassung aller Merkmale zueinander setzt eine identische Verteilung voraus. Sind alle Merkmalsverteilungen zum Beispiel mit einer Gauß-Verteilung approximierbar, wird das beste Klassifikationsresultat durch die Skalierung auf identische Mittelwertdifferenzen und identische Varianzen zwischen den klassenspezifischen Merkmalsvektoren jedes Merkmals m_i mit $i \in \{1, 2, \dots, M\}$ erreicht.

Die Annahme einer Gauß-Verteilung der Merkmalsvektoren ist bei vielen Merkmalen jedoch nicht vertretbar. Dies wird zum Beispiel an den Histogrammen in Abbildung A.1 und A.10 deutlich (siehe Anhang A). Eines der wichtigsten Designkriterien bei der Merkmalsextraktion ist daher die inhärente Normierung jedes Merkmals (zum Beispiel die Normierung der Fläche mit dem Objektumfang, wie sie in Gleichung (5.7) angewandt wurde) und die anschließende Skalierung auf einen global einheitlichen Wertebereich (zum Beispiel 0 bis 1). Falls diese abschließende Skalierung nicht gegeben ist und die Wertebereiche der Merkmale stark voneinander abweichen, wird ein suboptimales Klassifikationsergebnis erzielt.

7.2 DETEKTIONSRATE

Die Regionendetektion des PMD-Systems ermöglicht die Objektdetektion, also eine Detektion von Hindernissen, die sich im Erfassungsbereich der Sensorik befinden. Diese Detektion erfordert keinerlei vor- oder nachgeschaltete Klassifikation.

Zur Beschreibung der Detektionsraten muss ein Bereich definiert werden, in dem eine Detektion erwartet wird. Wir definieren daher einen Kollisionsbereich. Alle Objekte, die sich innerhalb dieses Kollisionsbereiches befinden, werden (sofern sie diesen nicht wieder verlassen) mit dem Fahrzeug kollidieren (siehe dazu Abbildung 7.2). Betrachtet man ausschließlich das Verhältnis der im Kollisionsbereich detektierten Objekte zu den insgesamt im Kollisionsbereich befindlichen Objekten, so erhält man die Detektionsrate d . Beschränkt man sich auf die Detektionsraten einzelner Klassen, so erhält man die Detektionsrate d_p für Fußgänger und d_n für Nicht-Fußgänger. Von den insgesamt 102 im Kollisionsbereich befindlichen Objekten (davon 35 Fußgänger und 67 Nicht-Fußgänger) wurden 1 Fußgän-

ger und 6 Nicht-Fußgänger nicht detektiert. Die daraus resultierenden Detektionsraten sind in Tabelle 7.1 aufgelistet. Diese hohen Detektionsraten zeigen, dass der Ansatz der Objektdetektion mit dem PMD-System vielversprechend ist.

Detektionsrate d	93,1 %
Detektionsrate für Fußgänger d_p	97,1 %
Detektionsrate für Nicht-Fußgänger d_n	91,0 %

TABELLE 7.1: Zusammenfassung der Detektionraten für alle kollisionsrelevanten Objekte

Für eine mögliche Nichtdetektion von Fußgängern gibt es verschiedene Ursachen, die in Tabelle 7.2 zusammengefasst sind. Ursache eins und zwei sind algorithmisch in PMD-TRACKING und PMD-REGIONENDETEKTION zu beeinflussen. Fall drei kann nur durch eine Hardwareverbesserung abgeschwächt, in keinem Fall jedoch vollständig eliminiert werden. Mögliche Lösungsansätze dazu sind eine Erhöhung der Strahlungsleistung der IR-Strahlungsquellen oder eine Erhöhung der Ladungstrennung.

Die Entstehung von Geisterzielen ist, vor allem im Vergleich zu Stereokamerasystemen, deutlich deterministischer. Die Möglichkeiten sind in Tabelle 7.3 aufgelistet. Überreichweiten, wie schon in Abschnitt 4.1.1 erläutert, entstehen dann, wenn ein Objekt durch seine Größe und hohe Reflektivität in einer Entfernung detektiert wird, die größer als der Eindeutigkeitsbereich ist. Algorithmen zur Vermeidung der Überreichweiten sind denkbar, wurden jedoch im Rahmen dieser Arbeit nicht entwickelt, da in den untersuchten innenstädtischen Bereichen solche Hindernisse nicht vorgekommen sind. Eine weitere Ursache für Geisterziele kann Interferenz zwischen zwei PMD-Systemen sein. Auch diese störenden Effekte können behandelt werden, jedoch ist dies ebenfalls nicht Teil der vorliegenden Arbeit.

7.3 KLASSIFIKATIONS RATEN

Die Leistungsfähigkeit der Klassifikation wurde durch die Verwendung einer 5-fachen stratifizierten Kreuzvalidierung auf einem Trainingsdatensatz (Datensatz #0₆) ermittelt und zusätzlich auf Validierungsdatensätzen (Datensätze #1 bis #14) überprüft. Die Häufigkeitsbeschreibung der richtig-positiven, richtig-negativen, falsch-positiven und falsch-negativen Einstufungen führt zu den Werten Sensitivität r_p , Spezifität r_n und Falsch-Positiv-Rate f_p (siehe dazu auch Tabelle 6.4). Als Zielfunktion wird der Youden-Index Y_k aus Gleichung (6.9) verwendet. Die Optimierung dieses Index lie-

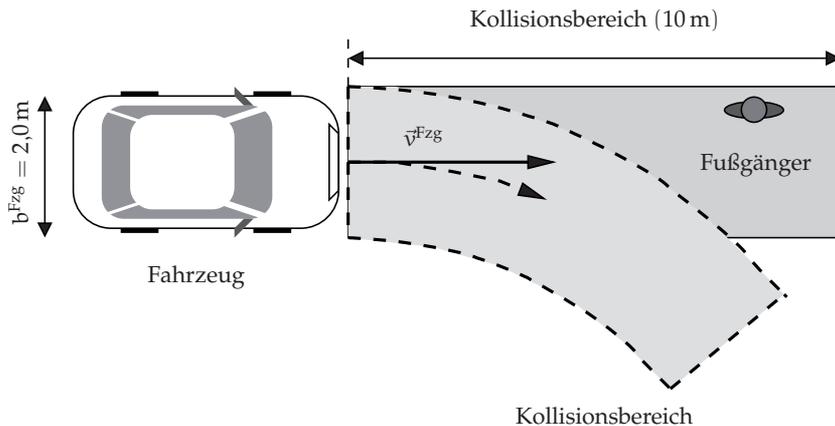


ABBILDUNG 7.2: Definierter Kollisionsbereich (prädizierter Fahrschlauch) für Geradeaus- und Kurvenfahrt

Nichtdetektion von Fußgängern möglich durch ...

- ▶ zu hohe laterale oder relative Geschwindigkeit
- ▶ zu geringen Abstand zu anderen Objekten
- ▶ zu geringe Reflektivität der Kleidung

TABELLE 7.2: Gründe für die Nichtdetektion von Fußgängern

Entstehung von Geisterzielen möglich durch ...

- ▶ Überreichweiten
- ▶ Interferenz mit anderen PMD-Systemen
- ▶ Gischt

TABELLE 7.3: Gründe für die Entstehung von Geisterzielen

fert die beste Klassifikationsleistung unter Berücksichtigung einer Gleichgewichtung der Sensitivität und Spezifität, also einer Gleichgewichtung der Empfindlichkeit und Robustheit des Klassifikationssystems.

Eine Zusammenfassung der Ergebnisse zeigt Tabelle 7.4. PRESUME 1 und PRESUME 2 beschreiben die Klassifikation aller Objekte, die sich im Erfassungsbereich der Sensorik befinden. Sie unterscheiden sich lediglich in der Höhe der klassenspezifischen Bestrafung. PRESUME 2* beschränkt sich zusätzlich auf die unfallrelevanten Objekte, also auf die Objekte innerhalb des Kollisionsbereiches.

Die vollständigen Validierungen sind in den Tabellen 7.10, 7.11 und

System	Sensitivität r_p	Falsch-Positiv-Rate f_p
PreSUME 1	91,2 %	3,0 %
PreSUME 2	69,0 %	0,7 %
PreSUME 2*	84,0 %	< 0,086 %

TABELLE 7.4: Zusammenfassung der Klassifikationsergebnisse

7.12 beschrieben. Details über die verwendeten Klassifikatoren mit Beschreibung der Parameter enthält Tabelle 7.5. Diese SVM-Parameter wurden durch die Optimierungsschritte im Parameterfitting hergeleitet und liefern die nahezu optimalen Klassifikatoren. Die Parameter zwischen PreSUME 2 und PreSUME 2* unterscheiden sich nicht.

System		PreSUME 1	PreSUME 2
Datensatz		0 ₄	0 ₄
Klassifikation		SVM	SVM
Kernel	K	Gauß	Gauß
Kernelparameter	σ	1,5	1,5
Bestrafungsgewichtung	C	8,0	8,0
SV-Auswahl	ϵ	0,01	0,01
klassenspez. Bestrafung	g_ϵ	1	30
Klassifikationshistorie	H	1	1

TABELLE 7.5: Verwendete Klassifikatoren und deren Parameter

7.3.1 Einfluss der SVM-Parameter

Um einen tieferen Einblick in die Einflüsse der einzelnen Parameter der SVM auf die Trainings- und Testdaten zu gewinnen, wurde ein Brute-Force Ansatz mit reduzierter Anzahl verwendeter Objekte durchgeführt. Wählt man aus dem Datensatz nur 1/20-stel der Objekte aus, so ist zwar die Aussagekraft des Absolutwertes des berechneten Gütekriteriums herabgesetzt, jedoch bleibt eine grundsätzliche Aussagekraft der relativen Änderungen des Kriteriums in Abhängigkeit der Parameterauswahl erhalten. Da es sich um drei veränderbare Parameter (ϵ , C und σ) handelt, die in jeder Kombination in einem Gütewert Y_k resultieren, müssen 4 Dimensionen dargestellt werden. Eine farbcodierte Darstellung des Y_k -Wertes in einem 3D-Parameterraum ist jedoch zu unübersichtlich, daher wurde die Dimensionalität der Darstellung reduziert. Das Ergebnis ist in Anhang B in den Abbildungen B.1 bis B.3 (für ein reines PMD-System) und B.4 bis B.6 (für

PREsUME) visualisiert. Die Abbildungen zeigen Y_k in Abhängigkeit von jeweils 2 Parametern mit konstantem dritten Parameter in Form einer Höhenkarte. Ein höherer Wert in der Ebene weist auf eine grundsätzlich bessere Parameterkombination hin.

7.3.2 Ergänzende Ergebnisvalidierung

Zur Validierung der Ergebnisse wurden weitere Untersuchungen vorgenommen. Es fanden zusätzlich Testfahrten statt, bei denen die prinzipielle Leistungsfähigkeit ständig von einem Experten beobachtet wurde. Die daraus gewonnenen Erkenntnisse belaufen sich hauptsächlich auf die Beschreibung der Objekten, die systematisch richtig klassifiziert beziehungsweise typischerweise falsch klassifiziert wurden. Diese Objekttypen sind in Tabelle 7.6 aufgelistet. Der Umfang dieser Bewertungsfahrten beläuft sich auf ungefähr 80 Stunden.

Klassifikation meistens richtig

- ▶ einzelne Fußgänger
- ▶ Radfahrer von hinten
- ▶ vollständig sichtbare Fahrzeuge

Klassifikation häufig richtig

- ▶ Fußgängergruppen
- ▶ Radfahrer von der Seite
- ▶ teilverdeckte Fahrzeuge bzw. Fahrzeughecks

Klassifikation meistens falsch

- ▶ niedrig montierte Verkehrsschilder
- ▶ kleine Werbeplakate (Größe etwa 80 cm × 50 cm)

TABELLE 7.6: Objekttypen, die »meistens richtig«, »häufig richtig« beziehungsweise »meistens falsch« klassifiziert werden

7.4 ERKENNUNGSRATEN

Die Beschreibung der Leistungsfähigkeit von PREsUME aus den vorangegangenen beiden Abschnitten in Bezug auf die Detektions- und Klassifikationsraten stellen im Prinzip nur Teilergebnisse dar. Sie beschreiben die für die Erkennung wichtigen Teilaspekte separat. Für Fahrzeugfunktionen wie das FSS spielen jedoch die *Gesamterkennungsraten* aller *unfallrelevanten Objekte* eine Rolle, also sowohl die Kombination aus Detektion und Klassifikation als auch eine Beschränkung auf die Objekte, die im Falle eines Unfalls beteiligt wären. Desweiteren sind Aussagen über Klassifikationsraten

auch nur dann vergleichbar, wenn eine definierte Beschreibung des Erfassungsbereiches und der Art der Detektion vorliegt. Wir kombinieren daher die Ergebnisse aus den Abschnitten 7.2 und 7.3 unter Berücksichtigung der unfallrelevanten Objekte. Es ergeben sich die Werte wie in Abbildung 7.3 gezeigt.

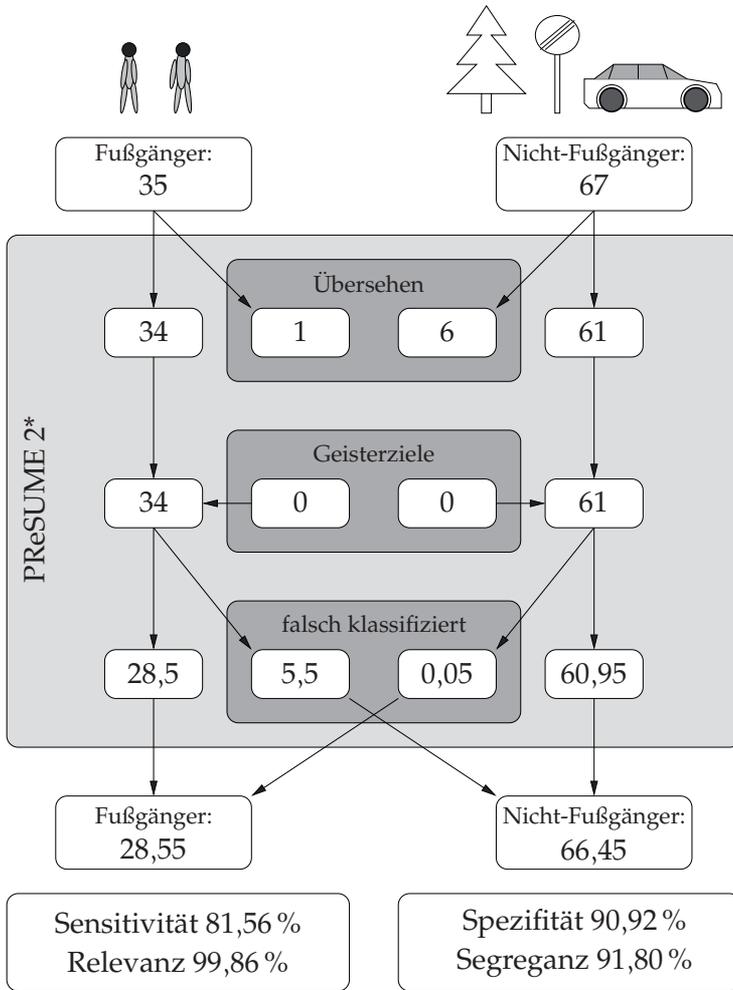


ABBILDUNG 7.3: Analyse der Leistungsfähigkeit von PReSUME 2*: oben ist die tatsächliche Anzahl unfallrelevanter Objekte gegeben, unten die von PReSUME 2* erkannten Fußgänger und Nicht-Fußgänger

Diese Ergebnisse entsprechen einer *Worst Case Abschätzung*. Dies liegt an folgender Vereinfachung, die bei Klassifikationsraten von > 50% gelten. Anstatt die Falschklassifikationsrate auf einzelne Samples von Objekten

ten zu beziehen, wird sie absolut auf ganze Hindernisse bezogen. Ein Beispiel: Bei einer Sensitivität von 80 % werden in Realität alle Fußgänger in 80 % aller Samples richtig erkannt. Dies bedeutet, dass im besten Fall – zum Beispiel durch Hinzunahme eines Klassifikationstrackings – alle 5 Fußgänger richtig klassifiziert werden, oder sich im schlechtesten Fall die 20 % Falschklassifikationsrate auf einen einzigen Fußgänger konzentrieren, somit also nur 4 Fußgänger richtig klassifiziert werden. Genau diese Worst Case Abschätzung wurde hier angenommen.

Die Ergebnisse zeigen, dass im Validierungsdatensatz durchschnittlich 0,14 % der von PReSUME 2* erkannten Fußgänger in Wahrheit Nicht-Fußgänger sind (99,86 % Relevanz). Rund 8,2 % der vom System erkannten Nicht-Fußgänger sind hingegen tatsächlich Fußgänger (91,8 % Segreganz).

Die Werte für Relevanz und Segreganz ergeben sich zwar aus den Klassifikationsergebnissen aus Tabelle 7.4, jedoch nur unter Hinzunahme einer bestimmten Klassenhäufigkeiten $P(S_1)$ und $P(S_2)$ und gegebenen klassenspezifischen Detektionsraten. Sind diese zusätzlichen Informationen nicht gegeben, so können nur die Spezifität und die Sensitivität berechnet werden.

7.5 ERFORDERLICHE RECHENZEIT

Um eine Abschätzung über den Rechenaufwand der Erkennung und die Echtzeitfähigkeit des Gesamtsystems geben zu können, wurden eine Reihe von Laufzeitmessungen durchgeführt. Als Referenzsystem dient dabei ein Intel Core(TM)2 1,86 GHz PC mit 1 GByte RAM mit dem Betriebssystem Microsoft Windows¹ XP. Die Zeitbetrachtung bezieht sich auf die reine Datenverarbeitung ohne Initialisierung der Filter und ist über alle Objekte der Datensätze #1 bis #14 ermittelt, die eine gültige Klasse besitzen und nicht durch die Vorklassifikation aussortiert wurden. Tabelle 7.7 zeigt die Rechenzeiten der einzelnen Filter an. Die Schwankungen der Rechenzeiten sind auf die unterschiedliche Verarbeitungszeit zwischen großen und kleinen Objekten² sowie den Nichtdeterminismus des verwendeten Betriebssystems, zum Beispiel bei Dateisystemzugriffen, zurückzuführen. Die Berechnung der PMD-Regionen und das PMD-Tracking finden bereits im PMD-Sensor statt und werden daher in dieser Betrachtung ignoriert.

Aus Tabelle 7.7 ist zu erkennen, dass besonders die Berechnung der PMD- und der MV-Merkmale wenig Rechenzeit benötigt. Dies ist ein zentrales Ergebnis der vorliegenden Arbeit, da die echtzeitfähigkeit der Bild-

¹Microsoft Windows ist ein eingetragenes Markenzeichen der Firma Microsoft.

²Mit *groß* und *klein* wird hier die Anzahl an Pixeln referenziert, die ein Objekt im Bild beansprucht.

verarbeitungsalgorithmen ein wichtiges Erfolgskriterium für den Serieneinsatz darstellt.

Betrachten man die Gesamtrechenzeit, so setzt sich diese aus der Kernrechenzeit und einem Overhead, der durch das ADTF und die Datenhaltung sowie -übergabe entsteht, zusammen. Da die Verarbeitungskette aber nicht von jeweils einem einzelnen Objekt, sondern von jeweils einem Paket von Objekten angesteuert wird, wird in einem Durchlauf das gesamte Paket verarbeitet. Dies sind alle Objekte, die sich zu einem Zeitpunkt im Erfassungsbereich der Sensorik befinden. Rein statistisch werden durchschnittlich 4 Objekte pro Paket verarbeitet, wobei 2 der 4 Objekte bereits vom Vorklassifikator aussortiert werden und somit keinen nennenswerten Einfluss auf die Gesamtrechenzeit T_g haben. Aus diesem Grund sind in Tabelle 7.8 nicht so sehr die durchschnittlichen und die minimalen Rechenzeiten, sondern die maximalen Rechenzeiten T_k , T_g und T_a von Bedeutung. Der abgeschätzte Overhead T_a beläuft sich auf circa 40% ... 55% der reinen Verarbeitungszeit, also der Gesamtkernrechenzeit T_g .

Bei maximaler Ausnutzung der zur Verfügung stehenden Rechenkapazität wären also Systemupdateraten von deutlich über 100 Hz möglich. Die Worst Case Abschätzung von knapp 16 Hz für $T_{g,max}$ beziehungsweise 25 Hz für $T_{k,max}$, zeigt die geringen Anforderungen von PRESUME an die Rechenkapazität. Unter der Einschränkung der verwendeten PC Hard- und Software kann es als quasi echtzeitfähig bezeichnet werden.

7.6 VERGLEICH MIT INTERNATIONALEN FORSCHUNGS- & ENTWICKLUNGSAKTIVITÄTEN

Die Leistungsfähigkeit von PRESUME ist auch im Vergleich mit den Systemen anderen Forschungs- und Entwicklungsabteilungen bemerkenswert. Die folgenden Abschnitte geben einen Überblick über den Stand der Technik bezüglich der untersuchten Verfahren zur Klassifikation von Fußgängern und den dabei erzielten Ergebnissen. Betrachtet man die Framerate,

Filter	Mittelwert	Std.-Abw.	Minimal	Maximal
Vorklassifikator	0,002 ms	0,003 ms	< 0,001 ms	0,248 ms
Segmentierung	0,09 ms	0,07 ms	0,035 ms	2,20 ms
PMD-Merkmale	0,29 ms	0,07 ms	0,16 ms	4,12 ms
MV-Merkmale	0,78 ms	0,48 ms	0,08 ms	6,63 ms
Hauptklassifikation	1,18 ms	0,13 ms	1,08 ms	6,85 ms
Gesamtzeit	2,34 ms		1,35 ms	20,06 ms

TABELLE 7.7: Rechenzeiten für einzelne Verarbeitungsschritte pro Objekt

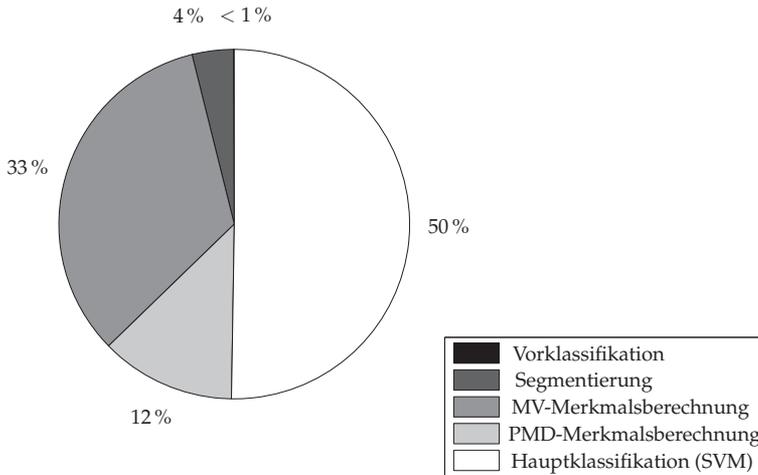


ABBILDUNG 7.4: Verhältnis der Rechenzeiten der wichtigsten Verarbeitungsschritte für einzelne Objekte

die Klassifikationsgüten und den Umfang an verwendeten Messdaten, so zeigt sich PReSUME als stärkstes Erkennungssystem. Lediglich in der maximalen Reichweite muss es sich anderen Systeme geschlagen geben. Da aber die Anforderungen an die Reichweite eines autonom bremsenden Systems deutlich geringer als die eines warnenden Systems ist, spielt dieses Kriterium keine so wesentliche Rolle.

Filter	Mittelwert	Minimal	Maximal
Vorklassifikator	0,003 ms	< 0,001 ms	0,488 ms
Segmentierung	0,10 ms	0,016 ms	4,33 ms
PMD-Merkmale	0,172 ms	< 0,001 ms	8,12 ms
MV-Merkmale	0,73 ms	< 0,001 ms	13,05 ms
Hauptklassifikation	0,69 ms	< 0,001 ms	13,49 ms
Gesamtkernrechenzeit T_k	1,70 ms	0,02 ms	39,47 ms
Gesamtrechenzeit T_g	3,49 ms	0,04 ms	63,76 ms
Overhead $T_a = T_g - T_k$	1,79 ms	0,02 ms	24,29 ms

TABELLE 7.8: Rechenzeiten für einzelne Verarbeitungsschritte und die Gesamtverarbeitung pro Paket von Objekten

7.6.1 *Universität Parma*

Die Arbeitsgruppe um Prof. Dr. Alberto Broggi³ forscht seit vielen Jahren an der Fußgängererkennung auf Basis optischer Sensoren.

[69] beschreibt einen Ansatz zur Fußgängererkennung basierend auf einem monokularen IR-System. Es wurde in Kooperation mit der Elektronik Forschung der Volkswagen AG entwickelt. Die Algorithmik der Bildverarbeitung sucht nach symmetrischen und warmen Objekten mit spezifischen Größen und Seitenverhältnissen der Bounding Boxes. Es wird ein ebener Straßenverlauf vor dem Fahrzeug angenommen, um sowohl die Größe der Bounding Boxen, als auch deren Positionen und damit den Abstand des Fußgängers vom Fahrzeug bestimmen zu können. Nach Angaben der Autoren ist diese Annahme bis zu einer Entfernung von 20 m auch dann noch erlaubt, wenn die Straße einen hügeligen Verlauf annimmt. Erst bei Entfernungen größer 20 m sei mit Fehlern zu rechnen, die die (allgemeine) Ungenauigkeit des Messverfahrens merklich beeinflussen. Die Reichweite der Erkennung setzt sich aus 2 Bereichen zusammen: einem Nahbereich von 7 m bis 20 m und einem Fernbereich von 20 m bis 43,5 m. Die kürzeste Detektionsentfernung ist durch die Einschränkung definiert, dass die Straße und damit der Fußpunkt der Objekte sichtbar sein muss. Es kann eine Bildwiederholrate von knapp 8 Hz auf einem 1,8 GHz PC mit 512 MByte Arbeitsspeicher erreicht werden. Über die Güte der Erkennung wurde keine Angabe gemacht.

In [70] wird ein Ansatz mit zwei FIR-Kameras beschrieben. Dieser erreicht eine Sensitivität von etwas über 80 % bei gleichzeitiger Falsch-Positiv-Rate von 10 %. Für die Detektion und Klassifikation wurden die Flächen im Amplitudenbild (also den Informationen aus der Objektwärme) ein Kantendetektor und das V-Disparitätenbild benutzt, aus denen die Merkmale Objektgröße, Abstand, Seitenverhältnis und ein Indikator für das »Vorhandensein eines Kopfes« extrahiert wurden. Die alleinige Verwendung der warmen Gebiete führte zu einer Richtig-Klassifikationsrate von 70 %. Die Rechenzeit für die BV beläuft sich auf 84 ms mit einem Pentium IV Prozessor mit 2,8 GHz Taktrate und 1 GByte Arbeitsspeicher.

Eine weitere relevante Arbeit ist in [71] beschrieben. Dort werden die Aspekte für ein Werkzeug zur Evaluierung der Leistungsfähigkeit detailliert. Unter Anderem wird die ROC-Kurve zur Performanceanalyse von Klassifikationsfilter erklärt. Diese trägt die Sensitivität über der Falsch-Positiv-Rate auf (siehe dazu auch Abschnitt 6.4). Zusätzlich wird die Genauigkeit der Detektion, also das Zuordnungsproblem zwischen den realen und den vom Videosystem ausgewählten Bereichen im Bild, mit in den Lösungsansatz einbezogen.

³aus <http://www.ce.unipr.it/people/broggi/>, zuletzt besucht im April 2008

Den aktuellen Stand der Technik präsentiert [72]. Der dort verwendete Ansatz beruht wiederum auf einem Stereo-FIR-System, auf dem eine Bildverarbeitung implementiert ist, die nach warmen Bildbereichen und Kanten sucht und diese als Merkmale in das Disparitätenbild einfließen lässt. Anschließend werden typische morphologische und thermische Eigenschaften des Kopfes zur Validierung verwendet. Es finden keine von der Zeit oder von Bewegung abhängigen Merkmalsextraktionen statt. Unter Verwendung eines Pentium IV mit 2,8 GHz und 1 GByte RAM wurde eine Systemtaktzahl von knapp 12 Hz und eine Korrektorklassifikationsrate von etwas über 80 % bei gleichzeitiger Fehlalarmrate von knapp 10 % erreicht.

7.6.2 Technische Universität Cluj-Napoca

An der rumänischen Universität Cluj-Napoca (zu deutsch: Klausenburg) wird ein auf einer Stereokamera basierendes System zur Klassifikation von Fußgängern und Radfahrern entwickelt [36]. Dabei werden die Bewegungen der Objekte im Raum analysiert und es wird ein 3D-Bewegungsfeld erzeugt. Die extrahierten Merkmale sind die Objektbreite, -höhe, -geschwindigkeit und weitere Charakteristika über die Bewegung selbst. Diese werden durch die Bewegungssignatur und die Periodizität der Bewegungssignatur beschrieben. Die Bewegungssignatur entspricht der Varianz der Bewegungsvektoren des Objektes. Es kann gezeigt werden, dass sich bewegende Fußgänger eine hohe Bewegungssignatur aufweisen, da die Arme, Beine und der Körper unterschiedliche Bewegungen ausführen. Bei starren Objekten tritt dieser Effekt nicht auf. Für die Klassifikation wird ein Bayes-Klassifikator benutzt.

Die Ergebnisse aus [36] zeigen, dass eine Korrektorklassifikationsrate von 85 % bei Testszenen mit 600 Objekten erreicht wurden. Von den 600 Objekten waren 40,5 % Fußgänger. Die Reichweite wurde dabei auf 35 m begrenzt. Die echtzeitfähige Klassifikation wurde auf einem Intel Core 2 Duo Prozessor mit 2,4 GHz E6600 und 20 fps ausgeführt. Für die rechenintensive Berechnung des Tiefenbildes wird eine spezielle Hardware eingesetzt.

7.6.3 Technische Universität Hamburg-Harburg

In [73] sind einige Ansätze zur Fußgängerklassifikation auf Basis von Radardaten beschrieben. Dort kommt ein FPGA-basierter PD-Radar zum Einsatz. Die Hauptmerkmale, die zur Klassifikation herangezogen werden, basieren auf der periodischen Änderung der Ausdehnung und Periodizität der Geschwindigkeiten radial auf den Sensor zu oder sich von ihm weg bewegender, gehender Fußgänger. Das Ergebnis einer Ganganalyse von Fußgängern zeigt, dass die typische Werte bei einer Gehbewegung relativ

wenig streuen. Mittelwerte der Schrittlänge liegen bei ca. 75 cm, die mittlere Geschwindigkeit um 5 km/h bis 5,3 km/h und die Schrittfrequenz recht konstant um 1,9 Hz. Das auffälligste Merkmal, das aus den Radardaten extrahiert werden kann, ist die Geschwindigkeitsschwingung des Rumpfes, also die periodische Änderung der Absolutgeschwindigkeit des Körpers der Person.

7.6.4 Technische Universität Wuppertal

In Zusammenarbeit mit Delphi Delco Electronics Europe und der Delphi Deutschland GmbH entstand ein Fußgängererkennungssystem, welches auf einer reinen Monovideo-Kamera basiert [74]. Ausgehend von einer Objektdetektion, basierend auf einem virtuellen Stereo-Ansatz und Inverse Perspective Mapping (IPM) wird eine Bildstabilisierung und eine auf morphologischen Parametern bezogene Erkennung durchgeführt. Besonderer Wert wurde auf eine schnelle Datenverarbeitung gelegt, so dass das Gesamtsystem auf einem 3 GHz PC mit 64 Bildern pro Sekunde arbeitet. Die Fußgängererkennung funktioniert prinzipiell bis zu einer Entfernung von 50 m, getestet wurde auf einem Datensatz von 50 000 Bildern. Es wurden in dieser Veröffentlichung jedoch keine Erkennungsraten angegeben.

7.6.5 Australian National University

In [75] wird ein Stereovideoansatz zur Erkennung von Fußgängern beschrieben. Die zur Detektion benötigten Merkmale werden durch eine Höhenkarte und das Geschwindigkeitsdisparitätenbild geliefert. Ein Vorklassifikator entfernt alle Objekte, deren Größe ein Minimum unterschreiten oder ein Maximum übersteigen. Formbasierte Merkmale, die mit einem 3×3 Sobeloperator erzeugt werden, fließen in die Klassifikation mit einer SVM ein. Es werden zwei SVM eingesetzt, da eine Variante für die Frontansicht von Fußgängern und eine für die Seitenansicht optimiert ist. Die Autoren gelangen zu der Erkenntnis, dass, anders als bei reiner 2D-BV, ein 3D-Ansatz basierend auf der Generierung eines Tiefbildes des Stereosystems ein besseres Verständnis der Verkehrsszene ermöglicht und somit die Fehlklassifikationsrate reduziert werden kann. Dies kann zum Beispiel bei Abbildungen von Personen auf Leinwänden am Straßenrand relevant sein.

Das Stereovideo-System besteht aus zwei Sony EX470L Kameras mit einer Basisweite von ungefähr 30 cm. Die Bildverarbeitung wird von einem 2,4 GHz Pentium 4 PC geleistet, der die 8 Bit Graustufenbilder mit einer Auflösung von 320×240 Bildpunkten in einer durchschnittlichen Rechenzeit von 43 ms verarbeitet. Das Ergebnis der Klassifikation auf 2400 Test-

bildern zeigt, dass circa 83,5 % der Fußgänger richtig erkannt wurden, die Falsch-Positiv-Rate liegt bei 0,4 %.

7.6.6 Daimler AG

Forschungsabteilungen der Daimler AG betreiben seit vielen Jahren intensive Forschungs- und Entwicklungsarbeiten auf dem Gebiet der Fußgängererkennung. Besonders hervorzuheben sind dabei die Arbeiten von Prof. Dr. Dariu Gavrilă⁴, der einen wesentlichen Beitrag zur Fußgängererkennung mit Videosystemen in Forschungsprojekten geleistet hat. Einen umfassenden Überblick über die Arbeiten liefert [52]. Dort wird einerseits das im Preventive Safety for Unprotected Road User (PROTECTOR) Projekt entwickelte Stereovideo System beschrieben. Zusätzlich wird ein Verfahren zur Optimierung der im Stereosystem verwendeten Parameter vorgestellt, das auf vollständige und sehr ressourcenintensive Trainings- und Testdurchläufe verzichten kann. Darüber hinaus wird eine Testmethodik präsentiert, die die Messungen im zweidimensionalen Videobild in das dreidimensionale Umfeld abbildet, damit ein besserer Bezug zu einer darauf aufbauenden Fahrzeugfunktion möglich ist. Die Tests wurden sowohl auf einem Prüfgelände, als auch im Straßenverkehr durchgeführt [76].

Das dort vorgestellte Stereosystem berechnet ein Tiefenbild, aus dem potenzielle Fußgängerobjekte extrahiert werden. Darauf baut ein Chamfer-System mit formbasierten Merkmalen auf. Dieses arbeitet pixelbasiert – das heißt nicht konturbasiert – und sucht im Videobild nach prototypischen Vorlagen. Diese sogenannten prototype templates werden offline aus einer großen Sammlung an möglichen Objekttemplates generiert. Da nicht nach jeder Templatevorlage, sondern hierarchisch nach den Repräsentanten verschiedener Cluster gesucht wird, ist eine echtzeitfähige Anwendung möglich. Derzeit wird das Verfahren bei der Verkehrszeichen- und Fußgängererkennung eingesetzt. Anschließend verifiziert ein Texturmerkmalmodul mit Hilfe eines Künstliche Neuronale Netze (KNN) die Detektionen des Chamfer-Moduls und reduziert die Anzahl ausgewählter Regionen. Ein weiterer Verifikationsschritt findet durch die Analyse der Kreuzkorrelation zwischen beiden Videobildern statt. Abschließend wird ein Tracking und eine Risikoabschätzung durchgeführt. Wie bereits erwähnt kann ein vollständiger Test, das heißt die Optimierung einer Zielfunktion in einem sehr großen Parameterraum, aus Kapazitätsgründen oft nicht durchgeführt werden. In diesem Fall kann eine sequentielle Optimierung einzelner aufeinanderfolgender Module erfolgen [77].

⁴aus <http://www.homepageGavrila.de>, zuletzt besucht im Dezember 2007

7.6.7 *Toyota Central R & D Labs*

Im Sensing System Lab der Toyota R & D Abteilung wurden Untersuchungen zur Fußgängerklassifikation mit Hilfe von Millimeterwellenradar und Stereovideo durchgeführt. Dort wird ein Fusionsansatz auf Objektebene angewandt, der mit Hilfe von Wahrscheinlichkeiten über die Existenz, Nicht-Existenz und Unsicherheit der einzelnen Sensoren zu einer Fusionsaussage kommt. Die Klassifikationsfähigkeit dieses Ansatzes, in [78] als Detektionsrate bezeichnet, wird mit 60 % bis 80 % bei einer Entfernung von 10 m bis 30 m angegeben. Durch den Fusionsansatz kann sowohl die Rate der detektierten als auch die Anzahl der übersehenen Fußgänger gesteigert werden. Die Datenbasis zur Evaluierung besteht aus 20 Szenen mit einer nicht bekannten Anzahl an Umfeldobjekten.

7.6.8 *Nissan*

Nissan arbeitet zur Zeit an einem Sicherheitssystem, das den Fahrer bei herannahenden Fußgängern warnt. Die Information über die Fußgängerposition wird mit Hilfe eines Global Positioning System (GPS) Systems im Mobiltelefon des Fußgängers und einer zusätzlichen Übertragung der Fußgängerposition an das Fahrzeugsystem gewonnen. Voraussetzung ist natürlich, dass der Fußgänger mit einem GPS-fähigen Mobiltelefon ausgestattet ist und diese Funktion auch verwendet. Eine großer Anteil der sich in Japan im Umlauf befindenden Mobiltelefone ist bereits kompatibel zu GPS. Über die Zuverlässigkeit oder die Präzision sind bisher keine Daten bekannt.

Datensatz	Sensitivität r_p	Spezifität r_n	Falsch-Positiv-Rate f_p	Bemerkung	Gewichtung
#0 ₆	92,0%	88,5%	11,5%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 1$
#0 ₆	89,8%	90,2%	9,8%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 10$
#0 ₆	88,0%	91,4%	8,6%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 15$
#0 ₆	83,0%	94,1%	5,9%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 20$
#0 ₆	77,2%	95,7%	4,3%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 25$
#0 ₆	68,5%	96,9%	3,1%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 30$
#0 ₆	45,0%	99,1%	0,9%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 35$
#0 ₆	32,4%	99,5%	0,5%	Mittelwerte über alle 5 Blöcke	$g_\varepsilon = 40$

TABELLE 7.9: Änderung der Leistungsfähigkeit von PResUME und jeweils unterschiedlicher klassenspezifischer Bestrafung g_ε bei der Kreuzvalidierung auf Datensatz #0₆

Datensatz	Sensitivität r_p	Spezifität r_n	Falsch-Positiv-Rate f_p	Szeneninhalt
#1	97,9%	97,8%	2,2%	Fußgänger zwischen parkenden Pkw
#2	88,8%	98,9%	1,1%	Fußgängerzone
#3	100,0%	100,0%	0,0%	Fußgänger bei Nacht
#4	85,5%	98,7%	1,3%	lange Fahrt durch eine Stadt
#5	100,0%	99,8%	0,2%	Einzelperson vor Fahrzeug
#6	95,7%	96,6%	3,4%	Personen in diversen Posen
#7	99,4%	99,5%	0,5%	Rollschuhfahrerinnen
#8	93,5%	99,9%	0,1%	Harderstraße
#9	82,4%	95,2%	5,1%	Milchstraße 1
#10	93,1%	99,4%	0,6%	Milchstraße 2
#11	93,9%	96,5%	3,5%	Rund um Polizei
#12	73,4%	96,2%	3,8%	Innenstadtrunde Nord
#13	87,2%	94,2%	5,8%	Kuperstraße
#14	86,4%	85,4%	14,6%	Plakat Harderstraße
\emptyset	91,2%	97,0%	3,0%	Durchschnittswert

TABELLE 7.10: Ergebnisse der Klassifikationen mit PRESUME 1

Datensatz	Sensitivität r_p	Spezifität r_n	Falsch-Positiv-Rate f_p	Szeneninhalt
#1	77,4%	99,4%	0,6%	Fußgänger zwischen parkenden Pkw
#2	70,8%	99,9%	0,1%	Fußgängerzone
#3	98,8%	100,0%	0,0%	Fußgänger bei Nacht
#4	53,2%	99,6%	0,4%	lange Fahrt durch eine Stadt
#5	97,6%	100,0%	0,0%	Einzelperson vor Fahrzeug
#6	89,4%	98,4%	1,6%	Personen in diversen Posen
#7	67,6%	100,0%	0,0%	Rollschuhfahrerinnen
#8	82,8%	100,0%	0,0%	Harderstraße
#9	54,2%	97,8%	2,2%	Milchstraße 1
#10	52,6%	99,8%	0,2%	Milchstraße 2
#11	54,6%	99,7%	0,3%	Rund um Polizei
#12	37,4%	98,0%	2,0%	Innenstadtrunde Nord
#13	62,5%	99,2%	0,8%	Kuperstraße
#14	66,5%	98,6%	1,4%	Plakat Harderstraße
\emptyset	69,0%	99,3%	0,7%	Durchschnittswert

TABELLE 7.11: Ergebnisse der Klassifikationen mit PResUME 2 (klassenspezifischer Bestrafung $g_\epsilon = 30$)

Datensatz	Sensitivität r_p	Spezifität r_n	Falsch-Positiv-Rate f_p	Szeneninhalt
#1	93,9%	99,9%	0,1%	Fußgänger zwischen parkenden Pkw
#2	100,0%	99,6%	0,4%	Fußgängerzone
#3	100,0%	100,0%	0,0%	Fußgänger bei Nacht
#4	82,0%	99,6%	0,4%	lange Fahrt durch eine Stadt
#5	100,0%	100,0%	0,0%	Einzelperson vor Fahrzeug
#6	90,5%	100,0%	0,0%	Personen in diversen Posen
#7	64,3%	100,0%	0,0%	Rollschuhfahrerinnen
#8	100,0%	100,0%	0,0%	Harderstraße
#9	69,8%	100,0%	0,0%	Milchstraße 1
#10	31,7%	100,0%	0,0%	Milchstraße 2
#11	97,2%	100,0%	0,0%	Rund um Polizei
#12	65,3%	99,7%	0,3%	Innenstadtrunde Nord
#13	100,0%	100,0%	0,0%	Kupferstraße
#14	81,6%	100,0%	0,0%	Plakat Harderstraße
\emptyset	84,0%	99,9%	< 0,086%	Durchschnittswert

TABELLE 7.12: Ergebnisse der Klassifikationen mit PReSUME 2* (klassenspezifischer Bestrafung $g_\epsilon = 30$) und Beschränkung auf die kollisionsrelevanten Objekte innerhalb des Fahrtschlauchs

ZUSAMMENFASSUNG UND AUSBLICK

Die Bedeutung von Fahrerassistenzsystemen mit aktivem Eingriff in das Fahrzeug zur Steigerung der Sicherheit im Straßenverkehr hat in den letzten Jahren stark zugenommen. Einen wesentlichen Beitrag zur Realisierung dieser Systeme können vorausschauende Sensoren leisten. Während Radarsensoren schon seit einigen Jahren in Kraftfahrzeugen eingesetzt werden, sind besonders optische Sensoren derzeit im Fokus der Entwicklung. Diese werden jedoch – im Gegensatz zu den Radarsensoren – häufig nur für eine reine Visualisierung oder für die Detektion von Fahrspuren eingesetzt, seltener für eine Umfeldwahrnehmung mit einer Erkennung unfallrelevanter Objekte. Ist nun ein Assistenzsystem mit einer Videosensorik und reiner Visualisierung umgesetzt, so erhöht sich im Allgemeinen die Menge an Informationen, die der Fahrer verarbeiten muss. Diese zusätzlichen Informationen können die Belastung des Fahrers erhöhen. Es ist somit nicht sichergestellt, dass die gewünschte Steigerung der Sicherheit tatsächlich erreicht wird.

Eine Analyse der Statistiken über Unfälle zwischen Pkw und Fußgängern zeigt, dass sich in vielen Unfallsituationen die betroffene Person vor dem Zusammenstoß in einem Bereich vor dem Fahrzeug befunden hat, der sich im Sichtfeld von typischen vorausschauenden Sensoren befindet. Um das Sichtfeld möglichst optimal auszurichten, ist die Positionierung der Sensorik im Bereich der Windschutzscheibe besonders gut geeignet. Weiterhin ist zu erkennen, dass sich durch einen aktiven Eingriff in das Fahrverhalten, zum Beispiel durch die Reduktion der Aufprallgeschwindigkeit, häufig die Verletzungsschwere des ungeschützten Verkehrsteilnehmers reduzieren ließe. Oft ist der für den Eingriff zur Verfügung stehende Zeitraum jedoch äußerst gering. Ist der verbleibende Zeitraum zwischen Sichtkontakt zum Unfallgegner und Primäraufprall so gering, dass die Reaktionszeit des Fahrers zum Eingriff nicht mehr ausreicht, so leitet sich unmittelbar das Ziel einer möglichst zuverlässigen Objektdetektion mit anschließender Situationsanalyse und automatischem Eingriff in das Fahrgeschehen durch ein aktives Fußgängerschutzsystem ab. Erfolgt diese Detektion des Umfeldobjektes mit anschließender korrekter Klassifikation des Umfeldobjektes in die Klasse Fußgänger oder die Klasse Nicht-Fußgänger, so kann das Schutzsystem die Aktoren, dazu zählen die Brem-

sen, eine aktive Haube oder weitere crash-aktive Formelemente, optimal ansteuern. Natürlich ist die Ansteuerung von Aktoren, wie beispielsweise der Bremsen, auch bei vielen Nicht-Fußgänger Objekten sinnvoll. Jedoch kann die Aktivierung von Aktoren zum *Weichmachen* der Fahrzeugfront zu einer Veränderung der Crash-Stabilität des Fahrzeugs führen, die bei einem Unfall mit einem *harten* Objekt nicht erwünscht ist. Ebenso können irrelevante Umfeldobjekte detektiert werden, für die ein Fahrzeugeingriff nicht lohnenswert ist. Denn jeder Fahrzeugeingriff birgt eine zusätzliche Gefahr eines Unfalls, zum Beispiel mit nachfolgenden Fahrzeugen.

In dieser Arbeit wird ein Photonic Mixer Device (PMD)-System und eine Monovideo (MV)-Kamera eingesetzt, um die Objektdetektion und die Klassifikation durchzuführen. Die dafür notwendige Merkmalsextraktion basiert auf diversen Bildverarbeitungsalgorithmen und analysiert die Größe, Formen und die Oberflächen der Objekte. Für die Klassifikation kommen zwei verschiedene, sequentiell arbeitende Algorithmen zum Einsatz. Dies sind ein Schwellwertentscheider und eine Support Vector Machine (SVM). Um die optimalen Parameter der SVM zu finden, stehen ein Brute-Force-Ansatz und ein genetischer Algorithmus zur Verfügung. Auf Basis der vorhandenen Messdaten konnte eine Klassifikationsrate von 91 % für Fußgänger und von 97 % für übrige Objekte nachgewiesen werden. Betrachtet man ausschließlich kollisionsrelevante Objekte, so wurde eine Detektionsrate von über 93 %, für Fußgänger sogar 97,1 %, erreicht. Unter Einsatz einer klassenspezifischen Bestrafung konnte für diese kollisionsrelevanten Objekte eine Spezifität von über 99,914 %, eine Falschalarmrate von unter 0,086 %, eine Sensitivität von 84,0 %, eine Relevanz von 99,86 % und eine Segreganz von 91,8 % erzielt werden.

Das erstmals im Rahmen dieser Arbeit entwickelte System mit der Bezeichnung Pedestrian Recognition System Using a Multi-Sensor Environment (PRESUME) ist neu und einmalig. Die hier untersuchte Sensorik zeigt sich als vorausschauende Sensorik für das Fußgängerschutzsystem (FSS) überaus geeignet. Die hohe Detektionsrate des PMD-Sensors, die vielfältigen Möglichkeiten bei der Merkmalsextraktion aus hochauflösenden Videodaten und die Möglichkeit zur schnellen Datenverarbeitung durch die spezielle Fusion auf Rohdatenbasis sind die herausragenden Eigenschaften dieses Systems. Für eine Fußgängerschutzfunktion, die sich an dem in dieser Arbeit beschriebenen Aktorikkonzept orientiert, kann die vorgestellte Sensorkombination und große Teile der Software nach Einschätzung des Autors als adäquate und ideale Lösung betrachtet werden.

Das Aktionskonzept mit autonomem Bremseneingriff kann trotz möglicher Falschklassifikation eingesetzt werden, da die Geisterzielrate, zumindest auf Basis des im Rahmen der Arbeit untersuchten Datenmaterials, zu Null angenommen werden kann. Somit würde das Fahrzeug bei einem Falsch-Positiv-Alarm zwar nicht auf einen Fußgänger reagieren, in jedem

Fall jedoch auf ein Nicht-Fußgängerobjekt, bei dem ein Bremsingriff ebenfalls gerechtfertigt wäre und damit akzeptabel sein könnte.

Natürlich können auch Optimierungsschritte der Teilkomponenten des Erkennungssystems untersucht werden. Für die Extraktion fußgängerrelevanter Merkmale bieten sich grundsätzlich Merkmale an, die auf Bildfolgen aufbauen. Die Herausforderung besteht dabei in der Entwicklung robuster Merkmale trotz bewegter Kamera und in der Aufnahme einer ausreichend großen Datenmenge, die typische Fußgängerbewegungsmuster enthält. Im Bereich der Klassifikation ist es denkbar, den hier verwendeten Klassifikator auf eine Kette mehrerer Klassifikatoren zu erweitern, wobei jeder einzelne Klassifikator auf eine spezielle Ansicht von Fußgängern oder diskrete Positionen im Sinne von Abständen trainiert werden kann [79].

Doch zusätzlich zu technischen Optimierungen sollten auch Richtlinien festgelegt werden, nach denen solche Gesamtsysteme bewertet werden können. Diese Richtlinien müssen die Art der Tests und die Testumfänge in der Art und Weise festlegen, dass eine Steigerung der Verkehrssicherheit beziehungsweise eine Minderung der Unfallfolgen für möglichst viele der relevanten Unfallszenarien nachgewiesen werden kann. Eine der größten Herausforderungen besteht im gleichzeitigen Erfüllen der gegensätzlichen Forderungen nach einem möglichst großem Testumfang, damit viele Verkehrssituationen adressiert werden können, aber auch einem möglichst kleinem Testumfang, damit eine vollständige Durchführung überhaupt möglich ist.

Diese Arbeit beschreibt also ein technisches System, das den Fahrer in speziellen kritischen Verkehrssituationen unterstützt. In den meisten Fällen beherrscht der Mensch diese Situationen, jedoch können schon kleine Aufmerksamkeitslücken die Reaktionszeit und die Präzision der Reaktion deutlich verschlechtern. Wir Menschen besitzen Sinnesorgane zur visuellen Informationsaufnahme aus der Umgebung inklusive der Möglichkeit zur Erkennung von Entfernungen. Auch der hier gewählte Sensorik-Ansatz liefert 3D- und 2D-Bildinformationen bewegter Sensoren. Unsere visuellen Fähigkeiten helfen uns bei der Wahrnehmung der Umwelt inklusive der Klassifizierung von Gegenständen, der Interpretation von Situationen und der Betrachtung komplexer Zusammenhänge, auch über Zeiträume hinweg. Diese Fähigkeiten haben wir durch eine Millionen von Jahren dauernde Evolution und durch mehrere Jahre dauerndes Training gewonnen. Die zukünftige Sensorentwicklung wird weiterhin immer robustere und präziser messende Sensoren hervorbringen. Sensordatenverarbeitung mit zunehmender Intelligenz, insbesondere durch Nachahmung natürlicher Prozesse, wird ebenfalls zur präziseren Umfeldwahrnehmung beitragen. Bei der Erfüllung dieser Wahrnehmungsaufgabe wird der Unterschied zwischen der menschlichen und der maschinellen Leistungsfa-

higkeit immer geringer. Es ist überaus wahrscheinlich, dass er gänzlich verschwinden wird und Maschinen sogar deutlich schneller und zuverlässiger arbeiten und Fahrzeuge steuern können. In vielen Bereichen, zum Beispiel in Teilen von Produktionsanlagen, ist dies ja bereits geschehen.

Die Nutzung von 3D-Informationen und die Hinzunahme von hochauflösenden 2D-Bilddaten beschreibt daher den richtigen Weg zur Umsetzung eines aktiven Fußgängerschutzsystems. Es ist jedoch noch viel Forschungs- und Entwicklungsarbeit zu leisten, bis vollständig unfallfreies Fahren auf Basis computergesteuerter Fahrerassistenz- und Sicherheitssysteme möglich ist.

A

HISTOGRAMME DER MERKMALE

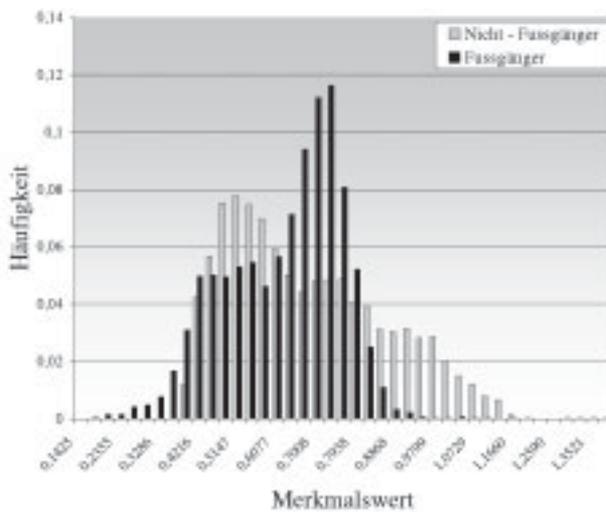
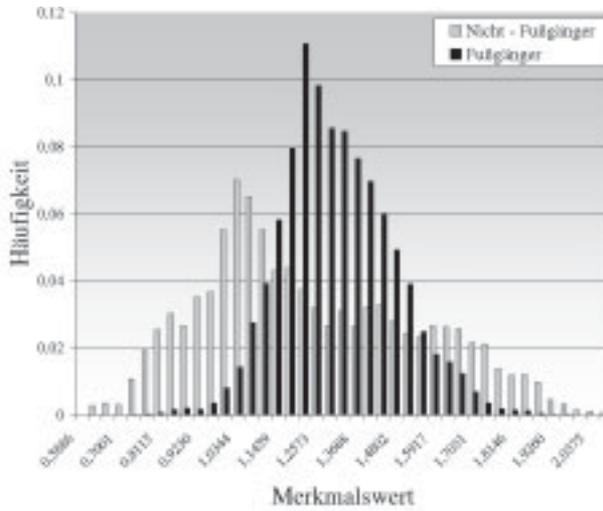


ABBILDUNG A.1: Histogramm über die Objekthöhe HEIGHT



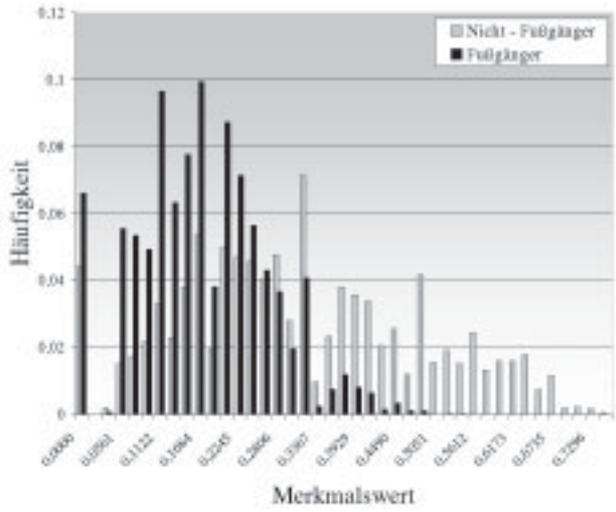


ABBILDUNG A.4: Histogramm über die Orientierung von horizontalen Kanten NUMBEREDGES0DEGREE

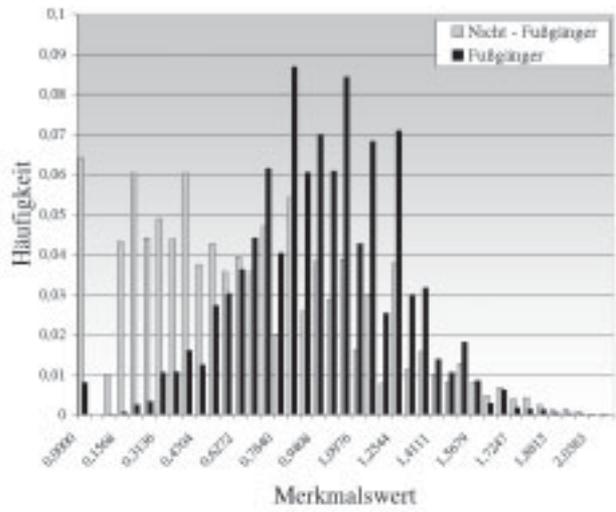


ABBILDUNG A.5: Histogramm über die Orientierung von vertikalen Kanten NUMBEREDGES90DEGREE

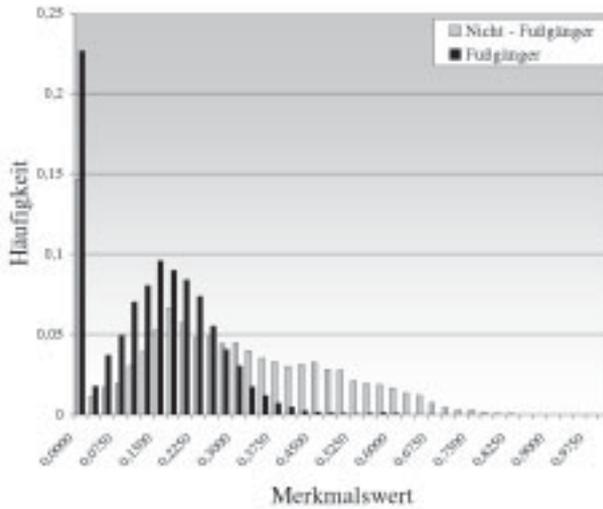


ABBILDUNG A.6: Histogramm über die Homogenität der Region HOMOGENEITY

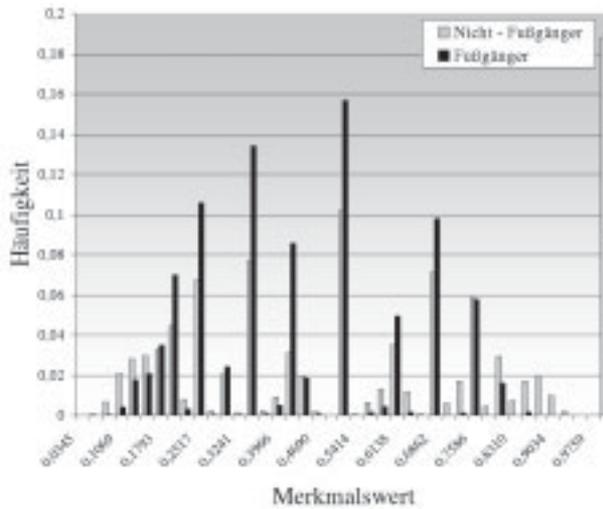


ABBILDUNG A.7: Histogramm über den Füllstand der PMD-Pixel in den obersten zwei Zeilen UPPERBORDERTOWIDTH

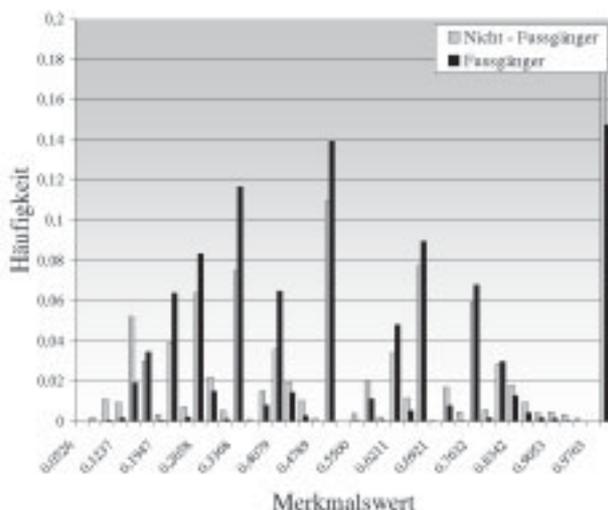


ABBILDUNG A.8: Histogramm über den Füllstand der PMD-Pixel in den untersten zwei Zeilen LOWERBORDERTOWIDTH

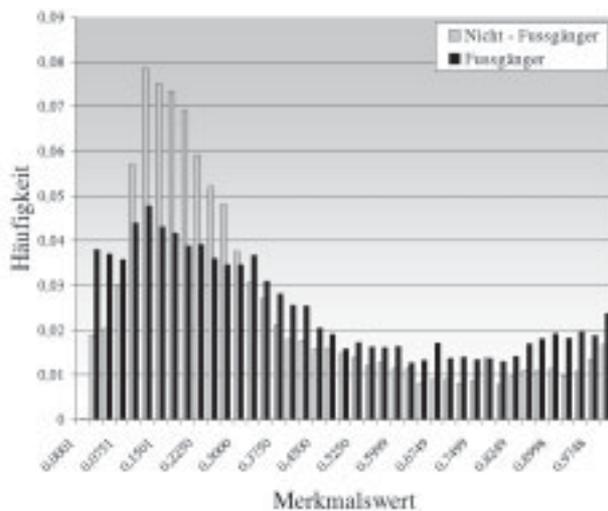


ABBILDUNG A.9: Histogramm über die Orientierung der Objektregion ORIENTATION

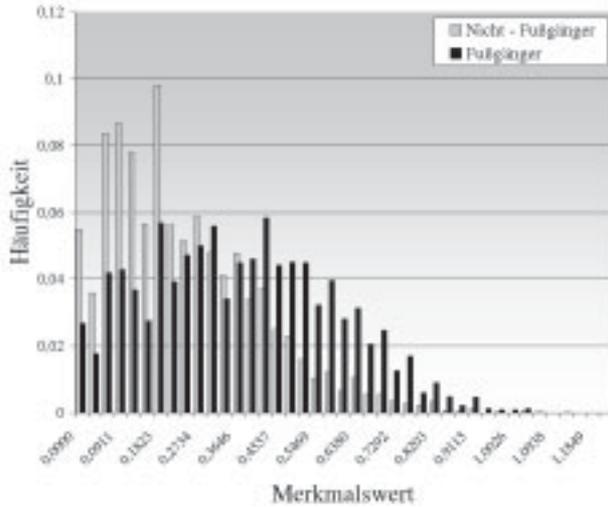


ABBILDUNG A.10: Histogramm über die Spaltenvarianz der PMD-Objektregion SQUARED_COLDIFFFROMAVG

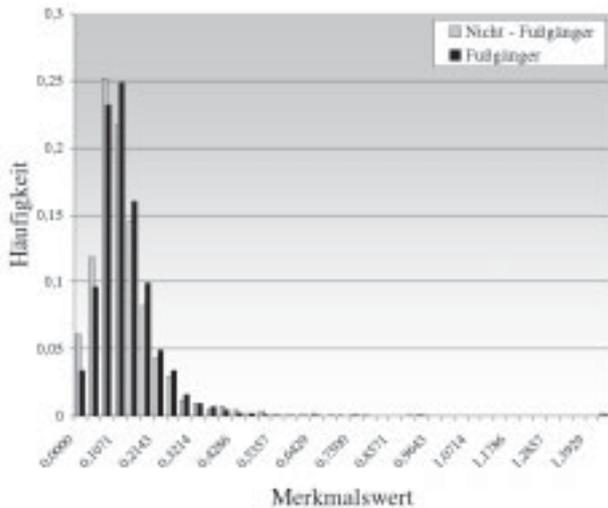


ABBILDUNG A.11: Histogramm über die Beträge des dritten Fourierkoeffizienten der 2D-PMD-Kontur FOURIERAMPLITUDE_1

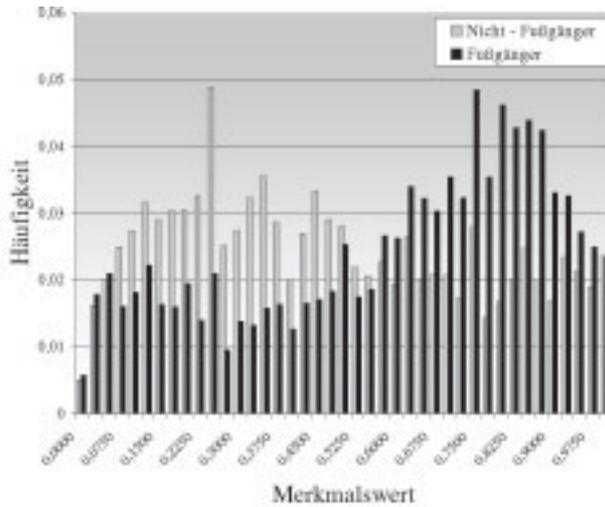


ABBILDUNG A.12: Histogramm über die Phasen des dritten Fourierkoeffizienten der 2D-PMD-Kontur FOURIERPHASE_1

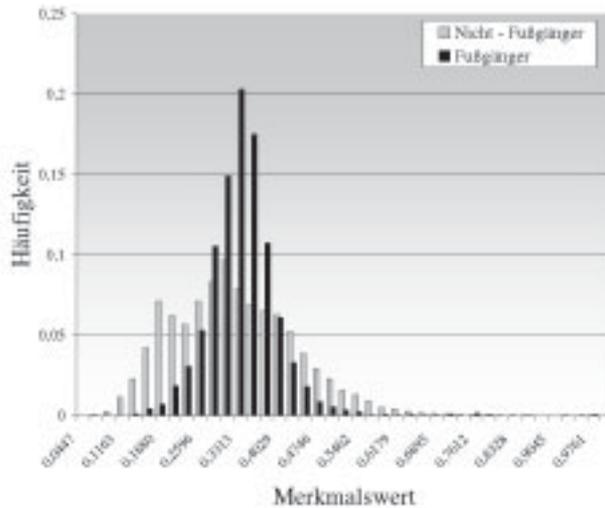


ABBILDUNG A.13: Histogramm über die Verhältnisse von Gradienten aus der MV-Bildregion (Merkmal 18 von EDGEORIENTHIST)

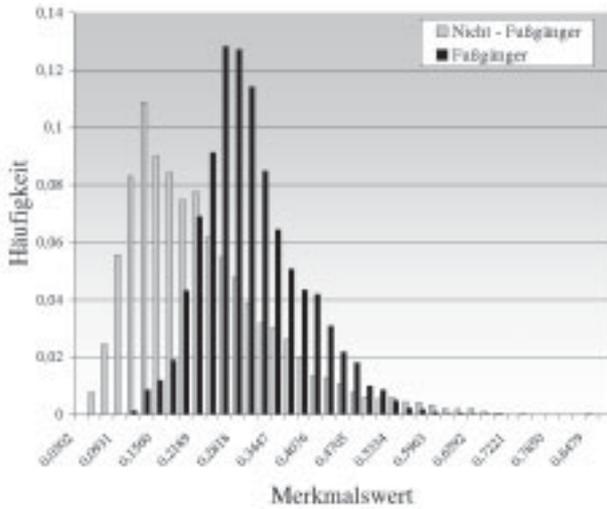


ABBILDUNG A.14: Histogramm über die Verhältnisse von Gradienten aus der MV-Bildregion (Merkmal 45 von EDGEORIENTHIST)

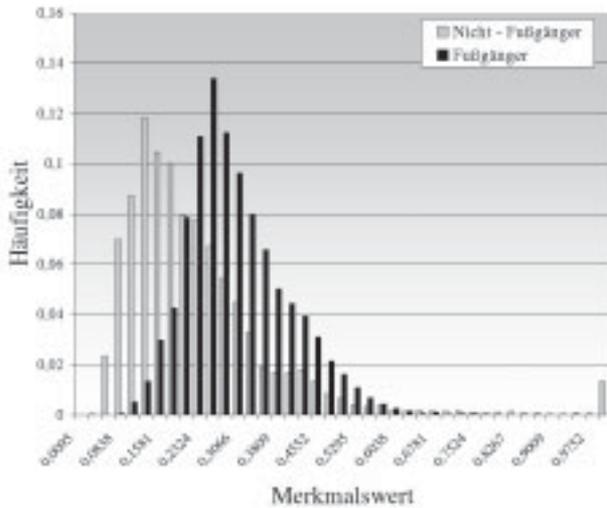


ABBILDUNG A.15: Histogramm über die Verhältnisse von Gradienten aus der MV-Bildregion (Merkmal 76 von SYMEDGEORIENTHIST)

B

EINFLUSS DER SVM-PARAMETER

Um den Gesamteinfluss der hier variierten SVM-Parameter auf die Klassifikationsleistung, gemessen am Gütekriterium Y_k , darzustellen, wurden die Parameter entsprechend Tabelle B.1 permutiert. Die folgenden Abbildungen zeigen nun die 3D-Visualisierungen für ein rein auf PMD-Merkmalen basierendes System und für PReSUME. Es wurde der Datensatz $1_{(1/20)}$ verwendet. Es wird angenommen, dass ein Transfer der Ergebnisse von Datensatz $1_{(1/20)}$ auf die Datensätze $1_{(1/10)}$, $1_{(1/6)}$, $1_{(1/4)}$ und 1 gültig ist.

Bestrafungsterm C

0,1; 2; 4; 6; 8; 10; 12; 14; 16; 18; 20; 25

Kernelparameter σ

0,2; 0,6; 0,8; 1,5; 2; 3; 5; 7; 9; 15

Schwellwert-Parameter ε

0,0001; 0,01; 0,1; 0,5; 1; 1,5

TABELLE B.1: Zu permutierende Parameter

LEISTUNGSFÄHIGKEIT DES PMD-SYSTEMS

Das Maximum von Y_k liegt bei $Y_{k,\max} = 0,7643$ für die Parameter

$$C = 2,0; \sigma = 0,8; \varepsilon = 0,01. \quad (\text{B.1})$$

Die übrigen Werte für Y_k sind in Abbildungen B.1 bis B.3 dargestellt.

LEISTUNGSFÄHIGKEIT VON PReSUME

Das Maximum von Y_k liegt bei $Y_{k,\max} = 0,7797$ für die Parameter

$$C = 2,0; \sigma = 1,5; \varepsilon = 0,0001. \quad (\text{B.2})$$

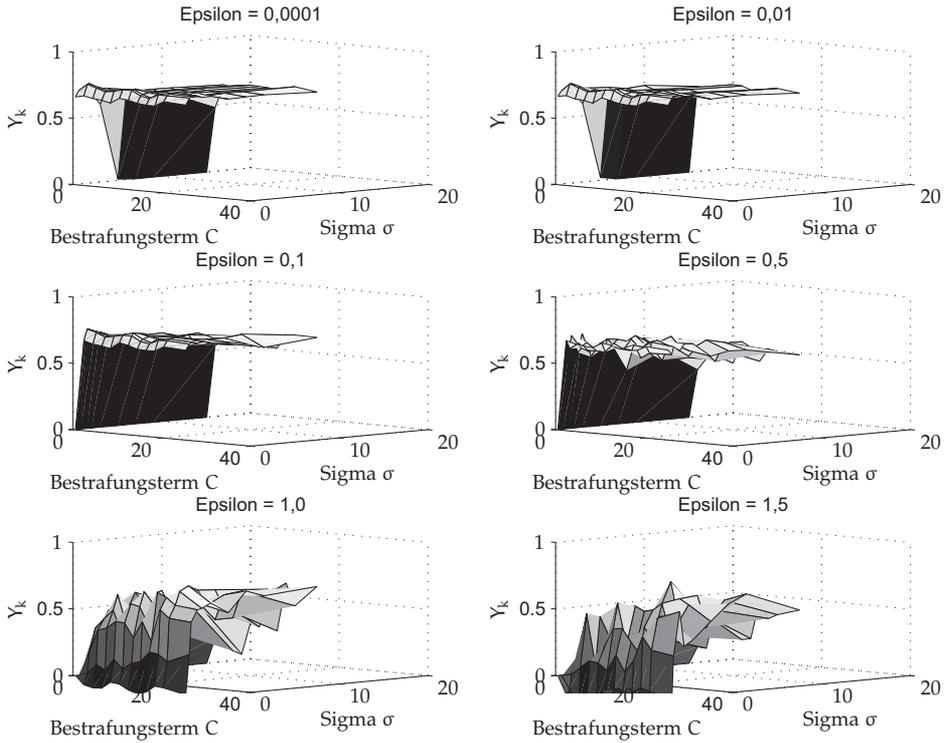


ABBILDUNG B.1: PMD-Klassifikation: Dargestellt sind Ebenenscharen über C und σ mit jeweils konstantem ε .

Die übrigen Werte für Y_k sind in Abbildungen B.4 bis B.6 dargestellt.

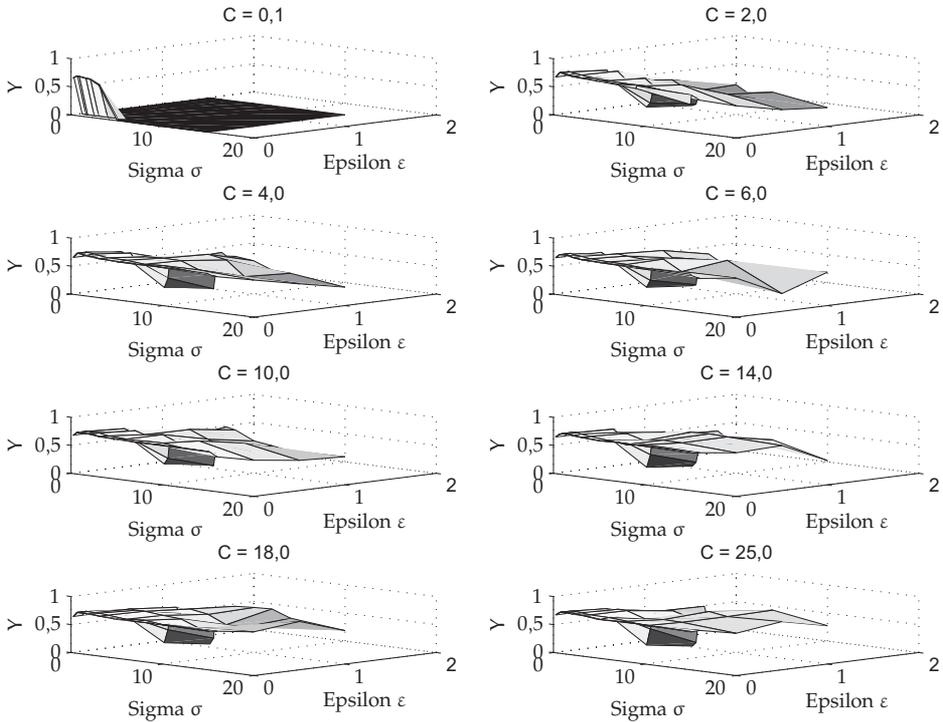


ABBILDUNG B.2: PMD-Klassifikation: Dargestellt sind Ebenenscharen über σ und ϵ mit jeweils konstantem C . Bei $\epsilon = 1$ werden so wenige Merkmalsvektoren zu Support-Vektoren gewählt, dass die Leistungsfähigkeit der Klassifikation auf dem unbekanntem Testdatensatz deutlich herabgesetzt ist.

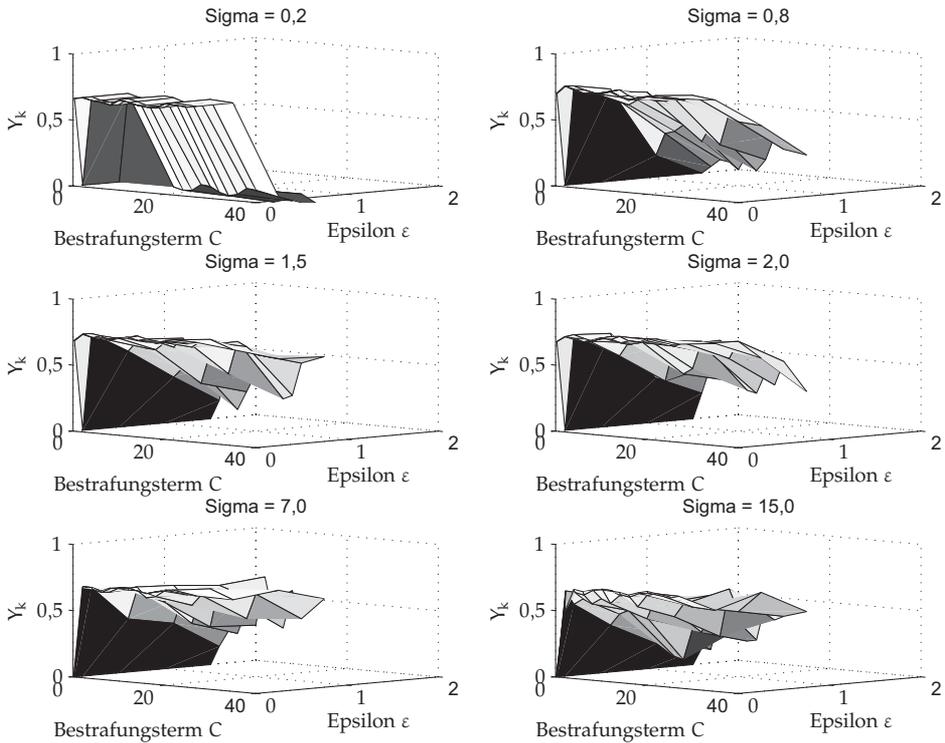


ABBILDUNG B.3: PMD-Klassifikation: Dargestellt sind Ebenenscharen über C und ϵ mit jeweils konstantem σ . Die wie ein gefaltetes Blatt wirkende Ebene enthält jedoch keine Periodizität in sich und keine feste Abhängigkeit zwischen den verschiedenen σ -Werten, sondern ist ein reiner Rauschprozess, der durch die hohen ϵ -Werte zustande kommt.

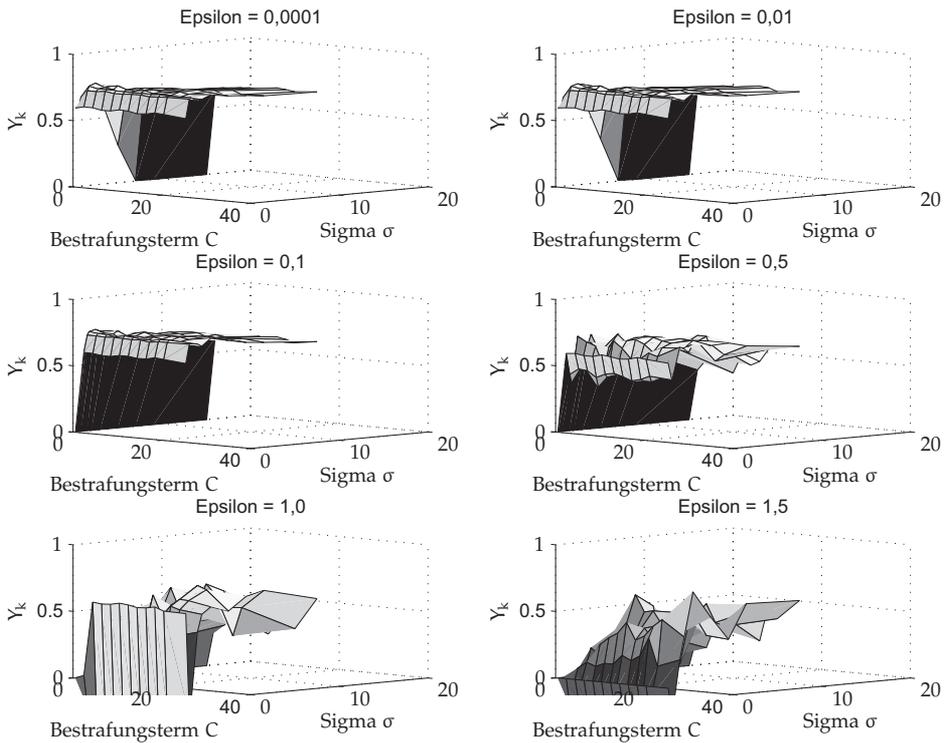


ABBILDUNG B.4: PReSUME: Dargestellt sind Ebenenscharen über C und σ mit jeweils konstantem ϵ .

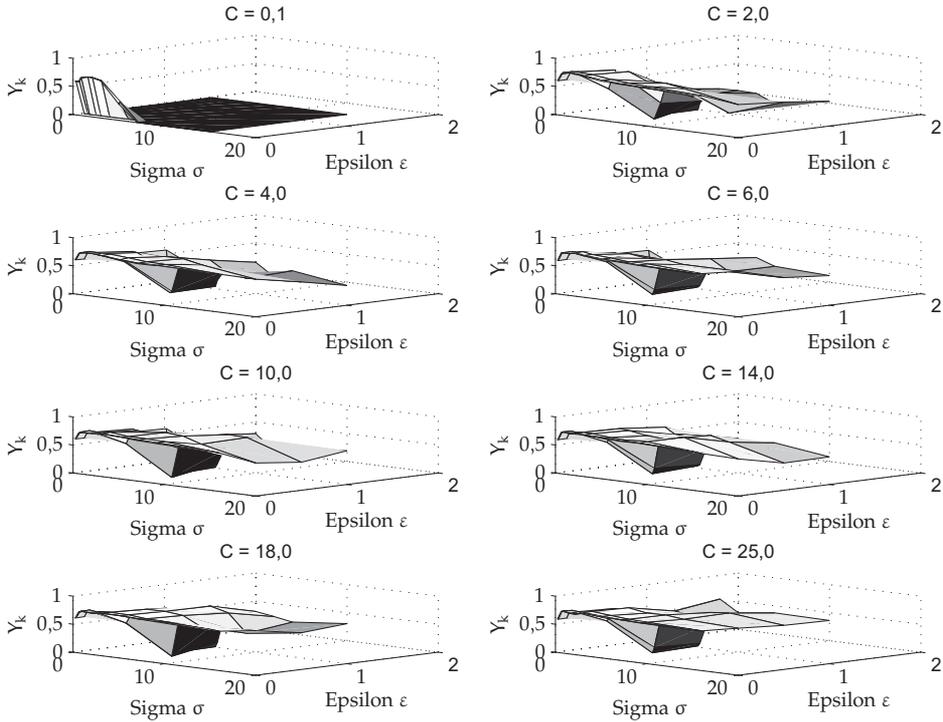


ABBILDUNG B.5: PReSUME: Dargestellt sind Ebenenscharen über σ und ϵ mit jeweils konstantem C . Bei $\epsilon = 1$ werden so wenige Merkmalsvektoren zu Support-Vektoren gewählt, dass die Leistungsfähigkeit der Klassifikation auf dem unbekanntem Testdatensatz deutlich herabgesetzt ist.

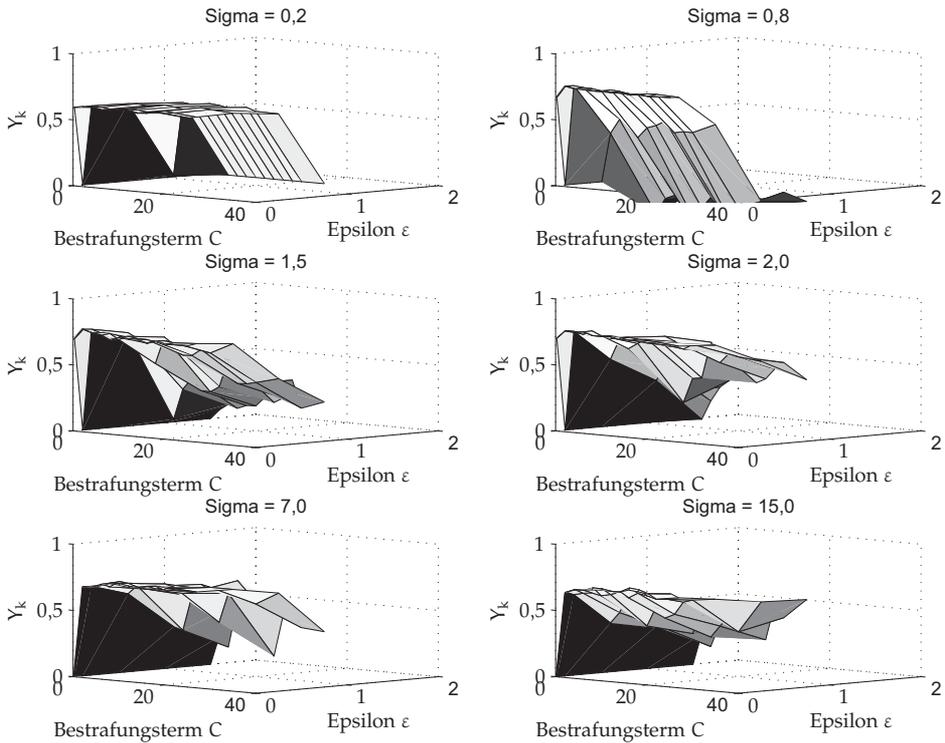


ABBILDUNG B.6: PReSUME: Dargestellt sind Ebenenscharen über C und ϵ mit jeweils konstantem σ . Die wie ein gefaltetes Blatt wirkende Ebene enthält jedoch keine Periodizität und keine feste Abhängigkeit zwischen den verschiedenen σ -Werten, sondern ist ein reiner Rauschprozess, der durch die hohen ϵ -Werte zustande kommt.

C

OPTIMIERUNG MIT GENETISCHEM ALGORITHMUS

Der in Abschnitt 6.3.3 beschriebene genetische Algorithmus wurde eingesetzt, um eine Beschleunigung der Parameteroptimierung zu erzielen. In Abbildung C.1 ist ein Durchlauf gezeigt, bei dem eine (suboptimale) Startpopulation durch zufällige Mutation, Rekombination und Selektionen zu einer (nahezu optimalen) Parameterkombination führt. Die Startpopulation ist in Tabelle C.1 aufgelistet, als Klassifikator wurde wiederum die SVM mit einem Gauß-Kernel verwendet.

	C	σ	ϵ
Individuum 1	100	500	0,001
Individuum 2	100	100	1,2
Individuum 3	1	1	1,5
Individuum 4	0,1	0,1	0,1

TABELLE C.1: Parameter der Startpopulation

Es wurde bei der Startpopulation darauf geachtet, dass sich die Parameter deutlich außerhalb guter Bereiche befinden. Durch Mutation, Rekombination und Selektion (in diesem Fall beinhaltet die Selektion das Überleben des aktuell besten Individuums) wird sehr schnell ein relativ gutes Ergebnis mit $Y_k \geq 0,7$ erzielt. Dies ist auf die Rekombination und auf die »Hochebene« zurückzuführen, die etwa in Abbildung B.4 oben zu erkennen ist. Im weiteren Verlauf wirkt sich die Mutation und Selektion auf eine fortlaufende Verbesserung der Fitness auf. Der genetische Algorithmus überzeugt durch seine schnelle Konvergenz in sinnvolle Bereiche, erweist sich jedoch als recht träge und suboptimal im weiteren Optimierungsprozess. Dies ist auf zwei Ursachen zurückzuführen. Erstens ist die Startpopulation recht klein gewählt, daher muss die Selektion sicherstellen, gute Individuen überleben zu lassen, was eine gewisse Trägheit verursacht. Zweitens erstreckt sich die angesprochene Hochebene in unserem Fall auf einen sehr großen Wertebereich, bei dem viele kleine lokale

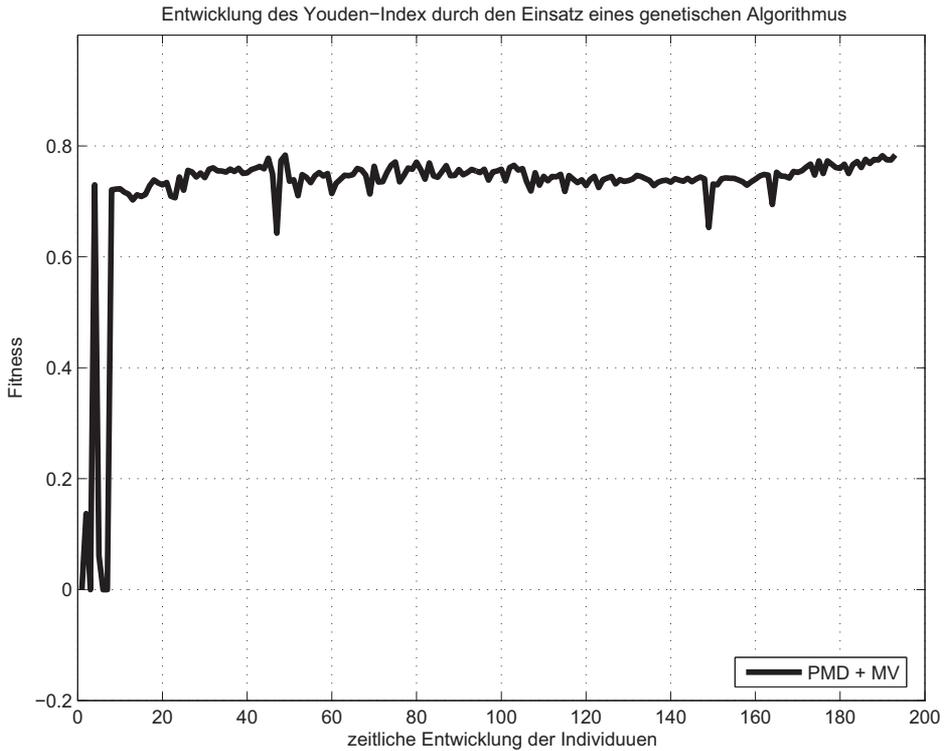


ABBILDUNG C.1: Entwicklung des Fitnesswertes Y_k beim Einsatz eines genetischen Algorithmus

Maxima existieren, die die Performance aller Optimierungsfunktionen erschwert.

D

SYSTEMGRENZEN DES VERWENDETEN SENSORSYSTEMS

Die gewonnenen Erkenntnisse über die theoretischen Grenzen des Sensorsystems inklusive der Sensordatenverarbeitung sind im Folgenden aufgelistet.

MAXIMALE REICHWEITE

Die Bildverarbeitungsmethoden der PMD-Merkmalsextraktion benötigen häufig Regionen mit einer Mindestgrößenanforderung des abgebildeten Objektes. Diese Anforderungen sehen eine Regionengröße von mindestens 3×5 Pixeln pro Objekt vor, um verwertbare Merkmale berechnen zu können. Quantitativ decken sich die Abschätzungen mit den Werten aus einer internen Untersuchung (unveröffentlichter Bericht). Als Konsequenz aus der minimalen Pixelzahl ergibt sich eine maximal erlaubte Entfernung eines Fußgängers gegebener Größe. Je nach gegebener Fußgängerbreite liegt dieser Abstand bei 15 m bis 20 m, eine Detaillierung liefert Abschnitt »Reichweite und Auflösung« auf den folgenden Seiten.

REFLEKTIVITÄT DER OBJEKTE

Neben der Objektgröße begrenzt die Reflektivität die Sichtweite für den PMD-Sensor. Fußgänger weisen typischerweise stark unterschiedliche Reflektivitäten auf, häufig aber einen relativ niedrigen Wert. Ist er zu niedrig, kann der Fußgänger nicht detektiert werden.

NUTZUNG DER GRÖSSENINFORMATION

Die Genauigkeit der Abmessungen ist von mehreren Faktoren abhängig. Bei größeren Entfernungen steigt das Quantisierungsrauschen der Objektgrößenmessung. Bei der Objektbreite ist der Sonderfall zu betrachten, dass ein Objekt den linken oder rechten Rand des Erfassungsbereiches des PMD-Sensors abdeckt, das heißt das mindestens ein Pixel der äußersten rechten

oder äußersten linken Spalte von der Objektregion belegt ist. Die Objektbreite kann lediglich als Mindestbreite interpretiert werden. Bei geringer Entfernung – und zwar innerhalb des damit definierten Nahbereiches – reicht nun der vertikale Erfassungsbereich zur vollständigen Erfassung eines durchschnittlich großen Fußgängers nicht mehr aus. In diesem Fall kann die Höhenangabe nur als Mindesthöhe interpretiert werden. Falls sich ein Objekt in diesem Nahbereich befindet, fehlen Details der Regionen am oberen und eventuell am unteren Rand, und der trainierte Klassifikator kann nur noch ein suboptimales Ergebnis liefern. Ein Lösungsansatz ist die Nutzung unabhängiger Klassifikatoren für den Nah- und Fernbereich.

REICHWEITE UND AUFLÖSUNG

Die Amplitudenwerte der PMD-Kamera bieten eine Umfeldwahrnehmung, die physikalisch ähnlich der eines Nahinfrarotsystems ist, allerdings mit einer deutlich geringeren Auflösung. Das Verhältnis der Pixelanzahl liegt bei 1 : 300 für eine 64×16 PMD-Kamera und eine Videokamera mit VGA Auflösung. Der Nachteil der kleineren Auflösung hat Einfluss auf die Bildverarbeitung, vor allem in der maximalen Entfernung der zu untersuchenden Objekte. Jedoch führt eine Erhöhung der Auflösung nicht immer zu einer Erhöhung der maximalen Reichweite. Dieser Abschnitt erklärt die Grenzen der Auflösungsvergrößerung eines PMD-Systems. Die Veränderung der Pixelzahl in der Kamera bewirkt zwei Effekte. Einerseits wird die maximale Reichweite des Systems durch eine kleinere Photodiodenfläche eines einzelnen Pixels reduziert, andererseits wird die maximale Reichweite des Systems erhöht, da für die notwendigen Bildverarbeitungsalgorithmen oft eine Mindestpixelanzahl pro Objekt benötigt wird.

Es sei ein Beispielobjekt (Fußgänger) mit folgenden Eigenschaften gegeben: Höhe $h = 1,8$ m, Breite $b = 0,5$ m, Reflektivität $\sigma = 20\%$ und feste Ausgangsleistung der Strahlungsquellen. Die Diagramme in Abbildung D.1 zeigen die Abhängigkeit der maximalen Reichweite d_{\max} des PMD-Systems als Funktion der Auflösung R . Der Wert R entspricht dabei dem Produkt aus der horizontalen und der vertikalen Auflösung. Zusätzlich sei ein konstanter horizontaler Öffnungswinkel α der Kamera gegeben.

Die obere Kurve in Abbildung D.1 beschreibt die maximale Sichtweite des PMD-Systems d_1 bezogen auf den Beispielfußgänger ohne Rücksicht auf eine Mindestpixelzahl p_0 . Eine höhere Auflösung führt zu einer kleineren Fläche pro Pixel, damit zu einer geringeren Empfindlichkeit und geringeren Reichweite. Die Werte sind sowohl durch Messungen, also auch durch Interpolation entstanden¹.

¹Die maximale Reichweite für sehr große und sehr kleine Auflösungen r_s wurden geschätzt.

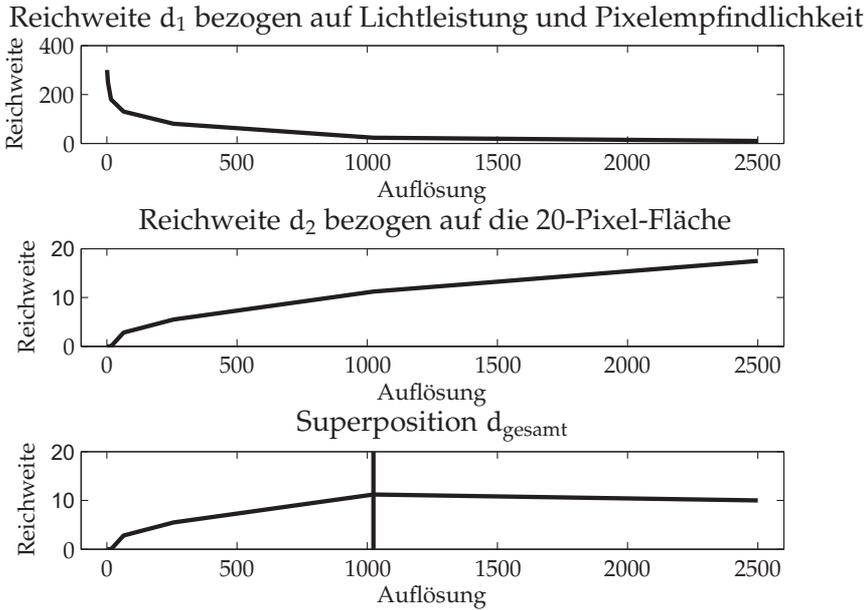


ABBILDUNG D.1: Bestimmung der optimalen Auflösung des PMD-Systems zur Fußgängerdetektion und -klassifikation

In der Mitte ist nun eine Kurve berechnet worden, die ausschließlich das Kriterium der Mindestpixelzahl $p_o = 20$ Pixel in Betracht zieht. Damit wird konkret auf die Objektgröße eingegangen. Für den Beispielfußgänger ergibt sich eine umso größere maximale Reichweite, je höher die Auflösung ist. Gleichungen (D.1) liefern genau diesen Zusammenhang, also die maximale Reichweite d_2 als Funktion der Objektbreite² b , des horizontalen Öffnungswinkels α und der Kameraauflösung R . Das Ergebnis dieser Kurve ist also eine höhere maximale Reichweite bei höherer Auflösung.

$$d_2 = \frac{b/3}{\tan \alpha}, \quad \alpha = \frac{55^\circ}{2\sqrt{r_s}} \tag{D.1}$$

Da für ein vollständiges System beide Bedingungen erfüllt sein müssen, ergibt sich die tatsächliche maximale Reichweite d_{gesamt} aus der Überlagerung der oberen und mittleren Funktion zu

$$d_{gesamt} = \min \{d_1, d_2\}. \tag{D.2}$$

²Es genügt, sich auf die Breite des Fußgängers zu beschränken, da diese immer kleiner als die Höhe ist.

Die untere Kurve in Abbildung D.1 zeigt d_{gesamt} über r_s . Es ergibt sich ein lokales Maximum. Für die oben angegebenen Größen, also den Eigenschaften des Beispielfußgängers und der Bestrahlungsstärke, liegt dieses lokale Maximum bei $r_s = 1000$ und damit sehr nahe der senkrechten Markierung, welche die Auflösung des hier verwendeten PMD-Systems zeigt. Demnach liegt die Auflösung in einem akzeptablen Bereich. Ohne Überarbeitung der Pixelempfindlichkeit oder Erhöhung der Bestrahlungsstärke wird die Leistungsfähigkeit des PMD-Systems von einer Verringerung oder Erhöhung der Auflösung kaum profitieren.

STRAHLDICHTE UND ÖFFNUNGSWINKEL

Unter der Annahme, dass die von den PMD-Strahlungsquellen bestrahlte Fläche sphärisch ist, siehe Abbildung D.2 links, also dass alle Objekte den gleichen Abstand haben, ist die bestrahlte Fläche A_o proportional zum Produkt der Öffnungswinkel $\theta_h \in [0^\circ; 180^\circ]$ und $\theta_v \in [0^\circ; 180^\circ]$. Es gilt also

$$A_o \propto \theta_h \theta_v \quad (\text{D.3})$$

für alle Öffnungswinkel $\theta_{h,v}$ mit $0 \leq \theta_{h,v} < \pi$.

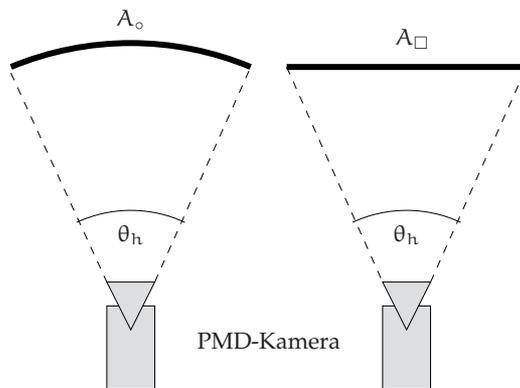


ABBILDUNG D.2: Schematische Abbildung von zwei PMD-Kameras mit horizontalem Öffnungswinkel θ_h , die eine sphärische Fläche (links) und eine ebene Wand (rechts) beobachten

Die Strahldichte L an der Kamera ist umgekehrt proportional zur bestrahlten Fläche. Da die Stärke der ausgesandten Infrarotstrahlung Φ der Strahlungsquellen konstant ist, aber eine vom Öffnungswinkel abhängige Fläche bestrahlt wird, gilt eine zur Strahldichte L umgekehrt proportionale

Abhängigkeit vom Öffnungswinkel

$$L_o \propto \frac{1}{\theta_h \theta_v}. \tag{D.4}$$

Unter der Annahme, dass die angestrahlte Fläche keine Kugelform hat, sondern einer geraden Wand entspricht, siehe Abbildung D.2 rechts, besteht folgende Abhängigkeit zwischen bestrahlter Fläche und Öffnungswinkel

$$A_{\square} \propto \tan\left(\frac{\theta_h}{2}\right) \tan\left(\frac{\theta_v}{2}\right). \tag{D.5}$$

Dies bedeutet sogar eine überproportionale Abhängigkeit vom Winkel, so dass unter dieser Annahme die Bestrahlungsstärke überproportional von einer Vergrößerung des Öffnungswinkel betroffen ist

$$L_{\square} \propto \frac{1}{\tan\left(\frac{\theta_h}{2}\right) \tan\left(\frac{\theta_v}{2}\right)}. \tag{D.6}$$

Abhängigkeit der Strahldichte vom Öffnungswinkel

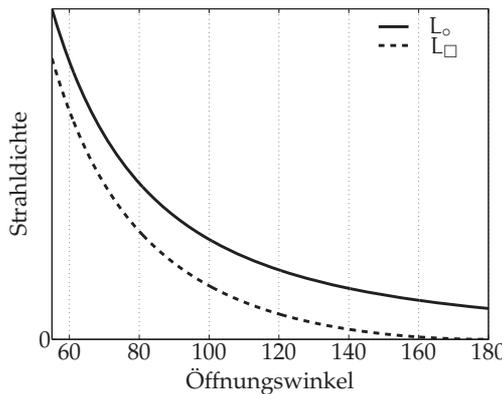


ABBILDUNG D.3: Abhängigkeit der Strahldichte in der PMD-Kamera vom Öffnungswinkel

Dies bedeutet, dass die Abstrahlcharakteristik der Strahlungsquellen auf den Öffnungswinkel der Kamera angepasst sein muss, und dass eine umgekehrt proportionale lineare Abhängigkeit der Strahldichte L vom horizontalen Öffnungswinkel θ_h und vertikalen Öffnungswinkel θ_v bleibt. Wie in Abbildung D.3 dargestellt, führt eine Vergrößerung des Öffnungswinkels also zu einem Verlust an Strahldichte in der Kamera, der im besten Fall proportional (durchgezogene Linie), tendenziell jedoch überproportional stark (gepunktete Linie) auftritt. Im letzteren Fall muss bei einem Öffnungswinkel von 180° eine unendlich große Fläche erfasst werden. Daher sinkt die Strahldichte auf den Wert Null.

ABKÜRZUNGSVERZEICHNIS

- A/D** Analog-zu-Digital
- ABS** Anti-Blockier-System
- ACC** Adaptive Cruise Control
- ACC FSR** Adaptive Cruise Control Full Speed Range
- ACEA** European Automobile Manufacturers Association
- ADAC** Allgemeiner Deutscher Automobil Club
- ADC** Analog Digital Converter
- ADR** Automatische Distanzregelung
- ADTF** Automotive Data and Time Triggered Framework
- AEC** Automatic Exposure Control
- AEV** Audi Electronics Venture GmbH
- AGC** Automatic Gain Control
- AIS** Abbreviated Injury Scale
- AKTIV** Adaptive und Kooperative Technologien für den Intelligenten Verkehr
- ALA** Audi Lane Assist
- ANB** Automatische Notbremse
- APCS** Advanced Pre Crash Safety System
- APS** Acoustic Parking System
- ARA** Abstandsregelanlage
- ASR** Anti-Schlupf-Regelung
- ASA** Audi Side Assist
- AUTOSAR** Automotive Open System Architecture
- AWV** Anhaltewegverkürzung

BMBF Bundesministerium für Bildung und Forschung
BoF Bereich ohne Fluchtmöglichkeit
BSD Blind Spot Detection
BV Bildverarbeitung
CAN Controller Area Network
CATI Computer-Assisted Telephone Interviewing
CCD Charge-Coupled Device
CMOS Complementary Metal-Oxide Semiconductor
CV Closing Velocity
CW Continuous Wave
DM Differenzenmoment
DRL Daytime Running Lights(Tagfahrlicht)
DSC Dynamic Stability Control
EDS Elektronische Differentialsperre
EDL Electronic Differential Lock
EKF Extended Kalman Filter
ELC Electronic Light Control
EuroNCAP European New Car Assessment Programme
ETSI European Telecommunications Standards Institute
ERM Empirical Risk Minimization
ESP Elektronische Stabilitätsprogramm
FAS Fahrerassistenzsysteme
FG Fußgänger
FMCW Frequency-Modulated Continuous Wave
FCC Federal Communications Commission
FIA Federation Internationale de l'Automobile
FIR Ferninfrarot
FPGA Field Programmable Gate-Array
FSS Fußgängerschutzsystem

GPS Global Positioning System
GRA Geschwindigkeitsregelanlage
INVENT Intelligent traffic and user-friendly technology
IPM Inverse Perspective Mapping
IR Infrarot
IRTAD International Road Traffic and Accident Database
ISS Injury Severity Score
JAMA Japan Automobile Manufacturers Association
KAMA Korea Automobile Manufacturers Association
KNN Künstliche Neuronale Netze
LED Light Emitting Diodes
LIDAR Light Detection and Ranging
LRR Long Range Radar
MAIS Maximum Abbreviated Injury Scale
MAP Maximum-A-Posteriori
MIT Massachusetts Institute of Technology
ML Maximum-Likelihood
ML-Schätzer Maximum-Likelihood-Schätzer
MSDF Multisensordatenfusion
MV Monovideo
nFG Nicht-Fußgänger
NIR Nahinfrarot
NN Neuronale Netze
NV Night Vision
PLA Parklenkassistent
PSD Prädiktive Streckendaten
PD Pulse Doppler
Pkw Personenkraftwagen
PMD Photonic Mixer Device

PReSUME Pedestrian Recognition System Using a Multi-Sensor Environment

PROTECTOR Preventive Safety for Unprotected Road User

RBF radiale Basisfunktionen

RegTP Regulierungsbehörde für Telekommunikation und Post

RFID Radio Frequency Identification

ROC Receiver Operating Characteristic

ROI Region-Of-Interest

SAVE-U Sensors and System Architecture for Vulnerable road Users protection

SDF Sensordatenfusion

SFR Schutz für Fußgänger und Radfahrer

SGLD Spatial Greylevel Dependence

SOM Self-Organizing Maps

SRR Short Range Radar

SfM Structure from Motion

SVM Support Vector Machine

TDoA Time Difference of Arrival

TOF Time Of Flight

TTC Time To Contact

UK Großbritannien und Nordirland

UMRR Universal Medium Range Radar

US Ultraschall

VBPP Vision Based Pedestrian Protection

VGA Video Graphics Array

WHO World Health Organization

WSK Wahrscheinlichkeit

LITERATURVERZEICHNIS

- [1] M.-M. Meinecke und M. A. Obojski, "Sicherheit für Fußgänger und Radfahrer – SFR," *Förderprojekt AKTIV – Schutz für Fußgänger und Radfahrer, Teilprojektbericht*, September 2007.
- [2] "Unfallgeschehen im Straßenverkehr," Statistisches Bundesamt, Destatis, Wiesbaden, Deutschland, S. 14, Juli 2007, artikelnummer 5462401089004.
- [3] "Frost and Sullivan, Enforcement of Pedestrian Protection Legislation to Drive Demand for Pedestrian Safety Systems," *Authoritative and Timely Global Business Information*, London, Februar 2006. [Online]. Verfügbar unter: <http://www.just-auto.com/ia.aspx?id=2375>. Zuletzt besucht im August 2008.
- [4] "Europäische Kommission, Initiative der Europäische Union zur Halbierung der Zahl der Verkehrstoten bis 2010." [Online]. Verfügbar unter: http://europa.eu.int/comm/transport/roadsafety/charter_de.htm. Zuletzt besucht im August 2008.
- [5] M. Kühn, R. Fröming und V. Schindler, *Fußgängerschutz*, 1. Edition. Springer Verlag, 2006.
- [6] J.-L. Maté und D. Zittlau, "Electronics For More Safety," *ATZ*, Band 1, Nr. 07-08/2006, Seite 578–585, 2006.
- [7] "Fußgängerschutz-Richtlinie 2003/102/EG der Europäischen Union," Europäische Kommission, Dezember 2003. [Online]. Verfügbar unter: http://europa.eu.int/comm/enterprise/automotive/directives/vehicles/dir2003_102_ce.htm. Zuletzt besucht im August 2008.
- [8] "Technische Vorschrift 2004/90/EC zur Ausführung von Artikel 3 der Richtlinie 2003/102/EG," Europäische Kommission, FEBRUAR 2004. [Online]. Verfügbar unter: http://europa.eu.int/comm/enterprise/automotive/directives/vehicles/dir2003_102_ce.htm. Zuletzt besucht im August 2008.

- [9] M.-M. Meinecke, M. A. Obojski, M. Töns und M. Dehesa, "SAVE-U: First Experiences with a Pre-Crash System for Enhancing Pedestrian Safety," *Intelligent Transport Systems and Services*, Juni 2005.
- [10] D. Gavrila, M.-M. Meinecke, S. Munder, M. Obojski und P. Marchal, "Demonstrator vehicles," *SAVE-U: EU Projekt Deliverable 21*, August 2005.
- [11] P. Marchal, M. Dehesa, D. M. Gavrila, M.-M. Meinecke, N. Skellern und R. Vinciguerra, "SAVE-U: Sensors and System Architecture for Vulnerable Road Users Protection – final report," August 2005.
- [12] M.-M. Meinecke, M. A. Obojski, M. Töns, R. Doerfler, P. Marchal, L. Letellier, D. Gavrila und R. Morris, "Approach for Protection of Vulnerable Road Users Using Sensor Fusion Techniques," *International Radar Symposium, Dresden, Deutschland*, September 2003.
- [13] "Honda Develops World's First Intelligent Night Vision System," Honda Motor Co., Ltd., Tokio, Japan, August 2004, JCN Newswire.
- [14] "Straßenunfälle in Bayern 2006," Bayerisches Landesamt für Statistik und Datenverarbeitung, Juni 2007, statistische Berichte.
- [15] H. Becker und M. Stanzel, "Verletzungsverteilungsschwere für Fußgänger nach Kollisionsgeschwindigkeit des Pkw," *unveröffentlichter Bericht*, Februar 2003.
- [16] M.-M. Meinecke, A. M. Obojski, D. Gavrila, E. Marc, R. Morris, M. Töns und L. Letellier, "Strategies in Terms of Vulnerable Road User Protection," *SAVE-U EU Projekt Deliverable D6*, Mai 2003.
- [17] M. Richter, D. Otte, H. C. Pape, S. Glueer, B. Koenemann und H. Tscherne, *Problematik der Verletzungen von Kindern und Jugendlichen im Straßenverkehr*, 8. Edition. Springer Berlin / Heidelberg, August 2001, Band 104.
- [18] M. Kaufmann, "Der Fahrer legt die Hände in den Schoß (2)," Februar 2005. [Online]. Verfügbar unter: <http://www.manager-magazin.de/life/auto/0,2828,342970-2,00.html>. Zuletzt besucht im August 2008.
- [19] V. Schindler und R. Fröming, "Appraisal of Active and Passive Safety Measures on a Common Scale Considerung Pedestrian Safety as Example," Workshop at the Technical University of Berlin, Deutschland, Technischer Bericht, 2005.

- [20] S. Rohr, R. Lind, R. Myers, W. Bauson, W. Kosiak und H. Yen, "An Integrated Approach to Automotive Safety Systems," *Delphi Automotive Systems*, 2000.
- [21] W. Vogt, "Arbeitsrechtliche Aspekte zu Fahrerassistenzsystemen," 3. Workshop Fahrerassistenzsysteme, Walting, Deutschland, Seite 33–40, April 2005.
- [22] S. Kopischke, "Entwicklung einer Notbremsfunktion mit Rapid Prototyping Methoden," Dissertation, Fakultät für Maschinenbau und Elektrotechnik der Technischen Universität Carolo-Wilhelmina zu Braunschweig, Deutschland, Februar 2000.
- [23] forsa, "Meinungen zu Dieselrußfiltern," Gesellschaft für Sozialforschung und statistische Analyse mbH, Juni 2004.
- [24] emnid, "Umfrage zum Thema Sicherheit," 2003. [Online]. Verfügbar unter: <http://www.tns-emnid.com/index.htm>. Zuletzt besucht im Mai 2003.
- [25] "Automobil Produktion, Autofahrer wollen mehr Sicherheit." [Online]. Verfügbar unter: <http://www.automagazine.de/ap/themen/02793/index.php>. Zuletzt besucht im August 2008.
- [26] "Verkehrstechnisches Institut der Deutschen Versicherer, Fahrerassistenzsysteme: Mehr Sicherheit im Straßenverkehr?" GDV Presseforum der Schaden- und Unfallversicherer in Hamburg, Technischer Bericht, Mai 2004, Seite 32.
- [27] SWR, "Fahrerassistenzsysteme: Die Vision vom unfallfreien Fahren." [Online]. Verfügbar unter: <http://www.swr.de/rasthaus/archiv/2004/10/16/beitrag1.html>. Zuletzt besucht im Oktober 2004.
- [28] "Pressedienst der Versicherungswirtschaft, Fahrer denkt – Elektronik lenkt: Mit Assistenzsystemen mehr Sicherheit oder neue Risiken?" Mai 2004.
- [29] B. Fardi, J. Douša, G. Wanielik, B. Elias und A. Barke, "Obstacle Detection and Pedestrian Recognition Using A 3D PMD Camera," *IEEE International Conference on Intelligent Vehicles*, Seite 225–230, Juni 2006.
- [30] B. Elias und P. Mähönen, "Pedestrian Recognition Based on 3D Image Data," *IEEE International Symposium on Industrial Electronics*, Seite 1406–1411, Juni 2007.

- [31] H. Heinol, Z. Xu, R. Schwarte, J. Olk und R. Klein, "Electrooptical Correlation Arrangement for Fast 3D Cameras: Properties and Facilities of this Electro-Optical Mixer Device," *SPIE EOS -Sensors, Sensor Systems, and Sensor Data Processing*, Seite 245–253, 1997.
- [32] X. Luang, "Experimental Investigation of Photonic Mixer Device and Development of TOF 3D Ranging Systems Based on PMD Technology," Dissertation, Universität-Gesamthochschule Siegen, Deutschland, November 2001.
- [33] R. Lange, "3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology," Dissertation, Universität Siegen, September 2000.
- [34] R. Schwarte, H. Heinol, Z. Xu, J. Olk und W. Tai, "Schnelle und einfache optische Formerfassung mit einem neuartigen Korrelations-Photodetektor-Array," *GMA-Bericht 30 - Optische Formerfassung, DGZfP-VDI/VDE-GMA*, Seite 199–209, 1997.
- [35] D. Göhring, "Digitalkameratechnologien – eine vergleichende Betrachtung CCD kontra CMOS," Humboldt Universität zu Berlin, August 2002.
- [36] S. Bota und S. Nedesvchi, "Multi-Feature Walking Pedestrian Detection Using Dense Stereo and Motion," *Workshop on Intelligent Transportation WIT, Hamburg, Deutschland*, März 2007.
- [37] F. Pagel, B. Elias, B. Giesler und D. Willersinn, "Analysis of Binocular Stereo Systems," *Workshop on Intelligent Transportation WIT, Hamburg, Deutschland*, März 2008.
- [38] F. Fölster, "Erfassung ausgedehnter Objekte durch ein Automobil-Radar," Dissertation, Technische Universität Hamburg-Harburg, Dezember 2006.
- [39] Hella, "Technical Information Electronics - Driver Assistance System," S. 12, 2003.
- [40] K. Fürstenberg und K. Dietmayer, "Object Tracking and Classification for Multiple Active Safety and Comfort Applications using a Multi-layer Laserscanner," *IEEE Intelligent Vehicles Symposium*, Juni 2004.
- [41] A. Taner, "Vergleich verschiedener Systeme zur Sichtverbesserung bei Nacht in Personenkraftwagen," Dissertation, Technische Universität München, Deutschland, März 2007.

- [42] D. Linzmeier, M. Mekhail, J. Dickmann und K. C. J. Dietmayer, "Pedestrian Protection with Thermopiles using an Occupancy Grid," Seite 1063–1068, Oktober 2004.
- [43] H. P. Moravec und A. Elfes, "High resolution maps from wide angle sonar," *Proceedings IEEE International Conference of Robotics and Automation*, Seite 116–121, 1985.
- [44] T. Ender, "Objekt- und Bewegungsdetektion in Verkehrsszenen mit Hilfe optischer Abstandssensorik," Masterarbeit, Fachhochschule Ingolstadt, Deutschland, Audi Electronics Venture GmbH, Januar 2005.
- [45] B. Elias, M. Roehder und P. Mähönen, "Improvement of a PMD Based Classification of Pedestrians in Road Traffic Scenarios," *Workshop on Intelligent Transportation WIT, Hamburg, Deutschland*, März 2008.
- [46] C.-E. Liedtke und M. Ender, *Wissensbasierte Bildverarbeitung*, 1. Edition. Springer Verlag, 1989.
- [47] W. Voß, *Taschenbuch der Statistik*, 2. Edition. Fachbuchverlag Leipzig im Carl Hanser Verlag, 2004.
- [48] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Transactions on Information Theory IT-8*, Seite 179–187, 1962.
- [49] S. Funck und S. Fuchs, "Entwicklung eines gestisch-intuitiven Mensch-Maschine-Interfaces auf Basis der videogestützten Erkennung von Handzeichen," Technischer Bericht, Fakultät Informatik, Institut für Künstliche Intelligenz, Technischer Bericht.
- [50] A. Broggi, M. Bertozzi, A. Fasciolo und M. Sechi, "Shape-based Pedestrian Detection," *IEEE Intelligent Vehicles Symposium*, 2000.
- [51] L. Zhao und C. E. Thorpe, "Stereo- and Neural Network-Based Pedestrian Detection," *IEEE Transactions on Intelligent Transportation Systems*, Band 1, Nr. 3, 2000.
- [52] D. M. Gavrila, J. Giebel und S. Munder, "Vision-Based Pedestrian Detection: The PROTECTOR System," *Proceedings of the IEEE Intelligent Vehicles Symposium*, 2004.
- [53] "Bildinformationssysteme, Teil 4," Vorlesung Bildverarbeitung von Prof. Ullrich Köthe, Universität Hamburg, 2005.
- [54] M. Laschkow, "Texturanalyse," *Hochschule Bremen & Alfred-Wegener-Institut für Polar- und Meeresforschung*, 1996.

- [55] L. Held, *Methoden der statistischen Inferenz - Likelihood und Bayes*, 1. Edition. Heidelberg: Spektrum Akademischer Verlag, 2008.
- [56] R. Aliev, K. W. Bonfig und F. Aliew, *Soft Computing – Eine grundlegende Einführung*. Berlin: Verlag Technik, 2000.
- [57] D. Hebb, "Organisation of Behavior," *Wiley Interscience*, 1949.
- [58] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge University Press, 1996.
- [59] F. Xu, X. Liu und K. Fujimura, "Pedestrian Detection and Tracking with Night Vision," *IEEE Intelligent Transportation Systems*, Band 1, Seite 63–71, März 2005.
- [60] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2. Edition. Springer Verlag, 1995.
- [61] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery 2*, Seite 121–167, 1998.
- [62] I. Bronštein und K. Semendjaev, *Taschenbuch der Mathematik*, 5. Edition. Harri Deutsch, 2001.
- [63] Z. Zhang, "A Flexible New Technique for Camera Calibration," *Technischer Bericht*, Dezember 2002, mSR-TR-98-71.
- [64] M. Kellermeier, "2D/3D Kamerakalibrierung in Vorentwicklung, Fertigung und Service," Masterarbeit, Fachhochschule Ingolstadt, Deutschland, Audi Electronics Venture GmbH, 2007.
- [65] R. Y. Tsai, "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, Band RA-3, Nr. 4, Seite 323–344, August 1987.
- [66] I. Gerdes, F. Klawonn und R. Kruse, *Evolutionäre Algorithmen: genetische Algorithmen - Strategien und Optimierungsverfahren - Beispielanwendungen*. Wiesbaden: Verlag Vieweg, 2004.
- [67] W.-M. Lippe, *Soft-computing: Mit neuronalen Netzen, Fuzzy-Logic und Evolutionären Algorithmen*. Berlin, Heidelberg: Verlag Springer, 2006.
- [68] W. Youden, "Index for Rating Diagnostic Tests," *Cancer*, Band 3, Nr. 1, Seite 32 – 35, 1950. [Online]. Verfügbar unter: [http://dx.doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](http://dx.doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3). Zuletzt besucht im März 2009.

- [69] A. Broggi, F. Fascioli, M. Carletti, T. Graf und M.-M. Meinecke, "A Multi-resolution Approach for Infrared Vision-based Pedestrian Detection," *IEEE Intelligent Vehicles Symposium*, Juni 2004.
- [70] M. Bertozzi, E. Binelli, A. Broggi und M. D. Rose, "Stereo Vision-Based Approaches for Pedestrian Detection," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Juni 2005.
- [71] M. Bertozzi, A. Broggi, P. Grisleri, A. Tibaldi und M. D. Rose, "A Tool for Vision-Based Pedestrian Detection Performance Evaluation," *Proceedings of the IEEE Intelligent Vehicles Symposium*, Juni 2004.
- [72] M. Bertozzi, A. Broggi, C. Caraffi, M. D. Rose, M. Felisa und G. Vezioni, "Pedestrian Detection by Means of Far-Infrared Stereo Vision," *Computer Vision and Image Understanding 106*, Seite 194–204, 2007.
- [73] H. Ritter, "Fußgängererkennung mit Radarsensoren," Masterarbeit, Technische Universität Hamburg-Harburg, Juni 2005.
- [74] G. Ma, S.-B. Park, A. Ioffe, S. Müller-Schneiders und A. Kummert, "A Real Time Object Detection Approach Applied to Reliable Pedestrian Detection," *IEEE Intelligent Vehicles Symposium, Istanbul, Türkei*, Seite 755–760, Juni 2007.
- [75] G. Grubb, A. Zilensky, L. Nilsson und M. Rilbe, "3D Vision Sensing for Improved Pedestrian Safety," *Proceedings of the IEEE Intelligent Vehicles Symposium*, Juni 2004.
- [76] D. M. Gavrila und S. Munder, "Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle," *International Journal of Computer Vision 73*, Seite 41–59, 2007.
- [77] D. M. Gavrila, M. Kunert und U. Lages, "A Multi-Sensor Approach for the Protection of Vulnerable Traffic Participants - The PROTECTOR Project," *IEEE Instrumentation and Measurement Technology Conference, Budapest, Ungarn*, Mai 2001.
- [78] K. Kidono, A. Takahashi und Y. Ninomiya, "Pedestrian Detection using Millimeter Wave Radar and Stereo Vision," *Workshop on Intelligent Transportation WIT, Hamburg, Deutschland*, März 2007.
- [79] F. Xu und K. Fujimura, "Pedestrian Detection and Tracking with Night Vision," *IEEE Intelligent Vehicles Symposium*, Band 1, Seite 21–30, Juni 2002.

